EE 4211 Computer Vision

Lecture 13: Review

Semester A, 2020-2021

Schedules

Week	Date	Topics
1	Sep. 4	Introduction/Imaging
2	Sep. 11	Image enhancement in spatial domain
3	Sep. 18	Image enhancement in frequency domain (HW1 out)
4	Sep. 25	Morphological processing
5	Oct. 2	Image restoration(HW1 due)
6	Oct. 9	Image restoration
7	Oct. 16	Midterm (no tutorials this week)
8	Oct. 23	Edge detection (HW2 out, illustrate the project)
9	Oct. 30	Image segmentation (HW2 due)
10	Nov. 6	Face recognition with PCA, LDA (tutorial on deep learning framework)
11	Nov. 13	Face recognition based on deep learning Image segmentation based on deep learning (tutorial on coding)
12	Nov. 20	Object detection with traditional methods (Quiz) Object detection based on deep learning
13	Nov. 27	Project presentation and summary
14	Dec. 4	Review and Summary

Project presentation

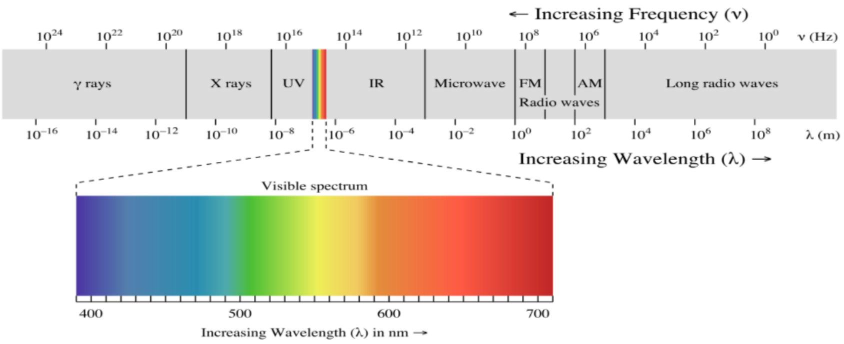
54777375	FANG Yucheng		5	Classification
55202608	HAO Yining		5	Classification
55203003	HE Jinwei		5	Classification
55200351	XU Tianxiao		5	Classification
55202909	YANG Qihan		5	Classification
55228071	CHAN Ding Wan		6	Segmentation
55234090	KWOK Wah Chung		6	Segmentation
55218956	LAW Yick Ming		6	Segmentation
55240232	TSANG Yiu Wai		6	Segmentation
54816829	WONG Ka Ho		6	Segmentation
55218354	AU Cheuk Ming		7	Classification
54845745	TSE Siu Ki		7	Classification
55218551	BUTT Ka Kiu		7	Classification
55240293	CHAN Ka Hung		7	Classification
55210290	SHAM Pui Hei		7	Classification
54818720	SIN Joy Leong		8	Segmentation
55293157	TANG Ho Yan Sophia		8	Segmentation
54400770	WONG Chak Kong		8	Segmentation
55231192	YAU Hiu Kwan		8	Segmentation

SID	NAME	GROUP		TASK
55228648	HE Guan Cheng		1	Segmentation
55234839	LAM King Hei		1	Segmentation
55224293	LEE Hung Hin		1	Segmentation
55216557	TSOI Hing Chung Sunny		1	Segmentation
55236833	WONG Yat Chau		1	Segmentation
55236476	LAM Hau Sze		2	Classification
55211875	LI Cheuk Yin		2	Classification
55235166	SIU Wai Fung		2	Classification
55099428	YU Tsz Ho		2	Classification
40132316	BARRIOS VIDAL Lucia Alejandra		2	Classification
54814615	CHAN Chung Hong		3	Classification
54398836	CHENG Kai Ming		3	Classification
55299045	LEE Man Lok		3	Classification
54845886	TANG Chun Ki		3	Classification
55297200	WONG Kai Chun		3	Classification
55237079	CHAN Cheuk Fung		4	Segmentation
55239915	HO Shing Fung		4	Segmentation
55238746	LEE Wing Yan		4	Segmentation
54850792	LUI Yu Yuet		4	Segmentation
55241843	NG Chak Ka		4	Segmentation

- 10 mins for presentation and 5 mins for Q & A
- 12:00 PM-1:00 PM presentation; 1:10 PM-2:10 PM presentation
- 2:20 PM-2:50 PM review

Lecture 1: Digital image fundamentals

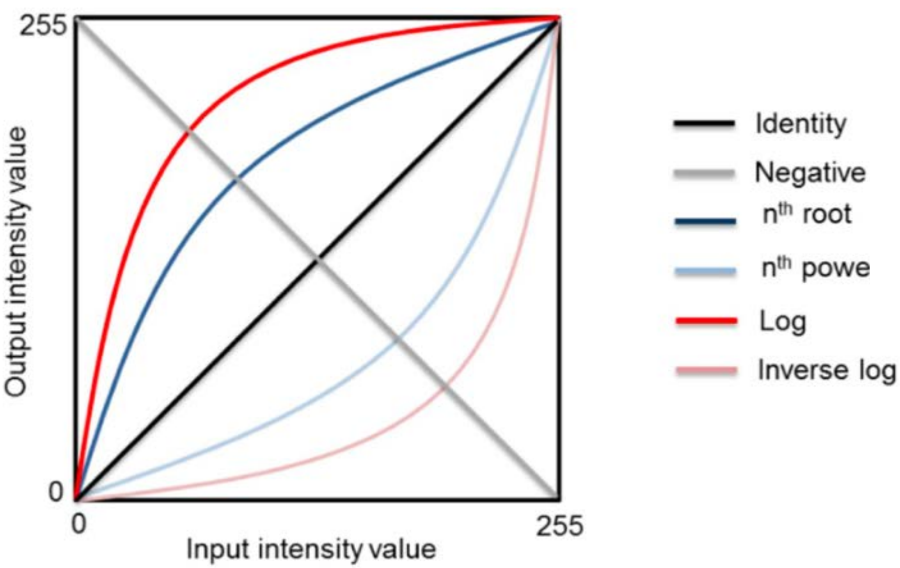
- Section A
  - Different image systems
  - Color Models: RGB, CMYK, HSV, YIQ, YCbCr
  - The differences, the suitable applications
- Section B
  - Image acquisition and formation
  - Resolution and size of objects
  - Spatial sampling and quantization

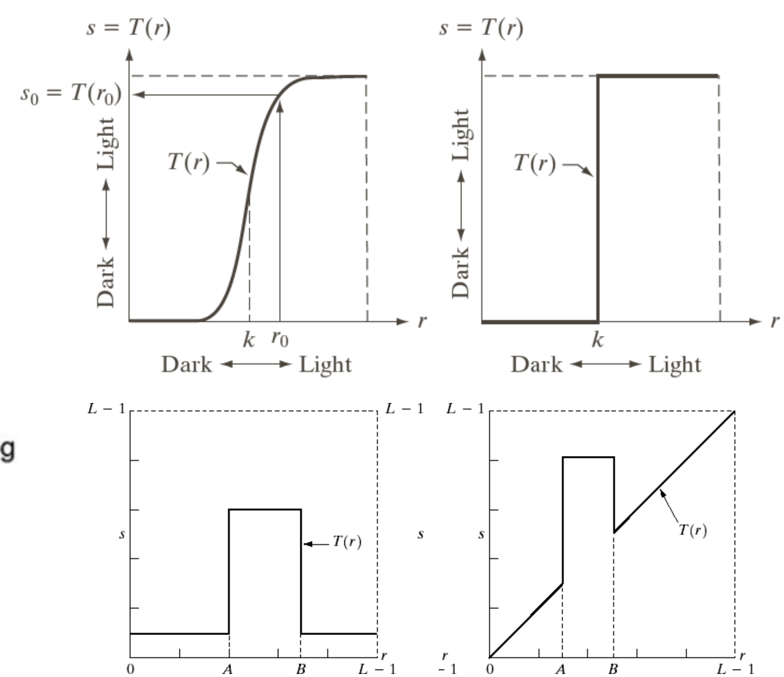


- For Example, a 8-bits 1920x1080 Full HD gray level image,
  - 1920 x 1080 x 8/8 = 2.0736 MB
  - 8bits=1Byte

Lecture 2: Image enhancement: Spatial Domain

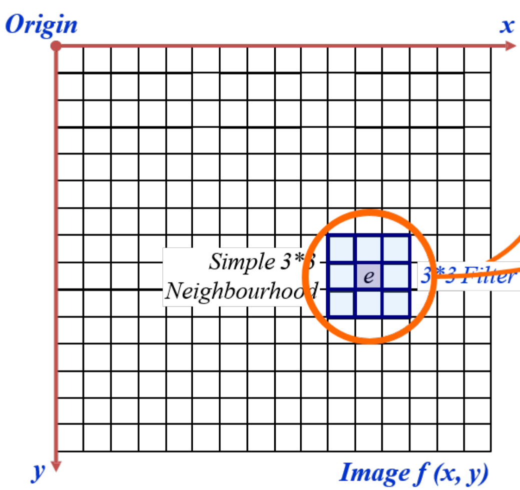
- Point processing – Gray values change without any knowledge of its surroundings (Part I)
  - Log, power-law, linear, piecewise linear (Contrast Stretching, Intensity-level Slicing)
  - Histogram Equalization (pdf->cdf->transformation) (calculate)

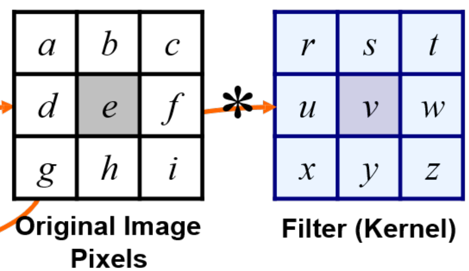




Lecture 2: Image enhancement: Spatial Domain

- Neighborhood processing (filtering) – Gray values change depending on the gray values in a small neighborhood of pixels around the given pixel (Part I)
  - Smoothing filters: Averaging filters, Order-Statistics filters
  - Median filters
  - Sharpening: Laplacian filters, Sobel filter





$$e_{processed} = v \cdot e + r \cdot a + s \cdot b + t \cdot c + u \cdot d + w \cdot f + x \cdot g + y \cdot h + z \cdot i$$

0	−1	0
−1	4	−1
0	−1	0

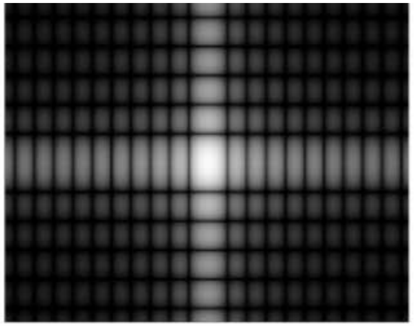
−1	−1	−1
−1	8	−1
−1	−1	−1

Lecture 3: Image enhancement: Frequency Domain

- Frequency transformation (matching image with spectrum)
- Image Enhancement in Frequency Domain (workflow)
- Filtering in Frequency Domain
  - Low-Pass Filtering (ideal, butterworth, gaussian): passes all frequencies with magnitudes below a specific level, and attenuates all frequencies above that level.
  - High-Pass Filtering: does the opposite.
  - Laplacian Filtering
  - Homomorphic Filtering (Principle, separating Illumination and Reflection)
  - Selective Filtering – Bandpass/Bandreject, Notch Filters (A “notch” filter rejects (or passes) frequencies at a specific point)

Important Fourier Transform Pairs

rectangle centred at origin
with sides of length X and Y



$f(x,y)$ $|F(u,v)|$

When X is larger than Y, then the value

$$F(u,v) = \int\int f(x,y)e^{-j2\pi(ux+vy)}dxdy,$$

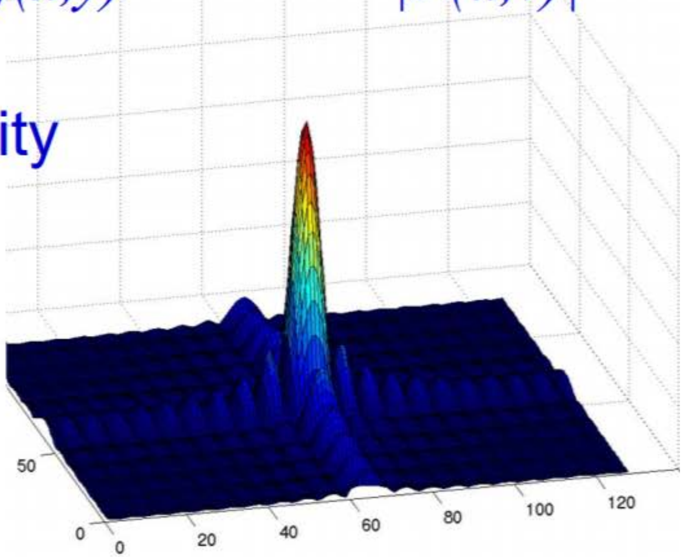
$$= \int_{-X/2}^{X/2} e^{-j2\pi ux}dx \int_{-Y/2}^{Y/2} e^{-j2\pi vy}dy,$$ separability

$$= \left[\frac{e^{-j2\pi ux}}{-j2\pi u}\right]_{-X/2}^{X/2} \left[\frac{e^{-j2\pi vy}}{-j2\pi v}\right]_{-Y/2}^{Y/2},$$

$$= \frac{1}{-j2\pi u}\left[e^{-juX} - e^{juX}\right] \frac{1}{-j2\pi v}\left[e^{-jvY} - e^{jvY}\right],$$

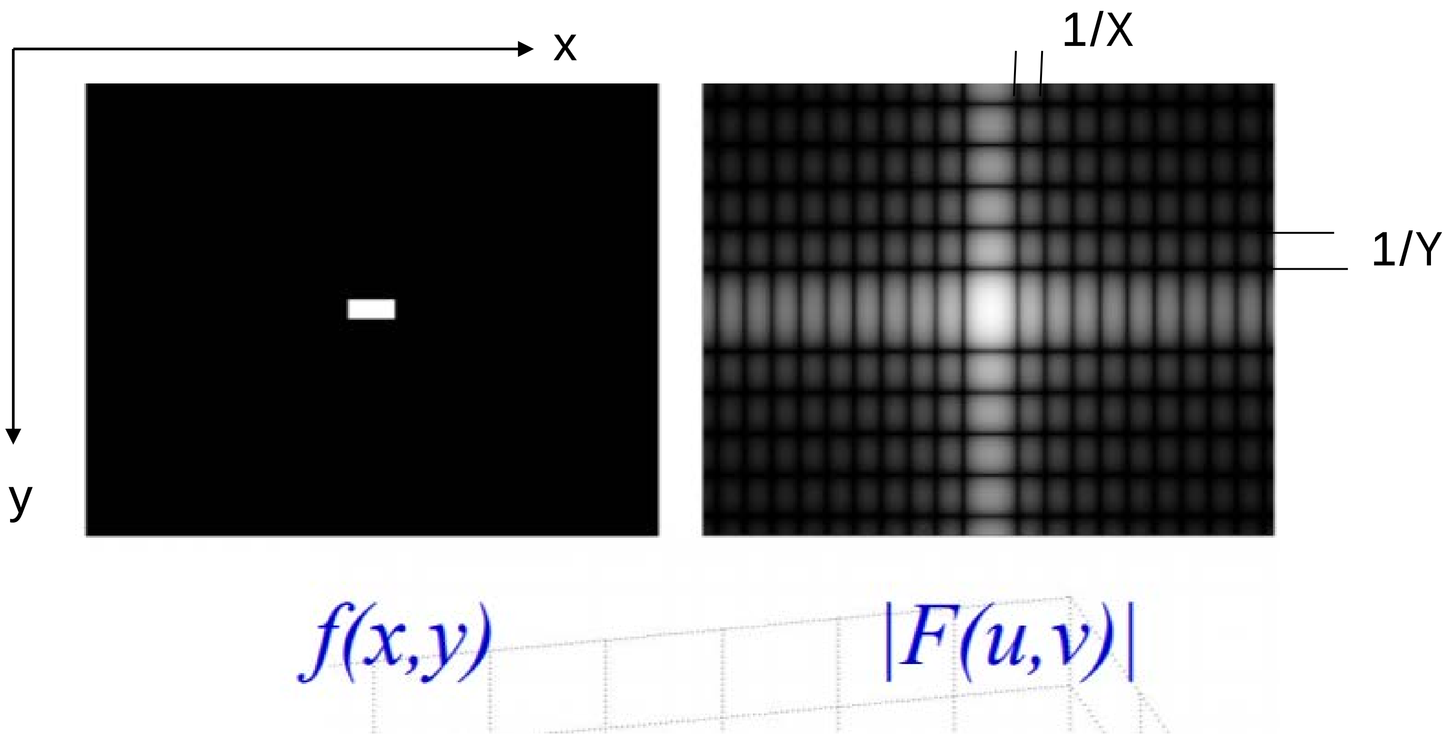
$$= XY\left[\frac{sin(\pi Xu)}{\pi Xu}\right]\left[\frac{sin(2\pi Yv)}{\pi Yv}\right]$$

$$= XY\mathrm{sinc}(\pi \mathrm{Xu})\mathrm{sinc}(\pi \mathrm{Yv}).$$



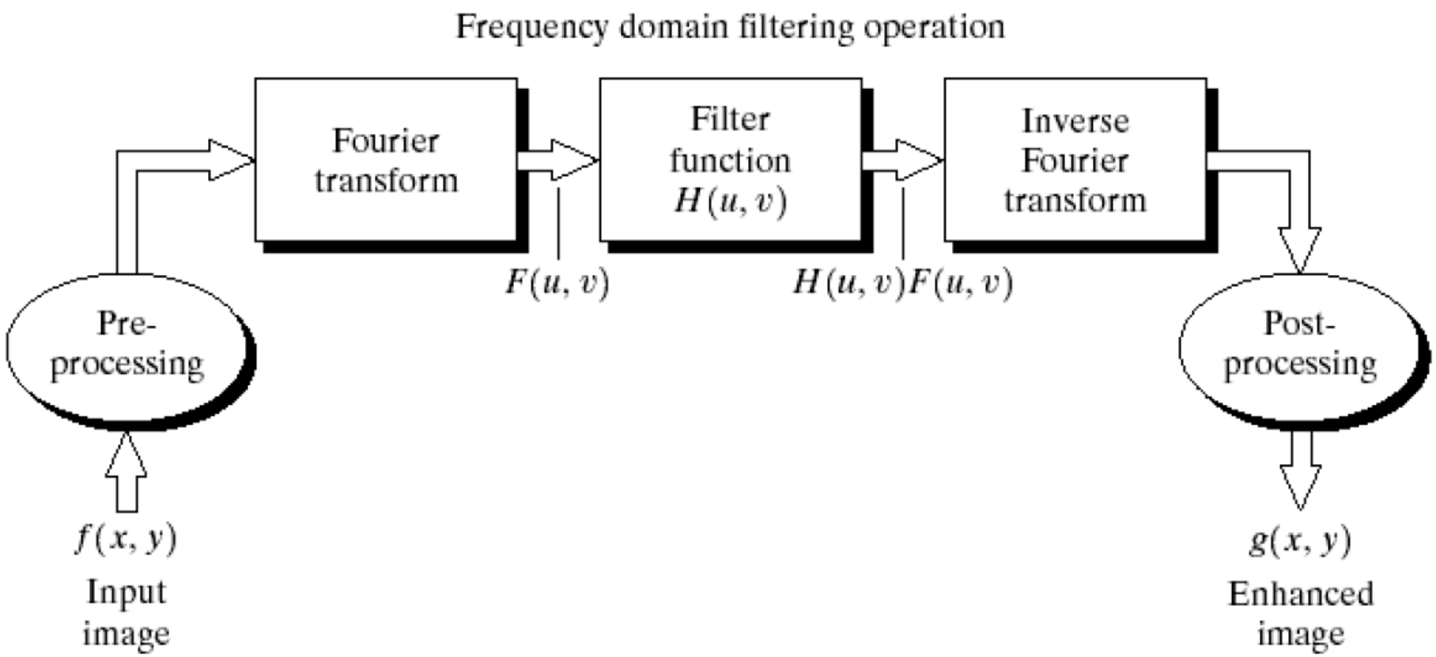
$|F(u,v)|$

Important Fourier Transform Pairs

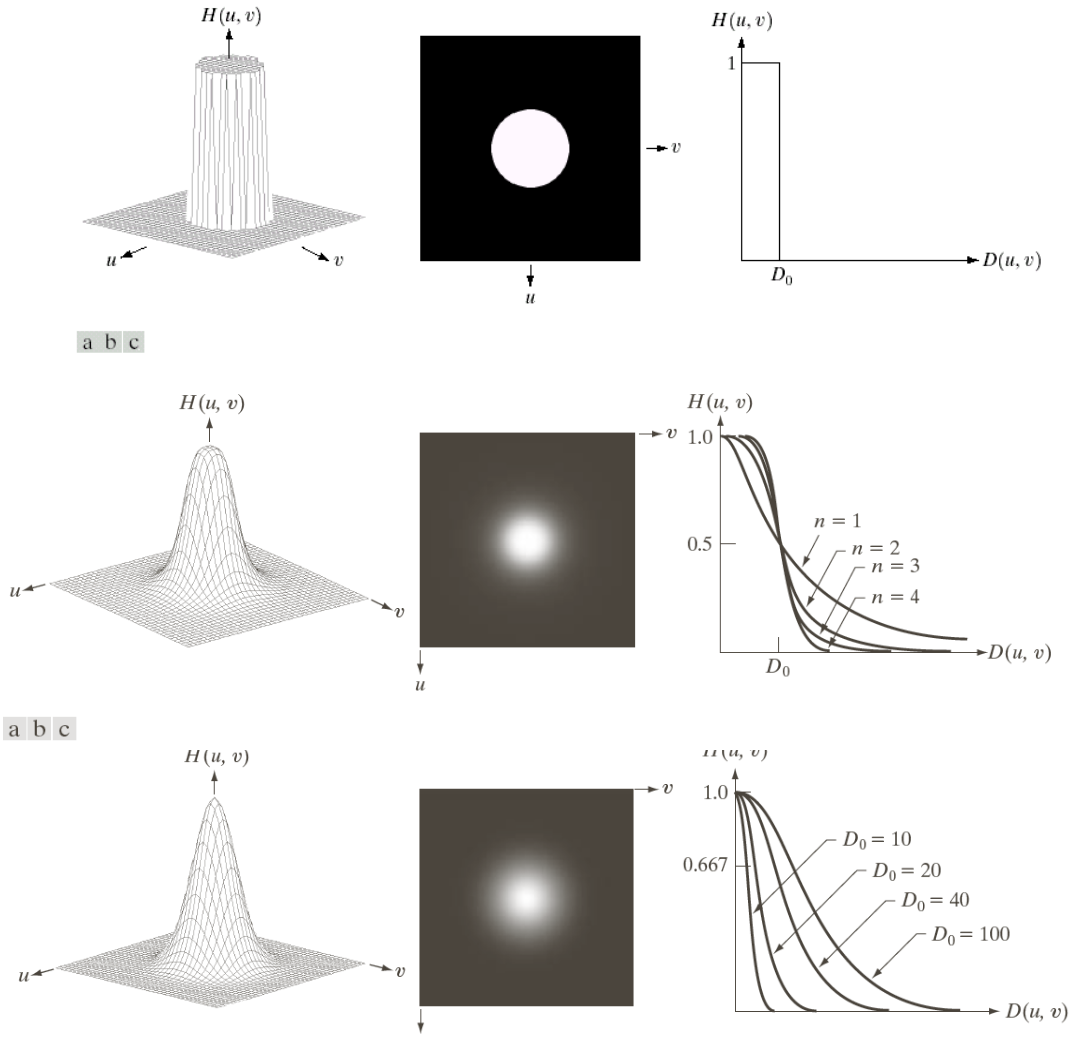


Lecture 3: Image enhancement: Frequency Domain

- Image Enhancement in Frequency Domain (workflow)



Comparison with different filter banks



a b c (a)Perspective plot of the transfer function
(b)Filter displayed as an image
(c)Filter radial cross sections for various values

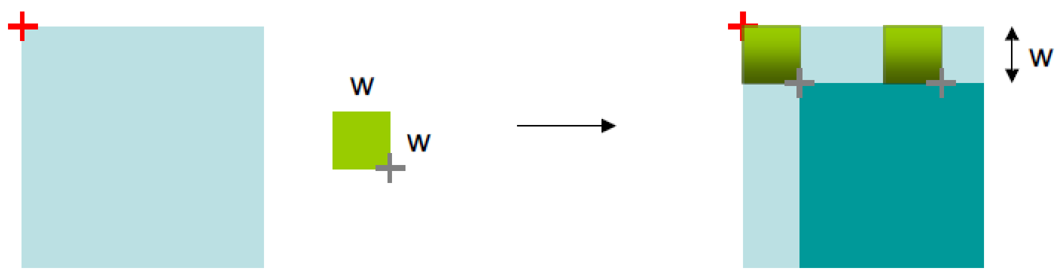
- Ideal low filter bank
  - all frequencies inside a circle of radius D_0 are passed with no attenuation
  - Presence of ripples/waves whenever there are boundaries in the image – “ringing effect”
- Butterworth lowpass filter
  - Reduces “ringing” while keeping clear cutoff
  - Tradeoff between amount of ringing and sharpness of cutoff
- Gaussian lowpass filter
  - No ringing, but allows high frequencies to pass

Lecture 4: Morphological processing

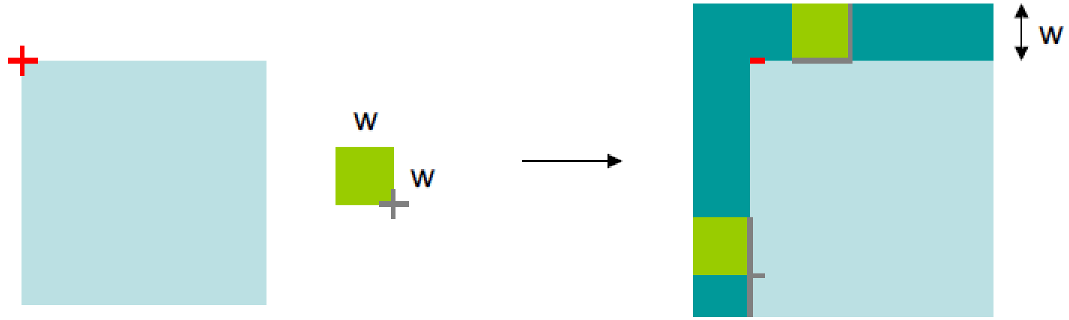
- Basic Morphological operations (calculate)
  - Hit: If any **ONE** of the ‘1’s in the SE is covered by the image, We say that the SE **hits** the image at the pixel position (the one on which the SE is centered).
  - Fit: If **ALL** of the ‘1’s in the SE are covered by the image, we say that the SE **fits** the image at the pixel position (the one on which the SE is centered).
  - Dilation
  - Erosion
  - Opening: Erosion followed by Dilation
  - Closing: Dilation followed by Erosion

Lecture 4: Morphological processing

- Applying **Fit** to an entire image is denoted Erosion
- Interpretation: shift B by z, if it is completely inside A, output a 1
- Applying **Hit** to an entire image is denoted Dilation
- Interpretation: reflect B, shift by z, if it overlaps with A, output a 1 at the center of B



$A \ominus B$



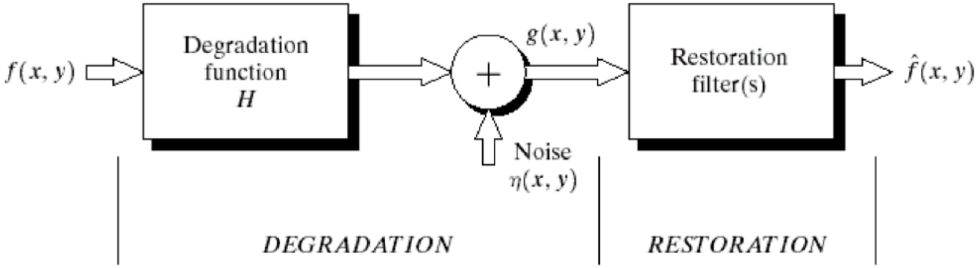
$A \oplus B$

Lecture 4: Morphological processing

- Morphological Algorithms (understand and can calculate with formula)
  - Boundary Extraction
  - Hole Filling
  - Connected Components
  - Skeletons

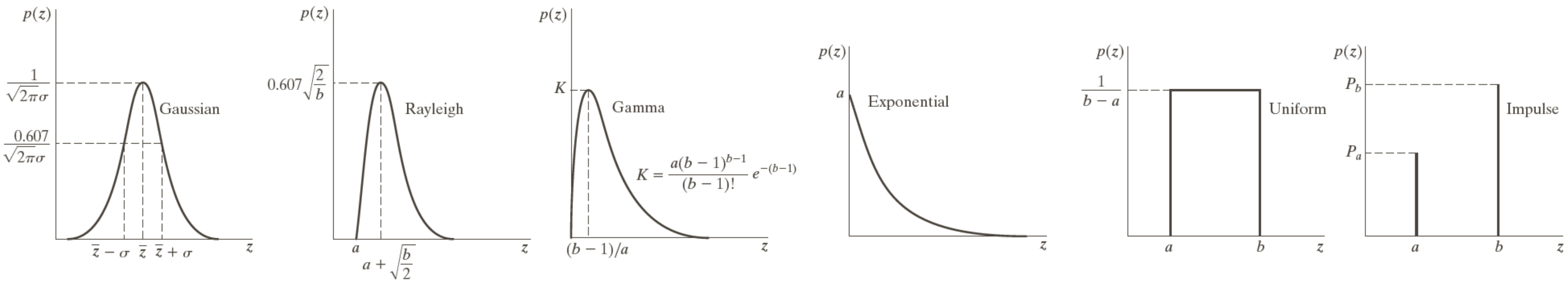
Lecture 6: Image restoration

- Degradation & Restoration Process Models (Principle)
  - Consists of 2 parts – Degradation function & Noise function



$$G(u,v) = H(u,v)F(u,v) + N(u,v)$$

- Noise Models (6 types)



Lecture 6: Image restoration

- Restoration in the presence of noise only (spatial filtering)
  - Mean filters: Arithmetic mean, Geometric Mean, Harmonic Mean Filter, Contraharmonic Mean
  - Order-statistics filters: min, max, median, midpoint, alpha-trimmed mean
  - Adaptive mean filters: preserve edge
- Estimation of the degeneration function
  - Observation
  - Experimentation
  - Modeling (motion blur, atmospheric blur)
- Inverse filtering (challenges and solutions) $G(u,v) = H(u,v)F(u,v) + N(u,v)$
- Wiener filtering (principle and differences, Find an estimate of the uncorrupted image such that the mean square error between them is minimum)

Example

- Given the Contraharmonic Mean Filter and the image patch, please calculate the filter results in the center

$$\hat{f}(x,y) = \frac{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q+1}}{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q}}$$

1	200	200
200	200	200
200	200	200

(a)

1	1	1
1	1	200
1	1	1

(b)

- (1) if Q=-1 for image patch a and b
- (2) if Q=1 for image patch a and b

Lecture 8: Segmentation – Categories

- Edge-based Segmentation
  - Finding boundary between adjacent regions
- Threshold-based Segmentation
  - Finding regions by grouping pixels of similar gray values
- Region-based Segmentation
  - Finding regions directly using growing or splitting
- Motion-based Segmentation
  - Finding regions by comparing successive frames of a video sequence to identify regions that correspond to moving objects

Lecture 8: Edge-based Segmentation

- Finding discontinuities (sharp, local changes in intensity) as boundary of regions
- Discontinuities in digital images (Spatial filters)
  - Point (Laplacian)
  - Line (horizontal, -45 degree, vertical, 45 degree)
  - Edges (Roberts, Prewitt, Sobel, LOG, Canny)
- Techniques
  - Point detection
  - Edge (pixel) detection
  - Edge formation from edge pixels – Edge linking, Hough Transform

Lecture 8: Edge-based Segmentation

- Finding discontinuities (sharp, local changes in intensity) as boundary of regions
- Discontinuities in digital images (Spatial filters)
  - Point (Laplacian)
  - Line (horizontal, -45 degree, vertical, 45 degree)

$$\begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}$$

  - Edges (Roberts, Prewitt, Sobel, LOG, Canny)

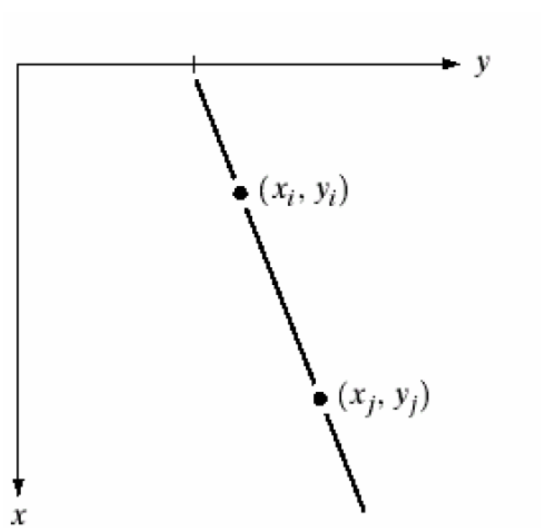
$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

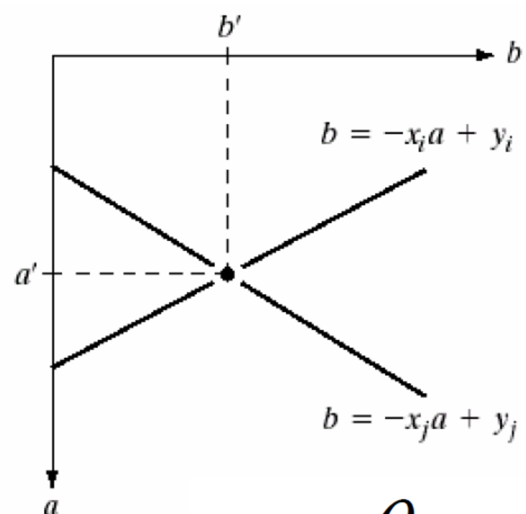
Non-maximal suppression
Double threshold

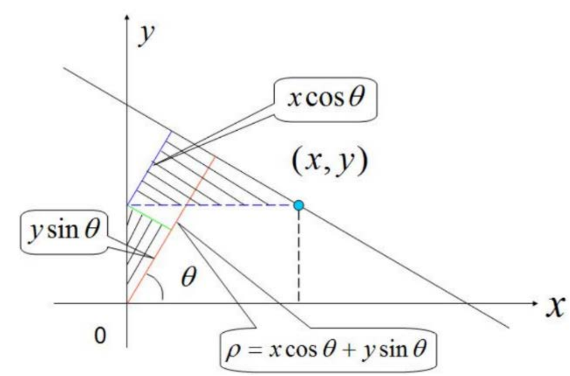
- Techniques
  - Point detection
  - Edge (pixel) detection
  - Edge formation from edge pixels – Edge linking, Hough Transform

Hough transform

- Motivation: isolated points
- Problem with this method: *a* approaches infinity as the line gets perpendicular to the *x* axis. (if a line is perpendicular to x axis, then this line is represented x=M. a->infinity)
- Solution: use the representation of the line as:



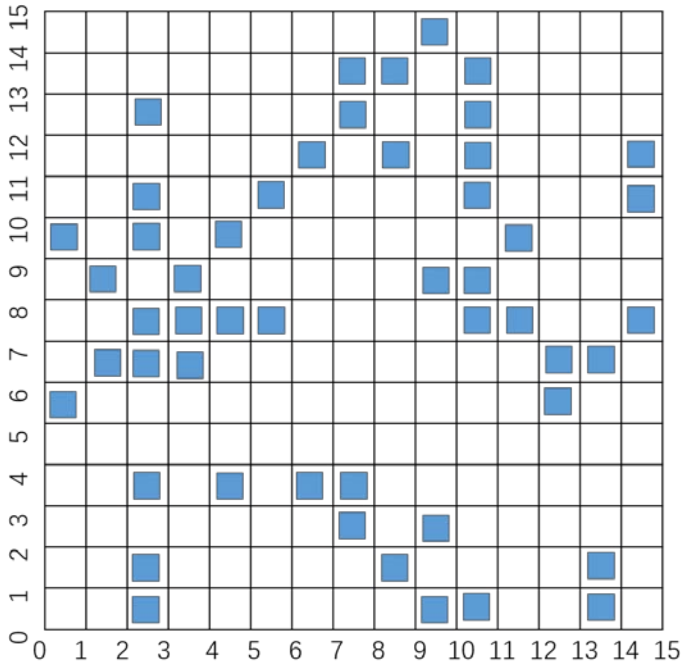




$$x\cos\theta + y\sin\theta = \rho$$

Example

- If we apply the Hough transform on the image below, what would be the maximum values for the accumulator cell in the (ρ, θ) space? What are the corresponding (ρ, θ) values.



Lecture 9: Threshold and region based segmentation

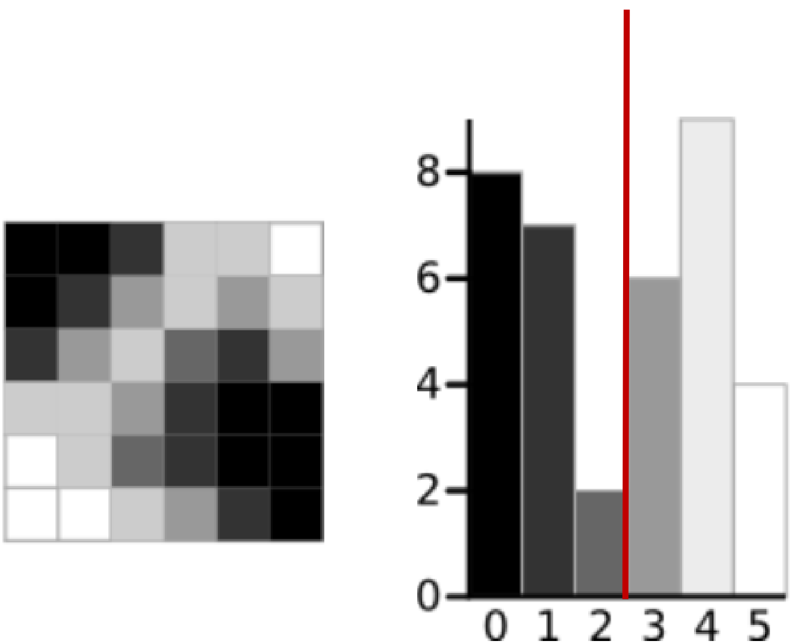
- Threshold based segmentation
  - Basic Global Thresholding
  - Optimum Global Thresholding (OSTU)
  - Multiple Thresholding
  - Variable Thresholding
- Region based segmentation
  - Region grow
  - K-means methods
- Understand principle of these segmentation methods and able to calculate the steps in the segmentation.

Algorithm: Basic Global Thresholding

- Select an initial estimate for the global threshold, *T*
- Threshold the image using *T*. This will produce two groups of pixels: G_1 consisting of all pixels with intensity values > *T*, and G_2 consisting of pixels with values ≤ *T*.
- Compute the average (mean) intensity values *m1* and *m2* for the pixels in G_1 and G_2, respectively.
- Compute a new threshold value: $T = \frac{1}{2}(m_1 + m_2)$
- Repeat Steps 2 through 4 until the difference between values of T in successive iterations is smaller than a predefined parameter ΔT.

Otsu's Method

- View thresholding as a statistical-decision theory problem
- Objective: Maximizes the between-class variance (minimize the within-class variance)



When the threshold is 3, the calculation of background

$$\text{Weight} \quad W_b = \frac{8+7+2}{36} = 0.4722$$

$$\text{Mean} \quad \mu_b = \frac{(0\times 8)+(1\times 7)+(2\times 2)}{17} = 0.6471$$

$$\begin{aligned}\text{Variance} \quad \sigma_b^2 &= \frac{((0-0.6471)^2\times 8)+((1-0.6471)^2\times 7)+((2-0.6471)^2\times 2)}{17} \\ &= \frac{(0.4187\times 8)+(0.1246\times 7)+(1.8304\times 2)}{17} \\ &= 0.4637\end{aligned}$$

When the threshold is 3, the calculation of foreground

$$\text{Weight} \quad W_f = \frac{6+9+4}{36} = 0.5278$$

$$\text{Mean} \quad \mu_f = \frac{(3\times 6)+(4\times 9)+(5\times 4)}{19} = 3.8947$$

$$\begin{aligned}\text{Variance} \quad \sigma_f^2 &= \frac{((3-3.8947)^2\times 6)+((4-3.8947)^2\times 9)+((5-3.8947)^2\times 4)}{19} \\ &= \frac{(4.8033\times 6)+(0.0997\times 9)+(4.8864\times 4)}{19} \\ &= 0.5152\end{aligned}$$

Within-Class Variance

$$\sigma_W^2 = W_b\,\sigma_b^2 + W_f\,\sigma_f^2 = 0.4722 * 0.4637 + 0.5278 * 0.5152 = 0.4909$$

K-means

- Partition the data points into K clusters randomly. Find the centroids of each cluster.
- For each data point:
  - Calculate the distance from the data point to each cluster.
  - Assign the data point to the closest cluster.
- Recompute the centroid of each cluster.
- Repeat steps 2 and 3 until there is no further change in the assignment of data points (or in the centroids).

$$\sum_{i \in \text{clusters}} \left\{ \sum_{j \in \text{elements of i'th cluster}} \left\| x_j - \mu_i \right\|^2 \right\}$$

K-means

- The input dataset is shown below, we first select (3,3) and (2,2) as cluster centers.
- Please calculate the process for theK-means method

No	X	Y
1	1	1
2	2	3
3	1	2
4	3	3
5	2	2
6	3	1

1. D1 $= \{(1\,,1), (2\,,2)\}$

$$= \sqrt{(2-1)^2 + (2-1)^2}$$

$$= 1.41$$

2. D1 $= \{(2\,,3), (2\,,2)\}$

$$= \sqrt{(2-2)^2 + (2-3)^2}$$

$$= 1$$

3. D1 $= \{(1\,,2), (2\,,2)\}$

$$= \sqrt{(2-1)^2 + (2-2)^2}$$

1. D2 $= \{(1\,,1), (3\,,3)\}$

$$= \sqrt{(3-1)^2 + (3-1)^2}$$

$$= 2.82$$

2. D2 $= \{(2\,,3), (3\,,3)\}$

$$= \sqrt{(3-2)^2 + (3-3)^2}$$

$$= 1$$

3. D2 $= \{(1\,,2), (3\,,3)\}$

$$= \sqrt{(3-1)^2 + (3-2)^2}$$

K-means

- The input dataset is shown below, we first select (3,3) and (2,2) as cluster centers.
- Please calculate the process for theK-means method

No	X	Y
1	1	1
2	2	3
3	1	2
4	3	3
5	2	2
6	3	1

4. D1 $= \{(3\,,3), (2\,,2)\}$

$$= \sqrt{(2-3)^2 + (2-3)^2}$$

$$= 1.41$$

5. D1 $= \{(2\,,2), (2\,,2)\}$

$$= \sqrt{(2-2)^2 + (2-2)^2}$$

$$= 0$$

6. D1 $= \{(3\,,1), (2\,,2)\}$

$$= \sqrt{(2-3)^2 + (2-1)^2}$$

$$= 1.41$$

4. D2 $= \{(3\,,3), (3\,,3)\}$

$$= \sqrt{(3-3)^2 + (3-3)^2}$$

$$= 0$$

5. D2 $= \{(2\,,2), (3\,,3)\}$

$$= \sqrt{(3-2)^2 + (3-2)^2}$$

$$= 1.41$$

6. D2 $= \{(3\,,1), (3\,,3)\}$

$$= \sqrt{(3-3)^2 + (3-1)^2}$$

$$= 2$$

K-means

No	X	Y
1	1	1
2	2	3
3	1	2
4	3	3
5	2	2
6	3	1

C1 = {(1 , 1), (1 , 2), (2 , 2), (3 , 1)}

C2 = {(2 , 3), (3 , 3)}

$$\text{Mean} = \left(\frac{x_1 + x_2 + \ldots x_n}{n}, \frac{y_1 + y_2 + \ldots y_n}{n}\right)$$

$$C1 = \left(\frac{1+1+2+3}{4}, \frac{1+2+2+1}{4}\right)$$

New C1 = (1.75 , 1.5)

$$C2 = \left(\frac{2+3}{2}, \frac{3+3}{2}\right)$$

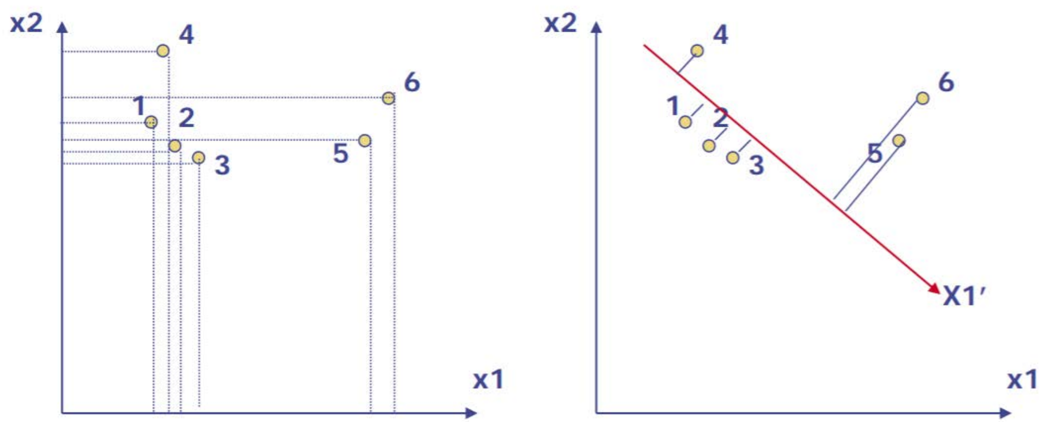
New C2 = (2.5, 3)

- Repeat until no points move

Lecture 10: face recognition

- Introduction to face recognition
- Principal Component Analysis (PCA)
  - converts a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components
  - Purpose: Identify the orientation with largest variance

$$\alpha_1 = \arg\max_{\alpha} \left(\mathrm{var}\left(\alpha^T \mathbf{X} \right) \right), \alpha \in \mathbb{R}^{p \times 1}$$



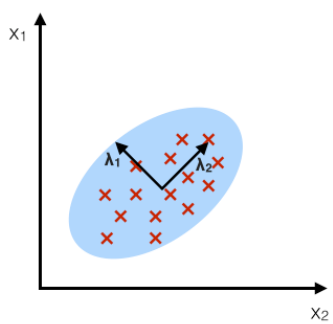
PCA projection

Linear Discriminant Analysis (LDA)

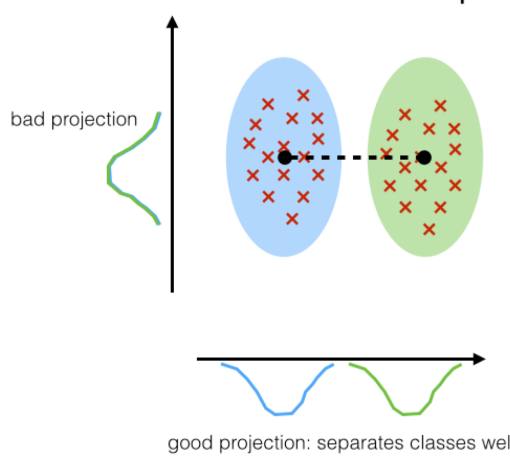
- Eigenfaces exploit the max scatter of the training images in face space
  - PCA
- Fisherfaces attempt to maximize the between class scatter, while minimizing the within class scatter.
  - Goal: find the best separation between two classes

$$J(w) = \frac{w^T S_B w}{w^T S_W w}$$

PCA:
component axes that maximize the variance



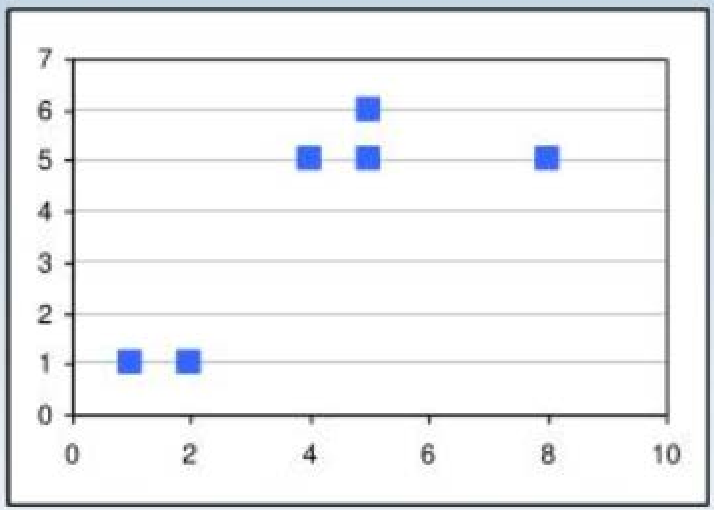
LDA:
maximizing the component axes for class-separation



PCA Example

1) Calculate Principal Component

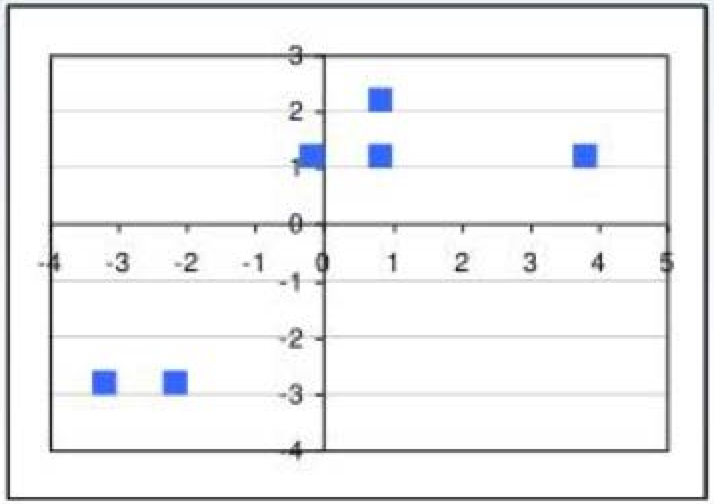
Step 1.1: Get some Data



Step 1.2: Subtract the mean

$\bar{x} = 4.17 \quad \bar{y} = 3.83$

Point	X	Y	$X - \bar{X}$	$Y - \bar{Y}$
A	1	1	-3.17	-2.83
B	2	1	-2.17	-2.83
C	4	5	-0.17	1.17
D	5	5	0.83	1.17
E	5	6	0.83	2.17
F	8	5	3.83	1.17



PCA Example

Step 1.3: Covariance matrix calculation

$$C = \begin{pmatrix} 5.139 & 3.694 \\ 3.694 & 4.139 \end{pmatrix}$$

Positive cov_{ij} values

→ x and y values increase together in dataset

Step 1.4: Eigenvectors and eigenvalues calculation –Principal axis

a) Calculate eigenvalues λ of matrix C

$$C - \lambda \cdot E = \begin{pmatrix} 5.139 - \lambda & 3.694 \\ 3.694 & 4.139 - \lambda \end{pmatrix}$$

Where E is identity matrix

The characteristic polynom is the determinant. The roots of the function, that appears if you set the polynom equals zero, are the eigenvalues

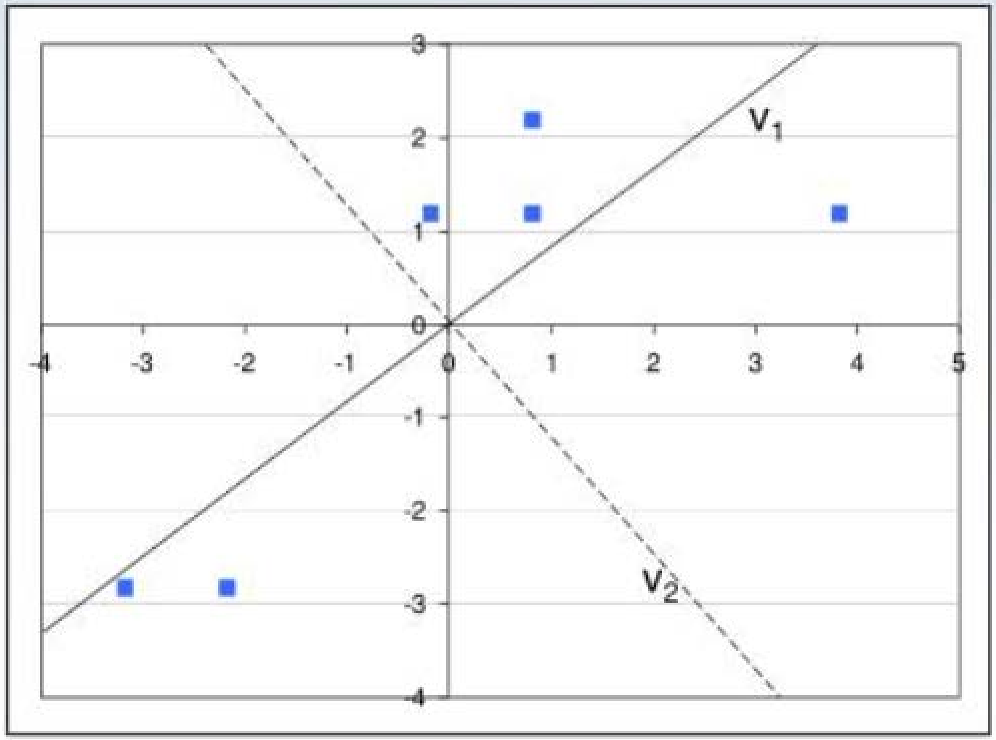
$$\det(C - \lambda \cdot E) = (5.139 - \lambda)(4.139 - \lambda) - (3.694)^2 = \lambda^2 - 9.278\lambda + 7.620 \Rightarrow \begin{array}{l} \lambda_1 = 8.367 \\ \lambda_2 = 0.911 \end{array}$$

PCA Example

b) Calculate eigenvectors v_1 and v_2 out of eigenvalues λ_1 and λ_2 via properties of eigenvectors (see matrix algebra background(3/3))

$$\begin{pmatrix} 5.139 & 3.694 \\ 3.694 & 4.139 \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = 8.367\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \qquad \begin{pmatrix} 5.139 & 3.694 \\ 3.694 & 4.139 \end{pmatrix}\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = 0.911\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

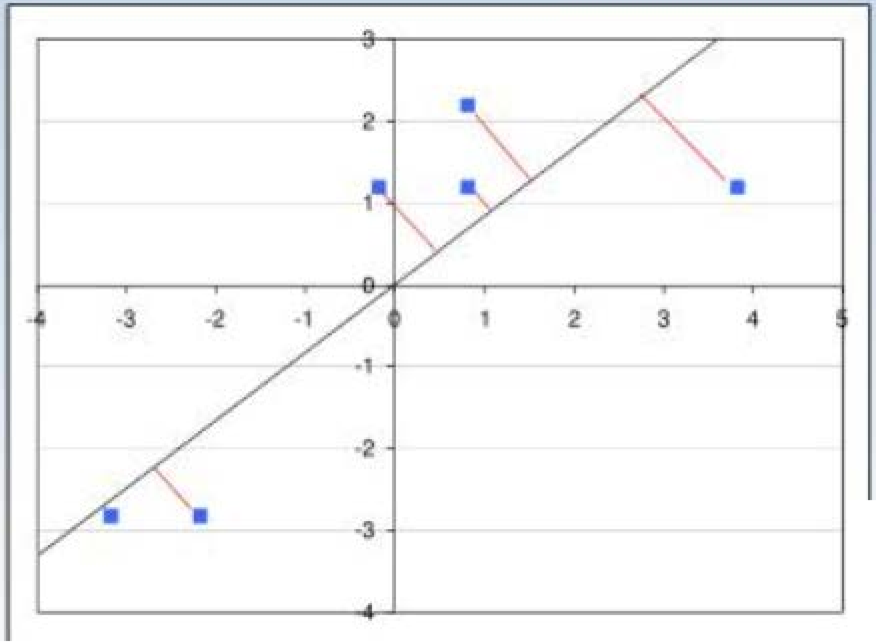
$$\Rightarrow \quad v_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} -0.753 \\ -0.658 \end{pmatrix} \qquad \Rightarrow \quad v_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.658 \\ -0.753 \end{pmatrix}$$



Eigenvector v_1 with highest eigenvalue fits the best. This is our principal component

PCA Example

2) Select the dividig point along the principal axis



Step 2.1: Calculate projections on principal axis

LDA Example

- For the given data below, please calculate the Within-class scatter matrix and Between-class scatter matrix.
- X1={(3,2),(2,3),(2,4),(5,3)}
- X2={(9,8),(8,9),(8,7),(7,8)}
- The LDA projection is then obtained as the solution of the generalized eigen value problem

$$S_W^{-1} S_B w = \lambda w$$

$$\Rightarrow \left| S_W^{-1} S_B - \lambda I \right| = 0$$

```
X1=[3,2;2,3;2,4;5,3];
X2=[9,8;8,9;8,7;7,8];
Mu1=mean(X1);
Mu2=mean(X2);
S1=cov(X1);
S2=cov(X2);
Sw=S1+S2;
Sb=(Mu1-Mu2)*(Mu1-Mu2)';
```

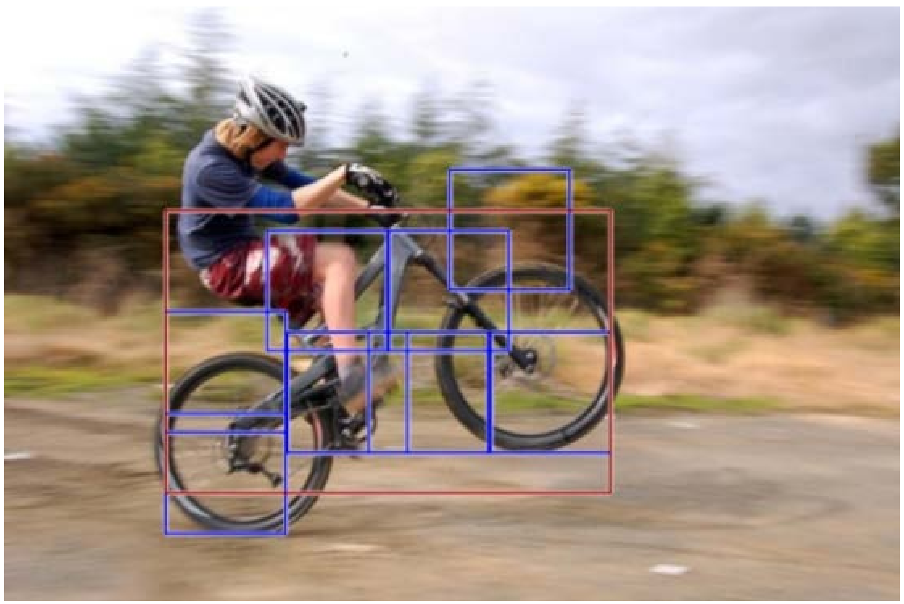
- Calculate eigenvalue and eigenvectors, where the eigenvector represent the project direction

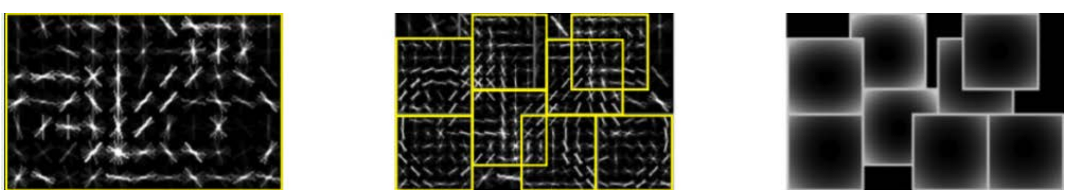
Lecture 11: face recognition

- Section a
  - Introduction of deep learning
  - DeepFace
  - FaceNet
- Section b: Deep learning in image segmentation
  - Models: Fully Convolutional Network, DeconvNet, SegNet, U-Net, PSPNet, DeepLab v1, v2, v3, v3+
  - Loss functions: **Cross entropy (CE),** Weighted cross entropy, Balanced cross entropy (BCE), Focal loss, Dice loss, Tversky loss
  - Able to choose suitable loss fuctions

Lecture 12: Detection

- HOG
- DPM





Scan image(s) at all scales and locations

↓

Extract features over windows

↓

Run window classifier at all locations

↓

Fuse multiple detections in 3-D position & scale space

↓

Object detections with bounding boxes

Lecture 12: Detection

- Section B
  - RCNN
  - Fast-RCNN
  - Faster-RCNN
  - Mask-RCNN