

# **INTRODUCTION TO AUDIOVISUAL PROCESSING**

## **Exercises v4.1 (Spring 2019-2020)**

**Signal Processing Group (Ferran Marqués, Francesc Rey)**

This document includes a set of problems covering the INTRODUCTION TO AUDIOVISUAL PROCESSING course of UPC. It combines four kinds of exercises:

- **Preliminary exercises**, which are examples of the background knowledge that students are supposed to have from previous courses,
- **Basic exercises**, which are mainly to get used to the basic concepts of each topic and to get some practice on them,
- **Exam exercises**, which are based on those exercises that have been proposed in previous exams and further develop the basic concepts, and
- **Further research exercises**, which are advanced exercises proposing open questions to be studied as a research topic.

The goal of this document is to help students assimilate the course contents. It also represents what may be expected from them at the mid-term control and the final exam.

*This document may include mistakes, typos, errors or inaccuracies.*

*Please report them, preferably by email, to [ferran.marques@upc.edu](mailto:ferran.marques@upc.edu) or [francesc.rey@upc.edu](mailto:francesc.rey@upc.edu)*

### **Table of Contents:**

Topic 1: Statistical signal modelling.....	2
Topic 2: Estimation Theory .....	11
Topic 3: Optimal and Adaptive Filtering .....	21
Topic 4: Signal Transforms.....	30

## Topic 1: Statistical signal modelling

### Preliminary Exercises

---

- 1.1. Demonstrate that  $\sigma_X^2 = \text{var}(X) = E\{[X - E\{X\}]^2\} = E\{X^2\} - E^2\{X\}$ .
- 1.2. Given the mean and variance of a random variable  $X$ , compute the mean and variance of  $Y = kX$ .
- 1.3. Given the pdf of a random variable  $X$  ( $f_X(x)$ ), compute the pdf of the sum of  $X$  and a constant value  $a$ :  $Z = X + a$ .  
Given the mean and variance of  $X$ , compute the mean and variance of  $Z$ .
- 1.4. Compute the mean and the variance of the sum of a set of independent random variables:  $Z = \sum_{i=1}^N X_i$
- 1.5. Compute the mean and the variance of the weighted sum of a set of independent random variables:  $Z = \sum_{i=1}^N w_i X_i$
- 1.6. Compute the mean and variance of the difference between two independent random variables:  $Z = X - Y$
- 1.7. Given two random variables  $(X, Y)$ , demonstrate that  $c_{X,Y} = r_{X,Y} - m_X m_Y$

### Basic Exercises

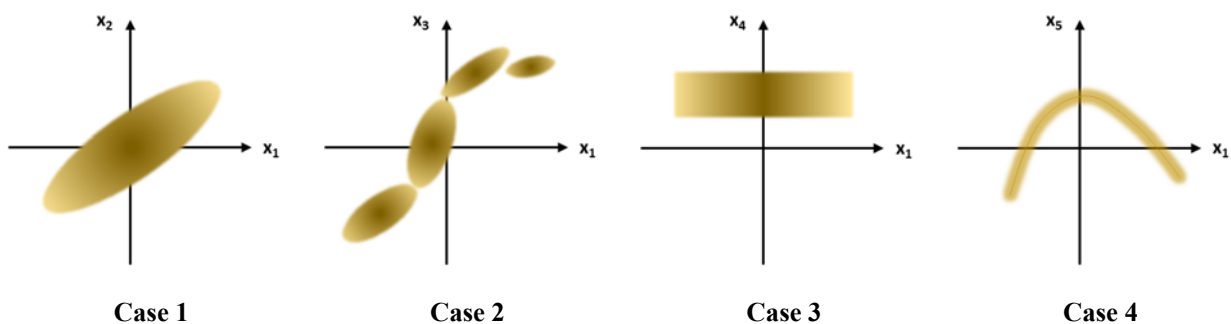
---

1.8. As first task at your new job in the multimedia company you are working for, you have been asked for reviewing an old piece of code that is a part of the company's basic coder. The soft is not correctly documented so it is complicated to understand the algorithm and the reasons for the various choices and steps in it. The first thing that surprises you is the fact that there are four different variables ( $x_2$ ,  $x_3$ ,  $x_4$ , and  $x_5$ ) that are linearly predicted based on an initial one ( $x_1$ ).

You suspect that this may come from successive patches that have been included in the code, so you want to analyze if these predictions are sound.

AT THIS POINT, YOU MAY PROPOSE YOUR OWN ANALYSIS STRATEGY

To do so, you collect sets of samples of all these variables at a large number of time instants. The following plots show the joint probability density function (pdf) of each of these predicted variables ( $x_2$ ,  $x_3$ ,  $x_4$ , and  $x_5$ ) and  $x_1$ .



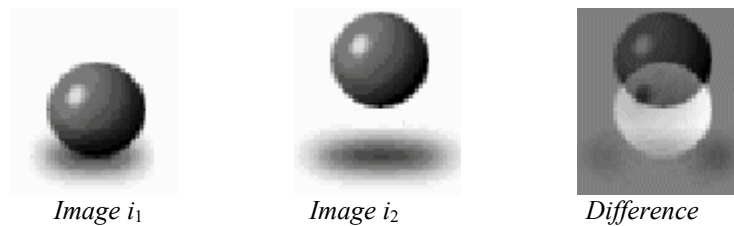
- a) **Discuss** in which of these four cases you would apply linear prediction to exploit the correlation between the variables. If there is any case on which you would not apply it, discuss which technique you would use.

- b) In order to quantify the usefulness of linear prediction in these cases, as well as in future cases, you decide to implement a routine that computes a measure of the linear dependency between two variables. Discuss the main features of the covariance, the correlation and the correlation coefficient and propose one measure for that task.
- c) Given the previous study, which are the following steps you will perform? Which will the content of your report to your team leader be?

**1.9.** The dynamic range of the difference between two images  $i_1$  and  $i_2$   $[-1,1]$  is higher than the dynamic ranges of the original images  $[0,1]$ .

a) Define by an equation the mapping  $s=f(r)$  that allows us to visualize the difference image with the same dynamic range as the original images such that positive (negative) values are displayed as light (dark) gray.

b) Which difference,  $s=f(i_2-i_1)$  or  $s=f(i_1-i_2)$ , is displayed as the third image of the following figure? (Justify your response)



**1.10.** Assuming that your input ( $r$ ) and output ( $s$ ) are discrete variables, discuss the implementation of a Range Transform using a Look-Up-Table. Write a pseudo-code of the operation.

**1.11.** Compute the Mean Square Error achieved when quantifying a variable  $X$ , that presents a uniform probability density function ( $\text{pdf}_X(x) = K, x \in [0, L\Delta]$ ), using a uniform quantizer with  $L$  levels. Assume that the representative for each quantization interval is the central value of the interval.

**1.12.** Show that a signal that is quantized with a uniform, scalar quantizer increases its quality (SNR) in 6 dBs with every additional bit used in the quantizer. Assumptions:

- The signal  $x[n]$  is always within the interval  $[-A_x, A_x]$ ,
- A *mid-rise* uniform quantizer of  $B$  bits is used,
- The signal  $x[n]$  is uniformly distributed within the quantization step  $D$ , and
- The signal power is approximated by  $s_x^2 = kA_x^2$  where  $k$  is a constant value that depends on the kind of signal.

**Hint:**  $\log_{10}[2] \approx 0,3$

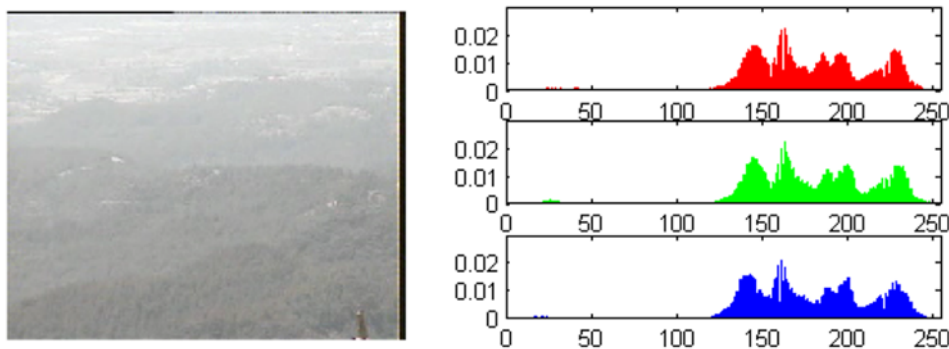
**1.13.** To transform the grey level values of an image whose range is  $[0, 255]$ , the following range transform  $s = r^2$  is proposed with normalized values ( $0 \leq r \leq 1$ ). Therefore, the levels of the input image must be divided by 255 before applying the transform. After transforming, the image is de-normalized to have an output image with natural values (including zero) in the range  $[0, 255]$ .

- a) Which are the zones of grey level expansion and compression?
- b) Can the transform be inverted?
- c) If this transform is used as first step before applying a 4-level mid-rise uniform quantizer, which are the intervals of the input image values that will be quantized to this four output levels?
- d) How could you build a non-uniform quantizer using this transform?

**1.14.** The following 4x4 image, initially quantized with 64 levels, is to be quantized with only 2 levels. Obtain the image after optimal quantization (with two levels) in the sense of minimum mean quadratic error. Which is the quadratic error achieved? Compare the result with the quadratic error that is obtained using a 2-level uniform quantization. Note that the resulting levels have to be integer values.

10	10	20	20
10	10	20	20
10	10	10	60
60	10	60	60

**1.15.** In forest fire monitoring, sometimes it is necessary to apply a range transform to palliate the effect in the image of the weather conditions. In the following figure we present a typical (blurred) image and the normalized histogram of its RGB components.

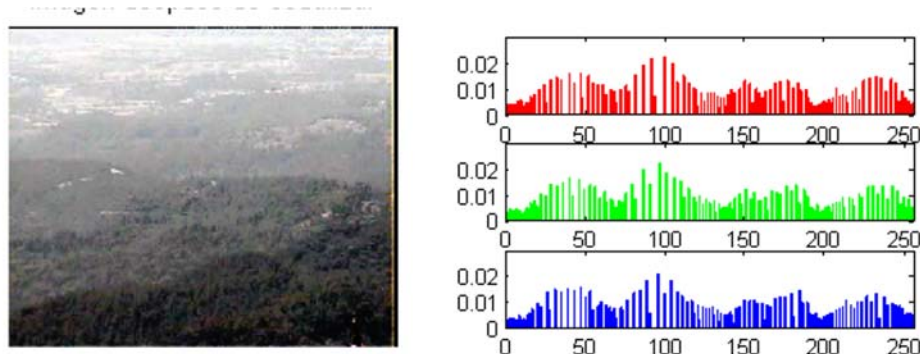


- Propose a range transform (no equalization) to improve the visualization of the image.
- Explain the necessary steps to equalize the image.

For the R component, the histogram has values equal to zero (0) for  $r \leq 124$ . In addition, the following 5 values of the normalized histogram are given:

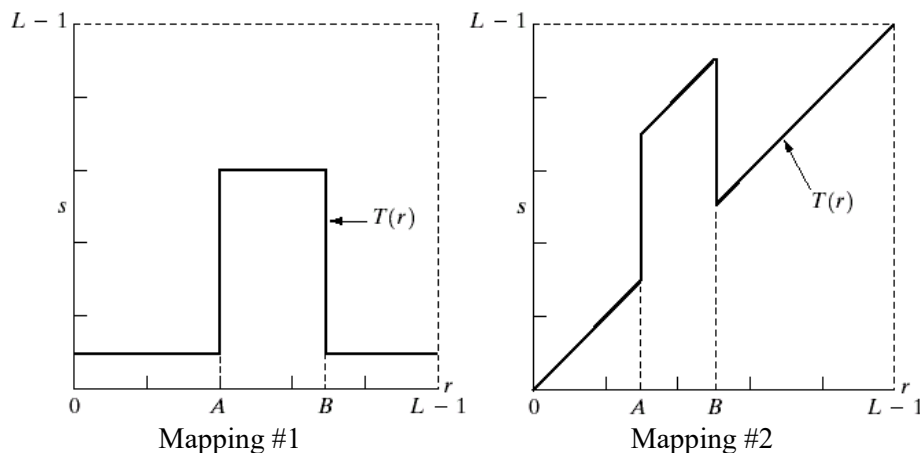
$p(125)$	$p(126)$	$p(127)$	$p(128)$	$p(129)$
0,001	0,001	0,0015	0	0,002

- Comment to which levels these five input levels will be mapped after equalization
- Discuss which transform (question (a) or (b)) has been applied to obtain the results presented in the following figure.



## Exam Exercises

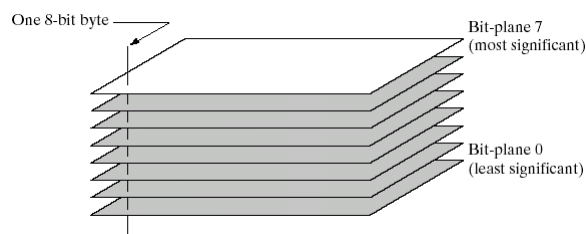
**1.16.** Gray level “slicing” is a special family of mapping transformations which consist in highlighting a specific range of gray levels in an image. Most are variations of the two basic approaches shown below:



Please answer the following questions as briefly as possible. Short (but complete) answers will be better marked:

1. What is the highlighted range of gray level values in Mapping #1 and Mapping #2?
2. Could you explain the basic difference between the images resulting from Mapping #1 and Mapping #2?
3. Could you think of a simple application (e.g. in X-ray or satellite imagery) where such mappings could be useful?

Instead of highlighting gray-level ranges, we might be interested in highlighting the contribution to the total appearance by specific bits. Suppose that each pixel in an image is represented by 8 bits. Imagine that the image is composed of eight 1-bit planes, ranging from bit-plane 0 for the least significant bit to bit-plane 7 for the most significant bit. In terms of 8-bit bytes, plane 0 contains all the lowest order bits in the bytes of the image pixels and plane 7 contains all the highest order bits, as shown in the figure below:

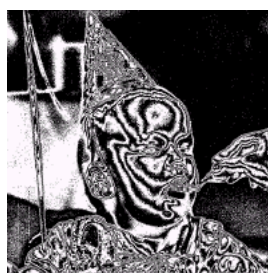


Separating a digital image into its bit-planes (bit-plane slicing) is useful for analyzing the relative importance played by each bit of the image.

4. What kind of image information is contained in the higher order bits (especially the top four)? What is left for the lower bit-planes?
5. The images below correspond to bit-planes 3 and 7 of the original image on the left. Could you tell which image corresponds to which bit-plane and why?



Original image



bit-plane X



bit-plane Y

6. Propose a set of gray-level slicing transformations  $s = T_i(r)$ ,  $0 \leq i < 8$  capable of producing all the individual bit-planes ( $i$ ) of an 8-bit monochrome image.
7. If the four lower bit-planes are set to zero and the image is reconstructed back only using the information remaining in the four higher bit-planes, what will the reconstructed image look like? What is this operation equivalent to?

**1.17.** The pixel values  $r$  of a gray level discrete image  $A$  are in range  $100 \leq r \leq 163$ . The image histogram  $h_A[r]$  is constant and can be expressed as:

$$h_A[r] = \begin{cases} 1 & 100 \leq r \leq 163 \\ 0 & \text{other } r \end{cases} \quad (1)$$

The gray level image  $B$  is the same size as image  $A$ , and its histogram  $h_B[r]$  has a unique non-zero value:

$$h_B[r] = \begin{cases} K & r = 64 \\ 0 & \text{other } r \end{cases} \quad (2)$$

- a) What is the size  $N \times N$  of the images, assuming they are square? What is the value of  $K$ ?
- b) Specify the histogram of the difference image  $A-B$ ,  $h_{A-B}[r]$ .
- c) Can the difference image be obtained through a mapping applied on  $A$ ? If yes, define precisely such mapping  $s=f(r)$ .
- d) Define the representation levels of the 2-bit quantizer optimal, in Mean Squared Error (MSE), for image  $A$ .
- e) Justify that the following images  $A$  and  $B$ :

$$i_A[m, n] = 8m + n + 100 \quad \forall [m, n] \quad (3)$$

$$i_B[m, n] = 64 \quad \forall [m, n] \quad (4)$$

where  $[m, n]$ ,  $0 \leq m, n < N$ , represent the image coordinates, have histograms corresponding to equations (1) and (2).

In the sequel, we will assume that the images  $A$  and  $B$  are defined by equations (3) and (4).

- g) A pseudo-color mapping is applied on image  $B$  to create image  $C$ . Compute the pixel values of image  $C$  assuming that the mapping is defined as:

$$LUT[r] = (R[r], G[r], B[r]), \text{ with } \begin{cases} R[r] = h_A[r] \\ G[r] = h_B[r] \\ B[r] = h_{A-B}[r] \end{cases} \quad \text{for } 0 \leq r \leq 255 \quad (5)$$

**Control IPAV 2019:** Se dispone de una imagen  $x[m, n]$  de tamaño  $256 \times 256$  cuantificada con un número de bits suficiente como para considerar que el nivel de gris toma valores continuos entre 0 y  $A$ . Con el objetivo de almacenar cada píxel con menos bits, se hace un “re-cuantificación” de la imagen mediante un cuantificador de 3 bits y margen dinámico  $[0, A]$ .

Se considera en primer lugar un cuantificador uniforme.

- a) Indica de forma gráfica los intervalos de cuantificación y los niveles de reconstrucción  $x_q$  a la salida del cuantificador, teniendo en cuenta que el margen dinámico del cuantificador es  $[0, A]$ .
- b) Calcula el tamaño en bits de la imagen cuantificada.
- c) Calcula la potencia del error de cuantificación en función de  $A$ . Explica las hipótesis utilizadas para el cálculo.

A continuación, se considera un cuantificador de Max Lloyd (cuantificador de mínimo error cuadrático medio) con el mismo margen dinámico  $[0, A]$ . Sus niveles de reconstrucción a la salida del cuantificador, para la imagen dada, son  $x_q = \{0.0954 \cdot A, 0.1458 \cdot A, 0.1949 \cdot A, 0.2415 \cdot A, 0.2770 \cdot A, 0.3090 \cdot A, 0.3509 \cdot A, 0.4025 \cdot A\}$ . La potencia del error de cuantificación en este caso es  $1.5A^2 \cdot 10^{-4}$

- d) A la vista de los niveles de reconstrucción escogidos, ¿qué se puede deducir de la función de densidad de probabilidad de los valores de la imagen  $x[m,n]$ ?
- e) ¿Cuál será el tamaño en bits de la imagen cuantificada?

A continuació representarem en una pantalla tant la imatge original  $x[m,n]$ , com les imatges quantificades  $x_{q,\text{uniforme}}[m,n]$  i  $x_{q,\text{Max\_Lloyd}}[m,n]$  (ver imatges de la hoja adjunta). Para la representació se utilizan 256 niveles de gris repartidos uniformemente entre 0 y 255 (el nivel 255 representa el valor A). Para  $x_{q,\text{Max\_Lloyd}}[m,n]$  el histograma normalizado,  $p[r]$ , que mide el uso de estos 256 niveles es:

$r$	0 ... 23	24	...	37	...	50	...	62	...	71	...	79	...	89	...	103	...	255
$p[r]$	0	0.155	0	0.087	0	0.147	0	0.153	0	0.142	0	0.146	0	0.085	0	0.085	0	0

- f) Calcula la transformació  $s=T(r)$  que permete ecualitzar el histograma de  $x_{q,\text{Max\_Lloyd}}[m,n]$ . Si has de aproximar el nou nivell per un nombre enter, redondea al enter més cercan.
- g) Finalment, indica de forma justificada a què imatge  $x[m,n]$ ,  $x_{q,\text{uniforme}}[m,n]$ ,  $x_{q,\text{MaxLloyd}}[m,n]$  i  $T(x_{q,\text{Max\_Lloyd}}[m,n])$ , corresponde cada una de les imatges de la hoja adjunta. ¿A què de aquestes imatges corresponen cada un dels dos histogrames representats?



Imagen 1

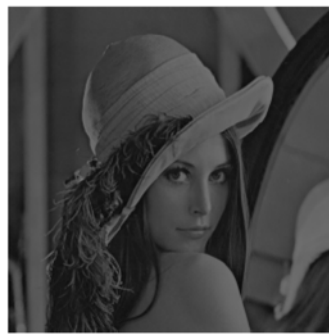


Imagen 2

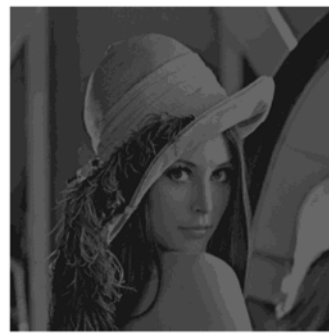
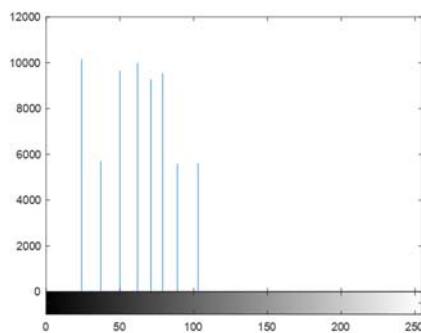


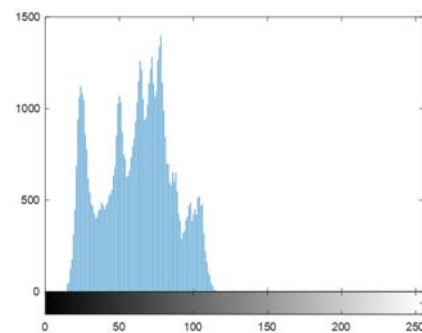
Imagen 3



Imagen 4



Histograma a



Histograma b

**Control Abril 2019:** A la teva empresa d'anàlisi de dades avui tens una reunió amb una representació de la Unitat de Dermatologia d'un dels hospitals universitaris de la teva ciutat. Per part de l'hospital, venen a la reunió la metge Cap de la Unitat de Dermatologia i el responsable TIC de la unitat.

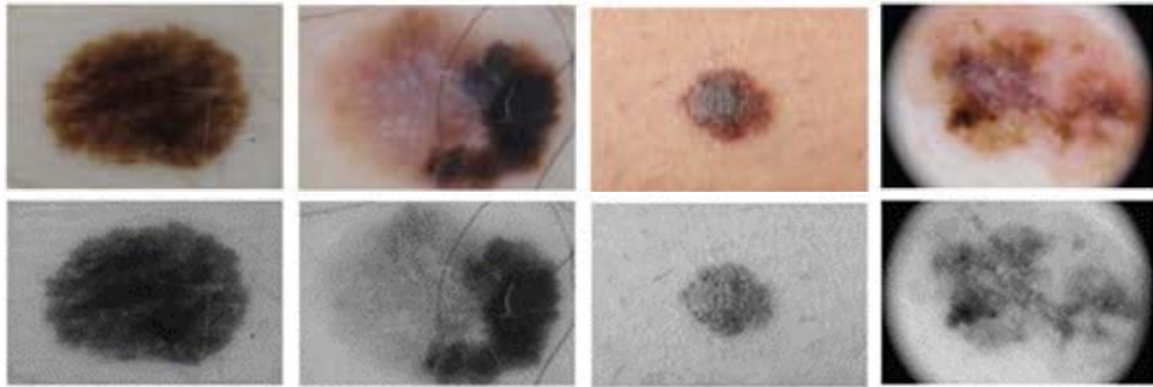
Aquesta unitat té un gran volum de dades clíniques dels milers de casos de melanoma (càncer de pell) que han tractat en els darrers anys, així com de la seva evolució. El que volen és poder crear una base de dades útil, aglutinant tota aquesta informació, que els permeti treballar amb aquestes dades i col·laborar amb la teva empresa per tal de fer models predictius de la gravetat i evolució de les lesions de pell que hagin de tractar en el futur.

Per tal de mostrar-vos el tipus de dades, us ensenyen amb una *tablet* la fitxa típica d'un malalt en la que es veu un gran volum de dades alfanumèriques i algunes imatges de les lesions, tant en color com en nivell de grisos. En aquesta



aplicació observes que, al costat de les imatges, hi ha un botó que permet activar dues possibilitats: “Stretched” i “Equalized”.

Li preguntes al responsable TIC sobre el significat d’aquest botó. Ell et comenta que no és expert en imatge, però que ha observat que moltes de les dermatoscòpies que prenen no utilitzen tot el marge dinàmic de la imatge. Així, amb aquest botó s’està fent una lleugera transformació a la imatge per a que el metge la pugui visualitzar millor: l’opció “Stretched” expandeix el marge dinàmic de manera lineal i la funció “Equalized” equalitza l’histograma. Et remarca que es tracta només d’una eina de visualització, i que les dades en memòria sempre es guarden amb els seus valors originals.



Exemples d’imatges en color i en nivells de grisos de la base de dades.

La Cap de la Unitat, veient el teu interès, et pregunta què vol dir exactament aquests canvis en les imatges.

- a) Explica-li a la metge què pretenen aquestes dues transformacions i la diferència entre les dues.

El responsable TIC et diu que, de fet, no està 100% segur de la implementació que ha fet de l’opció “Equalized”, ja que no acaba d’entendre alguns resultats que obté.

- b) Escribe amb pseudo-codi la funció “Equalized” la qual, a partir d’una imatge en nivell de grisos, et retorna una altra imatge amb el seu histograma equalitzat.

El responsable TIC, després d’analitzar breument el teu pseudo-codi, diu que sí, que això és el que ell ha implementat. Aleshores et pregunta pels resultats que obté d’aquesta funció i que no acaba d’entendre. En concret et pregunta:

- c) En imatges en nivell de grisos, i especialment quan la imatge té un fons negre (veure la quarta imatge en els exemples), el resultat de la funció “Equalized” no utilitza tampoc tot el marge dinàmic de la imatge de sortida. En concret ha vist que, sovint, els valors més propers a zero no són utilitzats. Raona si això és correcte o no. Proposa un exemple simple d’histograma que ajudi a il·lustrar el teu raonament.
- d) En imatges en color, el resultat d’aplicar la funció “Equalized” a cadascuna de les components de la imatge per separat (R, G, B) i després fusionar aquesta informació en una sola imatge dona resultats molt estranys i ha decidit no incloure aquesta possibilitat a l’aplicació. Raona per què succeeix això.

Finalment, la cap de la unitat et demana quina creus que ha de ser la manera de tractar les imatges abans de visualitzar-les per als metges: deixar-les amb el rang original, ampliar el rang mitjançant l’operador “Stretched” o ampliar-lo mitjançant l’operador “Equalized”.

- e) Raona-li a la metge quina opció creus que és la més adient per a la visualització per part dels metges. Centra el teu raonament en el cas d’imatges de nivells de grisos.



**Hints and Solutions**

---

- 1.2:**
- **Result:**  $m_Y = km_X$  and  $\sigma_Y^2 = k^2\sigma_X^2$
- 1.3:**
- **Hint1:** Model a deterministic parameter as a random variable. Which is its pdf?
  - **Hint2:** Which is the pdf of the sum of two independent random variables
  - **Result:**  $m_Z = m_X + a$  and  $\sigma_Z^2 = \sigma_X^2$
- 1.5:**
- **Hint1:** Use result in 1.2 and 1.4
  - **Result:**  $m_Z = \sum_{i=1}^N w_i m_i$  and  $\sigma_Z^2 = \sum_{i=1}^N w_i^2 \sigma_i^2$
- 1.6:**
- **Hint:** Use result in 1.5
  - **Result:**  $m_Z = m_X - m_Y$  and  $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$
- 1.8:**
- **Result(a):** Case 1: Linear prediction; Case 2: Piecewise linear prediction; Case 3: Independent variables; Case 4: Non-linear (parabolic) prediction
- 1.9:**
- **Result(a):**  $f(r) = \frac{1}{2} + \frac{r}{2}$
  - **Result(b):**  $s = f(i_2 - i_1)$
- 1.13:**
- **Hint(a):** Study the derivative of the transform
  - **Result(a):** Compression  $[0, \frac{1}{2})$  and Expansion  $(\frac{1}{2}, 1]$
  - **Result(c):**  $\left\{ \frac{1}{2\sqrt{2}}, \frac{\sqrt{3}}{2\sqrt{2}}, \frac{\sqrt{5}}{2\sqrt{2}}, \frac{\sqrt{7}}{2\sqrt{2}} \right\}$
  - **Result(d):** The inverse transform ( $f^{-1}(s)$ ) has to be applied onto the quantized values to recover the original values ( $r$ ).
- 1.14:**
- **Result(a):**  $10 \rightarrow 13; 20 \rightarrow 13; 60 \rightarrow 60 \Rightarrow \text{MSE} = 464$
  - **Result(b):**  $10 \rightarrow 16; 20 \rightarrow 16; 60 \rightarrow 48 \Rightarrow \text{MSE} = 928$
- 1.15:**
- **Result(a):** Assuming that histograms have zero values for values of ( $r$ ) below 125:  $s = \frac{255}{255-125}(r - 125)$
  - **Hint(c):** Check which probability value would have a uniform variable
  - **Result(c):**  $125 \rightarrow 0; 126 \rightarrow 0; 127 \rightarrow 0; 128 \rightarrow 0; 129 \rightarrow 1$
  - **Result(d):** Equalization
- 1.16:**
- **Result(1):**  $[A, B]$  in both cases.
  - **Result(5):**  $X = 3$  and  $Y = 7$
  - **Result(6):**  $T_7: [0, 127] \rightarrow 0$  and  $[128, 255] \rightarrow 1$ ;  $T_6: [0, 63] \cup [128, 191] \rightarrow 0$  and  $[64, 127] \cup [192, 255] \rightarrow 1$ ; ...
  - **Result(7):** Quantization with 16 representatives (4 bits)

**1.17:**

## Topic 2: Estimation Theory

### Preliminary Exercises

---

#### Basic Exercises

---

**2.1.** Suppose that we have  $N$  samples of a random process denoted by  $X[n] = A + W[n]$ , where  $A$  is an unknown constant and  $W[n]$  is a stationary, white noise ( $\sigma_W^2$ ). Consider the modified estimator:

$$\check{A} = a \frac{1}{N} \sum_{n=1}^N x[n]$$

for some constant  $a$ . Find  $a$  which results in the minimum MSE and discuss the result.

**2.2.** Suppose that we have  $N$  uncorrelated samples of a stationary random process denoted by  $X[n]$ . From these samples, we estimate the mean  $m_X$  and the variance  $\sigma_X^2$  of the process using the following expressions:

$$\hat{m}_X = \frac{1}{N} \sum_{n=1}^N x[n]$$

$$\hat{\sigma}_X^2 = \frac{1}{N} \sum_{n=1}^N (x[n] - \hat{m}_X)^2$$

Determine:

- a) The reasoning behind the definition of the variance estimator
- b) The bias of the variance estimator  $\hat{\sigma}_X^2$ .
- c) An unbiased estimator for the variance of the process  $\sigma_X^2$ .

**2.3.** In order to compute the area  $A$  of a rectangle, we measure the length of its sides  $a$  and  $b$ . The  $i$ -th measure of each side can be expressed as a function of the errors  $e_i$  and  $n_i$  as

$$\begin{aligned} x_i &= a + e_i, \\ y_i &= b + n_i. \end{aligned}$$

- a) Considering that the errors are independent, with zero mean and variances  $\sigma^2$ , determine the bias and the variance of the area estimator given by:  $\hat{A}_i = x_i y_i$ .

In order to reduce the variance of the estimates, we perform an average of  $N$  measures. The following two alternatives are proposed:

$$\begin{aligned} \hat{A}_1 &= \left( \frac{1}{N} \sum_{i=0}^{N-1} x_i \right) \left( \frac{1}{N} \sum_{i=0}^{N-1} y_i \right), \\ \hat{A}_2 &= \frac{1}{N} \sum_{i=0}^{N-1} x_i y_i. \end{aligned}$$

- b) Which estimator would you choose? Why?

**2.4.** We have obtained two measures  $z_i$  ( $i = 1, 2$ ) of a given magnitude  $x$  using two different procedures. Each procedure introduces an error denoted by  $v_i$ :

$$z_i = x + v_i.$$

Considering that the errors are mutually independent, zero mean Gaussian random variables with variance  $\sigma_i^2$  ( $i = 1, 2$ ), determine the ML estimator for  $x$  as a function of  $z_1$  and  $z_2$ . Generalize this estimator for the case of having available  $N$  independent measures of  $x$ . Compute its bias.

**2.5.** We have a realization of a random vector  $\underline{\mathbf{x}}$ . The  $N$  components of this vector take binary values 0 or 1 with probability  $p$  and  $1 - p$ , respectively. That is,  $\underline{\mathbf{x}} = (x_1, x_2, \dots, x_N)^T$ ,  $x_k \in \{0, 1\}$ . Answer the following questions, assuming that the pdf of the  $k$ -th element of vector  $\underline{\mathbf{x}}$  is given by:

$$f_{x_k}(x_k) = p^{x_k}(1 - p)^{1-x_k}$$

- What is the joint probability function  $f_{\underline{\mathbf{x}}}(\underline{\mathbf{x}})$  if the vector components are mutually independent?
- Determine the ML estimator of  $p$ .
- Is the estimator biased?
- Determine its variance and show that it is efficient.

**2.6.** Suppose that we have an observation of the form  $y = \sum_{n=1}^N x_n$ , where the samples  $\{x_n\}$  are real Gaussian random variables, statistically independent with zero mean and variance  $\sigma_x^2$ . Considering that  $N$  is a real continuous magnitude:

- Determine the ML estimator of  $N$ .  
**Hint:** As  $y$  is a Gaussian variable (addition of Gaussian variables), define its pdf computing its mean and variance.
- Analyze if the previous estimator is biased or unbiased.
- Obtain the variance of the previous ML estimator.
- Compute the mean square error of the previous ML estimator.
- Discuss whether the estimator is efficient or not.
- Comment how the ML estimator of  $N$  should be modified if the magnitude  $N$  is considered as a positive integer instead of a continuous magnitude.

**2.7.** We want to estimate the components of a vector parameter  $\underline{\boldsymbol{\theta}} = (\theta_1, \theta_2, \dots, \theta_p)^T$  and, to do so, we have  $N$  observations of a process  $X[n]$  that we model as  $X[n] = \underline{\mathbf{H}} \underline{\boldsymbol{\theta}} + W[n]$  and  $\underline{\mathbf{w}}: N(\underline{\mathbf{0}}, \underline{\mathbf{C}}_w)$ . Analyze if the Cramer-Rao Lower Bound provides an efficient estimator of  $\underline{\boldsymbol{\theta}}$ . Obtain the expression of the Fisher information matrix.

**Note:** Assume that  $\underline{\mathbf{H}}^T \underline{\mathbf{C}}_w^{-1} \underline{\mathbf{H}}$  is symmetric and invertible.

**2.8.** Given  $N$  samples of a process that can be modeled as  $\underline{\mathbf{x}} = \mu \underline{\mathbf{1}} + \underline{\mathbf{w}}$ , compute the MAP estimator of its mean ( $\mu$ ), knowing that it is a random variable with distribution  $N(\mu_m, \sigma_m^2)$ . Discuss the result for  $\sigma_m^2 \rightarrow 0$  and  $\sigma_m^2 \rightarrow \infty$  and for the case of white noise.

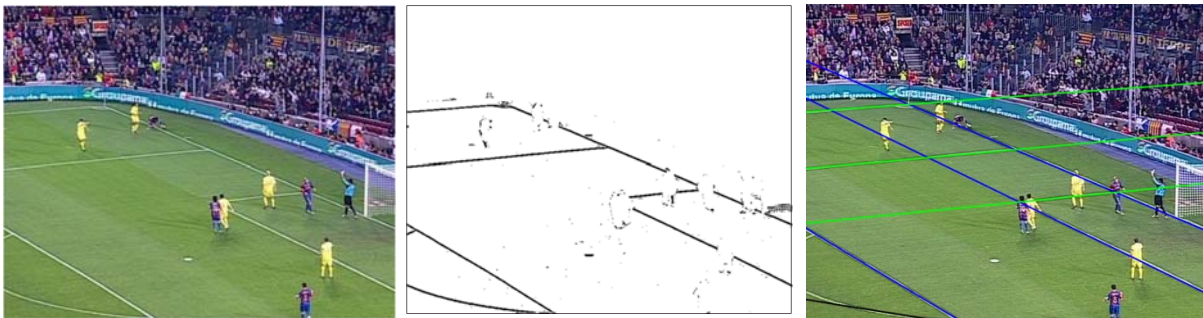
**Note:**  $W[n]$  is a **Gaussian, stationary, colored noise**.

## Exam Exercises

**Control Abril 2019:** En el primer dia a l'empresa de continguts multimèdia que has triat per fer les teves pràctiques d'estiu com a becari, tens una reunió amb l'equip amb el que treballaràs. El projecte en el que t'incorpores és d'anàlisi automàtica d'esdeveniments esportius i, més en concret, implica l'estimació de paràmetres a partir d'imatges de futbol. És per això que t'has dut a la feina els apunts de la carrera i, específicament, d'Introducció al Processat Audiovisual.

L'equip de treball està compostat per cinc persones: dos informàtics, dos enginyers de telecomunicació i un matemàtic. A la reunió ràpidament t'expliquen que per analitzar automàticament esdeveniments esportius mitjançant processat d'imatge, un primer pas habitual es determinar a l'escena els segments de línia que caracteritzen el terreny de joc. Per caracteritzar aquests segments, se sol buscar els paràmetres de les seves línies; és a dir, el pendent i l'ordenada a l'origen de la línia que inclou el segment.

Degut al soroll que introdueixen els sensors (càmeres) i a la presència d'objectes a l'escena (jugadors), la caracterització d'aquestes línies es planteja com un problema d'estimació. L'enginyera et diu que ells ja tenen un mètode per trobar el conjunt de punts que pertany a una línia determinada del terreny de joc i que aquest ha de ser el teu punt de partida. Per tal d'il·lustrar-te aquests conceptes et mostren les imatges següents:



**Imatge 1:** (a) Imatge original, (b) Imatge amb els punts candidats a pertànyer a les diferents línies del terreny de joc, (c) Resultat que s'espera de la teva feina: Línies estimades a partir d'aquests punts candidats

L'enginyer et comenta també que, com les càmeres que s'utilitzen en aquests tipus de retransmissions encara solen ser entrellaçades, les coordenades horitzontals dels punts detectats són molt més variables que les seves coordenades verticals. Per il·lustrar-te aquest problema et mostren la imatge 2.a.



**Imatge 2:** (a) Efecte de l'entrellaçat en la representació d'una línia del camp, que fa que els punts detectats presentin una component de soroll horitzontal superior a la vertical. (b) Imatge que compta la línia central del camp

La matemàtica et comenta que ells creuen que un model vàlid per caracteritzar els punts podria ser

$$x_i = \alpha y_i + \beta + \varepsilon_i$$

a on cada punt detectat ve donat per les coordenades  $(x_i, y_i)$ ,  $\alpha$  és la inversa del pendent de la recta,  $\beta$  es relaciona amb l'ordenada a l'origen i  $\varepsilon_i$  és l'error en la dimensió horitzontal, que es pot modelar com un soroll gaussià, de mitjana nul·la, independent entre mostres i de potència  $\sigma_\varepsilon^2$ .

A més, un dels informàtics et diu que per trobar els paràmetres d'una línia en concret, disposaràs de  $N$  punts, i que les seves coordenades  $x_i$  i  $y_i$  estan emmagatzemades respectivament en dos vectors  $\mathbf{x}$  i  $\mathbf{y}$ .

Aleshores et diuen que en aquest moment han de començar una altra reunió per la qual cosa t'han de deixar una hora o dues. Et mostren la teva taula i et comenten que després de la seva reunió et podran respondre els dubtes que tinguis.

Un cop sol, et dones compte de que encara no tens ordinador assignat i que, per tant, no pots ni buscar a internet els conceptes que t'han comentat i desconeixes, com ara "càmera entrellaçada". Per tant, decideixes analitzar el problema que t'han explicat des d'un punt de vista teòric i, per això, planteges el següent:

- a) Ja que t'han proposat un model gaussià pel soroll horitzontal ( $\varepsilon$ ), et proposes calcular, de manera independent per a cada paràmetre, el seu estimador no biaixat de mínima variància. Per això fas servir el fet de que la funció de densitat de probabilitat conjunta d' $N$  realitzacions d'una variable gaussiana real  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N$ , amb vector de mitjanes  $\underline{\mathbf{m}}_\varepsilon$  i matriu de covariàncies  $\underline{\mathbf{C}}_\varepsilon$  és

$$f(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N) = f(\underline{\varepsilon}) = \frac{1}{\sqrt{(2\pi)^N |\underline{\mathbf{C}}_\varepsilon|}} \exp\left(-\frac{1}{2}(\underline{\varepsilon} - \underline{\mathbf{m}}_\varepsilon)^T \underline{\mathbf{C}}_\varepsilon^{-1}(\underline{\varepsilon} - \underline{\mathbf{m}}_\varepsilon)\right)$$

i el següent resultat per a la fita de Cramer-Rao:

$$\frac{\delta}{\delta\theta} \ln(f(\underline{\mathbf{x}}; \theta)) = k(\hat{\theta} - \theta)$$

- b) Discuteixes els resultats, tant pel que fa als dos estimadors obtinguts com a les seves variàncies. En concret, interpretes la variància de l'estimador del paràmetre  $\alpha$  i analitzes, en cas de que hagi d'emmagatzemar un nombre reduït de punts per segment, quins punts triaries. Per tal d'il·lustrar el teu raonament, utilitzes com exemple una recta propera a vertical.

La discussió prèvia et fa veure que estaria bé analitzar l'estimador conjunt dels dos paràmetres. A més, el fet d'estudiar una línia en concret (per exemple, la central del camp -veure Imatge 2.b-), et fa pensar que segurament pots incloure en l'estimador informació a priori sobre el segment que estàs analitzant. Per arribar a plantejar un estimador que inclogui aquesta informació, segueixes els següents passos:

- c) Per poder proposar un estimador conjunt dels dos paràmetres, escrius en notació vectorial el model que caracteritza els punts. Així, discuteixes la validesa del següent model, comentant què representa la matriu  $\underline{\mathbf{Z}}$  i el vector  $\underline{\boldsymbol{\theta}}$ :

$$\underline{\mathbf{x}} = \underline{\mathbf{Z}} \underline{\boldsymbol{\theta}} + \underline{\boldsymbol{\varepsilon}}$$

- d) Amb aquest nou model, busques l'estimador de Màxima Versemblança (ML) dels paràmetres de la recta ( $\alpha, \beta$ ).
- e) Per tal de completar l'estudi, analitzes en quines condicions l'estimador que has obtingut és realitzable i si presenta biaix.

Ara, inclous la possibilitat de tenir un cert coneixement del pendent de la recta. Aquest coneixement, per la línia del mig del camp, el representes mitjançant el següent model gaussià del paràmetre  $\alpha$ :

$$f_\alpha(\alpha) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(-\frac{\alpha^2}{2\sigma_\alpha^2}\right)$$

- f) Fent servir aquest model, calcules ara un millor estimador dels paràmetres de la recta. Per això, expresses  $\alpha^2 = \underline{\boldsymbol{\theta}}^T \underline{\mathbf{e}}_1 \underline{\mathbf{e}}_1^T \underline{\boldsymbol{\theta}}$  a on  $\underline{\mathbf{e}}_1^T = [1, 0]$ .

Aleshores t'adones que en algunes situacions, malgrat no tinguis informació a priori d'un paràmetre concret, es pot aprofitar la informació que es té sobre com es relacionen parelles de segments de rectes en un terreny de futbol. Per exemple, es pot imposar en el procés d'optimització que dos segments determinats formen un cert angle:

g) Suposant que s'està analitzant dos conjunts de punts  $\{(x_{1i}, y_{1i}) \mid i = 1 \dots N\}$  i  $\{(x_{2j}, y_{2j}) \mid j = 1 \dots M\}$  i que podem imposar que aquestes dues rectes són ortogonals (veure imatge 2.b), proposes un sistema d'equacions que permet trobar l'estimació de Màxima Versemblança dels paràmetres de les dues rectes  $(\alpha_1, \beta_1)$  i  $(\alpha_2, \beta_2)$  sota la condició d'ortogonalitat. Com no tens temps de desenvolupar aquest idea en profunditat només planteges el sistema d'equacions però no el resols.

Quan els membres de l'equip tornen de la seva reunió, els hi presentes el que has estat desenvolupant. La resposta de l'equip és molt positiva i quan acaba la reunió i marxas cap a casa escoltes com, en broma, es barallen entre ells per fer-te de tutor durant les pràctiques.

**Final Juny 2019:** La teva empresa d'anàlisi de dades multimèdia està desenvolupant un projecte de monitorització de la vegetació del Parc Nacional del Teide mitjançant teledetecció. A la reunió amb el cap del Parc Nacional del Teide, el seu director us ha explicat que el parc presenta principalment quatre espècies vegetals:



Vista general de les espècies vegetals del Parc Nacional del Teide (Tenerife)



E1: *Spartocytisus supranubius*



E2: *Pterocephalus lasiospermus*



E3: *Descurania bourgaeana*



E4: *Pinus canariensis*

Donat el pressupost del projecte, la teva empresa es proposa utilitzar les imatges del sensor que està a bord del satèl·lit Worldview-2. Aquest sensor proporciona imatges multiespectrals de tal manera que, per a cada píxel de la imatge, es té informació de la resposta radiomètrica del terreny dins de l'espectre freqüencial dividit en vuit bandes; és a dir, cada píxel ve representat per un vector de dimensió  $N = 8$ .

Per realitzar aquest projecte, la teva empresa ha contractat a una biòloga amb formació TIC que us ha explicat que cada espècie botànica té una signatura espectral particular. Així, si es tingués un píxel de la imatge que representés una zona del terreny només contenint una espècie botànica (píxel mono-espècie), el vector representant d'aquest píxel ens donaria la signatura pròpia d'aquesta espècie. A la reunió d'avui us voleu centrar només en com obtenir aquestes signatures espectrals.

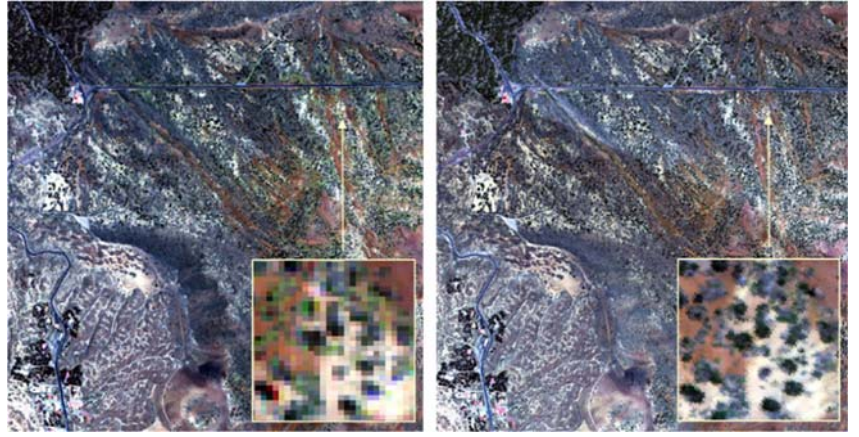
El sensor multiespectral del Worldview-2 no té una gran resolució espacial (cada píxel cobreix una àrea quadrada de 1,84 metres de costat; imatge a la pàgina següent) i, a més, se sap que presenta un cert error a les seves mesures (biaix a les bandes) que ha de calibrar-se. Per tant, per estimar les signatures espectrals amb la màxima qualitat possible, l'especialista en teledetecció de la teva empresa proposa fer servir un sensor molt més precís (el Pica-L, hiperespectral configurat a 150 bandes) que es farà volar sobre el terreny d'interès amb un dron.

A l'alçada a la que es vol fer volar el dron, aquest sensor es pot configurar a una resolució espacial de 0,1 metres (imatge a la pàgina següent). Aquesta combinació de major resolució freqüencial i espacial permet estimar la signatura espectral de l'espècie amb molta més precisió. Després, aquesta signatura hiperespectral (150 bandes molt estretes) es passa a signatura multiespectral (8 bandes amples) combinant linealment les bandes hiperespectrals.

Per temes pressupostaris, només podem fer volar el dron un cop i en una àrea petita del parc. En un sol vol del dron segurament no aconseguirem obtenir les signatures espectrals de totes les espècies (diferents ubicacions de les espècies dins del parc, diferents instants de floració). Ara bé, les signatures hiperespectrals que s'aconsegueixin serviran per, primer, obtenir unes signatures multiespectrals d'alta qualitat i, segon, per comparar-les amb les multiespectrals de la mateixa espècie que obtinguem amb el satèl·lit i així calibrar el sensor multiespectral.



Així, proposes la següent estratègia i modelat del problema. El superíndex  $A$  fa referència al conjunt de signatures espectrals que s'aconsegueixen estimar mitjançant la campanya amb el dron (sensor hiperespectral) i que es fan servir per calibrar el sensor multiespectral. El superíndex  $\bar{A}$  fa referència a la resta de signatures que han d'estimar-se mitjançant les imatges de satèl·lit (sensor multiespectral).



1. **Obtenció de les signatures hiperespectrals possibles amb un sol vol del dron:** Modelem cada signatura estimada amb el sensor hiperespectral com un vector de dimensió igual al nombre de bandes  $\hat{\mathbf{s}}_{h,k}^A$ , a on l'índex  $k$  determina a quina espècie fa referència la signatura freqüencial.
2. **Obtenció de signatures multiespectrals a partir de les hiperespectrals:** Amb la informació que donen els fabricants de les respostes freqüencials dels sensors a cada banda, s'usen aquestes signatures hiperespectrals per estimar les signatures multiespectrals possibles  $\hat{\mathbf{s}}_{m,k}^A = \mathbf{H}_k^T \hat{\mathbf{s}}_{h,k}^A$ .  
La signatura multiespectral de cada espècie del conjunt  $A$  es modela com  $\hat{\mathbf{s}}_{m,k}^A = \mathbf{s}_{m,k}^A + \boldsymbol{\varepsilon}$ . El vector  $\boldsymbol{\varepsilon}$  és un vector de mostres de soroll independents, de mitjana nul·la i amb variància molt baixa  $\sigma_{\varepsilon}^2$  i igual per a cada component (banda freqüencial).
3. **Estimació del biaix per calibrar el sensor:** S'estima el biaix dels sensors multiespectrals  $\mathbf{b}_m$  mitjançant la següent expressió, a on  $M_c$  és el nombre de mostres mono-espècie que s'han obtingut amb el sensor multiespectral de les espècies del conjunt  $A$  i que es fan servir pel calibrat.

$$\hat{\mathbf{b}}_m = \frac{1}{M_c} \sum_{i=1}^{M_c} (\mathbf{x}_{m,k}^{A,i} - \hat{\mathbf{s}}_{m,k}^A)$$

Les mostres mono-espècie del sensor multiespectral es modelen com  $\mathbf{x}_{m,k}^i = \mathbf{s}_{m,k} + \mathbf{w}^i$ . El vector  $\mathbf{w}$  és un vector de mostres de soroll independents, de mitjanes no nul·les ( $\mathbf{m}_w$ ) i amb variàncies  $\sigma_w^2$  iguals per a cada component. Les mostres de dos vectors de soroll diferents ( $\mathbf{w}^i$  i  $\mathbf{w}^j$ ) són també independents entre sí.

4. **Calibrat dels sensors i estimació de la resta de signatures multiespectrals:** L'estimació del biaix serveix per calibrar els sensors multiespectrals i obtenir una estimació de millor qualitat de les signatures espectrals del conjunt  $\bar{A}$ . La signatura multiespectral per a cada una de les espècies del conjunt  $\bar{A}$  es modela com:

$$\hat{\mathbf{s}}_{m,k}^{\bar{A}} = \frac{1}{M_e} \sum_{i=1}^{M_e} (\mathbf{x}_{m,k}^{\bar{A},i} - \hat{\mathbf{b}}_m)$$

Per comprovar la qualitat de l'estratègia proposada, fas el següent estudi:

- a) Calcules el biaix de l'estimador  $\hat{\mathbf{s}}_{m,k}^{\bar{A}}$  de les signatures multiespectrals pel conjunt  $\bar{A}$  d'espècies.
- b) Suposes que la presència del soroll  $\boldsymbol{\varepsilon}$  és menyspreable ( $\sigma_{\varepsilon}^2 \ll \sigma_w^2$ ) i calcules la matriu de covariància de l'estimador  $\hat{\mathbf{s}}_{m,k}^{\bar{A}}$  en funció de  $M_c$ ,  $M_e$  i  $\sigma_w^2$ . En el procés de simplificació del càlcul et dones compte que tota l'expressió pot quedar més senzilla si desenvolupes més el model del soroll  $\mathbf{w}$ .
- c) Per tal de reduir la variància de l'estimador i per un nombre total de mostres fix ( $M_T = M_c + M_e$ ), determines si és més interessant obtenir més mostres mono-espècie multiespectrals en el moment d'estimar el biaix per calibrar ( $M_c$  mostres de tipus  $A$ ) o en el moment d'estimar les signatures espectrals restants ( $M_e$  mostres de tipus  $\bar{A}$ ).

**Revaluació Juliol 2019:** Avui, a la teva empresa d'anàlisi de dades multimèdia tens la segona reunió per desenvolupar el projecte de monitorització de la vegetació del Parc Nacional del Teide mitjançant teledetecció. A la primera reunió veu establir l'estratègia per trobar les signatures espectrals de cadascuna de les quatre espècies vegetals dominants del terreny i del terra sense vegetació. Avui voleu discutir com heu de fer per caracteritzar els píxels de les imatges del parc obtingudes mitjançant satèl·lit:



Vista general de les espècies vegetals del Parc Nacional del Teide (Tenerife)



E1: *Spartocytisus supranubius*



E2: *Pterocephalus lasiospermus*



E3: *Descurania bourgaeana*



E4: *Pinus canariensis*

A aquesta fase del projecte s'utilitzaran només les imatges del sensor que està a bord del satèl·lit Worldview-2. Aquest sensor proporciona imatges multispectrals; és a dir, a cada píxel de la imatge s'emmagatzema la informació de la resposta radiomètrica del terreny dins del espectre freqüencial dividit en vuit bandes. Per tant, cada píxel està representat per un vector de dimensió  $N = 8$ .

El director del parc os recorda que la majoria de les agrupacions d'aquestes espècies són de mida petita i, en molts casos, apareixen espècies diferents molt properes en l'espai i, fins i tot, barrejades. Això fa que sigui difícil trobar píxels mono-espècie donada la resolució espacial del satèl·lit Worldview-2 a l'àrea d'estudi (cada píxel cobreix una àrea quadrada de 1,84 metres de costat).

Per tant, proposes caracteritzar cada píxel  $\underline{x}$  com una combinació lineal de les cinc firmes espectrals conegudes ( $\underline{s}_i$ ) més un soroll  $\underline{w}$ . Les mostres d'aquest soroll a cada banda són variables aleatòries gaussianes independents, de mitjana nul·la i de variància  $\sigma_w^2$ :

$$\underline{x} = \sum_{i=1}^5 a_i \underline{s}_i + \underline{w}$$

El paràmetre  $a_i$  s'anomena *abundància* de la signatura espectral  $\underline{s}_i$  a la mostra  $\underline{x}$ . Ara, per a cada píxel de la imatge, busques estimar quina combinació de signatures espectrals l'ha generat. Així, et proposes trobar en un primer pas l'estimador de màxima versemblança (ML):

1. Relaciones la següent expressió matricial  $\underline{x} = \underline{S} \underline{\theta} + \underline{w}$  amb les signatures espectrals i els paràmetres d'abundància. Per tenir una idea clara de què representa cada terme, dónes les dimensions de cada vector i matriu de l'expressió. Posteriorment obtens la funció de versemblança  $f(\underline{x}; \underline{\theta})$ .
2. Trobes l'estimador de màxima versemblança dels paràmetres  $\hat{\underline{\theta}}_{ML}$ .

La vostra companya biòloga us fa veure que, per a cada època de l'any i possiblement per a cada orientació del terreny, es pot tenir una informació inicial de la distribució de les abundàncies i us demana si no hi ha una manera d'incloure aquest coneixement a l'estimació.

D'aquesta manera, suposes que per a cadascuna de les situacions concretes (època de l'any i orientació del terreny) els paràmetres a estimar segueixen una distribució gaussiana i a on cada paràmetre és independent dels altres i ve caracteritzat per una mitjana i variància pròpia i coneguda ( $m_i$  i  $\sigma_i^2$ ). Així:

3. Trobes l'estimador de màxim a posterior dels paràmetres  $\hat{\underline{\theta}}_{MAP}$  per a una situació concreta (època de l'any i orientació del terreny) qualsevol.

La biòloga et diu que entén el desenvolupament matemàtic que has fet, però que potser li falta la intuïció de què s'està fent quan s'inclou aquesta informació a priori i et demana si pots donar-li un exemple que l'ajudi a entendre-ho.

Per tal de simplificar l'explicació, proposes el cas senzill en el que vols estimar un paràmetre real i positiu  $\theta$  que caracteritza estadísticament una certa població a partir d'un conjunt d'observacions  $\underline{x}$  d'aquesta població. Disposes de la funció  $f(\underline{x} | \theta)$  que et dona la probabilitat del conjunt de dades  $\underline{x}$  quan el paràmetre pren el valor  $\theta$ . (Nota: Per comoditat a l'hora de fer dibuixos, pren la funció  $f(\underline{x} | \theta)$  com aproximadament gaussiana amb variància  $\sigma^2$ ).

A més, suposes que tens informació a priori de la funció densitat de probabilitat del paràmetre a estimar  $f(\theta)$ . Per tal de discutir diferents escenaris, suposes aquestes tres situacions:

- La informació que tens del paràmetre  $f(\theta)$  és correcta i el model és com una variable aleatòria gaussiana de mitjana el valor del paràmetre i variància  $\sigma_\theta^2 < \sigma^2$ .
  - La informació que tens del paràmetre  $f(\theta)$  és incorrecta i el model és com una variable aleatòria gaussiana de mitjana menor (per exemple, la meitat) al valor del paràmetre i variància  $\sigma_\theta^2 < \sigma^2$ ,
  - La informació que tens del paràmetre  $f(\theta)$  és incorrecta i el model és com una variable aleatòria uniformement distribuïda entre zero i la meitat del valor real del paràmetre.
4. Desenvolupes l'explicació de l'estimació MAP basant-te en aquestes dades i aquests tres escenaris, recolzant-te en els dibuixos que consideris necessaris i discutint i arribant a determinar o aproximar en cada cas quin seria el valor resultant de l'estimació. (Nota: La biòloga té uns coneixements matemàtics similars als d'un estudiant del GCED que ha cursat profitosament del Q1 fins el Q3).

## Hints and Solutions

2.1:

- **Hint1:**  $\check{A} = a\hat{\theta}_N$
- **Hint2:**  $MSE(\hat{\theta}) = \sigma_{\hat{\theta}}^2 + B^2(\hat{\theta})$
- **Result:**  $a_{\text{opt}} = \frac{A^2}{A^2 + \sigma_w^2/N}$

2.2:

- **Hint1:** Model the process as  $X[n] = m_X + W[n]$
- **Hint2:** As you want to compute  $B(\hat{\sigma}_X^2) = |\sigma_X^2 - E\{\hat{\sigma}_X^2\}|$ , find first the actual value of the parameter to be estimated ( $\sigma_X^2$ ) under the previous model.
- **Hint3:**  $\hat{\sigma}_X^2 = \frac{1}{N} \sum_{n=1}^N (x[n] - \hat{m}_X)^2 = \frac{1}{N} \sum_{n=1}^N (x[n] - m_X - \hat{m}_X + m_X)^2$
- **Result1:** As the mean of the process ( $m_X$ ) is not unknown but just estimated, we have to use its estimated value ( $\hat{m}_X$ ) in the computation of the variance estimation ( $\hat{\sigma}_X^2$ )
- **Result2:**  $E\{\hat{\sigma}_X^2\} = \frac{N-1}{N} \sigma_W^2 = \frac{N-1}{N} \sigma_X^2 \Rightarrow \text{Biased}$
- **Result3:**  $\check{\sigma}_X^2 = \frac{1}{N-1} \sum_{n=1}^N (x[n] - \hat{m}_X)^2 \Rightarrow \text{Unbiased}$

2.3:

- **Result(a):**  $E\{\hat{A}_i\} = A$  and  $\text{var}(\hat{A}_i) = a^2 \sigma^2 + b^2 \sigma^2 + \sigma^4$
- **Result(b):**  $E\{\hat{A}_1\} = E\{\hat{A}_2\} = A$  and  $\text{var}(\hat{A}_1) = \frac{a^2}{N} \sigma^2 + \frac{b^2}{N} \sigma^2 + \frac{\sigma^4}{N^2}$  and  $\text{var}(\hat{A}_2) = \frac{1}{N} (a^2 \sigma^2 + b^2 \sigma^2 + \sigma^4)$
- **Result(c):**  $\text{var}(\hat{A}_1) < \text{var}(\hat{A}_2)$

2.4:

- **Result(a):**  $\hat{x}_{ML} = \frac{\frac{z_1 + z_2}{\sigma_1^2 + \sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} = \frac{\sigma_2^2 z_1 + \sigma_1^2 z_2}{\sigma_1^2 + \sigma_2^2}$
- **Result(b):**  $\hat{x}_{ML} = \frac{\sum_{i=1}^N \frac{z_i}{\sigma_i^2}}{\sum_{i=1}^N \frac{1}{\sigma_i^2}}$

2.5:

- **Result(b):**  $\hat{p}_{ML} = \frac{1}{N} \sum_k x_k$
- **Result(d):**  $\text{var}(\hat{p}_{ML}) = \frac{p(1-p)}{N}$

2.6:

- **Result(a):** Considering  $N$  a continuous magnitude  $\hat{N}_{ML} = \frac{y^2}{\sigma_x^2}$
- **Result(c):**  $\text{var}(\hat{N}_{ML}) = 2N^2$
- **Result(f):** Optimization through intensive search over  $N$

2.7:

- **Result:**  $\hat{\underline{\theta}}_{CR} = (\underline{\underline{\mathbf{H}}}^T \underline{\underline{\mathbf{C}}}_w^{-1} \underline{\underline{\mathbf{H}}})^{-1} \underline{\underline{\mathbf{H}}}^T \underline{\underline{\mathbf{C}}}_w^{-1} \underline{\underline{\mathbf{x}}}$



## Topic 3: Optimal and Adaptive Filtering

### Basic Exercises

---

**3.1.** There is a microphone situated inside the cockpit of a helicopter that captures the voice of the pilot contaminated by the noise coming from the helix. In order to cancel this noise, we have another signal coming from a different microphone situated close to the helix outside the cockpit.

- Sketch the Wiener architecture needed in order to recover the pilot voice minimizing the mean square error.
- Find the system of equations that allows computing the coefficients of the Wiener FIR filter of order  $M$ .
- Show that it would be possible to recover completely the voice signal if the noises of both microphones were sinusoidal signals (with the same frequency but different amplitudes and phase shifts). What would be the minimum required order of the filter in this case?

**3.2.** Consider the 2 channels of a stereo (real) audio signal:

- the left one:  $x[n] = x_v[n] + x_m[n]$
- the right one:  $y[n] = y_v[n] + y_m[n]$

so that both of them contain a voice component (sub-index  $v$ ) and a musical component (sub-index  $m$ ).

We would like to cancel the voice component of the right channel  $y_v[n]$  using the samples we have from the two signals,  $x[n]$  and  $y[n]$ , and assuming the following conditions:

- All components have zero mean.
- Each one of the voice components is uncorrelated with each of the two musical components.
- The musical components of the left and right channel are uncorrelated.
- The voice component of the right channel is a delayed version of the same component in the left channel:  $y_v[n] = x_v[n - D]$ .

Assuming that the signals are sampled with a sampling frequency of 8 kHz and that the delay between the voice components of both channels is 0.125 ms:

- Propose a scheme to cancel  $y_v[n]$  based on a Wiener FIR filter  $H(z)$ .
- Calculate the value of  $D$  (delay in samples) and indicate in a reasoned way what should be the minimum order  $M - 1$  of the Wiener filter.
- Derive the Wiener-Hopf equations for the design of the Wiener filter with the minimum order  $M - 1$ . Indicate how it would be possible to estimate, for the available signals  $x[n]$  and  $y[n]$ , the correlations that appear in the system of equations.

Assume now that the left channel components  $x_v[n]$  and  $x_m[n]$  have, respectively, mean powers  $\sigma_v^2$  and  $\sigma_m^2$ , and their correlation indexes are:

$$\frac{r_{x_v}[1]}{r_{x_v}[0]} = \frac{r_{x_m}[1]}{r_{x_m}[0]} = \rho$$

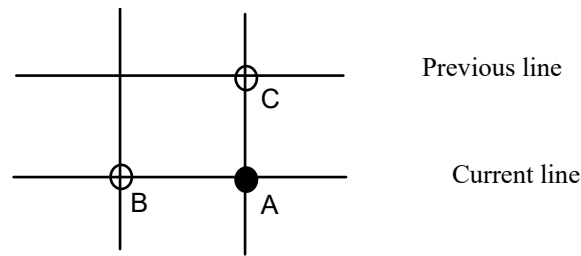
- Write the system of equations to calculate the coefficients of the Wiener filter in terms of the mean powers and correlation indexes given above.
- Solve the proposed equations in the previous point to calculate the coefficients  $h[n]$  of the filter  $H(z)$  of order  $M - 1$ .

- f) Reason the meaning of the expression of  $H(z)$  in the two following particular cases referred to the left channel:
- Case A:* Music power much lower than voice power.  
*Case B:* Music power much higher than voice power.
- g) Assuming case A, find the expression of the error power and reason its meaning.

**3.3.** Consider the real signal  $x[n] = s[n] + w[n]$ , where  $s[n]$  is an AR process that satisfies the recursive equation  $s[n] = 0.8s[n-1] + v[n]$ , where  $v[n]$  and  $w[n]$  are zero-mean white random processes, uncorrelated between them, and with powers  $\sigma_v^2 = 0.49$  and  $\sigma_w^2 = 1$ , respectively.

- Determine the autocorrelation sequences of  $s[n]$  and  $x[n]$ .
- Design a Wiener filter to be applied to  $x[n]$  with length  $M = 2$  to estimate  $s[n]$ .
- Compute the minimum mean square error if  $M = 2$ .

**3.4.** A second order linear predictor of a fixed image uses the pixels B and C to obtain a prediction of the value  $x_A$  of pixel A:



Suppose that the image is a realization of a zero-mean, spatial stationary random process where the correlation between consecutive horizontal and vertical pixels are the same with value  $\rho$  and the diagonal correlation is  $\rho^2$ . These values are normalized with respect to the power of the process  $\sigma^2$ .

- Determine the expression of the linear predictor and the prediction error.
- Find the equations that allow computing the optimal predictor coefficients, and solve them to obtain such coefficients.
- Compute the prediction gain  $G_2$  if this is defined as the quotient between the signal power and the prediction error power.
- Compute the gain  $G_1$  of a first order optimum predictor that predicts  $x_A$  from  $x_B$  (or  $x_C$ ). Show that  $G_1 < G_2$ .
- Compute the prediction gain  $G'_2$  of a second order predictor when all the coefficients are 0.5. Show that when  $\rho < 1/3$ ,  $G_1 > G'_2$ .



**3.5.** For a predictor with  $N$  coefficients based on the FIR Wiener filter, we want to study the effect produced by the misadjustment  $\mathbf{d}$  when obtaining the optimum coefficients  $\mathbf{h}_{\text{opt}}$ .  $\mathbf{d}$  can be considered as an implementation error, or the result of trying to adapt the filter for different statistics, for example.

- a) Show that the prediction error power is given by:

$$J = r_x[0] - \mathbf{h}^T \mathbf{r}_x - \mathbf{r}_x^T \mathbf{h} + \mathbf{h}^T \mathbf{R}_{xx} \mathbf{h},$$

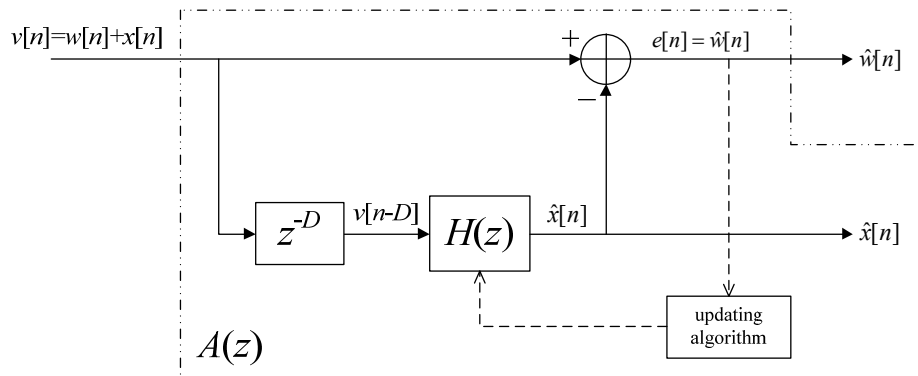
where  $x[n]$  is the input to the filter,  $\mathbf{r}_x = [r_x[-1], \dots, r_x[-N]]$ , and  $\mathbf{R}_{xx}$  is the autocorrelation matrix.

- b) If the filter coefficients finally implemented are not the optimal ones (for instance, due to the quantization of the coefficients) appears a misadjustment  $\mathbf{h} = \mathbf{h}_{\text{opt}} + \mathbf{d}$ , prove that  $J_d = J_{\min} + \mathbf{d}^H \mathbf{R}_{xx} \mathbf{d}$ .
- c) Compute  $J_d$  for the case of an order 1 predictor if the filter is not the optimal one  $h_1 = h_{\text{opt}} + d$ .
- d) For the particular case of a signal with correlation coefficient between samples  $\rho = r_x[1]/r_x[0] = 0.7$ , compute the resultant prediction gain if we choose  $h_1 = 0.825$ . Compute the experienced loss with respect to the optimum case  $h_1 = h_{\text{opt}}$ .

Consider now a predictor to which two inputs  $u[n]$  and  $v[n]$  with the same power can be applied indifferently. For a given predictor  $\mathbf{h}$ , suppose that the corresponding prediction error powers are  $J_u, J_v$ . If we consider the sum  $J_p = \beta J_u + (1 - \beta) J_v$ , where  $\beta$  ( $0 < \beta < 1$ ) is a weight parameter:

- e) Prove that  $J_p$  is minimum for a predictor  $\mathbf{h}_{\text{opt}} = \mathbf{R}^{-1} \mathbf{r}$ , if the elements of  $\mathbf{R}$  and  $\mathbf{r}$  are obtained as  $r[i] = \beta r_{uu}[i] + (1 - \beta) r_{vv}[i]$  where  $r[i]$  represents the auto-correlation function for each of the signals. This type of predictor is normally used to compute the  $\mathbf{h}_{\text{opt}}$  when the predictor must work with two different input signals. Notice that this is a clear case of misadjustment  $\mathbf{d}$  when obtaining the optimal coefficients  $\mathbf{h}_{\text{opt}}$  of the filter for each of the signals  $u[n]$  and  $v[n]$ .

**3.6.** The discrete process  $v[n]$  has two additive and uncorrelated components which must be separated: a broad-band component  $w[n]$  and a narrow-band component  $x[n]$ . As the band corresponding to  $x[n]$  is unknown and can change along time, it is necessary to implement an adaptive filter.



It is well known that the autocorrelation of a broad-band signal has a lower effective length than the one of a narrow-band signal. Using this property, the scheme of the figure above allows to separate both signals provided that the delay  $D$  is large enough for  $w[n]$  and  $w[n - D]$  to be considered as uncorrelated samples, but small enough for  $x[n]$  and  $x[n - D]$  to be correlated. In practice, we will use the smallest value of  $D$  that meets this constraint. Consider all signals to be real and zero-mean. If  $H(z)$  is a FIR filter with length  $M$ , i.e.,

$$H(z) = \sum_{k=0}^{M-1} h[k] z^{-k}.$$

- Show that the minimization of  $E[e[n]^2]$  is equivalent to the minimization of  $E[(x[n] - \hat{x}[n])^2]$ , so that the scheme of the figure is able to separate the two components  $w[n]$  and  $x[n]$ .
- Find the set of equations required to compute the coefficients of the filter  $H(z)$ .

In the following, we will assume that  $x[n]$  is a sinusoid with frequency  $\Omega_0$  and  $w[n]$  is white noise with variance  $\sigma^2$ .

- Choose the proper values for  $D$  and  $M$ , and write the coefficients of the global filter encompassed within the dashed line in the figure,  $A(z)$ , in terms of the coefficients of the filter  $H(z)$ . Note that  $A(z)$  is a prediction error filter, and give its impulse response.
- Write the expression of the LMS equation that updates the filter coefficients.
- If the power of the sinusoid is constant, reason how the speed of convergence and the misadjustment error change when the noise variance increases.

**Hint:** Show that the eigenvalues of  $v[n]$  are the eigenvalues of the sinusoid plus  $\sigma^2$ .

**3.7.** When we want to predict a non-zero mean correlated signal, i.e.,  $E[x[n]] = m$  (for example, an image) the affine predictor is a better alternative than the linear predictor, since it results in a smaller prediction error power. For the case of order 1, the equations are given by:

$$\begin{aligned}\hat{x}_p[n] &= h_p x[n-1] && \text{Linear predictor} \\ \hat{x}_a[n] &= h_a x[n-1] + b = [h_a \ b] \begin{bmatrix} x[n-1] \\ 1 \end{bmatrix} && \text{Affine predictor}\end{aligned}$$

- Which equation allows us to obtain the coefficient of the first order linear predictor?
- Obtain the equations that allow to compute  $h_a$  and  $b$ , from the minimization of  $E[e[n]^2] = E[(x[n] - \hat{x}_a[n])^2]$ .
- For the previous case, what is the expression of the minimum prediction error power?
- Suppose that we want to implement the first order affine predictor adaptively. Find the equations of the LMS from the instantaneous estimation (stochastic approximation) of the gradient.
- Which interval of  $\mu$  values assures the convergence?

**3.8.** Wiener filter architectures have some drawbacks when the data  $\mathbf{x}[n]$  is highly correlated. In such case, the autocorrelation matrix  $\mathbf{R} = E[\mathbf{x}[n]\mathbf{x}[n]^H]$  may be singular and not have inverse or may yield to very high values in the optimal vector coefficients  $\mathbf{h}[n]$  (high values could degenerate the performance of the algorithm, especially if we use finite arithmetic processors). In order to solve the problem of designing the filter coefficients  $\mathbf{h}[n]$  from the mentioned data and from a training sequence or reference  $d[n]$ , we need to introduce a penalization factor in the adaptation equation of the vector  $\mathbf{h}[n]$ . This can be expressed as the minimization of the following objective function:

$$\mathbf{h} = \arg \min_{\mathbf{h}} \{\xi(\mathbf{h}) = E[|e[n]|^2] + \alpha \|\mathbf{h}\|^2\}$$

where  $\alpha$  is a real scalar and constant along time and the error is  $e[n] = d[n] - \mathbf{h}^T \mathbf{x}[n]$ . Answer the following questions:

- Find the expression of the optimum filter that minimizes the objective function  $\xi(\mathbf{h})$  in stationary conditions.
- Discuss the role of constant  $\alpha$  in the solution found in previous section. Under what conditions would such constant adopt a positive, negative, or zero value?
- Obtain the equation of the adaptive filter based on the LMS (i.e., on the stochastic or instantaneous gradient of the objective function  $\xi(\mathbf{h})$ ).
- Prove that the adaptive filter found in the previous section converges in mean to the solution obtained in the first section a), i.e., it is fulfilled that

$$\lim_{n \rightarrow \infty} E[\mathbf{h}[n]] = \mathbf{h}_{opt}.$$

- e) Find the limits of the step-size  $\mu$  in which the adaptive filter converges to the desired solution.
- f) Determine an upper-bound of the convergence time for the previous adaptive algorithm.

Now suppose that  $\mathbf{h}[n]$  has only two coefficients  $\mathbf{h}[n] = [h_0[n] \ h_1[n]]^T$ ,  $\alpha = 0.5$ , and the cross-correlation vector and autocorrelation matrix are given by

$$\mathbf{r}_{xd} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{R}_{xx} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

- g) Sketch the error surface (contour line) as a function of  $h_0[n]$  and  $h_1[n]$ . Indicate clearly the minimum point and its value, the principal axis, and the direction of the contour lines.
- h) If  $\mathbf{h}[0] = 0$ , sketch approximately the evolution of  $h_0[n]$  and  $h_1[n]$  along time, for the values of  $\mu = 0.1$ ,  $\mu = 0.25$ , and  $\mu = 1$ .

### Hints and Solutions

**Hint AR(1):** For all exercises dealing with a process  $x[n]$  that can be modeled as an **order 1 auto-regressive process AR(1)** generated by a white noise  $w[n]$  with variance  $\sigma_w^2$ , the necessary expressions are:

- $x[n] = \rho x[n-1] + w[n]$
- $r_x[l] = \rho r_x[l-1] + \sigma_w^2 \delta[l]$
- $r_x[l] = \rho^{|l|} r_x[0]$
- $r_x[0] = \frac{\sigma_w^2}{1-\rho^2}$

#### 3.1:

- **Hint1:** The signal in the helix sensor can be defined as the engine signal in the helix ( $e_s[n]$ ) plus a white noise ( $w_s[n]$ ), whereas the signal in the microphone can be defined as the voice ( $v_m[n]$ ) plus the engine signal filtered by the cockpit system ( $e_m[n]$ ).
- **Result2:**  $(\underline{\mathbf{R}}_{e_s} + \underline{\mathbf{R}}_{w_s}) \underline{\mathbf{h}}_{opt} = \underline{\mathbf{r}}_{e_s e_m}$
- **Hint3a:** The difference between the engine signal in the helix sensor ( $e_s[n] = A \cos[w_o n + \phi]$ ) and the engine signal in the microphone ( $e_m[n] = B \cos[w_o n + \phi - \varphi]$ ) can be determined with two parameters (phase and amplitude). Therefore, we will need two equations to estimate these changes in the signal that are due to the cockpit effect.
- **Hint3b:** No white noise is present (noise have only sinusoidal components).
- **Hint3c:** If  $x[n] = A \cos[w_o n + \phi]$  and  $y[n] = B \cos[w_o n + \phi - \varphi]$ ,  $r_x[l] = \frac{A^2}{2} \cos[w_o l]$  and  $r_{xy}[l] = \frac{AB}{2} \cos[w_o l - \varphi]$
- **Result3:**  $\underline{\mathbf{h}}_{opt} = \frac{B}{A \sin w_o} \begin{bmatrix} \sin[w_o - \varphi] \\ \sin \varphi \end{bmatrix}$  and  $E\{(e[n])^2\} = \varepsilon = r_v[0]$

#### 3.2:

- **Hint2:** Sampling frequency 8 kHz implies 8000 samples per second. The time difference between consecutive samples is  $1/8000 = 0.125$  ms. Therefore, the delay is one sample and we need, at least, to process the current and the previous samples in the filter:  $M=2$ .
- **Result(d):**

$$\rho \sigma_v^2 = h_0(\sigma_m^2 + \sigma_v^2) + h_1(\rho \sigma_m^2 + \rho \sigma_v^2)$$

$$\sigma_v^2 = h_0(\rho \sigma_m^2 + \rho \sigma_v^2) + h_1(\sigma_m^2 + \sigma_v^2)$$

- **Result(e):**  $h_0 = 0$  and  $h_1 = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_m^2}$
- **Result(f):** Case A: if the musical component is small, the filter just takes the previous sample. Case B: if the voice component is small, the filter just tries to eliminate it
- **Result(g):**  $h_0 = 0$  and  $h_1 = 1$  and, therefore,  $E\{(e[n])^2\} = \varepsilon = \sigma_m^2$  as the right channel voice signal has been cancelled

3.3:

- **Result(a):**  $r_s[l] = \frac{49}{36} 0.8^{|l|}$  and  $r_x[l] = \frac{49}{36} 0.8^{|l|} + \delta[l]$
- **Result(c):**  $E\{(e[n])^2\} = 0.462$

3.4:

- **Hint:** The available data for the prediction are the B and C pixels, that can be included in a vector  $\underline{x}^T = (B, C)$  containing the observations, while the reference is pixel A.
- **Result(a, b, c):**  $h_B = h_C = \frac{\rho}{1+\rho^2}$ . The prediction error power is  $E\{(e[n])^2\} = \sigma^2 \frac{1-\rho^2}{1+\rho^2}$  and  $G_2 = \frac{1+\rho^2}{1-\rho^2}$
- **Result(d):**  $G_2' = \frac{1}{1.5-2\rho+0.5\rho^2}$

3.5:

- **Result(c):**  $J_d = r_x[0] \left(1 - \frac{r_x^2[1]}{r_x^2[0]} + d^2\right)$
- **Result(d):**  $G_p' = 1.9$

3.6:

- **Result(b):**
- **Hint(c):** See Hint3(a) in Exercise 3.1.
- **Result(c):**  $D = 1$  and  $M = 2$ .  $A(z) = 1 - h_0 z^{-1} - h_1 z^{-2}$
- **Hint(e):** Show that the eigenvalues of  $v[n]$  are the eigenvalues of the sinusoid plus  $\sigma^2$ .
- **Result(e):**  $\sigma^2 \uparrow \Rightarrow N_{iter} \downarrow$  and  $\sigma^2 \uparrow \Rightarrow D \uparrow$

3.7:

- **Result(a):**  $h_{opt} = \frac{r_x[1]}{r_x[0]}$  and  $E\{(e[n])^2\} = \varepsilon = r_x[0] \left(1 - \left(\frac{r_x[1]}{r_x[0]}\right)^2\right)$
- **Result(b):**  $h_{a_{opt}} = \frac{r_x[1]-m_x^2}{r_x[0]-m_x^2}$  and  $m_{x_{opt}} = \frac{r_x[0]-r_x[0]}{r_x[0]-m_x^2} m_x$
- **Result(c):** Defining  $c_x[l] = (r_x[l] - m_x^2)$ , then  $\varepsilon = c_x[0] \left(1 - \left(\frac{c_x[1]}{c_x[0]}\right)^2\right)$
- **Result(d):**  $h_a[n+1] = h_a[n] + \mu x[n-1]e[n]$  and  $b[n+1] = b[n] + \mu e[n]$
- **Result(e):**  $0 < \mu < \frac{2}{r_x[0]+1}$

3.8:

- **Result(a):**  $(\underline{\mathbf{R}}_x + \alpha \underline{\mathbf{I}}) \underline{\mathbf{h}}_{opt} = \underline{\mathbf{r}}_{xd}$
- **Result(b):**  $\alpha > 0$  ensures that the matrix  $(\underline{\mathbf{R}}_x + \alpha \underline{\mathbf{I}})$  can be inverted
- **Result(c):**  $\underline{\mathbf{h}}[n+1] = (1 - \mu\alpha) \underline{\mathbf{h}}[n] + \mu x[n] e[n]$
- **Result(e, f):**  $0 < \mu < \frac{2}{\lambda_{max} + \alpha}$  and  $N_{iter} \propto \frac{\lambda_{max} + \alpha}{\lambda_{min} + \alpha}$

## Exam Exercises

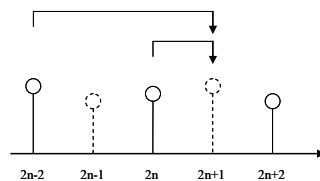
**Final 2019:** Sigui un senyal real  $x[n]$  que es pot modelar com un procés AR(1) a on la potència del soroll blanc generador és coneguda  $\sigma_w^2$  i el coeficient de correlació entre mostres consecutives de senyals és

$$\rho = \frac{r_x[l+1]}{r_x[l]}$$

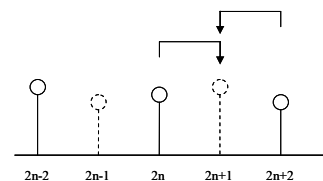
1) Amb les equacions de Yule-Walker, troba una expressió de les mostres de l'autocorrelació de  $x[n]$  en funció de la potència de  $x[n]$  i el coeficient  $\rho$ . Determina el valor de la potència del procés  $x[n]$ . Discuteix el rang de valors possibles del coeficient  $\rho$ .

D'aquest senyal es té una versió  $z[n]$  (corresponent a les mostres parelles del senyal  $x[n]$ ) que s'ha obtingut eliminant les mostres senars del senyal  $x[n]$ . Vegeu la figura adjunta que representa  $x[n]$  on les mostres discontinues corresponen a les mostres senars eliminades.

Per reconstruir el senyal  $x[n]$  a partir de les mostres del senyal  $z[n]$  es proposa estudiar dos predictors diferents, tots dos de 2 coeficients. Es farà servir en els dos casos la solució amb error quadràtic mitjà mínim i es compararan tant la seva versió adaptativa com no adaptativa.



Predictor A



Predictor B

2) Troba els coeficients del predictor A que permet predir les mostres senars de  $x[n]$  a partir de les mostres de  $z[n]$ . Justifica el resultat. Troba la potència de l'error de predicció que s'obté.

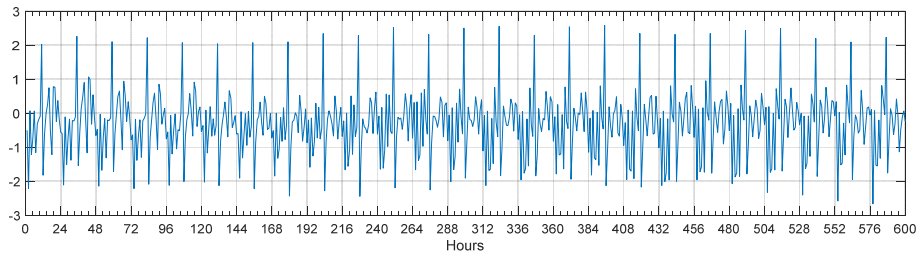
3) Troba els coeficients del predictor B que permet predir les mostres senars de  $x[n]$  a partir de les mostres de  $z[n]$ . Troba la potència de l'error de predicció que s'obté amb aquest predictor. Justifica els dos resultats obtinguts, comparant-los amb els obtinguts a l'apartat 2.

Si la solució dels dos predictors s'ha obtingut amb un sistema adaptatiu que utilitza l'algorisme LMS:

4) Troba pels dos casos el rang de valors del paràmetre  $\mu$  del sistema que assegura la convergència a la solució de Wiener.

5) Discuteix la velocitat de convergència en funció del paràmetre  $\rho$ .

**Revaluació 2019:** Sovint, les sèries temporals (senyals que contenen les dades, observacions o valors mesurats en determinats moments del temps) presenten comportaments cíclics al llarg del temps. La figura adjunta, on pot veure's un comportament cíclic diari, n'és un cas.



Se'ns planteja la necessitat d'analitzar la seqüència anterior per poder capturar el comportament cíclic observat i predir valors futurs del senyal. Per realitzar aquesta anàlisi és bàsic partir d'un model del senyal. Pel senyal  $x[n]$  de la figura es proposa el següent model:

$$x[n] = s_1[n] + s_2[n]$$

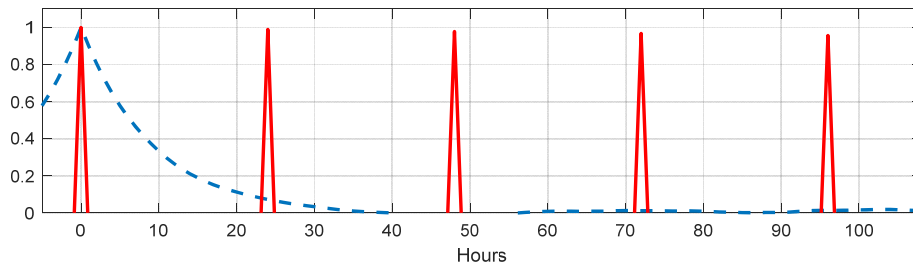
on el senyal  $s_1[n] = \rho_1 s_1[n-1] + w_1[n]$  és un procés autoregressiu d'ordre 1 i, pel seu costat, el senyal  $s_2[n] = \rho_2 s_2[n-24] + w_2[n]$  modela el comportament cíclic diari del senyal. Els senyals  $w_1[n]$  i  $w_2[n]$  són dos processos blancs, de mitjana nul·la, incorrelats entre ells i de potències  $\sigma_{w1}^2$  i  $\sigma_{w2}^2$ , respectivament.

1. Sabent que les autocorrelacions dels processos  $s_1[n]$  i  $s_2[n]$  obeeixen a la següents equacions de recurrència:

$$r_1[m] = \rho_1 r_1[m-1] + \sigma_{w1}^2 \delta[m]$$

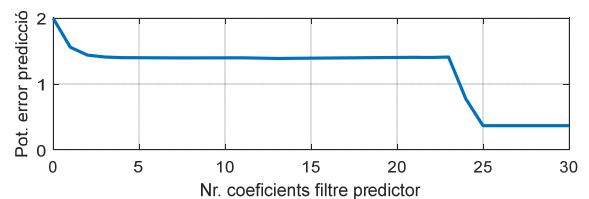
$$r_2[m] = \rho_2 r_2[m-24] + \sigma_{w2}^2 \delta[m]$$

- Calcula l'autocorrelació del senyal suma  $x[n]$ .
  - Determina el valor de la variància  $\sigma_{w1}^2$  en funció de  $\rho_1$  i de la potència de  $s_1[n]$ .
  - Determina el valor de la variància  $\sigma_{w2}^2$  en funció de  $\rho_2$  i de la potència de  $s_2[n]$ .
2. La següent figura mostra l'autocorrelació dels senyals  $s_1[n]$ ,  $s_2[n]$  pel cas de que la potència dels senyals sigui unitària i per  $\rho_1 = 0.9$  i  $\rho_2 = 0.99$ . Indica, justificant correctament la resposta, quina gràfica correspon a cadascuna de les autocorrelacions.



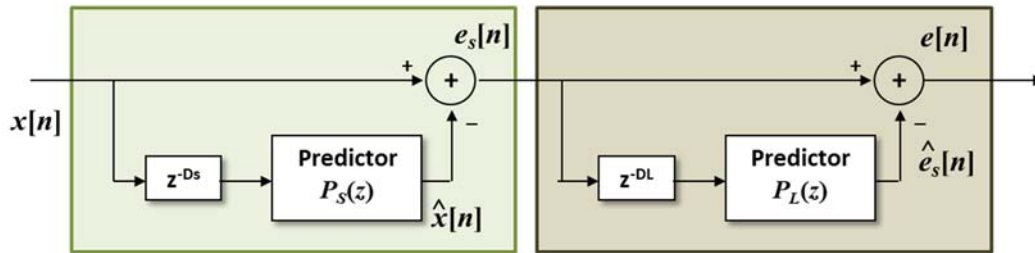
Un cop coneguda l'autocorrelació de la seqüència, et proposes dissenyar un predictor que et serveixi per determinar valors futurs de la seqüència. A la pregunta ¿quants coeficients  $Q$  hauria de tenir el filtre predictor? no saps ben bé que respondre i decideixes realitzar una simulació per diferent nombre de coeficients. El resultat és el de la figura adjunta.

3. Et preguntes a què es deuen les dues davallades que apareixen en la potència de l'error de predicció i perquè la primera (al voltant de 1-2 coeficients) és menor que la segona. Pots explicar el comportament del predictor en funció del nombre de coeficients del filtre  $Q$ ?



Si bé comproves que el predictor que has dissenyat prediu força bé el senyal  $x[n]$ , hi veus dos problemes: el primer és que necessites un nombre elevat de coeficients del filtre ( $Q > 25$ ), i el segon que no has estat capaç de separar els comportaments cíclic i no cíclic del senyal (això és,  $s_1[n]$ ,  $s_2[n]$ ). Reproduint esquemes similars als emprats per

codificar senyals de veu se't planteja dissenyar un predictor com la combinació de dos filtres predictors en sèrie, un primer filtre que prediu a curt termini el senyal (Predictor  $P_S$ ) i un segon filtre que prediu el senyal a llarg termini (Predictor  $P_L$ ). Cadascun dels dos filtres predictors correspon a un filtre de Wiener que es dissenyarà a continuació.

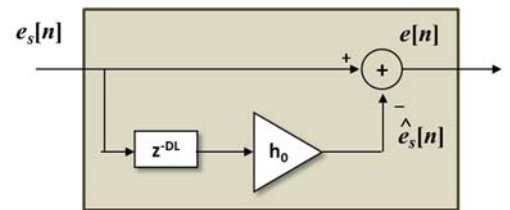


#### Disseny del filtre predictor a curt termini:

4. El filtre predictor a curt termini correspondrà a un filtre de Wiener de longitud  $Q = 2$  coeficients que vol predir la mostra  $x[n]$  a partir de les dues mostres anteriors ( $D_S = 1$ ). Troba els coeficients del predictor que permeten predir aquest senyal. Troba la potència de l'error de predicció que s'obté amb aquest predictor. Nota: Utilitza que la potència dels processos  $s_1[n]$  i  $s_2[n]$  és en ambdós casos unitària i deixa les expressions en funció de  $\rho_1$  i  $\rho_2$ .

#### Disseny del filtre predictor a llarg termini:

5. El filtre predictor a llarg termini correspondrà a un filtre de Wiener amb un sol coeficient  $Q = 1$  que vol predir la mostra  $x[n]$  a partir de la mostra  $x[n - D_L]$ . Troba el valor del coeficient del predictor (en funció de  $D_L$ ) que minimitza l'error de predicció. Pots intuir quin ha de ser el valor de  $D_L$ ? Pel valor de  $D_L$  òptim troba la potència de l'error de predicció que s'obté amb aquest predictor.



6. Mitjançant aquest plantejament d'utilitzar dos predictors, un a curt termini i un a llarg termini, has aconseguit separar els senyals  $s_1[n]$  i  $s_2[n]$ ? En cas afirmatiu, en quins punts de l'esquema es troben?



## Topic 4: Signal Transforms

### Basic Exercises

#### 4.1.

1. Which is the relationship between the Fourier transforms  $X(e^{j\omega_x}, e^{j\omega_y})$  of the images A and B? Note that B has been obtained by a symmetry with respect to the central vertical axis.
2. Which is the relationship between the Fourier transforms  $X(e^{j\omega_x}, e^{j\omega_y})$  of the images A and C? Note that C has been obtained by a symmetry with respect to the central point of the image.



Image A



Image B



Image C

4.2. Consider the following image  $x[m, n]$  (of size  $N \times N$ ):

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

We compute its 2D-DFT  $X[k, l]$  (size  $N \times N$ ) and multiply each coefficient by a complex exponential:  $Y[k, l] = X[k, l]e^{-j2\pi(\frac{k}{4} + \frac{l}{2})}$ . Define the image  $y[m, n]$  that corresponds to the inverse DFT of  $Y[k, l]$ .

4.3. In the Figure 1, an image  $x[m, n]$  of 96x96 pixels is presented. We apply different linear transforms to this image to code it with different compression values.

We compute its 2D-DFT  $X[k, l]$  (size  $N \times N$ ) and multiply each coefficient by a complex exponential:  $Y[k, l] = X[k, l]e^{-j2\pi(\frac{k}{4} + \frac{l}{2})}$ . Define the image  $y[m, n]$  that corresponds to the inverse DFT of  $Y[k, l]$ .

First, we analyze the use of DCT and, therefore, we divide the image in 12x12 blocks of 8x8 pixels

- a) Why do you think that the block decomposition is useful?
- b) In every block, how many DCT coefficients are necessary in order not to include information losses?
- c) In Figures 2a), 2b) and 2c), we present the images that have been obtained only using two DCT coefficients. In all three cases, a coefficient is that of maximal energy ( $X_{0,0}$ ). Discuss which is the other coefficient that has been used. The possibilities are:  $X_{0,1}$ ,  $X_{1,0}$  or the second coefficient with highest energy.



Figure 1- Original image

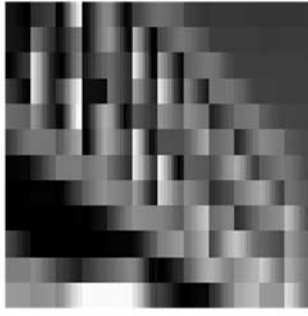


Figure 2a)

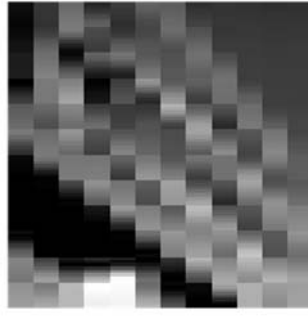


Fig. 2b)

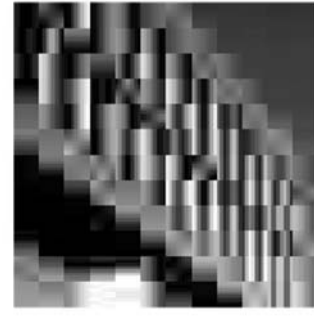


Fig. 2c)

- d) Describe the resulting image if only the coefficient of highest energy is used in its reconstruction.
- e) In the previous case, which will the minimum possible PSNR be? Assume that grey level values are in the  $[0, 255]$  range.

#### 4.4.

1. Define the impulse and frequency responses of the two Sobel filters.
2. Explain their use to detect contours in an image.
3. We want to estimate the horizontal and vertical second derivatives of an image ( $\frac{\partial^2 f(x,y)}{\partial x^2}$  and  $\frac{\partial^2 f(x,y)}{\partial y^2}$ , respectively). Towards this goal, we decide to apply twice every Sobel filter. Justify this strategy and compute the impulse and frequency responses corresponding to the iteration of each Sobel filter.
4. Finally, and in analogous way as with the Laplacian filter, we merge both derivatives:  $\frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}$ . Define the impulse response of the equivalent filter and propose a strategy for, using that filter, identify the position of the contours in an image.

#### 4.5. We use the following basis to represent 2x2 image blocks:

$$a_{00}[m,n] = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad a_{01}[m,n] = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \quad a_{10}[m,n] = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \quad a_{11}[m,n] = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

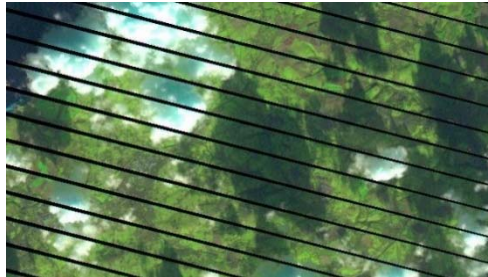
- a) Show that it is an orthonormal basis.
- b) Compute the transformed coefficients of the image  $I[m,n] = \begin{bmatrix} 4 & 8 \\ 3 & 5 \end{bmatrix}$
- c) Find the reconstructed image using the two coefficients with greatest magnitude.
- d) Compare the original and the reconstructed images and comment the approximation effect obtained the two most significant coefficients.

#### 4.6. Let us consider the image and the filter defined in the following matrices. We want to compute the image $y[m,n]$ resulting from filtering $x[m,n]$ with $h[m,n]$ using the zero-padding and the mirroring techniques. Compute the resulting image when using the transpose filter $h^T[m,n]$ . Note that the filter is separable.

$$x[m,n] = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \quad h[m,n] = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

## Exam Exercises

**Final 2019:** En el dossier que utilitza el director del Parc Nacional del Teide per mostrar-vos exemples i dades, has vist que hi ha imatges de satèl·lit del parc més antigues. Li preguntes si les volen fer servir i et comenta que sí, que els hi agradaria fer un estudi retrospectiu de com ha evolucionat el parc. Ara bé, et diu que moltes de les imatges que tenen no són molt bones i, en concret, presenten una interferència amb un patró constant, tal i com es mostra a la figura.



Imatge de satèl·lit amb interferència deguda a un sensor malmès.

Comentes al director del parc que estudiareu el problema. Quan acaba la reunió, et teu equip et pregunta com penses atacar aquest problema. Per debatre amb l'equip la idea que tens, modeles el senyal observat  $y[m, n]$  com el producte de dues imatges: la imatge original  $x[m, n]$  i la que presenta aquest patró d'interferència  $i[m, n]$ .

Com que vols simplificar el debat i centrar-ho amb el que et sembla important, imposes que la imatge és de mida  $N \times M$ , el patró d'interferència horitzontal i de gruix un píxel i que la imatge té un nombre  $N_p$  exacte de períodes del patró d'interferència els quals són de durada  $L_p$ , tal que  $N = N_p L_p$ .

- a) Aplicant la propietat de separabilitat, dona l'expressió de la imatge que modela el patró d'interferència  $i[m, n] = i_x[m]i_y[n]$ . Escriu la Transformada de Fourier Discreta (DFT) del senyal observat  $y[m, n]$  en funció de la DFT de la imatge original  $x[m, n]$  i de la DFT del patró d'interferència  $i[m, n] = i_x[m]i_y[n]$ . Calcula i dibuixa la DFT del patró d'interferència  $i[m, n] = i_x[m]i_y[n]$  i, després, calcula la del senyal observat  $y[m, n]$ . Interpreta el resultat (si no pots calcular la DFT però creus que pots intuir i interpretar el resultat que s'obindrà, fes-ho).

Els membres de l'equip veuen clara la complexitat de l'anàlisi del senyal en l'àmbit freqüencial i, per tant, proposen atacar el problema amb un filtre lineal i invariant en l'espai definit en l'àmbit espacial.

- b) Discuteix quins tipus de filtre lineal i invariant en l'espai faries servir per reduir la presència d'aquesta interferència simplificada que has proposat. Raona quin filtre implementaries i quins punts febles presenta per a aquesta aplicació concreta.

Un membre de l'equip et pregunta com canviaria el problema i la seva solució si no s'imposessin algunes de les simplificacions que has fet.

- c) Raona (no fa falta que ho calculis, però sí que ho justifiquis) com canviaria la DFT del senyal i la solució proposada si la interferència tingués inclinació diferent de l'horitzontal.
- d) Raona (no fa falta que ho calculis, però sí que ho justifiquis) com canviaria la DFT del senyal i la solució proposada si la interferència fos horitzontal, però de gruix  $G$  píxels.

Finalment, un altre membre de l'equip proposa utilitzar un filtre que no sigui lineal o/ni invariant en l'espai.

- e) Discuteix quina proposta faries si poguessis fer servir un filtre que no fos lineal o/ni invariant en l'espai.

**Revaluació 2019:** A l'empresa d'anàlisi de dades multimèdia a la que treballes us han arribat darrerament diversos projectes que impliquen l'emmagatzematge d'imatges i, alguns d'ells, tenen requeriments molt diferents. Així, avui teniu una reunió per analitzar si fer servir la mateixa tècnica de compressió per tots els projectes o si n'hi ha algun que hagi de tenir una tècnica pròpia. Les propostes actuals que esteu analitzant són tècniques basades en la DCT o la KLT.

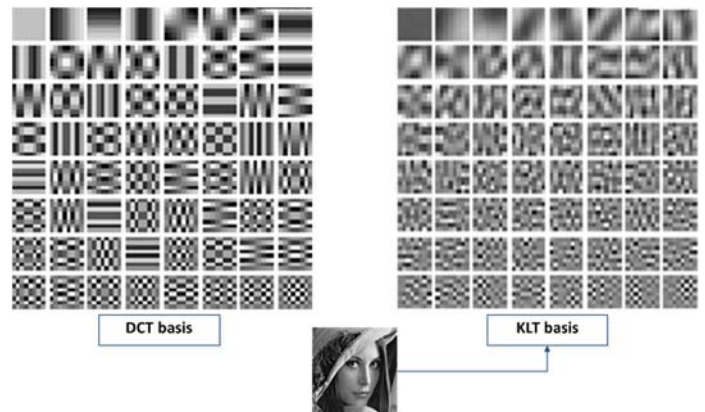
Abans de la reunió, i per poder presentar la teva opinió, analitzeu els diferents projectes i en trieu dos: un d'anàlisi de dades mèdiques a on us demanen que treballeu amb imatges de raigs X i un altre per una empresa de viatges a on, entre d'altres coses, els usuaris poden generar-se un àlbum multimèdia resum de les seves vacances. Et demanen que comencis tu la reunió, a la qual han vingut tècnics i comercials.

1. Expliqueu per a aquesta audiència com funciona un codificador d'imatge basat en blocs i, en aquest context, expliqueu les característiques bàsiques de les dues transformades: la DCT i la KLT. Finalment, raoneu per a cadascun dels dos projectes que has triat, quina transformada proposaries.

A l'equip hi ha una matemàtica que et demana si pots explicar la raó per no triar sempre un sistema basat en la KLT, donat que aquesta transformada és òptima.

2. Ajudant-te de la informació que tens a la Figura, li raoneu la resposta. (Nota: Explica clarament què representa la informació a aquesta figura).

L'informàtic que ha de fer-se càrrec de la programació de la solució basada en KLT et demana si pots explicar-li clarament com ha d'actuar el seu programa:



3. Escriu el pseudo-codi del programa que, a partir d'una base de dades d'imatges sense comprimir, retorna la mateixa base de dades amb les imatges comprimides amb la KLT sobre blocs de  $N \times N$  píxels, així com la informació necessària per descomprimir-les.

La persona al càrrec de la solució basada en la DCT és una enginyera de telecomunicació (no de la menció audiovisual) que ha estat analitzant l'estàndard JPEG. L'enginyera té dubtes del pas de quantificació de l'estàndard. Et comenta que no veu la relació entre el que ella entén per un quantificador i el que proposa l'estàndard: aplicar una "Taula de quantificació" ("*Quantization table*") i un arrodoniment posterior:

$$X_Q[k, l] = \text{round}(X[k, l]/Q[k, l])$$

Per a ella, un quantificador és un tipus de transformació o mapejat ("*mapping*") que actua sobre els valors d'una variable o d'una funció, mostra a mostra, posant en correspondència un nombre ampli de possibles valors d'entrada amb un nombre menor de possibles valors de sortida.

4. Per tal d'explicar-li, fas servir un cas simplificat de l'estàndard JPEG. Així, utilitzes només la component de luminància de la imatge, la qual suposes que utilitza 8 bits per mostra, fas servir una transformada discreta cosinus unitària, imposes que les components AC tenen com a rang de valors  $[-128, 128]$  i utilitzes la següent taula de quantificació. Estudies el quantificador que es fa servir pel coeficient de DC (0,0) i pel coeficient d'AC (2, 2) i li comentes quin tipus de quantificador es té.

$$Q = \begin{bmatrix} 2 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$