

# Treballs pràctics de Processament Digital del Senyal Curs 2002-2003

Bartomeu Coll i José Luis Lisani

L'objectiu d'aquests treballs és avaluar els coneixements adquirits pels alumnes al llarg del curs. La qualificació obtinguda dels treballs, juntament amb la qualificació dels informes de les pràctiques de laboratori, conformaran la nota final de l'alumne.

Els treballs són individuals i consisteixen en l'escriptura de programes de Matlab que implementin alguns algoritmes clàssics de processament d'imatges.

Cada alumne haurà de presentar un informe que inclourà:

- Explicació de la base teòrica de l'algoritme.
- Descripció detallada de l'algoritme i de cada una de les funcions de Matlab escrites per implementar-ho.
- Nombrosos experiments que mostrin l'efecte de l'algoritme damunt diferents imatges. S'ha de provar l'efecte de la modificació dels paràmetres de l'algoritme i comentar els resultats obtinguts.
- En un disquette s'ha d'incloure el codi font del programa implementat així com les imatges utilitzades en la part experimental. Els professors avaluaran la claretat i senzillesa del codi programat.

Els treballs es presentaran davant la resta de la classe durant la darrera setmana de curs. Cada alumne disposarà d'un màxim d'una hora per explicar el seu algoritme, mostrar els resultats obtinguts i contestar a les preguntes dels professors i de la resta de companys.

Els treballs proposats són els següents:

1. Equalització d'histograma i filtratge homomòrfic.
2. Filtre de grà o aniquilador d'extrems (*extrema killer*).
3. Detector de contorns clàssic.
4. Segmentació d'imatges per creixement de regions (*region growing*).

Es poden emprar totes les funcions de Matlab necessàries per al desenvolupament dels algoritmes (per exemple, si s'ha de calcular una transformada de Fourier, se pot emprar la funció *fft2* de Matlab). Queden excloses, evidentment, aquelles funcions de Matlab que implementin directament algun dels algoritmes sol·licitats (queda prohibit per tant utilitzar la funció *histeq* per fer el primer treball).

## 1 Equalització d'histograma i filtratge homomòrfic

S'han de desenvolupar dos algorismes diferents, un per a l'equalització d'histograma i l'altre per al filtratge homomòrfic. La informació necessària per implementar els algorismes es subministrarà a l'alumne en forma de fotocòpies (*Digital Image Processing*, Gonzalez and Woods, Ed. Addison-Wesley, 1993. pp.171-180, 212-218).

## 2 Filtre de grà o aniquilador d'extrems (*extrema killer*)

Una manera de representar una imatge de nivell de gris  $u : \mathbb{R}^2 \rightarrow \mathbb{R}$  és considerar-la com una funció tri-dimensional, on l'alçada de cada píxel  $(x, y)$  vé donada pel seu nivell de gris. D'aquesta manera, la tripleta de valors  $(x, y, u(x, y))$  representen un punt de l'espai (veure la figura 1).

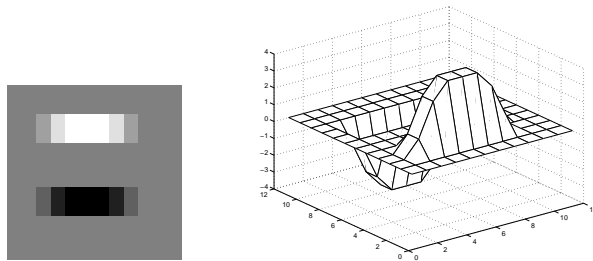


Figure 1: Imatge de nivell de gris i la seva representació tridimensional

Considerada d'aquesta manera, una imatge és equivalent a un conjunt de 'muntanyes' i 'valls', on les primeres representen les zones clares de la imatge i els segons les zones oscures. El filtre de grà o aniquilador d'extrems elimina els pics més punteguts i els avencs més petits d'aquesta representació de la imatge. Aquest pics i avencs són, formalment, els extrems de la imatge. El filtre depèn d'un únic paràmetre què és el l'àrea dels extrems a eliminar ( $A$ ).

La figura 2 representa l'efecte del filtre de grà per al cas uni-dimensional i amb un paràmetre  $A = 2$ , què en aquest cas representa la longitud màxima dels extrems a eliminar.

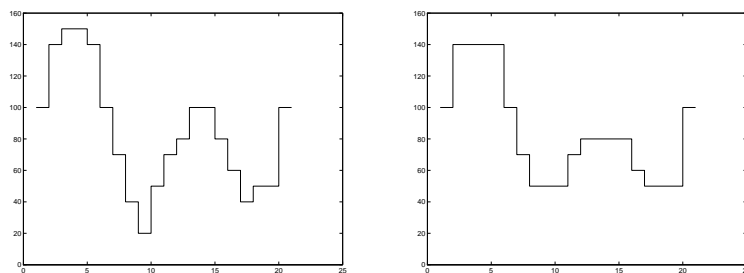


Figure 2: Funció uni-dimensional i funció filtrada amb un filtre de 'grà' de paràmetre 2

L'algoritme a implementar és el següent:

1. Trobar els píxels de la imatge què són extrems locals (màxims o mínims):  $(i, j)$  és mínim local (resp. màxim) si  $u(i, j) \leq u(i', j')$  (resp.  $u(i, j) \geq u(i', j')$ )  $\forall i' \in \{i-1, i+$

- $1\}$ ,  $j' \in \{j-1, j+1\}$ . (Estam considerant aquí 4-connectivitat).
2. Fer ‘crèixer’ cada un dels extrems trobats: trobar un grup de píxels connexos tals que el seu nivell de gris mínim (resp. màxim) sigui major (resp. menor) que el nivell de gris de tots els seus veïnats. El creixement de l’extrem es farà fins que es deixi de complir la condició d’extrem o fins que el nombre de píxels en el grup sigui més gran que el paràmetre  $A$ .
  3. Eliminar els extrems amb un nombre de píxels inferior o igual al paràmetre  $A$ : a aquests píxels s’els assignarà un nou nivell de gris igual al del seu veïnat amb nivell de gris més alt (en el cas dels màxims) o amb nivell de gris més baix (en el cas dels mínims).
  4. Repetir el procés per a tots els extrems de la imatge.

### 3 Detector de contorns clàssic (Marr-Hildreth, 1980)

**Introducció.** Dins el camp de la visió per ordinador, el tractament o anàlisi de les imatges és una de les tasques fonamentals. Pensem en el cas de la tomografia axial computada (TAC), la qual ens dóna informació sobre les estructures internes del cos humà. Poder entendre aquestes imatges requereix la identificació i modelització de les superfícies dels objectes que són presents en les estructures en 3D. En el cas de les imatges per satèl·lit, un dels problemes prové del renou present en la imatge mateixa, potser degut a problemes de captació o transmissió. En aquest cas s’intenta fer un filtratge o preprocessament, intentant eliminar aquest renou i conservar la informació inherent a la pròpia imatge. Altres aplicacions podrien ser el reconeixement automàtic de les formes (reconeixement automàtic de la signatura, el problema de la videovigilància, etc.), en el qual s’intenta extreure una sèrie de caràcters que ajudin en la identificació de l’objecte. Una imatge a nivell continu és una funció fitada que ens mesura a cada punt la intensitat (o energia) de la llum que incideix en aquesta part de l’objecte. És evident que diferents parts de la superfície dels objectes reflecteixen diferents intensitats de llum. Llavors la diferència entre una fotografia o imatge natural i una imatge digital és donada pel tipus de codificació. Per posar una fotografia dins la memòria d’un ordinador es divideix la imatge en petits trossos quadrats, que s’anomenen píxels i en cada un d’aquests trossos o píxels se li associa un número que representa la lluminància. Normalment el negre està codificat pel zero, el 1 representa un color un poc menys negre, el 2 és encara un poc menys negre que el 1, etc. Dins la convenció que s’utilitza en el món de la informàtica, el 255 representa la codificació del color blanc. Com a conseqüència, operar amb imatges digitals vol dir operar amb matrius, ja que tota la informació de la imatge és continguda dins una matriu de números.

**Pràctica del filtratge Multiescala** Detectar els diferents objectes dins una imatge vol dir detectar els contorns dels objectes que hi apareixen. Pensem que una fotografia és una projecció d’una escena 3D i per tant, és complicat descriure els diferents objectes que hi són presents, degut als efectes de les oclusions, ombres, etc. Un contorn va associat a punt de discontinuïtat de la imatge. Llavors, a nivell digital es treballa amb les “màscare” que no són més que discretitzacions numèriques d’operadors matemàtics. En el cas dels contorns d’una imatge  $u$ , direm que el píxel  $(i, j)$  pertany a un contorn si  $\Delta u(i, j) = 0$  i si  $|\nabla u(i, j)|$  és prou gran. En aquest cas la màscara, centrada al punt central  $(i, j)$ , on  $i, j \in 0, 1, \dots, N$ , que ens descriu la laplaciana ve donada per

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

i la que descriu el gradient pel càlcul de  $u_x$  i  $u_y$ . L'aproximació de  $u_x$  i  $u_y$  ve donada per

$$u_x = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

i la màscara per al càlcul de la  $u_y$

$$u_y = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Llavors  $\nabla u = (u_x, u_y)$ .

La teoria del "scale-space" o espai-escala, introduïda per Marr-Witkin (1982), La idea bàsica és que la informació de la imatge està continguda en diferents escales que ens donen les diferents característiques de la imatge a diferents nivells de resolució. Per això, farem un filtratge de la imatge en diferents nivells i en cada nivell, calcularem els contorns o característiques associades a dit nivell.

Un dels filtres més comuns ve donat per la convolució de la imatge amb la funció gaussiana. En aquest cas es pot construir una màscara que centrada al pixel  $(i, j)$  simuli numèricament la convolució amb la gaussiana. Bàsicament aquesta operació ens dona una mitja ponderada al voltant del pixel  $(i, j)$ , i per tant, quan fem aquesta operació, "regularitzem" o "suavitzem" la imatge. Definim una successió de màscares ponderades de la següent manera:

$$M_1 = \frac{1}{6} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

$$M_2 = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

$$M_3 = \frac{1}{36} \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 4 & 2 & 0 \\ 1 & 4 & 8 & 4 & 1 \\ 0 & 2 & 4 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

$$M_4 = \frac{1}{80} \begin{pmatrix} 0 & 1 & 2 & 1 & 0 \\ 1 & 4 & 8 & 4 & 1 \\ 2 & 8 & 16 & 8 & 2 \\ 1 & 4 & 8 & 4 & 1 \\ 0 & 1 & 2 & 1 & 0 \end{pmatrix}.$$

També es podria pensar de fer realment la mitja dels valors, i no la mitja ponderada, com és el cas de la gaussiana. En aquest cas el suavitzat es faria més ràpid, ja que no hi ha informació privilegiada de cap dels pixels de l'entorn.

### Part 1: Filtratge gaussià (sense direccions privilegiades)

(i) Sigui  $I_0$  la imatge original, i fem la convolució amb les diferents màscares introduïdes abans. S'obtenen les noves imatges  $I_1, I_2, I_3, I_4$ .

(ii) Per a cada imatge  $I_i$ , calculeu els seus contorns, és a dir: Passeu la màscara  $u_x, u_y$  i calculeu la imatge del mòdul del valor absolut  $Mod_i = u_x^2 + u_y^2$ . Feu el mateix amb la laplaciana i obtindreu  $L_i$ . Llavors definiu la nova imatge dels contorns com a  $Co_i(i, j)$ , on  $Co_i(i, j) = 0$  si tenim un contorn, és a dir, si  $Mod_i(i, j) > K$  i  $|L_i(i, j)| < \epsilon$ , on  $k, \epsilon$  són paràmetres prefixats, i  $Co_i(i, j) = 255$  en cas contrari (de no contorn).

### Part 2: Filtratge condicionat

Definim les següents màscares ponderades per direccions, on entenem que a nivell discret, tenim 4 direccions principals.

$$v_1 = \frac{1}{4} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$v_2 = \frac{1}{4} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

$$v_3 = \frac{1}{4} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

$$v_4 = \frac{1}{4} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Per tant podem definir un partició del domini de les direccions de la següent manera  $D_1 = [-\frac{11}{8}\pi, -\frac{3}{8}\pi], D_2 = [-\frac{3}{8}\pi, -\frac{1}{8}\pi], D_3 = [-\frac{1}{8}\pi, \frac{1}{8}\pi], D_4 = [\frac{1}{8}\pi, \frac{3}{8}\pi], D_5 = [\frac{3}{8}\pi, \frac{5}{8}\pi]$ . Llavors donada una direcció  $v$ , existeixen dues direccions, les de més a prop  $v_i, v_j$  tal que  $v = \theta_1 v_1 + \theta_2 v_2$ , on  $\theta_1 + \theta_2 = 1$ . Feu els mateixos passos que per la pràctica del filtratge gaussià.

La informació necessària per implementar els algoritmes es subministrarà a l'alumne en forma de fotocòpies (*Digital Image Processing*, Gonzalez and Woods, Ed. Addison-Wesley, 1993. pp.416-425).

## 4 Segmentació d'imatges per creixement de regions (*region growing*)

**Pràctica de Segmentació** L'operació de segmentació intenta dividir la imatge en les diferents regions disjunctes homegènies respecte a una propietat (segons un criteri de similitud (nivell de gris, textura, color, etc)), generalment relacionada amb el valor del nivell de gris (Per exemple, un criteri podria ser que la diferència absoluta en nivell de gris entre la "llavor" i el pixel candidat, no excedís d'un 10 per cent de la diferència entre el nivell de gris màxim i el mínim de la imatge sencera, i que el pixel candidat fos un veïnat de la regió que hi formarà part). El mètode més simple és l'anomenat creixement de regions o "region growing".

**Part 1** Feu un programa per segmentar imatges, basat en aquest model. Genèricament l'algorisme seria:

- a) Començar amb que tot pixel és una regió diferent.
- b) Fer una unió de tots els parells de regions que en “adjuntar-les” millora la segmentació, a partir d'uns pixels inicials o llavors, per la característica o propietat que agafam. Per exemple, si el valor mig del nivell de gris difereix tan sols en un valor  $\lambda = 2$ . Pensau en altres criteris.
- c) Agafar un  $\lambda$  més gran i iterar el procés.

**Formulació variacional** Existeix una formulació variacional del problema, suposant que les imatges que s'obtenen són constants a trossos. Aquest es un mètode que minimitza una certa energia, conegut com el funcional de Mumford-Shah:

$$E(u, K) = \lambda^2 \int_K d\sigma + \int_{\Omega-K} (u - u_0)^2 dx$$

on  $u_0$  és la imatge original. En aquest funcional,  $K$  és el la unió de corbes en que segmenta la imatge,  $\Omega$  és el domini de la imatge,  $u_0$  és la imatge incial que es vol segmentar i  $u$  és la imatge segmentada en el conjunt de corbes  $K$ . El primer terme minimitza la llargària de les fronteres de les regions mentre que el segon s'anomena terme de fidelitat, ja que no permet que la imatge resultant de la segmentació sigui molt “diferent” de la original.

**Part 2** d) Feu un programa basat en un mètode numèric que ens resol de manera iterativa el problema d'optimització, per exemple, del gradient, on poguem minimitzar el funcional, i per tant, trobar la solució aproximada, segons  $\lambda$ , de la segmentació de la imatge original.

La informació necessària per implementar els algoritmes es subministrarà a l'alumne en forma de fotocòpies (*Digital Image Processing*, Gonzalez and Woods, Ed. Addison-Wesley, 1993. pp.458-462).