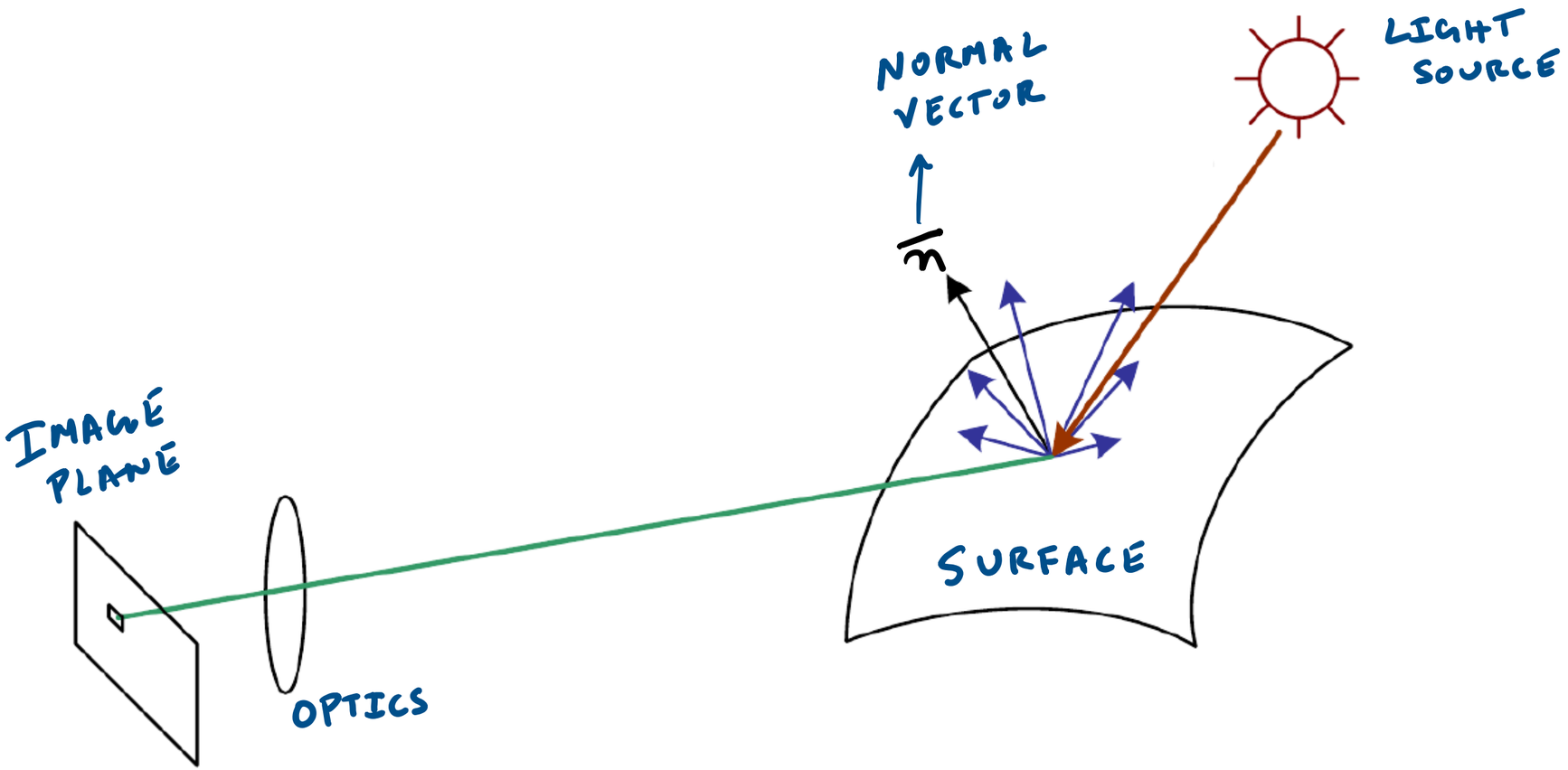
Image formation (part 3)

Lecture 4

COMP 6982/ENGI 9805: Computer Vision

Karteek Popuri

Video recording of Lecture 4 (Week 2) – 17 Jan 2024



When light hits the surface of an object in the scene, it gets REFLECTED in one or more directions

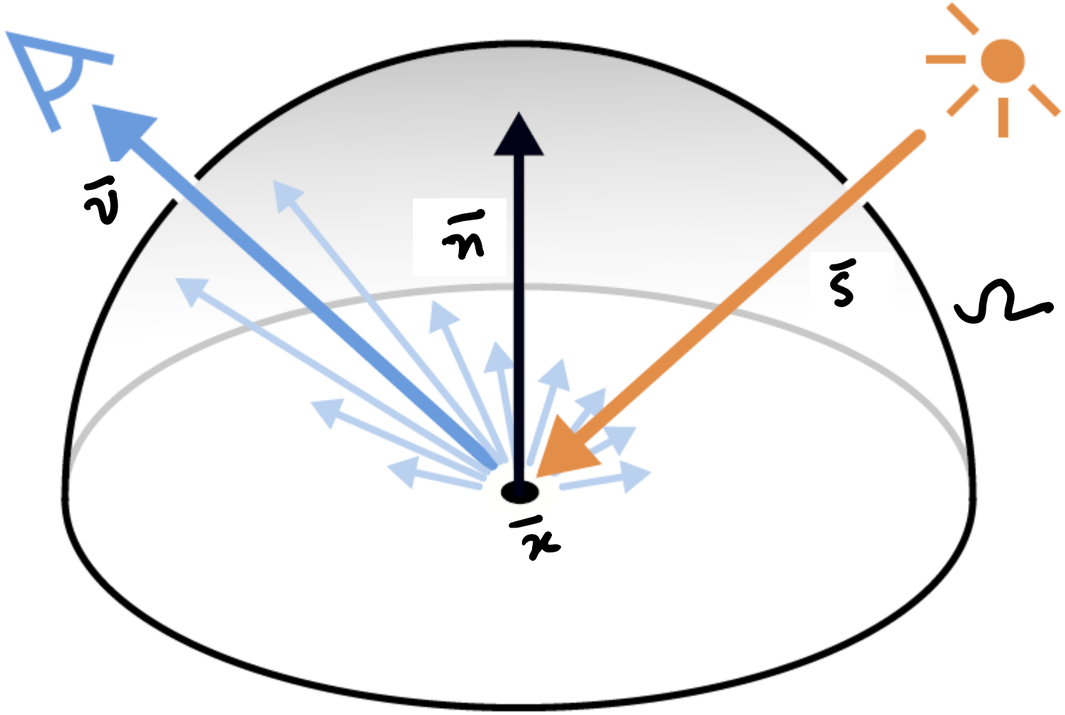
BIDIRECTIONAL REFLECTANCE DISTRIBUTION FUNCTION

FORESHORTENING

$$L_{OUT}(\bar{x}, \bar{v}, \lambda) = L_{EMIT}(\bar{x}, \bar{v}, \lambda) + \int_{\Omega} BRDF(\bar{x}, \bar{s}, \bar{v}, \lambda)\, L_{IN}(\bar{x}, \bar{s}, \lambda)\, (-\bar{n} \cdot \bar{s})\, d\bar{s}$$

WAVELENGTH

> 0 ONLY for light emitting surfaces



NON-NEGATIVE:

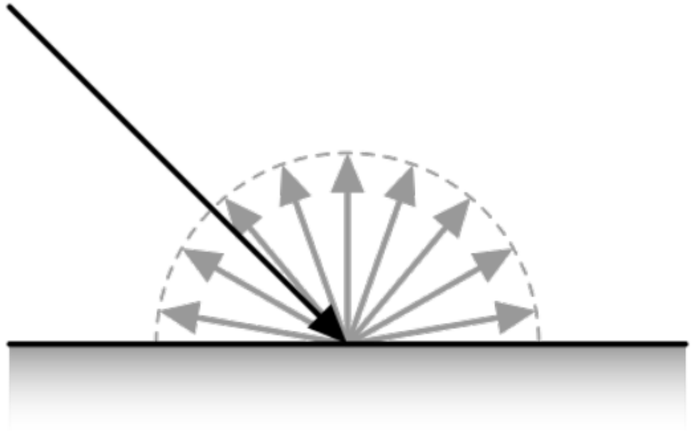
$BRDF(\bar{x}, \bar{s}, \bar{v}, \lambda) \geq 0$

HELMHOLTZ RECIPROCITY:

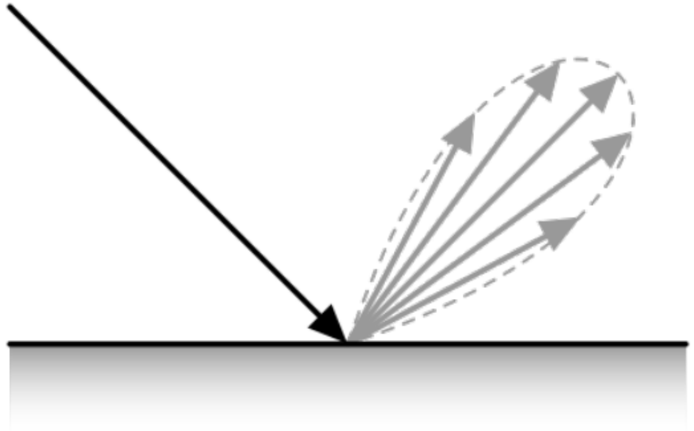
$BRDF(\bar{x}, \bar{s}, \bar{v}, \lambda) = BRDF(\bar{x}, \bar{v}, \bar{s}, \lambda)$

[ML-4360, Lecture 2]

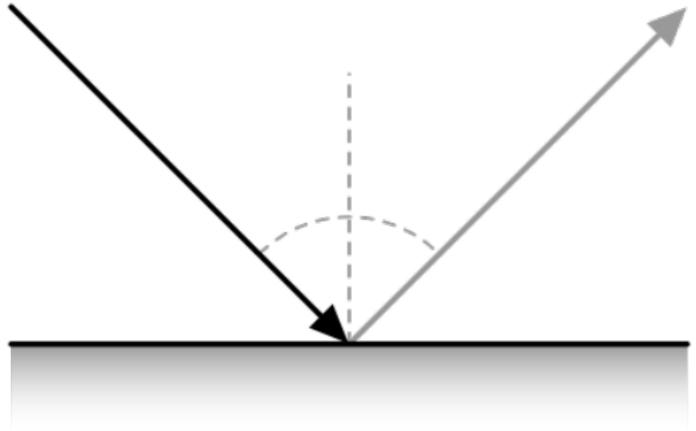
DIFFUSE
⇓
$BRDF(\bar{x}, \beta, \bar{v}, \lambda) = f(\lambda)$



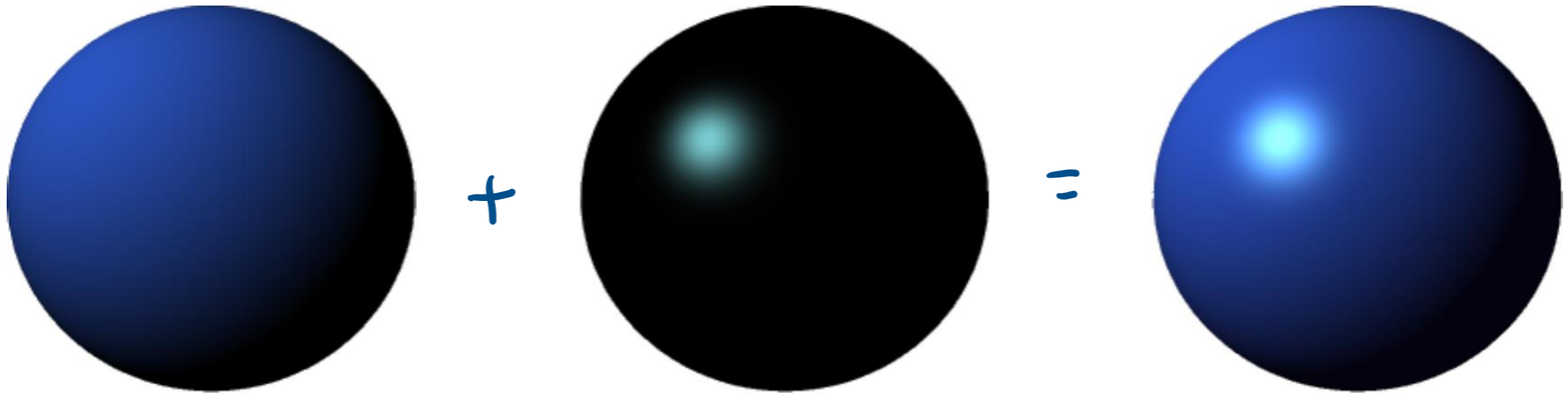
SPECULAR
⇓
BRDF depends STRONGLY on $\bar{v}$



MIRROR
⇓
BRDF is NON-ZERO ONLY for a particular $\bar{v}$



Real world BRDFs have BOTH DIFFUSE and SPECULAR components:



DIFFUSE + SPECULAR = COMBINED

[ML-4360, Lecture 2]

In general BRDFs can be very complex and spatially varying

"shallow" viewing angle

↑

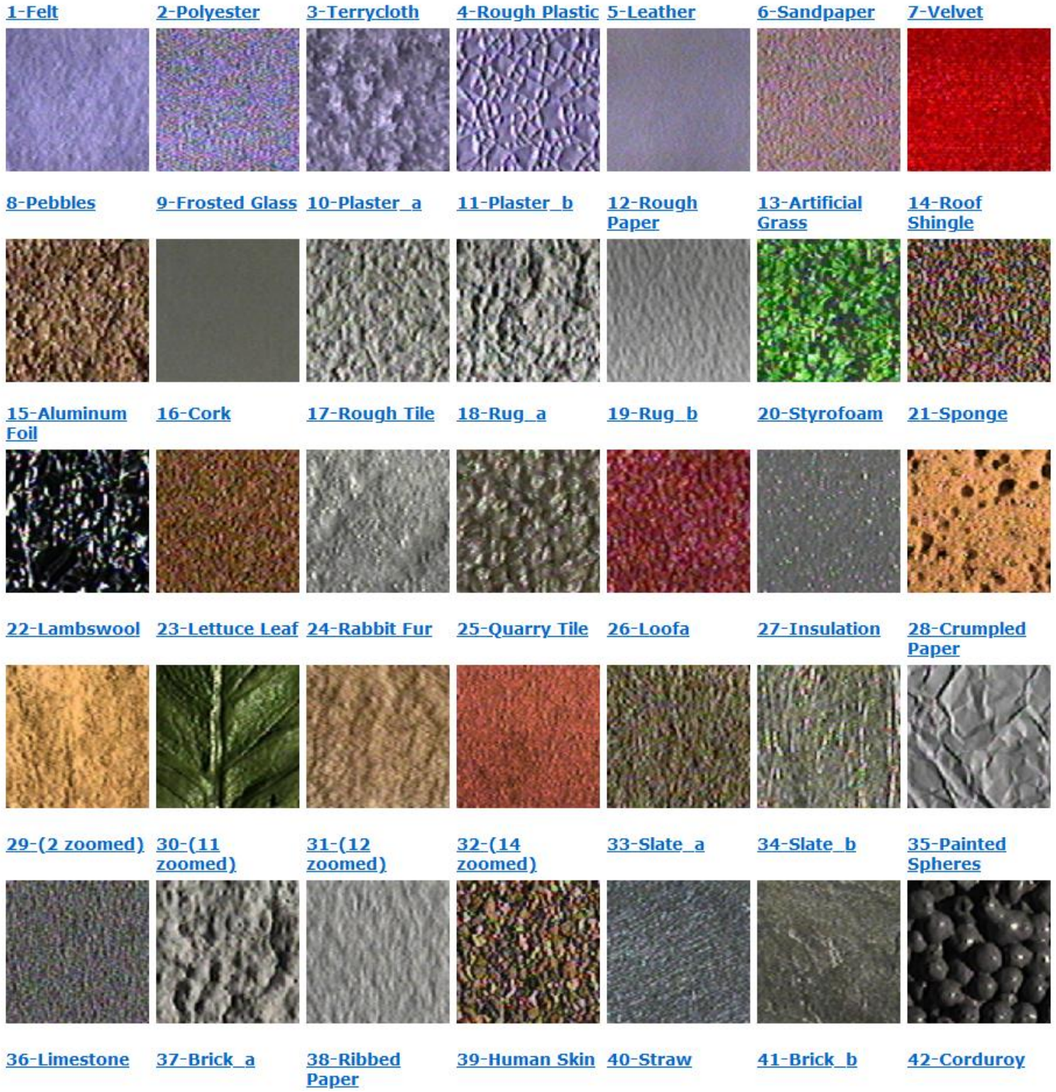
MIRROR LIKE REFLECTION ON THE TOP LAYER DOMINATES

DIFFUSE REFLECTION FROM THE BOTTOM LAYER CONTRIBUTES

↓

"steep" viewing angle

FRESNEL EFFECT: In LAYERED objects, the amount of light that reaches the BOTTOM layer is reduced owing to FRESNEL REFLECTION effect on the TOP layer



[CUReT database]

CuRRET:
Columbia-Utrecht Reflectance and Texture Database

Kristin J. Dana Bram Van Ginneken
Shree K. Nayar Jan J. Koenderink

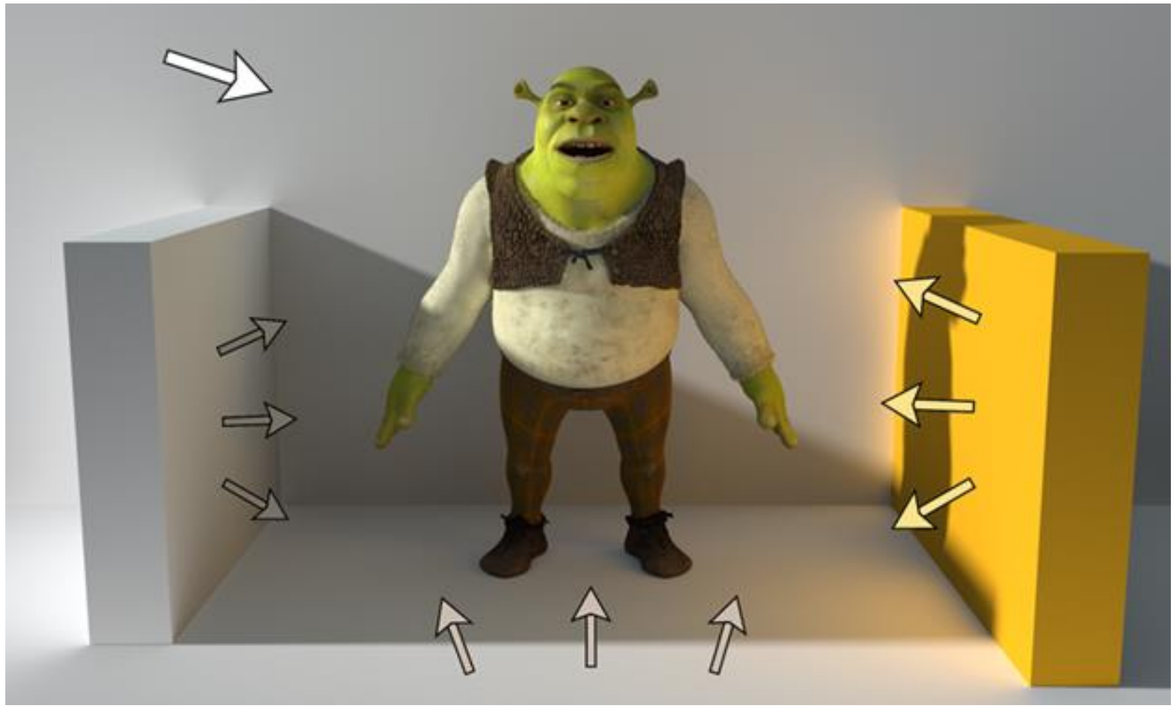
Researchers at Columbia University and Utrecht University have collaborated in an extensive investigation of the visual appearance of real-world surfaces. This joint effort, sponsored in part by REALISE of the European Commission, the National Science Foundation and by DARPA/ONR under the MURI Grant No. N00014-95-1-0601 has resulted in 3 databases:

1. **BRDF *(bidirectional reflectance distribution function)* database** with reflectance measurements for over 60 different samples, each observed with over 200 different combinations of viewing and illumination directions.

→ GONIOREFLECTOMETER

RENDERED WITH DIRECT LIGHTING

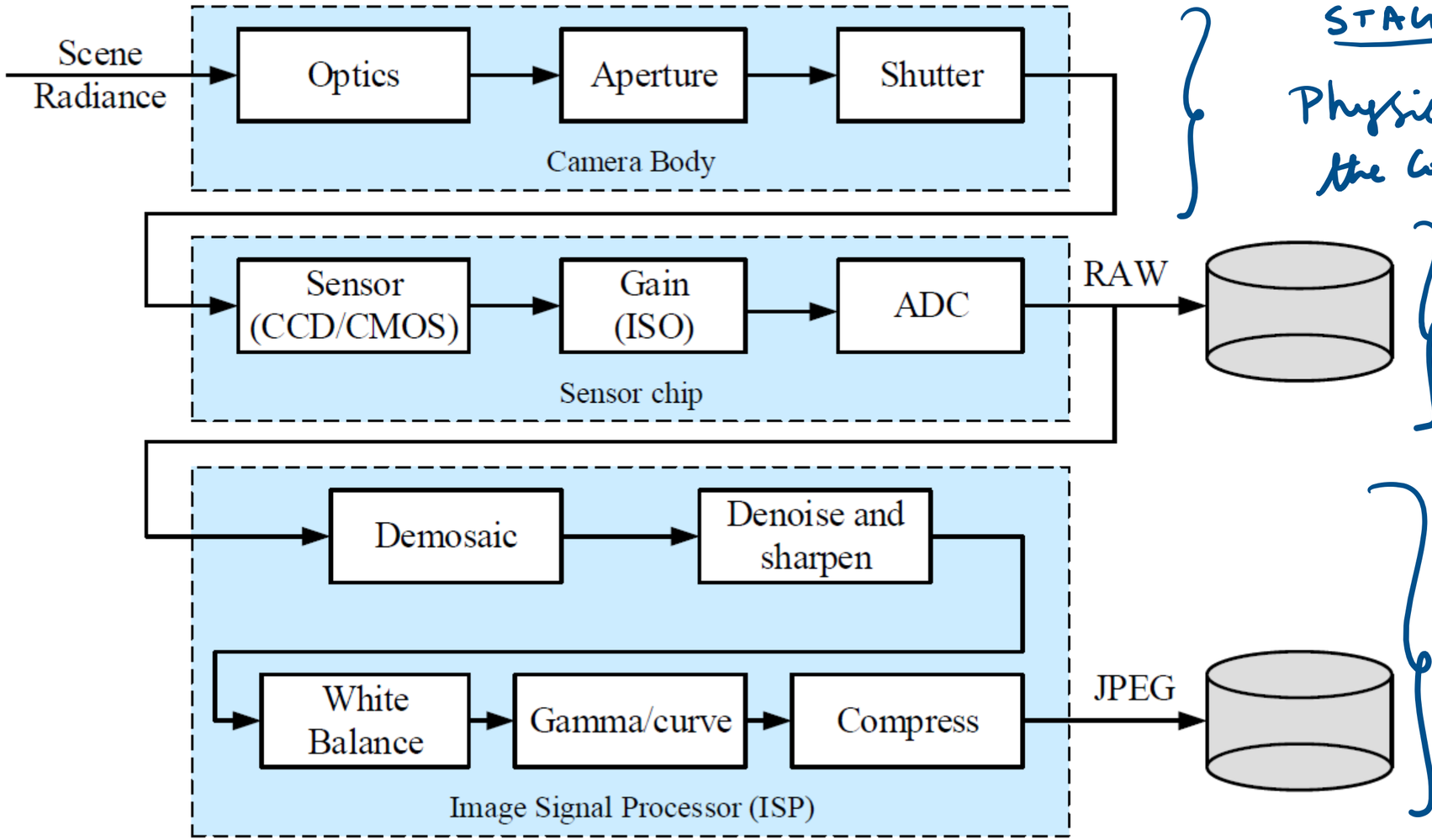
RENDERED WITH GLOBAL ILLUMINATION



In reality we have:

- SHADOWS owing to occluders
- light rays bouncing MULTIPLE times in the scene

} → GLOBAL ILLUMINATION EFFECTS



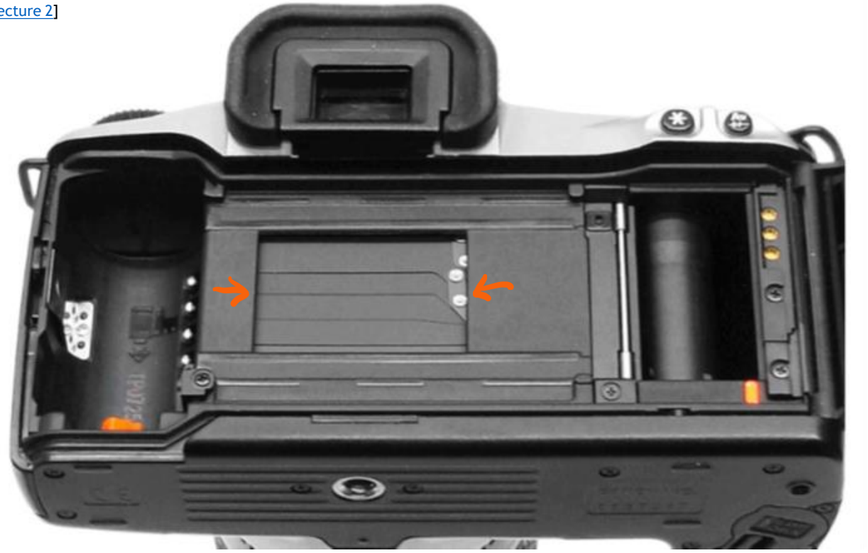
STAGE 1:
Physical light transport in the camera

STAGE 2:
Photon measurement and conversion on the sensor chip

STAGE 3:
Image signal processing and compression

[Szeliski 2022, Page 64]

[ML-4360, Lecture 2]



Shutter Speed
Controls
EXPOSURE TIME
|||
how much light
reaches the sensor

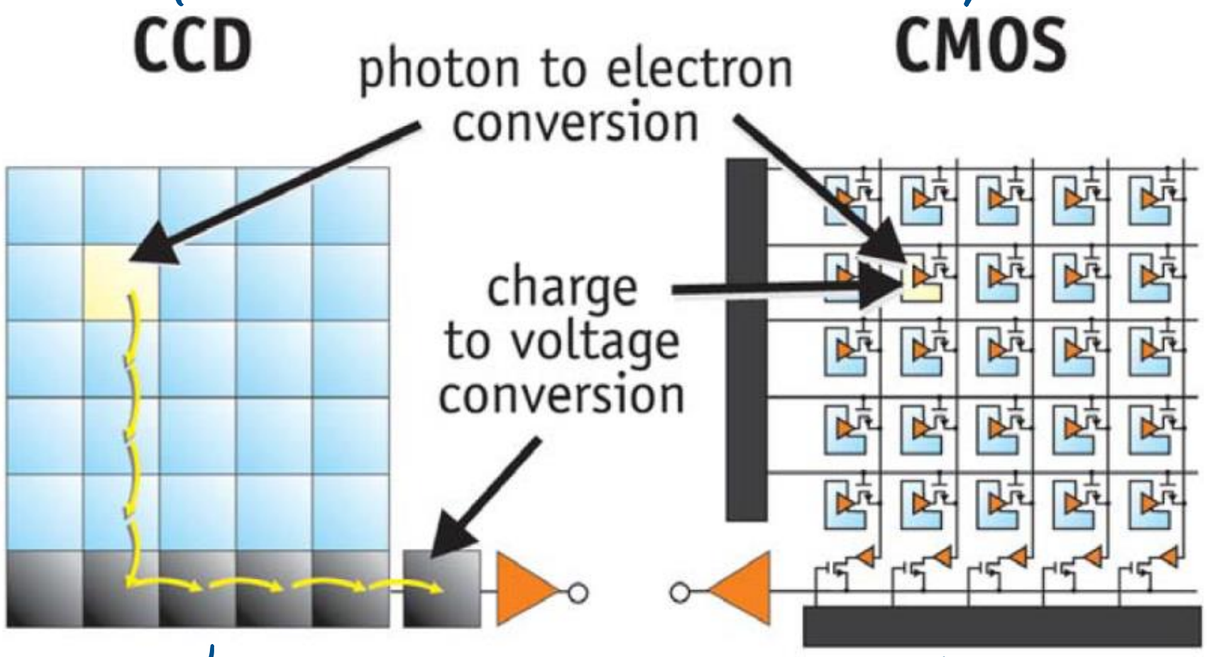
[https://www.studiobinder.com/blog/underexposed-photography-technique]

→ LONG EXPOSURE NIGHT SHOT

[https://www.eyexplore.com/blog/long-exposure-basic-how-to/]

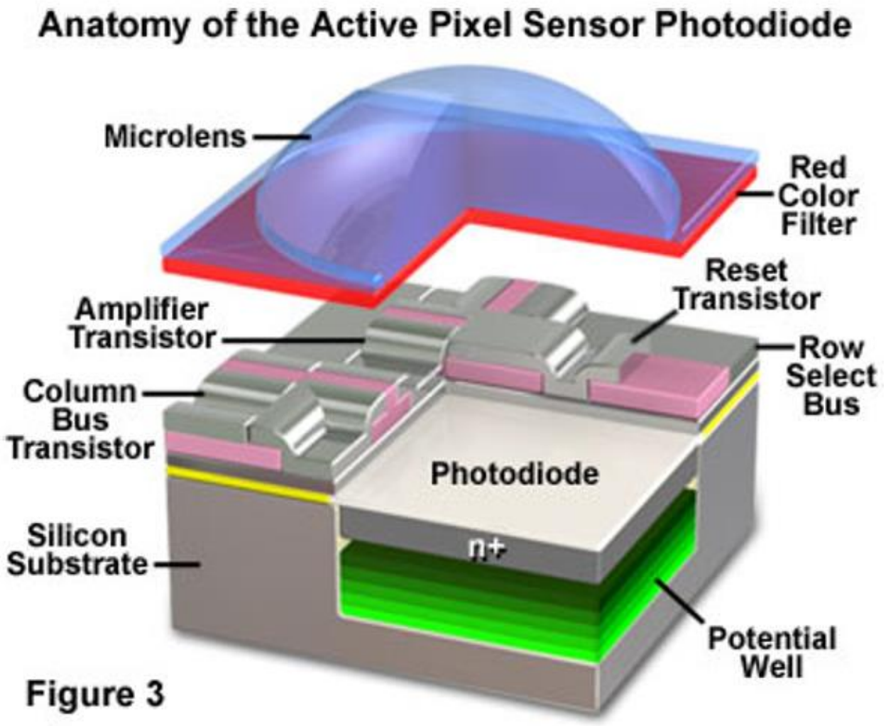
CHARGE-COUPLED DEVICE

COMPLEMENTARY METAL OXIDE ON SILICON



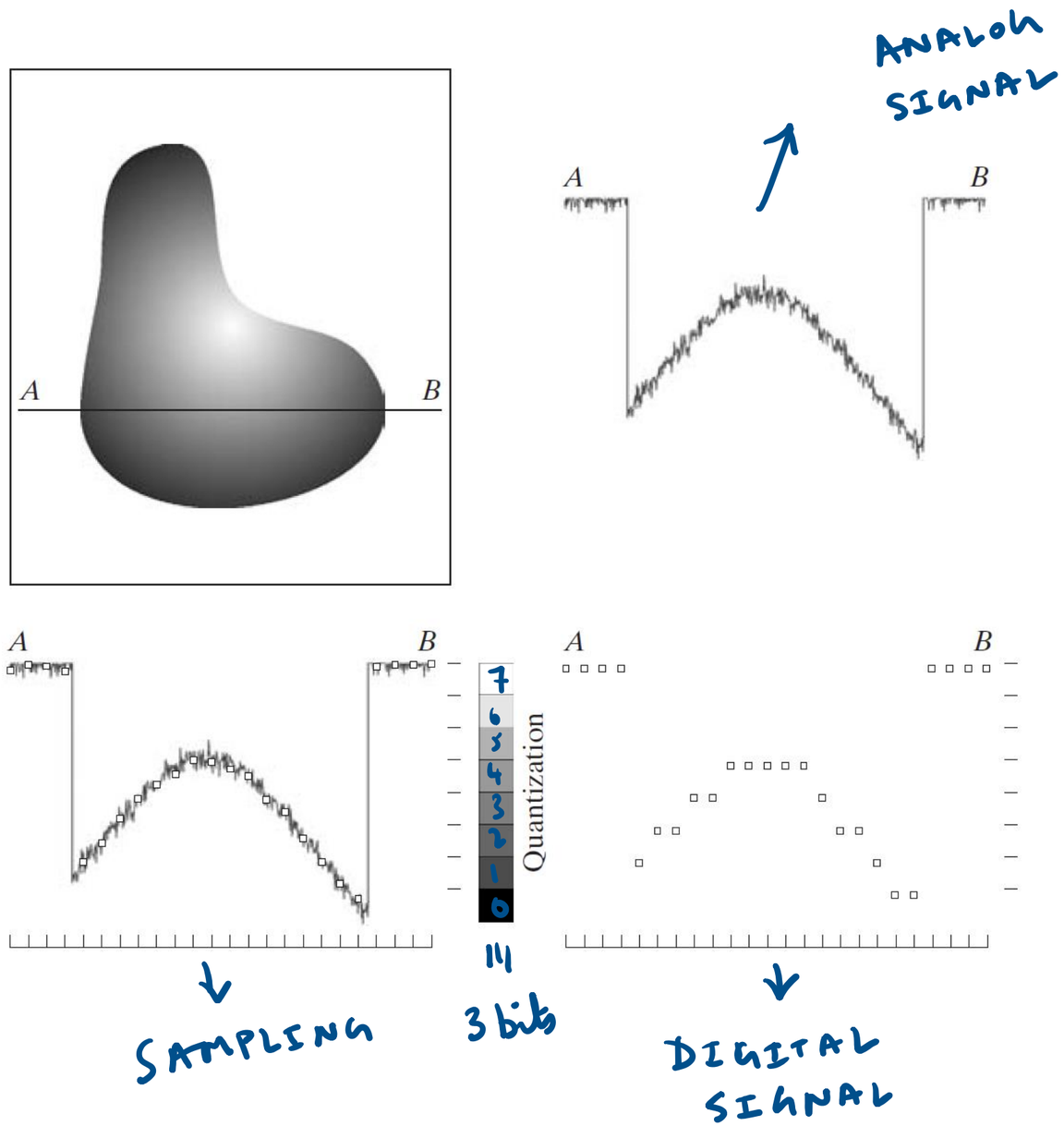
HIGHER QUALITY

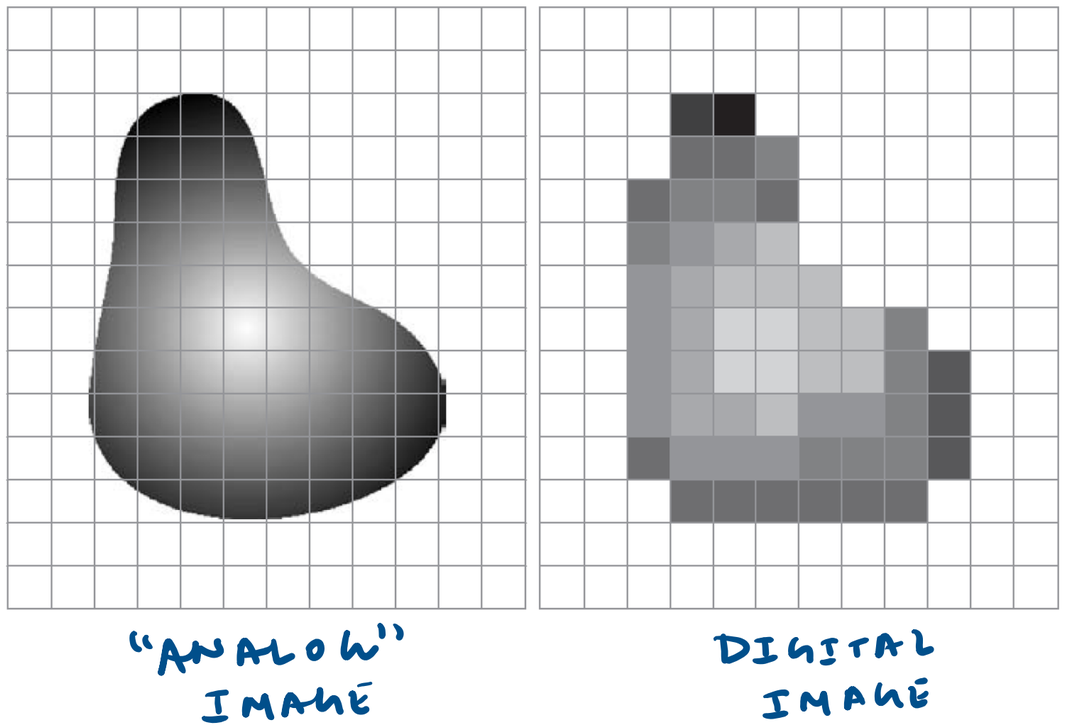
LOWER POWER CONSUMPTION



SEE HERE FOR A MORE DETAILED CCD VS CMOS COMPARISON:
https://www.scienceshot.com/post/cmos-vs-ccd-who-is-the-clear-winner

[Szeliski 2022, Page 65]

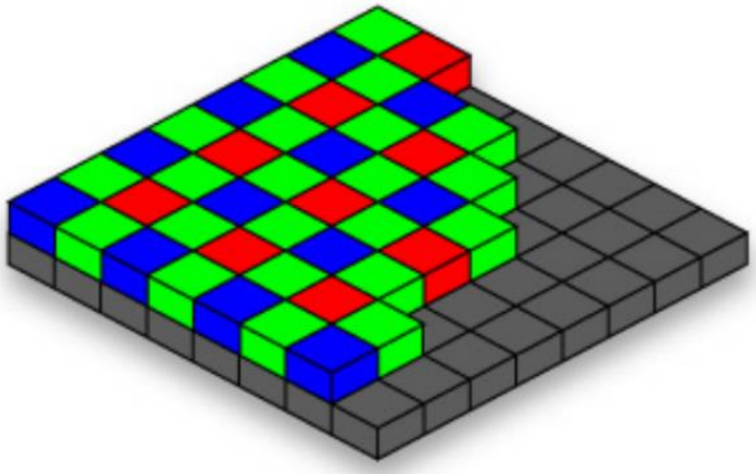


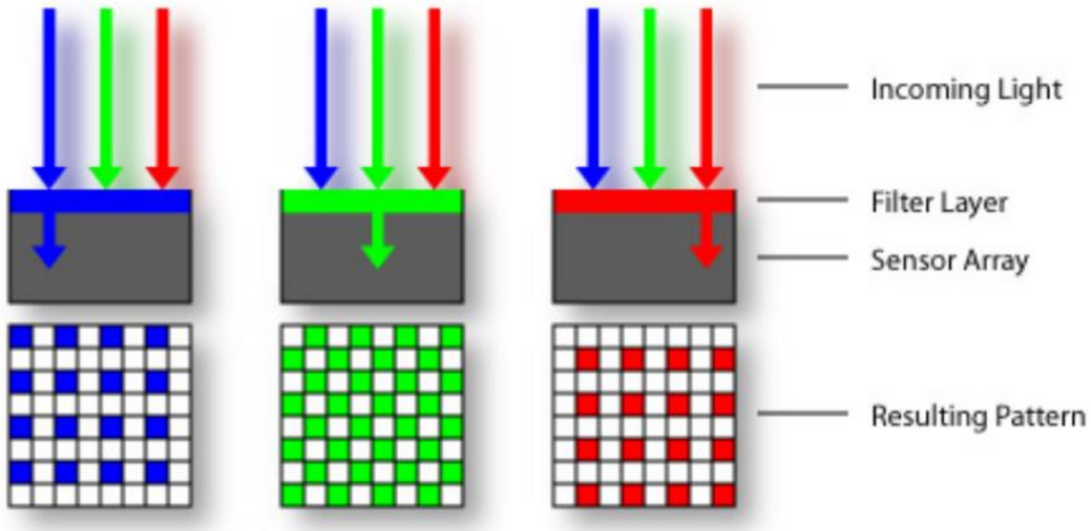


8 bits or 16 bits usually used

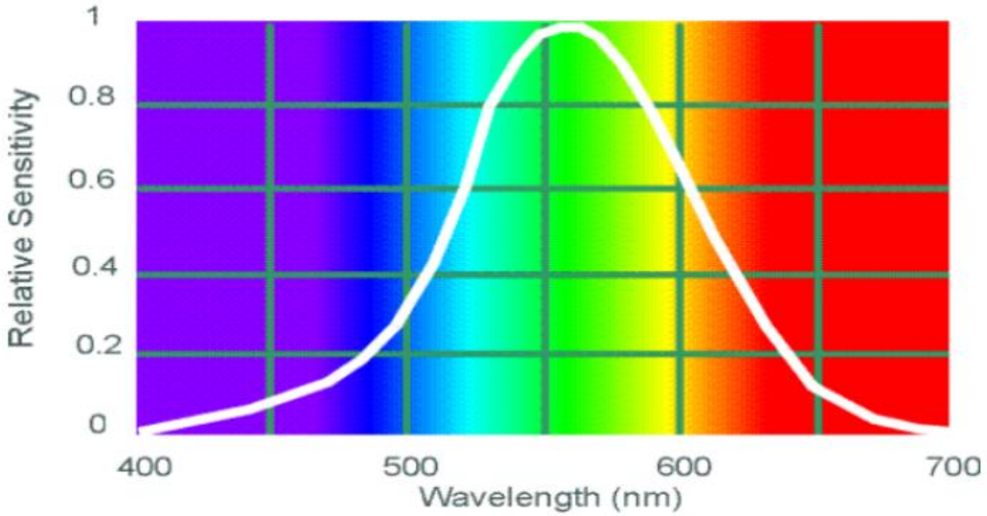
[Gonzalez and Woods 2008, Page 53, 54]

BAYER RGB PATTERN





HUMAN EYES are MORE sensitive to GREEN light



DEMOSAICING

G	R	G	R
B	G	B	G
G	R	G	R
B	G	B	G

rGb	Rgb	rGb	Rgb
rgB	rGb	rgB	rGb
rGb	Rgb	rGb	Rgb
rgB	rGb	rgB	rGb

[Szeliski 2022, Page 74]

[ML-4360, Lecture 2]

JPEG COMPRESSION

RAW IMAGE

DISCRETE COSINE TRANSFORM (DCT)

→ Coefficients are quantized to integers and stored using Huffman codes

INCREASING LEVEL OF JPEG COMPRESSION →

PEAK SIGNAL-TO-NOISE RATIO (PSNR)

$$MSE = \frac{1}{n} \sum_{x} \left[I(x) - \hat{I}(x) \right]^2$$

$$RMS = \sqrt{MSE}$$

$$PSNR = 10 \log_{10} \frac{I_{max}^2}{MSE}$$

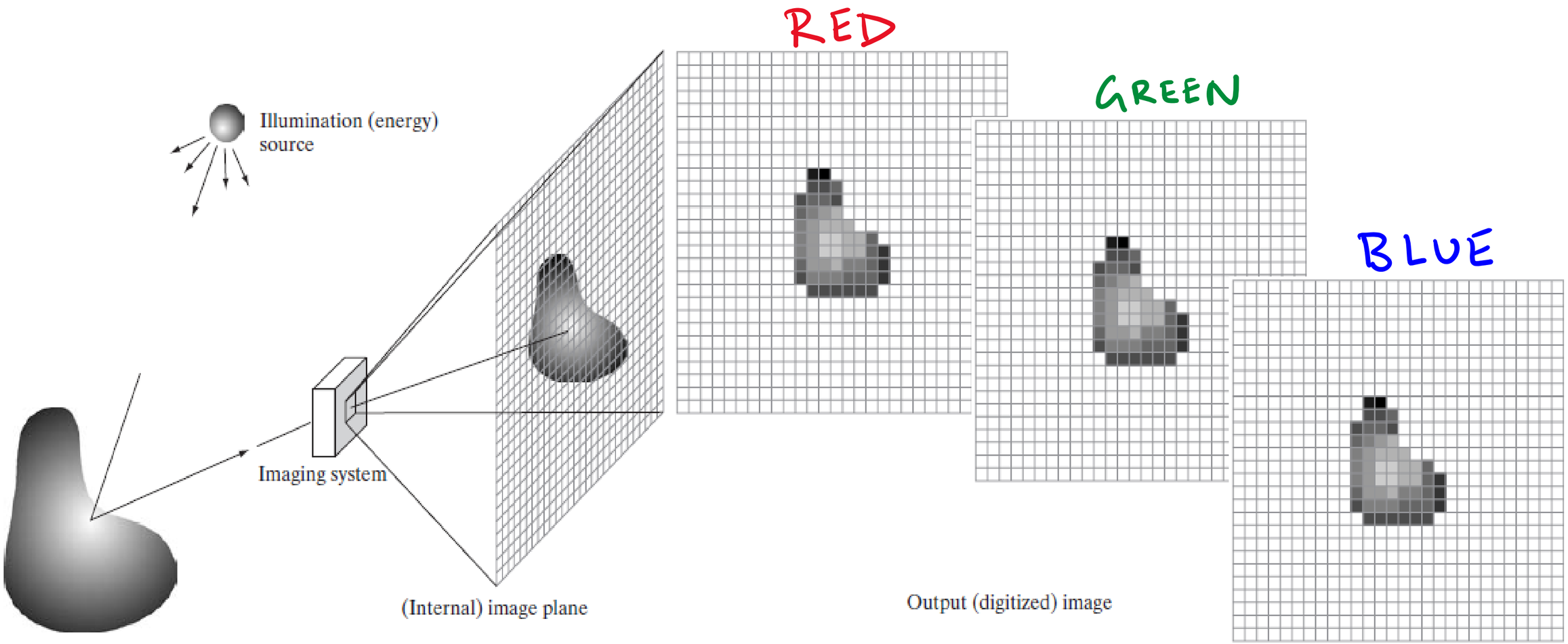
$$= 20 \log_{10} \frac{I_{max}}{\sqrt{RMS}}$$

$I_{max} = 255$ for 8 bit images

↓

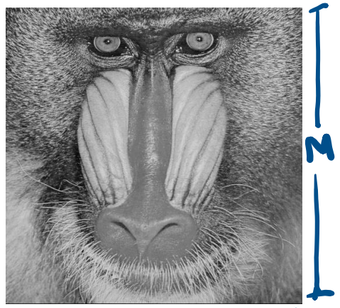
Metric for quantifying Quality of Compression

[ML-4360, Lecture 2]



[Gonzalez and Woods 2008, Page 51]

GRAYSCALE



M

N

$$\begin{bmatrix} 125 & 150 & \cdots & 35 \\ \vdots & \ddots & & \vdots \\ & & 213 & \\ 42 & 104 & \cdots & 220 \\ & 77 & & \end{bmatrix}$$

M

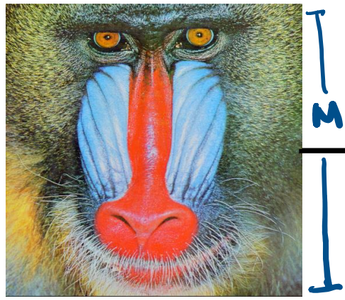
N

RESHAPE or FLATTEN

$p = MN$

$$[125 \ \ldots \ 42 \ \ 150 \ \ldots \ 104 \ \ 77 \ \ldots \ 213 \ \ldots \ 35 \ \ldots \ 220]$$

- $\frac{R+G+B}{3}$ or → "mean" | "average"
- $(0.2989 \times R) + (0.5870 \times G) + (0.1140 \times B)$ or
- $\frac{\max(R,G,B)+\min(R,G,B)}{2}$



M

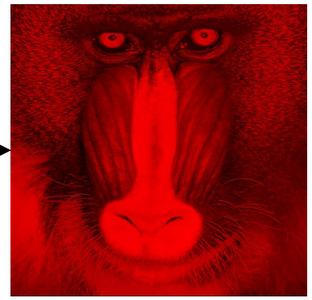
N

COLOUR (RGB)

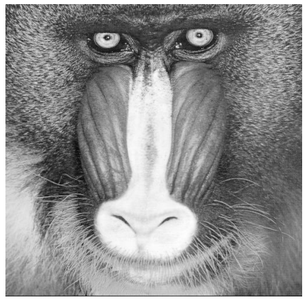
RED (R)

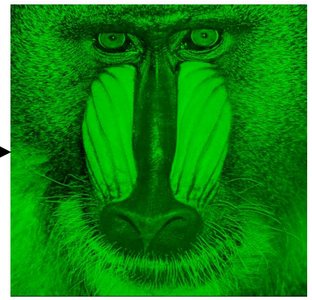
GREEN (G)

BLUE (B)

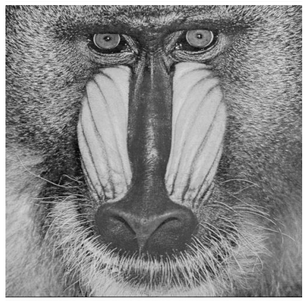


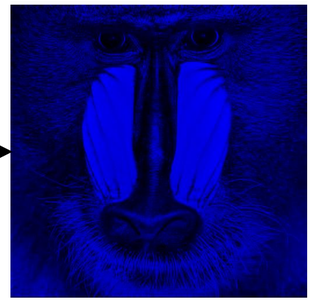
=



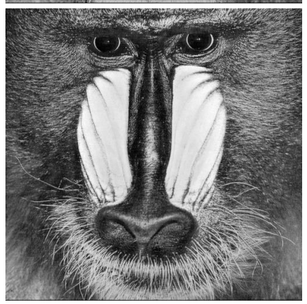


=





=



MN

→ [32 151 ... 222]

MN

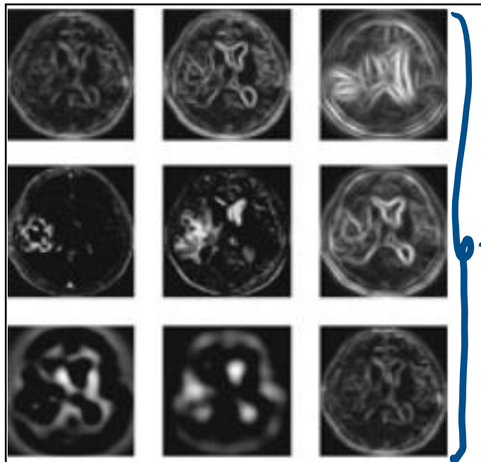
→ [55 112 ... 150]

MN

→ [71 105 ... 210]

→ $p = 3MN$

[32 151 ... 222 55 112 ... 150 71 150 ... 210]



MULTI-CHANNEL

In general, use a direct extension of approaches used for RGB images

May consider keeping ONLY one of the R|G|B channels, typically "G" as it is the LEAST noisy