影像演算法介紹

Lecture Notes: Introduction and Overview

Kai-Lung Hua (花凱龍)
Dept. of Computer Science and Information Engineering
National Taiwan University of Science and Technology

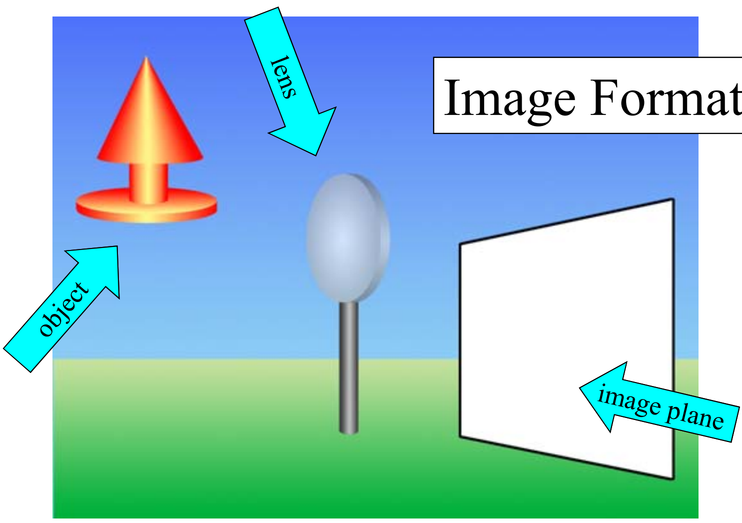
Introduction and Overview

This presentation is an overview of some of the ideas and techniques to be covered during the course.

Topics

1. image formation
2. point processing and equalization
3. color perception and transformation
4. Fourier transform
5. convolution and image filtering
6. frequency filtering
7. noise reduction
8. recent advances in image processing and computer vision

Image Formation



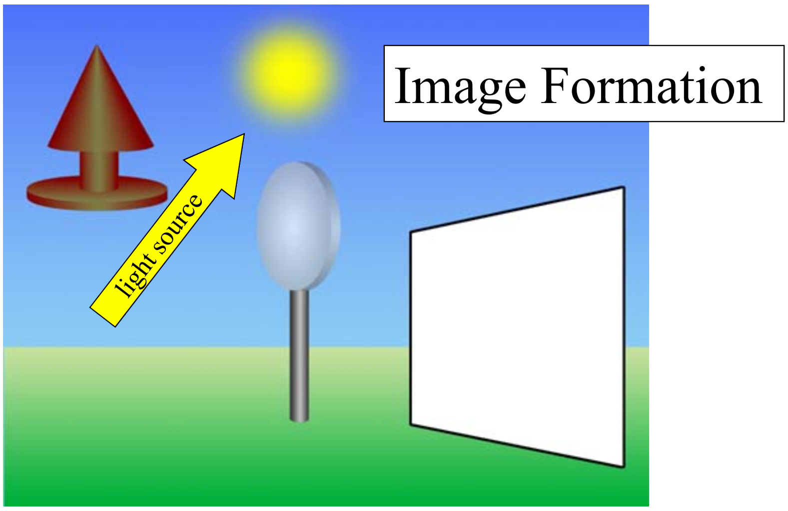


Image Formation
light source

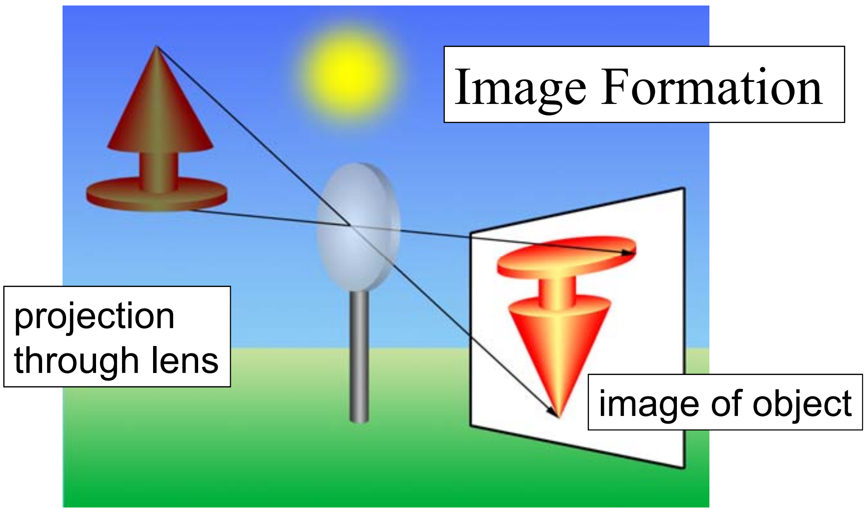


Image Formation
projection through lens
image of object

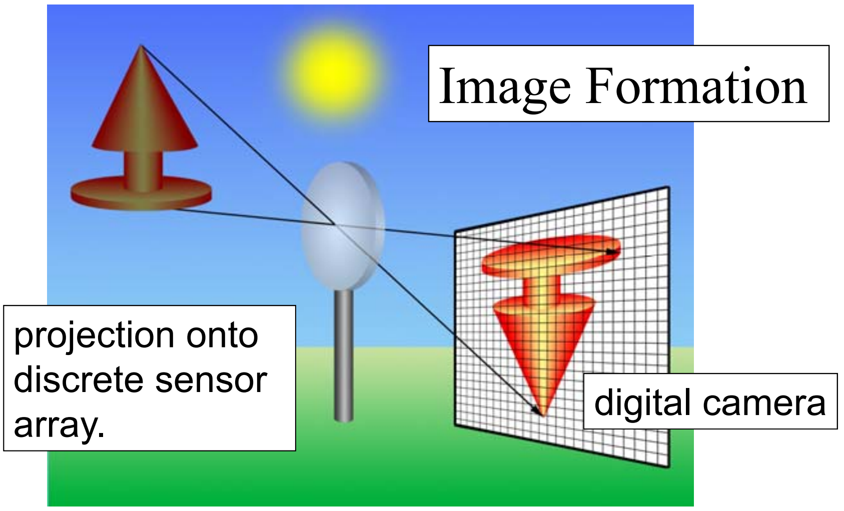


Image Formation
projection onto discrete sensor array.
digital camera

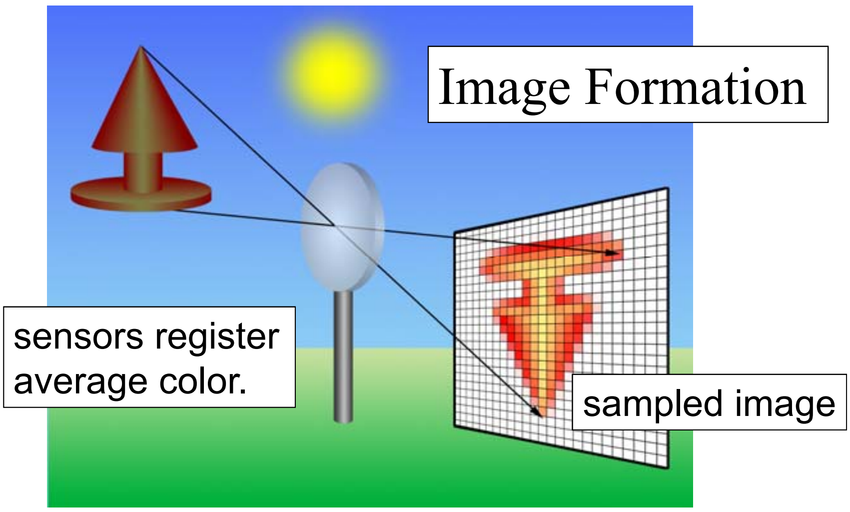
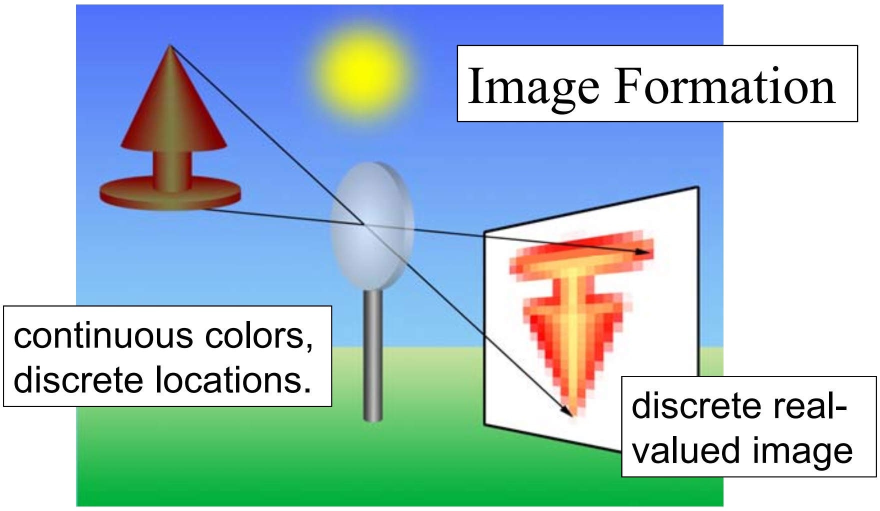
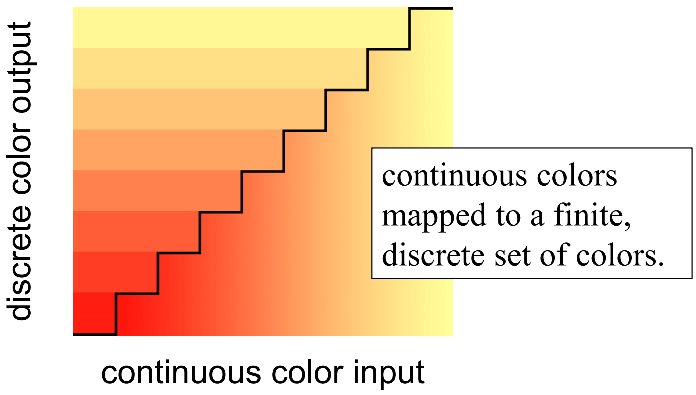


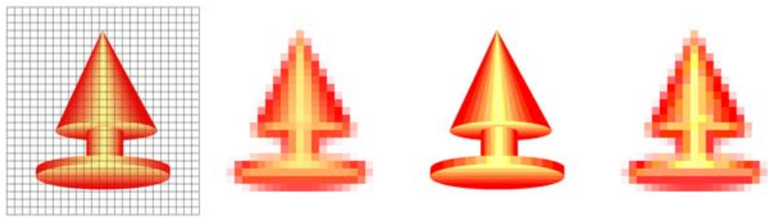
Image Formation
sensors register average color.
sampled image



Digital Image Formation: Quantization



Sampling and Quantization



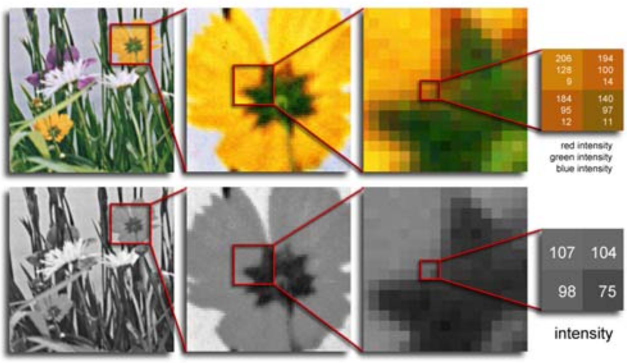
real image | sampled | quantized | sampled & quantized

Digital Image

Color images have 3 values per pixel; monochrome images have 1 value per pixel.

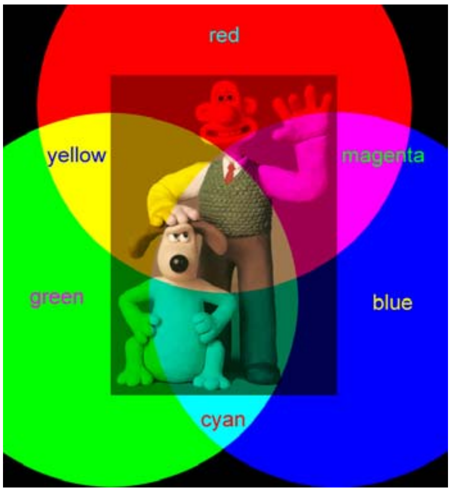
a grid of squares, each of which contains a single color

each square is called a pixel (for *picture element*)

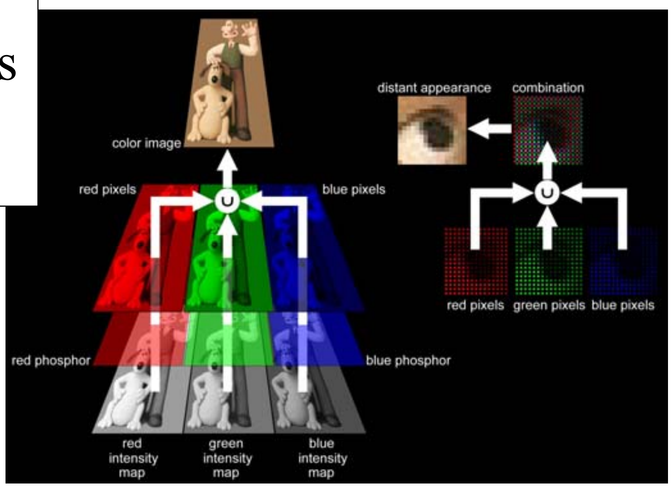


Color Images

- Are constructed from three intensity maps.
- Each intensity map is projected through a color filter (*e.g.,* red, green, or blue, or cyan, magenta, or yellow) to create a monochrome image.
- The intensity maps are overlaid to create a color image.
- Each pixel in a color image is a three element vector.



Color Images On a CRT



Point Processing

- gamma | - brightness | original | + brightness | + gamma

histogram mod | - contrast | original | + contrast | histogram EQ

Color Processing

requires some knowledge of how we see colors

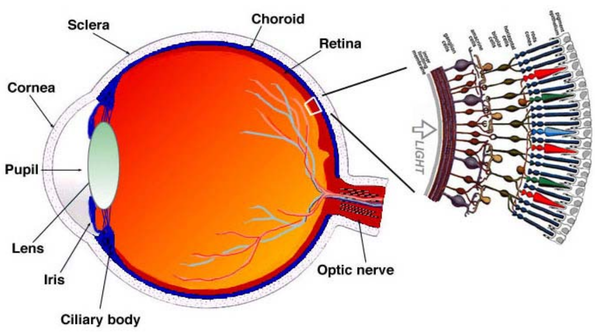
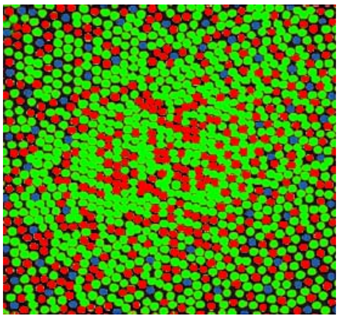


Fig. 1.1. A drawing of a section through the human eye with a schematic enlargement of the retina.

Eye’s Light Sensors

cone density near fovea



#(blue) << #(red) < #(green)

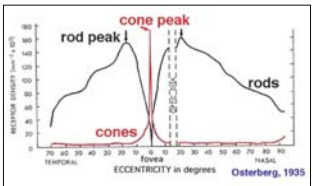


Fig. 20. Graph to show rod and cone densities along the horizontal meridian.

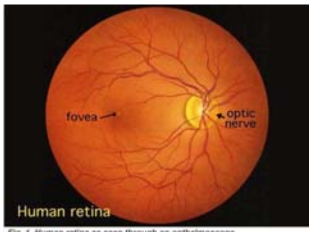
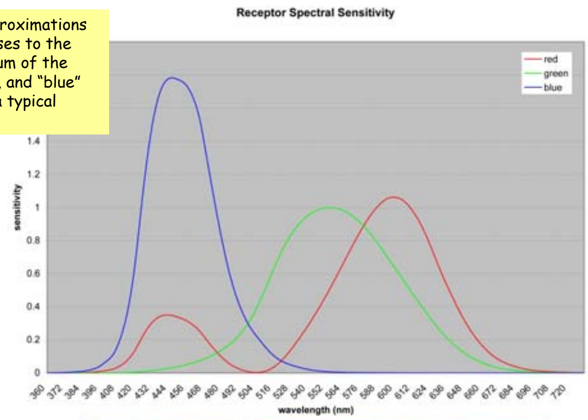


Fig. 1. Human retina as seen through an opthalmoscope.

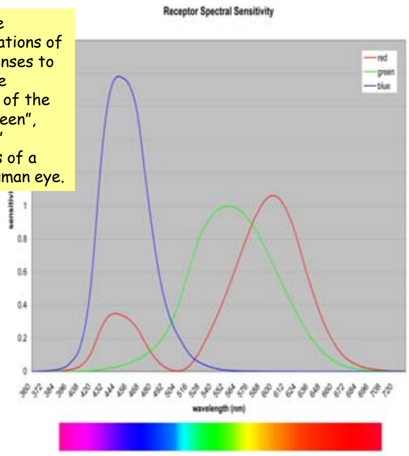
Color Sensing / Color Perception

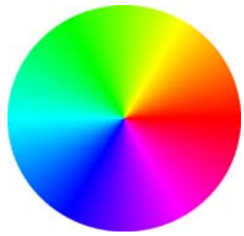
These are approximations of the responses to the visible spectrum of the “red”, “green”, and “blue” receptors of a typical human eye.



Color Sensing / Color Perception

These are approximations of the responses to the visible spectrum of the "red", "green", and "blue" receptors of a typical human eye.

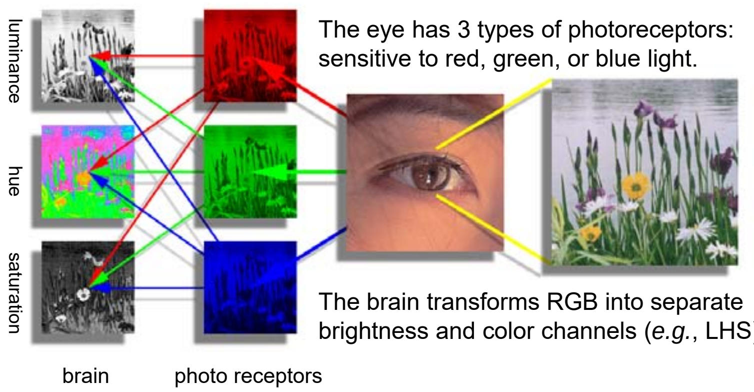




The simultaneous red + blue response causes us to perceive a continuous range of hues on a circle. No hue is greater than or less than any other hue.

Color Sensing / Color Perception

The eye has 3 types of photoreceptors: sensitive to red, green, or blue light.

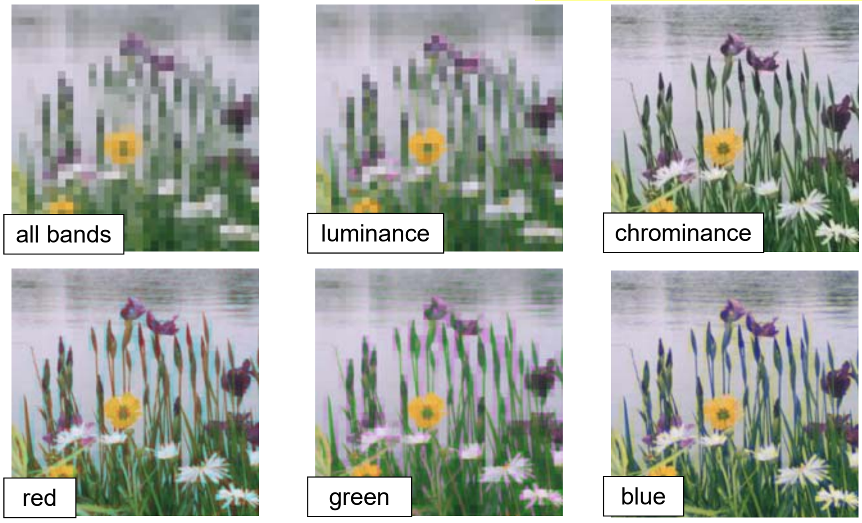


The brain transforms RGB into separate brightness and color channels (*e.g.*, LHS).

Color Perception

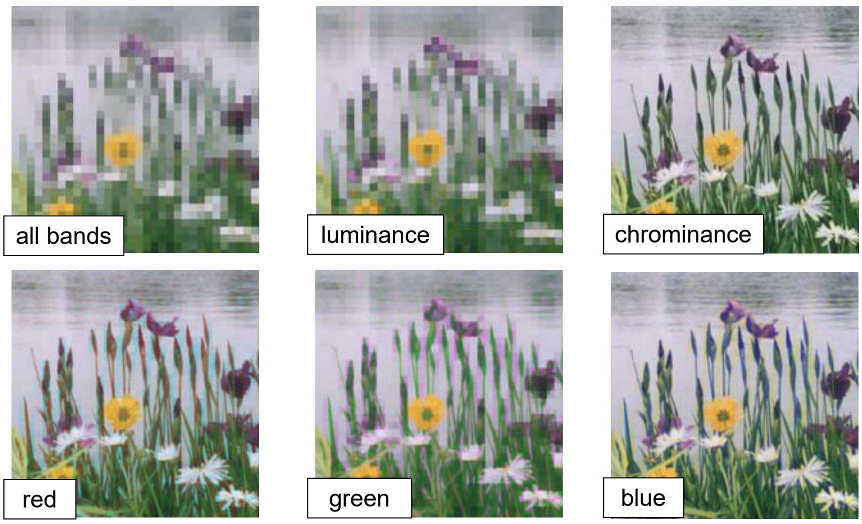
16x pixelization of:

luminance and chrominance (hue+saturation) are perceived with different resolutions, as are red, green and blue.



Color Perception

16x pixelization of:



Color Balance and Saturation

Uniform changes in color components result in change of tint.

E.g., if all G pixel values are multiplied by $\alpha > 1$ then the image takes a green cast.

Color Transformations

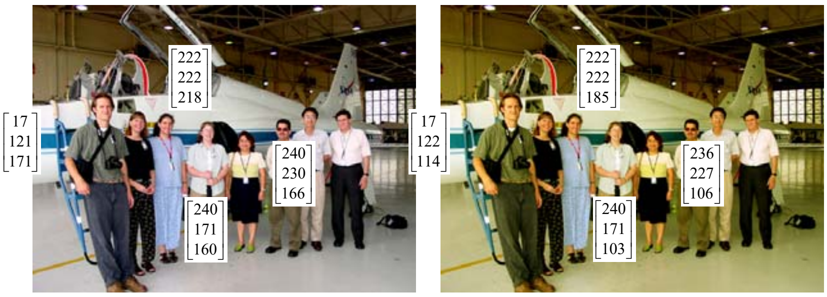


Image aging: a transformation, Φ, that mapped:

$$\begin{bmatrix}17\\122\\114\end{bmatrix} = \Phi\left\{\begin{bmatrix}17\\121\\171\end{bmatrix}\right\} \qquad \begin{bmatrix}222\\222\\185\end{bmatrix} = \Phi\left\{\begin{bmatrix}222\\222\\218\end{bmatrix}\right\} \qquad \begin{bmatrix}240\\171\\103\end{bmatrix} = \Phi\left\{\begin{bmatrix}240\\171\\160\end{bmatrix}\right\} \qquad \begin{bmatrix}236\\227\\106\end{bmatrix} = \Phi\left\{\begin{bmatrix}240\\230\\166\end{bmatrix}\right\}$$

The 2D Fourier Transform of a Digital Image

Let $I(r,c)$ be a single-band (intensity) digital image with R rows and C columns. Then, $I(r,c)$ has Fourier representation

$$I(r,c) = \sum_{u=0}^{R-1} \sum_{v=0}^{C-1} \boldsymbol{I}(u,v)\, e^{+i2\pi\left(\frac{ur}{R}+\frac{vc}{C}\right)},$$

where

$$\boldsymbol{I}(u,v) = \frac{1}{RC} \sum_{r=0}^{R-1} \sum_{c=0}^{C-1} I(r,c)\, e^{-i2\pi\left(\frac{ur}{R}+\frac{vc}{C}\right)}$$

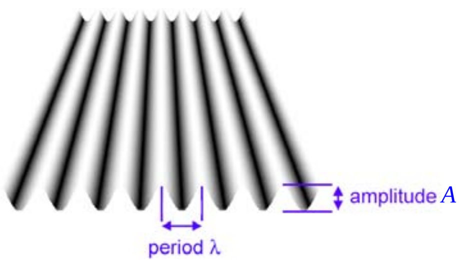
are the R x C Fourier coefficients.

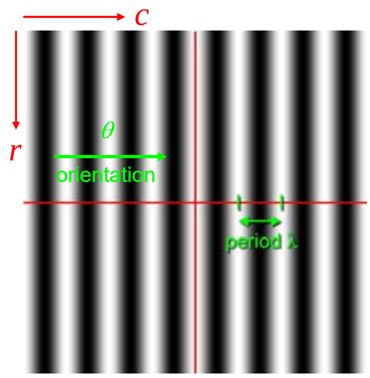
these complex exponentials are 2D sinusoids.

2D Sinusoids:

$$I(r,c) = \frac{A}{2}\left\{\cos\left[\frac{2\pi}{\lambda}\left(\frac{c}{C}\cos\theta - \frac{r}{R}\sin\theta\right)+\phi\right]+1\right\}$$

... are plane waves with grayscale amplitudes, periods in terms of lengths, ...

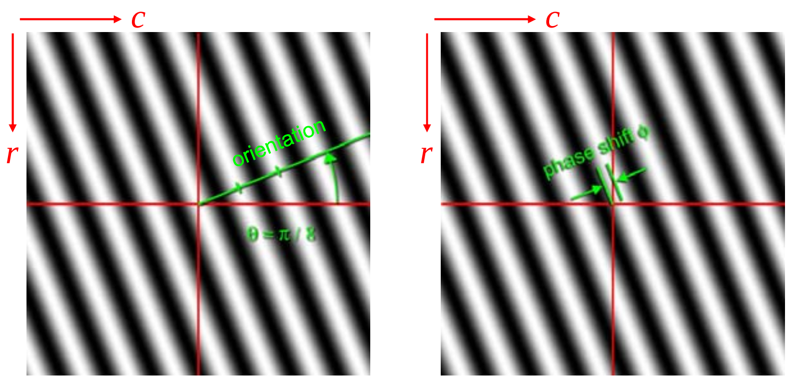




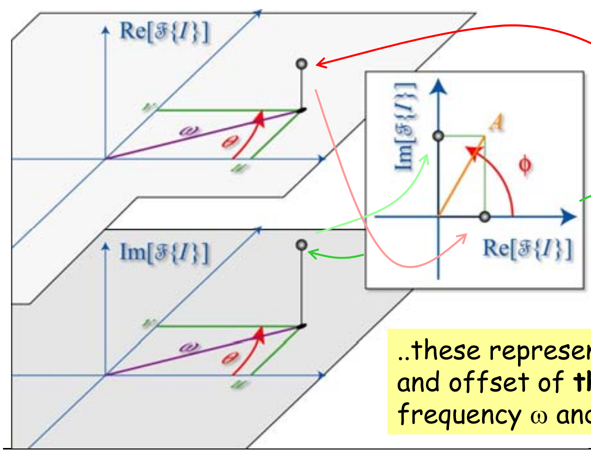
ϕ = phase shift

2D Sinusoids:

... specific orientations, and phase shifts.



The Value of a Fourier Coefficient …

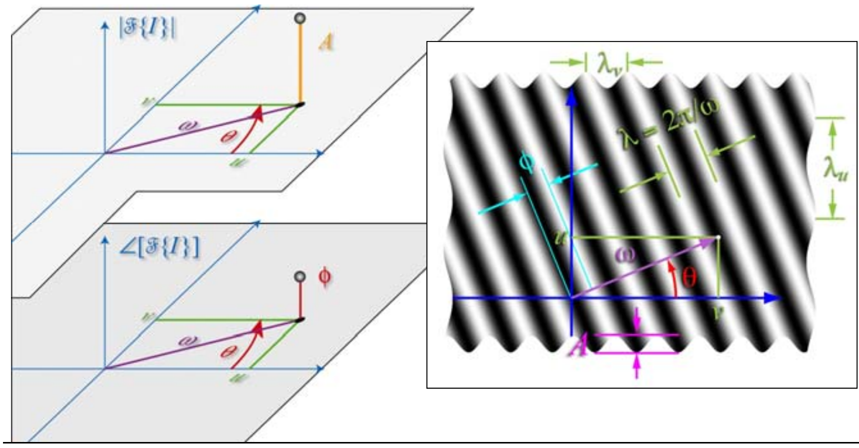


… is a complex number with a real part and an imaginary part.

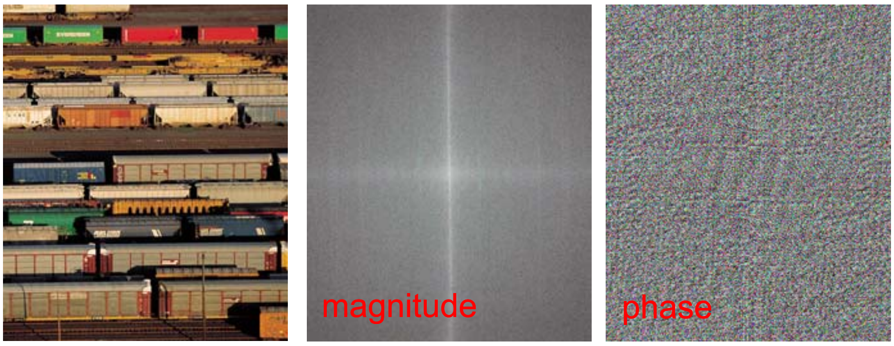
If you represent that number as a magnitude, A, and a phase, ϕ, …

..these represent the amplitude and offset of **the** sinusoid with frequency ω and direction θ.

The Sinusoid from the Fourier Coeff. at (u,v)



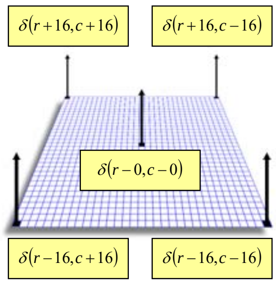
The Fourier Transform of an Image



I $|\mathcal{G}\{I\}|$ $\angle[\mathcal{G}\{I\}]$

Convolution

Sums of shifted and weighted copies of images or Fourier transforms.



Sum times 1/5

Convolution Property of the Fourier Transform

Let functions $f(r,c)$ and $g(r,c)$ have Fourier Transforms $F(u,v)$ and $G(u,v)$. Then,

$$\mathbf{F}\{f * g\} = F \cdot G.$$

Moreover,

$$\mathbf{F}\{f \cdot g\} = F * G.$$

$*$ represents convolution

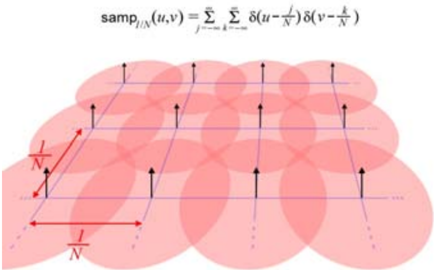
$\cdot$ represents pointwise multiplication

Then, a spatial convolution can be computed by

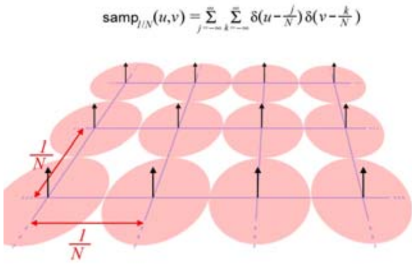
$$f * g = \mathbf{F}^{-1}\{F \cdot G\}.$$

The Fourier Transform of a product equals the convolution of the Fourier Transforms. Similarly, the Fourier Transform of a convolution is the product of the Fourier Transforms

Sampling, Aliasing, & Frequency Convolution

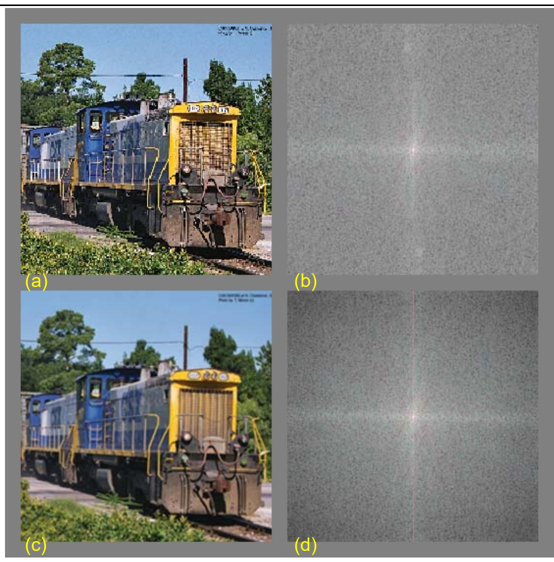


aliasing (the jaggies)



no aliasing (smooth lines)

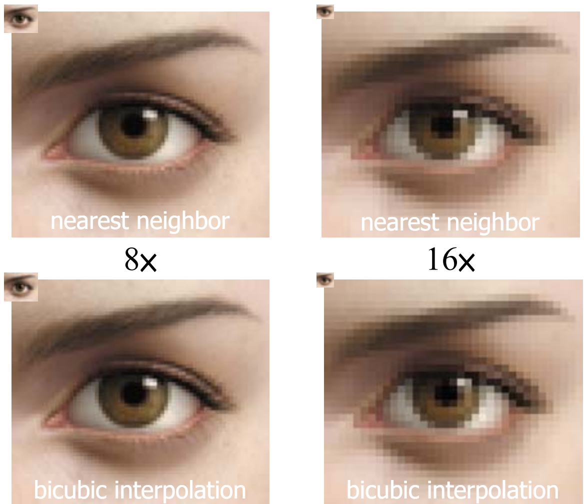
Sampling, Aliasing, & Frequency Convolution



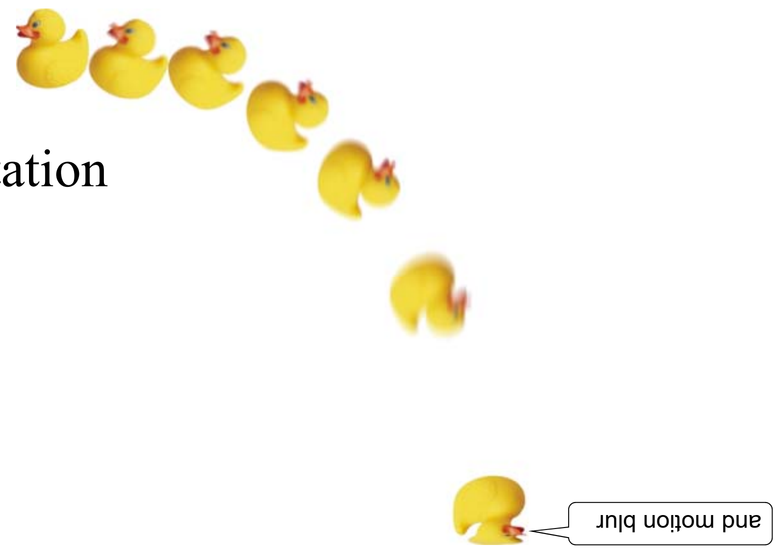
(a) aliased
(b) power spectrum
(c) unaliased
(d) power spectrum

Resampling

(resizing)



Rotation

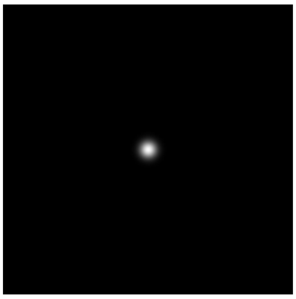


Frequency Domain (FD) Filtering

Image size: 512x512
SD filter sigma = 8

Original Image

Power Spectrum

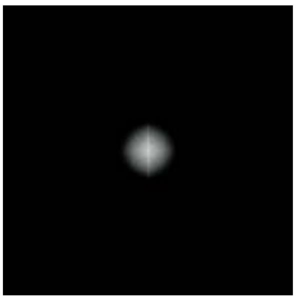


Gaussian LPF in FD

FD Filtering: Lowpass

Image size: 512x512
SD filter sigma = 8

Filtered Image

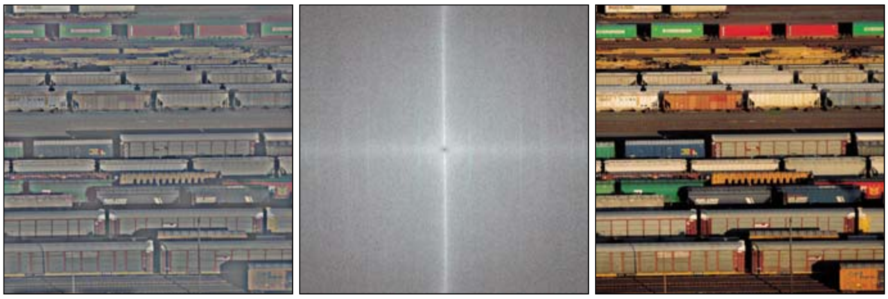


Filtered Power Spectrum

Original Image

FD Filtering: Highpass

Image size: 512x512
FD notch sigma = 8

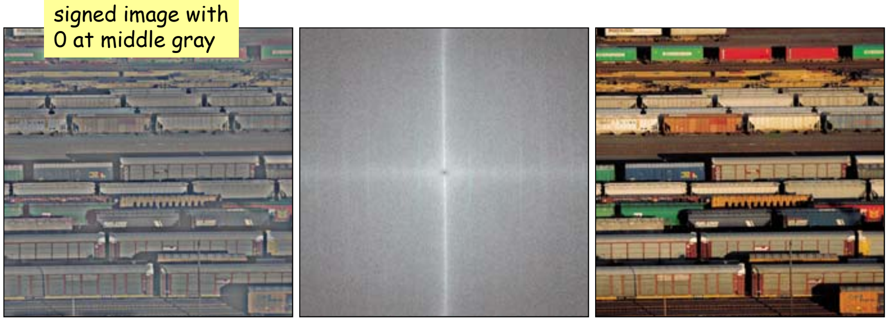


Filtered Image | Filtered Power Spectrum | Original Image

FD Filtering: Highpass

Image size: 512x512
FD notch sigma = 8

signed image with
0 at middle gray



Filtered Image | Filtered Power Spectrum | Original Image

Spatial Filtering

blurred

original

sharpened

Spatial Filtering

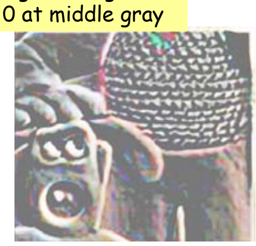
bandpass filter

original

unsharp masking

Spatial Filtering

signed image with
0 at middle gray



bandpass filter

original

unsharp masking

Motion Blur

regional

vertical

original

zoom

rotational

Noise Reduction

blurred image

color noise

color-only blur

Noise Reduction

blurred image

color noise

5x5 Wiener filter

Noise Reduction

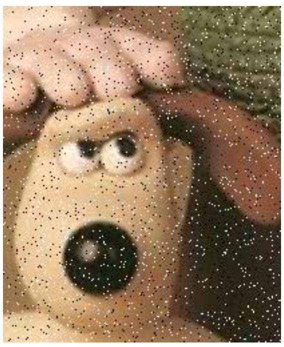
periodic noise

original

frequency tuned filter

Shot Noise or Salt & Pepper Noise

+ shot noise

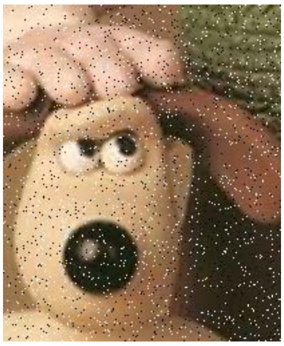


s&p noise

- shot noise

Nonlinear Filters: the Median

original



s&p noise

median filter

Nonlinear Filters: Min and Maxmin

+ shot noise

min filter

maxmin filter

Nonlinear Filters: Max and Minmax

- shot noise

max filter

minmax

Image Compression

Original image is 5244w x 4716h @ 1200 ppi: 127MBytes

Yoyogi Park, Tokyo, October 1999. Photo by Alan Peters.

Image Compression: JPEG

Image Compression: JPEG