AMRITA VISHWA VIDYAPEETHAM
DEEMED TO BE UNIVERSITY

20CYS212
Multimedia Processing
L01-L02 Feb 15, 2023

Introduction [Ch 01]
Digital Image Fundamentals [Ch 02]

Amritha P. P., Assistant Professor
Department of Cybersecurity

Amit Agarwal, Professor
Department of Electrical & Electronics Engineering | Department of Cybersecurity

Course Administration & Project [1/20]

L-T-P-C: 2-0-3-3
L Lecture | **T** Tutorial | **P** Practical | **C** Credit

Faculty

Amritha P. P. for Matlab, Text, Audio & Video Data & **Unit 3**

Basic Image compression methods: Simple coding schemes, Frequency based coding - Huffman coding, Relative encoding, Run length encoding, LZW compression - Image and video compression standards -MJPEG, MPEG2, MPEG4, H.264, H.26. Color image processing.

Amit Agarwal for the rest

Contact

+91 80563 75912 pp_amritha@cb.amrita.edu
+91 98679 10690 a_amit@cb.amrita.edu

Office Hours: Any day. Appointment only. Please message proposed:

Date	Time
Duration	Agenda

Course Administration & Project [2/20]

Syllabus

Introduction **[Ch 01]**

Digital Image Fundamentals **[Ch02]**

Intensity Transformations & Spatial Filtering **[Ch03]**

Filtering in the Frequency Domain **[Ch04]**

Wavelet and Other Image Transforms **[Ch07]**

Image Compression & Watermarking **[Ch08]**

Morphological Image Processing **[Ch09]**

Image Segmentation … *up to Superpixels* **[Ch10]**

Feature Extraction … *up to Region Feature Descriptors* if time permits **[Ch11]**

Image Pattern Classification … *up to Optimum (Bayes) Statistical Classifiers* if time permits **[Ch12]**

Book: Rafael C. Gonzalez & Richard E. Woods, *Digital Image Processing*, 4^{th} Edn,

Course Administration & Project [3/20]

Evaluation

1. *Periodical 1* **[10]** mainly multiple choice
2. *Periodical 2* **[10]** mainly multiple choice
3. *Continuous Assessment Theory*: **[15]** Group Project [max 3 ppl/group]
4. *Continuous Assessment Lab*: **[30]** Same Group Project as Above
5. Each person is expected to spend a minimum 40 hours on the project.
6. Feb 23 onwards, every Thu, every group will present ideas/challenges/next week plan on the project
7. *Final exam* **[35]** mainly multiple choice
8. No trick questions. No negative marking. ALL memorization-based formulae will be given in Q paper.

Attendance

1. Must be > 75% to be permitted to sit for final exam.
2. Lectures will not be repeated. If you miss then ask your friends what was covered.
3. Do not miss many classes - passing the course may turn hard.
4. Read regularly. Course is easy. But the material keeps building on what was already covered.

Course Administration & Project [4/20]

Eating & Drinking During Lectures

a. You can as long as you clean up the area before the end of the class.

b. Do not eat things that make sound.

Phones, Net Surfing & Game Playing

a. All mobile devices on silent/vibration or airplane mode during the lecture

b. Do not use in the class unless needed for the lecture.

c. If you really need to call or take a call then step outside of the class before speaking

Boredom

d. If you are bored then leave or feel free to sleep as long as you do not snore.

Copying

a. Getting caught for copying may trigger having to repeat the course in the following semester.
b. Feel free to discuss and google up project solutions on the internet.

Course Administration & Project [5/20]

Project 1: Detection and Classification of COTS Drones using Visible-band Signatures

Datasets

https://cord.cranfield.ac.uk/articles/dataset/Synthetic_Drone_Classification_Dataset/19423925

https://www.kaggle.com/dasmehdixtr/drone-dataset-uav

Papers

https://www.mdpi.com/2313-433X/8/8/218

https://www.mdpi.com/2226-4310/9/1/31

https://ieeexplore.ieee.org/document/8539442

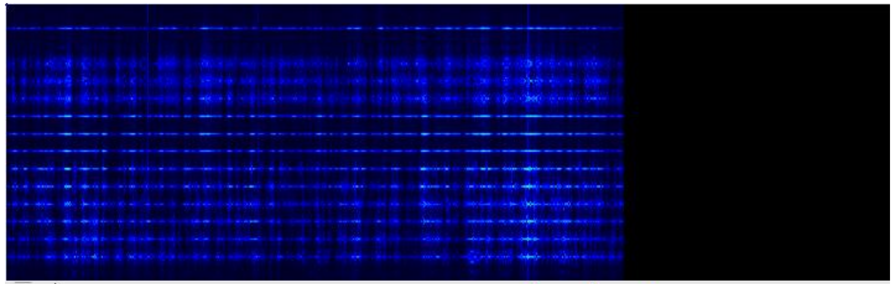
https://ieeexplore.ieee.org/document/9287909

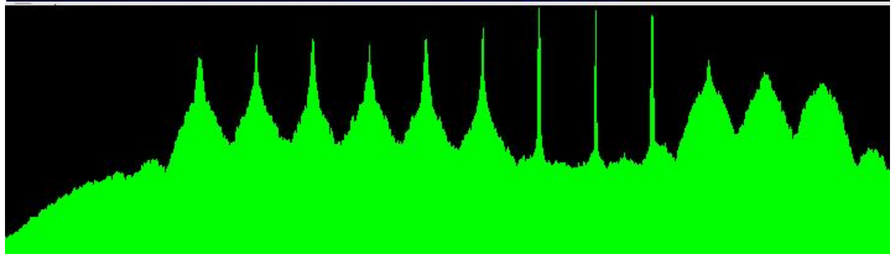
Course Administration & Project [6/20]

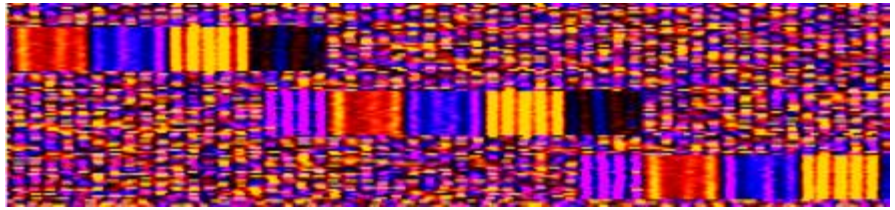
Project 2: High Frequency Radio Signal Analysis

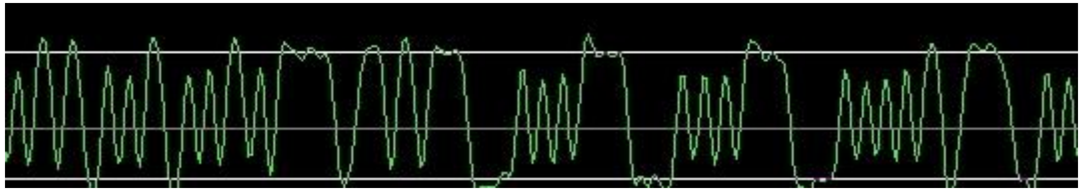
Datasets: We have plenty

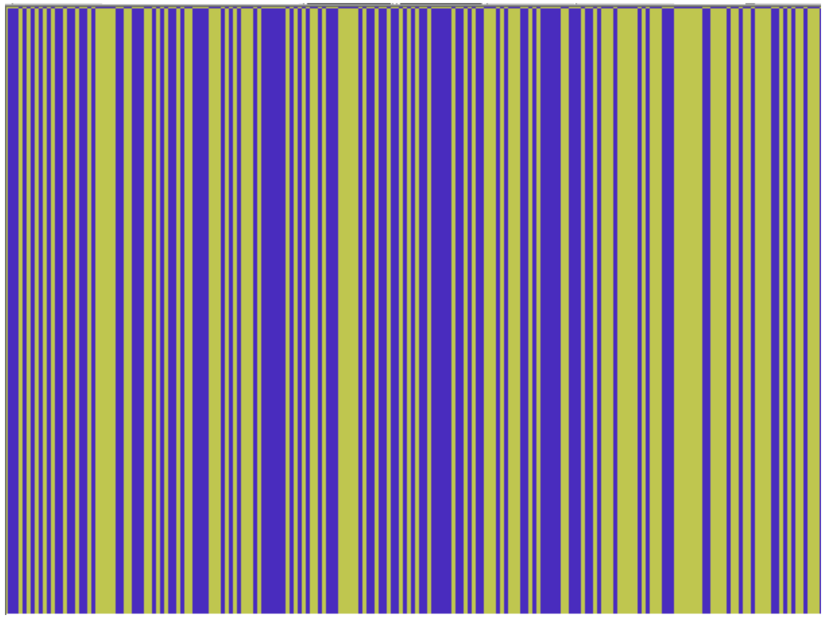
Papers: Plenty

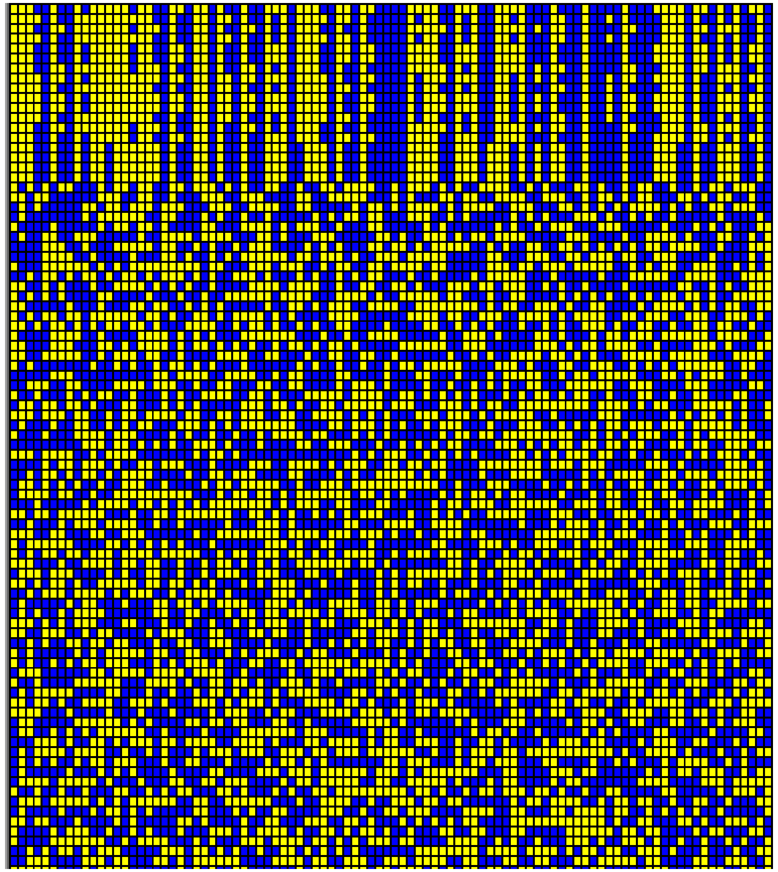


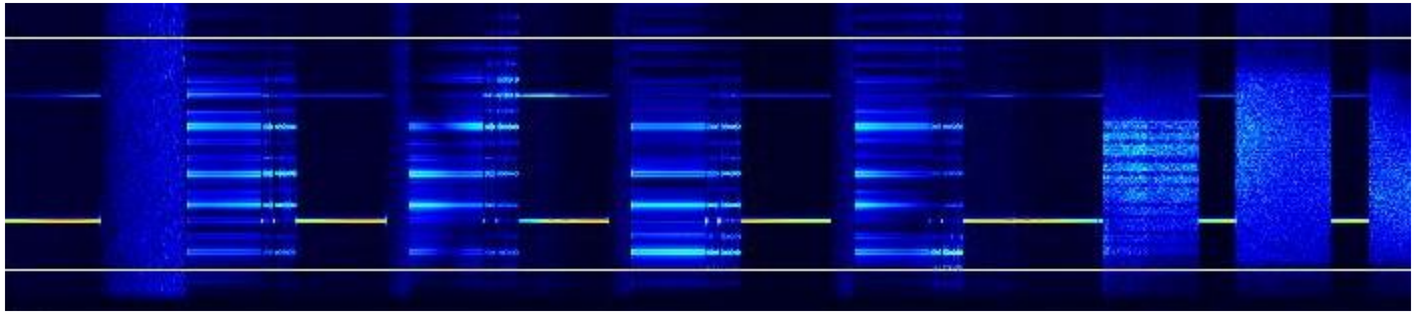












Course Administration & Project [7/20]

Project 3: Detection and Classification of COTS Drones using RF Signatures

Datasets

Papers

Course Administration & Project [8/20]

Project 4: Segregating COTS Drone RF Signal into FPV Video, Telemetry & Control

Datasets

Papers

Course Administration & Project [9/20]

Project 5: Visual Navigation of Drones in GPS-denied Urban Environments

Datasets

Papers

Course Administration & Project [10/20]

Project 6: Avoiding Drone Cable Strikes using RGB/IR Sensors

Datasets

Papers

Course Administration & Project [11/20]

Project 7: Urban Feature Classification from Drone Data

Datasets

Papers

Course Administration & Project [12/20]

Project 8: Apple Disease Detection from Leaf Images

Datasets

Papers

Course Administration & Project [13/20]

Project 9: Automated Apple Counting in Open Field Conditions

Datasets

Papers

Course Administration & Project [14/20]

Project 10: Automated Fuji Apple Size Estimation in Open Field Conditions

Datasets

Papers

Course Administration & Project [15/20]

Project 11: Ship Detection & Classification from Aerial Images

Datasets

Papers

1 Introduction

2 Digital Image Fundamentals

Introduction [1/13]

Image Processing Applications:

1) Improve pictorial representations for human/machine interpretation
2) Compression for storage, transmission and display
3) Steganography
4) Probabilistically secure watermarking

An image is a 2D function $f(x, y)$ where x & y are spatial coordinates and f is a mapping from (x, y) to color space (e.g., R, G, B intensity levels).

Pixel is an indivisible element of the (x, y) space.

Low-level processing: noise, contrast enhancement, and image sharpening

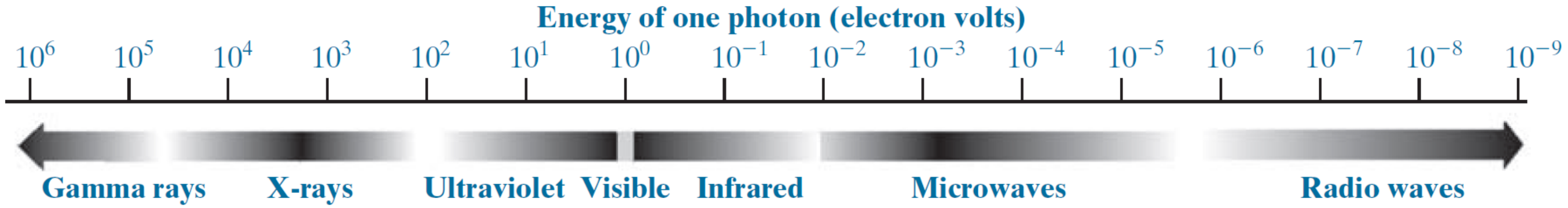
Mid-level processing: segmentation: intermediate outputs are image attributes – edges, contours; final outputs are regions/objects

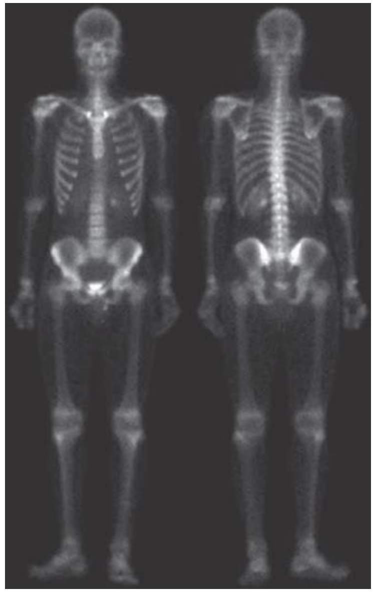
Higher-level processing: making sense of ensemble of objects or cognitive acts

Introduction [2/13]

History: Bartlane picture transmission system in the early 1920s reduced the time required to transport a picture across the Atlantic from more than a week to less than three hours.

It encodes in 5 grey levels; increased to 15 in 1929. Bottom-right photo is better due to higher resolution & tonal levels in just a 1 year.





Gamma-Ray Imaging is used in nuclear medicine & astronomy. The figure on left shows an image of a complete bone scan (bone scintigraphy) via gamma-ray imaging. Such images are used to check for bone pathology such as infections or tumor.

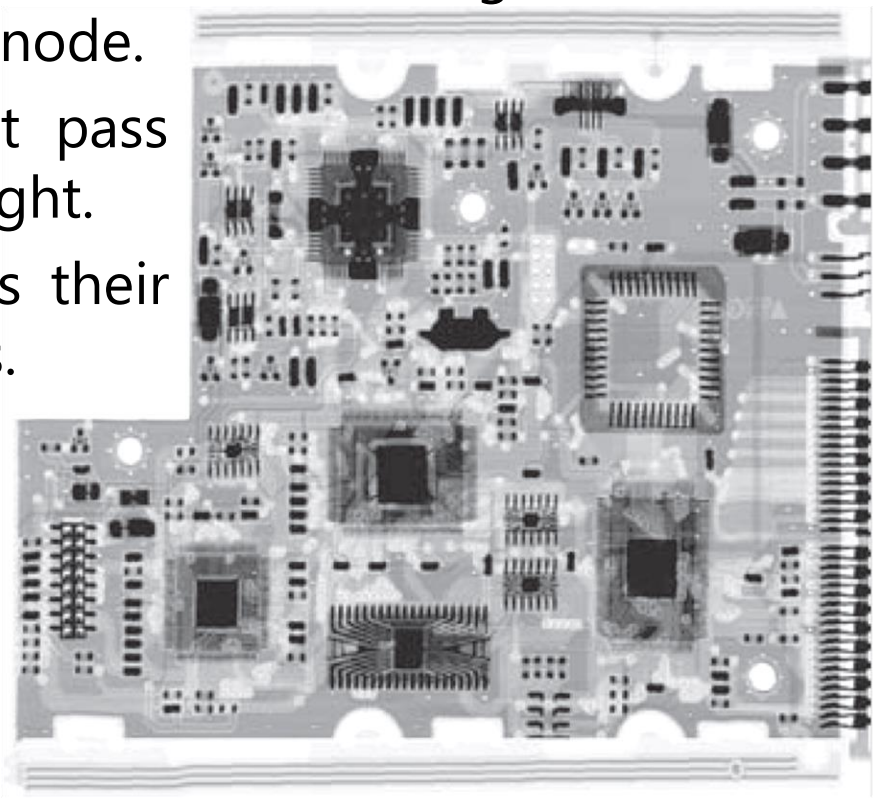
γ vs X-ray imaging: A radio-active nucleotide that emits γ rays is injected. It adsorbs onto the hydroxyapatite mineral of bone which is higher in bone regions with greater perfusion or growth. Thus cancer, fracture or infections are detected.

Introduction [3/13]

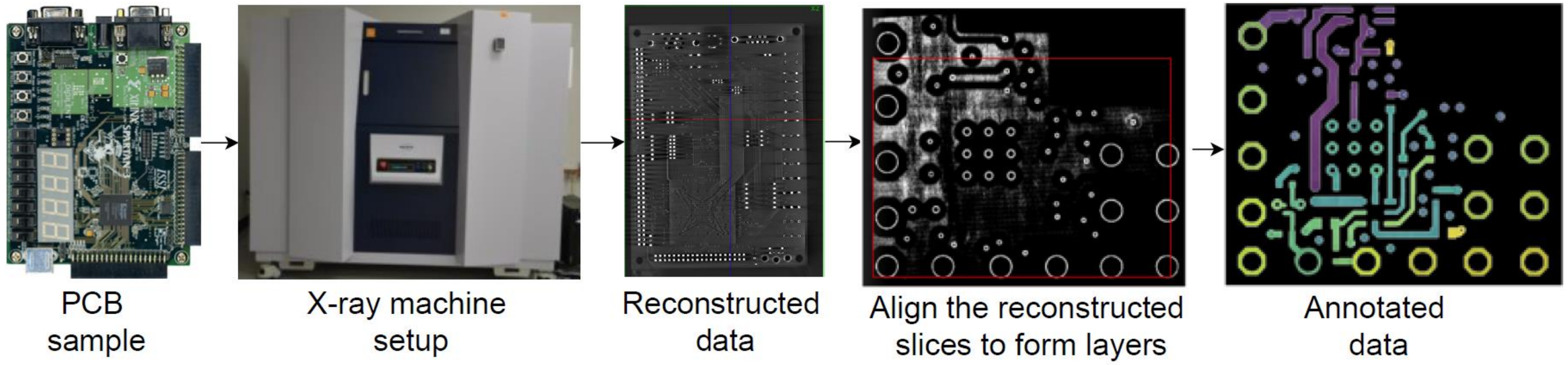
X-ray Imaging: X-rays for medical and industrial imaging are generated using an X-ray tube, which is a vacuum tube with a cathode and anode. The cathode is heated, causing free $\bar{e}$ to be released. These $\bar{e}$ flow at high speed to the positively charged anode.

When e strike a nucleus, X-rays are released. Those that pass through, fall on a phosphor screen. It converts the energy to light.

The figure shows X-ray image of an IC. X-rays are used as their wavelength is < 3 nm vs transistor gap of 3-7 nm in current ICs.



FICS PCB X-ray Dataset: https://gitlab.com/s3a/s3a; ReadMe https://paperswithcode.com/dataset/fics-pcb-image-collection-fpic & paper https://eprint.iacr.org/2022/924.pdf



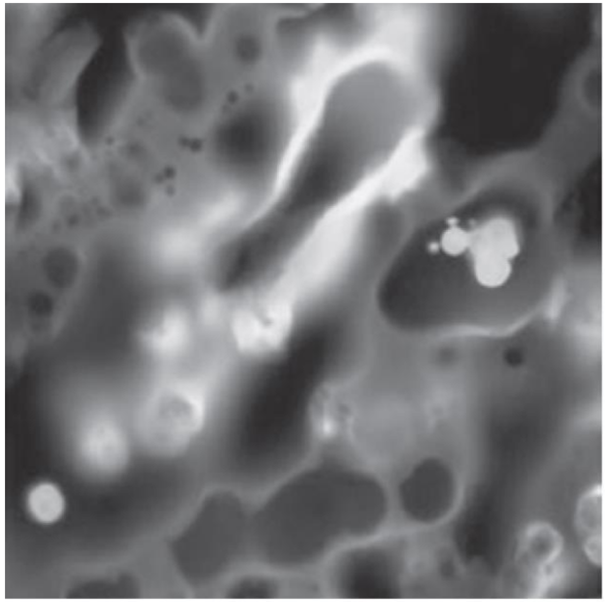
PROCESS OF PCB IC DATASET GENERATION

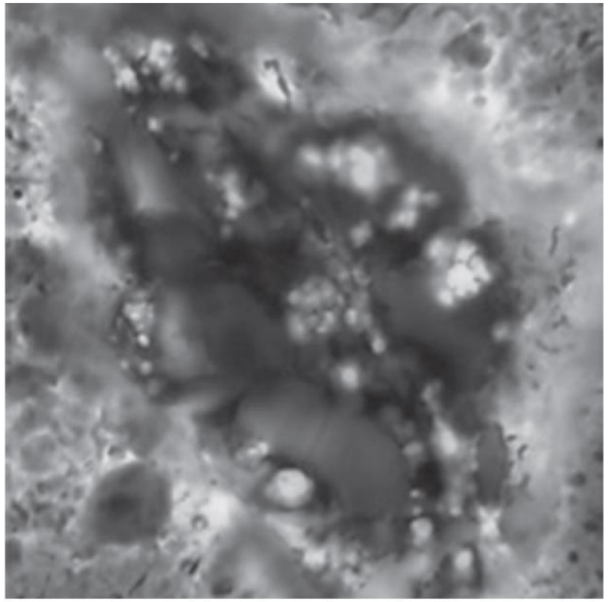
More comprehensive data is at https://www.trust-hub.org/#/data

Introduction [4/13]

UV-ray Imaging I: UV Fluorescence microscopy is used to study materials that can be made to naturally fluoresce (primary fluorescence) or when treated with chemicals capable of fluorescing (secondary fluorescence).

When a UV photon collides with an $\bar{e}$ of a fluorescent material, the $\bar{e}$ elevates to higher energy orbits. Upon relaxation to a lower energy level, it emits photons in the visible (red) region.


NORMAL CORN


SMUT-INFEECTED CORN

As UV wavelength is 100-400 nm & wavelength for visible light is 400-700 nm, magnificent obtained is c175-700% more. However, UV light is not visible. What is visible is the relaxation light with much larger wavelength. Thus, if features of interest are materially separated, then UV microscopy will provide more accurate location then a visible band microscope will.

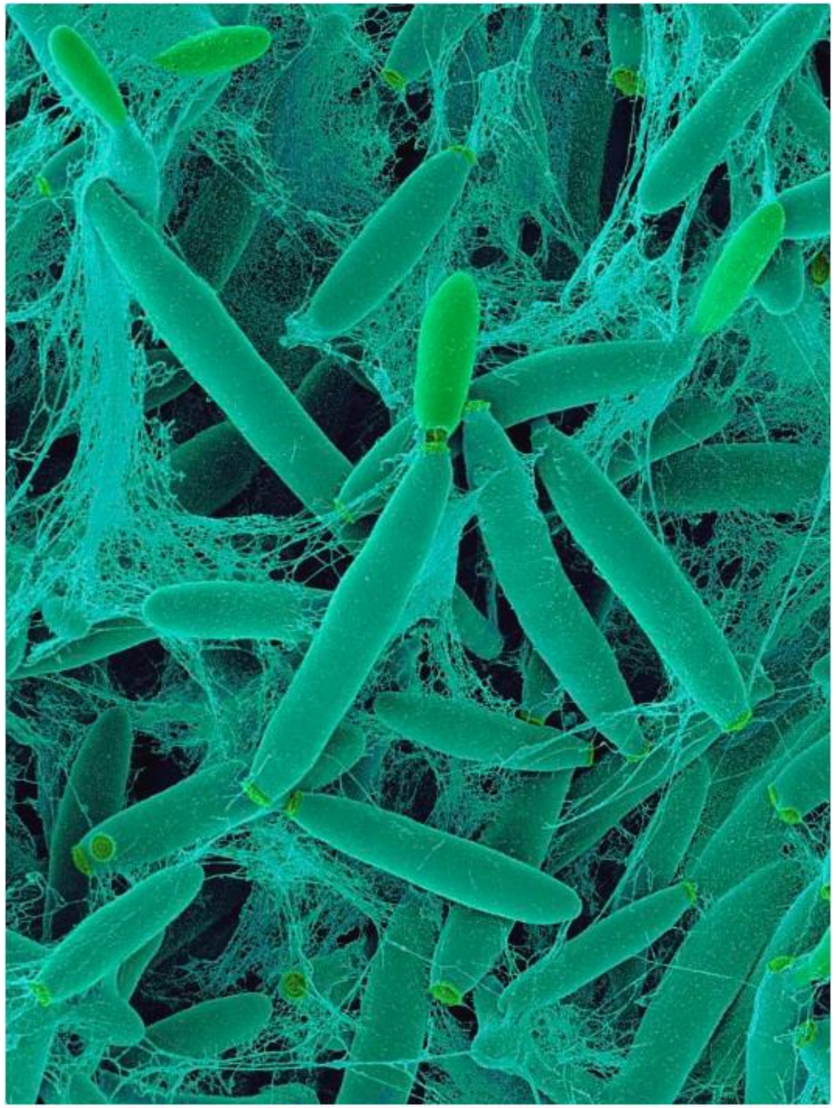
Two pre-processing are needed: i) distinguishing low energy red light from the high energy UV light and ii) estimate high resolution images using low energy visible light

Introduction [5/13]

UV-ray Imaging II: Why use UV Fluorescence microscopy which is 3-4x more expensive than visible light microscope in case of corm smut?

The reason is that i) the substrate fluoresces and ii) actionable detection requires early detection, i.e., when the smut infection has started, but is yet, in early stages, i.e., small in size.

See the infection in early stages in light green and its mucilaginous matrix in the figure on the right.

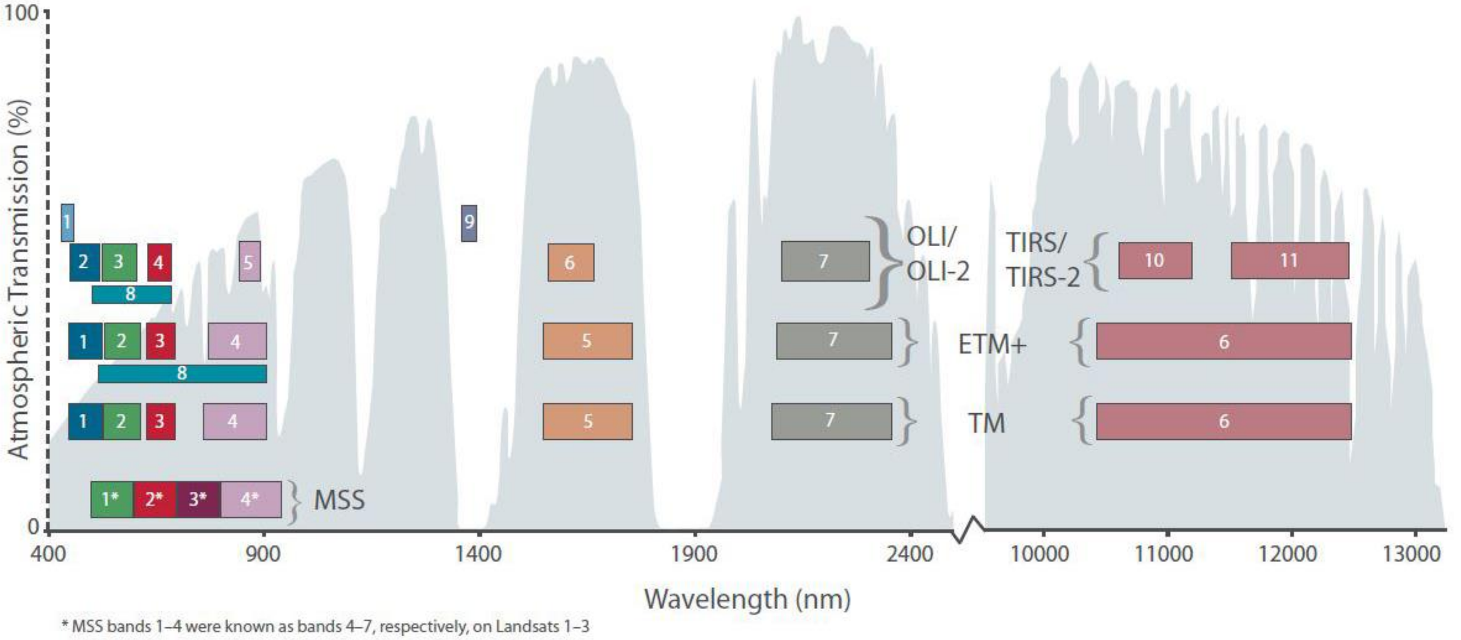


Corn Smut Fungus (ustilago Maydis)

Introduction [6/13]

Visible & IR Imaging: many examples … look at Landsat multispectral imaging satellite. Swathe size is c170 km NS x 183 km EW. What do you infer?

Bands	Wavelength (micrometers)	Resolution (meters)
Band 1 - Coastal aerosol	0.43-0.45	30
Band 2 - Blue	0.45-0.51	30
Band 3 - Green	0.53-0.59	30
Band 4 - Red	0.64-0.67	30
Band 5 - Near Infrared (NIR)	0.85-0.88	30
Band 6 - SWIR 1	1.57-1.65	30
Band 7 - SWIR 2	2.11-2.29	30
Band 8 - Panchromatic	0.50-0.68	15
Band 9 - Cirrus	1.36-1.38	30
Band 10 - Thermal Infrared (TIRS) 1	10.6-11.19	100
Band 11 - Thermal Infrared (TIRS) 2	11.50-12.51	100



Band 1: Coastal & Aerosol studies

Band 2: Bathymetric mapping, distinguishing soil from vegetation and deciduous from coniferous vegetation

Band 3: Assess plant vigor

Band 4: discriminate vegetation classes

Band 5: Biomass content & Shoreline estimation

Band 6: discriminate moisture content of soil & vegetation; penetrates thin clouds

Band 7: similar to Band 6

Band 8: *more spatial but less spectral* resolution

Band 9: Improves Cirrus cloud detection

Band 10/11: Same as Band 6 but, *less spatial & more spectral resolution*

Introduction [7/13]

Microwave Imaging: Can collect Radar Cross Section (RCS) information at both day & night, without much concern about the weather. It can see through vegetation, ice, clouds & dry sand.

Applications: See through walls, Estimate vegetation density, Plan drone paths in forests, Model watersheds, Plan tailing dam sites in forested regions

Radio Imaging I: Radio waves are very low energy waves - useful in medicine and astronomy. E.g. Magnetic Resonance Imaging (MRI): a patient is put in a powerful magnetic field and short radio pulses are passed through her. This excites her tissues to emit their own radio pulses. The strength of the signal and its origination location is used to produce a 2D image.

Compared to CT, MRI yields a better contrast in soft-tissue images of the brain or abdomen. MRI is unsafe for patients with medical implants.

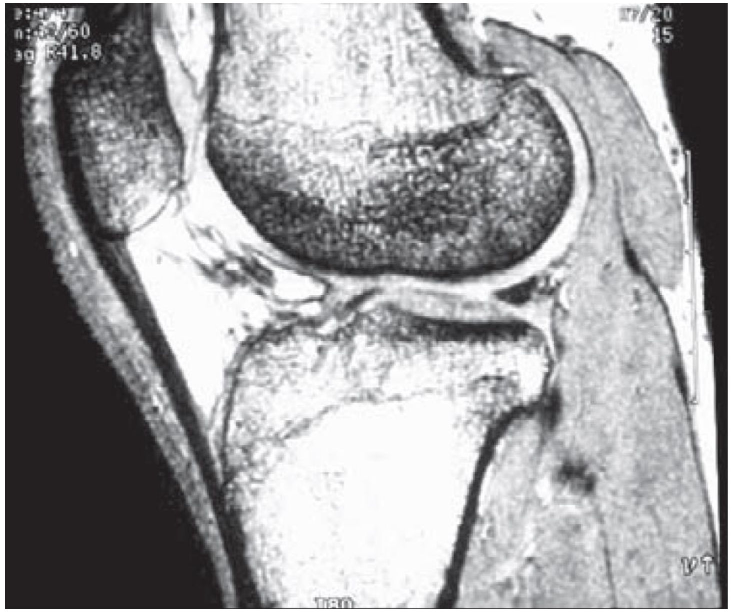
Introduction [8/13]

Radio Imaging II: The wavelength used is >1m. Then how does MRI generate an accurate image of body parts?

Some atomic nuclei can absorb radio waves when placed in a magnetic field. This causes magnetic spin polarization which induces a detectable RF signal.

63% of our body is hydrogen by volume, particularly in water and fat. ∴ most MRI scans map location of water and fat in the body. The radio wave signal induces a magnetic flux gradient which is detectable up to c2-8μm spatial resolution.

The gradient itself depends on tissue properties. Different combination of signal strengths and magnetic field strengths can be used to identify differences between specific tissues.

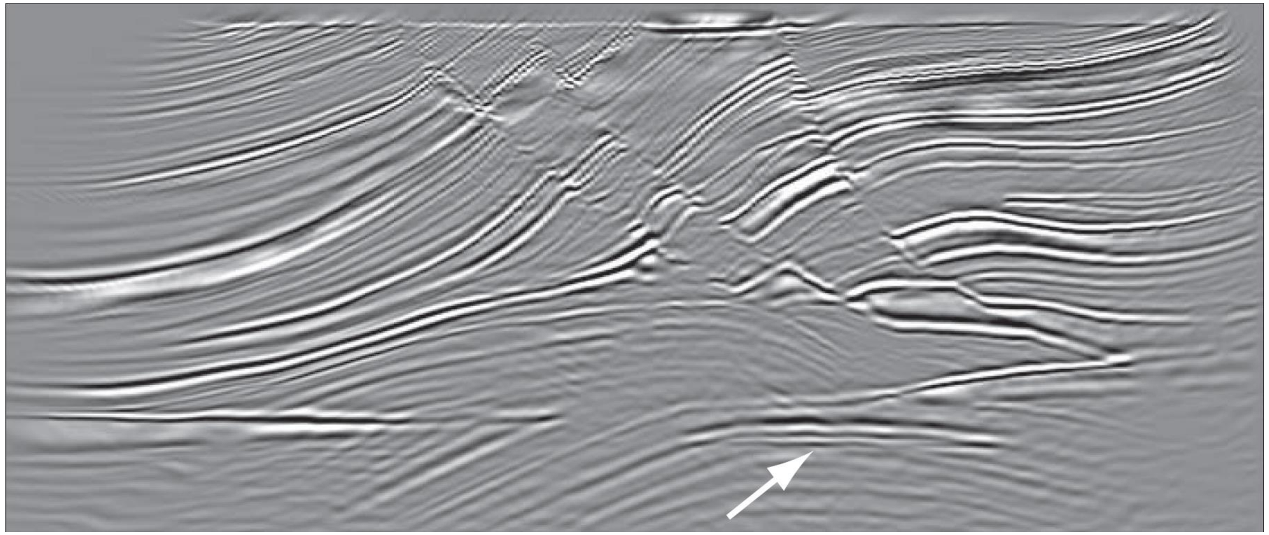


Introduction [9/13]

Sonar Imaging: From a few tens of Hz (at 10 Hz, λ = 34.3m in air or 153m in sea water) to few MHz (at 1 MHz, λ = 343μm in air or 2,000 μm in sea water). Compare this with blue light wavelength which is 450-495 nm. When will you use electromagnetic vs sonar imaging?

Applications: landmine detection, oil exploration, detect submarines, ocean bed mapping, etc.

The figure shows an ultrasonic image that shows presence of oil at a certain depth. This is read through 'bright spots'. The relevant bright spots are distinguished from others by the fact that others do not have a strong gradient vs depth.



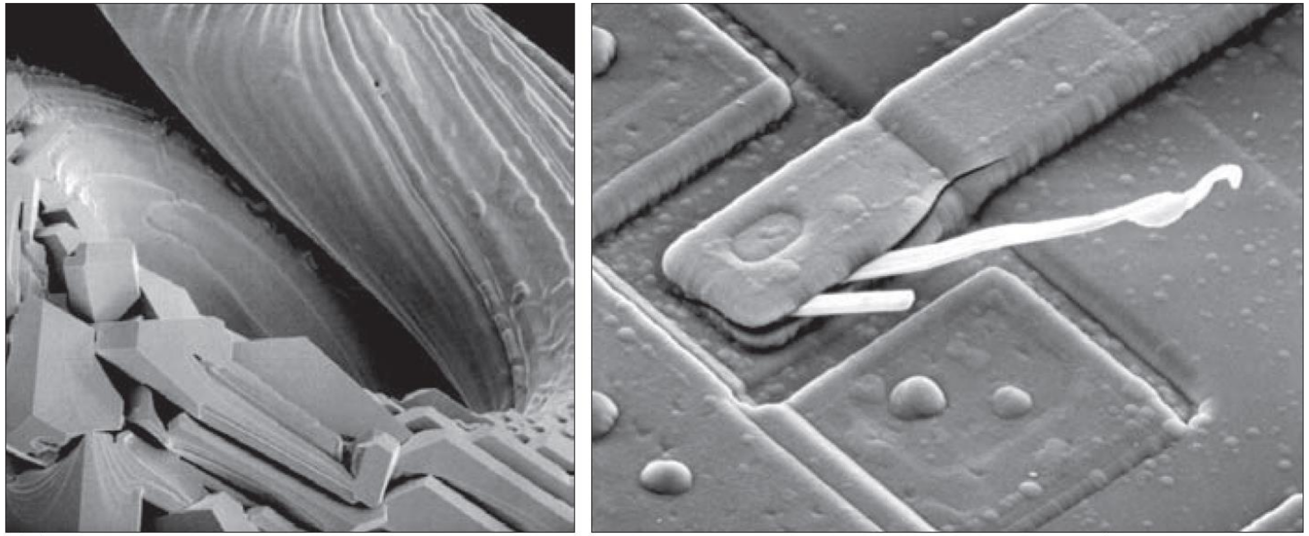
Note that the sound waves may reflect at multiple depths (the component that passes through will reflect with greater delay and will also be more feeble). Also, the reflection intensity will depend on the oil's constitution. Constructing, analyzing and reading ultrasonic oil exploration maps is not easy.

Introduction [10/13]

Transmission $\bar{e}$ Microscope (TEM) vs Scanning $\bar{e}$ Microscope (SEM): A TEM operates like a slide projector – a beam of $\bar{e}$ is transmitted through a specimen and the fraction which is not absorbed or reflected is projected to a phosphor screen. The interaction of $\bar{e}$ with phosphor produces visible light. TEMs can image a thin substrate.

An SEM scans the $\bar{e}$ beam and records interaction of beam and sample at each location. This produces dots on a phosphor screen. The image is formed by a raster scan of the beam. The $\bar{e}$ interact with a phosphor screen and produce light. SEMs can image bulky substrates.

Visible band microscopes yield magnifications c1000x. Electron microscopes have a magnification of 10,000x. The figure shows two SEM images of specimen failures due to thermal overload.



(a) 250 × SEM image of a tungsten filament following thermal failure (note the shattered pieces on the lower left). (b) 2500 × SEM image of a damaged integrated circuit. The white fibers are oxides resulting from thermal destruction.

Introduction [11/13]

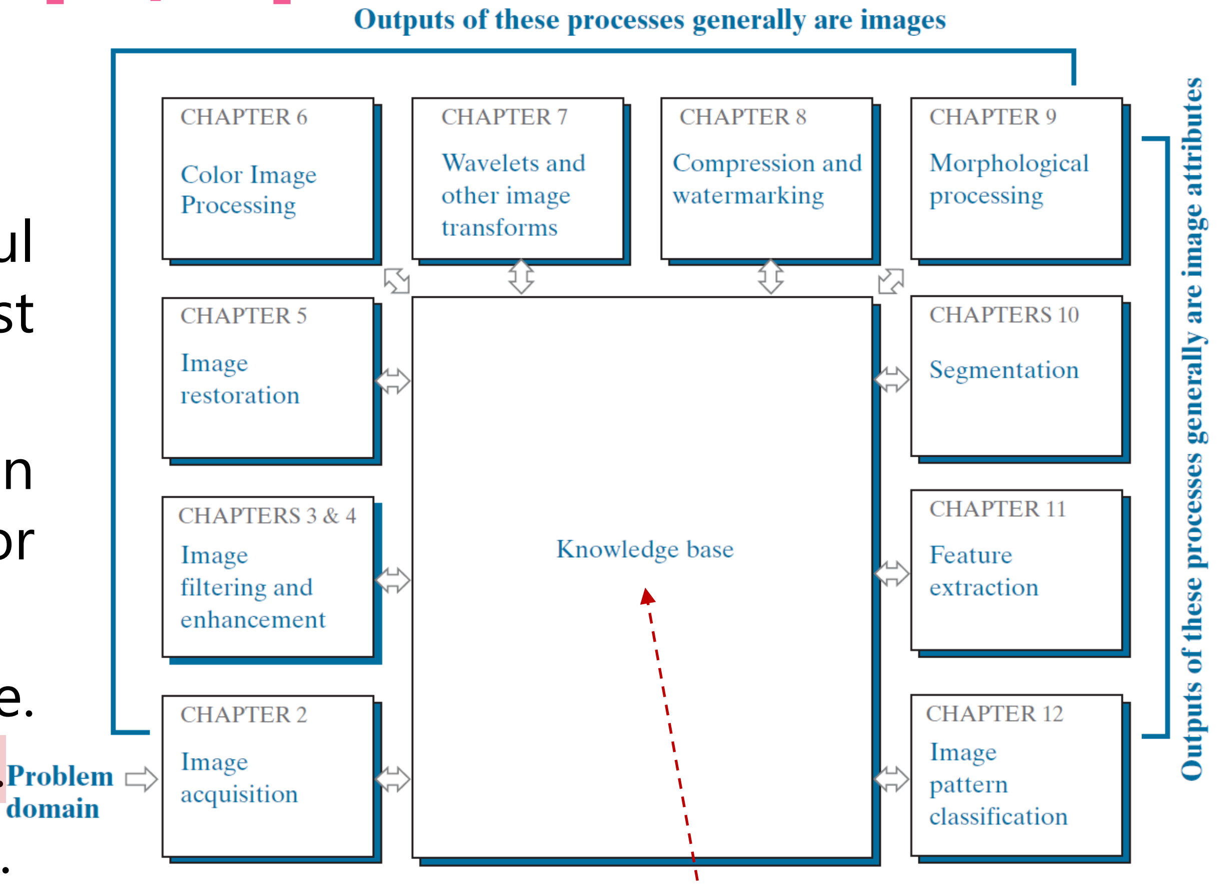
Steps in Digital Image Processing I:

Image acquisition

Image enhancement (*qualitative*) a method useful for enhancing X-ray images may not be the best for enhancing satellite images in the IR band.

Image restoration (*quantitative*) restoration techniques tend to be based on mathematical or probabilistic models of image degradation.

Color image processing may be used as a feature. We cover color models and basic color processing.



Finally, what about this?

Wavelets - used for multiresolution representation.

Compression is used for reducing storage and transmission bandwidth.

Morphological operations extract topological features of an image.

Segmentation cuts up an image into homogenous parts. Automated segmentation is hard.

Feature extraction maybe executed on segments. Classification Bayesian, DLN-based, etc.

Introduction [12/13]

Steps in Digital Image Processing II: can you identify the steps from last slide here?

Image acquisition platform is drone. Goal is to count the number of crocodiles in Ken River Basin, MP, India and, their sizes.

Introduction [13/13]

Components of a DIP:

Physical sensor to acquire the data, e.g. a photodiode – a Si-based material; use a filter to improve sensitivity)

Digitizer to digitize the data (e.g. CCD chip)

Specialized Hardware, e.g., Arithmetic & Logic Unit (ALU) to perform primitive tasks in parallel.

GPUs to perform intensive matrix operations in parallel.

PC for offline processing tasks.

Software that provides image processing primitives. You can write your own – but that is far more challenging and, time consuming.

Storage can be heavy for online computation; e.g. RCS computation for a missile seeker.

Frame buffers for rapid access to relatively small amount of memory to rapidly execute operations such as zoom, scroll or pan.

I/O elements can be challenging to design for intensive online applications.

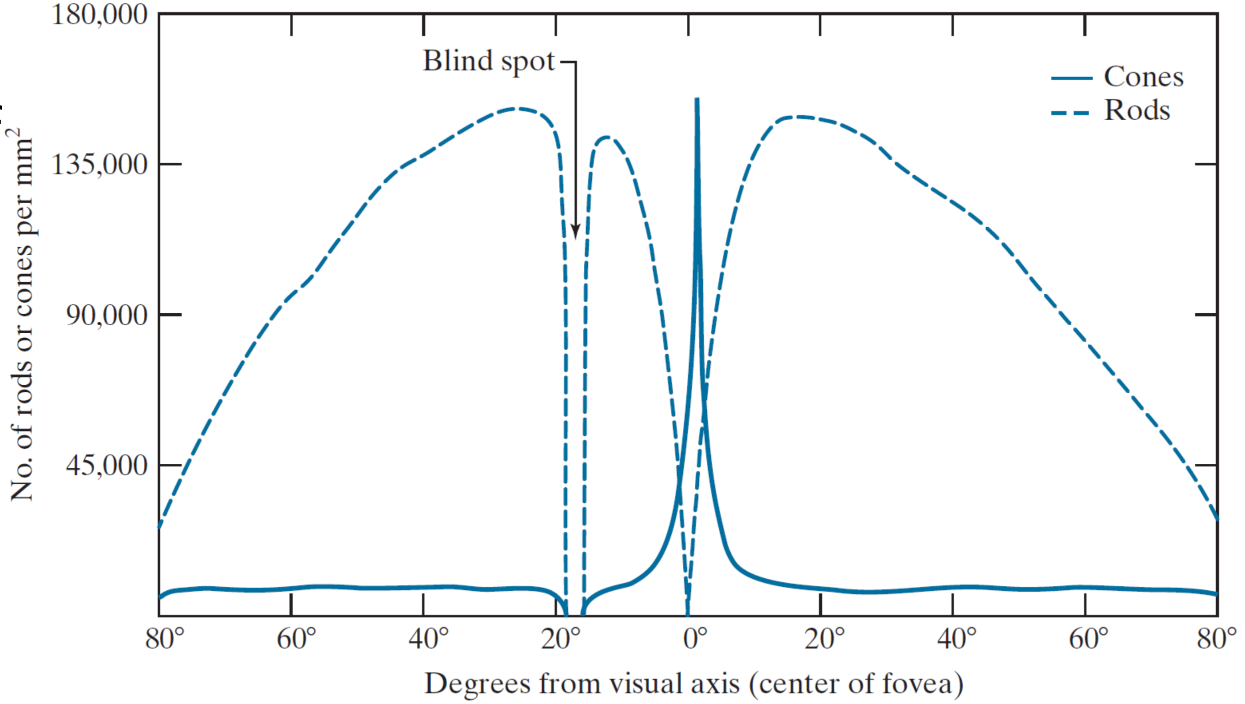
Display unit to display results in the same color or feature space as processing goal.

1 Introduction

2 Digital Image Fundamentals

Digital Image Fundamentals [1/21]

Elements of Visual Perceptions: We are interested in learning: i) physical limitations of human vision; ii) interaction between human vision & electronic imaging devices and; iii) adaptation to changes in image properties.



Eye lens comprises 60-70% water, 6% fat and more proteins than any other tissue in the eye.

Lens is covered by a slightly yellow pigmentation that rises with age (cause of cataract).

Lens absorbs 8% of visible spectrum light with higher absorption at higher frequencies. Excessive amount of either UV or IR light can damage the eye.

Rods (dim light vision) and Cones (bright light vision). # of rods >> # of cones. Several rods are connected to same nerve ending. ∴ spatial resolution is lower in dim light.

Blind spot is region of eye without any receptors.

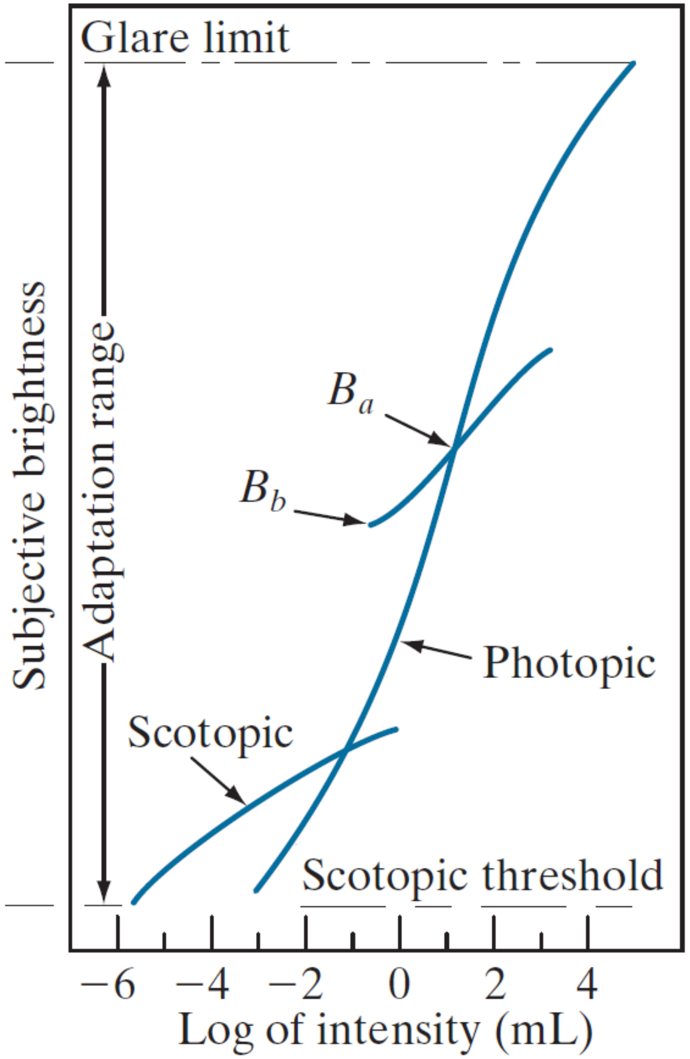
Digital Image Fundamentals [2/21]

Brightness Adaptation & Discrimination: The range of light intensity levels to which the human visual system can adapt is huge: $O(10^{10})$. However subjective brightness is not linear but logarithmic.

In photopic vision the range is 10^6. The visual system can however operate over only a small fraction of this range simultaneously. It accomplishes this variation by changing overall sensitivity – the phenomenon is called brightness adaptation.

Under a given set of conditions, brightness sensitivity is called brightness adaptation level, e.g. B_a. It is bounded lower by B_b which is the lower level of brightness below which the object appears black. The upper limit is usually 'far out'. Too high an intensity will simply raise the center point B_a.

We will ignore optical illusions such as Mach Band Effect or Simultaneous Contrast.



Digital Image Fundamentals [3/21]

Electromagnetic Spectrum: $E = hv$ & $c = \lambda v$. The perceived color is due to the wavelength of light reflected by the object in the visible range.

A **chromatic light source** is defined by:

i. frequency,

ii. radiance (*watt*),

iii. luminance (*lumen*) is similar to radiance but instead of measuring the actual radiated amount, it measures the perceived amount by a human

   i. *solar illuminance on Earth surface : c9*10^4 lm/m^2 on a clear day & c10^4 lm/m^2 on a clear day*

   ii. *full moon: 10^{-1} lm/m^2*

   iii. *Commercial office: c10^3 lm/m^2*

   $[L_{min}, L_{max}]$ is called intensity scale or grey scale. Common practice is to shift it $[0, 1]$. Depending on the application, the scale could be linear or log.

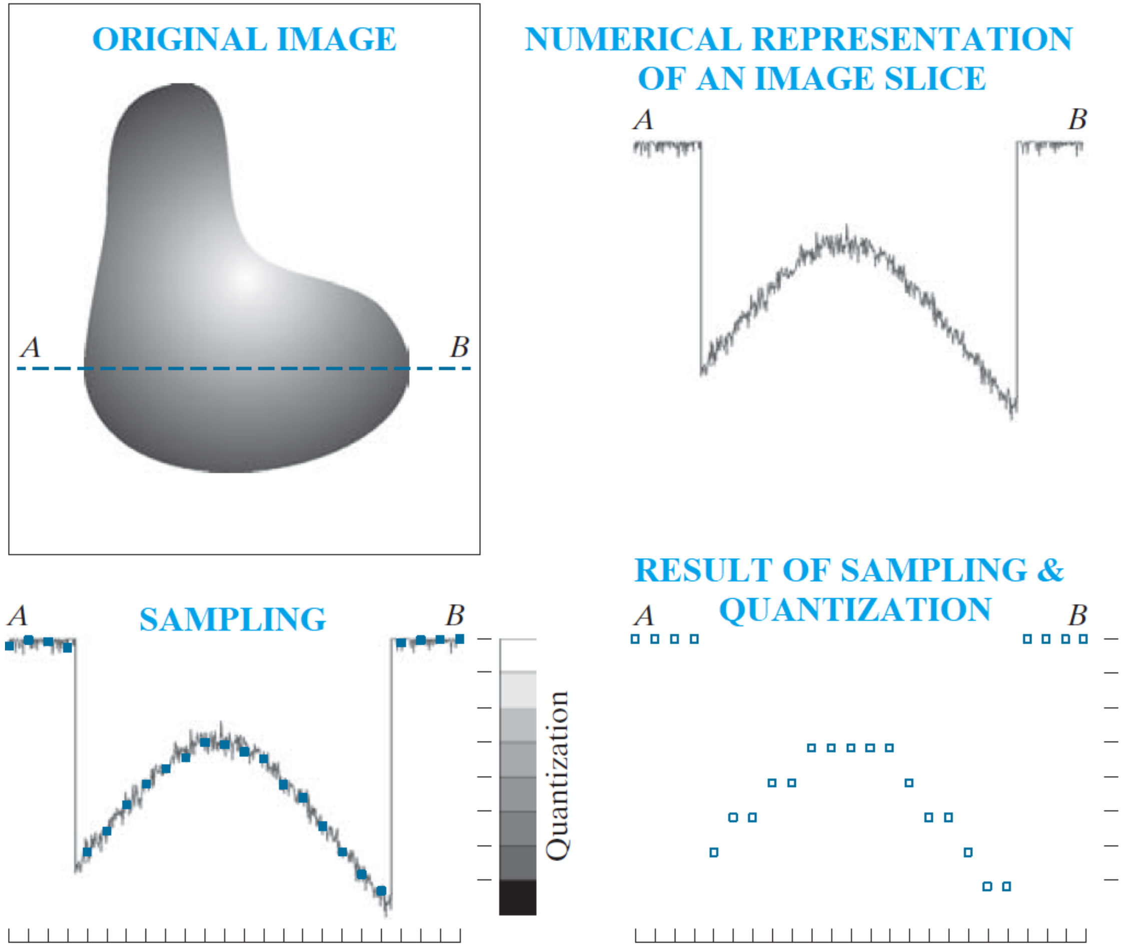
iv. brightness is a subjective descriptor of light perception. It embodies the achromatic notion of intensity and is one of the key factors in describing color sensation.

Digital Image Fundamentals [4/21]

$f(x,y) = i(x,y) * r(x,y)$ where $i(x,y)$ is the illumination component, $r(x,y)$ is the reflectance component; $0 \le f(x,y) < \infty$ & $0 \le r(x,y) < 1$

$r(x,y)$ typical values: 0.01 for black velvet, 0.65 for stainless steel, 0.90 for silver-plated metal, 0.93 for snow, >0.995 for mirrors

Sampling & Quantization: Digitizing the coordinate (amplitude) values is called sampling (quantization). x, y are called the spatial coordinates of the image and $f(x,y)$ is the spatial domain representation. Number of quantization levels are an integer power of 2: $L = 2^k$.



Is megapixel to do with sampling or quantization?

Why do phone cameras work very poorly at night?

How does an iPhone camera deliver exceptional image quality compared with the comparatively rather poor image acquisition system it has?

Digital Image Fundamentals [5/21]

Problem: A CCD element subtends a 0.33 x 0.33 arcsecond ($91.667*10^{-6}$ degree). It takes 1ms to register an image. The max rate the CCD device can transfer the picture to a storage is 7 frames/second. Distance between the camera and object of interest is 100 m. What is the max speed at which the camera can travel & take photos without blur?

Solution: 1ms to register an image may mean 1000 photos/sec. That surely isn't true. These images have to be transferred. So, the max rate is 7 frames/sec. How does one translate frame/sec to ms^{-1}?

At 100m, the FOV implies a horizontal distance of $2*100*\tan[91.667*10^{-6}/2)] = 0.00016$m. This must remain constant for at least 2*(1ms). ∴ max speed is $0.00016/(2*10^{-3}) = 0.08ms^{-1}$. which is quite small! This is independent of data transfer rate.

∴ post-processing with motion blur removal algorithms is needed. If real-time images are needed then, the speed of blur removal algorithms is the limiting factor.

A pixel is each element of $f(x, y)$ (see next slide).

Digital Image Fundamentals [6/21]

Pixel representation is an MxN matrix. By convention, the image origin is at the top right corner.

$$f(x,y)=\begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}$$

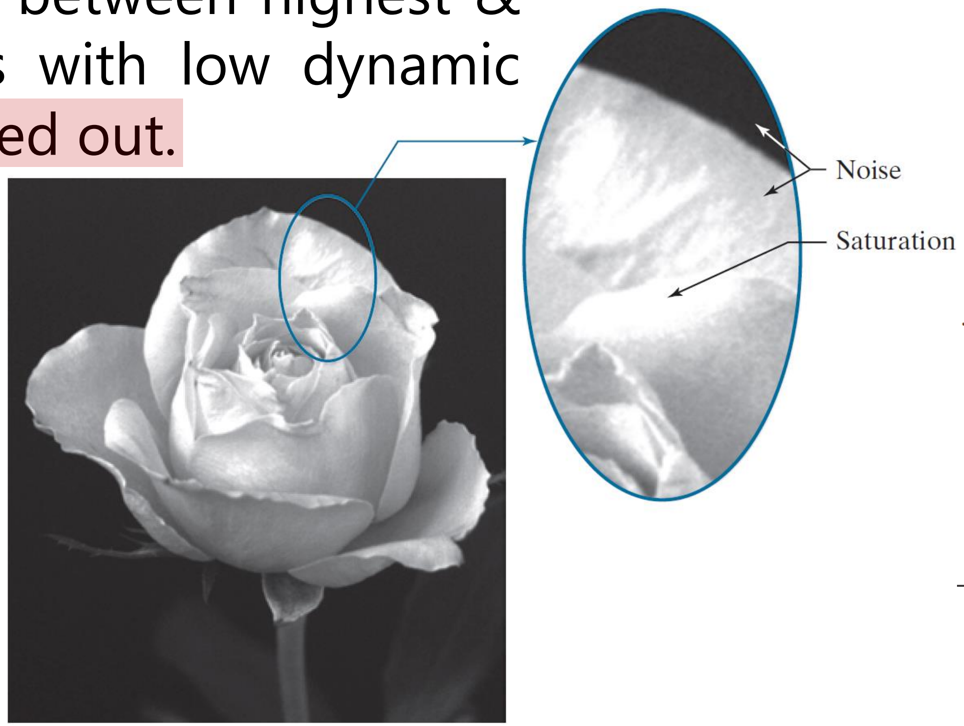
The 2nd image shows conventional coordinate system.

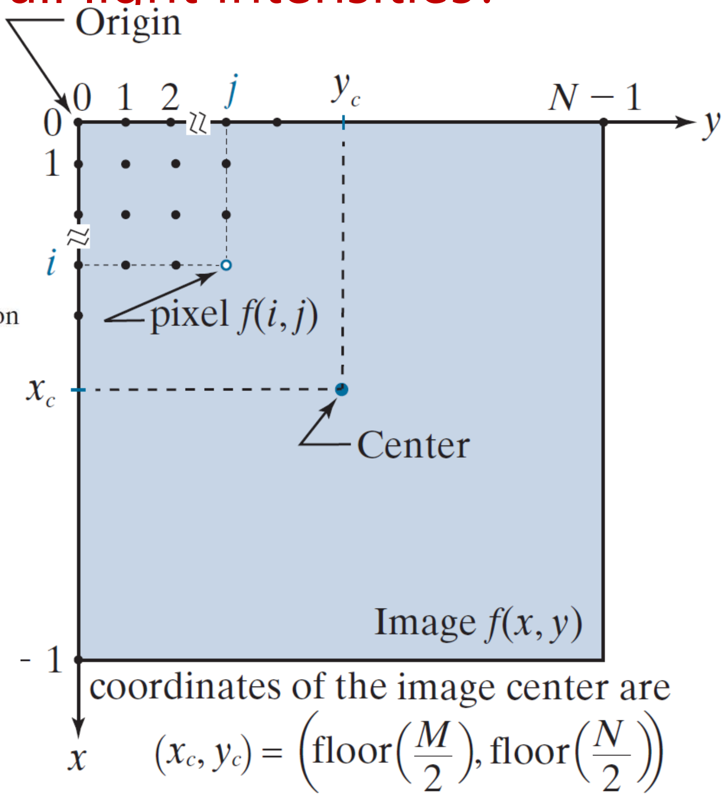
Range of values spanned by the grey scale is called dynamic range. The upper limit is determined by saturation and the lower by noise. Is noise present at all light intensities?

Image contrast is the difference between highest & lowest intensity levels. Images with low dynamic range appear either dull or washed out.

Image contrast is the difference between highest & lowest intensity levels. Images with low dynamic range appear either dull or washed out.

# of bytes needed to store an image B = M*N*k.



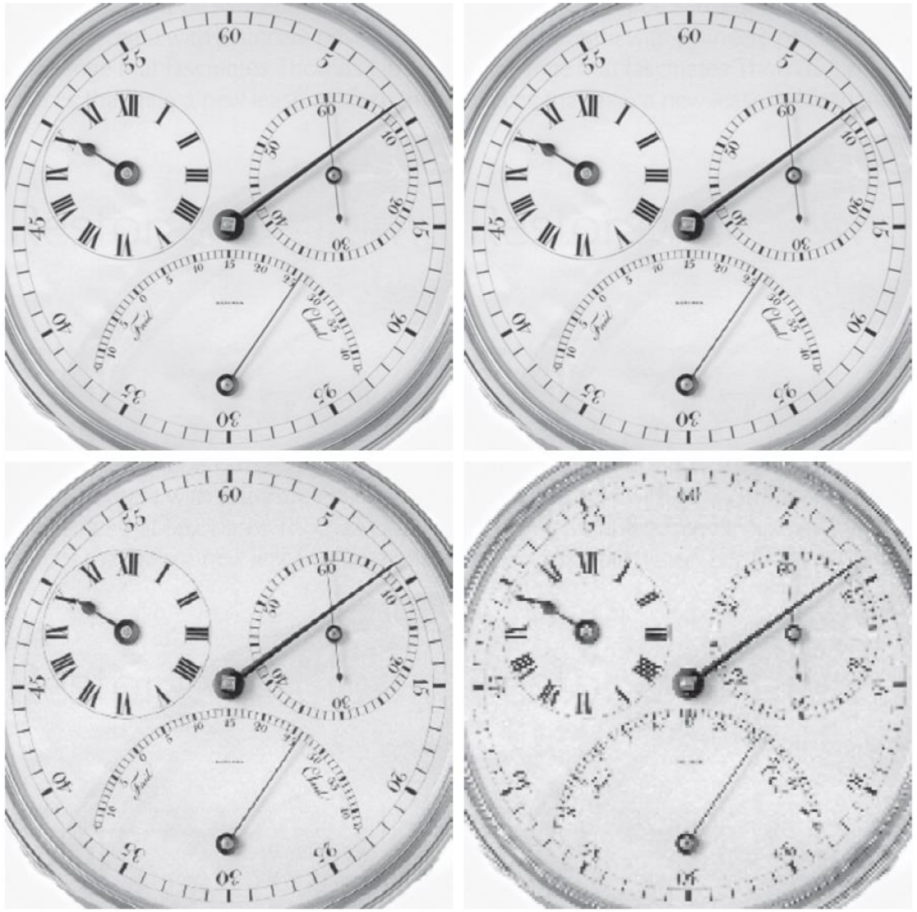


Digital Image Fundamentals [7/21]

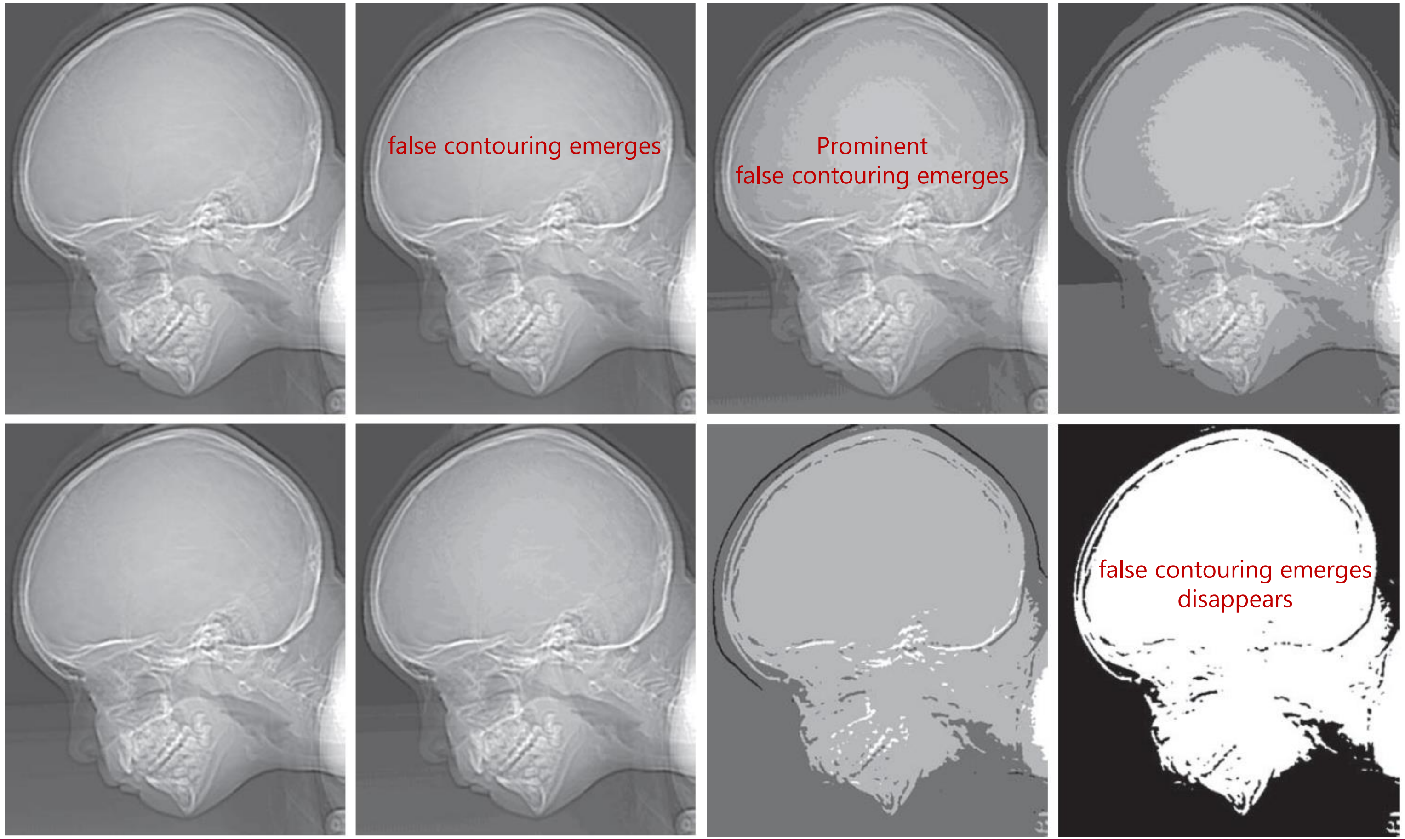
Megapixel is only a part of the story for spatial resolution. You also need to know sensor size. Why? For printing dots per inch (dpi) is a better measure.

Effect of poor resolution vs poor dynamic range

Spatial Resolution: 930, 300, 150, 72 dpi



DR: *256, 128, 64, 32, 16, 8, 4, 2 levels*



Digital Image Fundamentals [8/21]

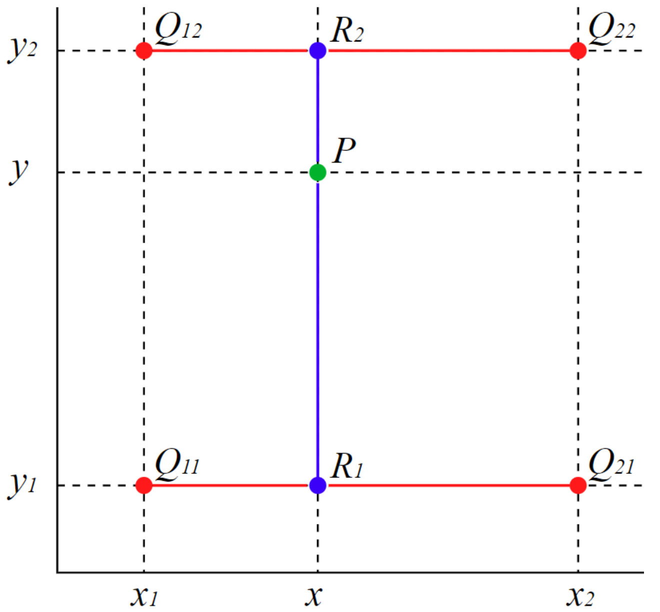
Image Interpolation: is the process of using known data to estimate values at unknown locations. It is used in zooming, shrinking, rotating, and geometrically correcting digital images. Let us zoom a 500x500 pixel image to a 750x750 pixel image.

Step 1: Create a 750x750 grid with same pixel spacing as that for 500x500.

Step 2: Reduce the spacing by 1.5 so that the 750x750 grid overlays the 500x500.

Step 3: For each smaller-sized pixel, assign it the intensity of the closest pixel based on 500x500.

All the variations arise in Step 3. Here, *nearest neighbor interpolation* was used. *Bilinear interpolation* is better but computationally more expensive. It uses 4 nearest neighbor points.



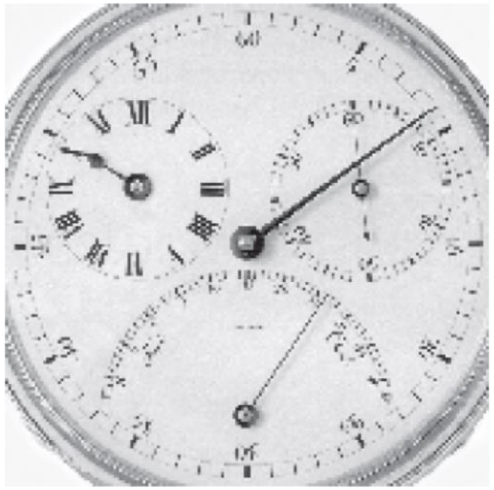
$$\frac{1}{(x_2-x_1)(y_2-y_1)}\begin{bmatrix}(x_2-x) & (x-x_1)\end{bmatrix}\begin{bmatrix}f(1,1) & f(1,2)\\ f(2,1) & f(2,2)\end{bmatrix}\begin{bmatrix}(y_2-y)\\ (y-y_1)\end{bmatrix}$$

Bicubic interpolation uses 16 nearest neighbor points. It is used in Adobe Photoshop/Corel Photopaint. Splines/Wavelets yields better interpolations - used in 3D zoom where difference is clearly noticeable.

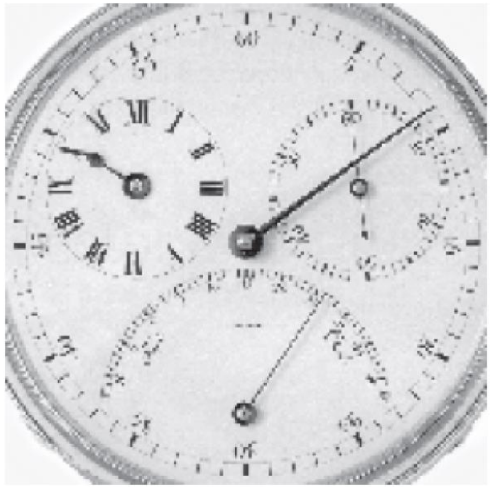
Digital Image Fundamentals [9/21]

Effect of image interpolation for zooming. The image that was reduced to 72 dpi is zoomed back to 930 dpi using nearest neighbor interpolation.

72 dpi used for HR reconstruction



930 dpi reconstructed with NNI



930 dpi reconstructed with BLI

930 dpi reconstructed with BCI

Original 930 dpi image

PIXEL NEIGHBORS: Neighbors of a pixel at (x, y) are:

Nearest neighbors: $N_4(P) = (x, +1\ y), (x - 1, y), (x, y + 1), (x, y - 1)$

Nearest diagonal neighbors: $N_D(P) = (x + 1, y + 1), (x - 1, y + 1), (x + 1, y - 1), (x - 1, y - 1)$

$N_8(P) = N_4(P) \cup N_D(P)$

Digital Image Fundamentals [10/21]

Distance Measures: D is a distance function over pixels p, q & s with coordinates (x, y), (u, v) & (w, z) IF: **a)** $\mathrm{D}(p, q) \geq 0$, **b)** $\mathrm{D}(p, q) = \mathrm{D}(q, p)$ & **c)** $\mathrm{D}(p, s) \leq \mathrm{D}(p, q) + \mathrm{D}(q, s)$

Euclidean Distance $D_e(p, q) = \sqrt[2]{(x-u)^2 + (y-v)^2}$

City-Block Distance $D_4(p, q) = |x-u| + |y-v|$

Chessboard Distance $D_8(p, q) = max(|x-u|, |y-v|)$

		2		
	2	1	2	
2	1	0	1	2
	2	1	2	
		2		

$D_4(p, q)$

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

$D_8(p, q)$

Elementwise vs Matrix Operations: Elementwise product symbol $\otimes$ or $\odot$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \otimes \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}$$

Matrix product $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \otimes \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$ Unless otherwise stated, for this course, we assume elementwise operations.

The addition/subtraction operations are defined only as elementwise, not as matrix operations.

Digital Image Fundamentals [11/21]

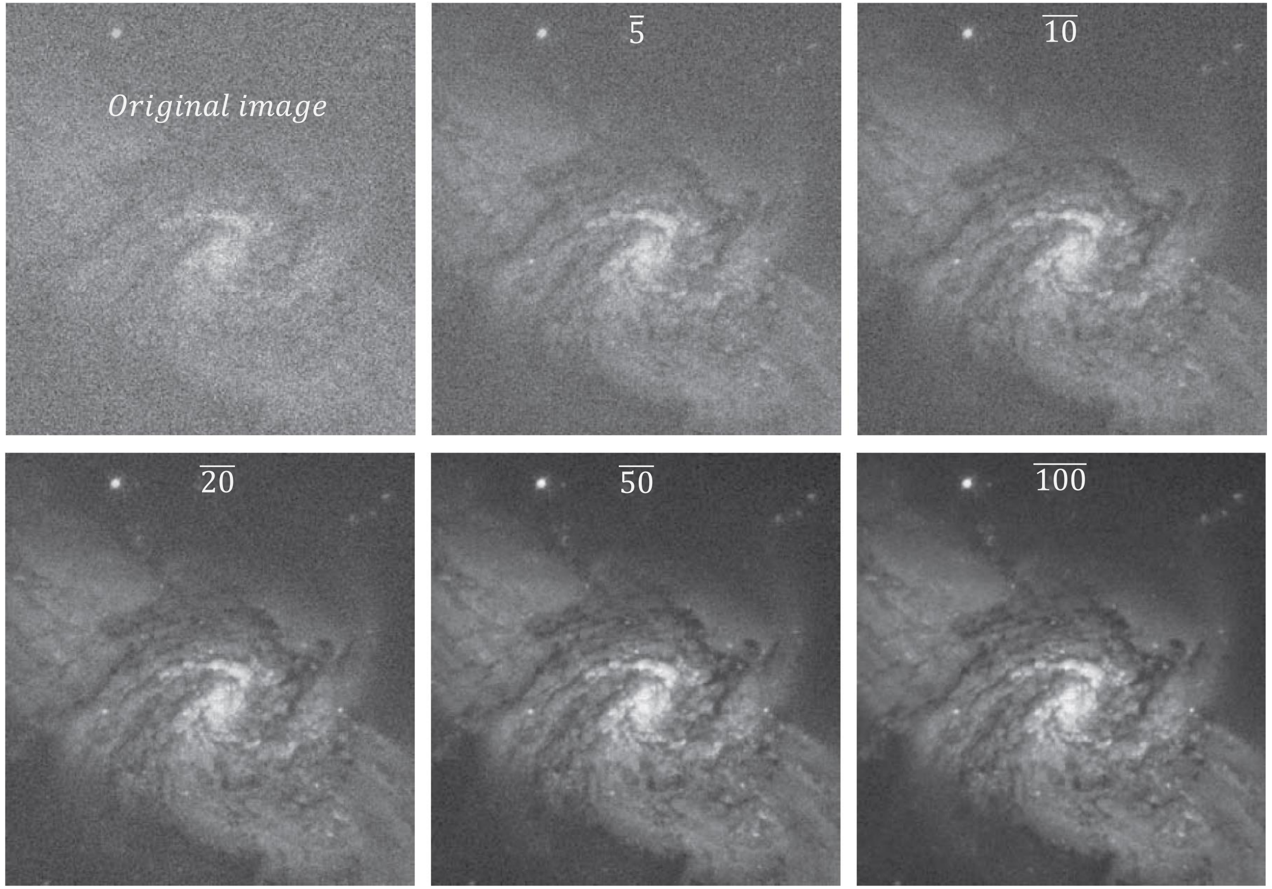
Image Addition (averaging) for Noise reduction: Suppose $\mathcal{H}[g(x,y)] = f(x,y) + \eta(x,y)$ where $f(x,y)$ is noiseless image and $\eta(x,y)$ is a noise with zero mean and uncorrelated with the image itself (a much stronger condition is statistical independence). If η indeed satisfies the conditions, then you can add several instances of the images and take the average to obtain the original noiseless image:

If η indeed satisfies the aforementioned conditions then you can add several instances of the images and take the average to obtain the original noiseless image:

$$\overline{\mathcal{H}[g(x,y)]} = \left(\frac{1}{k}\right)\sum_{i=1}^{k}\mathcal{H}[g_i(x,y)] = f(x,y)$$

&

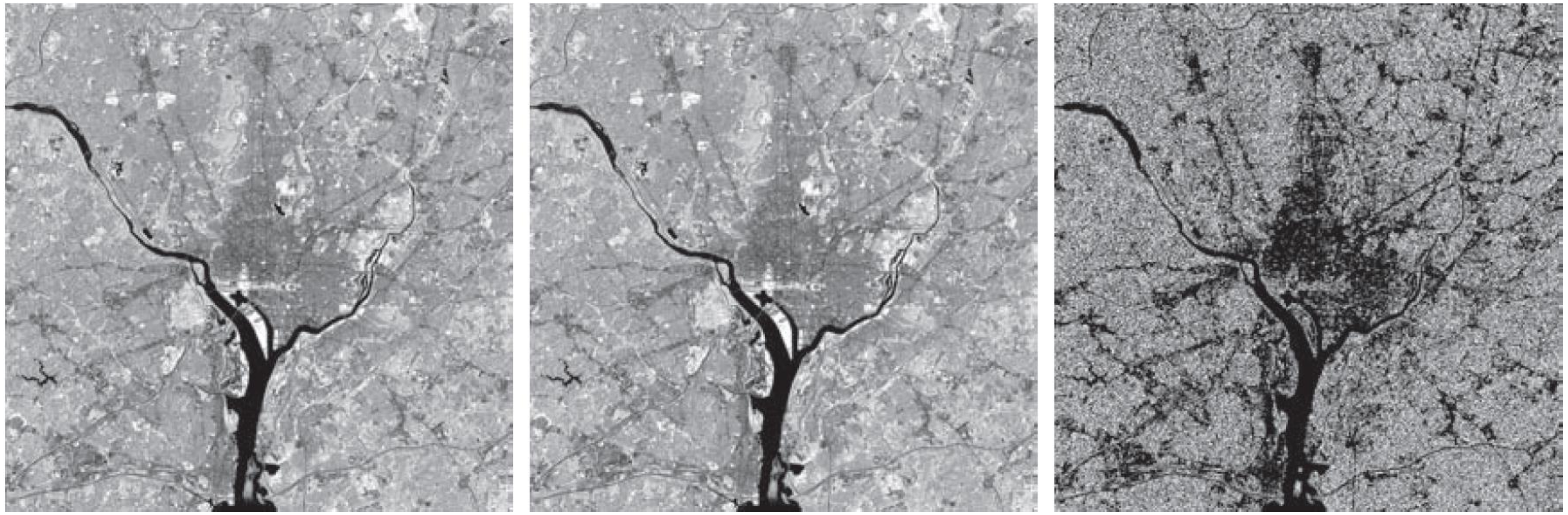
$\sigma^2_{\overline{\mathcal{H}[g(x,y)]}} = \left(\frac{1}{k}\right)\sigma^2_{\eta(x,y)}$. As k increases, the averaging draws the result of the operation closer to the original noiseless image. It is however necessary that each image taken, $g_i(x,y)$, must be spatially aligned.



Digital Image Fundamentals [12/21]

Subtracting Images for Comparison I: Image subtraction is used for enhancing differences between images.

In an 8-bit image, set LSB to 0 in the original image (L) to get (C) image. The two images are visually indistinguishable. (R) image is [1st image - 2nd image].

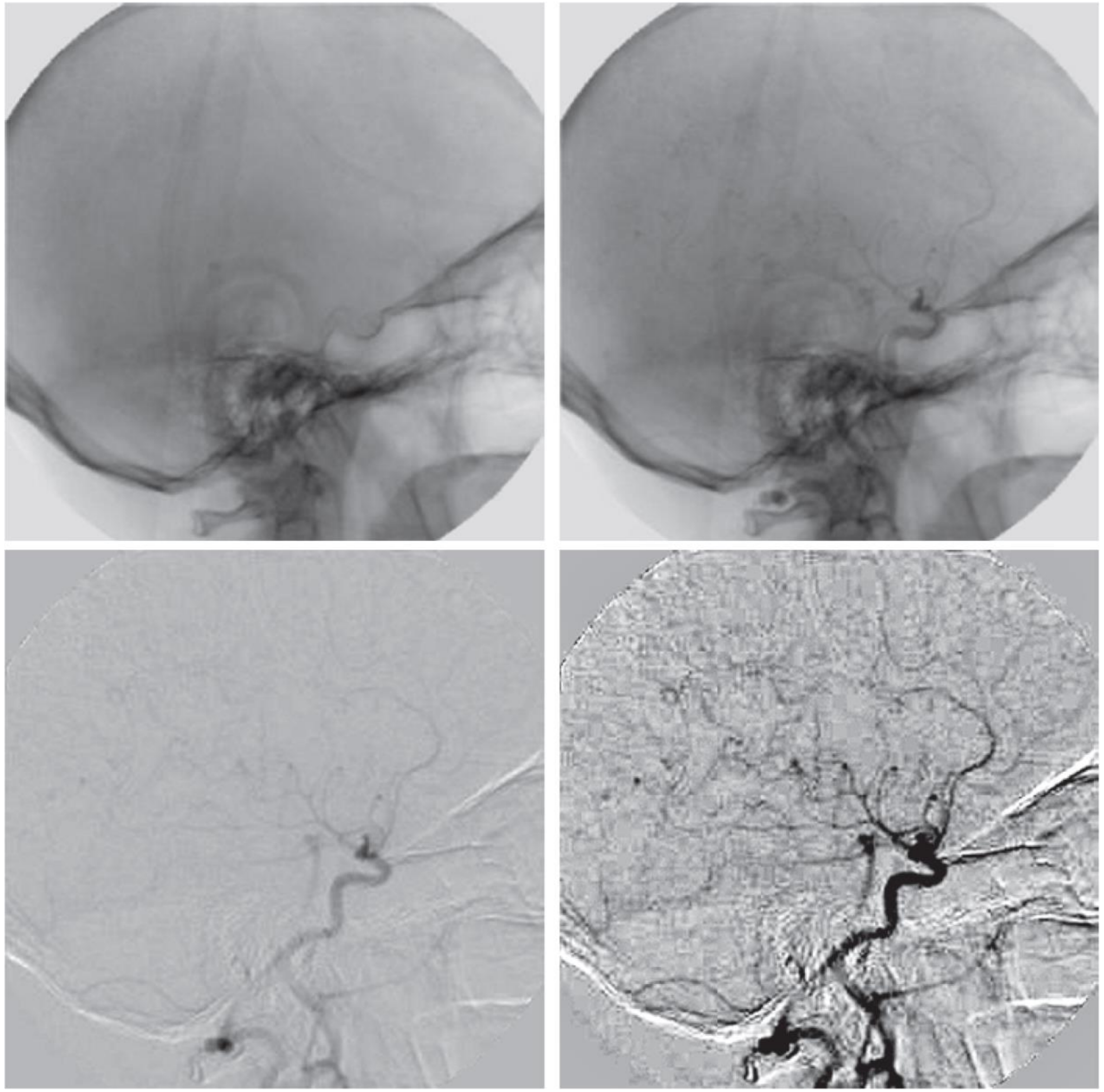


As another example, (L) image is difference between 930 and 72 dpi images. (C) image is difference between 930 & 150 dpi and (R) image is difference between 930 & 300 dpi.

Digital Image Fundamentals [13/21]

Subtracting Images for Comparison II: Last example of image subtraction for difference enhancement

a) Mask image (e.g. *angiography image for a patient under normal conditions*)
b) Test image (*image taken after injection of iodine*)
c) Difference image
d) Enhanced image (*sharpening & increase dynamic range*)



Digital Image Fundamentals [14/21]

Image Multiplication & Division for Shading Correction & Masking I: Suppose $g(x, y) = f(x, y)h(x, y)$ where $f(x, y)$ is noiseless image and $h(x, y)$ is a shading function. If $h(x, y)$ can be estimated, then $f(x, y)$ can be obtained by multiplying the sensed image, $g(x, y)$, with $[h(x, y)]^{-1}$.

$g(x, y)[h(x, y)]^{-1}$ is implemented as elementwise division.

How to get $h(x, y)$? If imaging system is accessible, then get the shading function by imaging an object of constant intensity. Else use Section 3.5 (*smoothing using low pass spatial filter*)/9.8 (*morphological smoothing*).

Another application is masking: multiplying an image by a mask image that has 1s in the region of interest (ROI) and 0s elsewhere.

Digital Image Fundamentals [15/21]

Set Theory: Be familiar with these relationships. How we actually define these, except for very simple cases such as union of an image and its complement, we will see later. At the moment just bear in mind that multiple definitions exist depending on the objective: e.g., union of two images A & B of the same resolution is:

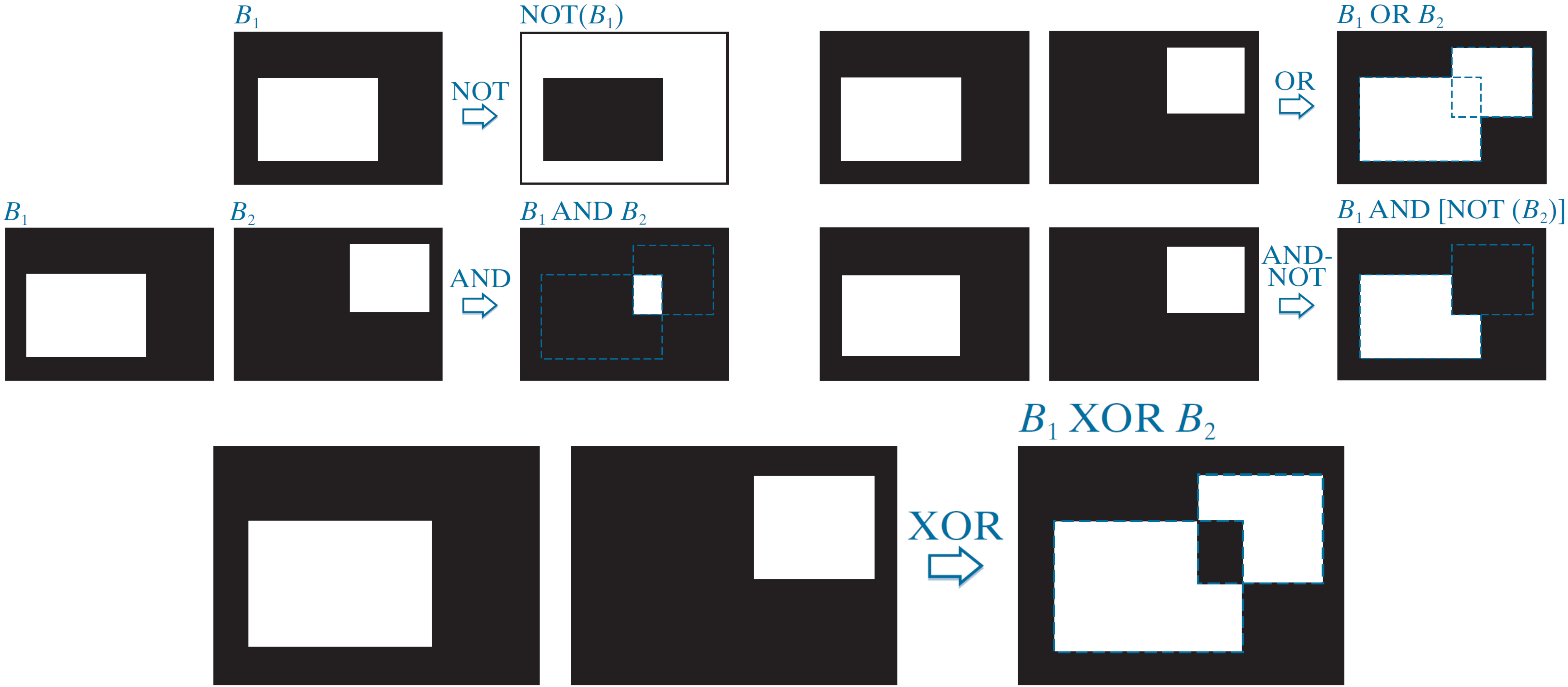
$$\left\{ \max_z(a, b) \mid a \in A, b \in B \right\}$$

But, defining it as $sum(a,b)$ or $min(a,b)$ is also equally valid as no natural definition exists.

Description	Expressions
Operations between the sample space and null sets	$\Omega^c = \varnothing;\ \varnothing^c = \Omega;\ \Omega \cup \varnothing = \Omega;\ \Omega \cap \varnothing = \varnothing$
Union and intersection with the null and sample space sets	$A \cup \varnothing = A;\ A \cap \varnothing = \varnothing;\ A \cup \Omega = \Omega;\ A \cap \Omega = A$
Union and intersection of a set with itself	$A \cup A = A;\ A \cap A = A$
Union and intersection of a set with its complement	$A \cup A^c = \Omega;\ A \cap A^c = \varnothing$
Commutative laws	$A \cup B = B \cup A$ $A \cap B = B \cap A$
Associative laws	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$
Distributive laws	$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
DeMorgan's laws	$(A \cup B)^c = A^c \cap B^c$ $(A \cap B)^c = A^c \cup B^c$

Digital Image Fundamentals [16/21]

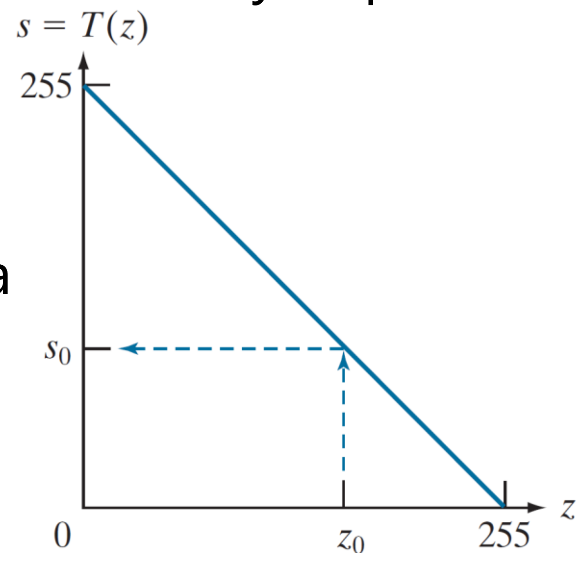
Logical Operations: DIY – it is intuitive. You can do these only on binary-valued images.



Digital Image Fundamentals [17/21]

Spatial Operations: Three types:

a) single-pixel operations: e.g. alter pixel intensity via $s = T(z)$ where z is the intensity of pixel in original image and s is the mapped intensity in the processed image.

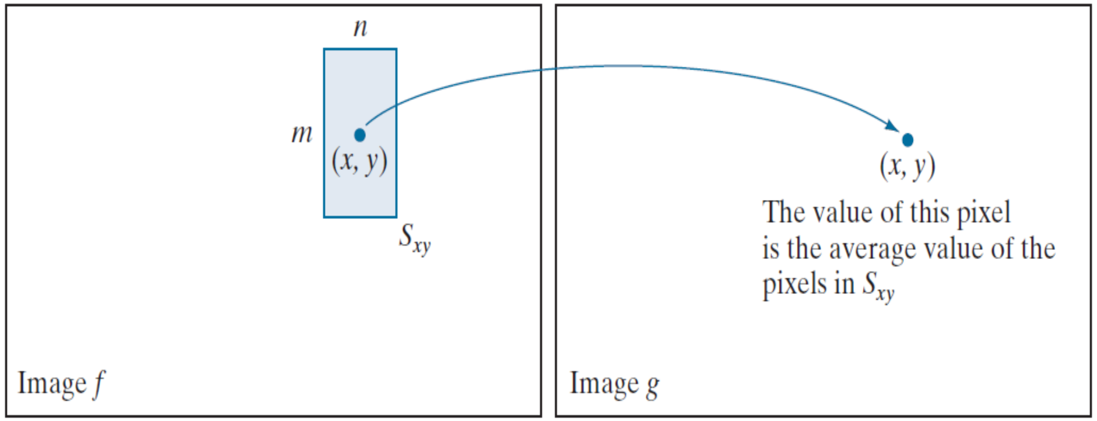


b) neighborhood operations: Let S_{xy} be set of coordinates of a neighborhood centered at (x, y) in an image f. Neighborhood processing yields an image g. E.g., if the processing function is averaging function in a rectangular neighborhood of size $m \times n$ on (x, y) - *bottom right image*.

$g(x, y) = \frac{1}{mn} \sum_{(r,c) \in S_{xy}} f(r, c)$ This can be used for local smoothing.

c) geometric spatial operations: it comprises a) spatial transformation and intensity assignment. $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \mathbf{k}$

This can implement affine transformations that include scale, translate, rotate & scale. All of these affine transformations are implemented using a single matrix $\mathbf{A}$. *See next slide …*



Affine Transformations:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{A}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Forward mapping (FM)

Inverse mapping (IM)

FM scans pixels of input image at each location (x, y) & computes (x', y'). This will unfortunately map more than one input pixels to the same output pixel. To see this, consider: $x' = a_{11}x + a_{12}y + a_{13}$; $y' = a_{21}x + a_{22}y + a_{23}$.

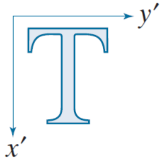
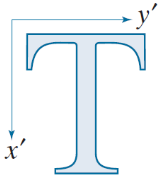
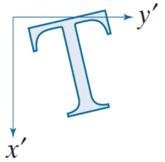
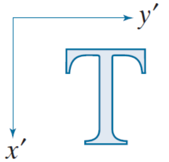
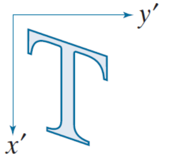
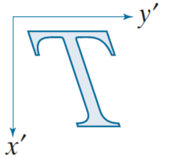
$\therefore\ x_1{}' = a_{11}x_1 + a_{12}y_1 + a_{13}$; $y_1{}' = a_{21}x_1 + a_{22}y_1 + a_{23}$

$x_2{}' = a_{11}x_2 + a_{12}y_2 + a_{13}$; $y_2{}' = a_{21}x_2 + a_{22}y_2 + a_{23}$

But, if $x_1' = x_2' = x$ & $y_1' = y_2' = y$ then

$a_{11}x_1 + a_{12}y_1 + a_{13} = a_{11}x_2 + a_{12}y_2 + a_{13}$ or

$a_{11}(x_1 - x_2) = -a_{12}(y_1 - y_2)$ or, $\frac{(x_1 - x_2)}{(y_1 - y_2)} = -\frac{a_{12}}{a_{11}}$. Similarly, $\frac{(x_1 - x_2)}{(y_1 - y_2)} = -\frac{a_{22}}{a_{21}}$. So, if x_1, x_2, y_1 & y_2 are thus related, two or more input pixels map to one output pixel.

Transformation Name	Affine Matrix, A	Coordinate Equations	Example
Identity	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = y$	
Scaling/Reflection (For reflection, set one scaling factor to −1 and the other to 0)	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = c_x x$ $y' = c_y y$	
Rotation (about the origin)	$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x\cos\theta - y\sin\theta$ $y' = x\sin\theta + y\cos\theta$	
Translation	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + t_x$ $y' = y + t_y$	
Shear (vertical)	$\begin{bmatrix} 1 & s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + s_v y$ $y' = y$	
Shear (horizontal)	$\begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = s_h x + y$	

Digital Image Fundamentals [19/21]

Affine Transformations: ∴ inverse mapping is preferred, i.e., compute $(x, y) = A^{-1}(x', y')$. MATLAB uses IM for spatial transformations.

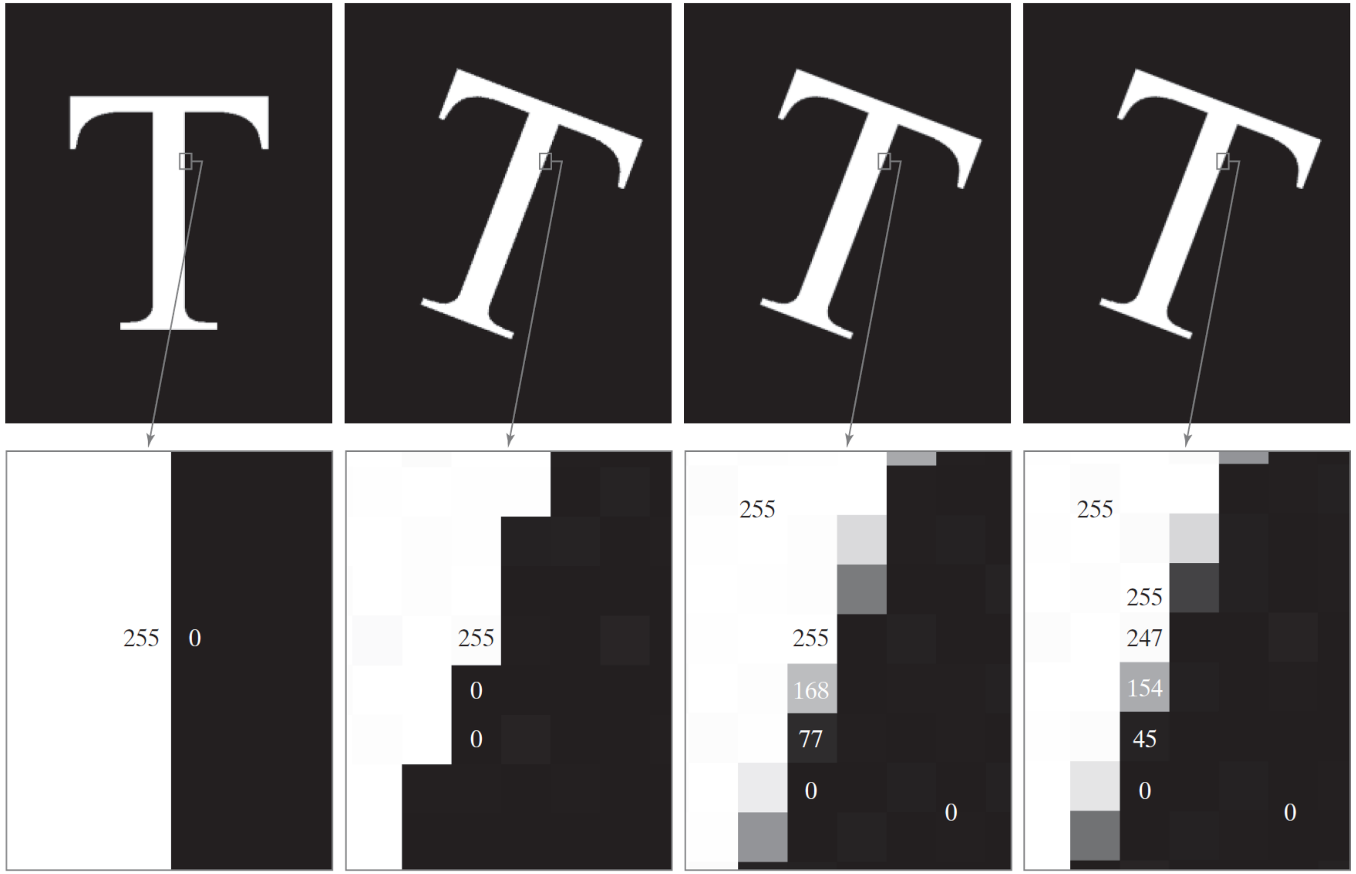
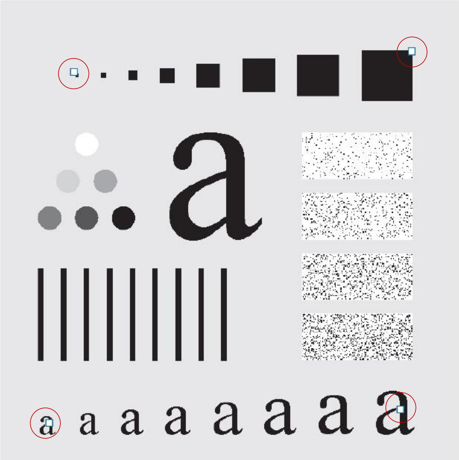


Image Rotation & Intensity Interpolation: Figure shows an image and its rotations by –21° (clockwise) with intensity assignments computed using NNI, BLI & BCI. The NNI produces the most jagged edges. While jagged edges may not be noticeable to the human eyes, it may cause a problem for an automatic edge-following robot.

Image Registration: Used to align multiple images to a reference image. Unlike simple rotation above, here the transformation fn needed to align the images is not known. E.g. aligning images taken by different sensors or, images of the same scene taken by a moving platform such as a satellite or a drone. Viewing angle, viewing distance, orientation, sensor resolution, object location, etc. could all change.

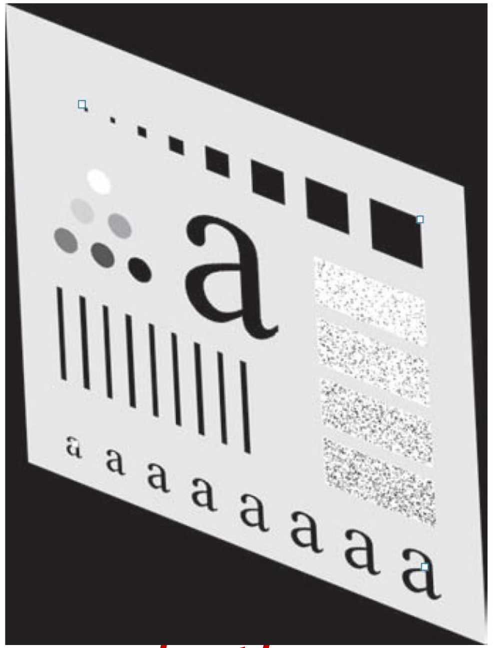
Digital Image Fundamentals [20/21]

Image Registration (... contd.): Control Points may be used to resolve the problem. These are points whose locations are precisely known in both the input and the reference images. Suppose we have 4 control points in an image. Then a bilinear approximation model is $x = c_1 v + c_2 w + c_3 vw + c_4$ & $y = c_5 v6w + c_7 vw + c_8$ where (v, w) are coordinates of control points in the input image and (x, y) in the reference image. With 4 control points we have 8 equations to get $c_1, c_2, \ldots, c_8$. This will map the coordinate pixels in the reference and the input images. We still need to perform intensity interpolation to assign intensity values to each pixel.

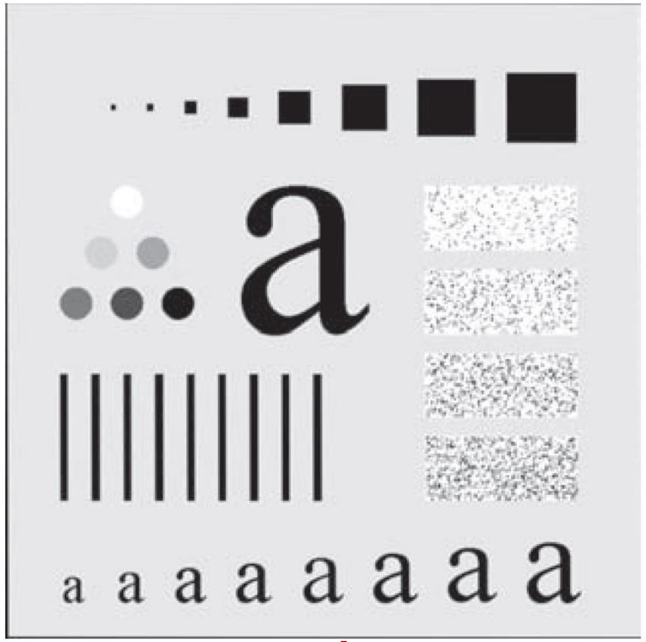


4 control points – blue squares highlighted with red circles

Reference Image



Input Image



Mapped Image

Difference from Original

Digital Image Fundamentals [21/21]

Vector Representation: Inner product of **a** & **b**: $\mathbf{a} \bullet \mathbf{b} \triangleq \mathbf{a}^T\mathbf{b} = a_1b_1 + a_2b_2 + \cdots + a_nb_n = \sum^{n} a_i b_i$

Euclidean vector norm is square root of inner product of a vector & its transpose $\|\mathbf{z}\| = \left(\mathbf{z}^T\mathbf{z}\right)^{\frac{1}{2}}$

Euclidean distance between **z** & **a** is

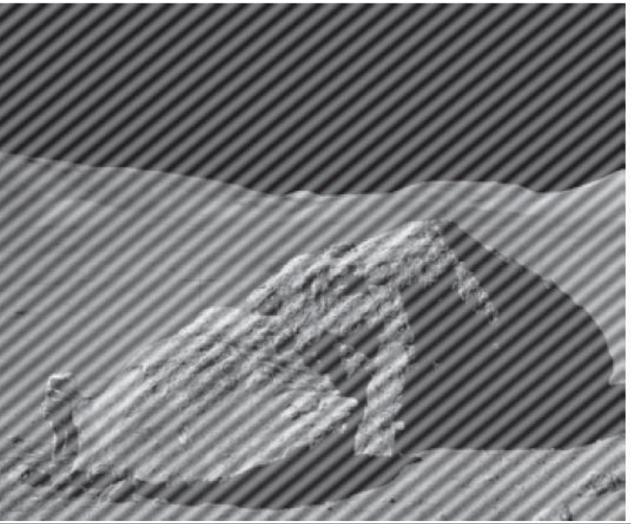
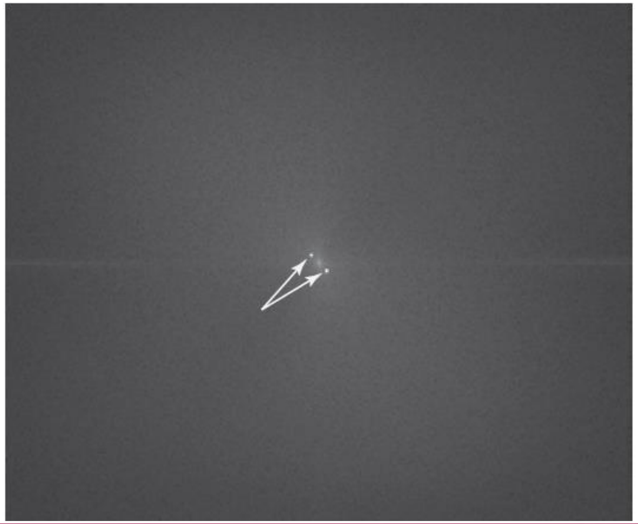
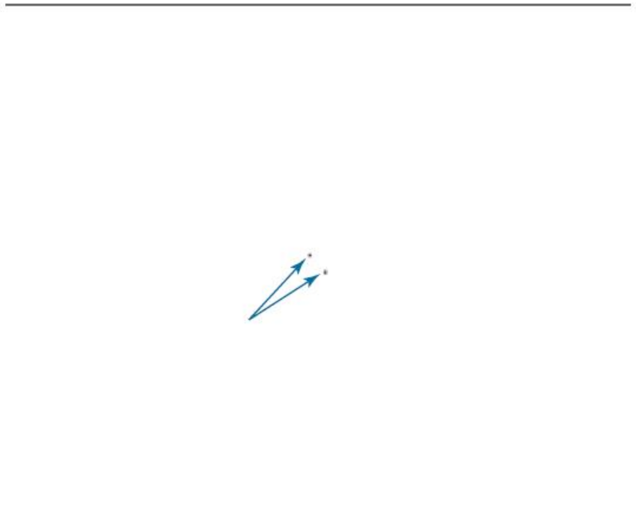
$$D(\mathbf{z},\mathbf{a}) = \|\mathbf{z}-\mathbf{a}\| = \left[(\mathbf{z}-\mathbf{a})^T(\mathbf{z}-\mathbf{a})\right]^{\frac{1}{2}} = \left[(z_1-a_1)^2 + (z_2-a_2)^2 + \cdots + (z_n-a_n)^2\right]^{\frac{1}{2}}$$

Discrete Fourier Transform (DFT) & its Inverse

$$T(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi(ux/M + vy/N)}$$

$$f(x,y) = \frac{1}{MN}\sum_{u=0}\sum_{v=0} T(u,v)e^{j2\pi(ux/M + vy/N)}$$

Figure shows image corruption by sinusoidal interference. It is corrected by subtracting in the FT domain.


AMRITA VISHWA VIDYAPEETHAM