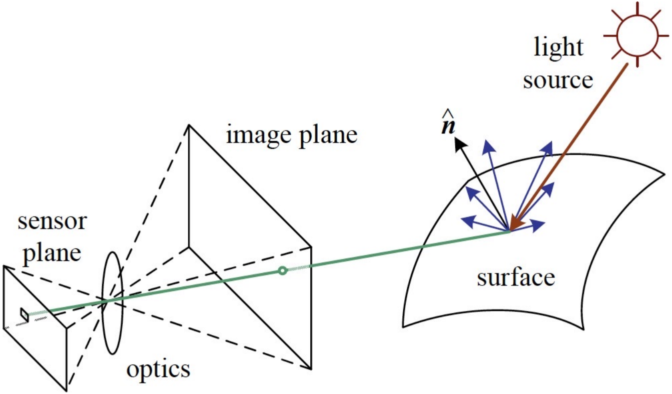
Vision System

Lecture 2

Part 1

- Image Formation

Image Formation: An Overview[1]

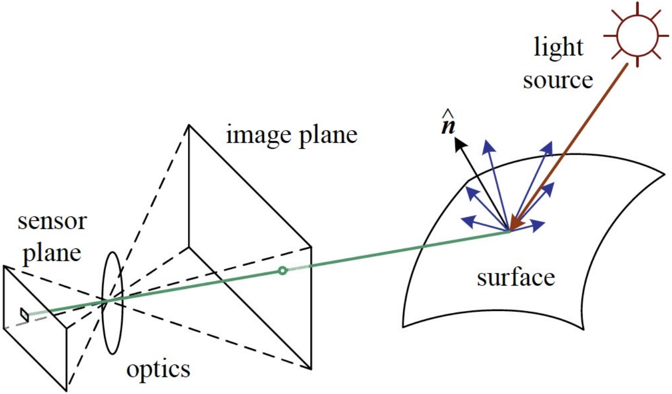


Factors

- Light source strength and direction
- Surface geometry, material and nearby surfaces
- Sensor capture properties
- Image representation and colour

[1]Credit: Szeliski, *Computer Vision: Algorithms and Applications*, 2010

Image Formation: An Overview[1]



Factors

- Light source strength and direction
- Surface geometry, material and nearby surfaces
- Sensor capture properties
- Image representation and colour

[1]Credit: Szeliski, *Computer Vision: Algorithms and Applications*, 2010

Related Topics

Geometry	2D Transformations, 3D Transformations, Camera Calibration, Distortion
Photometry	Lighting, Reflectance and Shading, Optics
Colour	Physics of Colour, Human Colour, Colour Representation
Sensor	Human Perception, Camera Design, Sampling and Aliasing, Compression

Related Topics

Geometry	2D Transformations, 3D Transformations, Camera Calibration, Distortion
Photometry	Lighting, Reflectance and Shading, Optics
Colour	Physics of Colour, Human Colour, Colour Representation
Sensor	Human Perception, Camera Design, Sampling and Aliasing, Compression

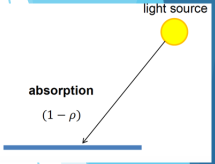
We will cover a few relevant topics from these

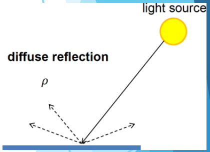
For a detailed understanding, read Chapters 1-5 of the book, *Computer Vision: A Modern Approach* by Forsyth and Ponce

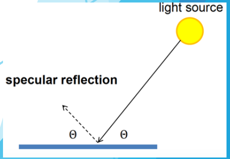
Models of Reflection

When light hits a surface:

- Some light is absorbed ($1-\rho$, ρ = albedo);
  - More absorbed for low albedos
- Some light is reflected diffusively, independent of viewing direction
  - E.g.: Brick, cloth, roughwood
  - **Lambert's cosine law:** Amount of reflected light proportional to $\cos(\theta)$
- Some light is reflected specularly, depends on viewing direction
  - E.g.: Mirror







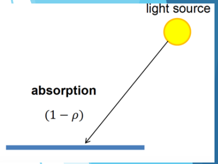
Credit: Derek Hoiem,UIUC

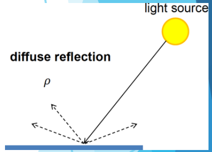
Models of Reflection

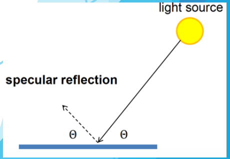
When light hits a surface:

- Some light is absorbed $(1-\rho, \rho = \text{albedo})$;
  - More absorbed for low albedos
- Some light is reflected diffusively, independent of viewing direction
  - E.g.: Brick, cloth, roughwood
  - **Lambert's cosine law:** Amount of reflected light proportional to $\cos(\theta)$
- Some light is reflected specularly, depends on viewing direction
  - E.g.: Mirror

Credit: Derek Hoiem,UIUC



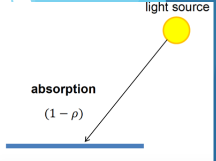


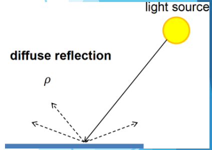


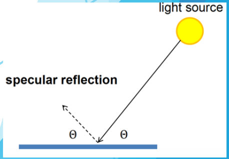
Models of Reflection

When light hits a surface:

- Some light is absorbed $(1 - \rho, \rho = \text{albedo})$;
  - More absorbed for low albedos
- Some light is reflected diffusively, independent of viewing direction
  - E.g.: Brick, cloth, roughwood
  - **Lambert's cosine law:** Amount of reflected light proportional to $\cos(\theta)$
- Some light is reflected specularly, depends on viewing direction
  - E.g.: Mirror







Credit: Derek Hoiem,UIUC

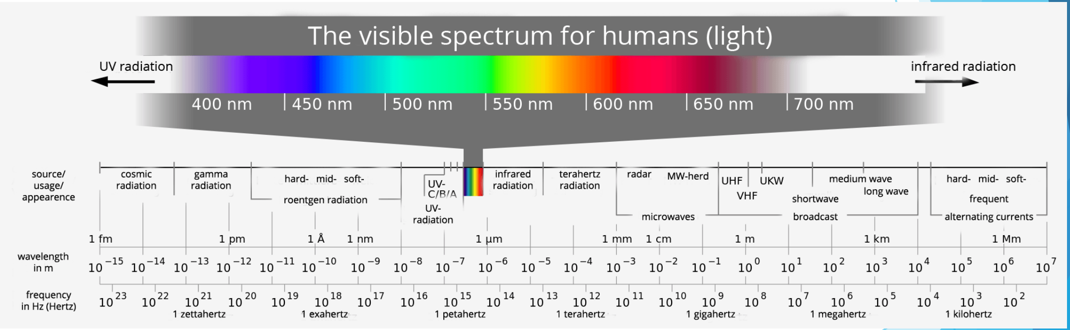
Models of Reflection

- Most surfaces have both specular and diffuse components
- Intensity depends on illumination angle because less light comes in at oblique angles
- Other possible effects:
    - Transparency
    - Refraction
    - Subsurface scattering
    - Fluorescence, phosphorescence
- BRDF - Bidirectional Reflectance Distribution Function: Model of local reflection that tells how bright a surface appears when viewed from one direction when light falls on it from another

Credit: Derek Hoiem,UIUC

Colour

Light is composed of a spectrum of wavelengths

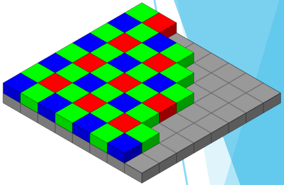


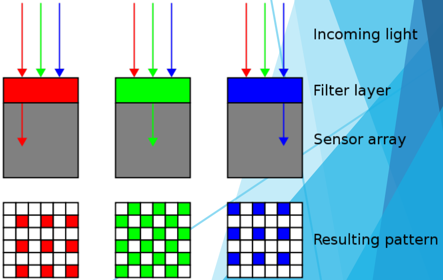
Coloured light arriving at sensor involves: (i) Colour of light source; (ii) Colour of surface

Credit: Electromagnetic spectrum by Horst Frank, Jailbird and Phrood. Under CC 3.0 License

Bayer Grid/Filter

- Bayer arrangement of color filters on a camera sensor
- Filter pattern is 50% green, 25% red and 25% blue
- To obtain full-colour image, **demosaicing** algorithms used - surrounding pixels used to estimate values for a particular pixel.





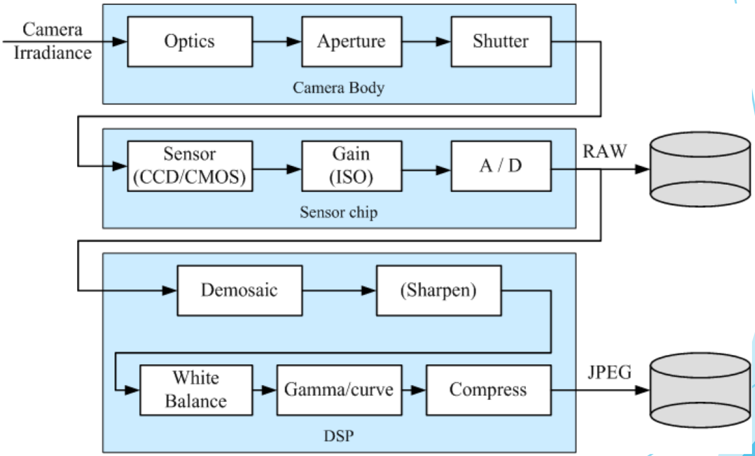
Credit: https://en.wikipedia.org/wiki/Bayer_filter

Question

On Colour

If visible light spectrum is VIBGYOR, why RGB colour representation?

Image Sensing Pipeline[2]



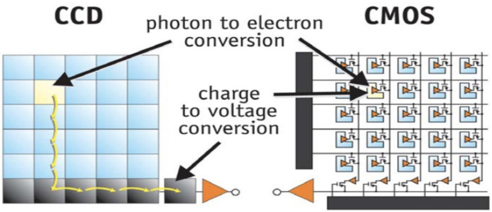
[2]Credit: Szeliski, *Computer Vision: Algorithms and Applications*, 2010

Digital Image Sensing

CCD vs CMOS: What’s the difference?

Digital Image Sensing

CCD vs CMOS: What’s the difference?



CCD

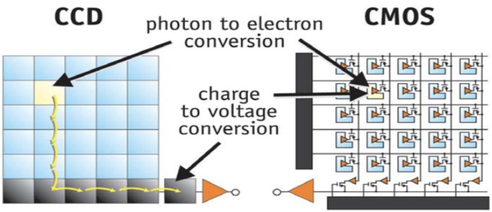
Move photogenerated charge from pixel to pixel, and convert it to voltage at output node

An analog-to-digital converter (ADC) then turns each pixel’s value into a digital value

Photo Credit: Litwiller, *CMOS vs. CCD: Maturing Technologies, Maturing Markets*, 2005

Digital Image Sensing

CCD vs CMOS: What's the difference?



CMOS

- CMOS convert charge to voltage inside each element
- Uses several transistors at each pixel to amplify and move the charge using more traditional wires
- CMOS signal is digital, so it needs no ADC

Digital Image Sensor Properties

Shutter speed:	Controls the amount of light reaching the sensor (also called *exposure time*)
Sampling pitch:	Physical spacing between adjacent sensor cells on the imaging chip
Fill factor:	Active sensing area size as a fraction of the theoretically available sensing area (product of horizontal and vertical sampling pitches)
Chip size:	Size/area of the chip
Analog gain:	Amplification of the sensed signal using automatic gain control (AGC) logic (controlled using ISO setting on cameras)
Sensor noise:	Noise from various sources in the sensing process
Resolution:	How many bits for each pixel, decided by analog-to-digital conversion module
Post-processing:	Digital image enhancement methods often used before compression and storage of captured image

Digital Image Sensor Properties

Shutter speed:	Controls the amount of light reaching the sensor (also called *exposure time*)
Sampling pitch:	Physical spacing between adjacent sensor cells on the imaging chip
Fill factor:	Active sensing area size as a fraction of the theoretically available sensing area (product of horizontal and vertical sampling pitches)
Chip size:	Size/area of the chip
Analog gain:	Amplification of the sensed signal using automatic gain control (AGC) logic (controlled using ISO setting on cameras)
Sensor noise:	Noise from various sources in the sensing process
Resolution:	How many bits for each pixel, decided by analog-to-digital conversion module
Post-processing:	Digital image enhancement methods often used before compression and storage of captured image

Digital Image Sensor Properties

Shutter speed:	Controls the amount of light reaching the sensor (also called *exposure time*)
Sampling pitch:	Physical spacing between adjacent sensor cells on the imaging chip
Fill factor:	Active sensing area size as a fraction of the theoretically available sensing area (product of horizontal and vertical sampling pitches)
Chip size:	Size/area of the chip
Analog gain:	Amplification of the sensed signal using automatic gain control (AGC) logic (controlled using ISO setting on cameras)
Sensor noise:	Noise from various sources in the sensing process
Resolution:	How many bits for each pixel, decided by analog-to-digital conversion module
Post-processing:	Digital image enhancement methods often used before compression and storage of captured image

Digital Image Sensor Properties

Shutter speed:	Controls the amount of light reaching the sensor (also called *exposure time*)
Sampling pitch:	Physical spacing between adjacent sensor cells on the imaging chip
Fill factor:	Active sensing area size as a fraction of the theoretically available sensing area (product of horizontal and vertical sampling pitches)
Chip size:	Size/area of the chip
Analog gain:	Amplification of the sensed signal using automatic gain control (AGC) logic (controlled using ISO setting on cameras)
Sensor noise:	Noise from various sources in the sensing process
Resolution:	How many bits for each pixel, decided by analog-to-digital conversion module
Post-processing:	Digital image enhancement methods often used before compression and storage of captured image

Digital Image Sensor Properties

Shutter speed:	Controls the amount of light reaching the sensor (also called *exposure time*)
Sampling pitch:	Physical spacing between adjacent sensor cells on the imaging chip
Fill factor:	Active sensing area size as a fraction of the theoretically available sensing area (product of horizontal and vertical sampling pitches)
Chip size:	Size/area of the chip
Analog gain:	Amplification of the sensed signal using automatic gain control (AGC) logic (controlled using ISO setting on cameras)
Sensor noise:	Noise from various sources in the sensing process
Resolution:	How many bits for each pixel, decided by analog-to-digital conversion module
Post-processing:	Digital image enhancement methods often used before compression and storage of captured image

Digital Image Sensor Properties

Shutter speed:	Controls the amount of light reaching the sensor (also called *exposure time*)
Sampling pitch:	Physical spacing between adjacent sensor cells on the imaging chip
Fill factor:	Active sensing area size as a fraction of the theoretically available sensing area (product of horizontal and vertical sampling pitches)
Chip size:	Size/area of the chip
Analog gain:	Amplification of the sensed signal using automatic gain control (AGC) logic (controlled using ISO setting on cameras)
Sensor noise:	Noise from various sources in the sensing process
Resolution:	How many bits for each pixel, decided by analog-to-digital conversion module
Post-processing:	Digital image enhancement methods often used before compression and storage of captured image

Digital Image Sensor Properties

Shutter speed:	Controls the amount of light reaching the sensor (also called *exposure time*)
Sampling pitch:	Physical spacing between adjacent sensor cells on the imaging chip
Fill factor:	Active sensing area size as a fraction of the theoretically available sensing area (product of horizontal and vertical sampling pitches)
Chip size:	Size/area of the chip
Analog gain:	Amplification of the sensed signal using automatic gain control (AGC) logic (controlled using ISO setting on cameras)
Sensor noise:	Noise from various sources in the sensing process
Resolution:	How many bits for each pixel, decided by analog-to-digital conversion module
Post-processing:	Digital image enhancement methods often used before compression and storage of captured image

Digital Image Sensor Properties

Shutter speed:	Controls the amount of light reaching the sensor (also called *exposure time*)
Sampling pitch:	Physical spacing between adjacent sensor cells on the imaging chip
Fill factor:	Active sensing area size as a fraction of the theoretically available sensing area (product of horizontal and vertical sampling pitches)
Chip size:	Size/area of the chip
Analog gain:	Amplification of the sensed signal using automatic gain control (AGC) logic (controlled using ISO setting on cameras)
Sensor noise:	Noise from various sources in the sensing process
Resolution:	How many bits for each pixel, decided by analog-to-digital conversion module
Post-processing:	Digital image enhancement methods often used before compression and storage of captured image

Popular Question: *"With smartphones, do you need DSLR cameras?"*[3]

[3]Source: VSBytes.com

Popular Question: *"With smartphones, do you need DSLR cameras?"*[3]

- **DSLR - Digital Single Lens Reflex camera:** Uses a mirror mechanism to reflect light from lens to a viewfinder, or let light fully pass onto image sensor by moving the mirror out of the way
- Essentially a comparison between mirror and mirrorless cameras
- *Pros of mirrorless cameras:* Accessibility, portability, Low cost
- *Pros of DSLRs:* Picture quality, Versatility and functionality, Physical shutter, Variable focal length/aperture

Credit: http://www.pixelrajeev.com/

[3]Source: VSBytes.com

Sampling and Aliasing

Credit: WikimediaCommons

- **Shannon's Sampling Theorem:** $f_s \geq 2f_{max}$, where f_s is sampling rate, and f_{max} is maximum frequency in signal, also called **Nyquist frequency**
- Frequencies above Nyquist frequency or when Shannon's sampling rate is not met $\Longrightarrow$ *aliasing* happens

Sampling and Aliasing

Credit: WikimediaCommons

- **Shannon's Sampling Theorem:** $f_s \geq 2f_{max}$, where f_s is sampling rate, and f_{max} is maximum frequency in signal, also called **Nyquist frequency**
- Frequencies above Nyquist frequency or when Shannon's sampling rate is not met $\Longrightarrow$ *aliasing* happens
- Why is aliasing bad?

Sampling and Aliasing

Credit: WikimediaCommons

- **Shannon's Sampling Theorem:** $f_s \geq 2f_{max}$, where f_s is sampling rate, and f_{max} is maximum frequency in signal, also called **Nyquist frequency**
- Frequencies above Nyquist frequency or when Shannon's sampling rate is not met $\Longrightarrow$ *aliasing* happens
- Why is aliasing bad? Creates issues while downsampling and upsampling an image
- More coming in later lectures

Colour Space Representations

- Popular colour spaces: RGB, CMYK
- Additive colours: R, G, B Subtractive
- colours: C, M, Y
- Other colour spaces: XYZ, YUV, Lab, YCbCr, HSV
- Standards established by Commission Internationale d'Eclairage (CIE)
- Understanding of colour spaces important in printing industry


Credit: Szeliski, Computer Vision: Algorithms and Applications, 2010

For more information:

- https://www.tutorialspoint.com/dip/introduction_to_color_spaces.htm
- https://ciechanow.ski/color-spaces/

Image Compression

- Last stage in a camera's processing pipeline
- Convert signal into YCbCr (or variant), compress luminance with higher fidelity than chrominance
- Most common compression technique: **Discrete Cosine Transform (DCT)**, used in MPEG and JPEG
- DCT, a variant of Discrete Fourier Transform - a reasonable approximation of eigendecomposition of image patches

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

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

where $\hat{I}$ is compressed version of I

Image Compression

- Last stage in a camera's processing pipeline
- Convert signal into YCbCr (or variant), compress luminance with higher fidelity than chrominance
- Most common compression technique: **Discrete Cosine Transform (DCT)**, used in MPEG and JPEG
- DCT, a variant of Discrete Fourier Transform - a reasonable approximation of eigendecomposition of image patches

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

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

where $\hat{I}$ is compressed version of I

Image Compression

- Last stage in a camera's processing pipeline
- Convert signal into YCbCr (or variant), compress luminance with higher fidelity than chrominance
- Most common compression technique: **Discrete Cosine Transform (DCT)**, used in MPEG and JPEG
- DCT, a variant of Discrete Fourier Transform - a reasonable approximation of eigendecomposition of image patches

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

$$\text{where } MSE = \frac{1}{n} \sum_x [I(x) - \hat{I}(x)]^2$$

where $\hat{I}$ is compressed version of I

Part 2

- Image Representation

Question

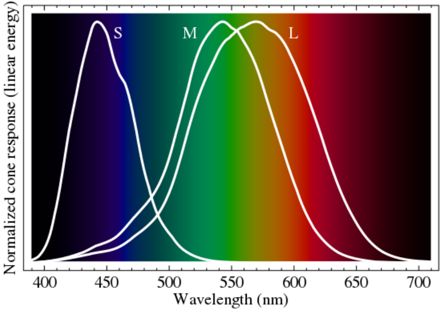
On Colour

If visible light spectrum is VIBGYOR, why RGB colour representation?

Question

On Colour

If visible light spectrum is VIBGYOR, why RGB colour representation?

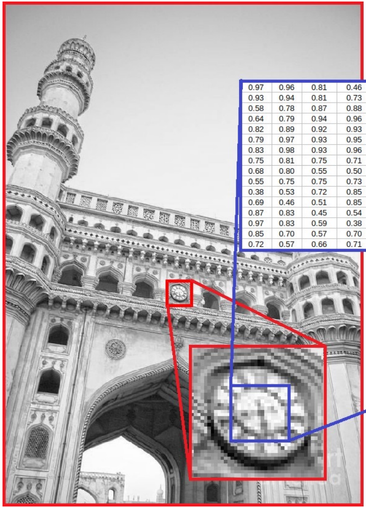


Credit: Derek Hoiem, UIUC

https://en.wikipedia.org/wiki/Color_vision

- Long (red), Medium (green), and Short (blue) cones, plus intensity rods
- Fun facts
  - "M" and "L" on the X-chromosome $\Longrightarrow$ men are more likely to be colour blind!
  - Some animals have 1 (night animals), 2 (e.g., dogs), 4 (fish, birds), 5 (pigeons, some reptiles/amphibians), or even 12 (mantis shrimp) types of cones

Image as a Matrix

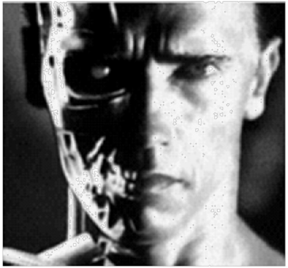


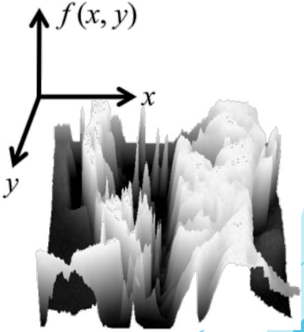
0.97	0.96	0.81	0.46	0.55	0.78	0.70	0.56	0.58	0.55	0.94	0.97	0.92	0.83	0.91	0.95
0.93	0.94	0.81	0.73	0.80	0.83	0.84	0.86	0.73	0.55	0.73	0.87	0.91	0.86	0.94	0.96
0.58	0.78	0.87	0.88	0.94	0.95	0.97	0.97	0.96	0.97	0.80	0.57	0.55	0.96	0.96	0.92
0.64	0.79	0.94	0.96	0.91	0.95	0.96	0.90	0.91	0.93	0.94	0.98	0.62	0.75	0.97	0.97
0.82	0.89	0.92	0.93	0.97	0.93	0.81	0.77	0.98	0.92	0.90	0.93	0.96	0.67	0.66	0.80
0.79	0.97	0.93	0.95	0.89	0.97	0.86	0.64	0.90	0.98	0.98	0.92	0.97	0.88	0.52	0.64
0.83	0.98	0.93	0.96	0.93	0.95	0.97	0.75	0.82	0.93	0.83	0.69	0.92	0.93	0.86	0.77
0.75	0.81	0.75	0.71	0.85	0.77	0.83	0.55	0.51	0.88	0.86	0.77	0.76	0.97	0.94	0.69
0.68	0.80	0.55	0.50	0.78	0.77	0.81	0.59	0.53	0.92	0.95	0.91	0.90	0.95	0.97	0.60
0.55	0.75	0.75	0.73	0.75	0.86	0.95	0.83	0.67	0.89	0.97	0.93	0.93	0.93	0.97	0.74
0.38	0.53	0.72	0.85	0.90	0.91	0.93	0.90	0.66	0.70	0.92	0.95	0.97	0.96	0.90	0.72
0.69	0.46	0.51	0.85	0.96	0.92	0.90	0.83	0.55	0.54	0.84	0.94	0.89	0.88	0.89	0.69
0.87	0.83	0.45	0.54	0.75	0.85	0.97	0.91	0.63	0.61	0.84	0.93	0.79	0.70	0.66	0.40
0.97	0.83	0.59	0.38	0.52	0.58	0.76	0.83	0.72	0.59	0.69	0.75	0.62	0.54	0.47	0.61
0.85	0.70	0.57	0.70	0.61	0.55	0.47	0.58	0.64	0.49	0.60	0.58	0.77	0.88	0.59	0.54
0.72	0.57	0.66	0.71	0.93	0.96	0.64	0.54	0.61	0.60	0.82	0.65	0.77	0.94	0.93	0.80

- Common to use one byte per value: 0 = black, 255 = white
- One such matrix for every channel in colour images

Image as a Function

- We can think of a (grayscale) image as a function $f : \mathrm{R}^2 \rightarrow \mathrm{R}$, giving the intensity at position (x, y)
- A digital image is a discrete (sampled, quantized) version of this function





Credit: Noah Snavely, Cornell Univ

Image Transformations

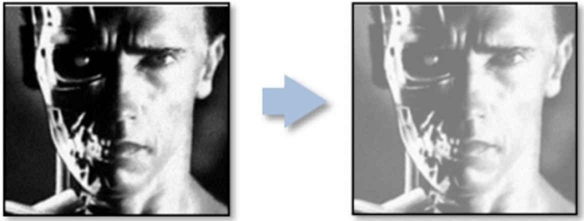
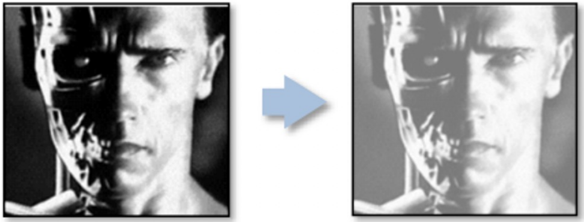


Image Transformations

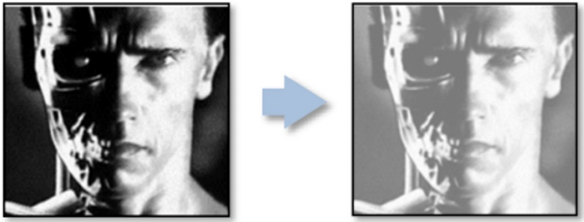
$$\hat{I}(x,y) = I(x,y) + 20$$

Image Transformations



$$\hat{I}(x,y) = I(x,y) + 20$$

Image Transformations

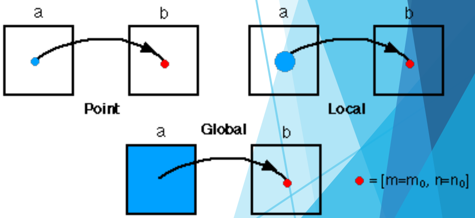


$$\hat{I}(x,y) = I(x,y) + 20$$

$$\hat{I}(x,y) = I(-x,y)$$

Image Processing Operations

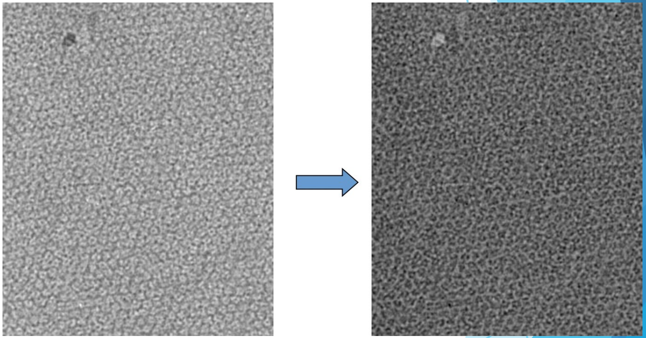
- **Point Operations**
  - Output value at (m_0, n_0) is dependent only on the input value at the same coordinate
  - *Complexity/pixel:* Constant
- **Local Operations**
  - Output value at (m_0, n_0) is dependent on input values in a $p \times p$ neighborhood of that same coordinate
  - *Complexity/pixel:* p^2
- **Global Operations**
  - Output value at (m_0, n_0) is dependent on on all the values in the input $N \times N$ image
  - *Complexity/pixel:* N^2



Point Operations: Example

- Image Enhancement:
  Reversing the contrast

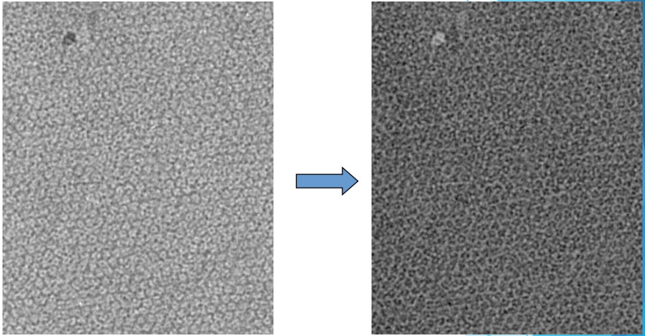
- How?



Point Operations: Example

- Image Enhancement:
  Reversing the contrast

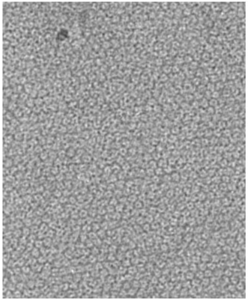
- How?

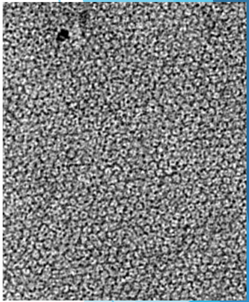


$$\hat{I}(m_0, n_0) = I_{MAX} - I(m_0, n_0) + I_{MIN}$$

Point Operations: Another Example

- Image Enhancement: Stretching the contrast
- How?

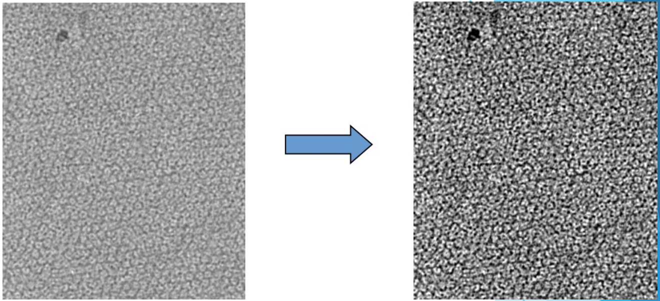




Point Operations: Another Example

- Image Enhancement:
  Stretching the contrast

- How?



Linear Contrast Stretching

$$\hat{I}(m_0, n_0) =$$

$$I(m_0, n_0) - \min_{x,y} I(x, y) \quad * \; (I_{MAX} - I_{MIN})/(\max_{x,y} I(x, y) - \min_{x,y} I(x, y)) \quad + I_{MIN}$$

Going Beyond Linear Stretching

Question

Heard about **Histogram Equalization**? Read about it, homework!

How Useful are Point Operations?

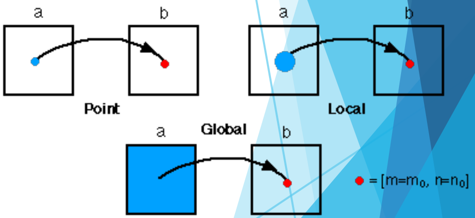
- A single point (or pixel)’s intensity is influenced by multiple factors, and may not tell us everything
  - Light source strength and direction
  - Surface geometry, material and nearby surfaces Sensor capture properties
  - Image representation and colour
- Given a camera and a still scene, how do you reduce noise using point operations?

How Useful are Point Operations?

- A single point (or pixel)'s intensity is influenced by multiple factors, and may not tell us everything
  - Light source strength and direction
  - Surface geometry, material and nearby surfaces Sensor capture properties
  - Image representation and colour
- Given a camera and a still scene, how do you reduce noise using point operations? Take many images, and average them!
- You need local operations otherwise. What is the local operation?

Image Processing Operations

- Point Operations
  - Output value at (m_0, n_0) is dependent only on the input value at the same coordinate
  - *Complexity/pixel:* Constant
- Local Operations
  - Output value at (m_0, n_0) is dependent on input values in a $p \times p$ neighborhood of that same coordinate
  - *Complexity/pixel:* p^2
- Global Operations
  - Output value at (m_0, n_0) is dependent on on all the values in the input $N \times N$ image
  - *Complexity/pixel:* N^2



Local Operation Examples: Moving Average

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

	0								

Credit: Steve Seitz, Univ of Washington

Local Operation Examples: Moving Average

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

	0	10							

Credit: Steve Seitz, Univ of Washington

Local Operation Examples: Moving Average

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

	0	10	20						

Credit: Steve Seitz, Univ of Washington

Local Operation Examples: Moving Average

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

	0	10	20	30					

Credit: Steve Seitz, Univ of Washington

Local Operation Examples: Moving Average

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

	0	10	20	30	30				

Credit: Steve Seitz, Univ of Washington

Local Operation Examples: Moving Average

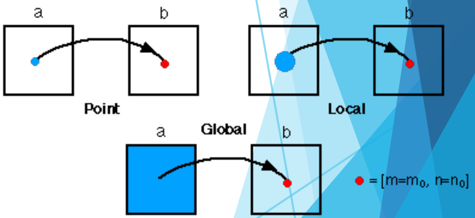
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

	0	10	20	30	30	30	20	10	
	0	20	40	60	60	60	40	20	
	0	30	60	90	90	90	60	30	
	0	30	50	80	80	90	60	30	
	0	30	50	80	80	90	60	30	
	0	20	30	50	50	60	40	20	
	10	20	30	30	30	30	20	10	
	10	10	10	0	0	0	0	0	

Credit: Steve Seitz, Univ of Washington

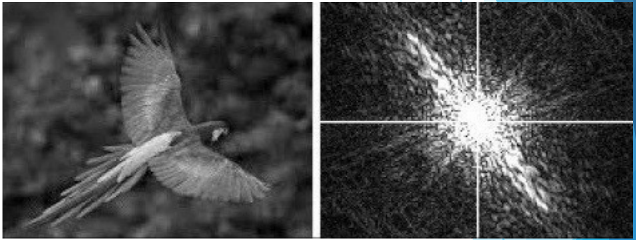
Image Processing Operations

- Point Operations
  - Output value at (m_0, n_0) is dependent only on the input value at the same coordinate
  - *Complexity/pixel:* Constant
- Local Operations
  - Output value at (m_0, n_0) is dependent on input values in a $p \times p$ neighborhood of that same coordinate
  - *Complexity/pixel:* p^2
- Global Operations
  - Output value at (m_0, n_0) is dependent on on all the values in the input $N \times N$ image
  - *Complexity/pixel:* N^2



Global Operations: Examples

- Image coordinate transformations, e.g. Fourier transform
- We will see more of this later



Credit: Mathworks MATLAB Toolbox

Part 3

- Feature Extraction for Image Computing

Contents

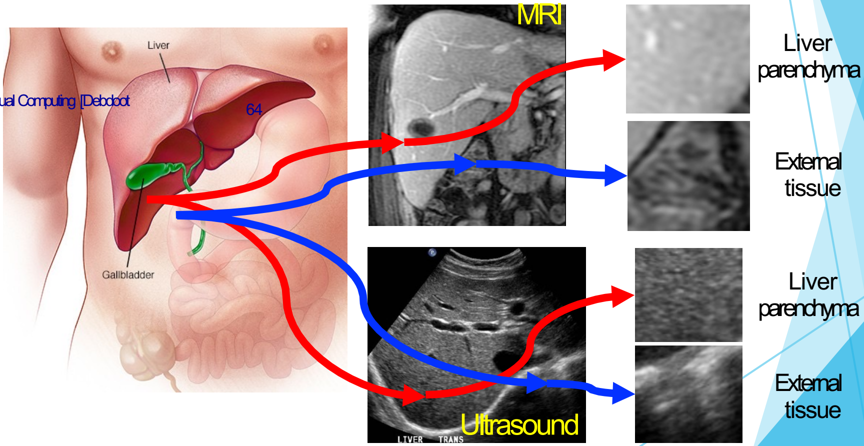
- Introductory concepts
- Texture characterization – statistical vs. structural
- Co-occurrence matrices
- Orientation histograms
- Local binary patterns (LBP)
- Texture from Fourier features
- Wavelets

Introductory Concepts

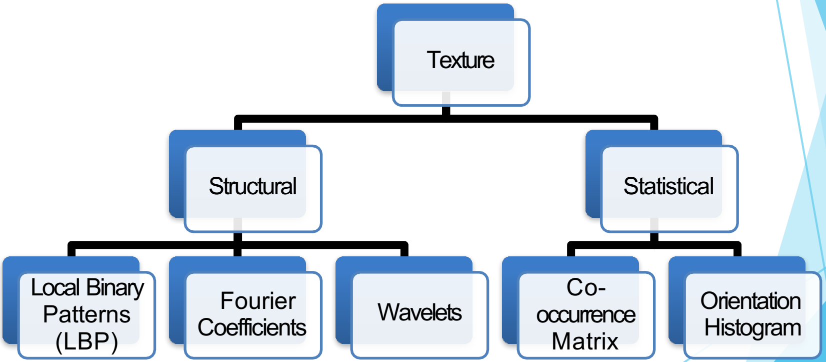
Texture
Local variation in intensity due to tissue heterogeneity

Noise
Uncertainty in image sensing by the instrument

Structural vs. Statistical Textures



Family of Texture Metrics



Local Binary Patterns

10	12	9
6	7	19
7	10	16

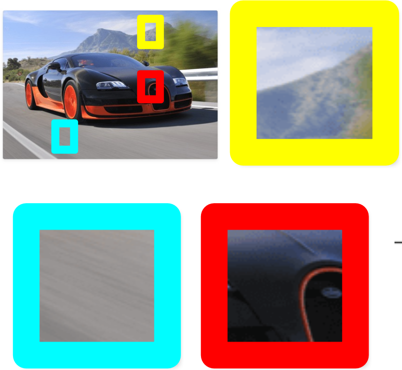
$$b_k = \begin{cases} 1 & if\ g_k \geq g(\mathbf{x}) \\ 0 & otherwise \end{cases}$$

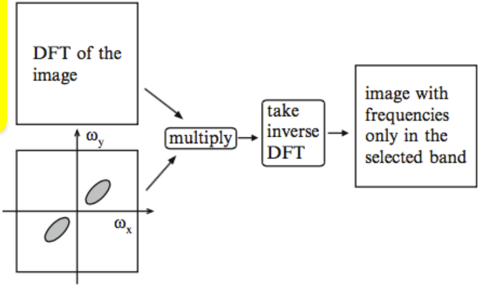
$$LBP_{ri}(\mathbf{x}) = min\{P_j\}$$

6	10	12
7	7	9
10	16	19

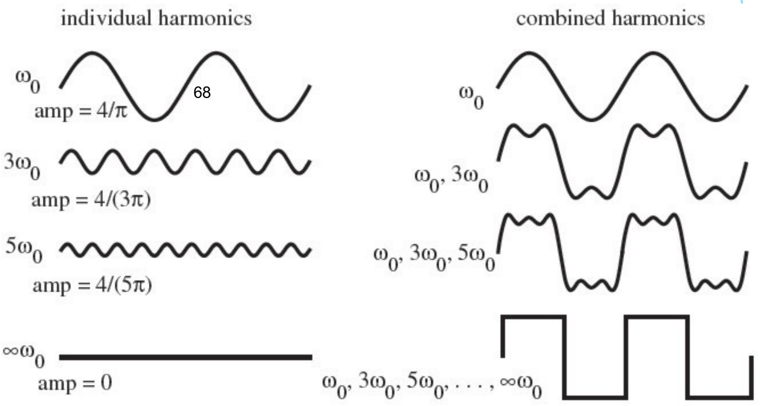
1	1	1	1	1	1	1	0	P_0=254
1	1	1	1	1	1	0	1	P_1=253
1	1	1	1	1	0	1	1	P_2=251
1	1	1	1	0	1	1	1	P_3=247
1	1	1	0	1	1	1	1	P_4=239
1	1	0	1	1	1	1	1	P_5=223
1	0	1	1	1	1	1	1	P_6=191
0	1	1	1	1	1	1	1	P_7=127

Texture from Fourier Features





Wavelet Texture Descriptors



Laws Masks

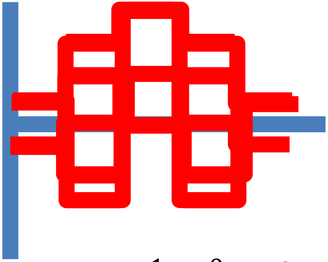
$Level\colon L_5 = [1, 4, 6, 4, 1]$

$Edge\colon E_5 = [-1, -2, 0, 2, 1]$

$Spot\colon S_5 = [-1, 0, 2, 0, -1]$

$Wave\colon W_5 = [-1, 2, 0, -2, 1]$

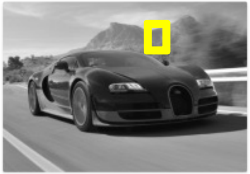
$Ripple\colon R_5 = [1, -4, 6, -4, 1]$

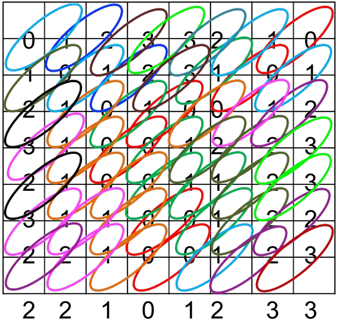


$$L_5^T \times S_5 = \begin{bmatrix} -1 & 0 & 2 & 0 & -1 \\ -4 & 0 & 8 & 0 & -4 \\ -6 & 0 & 12 & 0 & -6 \\ -4 & 0 & 8 & 0 & -4 \\ -1 & 0 & 2 & 0 & -1 \end{bmatrix}$$

$$E_5^T \times S_5 = \begin{bmatrix} 1 & 0 & -2 & 0 & 1 \\ 2 & 0 & -4 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & 0 & 4 & 0 & -2 \\ -1 & 0 & 2 & 0 & -1 \end{bmatrix}$$

Co-occurrence Matrices





	0	1	2	3
0	6	6	2	0
1	7	5	4	2
2	1	5	4	3
3	0	2	1	1

Operator: NE 1px

	0	1	2	3
0	6/49	6/49	2/49	0
1	1/7	5/49	4/49	2/49
2	1/49	5/49	4/49	3/49
3	0	2/49	1/49	1/49

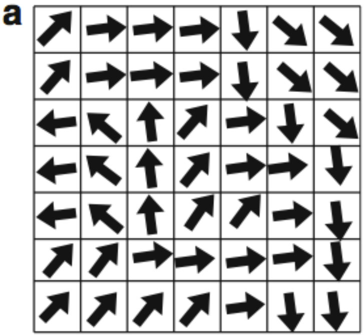
Probability

$$Energy = \sum_{a,b} P_{\varphi,d}^{2}(a,b)$$

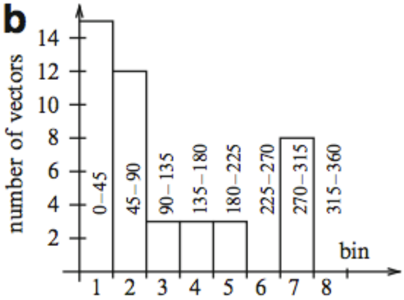
$$Entropy = \sum_{a,b} P_{\varphi,d}(a,b)\log\left(P_{\varphi,d}(a,b)\right)$$

$$Contrast = \sum_{a,b} |a-b|^{\kappa} P_{\varphi,d}^{\lambda}(a,b)$$

Orientation Histogram



Local orientation in an image



Histogram of orientation

Homework

Readings

- Chapter 3.1, Szeliski, *Computer Vision: Algorithms and Applications*
- M. Petrou and P. G. Sevilla, Image Processing Dealing with Texture, John Wiley and Sons, Ltd. 2006.

Questions to Answer

What is histogram equalization, and how do you derive its formula?

References

- Richard Szeliski. *Computer Vision: Algorithms and Applications*. Texts in Computer Science. London: Springer-Verlag, 2011.
- *Hoiem, Derek, CS 543 - Computer Vision (Spring 2011)*. URL: https://courses.engr.illinois.edu/cs543/sp2017/