Advanced Image Processing

CENTRE UNIVERSITAIRE D'INFORMATIQUE
UNIVERSITE DE GENEVE

Copyright notice

This course uses in part some materials from:

- A. K. Jain, Fundamentals of Digital Image Processing, Prentice-Hall,1989.
- P. Moulin, Lecture Notes, Advanced Image Processing ECE497PM, UIUC, 1998.
- M. Vetterli, Lecture Notes: Advanced Signal Processing: Wavelets and Applications, EPFL, 1999.
- M. Unser, Lecture Notes: Image Processing, EPFL.
- B. Girod, Lecture Notes: EE368: Digital Image Processing, Stanford, 2000.

Advanced Image Processing

Introduction

S. Voloshynovskiy

Slava

http://cui.unige.ch/~aip/

Grading Policy

To decide

10 weeks

2-3 weeks

1) paper

2) Matlab/Python 3)

The final grade will be determined based on a final course project and a final exam.

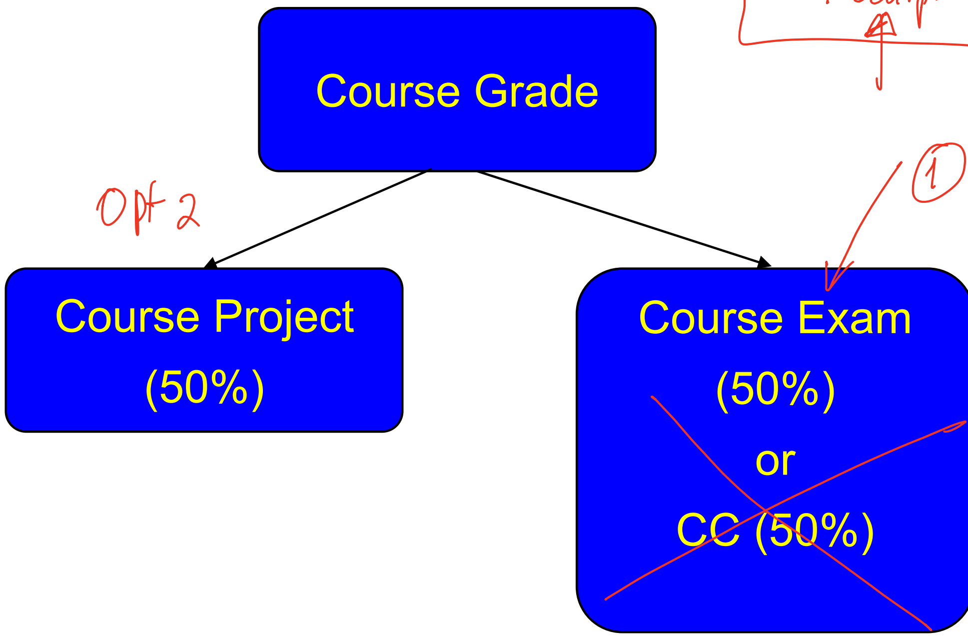
Oral —	2/3	1/2
TP/Cou.p.	1/3	1/2

opt 1

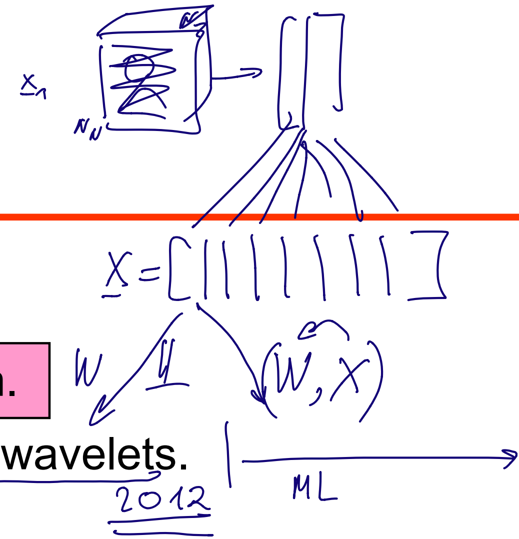
TPs

Opt 2

1



Course Outline



1. ■ Recall of Linear Algebra.
2. ■ Introduction. Human Visual System.
3. ■ Image Presentation: pyramids and wavelets.
4. ■ Random Signals.
5. ■ Image Modeling.
   ■ Image Sensor Models. Noise Models.
6. ■ Image Denoising.
   ■ Image Restoration.
   ■ Image Compression.
   ■ Video Modeling and Compression.
   ■ Digital Data Hiding. → MM Sec & Privacy

AIP

2012

no books → review papers, YouTube, invited talks

Recommended books

- A. K. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, 1989.
- A. Bovik, Handbook of Image & Video Processing, Academic Press, 2000.
- H. Stark and J. W. Woods, Probability, Random Processes, and Estimation Theory for Engineers, Prentice-Hall, 1994.
- A. Gersho and R. M. Gray, Vector Quantization and Signal Compression, Kluwer, 1992.
- M. Vetterli and J. Kovacevic, Wavelets and Subband Coding, Prentice-Hall, 1995.

1D DSP → 2D IP → ≡ Data Science

Introduction: Digital Signal Processing (DSP)

DSP Applications

- **Information Transmission/Storage/Retrieval**
  - **Source/Channel Coding**
    - Speech/image coding
    - Data compression
    - Communication over noisy channels
  - **Channel Equalization**
    - Data communication on adverse channels
- **Information extraction**
  - **Parameter estimation**
    - Spectral analysis
    - radar and sonar
    - medical imaging
    - geophysics
    - astronomy
    - multimedia
  - **Signal Classification**
    - Speech, image, character recognition signal, object detection

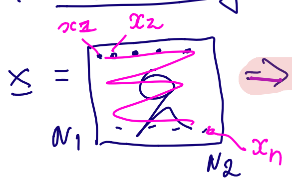
Parameter estimation → Channel Equalization; Parameter estimation → Signal Classification

Introduction: Digital Image Processing (DIP)

Image Processing covers:

1
- denoising
- restoration
- interpolation/extrapolation
- compression
- detection
- recognition
- data hiding
- data retrieval

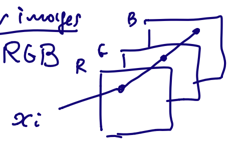
1. Denoising = noise removal



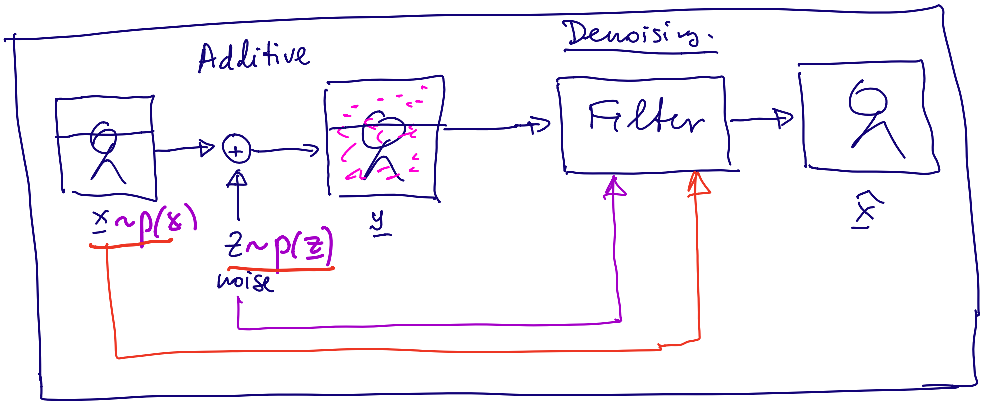
$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix}$, $n = N_1 \times N_2$

Remark: ▷ Grayscale: $x_i \in \{0, \dots, 255\}$, 8 bit

▷ Color images
• RGB



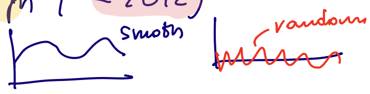
$x_i = [\underbrace{\dots}_{8}\,|\,\underbrace{\dots}_{8}\,|\,\underbrace{\dots}_{8}]$ (R, B, G), 24 bits



Denoising filter design

▸ Knowledge-based design (< 2012)

∘ deterministic



∘ probabilistic or stochastic

• $p(\underline{x})$ – model
• $p(\underline{z})$ – model
$\Rightarrow$ Filter

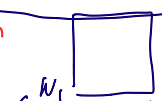
(+) • generic
∘ optimal given $p(\underline{x})$ & $p(\underline{z})$
∘ easy to derive & implement

(−) $\left.\begin{matrix} p(\underline{x}) \\ p(\underline{z}) \end{matrix}\right\}$ to be known ("Hand-crafted")

∘ difficult → impossible
• high dim. data $n \uparrow$.

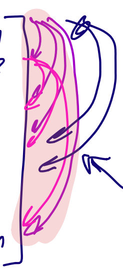
$\boxed{p(\underline{x}) = p(x_1, x_2, \dots, x_n)}$ = analytic (formula)

joint distribution



N_1 3'000, N_2 4'000

$n = N_1 \times N_2 = 3k \times 4k = 12$ Mio
$= 12'000'000$



$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

Chain rule for probability

$$\boxed{p(x_1, x_2, \dots, x_n) = p(x_1) \cdot p(x_2|x_1) \cdot p(x_3|x_2, x_1) \times \dots \times p(x_n|x_{n-1}, x_{n-2}, \dots, x_1)}$$

complexity :(

Simplifications

1) Extreme case: assume independence

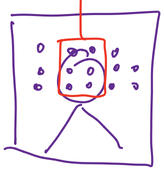
$X_i \perp\!\!\!\perp X_j$

Chain rule for probability

$$\boxed{p(x_1, x_2, \dots, x_n) = p(x_1)\cdot p(x_2|\cancel{x_1})\cdot p(x_3|\cancel{x_2, x_1}) \times \dots \times p(x_n|\cancel{x_{n-1}, x_{n-2}, \dots, x_1})}$$

$$\Rightarrow p(x_1)\cdot p(x_2) \cdots p(x_n) = \prod_{i=1}^{n} p(x_i).$$

"related"

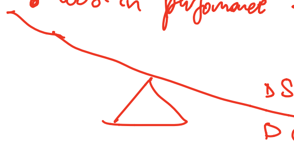


$\perp\!\!\!\perp$ why not?

$\boxed{\text{neglect}}$

$X_i \sim p(x_i) \sim$ (Gaussian with μ_i, σ_i^2)

- loose in "performance" $\neq$ model assumed $\equiv$ true model of data



- simplicity
- complexity

Objectives behind the transformation A

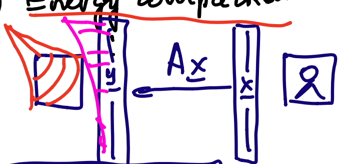
① $y_i \perp\!\!\!\perp / \approx y_j$ - "almost"

→ "uncorrelated" $\boxed{\text{wiki}}$

$\boxed{f(x_1, x_2) = \text{wiki}}$ ρ

bivariate Gaussian

② Energy compaction



$\boxed{\text{"like PCA"}}$

③ $Y_i \sim \mathcal{N}(\mu_i, \sigma_{Y_i}^2)$

$\mu_i \approx 0$

$\boxed{\text{Gaussian } \Delta}$

Solution to the optimal filtering

→ in analitical form.

↓

fast (low complexity).

2) Real world $\equiv$ image processing

y_{i0}

$p(\underline{y}) = p(y_1, y_2, \dots, y_n)$

$\simeq \prod_{i=1}^{n} p(y_i)$

y_n

$y_2 = \underline{a}_2^T \underline{x}$

$y_n = \underline{a}_n^T \underline{x}$

$= \sum_{j=1}^{n} a_{nj} x_j.$

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\underline{y}} = \underbrace{\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ \\ n \end{array}\begin{bmatrix} \underline{a}_1^T \\ \hline \\ \vdots \\ \underline{a}_n^T \end{bmatrix}}_{A\underline{x}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{\underline{x}}$$

Real image ⇒ a lot of dependency between pixels

x_1, x_2, x_n

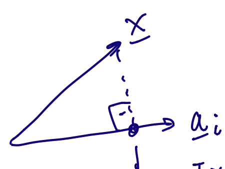
$p(\underline{x}) = p(x_1, x_2, \dots, x_n)$

$= p(x_1)\, p(x_2|x_1) \dots$

$\dots p(x_n|x_{n-1}, \dots, x_1)$

complex

Recall



$y_i = \underline{a}_i^T \underline{x} = \underline{x}^T \underline{a}_i$

complexity

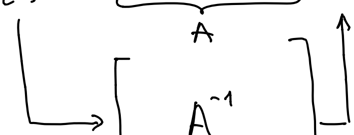
$A\underline{x}$

↓

② fast

complexity of computation

$$\begin{bmatrix} \\ \underline{y} \\ \\ \end{bmatrix} = \underbrace{\begin{bmatrix} \hline \\ \vdots \\ \hline \end{bmatrix}}_{A} \begin{bmatrix} \\ \underline{x} \\ \\ \end{bmatrix}$$



A^{-1}

invertability

Properties of transform A ①

$$\boxed{\begin{aligned} A^T &= A^{-1} \\ A^T A &= A A^T = I \end{aligned}}$$

over-determined

"A^{-1}" - if it exists.

under-determined

$A = \underset{m}{\left[\;\right]}^{n}$

$A = \underset{m}{\left[\;\right]}^{n}$

$A = \underset{m=n}{\left[\;\right]}^{n}$

"A^{-1}" $= V_k \Sigma_k^{-1} U_k^T$

$\left\{\frac{1}{\lambda_i}\right\}$

k

SVD: any $\boxed{A = U \Sigma V^T}$

EVD (PCA)

$\{\lambda_i\}$

$k = \min(m, n)$

Regularization

$\left\{\frac{1}{\lambda_i + \alpha}\right\}$

small const $\alpha \sim 10^{-3}$

Properties ⓪ ① invert. ② Comp. Cost < 2012

⓪ ≠ → assume → Not true

assume → $0/0$ performance

$X = A$: DCT, DFT, DWT → review

▶ Data-based design of A. ≡ machine learning or artif. intelligence VS

(Data-driven)

$p(x)$ ↓ ≡

D given ↓

$x \rightarrow A \rightarrow y$ $\prod p(y_i)$

from where one can take A?

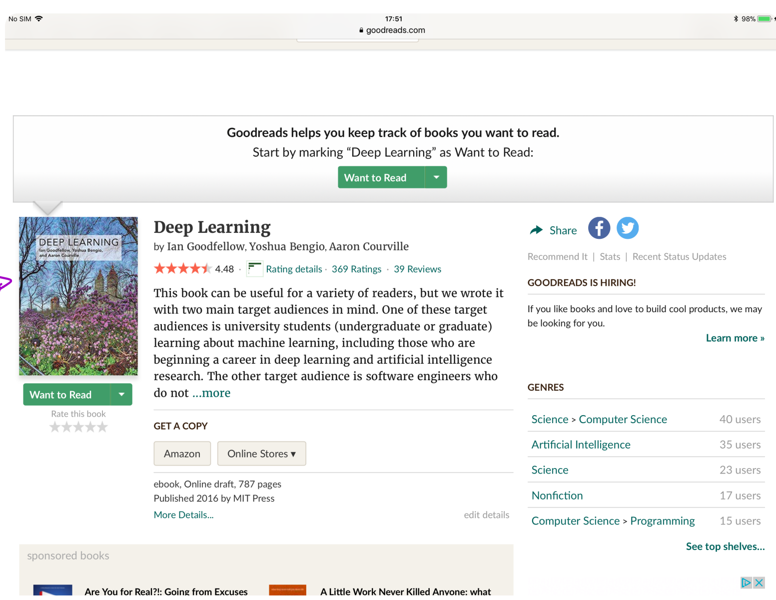
Data dependent.
- SVD
- PCA
- Deep Nets

$A = f(X)$

Training data $\underline{X}_{TR} = [\, x_1 \; \cdots \; x_M \,]$

⚠ However, if our data X have special properties, eg. AR(1), then $A_{DCT} = A_{\substack{PCA \\ SVD}}$.
$\rho \to 1$

Book → ML → Deep Learning


No SIM 17:51 98%
goodreads.com

Goodreads helps you keep track of books you want to read.
Start by marking "Deep Learning" as Want to Read:
Want to Read

Deep Learning
by Ian Goodfellow, Yoshua Bengio, Aaron Courville
4.48 · Rating details · 369 Ratings · 39 Reviews

This book can be useful for a variety of readers, but we wrote it with two main target audiences in mind. One of these target audiences is university students (undergraduate or graduate) learning about machine learning, including those who are beginning a career in deep learning and artificial intelligence research. The other target audience is software engineers who do not ...more

Want to Read
Rate this book

GET A COPY
Amazon | Online Stores ▾

ebook, Online draft, 787 pages
Published 2016 by MIT Press
More Details... edit details

Share
Recommend It | Stats | Recent Status Updates

GOODREADS IS HIRING!
If you like books and love to build cool products, we may be looking for you.
Learn more »

GENRES

Science > Computer Science	40 users
Artificial Intelligence	35 users
Science	23 users
Nonfiction	17 users
Computer Science > Programming	15 users

See top shelves...

DIP: Main Concepts

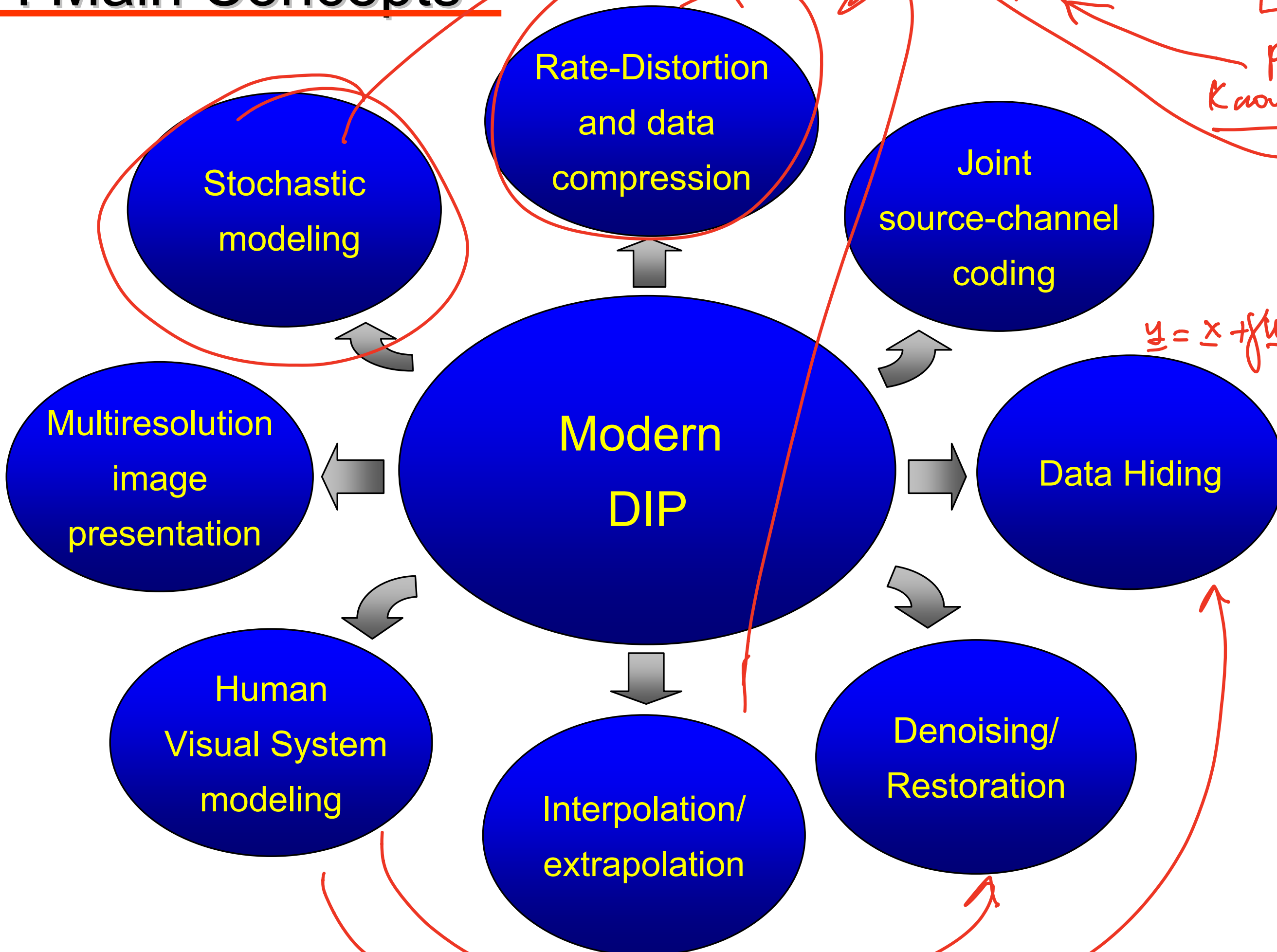
Data-driven

A

p(x)

Knowledge

$\underline{y} = \underline{x} + f(\underline{w}) \odot \underline{M}$



Goal of the course

Despite different applications, existing solutions and used particular methods, there are a few powerful, basic principles that can be used to guide the design of image processing algorithms.

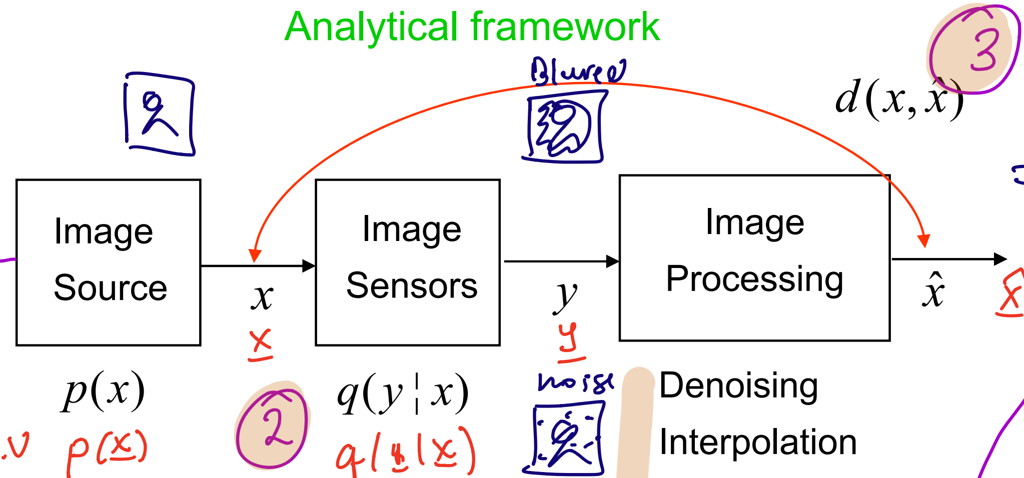
The goal of this course is to introduce these concepts and to investigate their applicability to Image Processing.

Goal of the course

In the scope of the course we will:

- study optimal processing techniques;
- learn how to recognize the implicit assumptions that lie behind suboptimal techniques;
- understand the fundamental connections between optimal techniques in apparently unrelated areas such as image restoration and interpolation/extrapolation, image denoising and lossy image compression.

Analytical Framework



- Denoising
- Interpolation
- Restoration
- Compression
- Image Analysis
- Detection
- Recognition

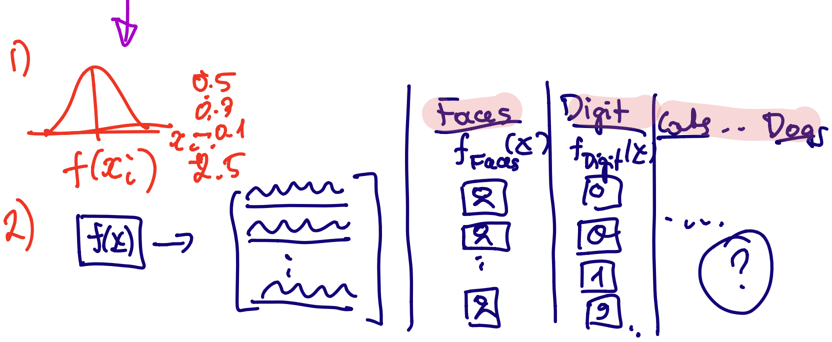
① ② ③

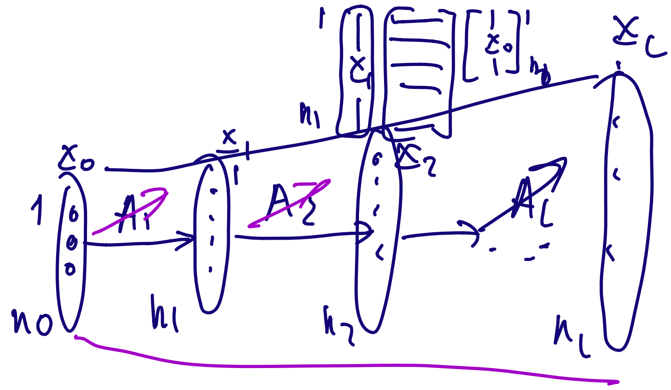
- d.R.V $p(\underline{x})$
- c.R.V $f(\underline{x})$

$q(\underline{y} \mid \underline{x})$

$d(\underline{x}, \underline{\hat{x}}) = \|\underline{x} - \underline{\hat{x}}\|_2^2 = \left(\sqrt{\sum (x_i - \hat{x}_i)^2}\right)^2$

$\ell_2 ; \ell_1 ; \ell_0$





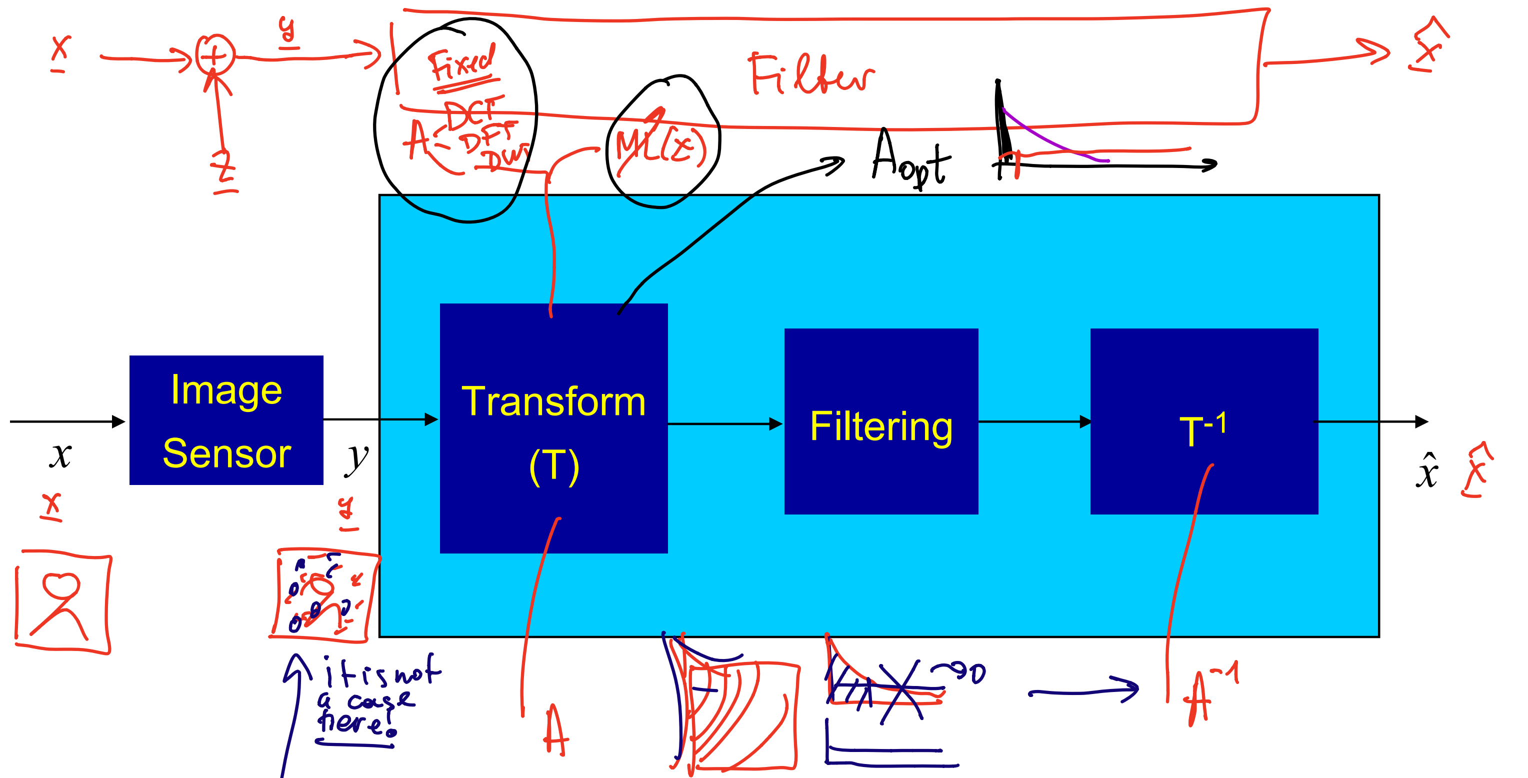
≤2012 (Aero Simoncelli) 2014 < GAN
VAE

https://arxiv.org/pdf/1606.05908.pdf

http://www.cns.nyu.edu/~lcv/texture/ (Texture synthesis)

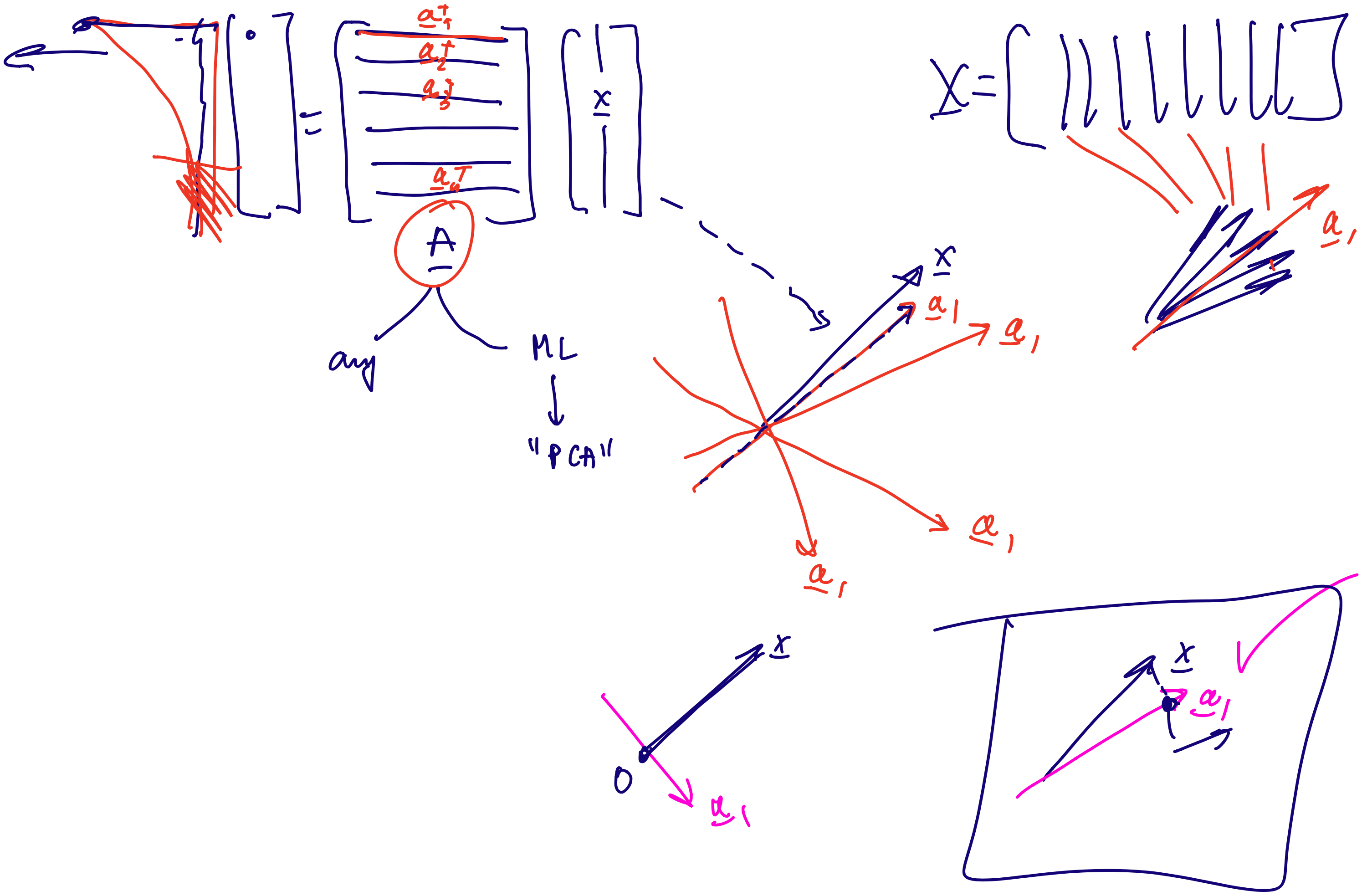
A. Dosovitsky (CVPR'15).
Boston

Restoration, Denoising and Enhancement

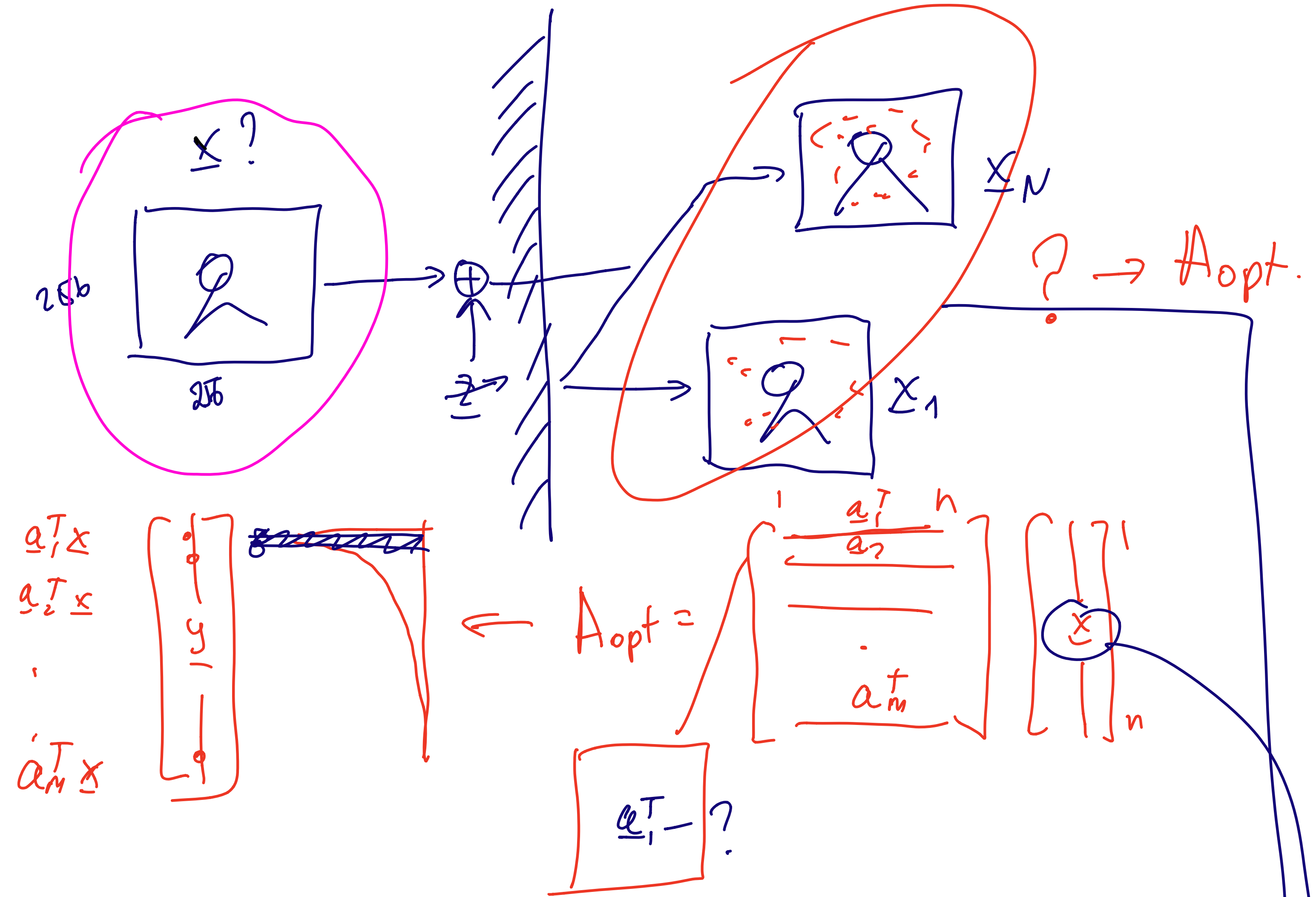


▷Separability (in terms of frequencies) of image & noise.

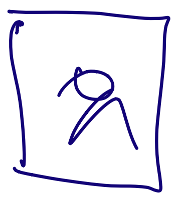
$A \xleftarrow{ML} X$ (A & X are related), then

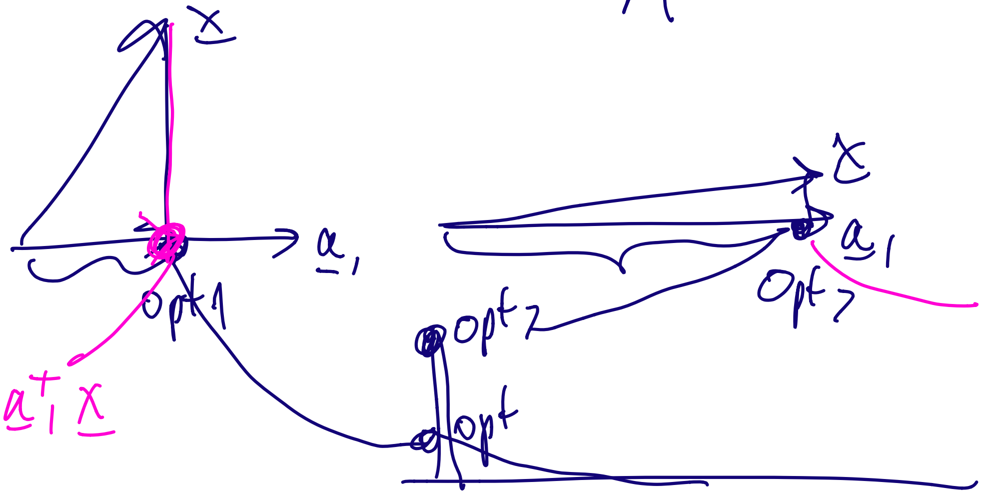


a_1^T
a_2^T
a_3^T
a_4^T
x
A
any
ML
"PCA"
$X =$
x
a_1
O



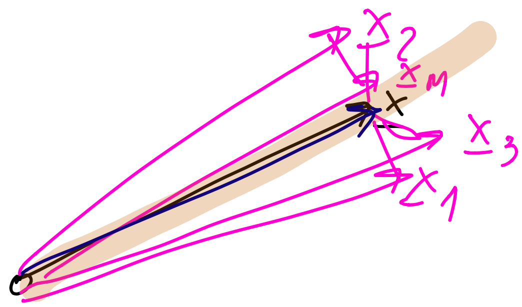
$$\underline{X} = (x_1, \ldots, x_M)$$

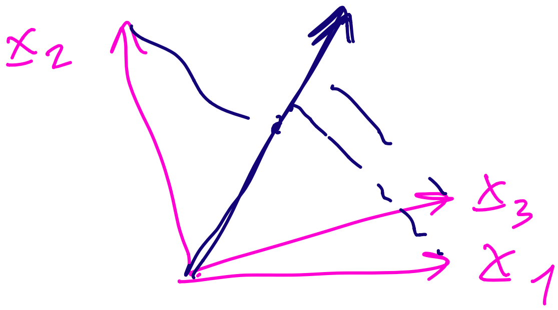
Opt 1) $\underline{a}_1^T \approx \frac{\underline{x}_1 + \underline{x}_2 + \cdots x_M}{M} \xrightarrow[M \to \infty]{}$  $\hat{x}$



$$\underline{a}_1^T \underline{x} = \sum_{i=1}^{n} a_{1i} x_i$$

opt 2) PCA $\to \underline{a}_1^T$





ML, if my training data is very non-uniform,
$\underline{a}_1^T \approx$ quasi-optimal

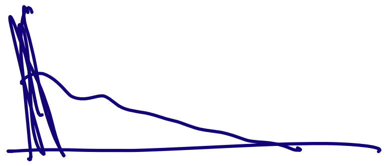


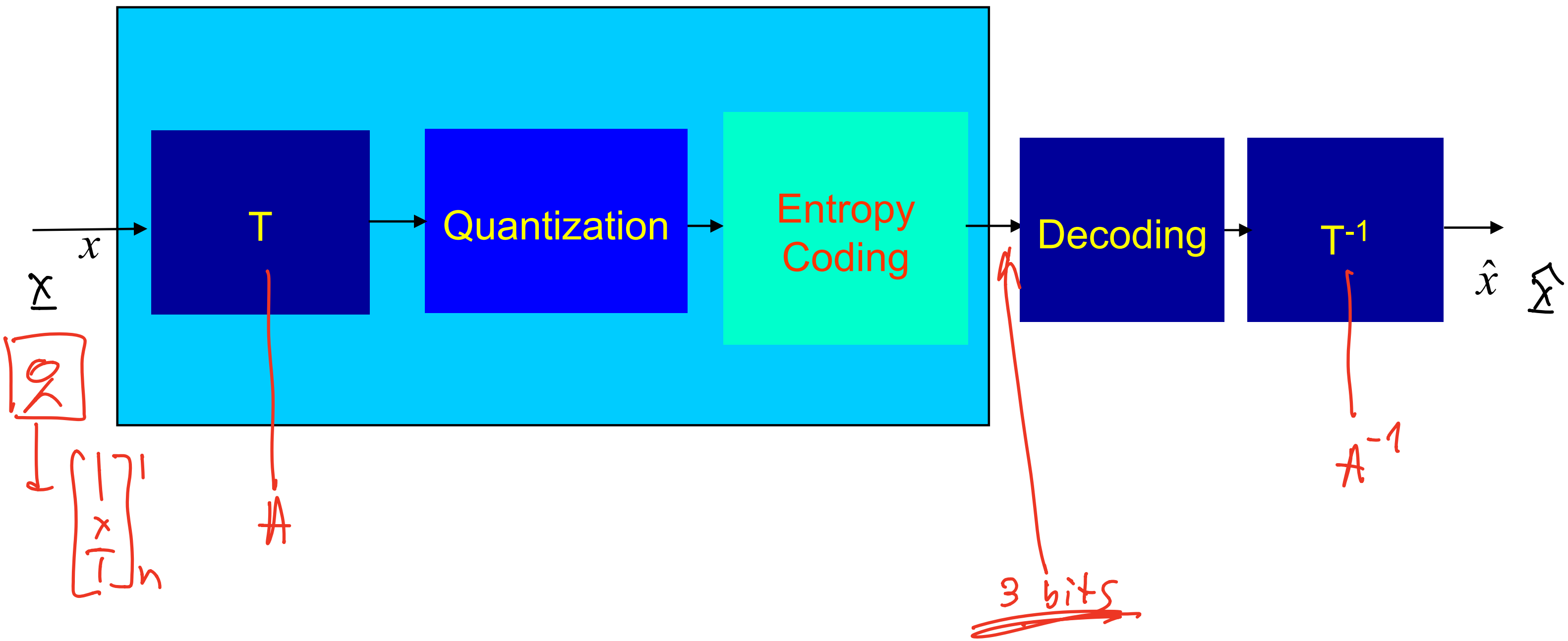
Image Restoration/Denoising

There is some uncertainty about the image x of interest, given the measurements y . This uncertainty is typically due to thermal noise or photon noise, or blurring introduced by the sensors.

This relationship can be modeled by a conditional probability distribution $q(y \mid x)$.

Image restoration/denoising: Compute $\hat{x}$ that minimizes the expected distortion $Ed\left(x,\hat{x}\right)$.

Image Compression



Vector quantization — ML-approach.

$\log_2 8 = 3$ bits (2^3)

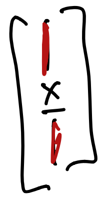
$X_{TR} = \left[\ |\ |\ |\ \cdots\ |\ |\ \right]$ ($n \times M$)

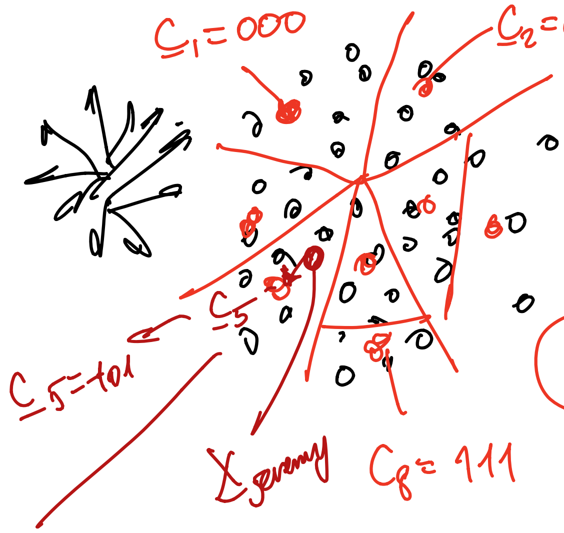
Codebook $c_1\ c_2\ \cdots\ c_8$

$\mathcal{C} = \left[\ |\ |\ |\ |\ |\ |\ |\ |\ \right]$

$x =$

$\rightarrow$





8 centroids

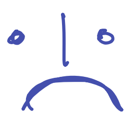
$2 \times 2 \times 2$

$2^3 \rightarrow$ 3 bits.

K-means

$d(x, c_5) = \|x - c_5\|_2^2$

(−)



$\rightarrow$ ① Complexity

② Memory

③ Training $\rightarrow X_{TR} \leftarrow x$

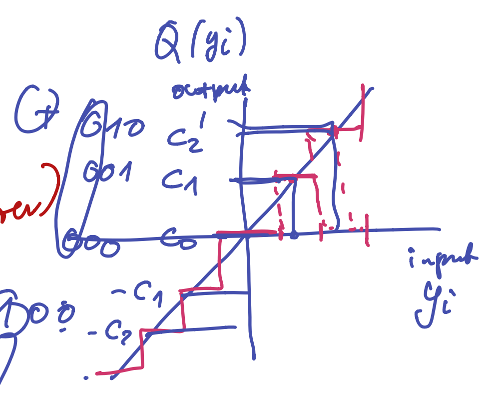
② Solution

$$\left[\ y\ \right] = \left[\ A\ \right] \left[\ x\ \right]$$

A

001|010|...

Scalar quantiser (vs Vector quantiser)

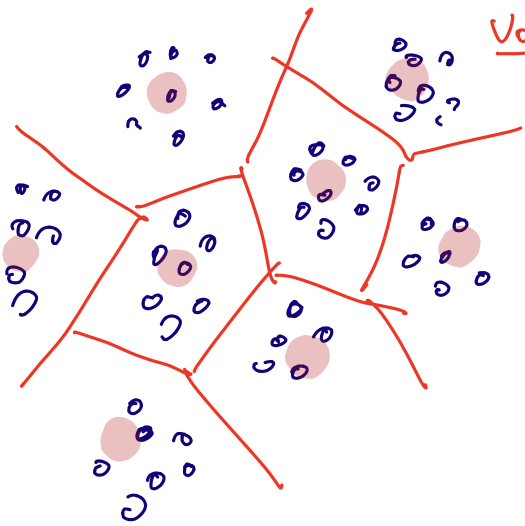


JPEG — DCT $= A$

— DWT $= A$

JPEG2000

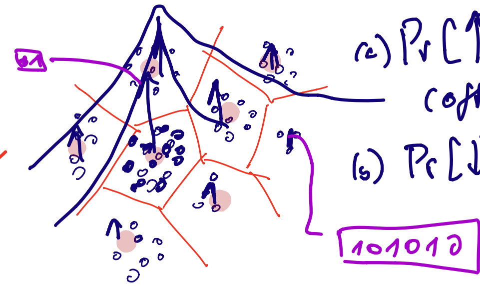
⚠ Centroid label encoding



Voronoi

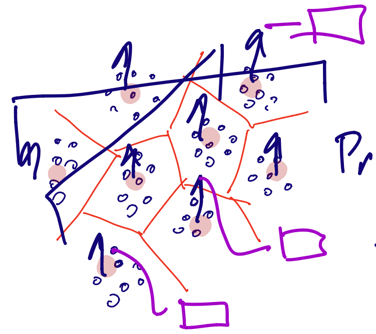
[Matlab python]

Labels



(a) $\Pr[\uparrow]$ = # bits small (often)

(b) $\Pr[\downarrow]$ = # bits big.



$\Pr[\text{of } c_i \text{ activation}] = \frac{1}{K}$.

K-means

K = 8 clusters

Image Compression

As well, there is a great variability among the possible images x .

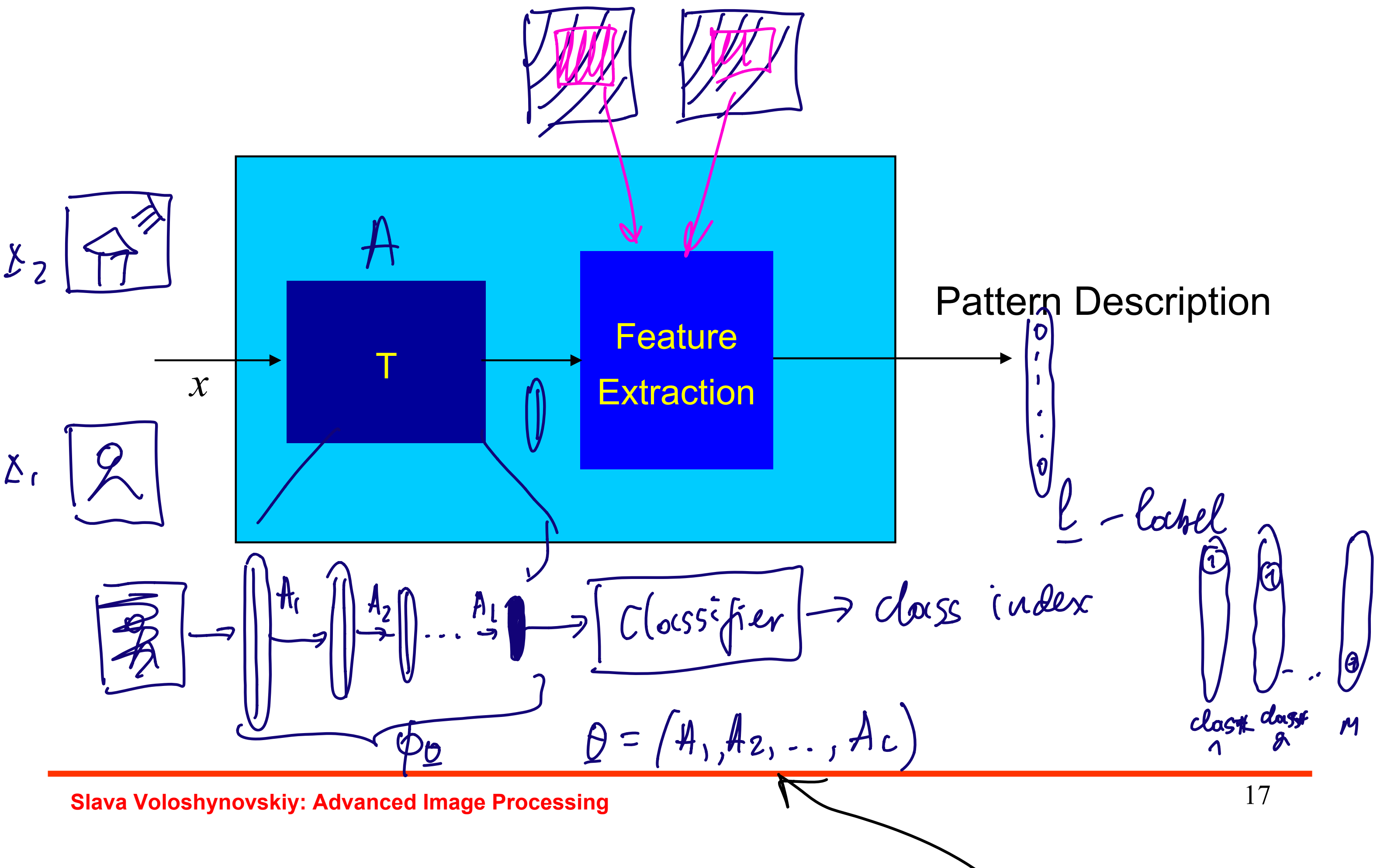
Consider the compression problem, assuming perfect sensor data for simplicity $(y = x)$.

Then $E\ell(\hat{x})$ represents an average over all possible input images x.

To evaluate $E\ell(\hat{x})$, we need a model with a prior distribution $p(x)$.

Image compression: Compute $\hat{x}$ that minimizes the expected distortion $Ed(x, \hat{x})$ subject to rate $E\ell(\hat{x}) \leq R$ constraint, where $\ell(\hat{x})$ is the length of the code word used to encode $\hat{x}$.

Image Analysis



F. Extraction

Classical Methods

$X \parallel A$

↓

manually

/hand-crafted features/

$\nabla_\theta(\cdot) = 0$

autograd

- SGD
- ADAM

2014

ML-based methods

$A = f(X)$

automatic feature extraction

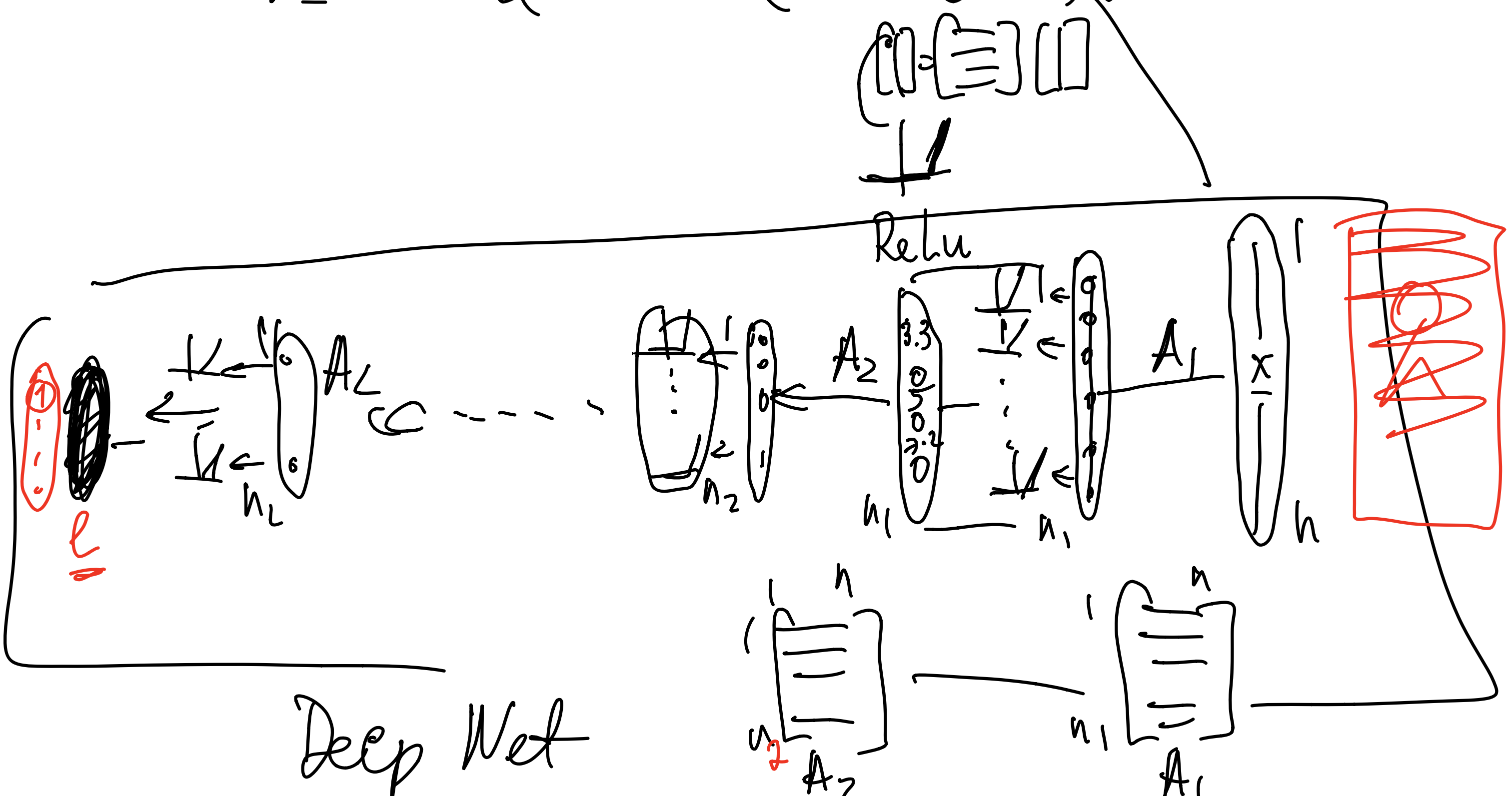
▷ Training data $\{x_i, l_i\}$

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N} L(l_i, \phi_\theta(x_i))$$

cost

$$(\hat{A}_1, \hat{A}_2, \dots, \hat{A}_L)$$

$$\phi_\theta(x) = G_L(A_L \dots G_2(A_2 G_1(A_1 x)))$$



Deep Net

Stochastic DIP

①   ②   ③

Once $p(x)$, $q(y \mid x)$ and $d(x,\hat{x})$ are specified, optimal solutions can in principle be obtained as the solution to an optimization problem.

The primary difficulty consists in identifying realistic models for $p(x)$ and $q(y \mid x)$ and $d(x,\hat{x})$ in making the optimization problem tractable (this introduces practical constrains on the possible models $p(x)$, $q(y \mid x)$ and $d(x,\hat{x})$).

Stochastic DIP

<u>Practical motivation</u>:

- Modeling of $p(x)$ is a difficult theoretical/practical problem.

- The physics of the sensors determine $q(y \mid x)$.

- The choice of a suitable $d(x, \hat{x})$ requires some knowledge of the Human Visual System.

Conclusions

A solid training in random processes is needed to formulate, understand and analyze statistical models.

Some familiarity with Image Processing is necessary to develop the necessary intuition about the design of statistical models.