

**UNIVERSITAT POLITÈCNICA DE CATALUNYA**

*Departament d'Enginyeria de Sistemes, Automàtica i Informàtica Industrial*

**Mètode d'Extracció Multiparamètrica  
de Característiques de Textura  
Orientat a la Segmentació d'Imatges**

Autor: Antoni Grau i Saldes  
Directora: Àlicia Casals i Gelpí

Barcelona, maig de 1997

## CAPÍTOL 4. CLASSIFICACIÓ I SEGMENTACIÓ.

L'objectiu d'aquest capítol és determinar, aplicar i millorar el mètode que utilitzarem en la classificació dels tèxels per obtenir les classes de textura que permetran efectuar la segmentació de la imatge original. Per aquest motiu, en aquest capítol es fa una breu ressenya de la metodologia de reconeixement de patrons (*pattern recognition*) i s'escull un mètode de classificació que s'adapta als requeriments imposats al sistema d'extracció de característiques de textura proposat: el baix cost computacional. El fet que els algorismes de classificació només treballen amb la informació de l'espai de característiques, és a dir, només tenint en compte les dades corresponents a les característiques extretes, fa que la classificació no sigui la desitjada. Per aquest motiu s'introdueix l'ús d'un mètode variacional per reduir els errors en la classificació aprofitant la informació sobre la relació espacial entre els píxels de les diferents classes, afavorint o penalitzant l'agrupació de píxels veïns o propers. Finalment es mostra un exemple de la millora introduïda.

Al final del capítol es realitza un estudi de diferents algorismes d'anàlisi de textura on es compara el mètode presentat en aquesta tesi en front d'altres mètodes àmpliament reconeguts. En aquest estudi es mostra el grau d'incertesa en la classificació respecte al temps de càlcul de les característiques de textura i es veurà que el nostre mètode té un grau d'incertesa molt elevat tot i les restriccions imposades per a poder operar a temps real.

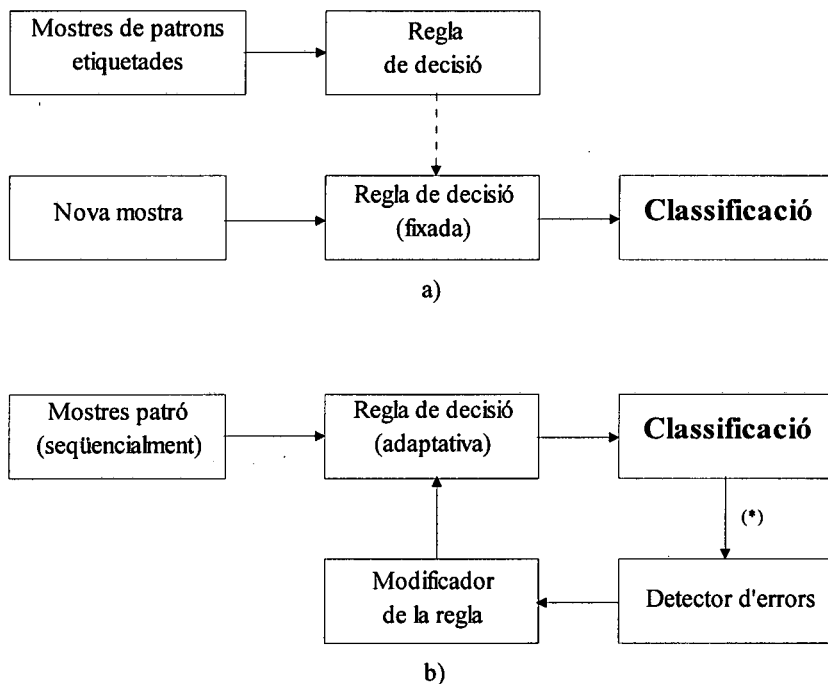
### 4.1. Introducció al reconeixement de patrons.

En el seu sentit més ampli, els patrons són els mitjans pels quals podem reconèixer els objectes que ens permetran interpretar el món. Quan un nadó neix aprèn a diferenciar el patró visual de la mare i del pare, diferencia patrons en els sons, patrons tàctils en el fred i el calor. A mida que va creixent, es va refinant el detall en el reconeixement de patrons i pot diferenciar entre el cant d'un rossinyol i d'una cadenera o entre l'olor de la farigola i del fonoll. En altres àmbits, també es pot abstrure la discriminació sensorial a patrons indirectes. Un sociòleg troba patrons en les dades que analitza, un metge detecta patrons en els seus malalts.

El procés del reconeixement de patrons consisteix en usar regles fixes o variables que s'han après de l'experiència per discriminar uns patrons d'uns altres. En el cas de discriminar entre el color vermell i el color blau, els humans ho poden realitzar sense dubte, usant regles fixes que realment existeixen, però que no són conegudes. L'home ha desenvolupat aquesta habilitat discriminatòria a partir de l'experiència. En algun moment, l'individu s'ha exposat repetidament a mostres del color vermell i a mostres del color blau i se li han indicat quines pertanyien a cada color. A partir de l'examen d'aquestes mostres etiquetades, adquireix el coneixement que permet elaborar la regla de decisió.

En el reconeixement de patrons hi ha dos aspectes importants: primer, desenvolupar una regla de decisió i segon, usar-la. El reconeixement real s'utilitza a l'hora d'usar la regla; mentre que el patró es defineix en el procés d'aprenentatge mitjançant mostres etiquetades. En el reconeixement de patrons matemàtic es desitja trobar una regla de decisió que pugui classificar exemples de patrons ràpidament. Normalment, no serà crític el temps emprat en l'aprenentatge del patró, és a dir, en la derivació de la regla de decisió.

Els dos estadis del reconeixement de patrons, derivació de la regla de decisió i l'ús d'aquesta, es poden realitzar concurrent o seqüencialment. En la figura 4.1 es poden veure les dues estratègies per al reconeixement.



**Figura 4.1.** Estratègies de reconeixement: a) aprenentatge abans del reconeixement; b) aprenentatge i reconeixement concurrents. Basada en [MEI72].

En la figura 4.1.a) es mostra el diagrama del procediment de reconeixement de patrons seqüencial en dues fases. Primer es recullen les mostres etiquetades de patrons i d'elles s'en deriva la millor regla de decisió basada en aquestes mostres. Aquesta regla s'usa sense canvis per a classificar les mostres no etiquetades.

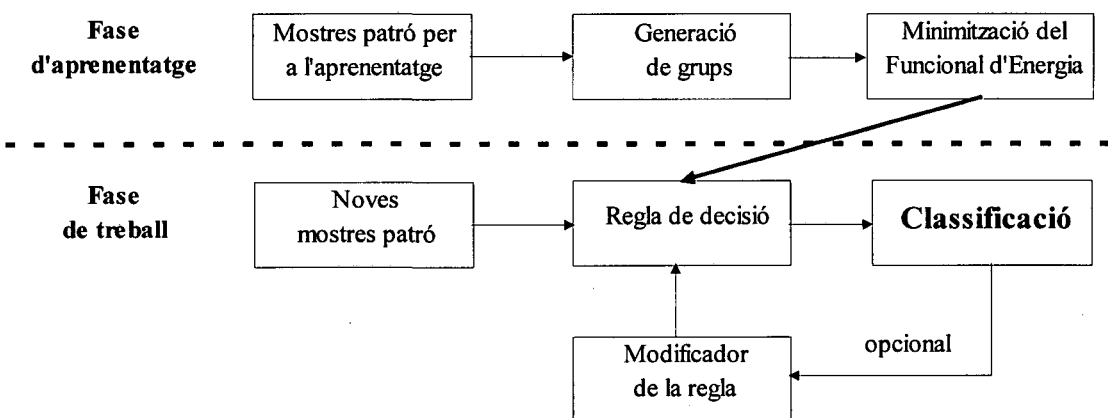
La figura 4.1.b) indica el cas on la regla de decisió es modifica a mesura que s'està usant. En aquest cas, es presenta una mostra i es classifica; un detector d'errors indica si la classificació ha estat correcta i, aleshores, la regla de decisió es modifica o no es modifica segons el cas. Si el detector d'errors s'anul·la a partir d'algun moment, tallant per (\*), i la regla de decisió es continua usant sense canvis, el procediment és efectivament seqüencial i correspon al que s'ha representat en la figura 4.1.a). Aquest procediment de realimentació és un tipus d'aprenentatge, ja que el rendiment del sistema millora presumiblement amb el temps.

Totes aquestes estratègies funcionen correctament quan es disposa de mostres que han estat obtingudes amb les condicions adequades. En el nostre cas hem fixat com a objectiu l'obtenció de les mostres a molt alta velocitat d'operació, per la qual cosa apareix un compromís entre la qualitat d'aquestes mostres i el temps de càlcul en la seva obtenció. Atès aquest objectiu, proposem una nova estratègia per a les fases d'aprenentatge i treball tal com es mostra en la figura 4.2.

La fase d'aprenentatge serà aquella en la qual es crearan els grups de mostres (elements de textura) a partir de la informació extreta sobre un conjunt d'imatges. Per millorar el contingut d'aquests grups de textura, es passarà a un segon mòdul on aplicarem un funcional del tipus energia sobre aquest agrupament tenint en compte no només la informació de l'espai de característiques sinó també la informació sobre el domini de la imatge 2D. Aquesta fase serà un procés *off-line* i no es tindrà en compte el temps per dur-la a terme.

Quan s'han creat els grups de textura amb les imatges d'aprenentatge, es passa a una segona fase o fase de treball, on per cada mostra d'entrada se li assigna un grup dels creats anteriorment. Aquesta fase serà un procés *on-line* i el temps d'assignació de mostres als diferents grups serà crític. Per tant, plantejarem una regla d'assignació que treballa en temps real. Concurrentment amb el reconeixement, pot passar que les mostres d'entrada evolucionin de manera que els grups creats en la primera fase quedin desplaçats o obsolets, de manera que, opcionalment, es poden actualitzar els grups amb les noves mostres que arriben. Aquest procés serà d'aprenentatge incremental i reconeixement en concurrència.

Si l'aplicació no requereix operar a velocitat de temps real, es pot aplicar el funcional d'energia en la fase de treball, donada la millora que produeix sobre la segmentació resultant, tal com es veurà en el següent capítol de resultats.



**Figura 4.2.** Diagrama de l'estratègia de classificació.

Fase d'aprenentatge: aprendre abans de classificar.

Fase de treball: aprendre incrementalment i classificació en concurrència.

En la figura 4.2 es poden veure esquemàticament les dues fases del sistema de reconeixement i els seus mòduls. Més endavant, en aquest mateix capítol, es desenvoluparà detalladament aquest sistema.

La derivació de la regla de decisió és un factor clau en el reconeixement de patrons. La figura 4.3 mostra el procés necessari per derivar aquesta regla de decisió. El sistema físic és la font de les dades per a qualsevol problema pràctic. Aquestes dades pertanyen a l'*espai de mesures*. Una mostra d'un patró ve representada per uns valors específics de totes les mesures, corresponent a un punt en l'espai de mesures. Els algorismes de classificació no s'apliquen però sobre l'espai de mesures sinó sobre un altre espai de dimensions, habitualment, més reduïdes. Aquest procés de *selecció de característiques* consisteix en descriure una mostra en l'espai de mesures per un conjunt finit i més petit de valors anomenats característiques. Una característica és una component d'un punt sobre l'*espai de patrons* el qual compleix que 1) és de dimensió finita i relativament baixa i 2) conté suficient informació com per realitzar la classificació satisfactòriament. Segons el sistema físic, l'espai de mesures pot ser idèntic a l'espai de patrons. Finalment, cal desenvolupar una regla de decisió en base a un conjunt finit de mostres etiquetades, la qual podrà classificar un punt en l'espai de patrons corresponent a una mostra no etiquetada. El procés de derivar la regla de decisió s'anomena *classificació de patrons*. Les dues components principals en el reconeixement de patrons són la selecció de característiques i la classificació de patrons [MEI72].



**Figura 4.3.** Derivació de la regla de decisió.

Els algorismes de classificació de patrons es divideixen sovint en paramètrics i no paramètrics. Un mètode paramètric de classificació defineix la funció discriminant com una classe de densitats de probabilitat integrada per un nombre relativament petit de paràmetres. La majoria dels mètodes paramètrics parteixen de l'assumpció que cada classe de patrons prové d'una distribució Gaussiana o normal, on els paràmetres són la mitjana i les covariàncies. Aquesta funció discriminant ens indicarà si una mostra pertany o no pertany a una classe. Els mètodes no paramètrics, malgrat el seu nom, sovint utilitzen funcions discriminants parametritzades, com ara els coeficients d'un polinomi multivariable d'un determinat grau. El que significa el terme no-paramètric és que no s'assumeix cap forma convencional de distribució de probabilitat.

Un conjunt de mostres de complexitat raonable està format normalment per subconjunts discriminables. Descriure i situar en l'espai de patrons aquests subconjunts dona una descripció de l'estructura de les dades la qual és útil en sí mateixa i la qual, també, simplifica o soluciona el problema de la classificació de patrons. El procés que descriu i situa aquests subconjunts s'anomena anàlisi de grups (*cluster analysis*).

La diferència entre la classificació de patrons i l'anàlisi de grups està en el fet que en l'anàlisi de grups les mostres no estan etiquetades. Això es pot interpretar, fonamentalment, de dues maneres:

- 1) Totes les mostres corresponen a mostres d'una sola classe i els grups que es generen corresponen a subclasses d'aquesta única classe. Per exemple, la lletra 'a' i la lletra 'A' són subclasses de la classe que consisteix en la primera lletra de l'alfabet. Es pot usar aquesta informació per estructurar una regla de decisió per al reconeixement de patrons.
- 2) Les mostres no tenen cap classe assignada. Els objectius d'utilitzar l'anàlisi de grups poden ser diversos: trobar grups "naturals" corresponents a classes "naturals" dins les dades; dividir les mostres en grups més reduïts per aplicar un algorisme de classificació local dins cada grup; reconstruir una densitat de probabilitat a partir de les mostres o simplement comprendre millor el sistema físic del qual provenen les mostres classificades.

El nostre cas correspon a l'anàlisi de grups donat que no existeixen mostres etiquetades. Totes les mostres que arriben al mòdul d'anàlisi de grups no tenen cap grup assignat prèviament, ni en la fase d'aprenentatge ni en la fase de treball i, de fet, buscarem els grups naturals que corresponen a textures semblants.

## **4.2. Agrupament de mostres.**

Abans d'aplicar qualsevol procediment d'agrupament de mostres, s'han de tenir en compte dos problemes importants. Primer, cal decidir una mesura de semblança entre mostres, per això es requereix definir què s'entén per semblança. Segon, cal especificar un procediment per formar els grups basat en la mesura de semblança escollida. Aquests seran els objectius que buscarem a continuació.

### **4.2.1. Mesures de semblança.**

Per l'ús de qualsevol tècnica d'agrupament serà fonamental el càlcul d'una mesura de semblança o distància entre les respectives mostres a l'espai de característiques. La selecció d'una mesura representa un problema interessant en l'anàlisi de grups. Aquestes mesures es poden separar en dos classes segons la qualitat de les dades disponibles:

- mesures tipus distància: s'usaran quan les dades tinguin propietats quantitatives (ja siguin discretes o contínues).
- mesures tipus *matching*: s'usaran quan les dades tinguin components qualitius.

#### 4.2.1.1. Mesures tipus distància.

Suposem que les dades s'han obtingut amb  $n$  mostres. Cada mostra es representarà per un vector d'observacions  $\mathbf{X}' = (X_1, X_2, \dots, X_p)$  amb  $p$  variables. Usarem la notació  $\mathbf{X}'_i = (X_{i1}, X_{i2}, \dots, X_{ip})$  per indicar la mostra  $i$ -èsima.

Direm que  $d(\mathbf{X}_i, \mathbf{X}_j)$  és una mesura de distància entre dos punts si compleix que:

1.  $d(\mathbf{X}_i, \mathbf{X}_j) = 0 \Leftrightarrow \mathbf{X}_i = \mathbf{X}_j$ .
2.  $d(\mathbf{X}_i, \mathbf{X}_j) = d(\mathbf{X}_j, \mathbf{X}_i)$
3.  $d(\mathbf{X}_i, \mathbf{X}_k) \leq d(\mathbf{X}_i, \mathbf{X}_j) + d(\mathbf{X}_j, \mathbf{X}_k)$

Moltes mesures de distància són casos especials de la mètrica Minkowski definida com

$$d_{ij} = \left\{ \sum_{k=1}^p |X_{ik} - X_{jk}|^r \right\}^{1/r}$$

on  $d_{ij}$  és la distància entre dos mostres  $i$  i  $j$ . Si utilitzem  $r = 2$ , tenim la coneguda distància Euclidiana:

$$d_{ij} = \left\{ \sum_{k=1}^p (X_{ik} - X_{jk})^2 \right\}^{1/2}$$

Si  $r = 1$ , llavors tenim

$$d_{ij} = \sum_{k=1}^p |X_{ik} - X_{jk}|$$

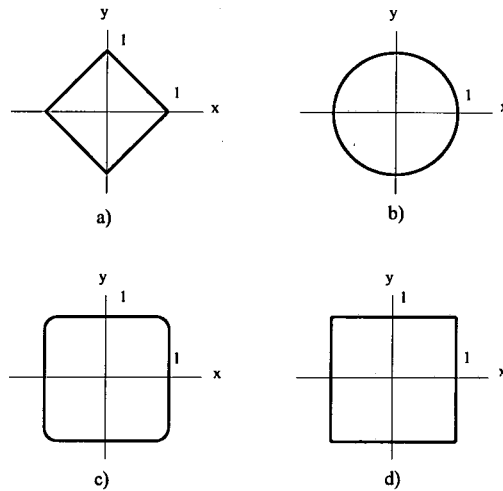
a la qual ens referim com mètrica absoluta [DIL84].

La figura 4.4 mostra diferents representacions de la funció distància per valors d' $r$  diferents en la mètrica Minkowski. Cal recordar que la distància Euclidiana no és invariant a canvis d'escala, per la qual cosa l'ús de dades directes obtingudes a partir d'un sistema físic s'han de mirar amb escepticisme quan es calculin distàncies entre mostres, ja que la distància pot estar distorsionada. Cal normalitzar aquestes dades abans de calcular la distància entre elles.

Una altra mesura legítima de distància entre dos punts ve donada per

$$d(\mathbf{X}_i, \mathbf{X}_j) = (\mathbf{X}_i - \mathbf{X}_j)' \mathbf{S}^{-1} (\mathbf{X}_i - \mathbf{X}_j)$$

on  $\mathbf{S}$  és una matriu simètrica invertible definida positiva i  $\mathbf{X}_i$  i  $\mathbf{X}_j$  són els respectius vectors d'observacions de les mostres  $i$  i  $j$ .



**Figura 4.4.** Representació del tots els punts a distància 1 de l'origen per a)  $r = 1$ ; b)  $r = 2$ ; c)  $r > 2$ ; d)  $r = \infty$ .

#### 4.2.1.2. Mesures tipus *matching*.

Les mesures tipus *matching* es coneixen millor com coeficients d'associació. Seran apropiades quan les dades estan escalades nominalment i s'apliquen quan les observacions prenen generalment els valors 0 o 1. Les mesures de semblança estan basades en el raonament que dues mostres són semblants quan comparteixen certs atributs.

#### 4.2.2. Tècniques d'agrupament.

Després d'haver escollit una mesura de semblança apropiada, el següent pas és seleccionar un tipus particular d'algorisme. Al llarg del temps, s'han proposat una gran quantitat d'algorismes que, principalment, es poden dividir en dos grups:

- tècniques jeràrquiques: Es van creant els grups en diferents etapes fins a obtenir un únic grup. El nombre de grups total s'assigna al final de l'agrupament.
- tècniques de particionament: es formen grups optimitzant algun criteri específic. El nombre de grups s'assigna abans de l'agrupament.

Existeix un altre enfocament per a la classificació o per buscar tipologies de mostres (objectes o individus): són els mètodes gràfics. Aquests mètodes són un mitjà extremadament útil i flexible per explicar, interpretar i analitzar dades a través de punts, línies, àrees i altres formes geomètriques o figures. Els mètodes gràfics més coneguts són els *glyphs*, els *metroglyphs*, sèries de Fourier i cares de Chernoff [CHE73]. La diferència entre aquests mètodes i les altres dues tècniques, jerarquia i particionament, està en que no s'usa cap algorisme computacional formal per generar els grups [DIL84].



#### 4.2.2.1. Tècniques jeràrquiques.

Els algorismes que pertanyen a aquestes tècniques duen a terme fusions o divisions de les dades. Una de les principals característiques que distingeix les tècniques jeràrquiques d'altres algorismes d'agrupament és que la col·locació d'una mostra en un grup és irrevocable, és a dir, una vegada una mostra s'afegeix a un grup, aquesta no s'elimina mai i es no es recol·loca amb d'altres mostres pertanyents a un altre grup. Els mètodes *aglomeratius* actuen formant una sèrie de fusions de les  $n$  mostres en grups. Els mètodes *divisius* parteixen el conjunt de les  $n$  mostres en subdivisions cada cop més fines. Els mètodes aglomeratius finalitzen el seu procés amb un únic grup al qual pertanyen totes les mostres, mentre que els mètodes divisius hauran dividit al final les dades de manera que cada mostra formarà el seu propi grup. En qualsevol cas, l'objectiu obvi és saber on aturar el procés de divisió o aglomeració per obtenir el nombre òptim de grups. Per ajudar a aquest procés, tant dels mètodes aglomeratius com dels divisius, la sortida és típicament recollida en un *dendograma*, el qual és un diagrama en forma d'arbre bidimensional il·lustrant les fusions o particions que s'han efectuat a cada nivell successiu.

- mètodes aglomeratius. Tots aquests mètodes actuen de forma semblant: Cada mostra comença en el seu propi grup. En el següent nivell es fusionen les mostres (grups) més properes, en el tercer nivell poden succeir dues coses: una nova mostra s'afegeix al grup que conté les dues mostres anteriors o bé es forma un altre grup de dues mostres. Aquesta decisió correspon a algun criteri d'assignació. El procés continua de manera semblant fins que es forma un grup que conté les  $n$  mostres. La diferència entre els diversos mètodes aglomeratius està en les alternatives definicions de distància o semblança usades en el criteri d'assignació.
- mètodes divisius: aquests mètodes comencen dividint totes les mostres en dos grups. El problema immediat és com realitzar la primera divisió. Amb  $n$  mostres, hi ha  $2^n - 1$  possibles maneres de formar subconjunts. Fins i tot, amb les actuals tecnologies dels computadors, el nombre de subconjunts a considerar s'ha de restringir en la majoria de les aplicacions reals. Una vegada es fa la primera divisió, les mostres es mouen d'un grup cap a un altre grup nou i es continuen fent divisions més fines sobre els grups ja formats. El que diferencia els diversos mètodes divisius és, en primer lloc, com es realitza la primera divisió i en segon lloc, com es subdivideixen els grups formats anteriorment.

#### 4.2.2.2. Tècniques de particionament.

A diferència de les tècniques jeràrquiques d'agrupament, els mètodes que realitzen una partició de les dades no requereixen que la col·locació d'una mostra en un grup sigui irrevocable. Això vol dir que les mostres es poden reassignar a un altre grup si la seva assignació inicial era inapropiada. Aquestes tècniques agrupen les dades basant-se en l'optimització d'un criteri formal i predefinit. En l'ús d'aquestes tècniques s'assumeix, habitualment, que el nombre de grups al final del procés és conegut i especificat prèviament, tot i que alguns mètodes permeten variar el nombre de grups al llarg de l'anàlisi. Les tècniques de particionament difereixen entre elles en 1) com s'inicien els

grups, 2) com s'assignen les mostres als grups, 3) com es reassignen a altres grups les mostres agrupades en passos anteriors.

### 4.3. Algorisme d'agrupament.

En aquest apartat es busca un algorisme d'agrupament que s'adapti bé a les nostres necessitats i aquest serà el que usarem per agrupar totes les mostres d'entrada. En qualsevol aplicació en la qual es requereixi agrupar unes mostres, la feina no s'acaba aquí, sinó que després d'haver generat els grups, es voldrà usar aquesta informació per continuar agrupant més mostres o, el que és el mateix, assignar una nova mostra d'entrada a un grup ja confeccionat amb anterioritat. Per aquest motiu els algorismes tenen dues fases: *aprenentatge* i *treball*. La fase d'aprenentatge serà costosa en temps ja que s'han de crear inicialment tots els grups. Aquesta fase serà supervisada i l'usuari decideix el nombre de grups a crear, sent *off-line*. Quan s'hagin creat tots els grups, la fase d'aprenentatge s'acaba i comença llavors la fase de treball on usarem la informació anterior per agrupar ràpidament qualsevol mostra a l'entrada. Aquesta fase pretenem que sigui *on-line* i per aconseguir-ho dissenyarem una arquitectura específica que permeti implementar l'algorisme de parametrització i classificació.

En el nostre cas en concret, tenim per cada imatge 32x32 elements de textura o mostres, que fan un total de 1024 mostres per imatge. Serà amb aquestes mostres que crearem grups en la primera fase d'aprenentatge. No tenim coneixement previ de a quin grup pertanyen les mostres ni tampoc tenim informació prèvia de quants grups volem crear. Per aquests motius usarem un algorisme d'agrupament jeràrquic aglomeratiu àmpliament estudiat i provat com és l'error de la suma de quadrats (E.S.S).

Els grups generats constituïran la informació que necessitarem per assignar noves mostres d'altres imatges, semblants a la imatge d'aprenentatge, a grups ja existents per segmentar, en una fase final, la imatge d'entrada.

#### 4.3.1. Mètode de l'error de la suma de quadrats de Ward. Fase d'aprenentatge.

[WAR63] proposa un mètode per a la formació de grups que està basat en la pèrdua d'informació que resulta de crear grups de mostres mesurant la suma total del quadrat de les desviacions de cada mostra respecte de la mitjana del grup al qual pertany. La regla d'assignació es basa en l'increment de l'error de la suma de quadrats induït per la combinació de cada parell de grups possible. Aquest valor, que s'anomena E.S.S. (*error sum of squares*), s'usa com a una funció objectiu.

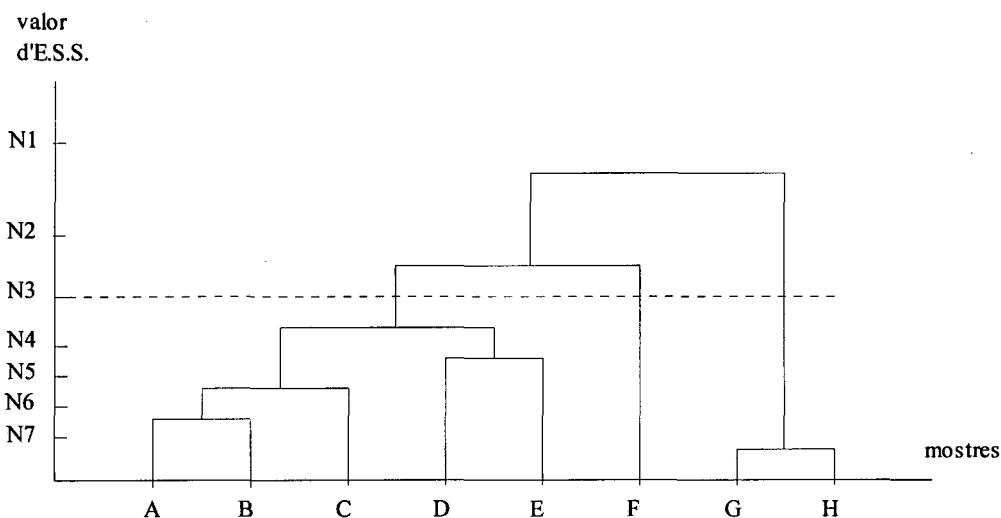
Aquest algorisme desenvolupat per Ward s'aplica juntament amb un procediment d'agrupament jeràrquic. El procés comença considerant  $K$  grups de mostres, amb una mostra per grup. El primer grup es forma seleccionant dos grups d'entre els  $K$  grups que, una vegada units, produeixi la menor alteració en el valor de la funció objectiu. Aleshores, aquest conjunt de  $K-1$  grups és reexaminat per a determinar els següents dos grups d'entre els  $K-1$  que, una vegada units, minimitzaran l'increment de la funció

objectiu. Els  $K$  grups inicials es van reduint sistemàticament de  $K$  a  $K-1$ , a  $K-2$  i fins a un grup en el transcurs d'aquest procediment d'agrupament jeràrquic. En cada nivell del procediment (des de  $K$  fins a 1) s'avalua el valor de la funció objectiu. Els canvis en els valors d'aquesta funció d'un nivell a un altre donaran una informació important per a la determinació del nombre de grups "naturals" per a les  $K$  mostres. Cal remarcar que el procediment jeràrquic pot no conduir a un valor òptim de la funció objectiu per a un nombre específic de grups. No obstant, es donarà una solució no òptima només en aquelles circumstàncies on l'agrupament "natural" no estigui ben definit.

Per determinar l'E.S.S. es calcula la següent expressió

$$E.S.S = \sum_{j=1}^k \left( \sum_{i=1}^{n_j} X_{ij}^2 - \frac{1}{n_j} \left( \sum_{i=1}^{n_j} X_{ij} \right)^2 \right) \quad (4.1)$$

on  $X_{ij}$  és el valor de la mostra  $i$ -èsima del grup  $j$ -èsim,  $k$  és el nombre total de grups en cada nivell ( $K, K-1, \dots, 1$ ), i  $n_j$  és el nombre de mostres en el grup  $j$ -èsim. Com que les mostres són vectors, quan ens referirem al valor d'E.S.S., ens referirem a la seva norma. El valor d'E.S.S. és zero en el primer nivell, ja que cada mostra pertany a un grup diferent. El valor d'E.S.S. serà màxim en el darrer nivell on totes les mostres pertanyen a un únic grup. El seguiment de l'agrupament es registra mitjançant un dendograma, el qual ens indica en cada nivell el nombre de grups formats, quines mostres pertanyen a cada grup i el valor d'E.S.S. per a cada nivell. En la figura 4.5 es veu la forma d'un dendograma.



**Figura 4.5.** Dendograma per al mètode E.S.S. de Ward.

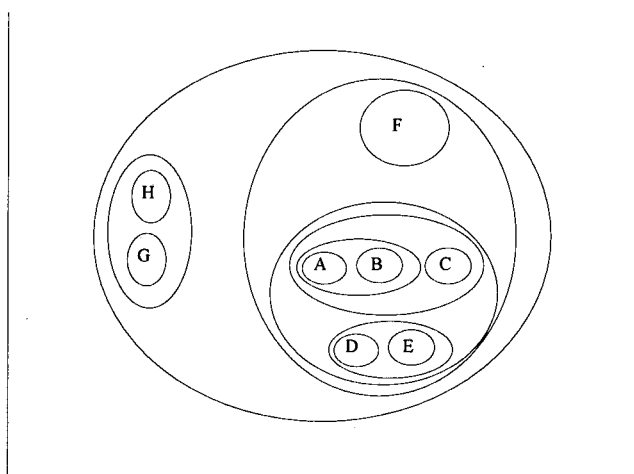
Com a exemple per veure la utilitat d'un dendograma, suposem que tenim les mostres que apareixen en l'eix de les abscisses (A, B, C, D, E, F, G i H). En l'instant inicial, hi ha tants grups com mostres i el valor de la funció E.S.S. és zero. En el nivell més baix del dendograma, cal agrupar dues mostres de manera que es produeixi la mínima alteració d'E.S.S. Després de fer totes les combinacions, s'escull l'agrupació de les mostres G i H que dona el valor N7. A continuació, es torna a buscar la combinació

que doni la menor alteració amb un grup menys i resulta que dona el valor  $N_6$  per a l'agrupació dels grups que contenen les mostres A i B. Es torna a aplicar el mateix algorisme però sobre un grup menys, doncs A i B ja estan en el mateix grup. S'acaba l'algorisme quan s'ajunten el grup que conté a les mostres A, B, C, D, E i F i el grup que conté les mostres G i H. Llavors, en el nivell superior de la jerarquia, tenim un únic grup contenint a totes les mostres i amb valor màxim d'E.S.S.  $N_1$ .

Només després de construir tot el dendograma estem en condicions d'escollir el nombre de grups que ens interessa. En aquest exemple, suposem que tallem l'arbre pel nivell  $N_3$ . Els grups que es generen són:

- Tall al nivell  $N_3$ . Tres grups: {A,B,C,D,E}, {F} i {G, H}.

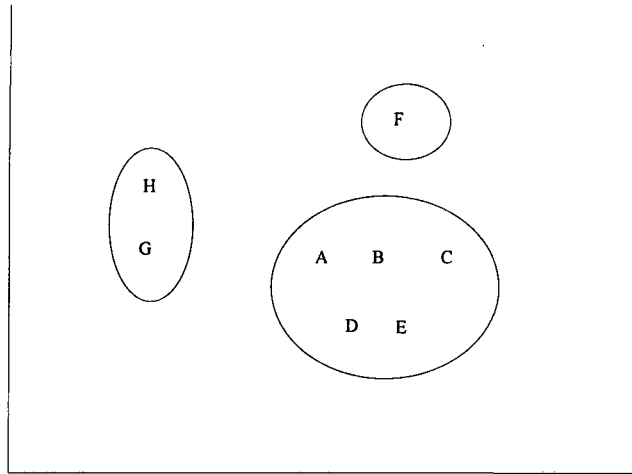
En la figura 4.6 es pot veure una possible representació de les mostres en l'espai de patrons i el seu agrupament jeràrquic. Els diferents nivells de la jerarquia es marquen mitjançant fronteres entre mostres.



**Figura 4.6.** Agrupament jeràrquic de mostres.

Dependrà de diversos factors el fet de tallar en un nivell o un altre de l'arbre. Si el resultat de l'agrupació no és representable visualment ens podem guiar, per exemple, amb la distància entre centres dels grups i les seves desviacions per decidir si ja tenim una bona agrupació. Però, si el resultat de l'agrupació es pot representar visualment, com en el nostre cas, la decisió del nivell de tall de l'arbre dependrà més de factors qualitius que no pas de factors quantitius com la distància entre centres. Al dir factor qualitatiu em refereixo a com queda sobre la imatge original l'agrupació en un nivell determinat. Això es veurà quan fem un exemple d'agrupament en el proper apartat.

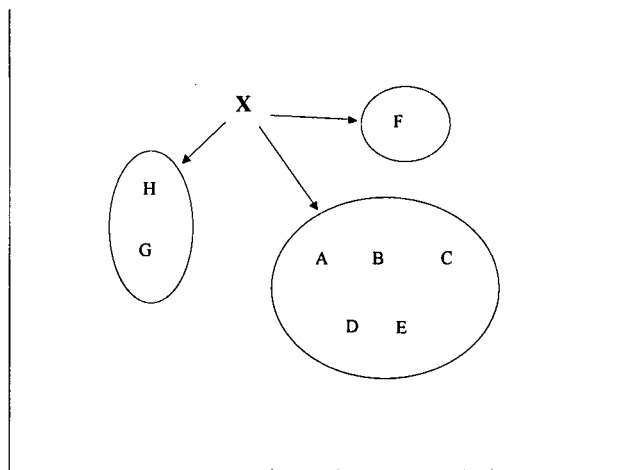
La fase d'aprenentatge s'acaba quan es decideix un nivell determinat en l'arbre que dona un agrupament que satisfà les necessitats que ens plantejava el problema. En aquest moment final de la fase d'aprenentatge, els grups escollits tenen un cert nombre de mostres, tenen un centre de masses i tenen una desviació respecte el seu centre. Seguint amb l'exemple anterior, suposem que escollim com a tall el nivell  $N_3$ . L'agrupament sobre l'espai de patrons és el de la figura 4.7, amb 3 grups.



**Figura 4.7.** Agrupament després d'escollir el nivell de tall N3 en el dendograma.

### 4.3.2. Fase de treball.

En aquesta segona fase de l'agrupament, es requereix que els grups ja estiguin formats. Així doncs, per cada nova mostra que entri al sistema se li assigna un grup dels formats anteriorment en la fase d'aprenentatge. Gràficament, això es pot veure en la figura 4.8.



**Figura 4.8.** Assignació d'una mostra a un grup.

Una mostra **X** que arriba al sistema es troba situada en un punt qualsevol de l'espai de patrons. Llavors, s'ha d'agrupar en algun dels grups que ja existeixen. No es permet la formació de nous grups a partir de noves mostres arribades en aquesta fase. S'ha d'escollir algun mecanisme que ens permeti fer aquesta assignació en el menor temps possible, donat que aquesta fase és *on-line*. Per aquest motiu, escollim com a regla d'assignació la mínima distància als centres de cada grup.

$$\text{Regla d'assignació} = \{ X \in i \mid \min\{d(X, C_i)\} \} \quad (4.2)$$

on  $i$  és el grup al qual s'assigna aquesta mostra i  $C_i$  és el centre de masses del grup  $i$ . Aquest tipus d'assignació parteix l'espai generant polígons de Voronoi.

Per a cada mostra d'entrada, calculem totes les distàncies entre la mostra i els centres de masses dels grups. S'assigna la mostra al grup on la distància sigui menor.

En qualsevol procés *on-line* és molt probable que les imatges vagin variant amb el temps. Per aquest motiu, si el canvi entre imatges no és molt bruscat sinó que es realitza suaument, existeix la possibilitat d'actualitzar el centre de masses dels grups a mida que es van agrupant les mostres. Per cada nova mostra que arriba al sistema s'assigna a un grup determinat mitjançant la regla (4.2) i a continuació es torna a calcular el centre de masses per aquest grup.

Per no perdre temps recalculant el centre de masses del grup amb un típic esquema com (4.3), aquest s'actualitza amb l'equació (4.4) on el temps de càlcul es redueix a dues operacions.

Per al càlcul del centre de masses d'un grup  $C_i$  usem:

$$C_i = \frac{\sum_{j=0}^N x_j}{N} \quad (4.3)$$

En aquest cas, s'han de sumar totes les  $N$  mostres del grup  $i$  i dividir-les pel nombre d'elements al grup. Aquest esquema és excessivament costós en temps, per la qual cosa volem trobar quin valor  $V$  cal afegir al centre de masses d'un grup amb  $N-1$  elements per obtenir el valor del centre de masses actualitzat per  $N$  elements. Així doncs, quan arribi una nova mostra a un grup, no caldrà tornar a recalculer el centre de masses, sinó que només sumant un valor ja estarà actualitzat el seu valor. Suposem que  $x_j$  és l'element  $j$ -èssim d'un grup i  $N$  és el nombre d'elements del grup després d'haver afegit la nova mostra arribada. Recordem que  $x_j$  és un vector.

$$\frac{\sum_{N-1} x_j}{N-1} + V = \frac{\sum_N x_j}{N}$$

Després de desenvolupar l'equació anterior, si aïllem la  $V$  ens queda que

$$V = \frac{x_N - \bar{x}_{N-1}}{N}$$

on  $x_N$  és la nova mostra,  $\bar{x}_{N-1}$  és el centre de masses del grup amb  $N-1$  elements i  $\bar{x}_N$  és el centre de masses pel grup amb  $N$  elements.

$$\bar{x}_N = \bar{x}_{N-1} + \frac{x_N - \bar{x}_{N-1}}{N} \quad (4.4)$$

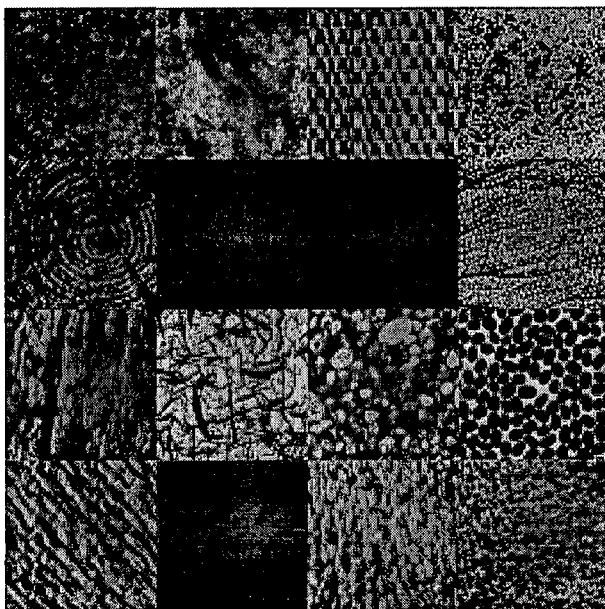
Així doncs, per actualitzar el centre de masses d'un grup amb una nova mostra, només cal sumar al centre anterior el terme  $V$  i obtenim el nou centre de masses amb un grup d' $N$  elements, segons l'equació (4.4).

### 4.3.3. Exemple d'agrupament jeràrquic.

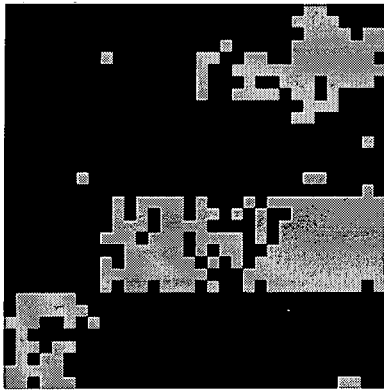
En el següent exemple, veurem com s'agrupen els elements de textura d'una imatge jeràrquicament. Es pren com a imatge d'entrada el següent mosaic d'imatges de textura de la figura 4.9.

Aquesta imatge s'ha dividit en 1024 (32x32) elements bàsics de textura per la qual cosa l'agrupament comença amb 1024 mostres i continuaria, consecutivament en la fase operativa *on-line*, amb 1024 grups per imatge, un grup per mostra. Es comença per aplicar l'algorisme E.S.S. de Ward, ajuntant dues mostres per formar un altre grup alterant mínimament l'error en la suma de quadrats. Aquest algorisme s'aplica iterativament fins arribar a obtenir un únic grup amb les 1024 mostres. Òbviament no es mostren tots els passos, sinó que ens fixarem només en els més significatius, que per altra banda, són els darrers passos en l'algorisme. Tampoc dibuixarem el dendograma per raons també òbvies d'espai.

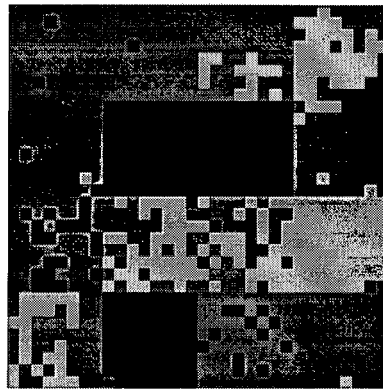
Al no poder dibuixar el dendograma, haurem de veure el resultat de l'agrupament visualment, és a dir, representant cada grup per un nivell de gris diferents i creant una imatge de sortida representativa d'aquest agrupament. Cada imatge (b), c), d) e), f), g), h) i i)) representa l'agrupament tallant un nivell més avall en el dendograma respectivament, començant en b) amb dos grups i acabant en i) amb 9 grups. Només observant les imatges de sortida podem fixar el nivell pel qual tallarem el dendograma. És amb aquesta tècnica qualitativa amb la qual escollirem el nombre de grups amb els quals agrupar la imatge d'entrada a partir de la seva informació textural.



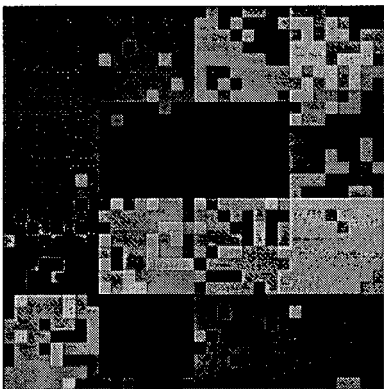
a) Imatge original [BRO66].



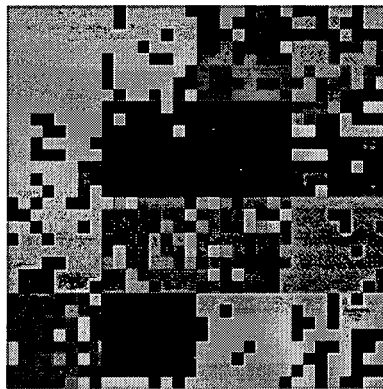
b) 2 grups,



c) 3 grups,



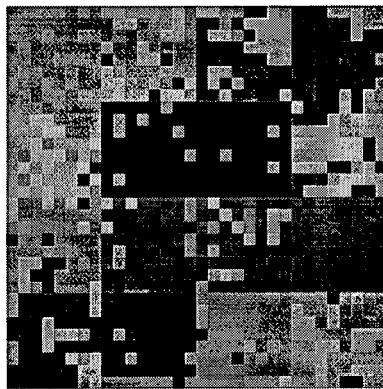
d) 4 grups,



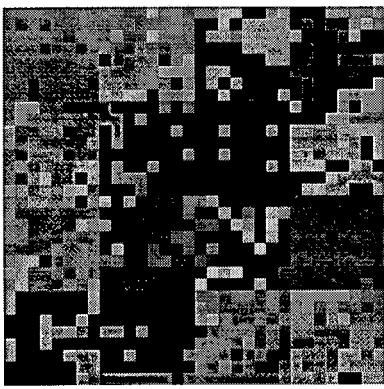
e) 5 grups,



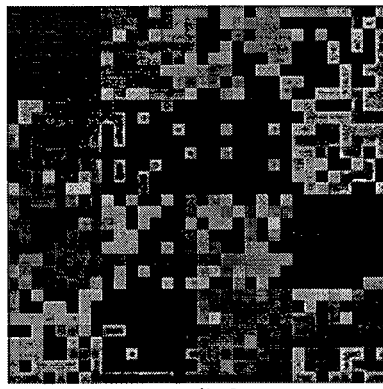
f) 6 grups,



g) 7 grups,



h) 8 grups,



i) 9 grups,

**Figura 4.9.** Agrupament jeràrquic de la imatge original a) en diferents grups.



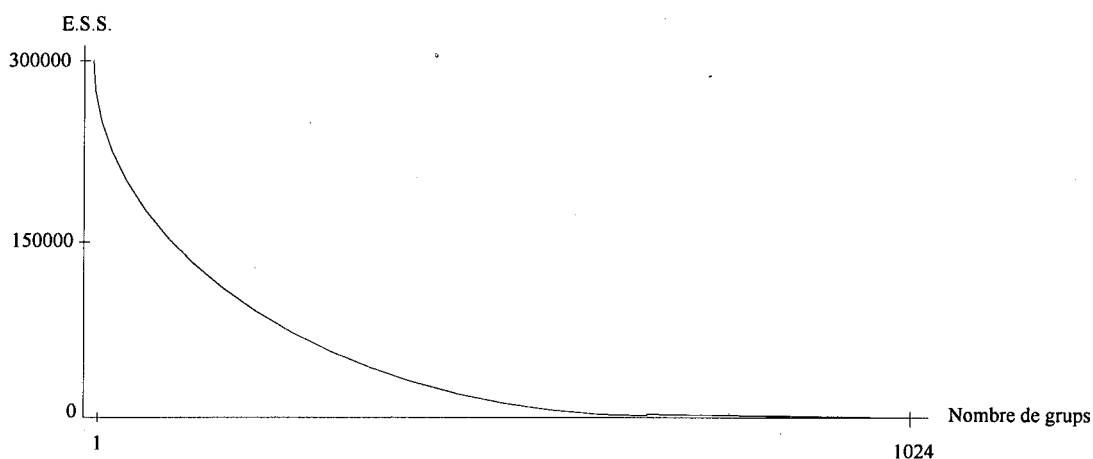
Després d'haver vist totes les imatges, perceptivament no es veu una segmentació correcta. Si féssim un estudi quantitatiu no ens podríem fer una idea de quant bona ha estat la classificació, perquè entre altres aspectes hi ha més textures diferents que grups de textura i la taxa d'error en la classificació no seria significativa. Per tant, és amb l'observació perceptiva que ens adonem a quin nivell tallar el dendograma. Si ens quedem amb pocs grups (2, 3, 4) hi ha compactació de regions de textura però poca segmentació. Si s'escull una classificació amb molts grups (8, 9) existeix una excessiva fragmentació de les regions de textura.

La taula 4.1. ens dóna quantitativament el nombre d'elements de textura per cada grups generat.

Nb de grups	Grup	Nb. d'elements per grup
2	0	768
	1	256
6	0	374
	1	120
	2	186
	3	99
	4	136
	5	109
9	0	166
	1	60
	2	96
	3	68
	4	115
	5	186
	6	136
	7	109
	8	88

**Taula 4.1.** Resultat de la classificació per a diferents grups.

Tallem al sisè nivell del dendograma i ens quedem amb els grups formats per la imatge f) amb 6 grups. Per a cada imatge, el nombre de grups que poden aparèixer serà diferent i per tant la fase d'aprenentatge serà supervisada. Això es pot veure també en el valor de l'E.S.S. per a cada grup. Si construïm la gràfica que ens relaciona el valor de l'error de la suma de quadrats en l'algorisme d'agrupament amb el nombre de grups que hi ha en cada moment, aquesta ens marcarà sobre quin nombre de grups és més significatiu tallar el dendograma. Això es pot veure en la figura 4.10. En l'eix de les abscisses tenim el nombre de grups que va creant l'algorisme d'agrupament i en l'eix de les ordenades es representa el valor de l'E.S.S. que ens generen el nombre de grups determinat.



**Figura 4.10.** Relació entre el valor de la norma d'E.S.S. i el nombre de grups creats en l'agrupament jeràrquic.

Per la pròpia natura de la funció en la qual es basa l'algorisme d'agrupament, la gràfica de relació entre el nombre de grups i el valor de l'E.S.S sempre té aproximadament la mateixa forma. Després de realitzar proves amb moltes imatges (més de 20 mosaics), en la figura es mostra una d'aquestes gràfiques. A mida que van quedant menys grups, la dispersió de les mostres que els formen és major, fins que el límit de dispersió arriba quan només hi ha un grup. Mirant purament la gràfica no es veu cap lloc decisiu per on tallar el dendograma. Per aquest motiu ens fixarem més en els resultats sobre les imatges que no en el valor que pren la funció E.S.S.

Segons el que s'ha vist en la secció anterior, després de la generació dels grups de textura ja es pot passar de la fase d'aprenentatge a la fase de treball, però perceptivament es veu que el resultat de l'agrupament no és del tot correcte, sigui quin sigui el nivell pel qual tallem el dendograma. Cal algun mecanisme que ens millori l'agrupament de la imatge d'entrada. Aquest mecanisme el veiem a continuació en la secció 4.4.

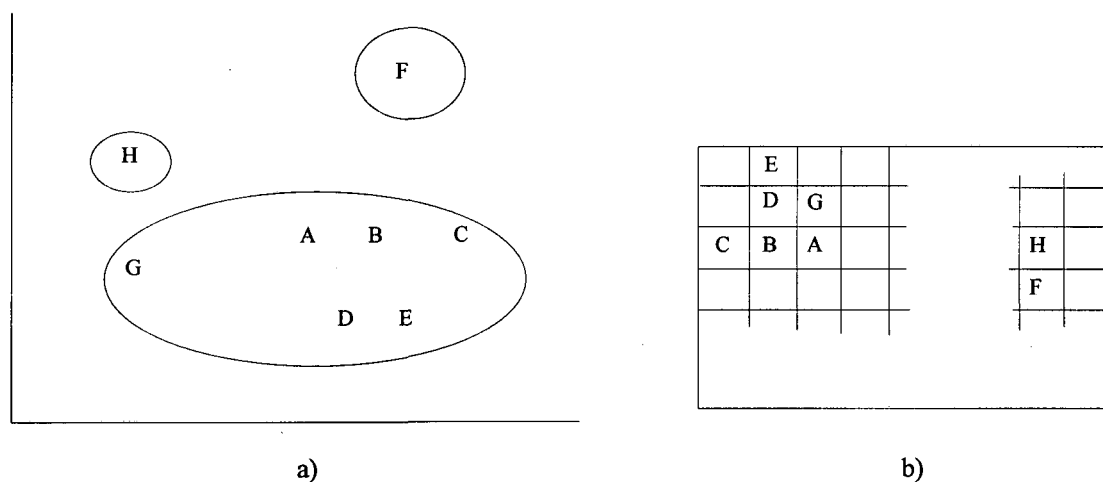
#### 4.4. Ús del funcional d'energia.

El problema de la segmentació és subdividir una imatge en regions on en cadascuna d'elles les propietats de l'escena visualitzada siguin relativament uniformes. El fet d'usar algorismes de *clustering* per agrupar punts en un espai  $\mathbb{R}^n$  de característiques de textura, no és una garantia suficient per assegurar que arribarem al final del procés de segmentació amb èxit. Diversos motius poden fer fracassar la correcta segmentació de regions uniformes en la imatge. Per millorar aquesta segmentació estudiarem el problema mitjançant un enfocament variacional [MUM89], [SHA92], [MAS92], [ALV93]. Aquest enfocament ve motivat en part per les errades ocasionals en els mètodes tradicionals, basats exclusivament en l'agrupament de punts formats a partir de característiques extrems de la imatge sense tenir en compte la seva disposició geomètrica en la imatge. La idea general és definir un funcional d'energia sobre un

conjunt de possibles segmentacions en termes de mesures penalitzadores que corresponen a diverses propietats desitjades en una correcta segmentació.

Intuïtivament, podem usar la figura 4.7 per entendre gràficament com actua el funcional d'energia. En la figura 4.7 s'ha vist el resultat de l'agrupament després d'haver fixat un nivell en el dendograma. En aquest exemple s'han creat tres grups amb la informació provinent de l'espai de característiques. Però pot succeir que algunes mostres, tot i minimitzar l'error de la suma de quadrats en l'algorisme, es classifiquin en algun grup erroni donat que no s'havien calculat correctament els seus paràmetres de textura. Per aquest motiu, és convenient afegir informació addicional. El funcional d'energia té en compte la posició dels elements de textura sobre la imatge i els recllassificarà si ho creu convenient. En la figura 4.11 podem veure aquest efecte.

La mostra G s'havia classificat en un grup juntament amb la mostra H per la seva proximitat en l'espai de característiques. Després d'aplicar el funcional d'energia sobre aquest espai i tenint en compte la posició de les mostres sobre la imatge, hi ha hagut una recllassificació de la mostra G cap a un altre grup no tant proper en l'espai de característiques però potser més proper sobre l'espai 2D del domini de la imatge i que fa que el funcional disminueixi el seu valor. Podria succeir també que tot i que les mostres F i H està molt properes en l'espai de la imatge no s'arribessin a agrupar sota una mateixa classe donada la gran diferència sobre l'espai de característiques ja que la possible agrupació no disminuiria el valor del funcional. Després d'aquesta reflexió sobre el cas hipotètic de la figura, es veu que realment poden succeir els dos casos extrems: agrupació de mostres llunyanes en l'espai de característiques o agrupació de mostres llunyanes en l'espai de la imatge. El que compta és minimitzar el funcional d'energia.



**Figura 4.11.** Resultat de l'agrupament després d'aplicar el funcional d'energia. a) classificació en l'espai de característiques; b) posició sobre la imatge de les mostres.

En particular, el funcional que hem escollit, presentat a [GRA97c], és el següent

$$J(\bar{x}, F) = \sum_{k=1}^M \sum_{j \in C_k} \|\bar{x}_j - \bar{x}_k\|^2 + \lambda |F| \quad (4.5)$$

on

- $\bar{x}$  és el conjunt de totes les mostres de l'espai de característiques
- $C_k$  és el nombre de mostres del grup  $k$
- $M$  és el nombre de grups
- $\lambda$  és un pes
- $F$  és la unió de les fronteres entre regions de la imatge
- $|F|$  és la longitud d' $F$
- $\bar{x}_j$  és una mostra del grup  $k$  que es defineix com  $\bar{x}_j = (x_j^1, x_j^2, x_j^3, x_j^4, x_j^5)$  on cada component del vector és una característica de textura. Per referir-nos a una mostra (o element) eliminarem l'indicatiu de vector per senzillesa.
- $\bar{x}_k$  és la mitjana del grup  $k$  calculada de la següent forma

$$\bar{x}_k = \frac{1}{|C_k|} \sum_{j \in C_k} x_j \quad \text{on } |C_k| \text{ és el nombre d'elements en el grup } k$$

El funcional treballa a partir de la informació que obté de dos espais totalment diferenciats. Aquests espais són l'espai de característiques de textura i l'espai de la imatge bidimensional. El primer terme del funcional mesura el grau de dispersió dels elements dins de cada grup al qual pertanyen. El segon terme quantifica les fronteres que existeixen entre regions de la imatge que pertanyen a grups diferents tal i com s'han classificat amb anterioritat. La suma ponderada d'aquests dos valors és el valor del funcional d'energia. Aquesta serà la funció de cost objectiu que s'ha de minimitzar per tal de trobar la mínima dispersió dels elements dins els grups juntament amb el fet de trobar les mínimes fronteres entre regions de la imatge. Aquesta mesura també indica el grau de dispersió entre regions de la imatge.

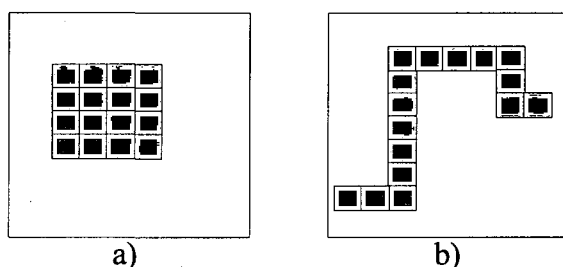
Aquests dos termes són oposats entre ells. En el primer terme, si la dispersió ha de ser mínima aleshores cada element ha de formar un grup. Amb això, tenim que hi ha tants grups com elements i aquest terme que mesura la dispersió en l'espai de característiques pren el seu valor mínim (zero). Per altra banda, si cada element forma un grup, tenim que les fronteres entre elements de la imatge que pertanyen a grups diferents seran màximes. Existirà una frontera entre qualsevol element de la imatge amb qualsevol altre veí seu.

Suposem la situació totalment contrària. Volem que el segon terme prengui el valor mínim. Això implica que no han d'existir fronteres entre regions de la imatge (valor zero). Per aconseguir-ho, cal que tots els elements pertanyin al mateix grup de textura.

Pel que fa al primer terme del funcional, si tots els elements pertanyen al mateix grup tenim que la dispersió amb aquest únic grup és màxima.

La solució al problema està en algun punt intermig entre aquestes dues situacions anteriors oposades. Cal arribar a un compromís entre aconseguir la mínima dispersió dels elements en l'espai de característiques (màxima compactació dels grups) i també fer mínimes les fronteres entre regions de textura de la imatge formades per aquests elements (màxima compactació de les regions sobre la imatge).

En la figura 4.12 es pot veure què entenem per màxima compactació en les regions de textura sobre la imatge.



**Figura 4.12.** Compactació de regions en la imatge: a) àrea amb mínimes fronteres; b) mateixa àrea amb màximes fronteres.

En la figura 4.12, la regió de textura semblant en l'àrea a) té la disposició geomètrica que minimitza les seves fronteres, per l'altra banda, la regió b) té la mateixa àrea que la regió en a) però la seva disposició sobre la imatge fa que el nombre de fronteres que presenta sigui màxim. Per tant, el funcional penalitzarà aquests tipus de regions amb gran nombre de fronteres.

Després de definir el funcional d'energia i veure'n els requeriments i objectius, descrivim a continuació l'algorisme usat per arribar al parell  $(\bar{x}, F)$  que minimitza el funcional.

#### **Algorisme per minimitzar el funcional d'energia.**

L'algorisme utilitzat per minimitzar el funcional és el següent:

- 1) Inicialització: cada tèxel està assignat prèviament a un grup,
- 2) escollir aleatòriament un tèxel  $t_{xy}$  de la imatge,
- 3) comprovar si algun veí al tèxel  $t_{xy}$  sobre la imatge pertany a una classe diferent a la seva en l'espai de característiques. Si no existeix cap veí que compleixi això tornar al pas 1,
- 4) afegir temporalment a la classe del veí el tèxel  $t_{xy}$  escollit i calcular el funcional  $J(\bar{x}, F)$ ,
- 5) si el valor del funcional s'ha reduït, llavors incorporar definitivament el tèxel  $t_{xy}$  a la classe del seu veí, actualitzar les mitjanes i el nombre d'elements dels grups i tornar al pas 1,

- 6) si després de provar amb tots els tèxels no s'ha reduït el valor del funcional, estarem sobre un mínim i acaba l'algorisme.

Cap algorisme troba el mínim absolut a aquest problema, ni teòrica ni pràcticament, a causa de la no-linealitat del problema i també per l'existència de mínims locals, però això no representa un problema excessivament preocupant perquè hi ha valors de  $(\vec{x}, F)$  que són suficientment vàlids.

De fet, els enfocaments variacionals estan en un grau de desenvolupament molt precari, tot i que hi ha diversos autors que treballen en aquest tema (mencionats anteriorment). Algun autor, com ara [LEE92], presenta un model computacional que segmenta imatges basant-se en propietats de textura extretes amb *wavelets* i minimització de funcionals d'energia. Tot i haver algun treball en el tema, aquesta poca evolució en el tema fa que existeixin moltes qüestions obertes com ara:

- com es varia el pes  $\lambda$  en el context del problema de visió?
- hi ha algun algorisme pràctic per minimitzar el funcional?
- com es comporta en presència de soroll?
- la formulació és correcta per solucionar problemes de visió?

L'optimització de funcionals d'energia dóna, des del punt de vista conceptual, una eina atractiva per al processat d'imatges. Per tant no sorprèn que hagi estat el punt de partida per a molts investigadors [HAA94], [HER91]. L'estudi del funcional d'energia presentat en aquest capítol és una ampliació i variació d'un dels models probabilístics més estudiats actualment relacionat amb els models d'energia: és el model d'Ising. Aquest model ve directament de la mecànica estadística aplicada a la física i s'usa per a descriure el comportament de cristalls bidimensionals d'àtoms de ferro subjectes a un camp magnètic extern. Aquest model aplicat a la visió per computador serveix per a modelitzar imatges que estan formades per un conjunt de taques blanques sobre un fons negre (o a l'inrevés) i sobre la qual es busca descriure les taques blanques (conjunt  $S$ ) mitjançant una imatge binària. El model que s'usa és senzill: volem  $S$  tal que el seu perímetre sigui el més petit possible, com menor sigui el perímetre, més probable serà que  $S$  estigui dins el model. Per una altra banda, se suposa que ni el fons de la imatge ni els objectes tenen un nivell de gris uniforme, per la qual cosa no serveix un esquema de segmentació per llindar per obtenir les imatges binàries buscades, donat que apareixerien forats sobre els objectes i soroll sobre el fons; precisament el model intenta evitar això. Es crea un funcional d'energia que intenta agrupar regions per semblança amb el seu nivell de gris però amb perímetre mínim.

Existeixen molt models probabilístics diferents per estudiar problemes relacionats amb funcionals d'energia i encara hi ha un gran camp a recórrer amb aquestes tècniques.

En concret, l'ús del funcional en la fase d'aprenentatge és delicat. Suposem el següent exemple: en una certa imatge hi ha una gran àrea que correspon a un camp, en mig del camp apareix una casa amb una textura diferent. El procés d'agrupament detecta aquesta textura diferent i genera una nova classe amb aquests elements corresponent a la casa. El fet d'aplicar el funcional després d'aquest agrupament pot provocar que els

elements de textura corresponents a la casa siguin inclòsos en el grup de textura que caracteritza el camp, amb la conseqüent desaparició d'aquest grup de textura no desitjat, afectant a la fase de treball posterior, però modificant lleugerament el centre de massa del grup que caracteritza aquesta textura.

Però suposem que els elements intrusos dins el camp fossin elements de soroll que s'havien introduït provocant la generació d'un nou grup de textura totalment erroni. El funcional actua de filtre i elimina aquests elements no desitjats.

El fet que succeeixi una situació o una altra dependrà de quant fort, marcat i separat sigui el grup de textura intrús, provocant que el funcional el fusioni o no amb un altre grup de textura veí sobre la imatge. La supervisió també pot forçar aquesta absorció en funció del valor del pes  $\lambda$ , però amb efectes colaterals no desitjats que repercuteixen sobre altres grups de textura.

A continuació es mostra el resultat d'aplicar un mètode per millorar l'agrupament que, en concret, és l'ús del funcional d'energia definit anteriorment per segmentar les regions de textura semblant sobre la imatge.

#### **4.5. Millora del procés d'agrupament amb la utilització del funcional d'energia.**

L'ús del funcional d'energia millora àmpliament els resultats obtinguts en la classificació a partir de mètodes convencionals en mostres en que es disposa de diferents tipus d'informació.

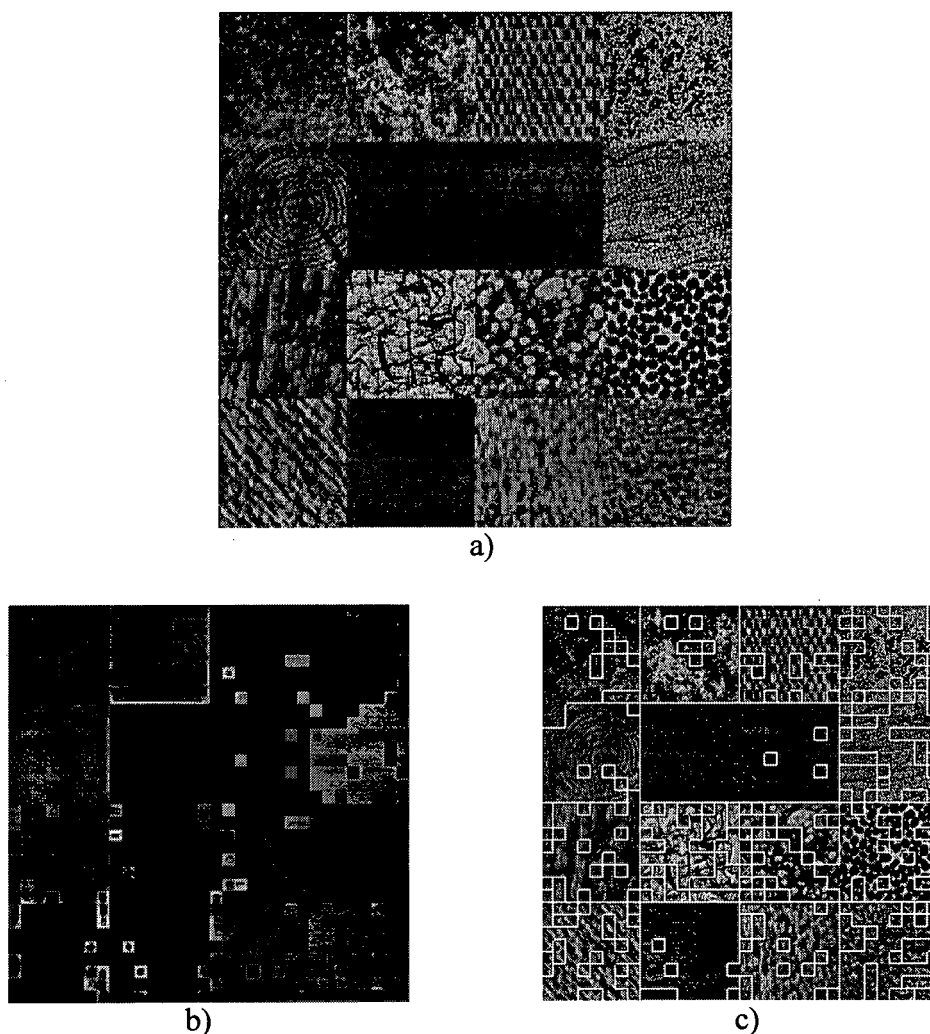
El funcional d'energia, representat per l'equació (4.5), és un classificador per ell mateix i sempre s'haurà de minimitzar el seu valor. Però en aquest cas, l'usarem en un pas posterior a la classificació inicial realitzada per (4.1). Dit d'una altra manera, comencem a partir d'unes condicions inicials i el funcional ens servirà per polir moltes de les incorreccions o males classificacions que el classificador convencional no ha pogut resoldre només amb les dades de l'espai de característiques de textura. La raó de no usar el funcional com a classificador exclusiu és ben senzilla: les dades inicials, amb les quals comencem a minimitzar el funcional, fan que aquest arribi a un mínim local que no ens interessa. Quan això succeeix no es pot continuar el procés de minimització: hem de sortir d'aquest mínim local i començar en un altre lloc. Més endavant, en la figura 4.16, es fa una breu discussió sobre aquest aspecte.

Per resoldre aquest problema i iniciar el procés de minimització en un altre punt, les dades es classifiquen amb el classificador segons l'equació (4.1).

Els passos a seguir són els següents:

1. S'inicia l'agrupament amb el classificador segons (4.1).
2. Es determina, supervisadament, un nivell de l'arbre pel qual tallar creant  $k$  grups.
3. A partir de les classes generades, s'aplica el funcional i es minimitza el seu valor partint de les dades classificades en el punt 2.
4. Quan s'arriba a un mínim local que ens interessi, s'acaba el procés global de classificació, segmentant i fixant les fronteres entre textures de classes diferents.

En aquesta secció veurem la millora en la segmentació per l'ús del funcional d'energia.



**Figura 4.13.** Classificació i en la segmentació d'una imatge abans d'usar un funcional d'energia: a) Imatge original, b) classificació per *clustering*, c) segmentació.

Sobre una imatge d'entrada, fig. 4.13 a), s'extreuen les característiques paramètriques de textura i es classifiquen mitjançant l'algorisme explicat en l'apartat 4.3 anterior. La sortida de la classificació són les imatges b) i c) en la figura 4.13. La imatge b) és la mateixa que es mostrava en la figura 4.9 f) on, recordant-ho, es va tallar el



dendograma al sisè nivell generant 6 grups de textura diferents. La imatge en la figura 4.13 c) és el resultat de la segmentació mostrant les fronteres entre textures superposades sobre la imatge original, per la seva millor interpretació.

Tal com es pot apreciar en c), s'han produït moltes males classificacions a causa dels propis valors dels tèxels que s'han calculat en l'extractor de característiques de textura. Serà sobre aquestes dades que apliquem la nova classificació amb el funcional (4.5).

Abans d'aplicar l'algorisme de minimització del funcional sobre les dades que produeixen la segmentació mostrada en la imatge c), cal fer algunes consideracions sobre el funcional que, en definitiva, ens conduiran a trobar un valor del pes  $\lambda$  que faci que el procés de minimització arribi a algun mínim local que sigui del nostre interès.

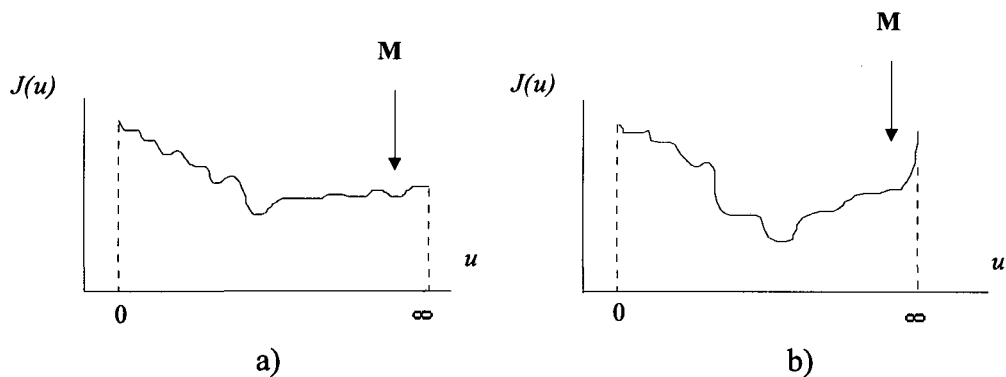
El funcional d'energia definit anteriorment,  $J(\vec{x}, F)$ , té dos límits superiors que anomenarem  $J_0$  i  $J_\infty$ . Existiran dues condicions que produiran aquests valors:

1. Primer terme del funcional igual a zero. Límit  $J_\infty$ . Aquesta situació es produirà quan existeixin tantes mostres com classes. Cada centre de la classe és la mateixa mostra per la qual cosa, la desviació de la mostra respecte al centre de la classe és zero. Però al mateix temps que cada mostra és una classe, en el domini imatge cada tèxel té una frontera amb els seus veïns, per la qual cosa aquest terme serà màxim.
2. Segon terme del funcional igual a zero. Límit  $J_0$ . Cap tèxel en la imatge té fronteres amb els seus veïns, per tant la suma de fronteres és zero. Però en el domini 5-dimensional de l'espai de característiques això implica que totes les mostres pertanyen a la mateixa classe. Quan això succeeix, hi haurà un únic centre de la classe i la desviació de totes les mostres respecte a aquest centre serà màxima.

Qualsevol altra situació, com ara que existeixi més d'una classe o que algun tèxel tingui veïns que pertanyen a la seva mateixa classe, estarà entremig dels dos casos patològics anteriors.

Serà a partir d'un exemple on veurem com evolucionen tant el conjunt de mostres dins els seus grups així com les fronteres entre aquests grups. L'únic grau de llibertat que disposem és un pes  $\lambda$ , però tampoc no és gaire significatiu. El que realment condiciona el comportament del funcional és la inicialització dels grups.

El pes  $\lambda$  ens marcarà la diferència entre un extrem i l'altre del funcional. Els dos termes han d'estar al mateix nivell o el que és el mateix, els dos límits han de valer el mateix donat que són màxims del funcional. En la següent figura veiem gràficament aquest concepte. Per simplicitat en la representació, suposem que només hi ha una variable ( $u$ ).



**Figura 4.14.** Importància del pes  $\lambda$ . a) Límits desiguals; b) límits normalitzats.

Suposem que el pes  $\lambda$  val 1. Això vol dir que els dos termes del funcional no estan normalitzats i seria el cas corresponent a la figura 4.14.a). Si s'inicia el procés de minimització del funcional  $J$  pot succeir que ens quedem aturats al mínim local **M**.

Suposem ara que hem trobat el valor del pes  $\lambda$  que fa que els dos termes estiguin normalitzats. Això vol dir que els extrems del funcional valen el mateix, fig. 4.14 b), i això equival a estirar (com si es tractés d'una corda) tota la funció  $J$  amb la qual cosa també s'han estirat o accentuat els valors màxims i mínims. En una situació inicial exactament igual a l'anterior, el que abans era un mínim local **M** ara ja no és tal mínim i podem continuar minimitzant  $J$  fins a trobar un valor de mínim local molt millor que el valor que s'havia aconseguit sense normalitzar els extrems.

Per poder començar el procés de minimització del funcional, cal escollir el valor del pes  $\lambda$ , que sortirà de la normalització per a la imatge inicial, fig. 4.13.c), dels extrems del funcional. Aquest valor serà

$$J_0 \lambda = J_\infty$$

$$1920 \lambda = 259488$$

$$\lambda = 259488/1920$$

$$\lambda = 135.15$$

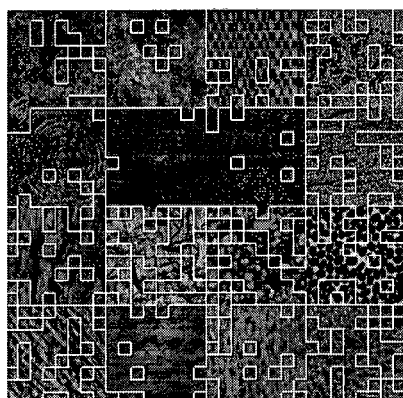
S'ha fixat aquest valor per la  $\lambda$  i hem iniciat el procés de minimització del funcional d'energia amb els valors inicials de la segmentació tal com es mostra en la fig. 4.13. c). S'han recollit algunes segmentacions intermèdies per veure l'evolució de la classificació en els diferents passos iteratius implementats en l'execució de l'algorisme. Els resultats es mostren en les imatges de la figura 4.15 a), b), c), d), e) i f). En cada imatge es veu com existeixen menys fronteres entre regions i la segmentació es va definint poc a poc. El resultat final és la imatge f) on el procés de minimització finalitza donat que no es pot trobar cap més valor inferior per al funcional  $J$ .

En la taula 4.2 s'ha recollit numèricament aquest procés de minimització. En la primera i segona files es mostra el valor dels extrems del funcional per a una imatge sense classificar. Això vol dir que la desviació de totes les mostres respecte al seu centre de masses per a una única classe és el valor  $J_0$ ; mentre que el valor de les fronteres per la mateixa imatge sense classificar però amb una mostra per grup és  $J_\infty$ . En la columna que fa referència al segon terme (valor de les fronteres) apareixen dos números: el primer és el valor total del terme ( $\lambda|F|$ ), mentre que el segon valor entre parèntesi és el número real de fronteres entre regions diferents.

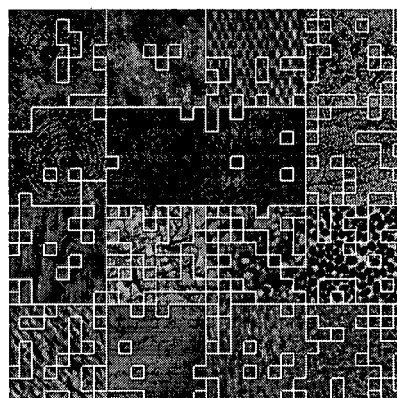
Segmentació	Primer terme $\bar{x}$	Segon terme $ F $	Funcional $J(\bar{x}, F)$
$J_0$	259.488	0	259.488
$J_\infty$	0	259.488 (1920)	259.488
Fig. 4.15 a)	89.552	105.165 (779)	194.717
b)	92.155	98.415 (729)	190.570
c)	94.894	91.260 (676)	186.154
d)	97.215	82.350 (610)	179.565
e)	100.254	69.930 (518)	170.184
f)	109.632	48.060 (356)	157.692

**Taula 4.2.** Minimització del funcional d'energia segons el pes  $\lambda$ .

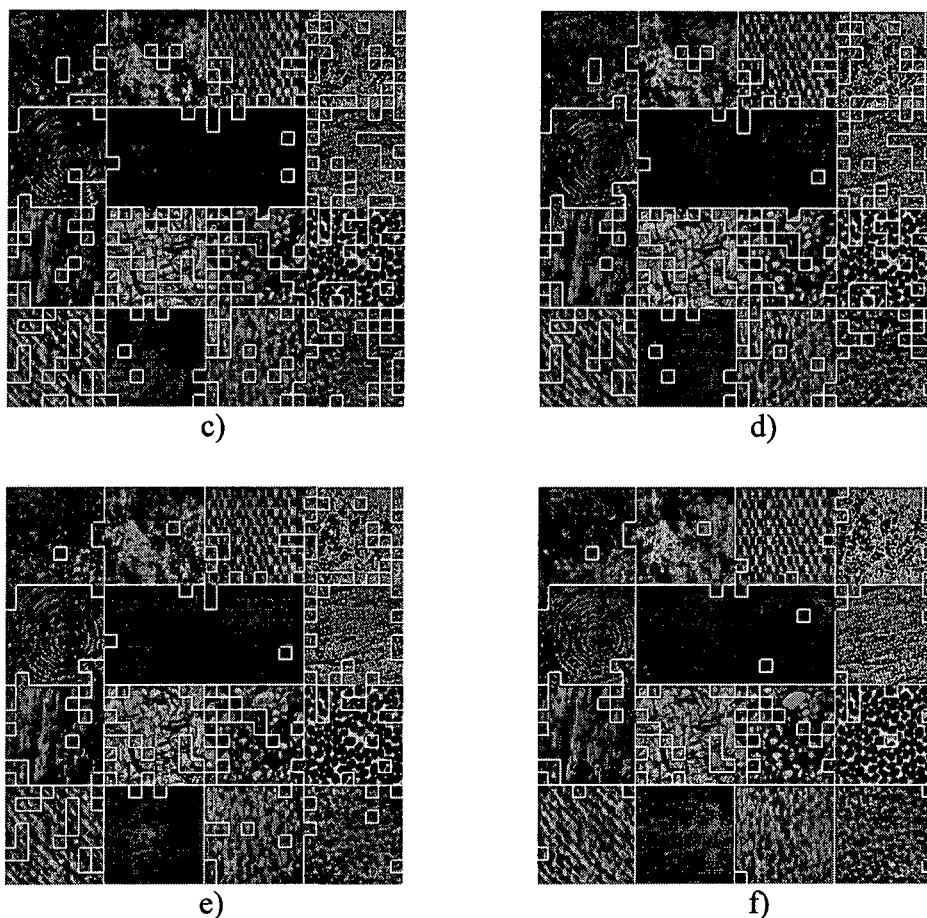
A continuació es poden veure el resultat de la minimització i alguns dels passos vàlids en aquest procés. Les imatges que es mostren en la figura 4.15 corresponen a iteracions ordenades de forma consecutiva en el temps però només hem mostrat algunes d'aquelles en les que el funcional d'energia tenia un valor inferior respecte a la imatge anterior.



a)



b)



**Figura 4.15.** Millora en la classificació i en la segmentació d'una imatge després d'usar un funcional d'energia iterativament. La imatge f) s'obté després de trobar un mínim local que s'accepta com a vàlid.

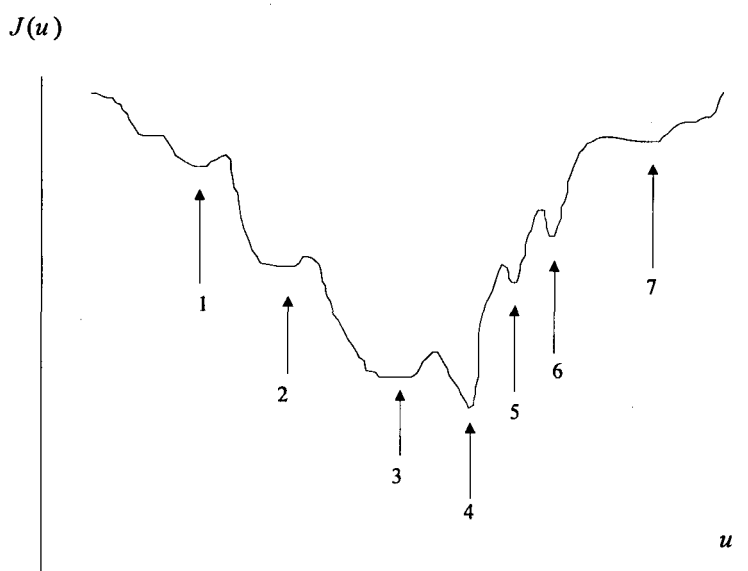
Existeix un compromís entre els dos termes mentre hi ha minimització del funcional. Donat que l'algorisme de classificació en (4.1) també tenia en compte les desviacions respecte a les mitjanes dels grups, l'estat final de la classificació (estat inicial per al funcional) tenia dispersió mínima. Per aquest motiu seria molt casual que aquest valor augmentés quan s'estan canviant mostres d'un grup a un altre. El primer terme augmenta a mida que s'està minimitzant el funcional. En canvi, s'està reduint el nombre de fronteres entre regions de la imatge a la vegada que s'està reduint el valor del funcional. Això és possible perquè el pes  $\lambda$  ha servit per normalitzar els dos valors extrems del funcional i permet que tot i que un valor augmenti (la dispersió en l'espai de característiques), el valor de les fronteres i, en definitiva, el valor del funcional es decremanti.

Només es mostren els intents vàlids de minimitzar el funcional. A part d'aquests hi ha hagut molts intents no vàlids on, tot i disminuir el valor de les fronteres, la dispersió en l'espai de característiques era excessiu i no s'aconseguia disminuir el valor del funcional.

Per simplicitat en el dibuix, imaginem la representació del funcional  $J(u)$  amb una sola variable. La seva possible representació gràfica podria ser la corba de la figura

4.16. En la gràfica es representen diversos mínims locals que bé podrien correspondre amb els que s'han calculat per les classificacions anteriors. Si el procés de minimització s'inicia sobre el punt assenyalat com 6, no hi haurà cap possible minimització, doncs ja estem sobre un mínim local i mai podrem arribar a un mínim inferior a aquest. Tal com es va comentar en l'apartat anterior, no se sap si un mínim és millor que un altre i només els resultats obtinguts poden determinar si aquell mínim és adequat o no.

En la gràfica de la figura 4.16 no s'ha representat intencionadament cap mínim absolut perquè és impossible saber si estem sobre un d'aquests mínims.



**Figura 4.16.** Representació de  $J(u)$ , on s'aprecia l'efecte dels mínims locals representats per les fletxes.

Els mètodes del gradient, i d'altres tècniques d'optimització, poden veure restringida la seva potència pel fet de trobar mínims locals [HER91], tal com acabem de veure anteriorment. Una possible solució usada àmpliament per sortir dels mínims locals és introduir soroll aleatori. Aquest soroll permetrà canviar ocasionalment el sentit de baixada en els passos que recorrem per arribar a la minimització de la funció cost. Aquesta tècnica s'anomena recuit simulat (*simulated annealing*) [KIR83], per l'analogia que guarda amb el recuit real dut a terme en el laboratori pels físics experimentals o els metal·lúrgics: refredar una mostra lentament de manera que trobi la fase estable a baixa temperatura on es poden realitzar les mesures. Aquesta fase de màxima estabilitat correspon a la màxima energia del material. Aplicant aquest concepte en el cas de la minimització de funcions, com més baix sigui el valor de la *temperatura*, menor serà la fluctuació positiva permesa en el sentit contrari de la minimització. Així doncs, com millor es minimitzi la funció de cost, més difícil serà sortir dels mínims locals.

## 4.6. Comparació teòrica d'algorismes de textura.

La motivació per la comparació d'algorismes d'anàlisi de textura es fa evident en dos casos clars: primer, quan l'investigador desenvolupa un nou mètode necessita saber el grau de funcionament del seu algorisme en relació a d'altres semblants i segon, quan existeix un problema sobre una aplicació concreta es vol seleccionar l'algorisme que millor s'adeqüi a la necessitat. El nostre cas concret correspondria al primer objectiu i avaluarem el mètode presentat en aquesta tesi enfront d'algorismes d'anàlisi de textura àmpliament provats. Aquesta avaluació es realitzarà en termes de qualitat en els resultats i en termes de temps de càlcul per obtenir-los.

### 4.6.1. Estudis comparatius d'algorismes de textura.

A mida que s'han anat desenvolupant nous algorismes d'anàlisi de textures, també han aparegut estudis sobre comparacions entre mètodes. Per citar-ne alguns dels més coneguts, Weszka *et al* [WES76] usen quatre conjunts de característiques per a classificar mostres de terrenys: espectre de potència de Fourier (anells i falques), estadístics de nivells de gris de segon ordre (matrius de co-ocurrència amb quatre característiques i quatre orientacions), estadístics de diferència de nivells de gris de primer ordre (mitjana i contrast en 8 distàncies) i estadístics de *run length* de nivells de gris (4 característiques en 4 direccions). Després, es classifiquen les característiques pel mètode del discriminant lineal de Fisher.

Un altre estudi comparatiu clàssic és el realitzat per Connors *et al* [CON80] on s'usava com a conjunt de característiques diversos mètodes: dependència espacial de nivells de gris (estadístics de segon ordre, 5 característiques), mètode de *run length* de nivells de gris amb 5 característiques, mètode de potència espectral (3 característiques) i el mètode de diferències de nivells de gris (5 característiques). Es generen les classes de textura amb un classificador de patrons i es comparen el resultats obtinguts.

Un altre treball també interessant és el presentat per Ojala *et al* en [OJA94]. Aquí és mostra l'avaluació de mesures de textura amb classificació segons diversos mètodes existents com ara: el mètode de la diferència de nivells de gris (4 característiques), mesures de textura de Laws (4 característiques amb màscares de 3x3), mesures d'autocorrelació (3 característiques) i un mètode relativament nou de patrons locals binaris amb mesures de contrast (2 característiques). Totes aquestes característiques són classificades pel mètode del veí més proper amb discriminació de Kullback [KUL68].

Més recentment, ha aparegut un treball, en [VEN96], on es comparen diferents algorismes d'anàlisi de textura els qual obtenen característiques que després seran avaluades en forma de comparació gràfica. Els mètodes comparats són: matrius de co-ocurrència (12 característiques), matrius de *run length* (7 característiques) i l'estadístic del ratio de variació de l'amplitud (AVRS) amb 6 característiques [ZUA90]. En aquest enfocament (AVRS), s'examina la distància entre un punt i un veí amb el mateix valor de nivell de gris en una direcció donada. Els seus elements són les freqüències de

distàncies  $d$  que se succeeixen en la direcció  $\theta$  entre dos píxels amb nivell de gris  $g$ . Aquestes distàncies entre els punts seleccionats de la imatge són mesurades i la freqüència s'acumula en una matriu. Els elements de la matriu  $A_\theta(g,d)$  indiquen el nombre de vegades que el valor de gris  $g$  es repeteix a la distància  $d$  sobre la direcció  $\theta$ .

#### 4.6.2. Mètodes a comparar.

En aquesta secció, presentem un estudi comparatiu entre diferents mètodes d'anàlisi de textures que parteixen de l'extracció de característiques que, en un procés posterior, serviran per classificar les textures que representen. Aquests mètodes a comparar són el nostre mètode d'extracció de característiques paramètriques de textura, un mètode estadístic de nivell de gris de segon ordre (en aquest cas, les matrius de co-ocurrència), un mètode estadístic de *run length* de nivells de gris  $i$ , finalment, un mètode de patrons binaris locals.

En el nostre procediment fixem diverses premisses:

1. Per obtenir cert nivell de fiabilitat dels algorismes, el joc de proves ha de ser gran. Amb poques imatges, la casualitat juga un paper massa important.
2. Les imatges a utilitzar per avaluar els diferents algorismes han de contenir textures molt generals i no es poden restringir a imatges amb un nombre realment limitat de textures.
3. La mètrica de la comparació és el percentatge de correcta classificació.

Pel motiu anteriors, usarem les imatges de l'àlbum de Brodatz, [BRO66]. Són imatges suficientment variades com per ser considerades textures generals, i en concret, usarem un mosaic amb 16 imatges on cada imatge ocupa 64x64 píxels representats amb 6 bits cadascun [PAR86].

#### 4.6.3. Característiques usades en la comparació.

A continuació descriurem els conjunts de característiques que utilitzarem en l'estudi de comparació, així com els algorismes amb els quals s'extreuen. Alguns d'aquests algorismes estan àmpliament explicats en el capítol 2 per la qual cosa no ens estendrem més del que sigui realment necessari.

##### 4.6.3.1. Matrius de co-ocurrència.

Segons [WES76], si la textura és direccional el grau de dispersió dels valors de la matriu de co-ocurrència al voltant de la diagonal pot variar segons la distància que s'analitza i es poden extreure mesures de dispersió per diverses distàncies  $d$ . També es pot avaluar la gruixudesa de la textura en termes de l'acumulació d'elements al voltant

de la diagonal. A partir de les consideracions anteriors, aquest estadístic de segon ordre de nivells de gris servirà per extreure la següent característica:

- Contrast:  $CON \equiv \sum (i-j)^2 P_{ij}$ . És essencialment el moment d'inèrcia de la matriu sobre la seva diagonal. És una mesura natural del grau de dispersió dels valors de la matriu.

Per a cada matriu, la distància entre píxels és

$$d \equiv (\Delta x, \Delta y) = (1,0), (0,1), (2,0), (0,2), (4,0), (0,4), (8,0), (0,8), \\ (1,1), (1,-1), (2,2), (2,-2), (3,3), (3,-3), (6,6), (6,-6)$$

Les primeres vuit distàncies són en direcció horitzontal i vertical ( $0^\circ$  i  $90^\circ$ ) a distància 1, 2, 4, i 8, mentre que les restants són en direccions diagonals ( $45^\circ$  i  $135^\circ$ ) a distàncies  $\sqrt{2}$ ,  $2\sqrt{2}$ ,  $3\sqrt{2} \approx 4$  i  $6\sqrt{2} \approx 8$ . Això fa un total de 16 característiques [DAV79], [DAV81].

#### 4.6.3.2. Matrius de run length de nivell de gris.

En una textura gruixuda s'espera que apareguin *runs* relativament llargs bastant sovint, mentre que una textura fina hauria de tenir *runs* curts. En una textura direccional, els *runs* depenen de la direcció d'avaluació. Les característiques que usem són les següents tal com les vàrem definir en el capítol 2:

- Èmfasi de *runs* llargs (RF2).
- No-uniformitat del nivell de gris (RF3).
- No-uniformitat de *run length* (RF4).
- Percentatge del *run* (RF5).

Aquestes quatre característiques es calculen en les direccions  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  i  $135^\circ$ . Això fa que treballem amb 16 característiques [GAL75].

#### 4.6.3.3 Patrons locals binaris.

Aquest va ser proposat per [OJA96] i s'ha explicat en l'apartat 2.1.8. El mètode dona una manera robusta de descriure patrons locals binaris (PLB) en una textura. Aquest algorisme es combina juntament amb una mesura de contrast calculant per a cada veïnatge la diferència de la mitjana dels nivells de gris dels elements que tenen el valor '1' amb la mitjana dels que tenen el valor '0'. Aquesta característica s'anomena PLB/C. Així mateix es crea una altra característica, PLB/COV, on es calcula la covariància per a cada veïnatge.

Aquestes són les dues característiques que es calcularan per totes les imatges i es classificaran per tal d'avaluar el resultat de la segmentació.



#### 4.6.4. Resultats de la comparació.

En primer lloc, es calculen les característiques abans mencionades per a totes les imatges i es du a terme la classificació amb l'algorisme que hem explicat en l'apartat 4.3.1 usant una única característica cada vegada. En la taula 4.3 es poden veure els resultats.

Encerts en la classificació usant cada característica individualment							
Run Length	Encerts (%)	Co-ocurrència	Encerts (%)	Patrons locals binaris	Encerts (%)	Paràmetres de textura	Encerts (%)
RF2(0)	62.9	CON(1,0)	72.7	PLB/C	62.7	L	31.6
RF2(45)	51.6	CON(0,1)	69.6	PLB/COV	71.0	A	29.7
RF2(90)	52.9	CON(2,0)	52.7			D	38.7
RF2(135)	63.4	CON(0,2)	54.6			G	29.9
RF3(0)	49.3	CON(4,0)	52.9			F	22.3
RF3(45)	47.6	CON(0,4)	55.6				
RF3(90)	50.0	CON(8,0)	51.0				
RF3(135)	52.1	CON(0,8)	53.6				
RF4(0)	61.9	CON(1,1)	62.4				
RF4(45)	60.1	CON(1,-1)	68.6				
RF4(90)	64.2	CON(2,2)	60.2				
RF4(135)	62.3	CON(2,-2)	68.7				
RF5(0)	50.1	CON(3,3)	60.1				
RF5(45)	47.2	CON(3,-3)	63.2				
RF5(90)	48.1	CON(6,6)	59.6				
RF5(135)	53.1	CON(6,-6)	64.0				

**Taula 4.3.** Resultat de la classificació amb característiques individuals.

Tal com es pot veure, els resultats són pobres. Els paràmetres de textura (linealitat, abruptitat, difuminat, granulositat i discontinuïtat) ofereixen individualment uns resultats pitjors que les altres característiques calculades també individualment. D'aquí es desprèn que s'han d'agrupar algunes característiques per millorar la classificació. En el cas dels paràmetres de textura l'agrupament no deixa dubtes: s'agafen tots els paràmetres per a la classificació. Amb els patrons binaris locals també escollim les dues característiques.

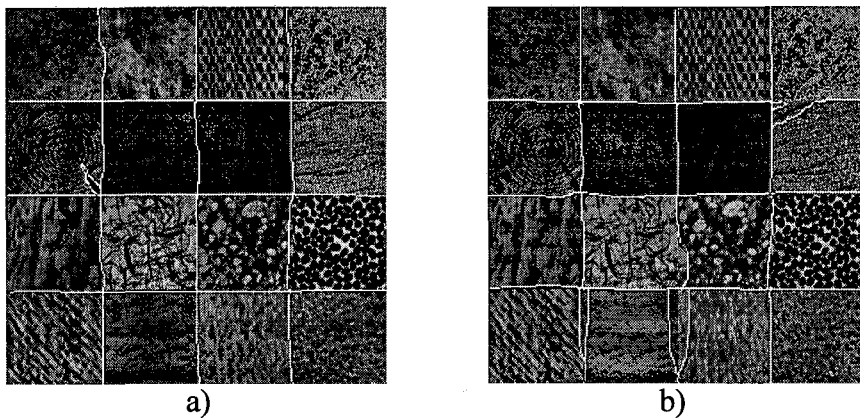
Pel que respecta als mètodes de *run length* i de matrius de co-ocurrència, s'escullen les millors parelles de característiques dels 120 parells que es poden obtenir amb cada mètode. En tots els estudis comparatius, mai s'escullen més de dues característiques donat que les combinacions creixen ràpidament i el temps de càlcul esdevé excessiu.

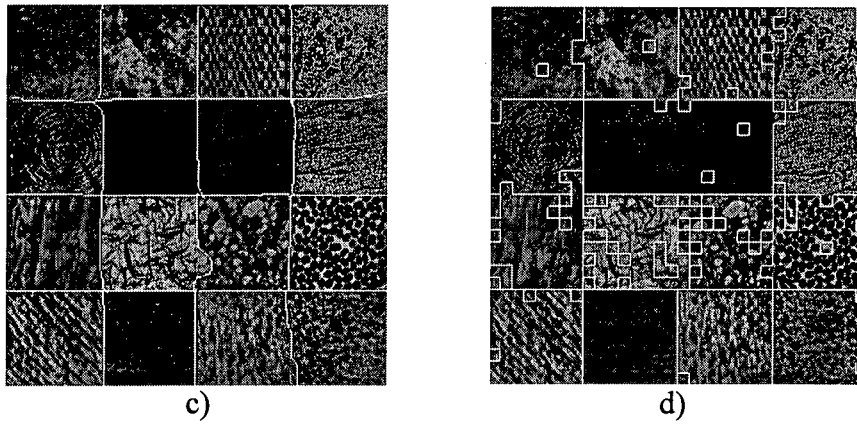
Amb totes les característiques seleccionades, es torna a classificar tota la informació i els resultats els podem veure en la taula 4.4.

Encerts en la classificació usant característiques agrupades		
Conjunt de característiques	Característiques	Encerts (%)
Run Length	RF2(135), RF2(0)	94.23
	RF2(135), RF2(90)	92.4
	RF4(0), RF4(135)	94.2
	RF4(90), RF4(135)	91.2
Co-ocurrència	CON(1,-1), CON(0,1)	96.77
	CON(1,-1), CON(0,1)	95.2
	CON(2,-2), CON(1,0)	94.9
	CON(1,0), CON(0,1)	94.6
Patrons locals binaris	PLB/C, PLB/COV	94.92
Paràmetres de textura	L, A, D, G, F	85.7

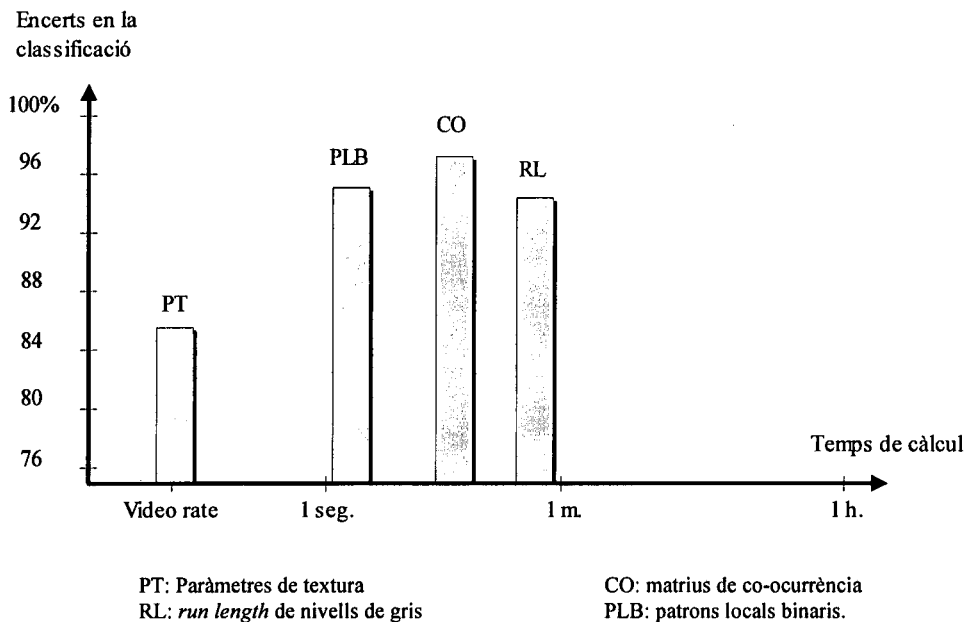
**Taula 4.4.** Resultat de la classificació amb característiques agrupades.

Les imatges de la figura 4.17 mostren els resultats de la classificació. Només es mostra el millor resultat per a cada conjunt de característiques. Els estadístics de segon ordre es comporten millor en direccions diagonals i verticals, la qual cosa indica certa presència d'aquestes estructures en les imatges. També es veu que les millors distàncies on s'analitzen els elements són distàncies curtes d'un píxel [CHE79]. Les característiques de *run length* són més sensibles al soroll i potser per això no es comporten tant bé com les matrius de co-ocurrència. Els patrons binaris locals donen bons resultats al ser invariants a les variacions de nivell de gris, tal com ja va veure [OJA94]. Els nostres paràmetres potser són els que donen uns resultats més fluixos respecte als altres, sobretot perquè les dues subimatges de la segona fila són considerades iguals. Això fa que el nombre d'errades en la classificació augmenti enormement, tot i que no és molt greu donat que per un observador humà tampoc no és evident la segmentació entre aquestes dues textures. Per tant, aquestes deficiències en la classificació són completament acceptables.





**Figura 4.17.** Resultat de la classificació segons els diversos conjunts de característiques de textura: a) estadístic de segon ordre (matrius de co-ocurrència); b) *run length* de nivells de gris; c) patrons locals binaris; d) paràmetres de textura.



**Figura 4.18.** Qualitat en la classificació respecte temps de càlcul.

El factor determinant per a la selecció del nostre mètode de paràmetres ha sigut el temps de càlcul necessari per tal de trobar les característiques anteriors. Això ho podem veure en la figura 4.18 on es relaciona el percentatge d'encerts en la classificació en front del temps de càlcul necessari per extreure les característiques.

El nostre sistema d'extracció multiparamètrica de característiques de textura està dissenyat per a treballar en temps real (considerem *video rate*) mentre que els altres mètodes requereixen d'ordres de magnitud superiors. En concret, el càlcul de les matrius de co-ocurrència més el càlcul de la característica  $CON(1,-1)$  i  $CON(0,1)$  necessita al voltant dels 30 segons, usant un ordinador personal (i486). Pel que fa a les matrius de *run length* de nivells de gris més el càlcul de les característiques  $RF2(135)$

i RF2(0) necessita més temps de càlcul (sobre els 45 segons). Per trobar les característiques PLB/C i PLB/COV el temps és bastant inferior respecte als dos mètodes anteriors, arribant-se a l'ordre d'alguns segons per combinar les dues característiques. Si pensem en mètodes estructurals, que aquí no s'han provat, el temps seria encara molt superior incrementant-se també la qualitat en la classificació.

En els treballs presentats en [CAT91] i [CAT92] hem aplicat les xarxes neuronals per a l'anàlisi de textures i segmentació de carreteres a partir de dos paràmetres de textura: punts singulars i desviació. En aquells treballs s'obtenien uns resultats molt satisfactoris arribant a un 92% i 95% de tèxels correctament classificats com carretera o no carretera en la discriminació de textures, després de realitzar aprenentatges amb més de 50 imatges diferents. En aquest aspecte dóna millors resultats que amb els cinc paràmetres de textura. En canvi, pel que fa al temps de processat i segmentació, el cost computacional era molt elevat pel fet d'usar una xarxa neuronal i el temps d'operabilitat no és comparable amb l'alta velocitat (*video rate*) que s'assoleix amb el mètode que es proposa en aquesta tesi.

En conclusió, després de veure l'anterior comparació no es pot dir quin mètode és millor o quin és pitjor. Segons les necessitats en l'aplicació real que es treballi, s'escollirà un mètode o un altre tenint en compte el compromís entre qualitat en la classificació o temps de càlcul de les característiques. En definitiva, el nostre mètode ha apostat per una ràpida extracció de característiques per segmentar imatges de textura en temps real.

#### **4.7. Comportament del procés de segmentació davant de soroll aleatori.**

Una de les proves indispensables a realitzar en qualsevol nou mètode de processat del senyal és observar el seu comportament davant la presència de soroll. La variació de la sortida del sistema abans i després de la introducció de soroll ens marcarà la seva robustesa.

Per analitzar el comportament i, en definitiva, la robustesa del nostre mètode d'extracció de característiques paramètriques de textura i la seva posterior classificació i segmentació, hem realitzat la següent prova:

1. S'ha generat una imatge sintètica per assegurar la mesura del grau d'encerts en la segmentació. Si no sabéssim el resultat de la segmentació sense soroll, difícilment podríem avaluar la taxa d'errors després d'introduir distorsions. Concretament, la imatge és generada a partir de les màscares que busquen patrons de textura en les imatges d'entrada, per la qual cosa la imatge sintètica contindrà regions de textura perfectament discriminables per a tots els paràmetres i la taxa d'encerts en la classificació serà del 100%.

2. Després de definir la imatge sobre la qual introduïrem el soroll, cal definir com serà aquest soroll i on s'aplicarà. El soroll serà una variable aleatòria que s'aplicarà a cada píxel de la imatge d'entrada. La quantificació d'aquest soroll serà el resultat de generar una variable aleatòria que segueix una distribució normal (gaussiana) amb mitjana zero i variància  $\sigma_w$ . Aquesta variable aleatòria se sumará a cada píxel de la imatge.

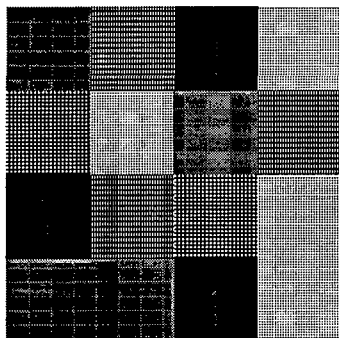
En concret, s'ha fet l'estudi del comportament del sistema amb dos imatges sorolloses amb una variància  $\sigma_w = 100$  i  $\sigma_w = 128$ , seguint el següent model d'addició de soroll aleatori sobre el domini de l'espai discret.

$$f(n1,n2) = I(n1,n2) + v(n1,n2)$$

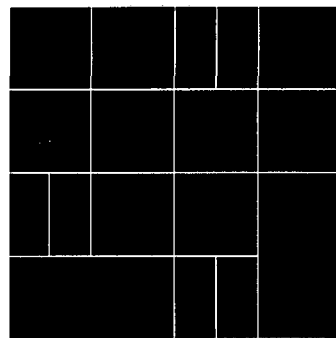
on la imatge sorollosa  $f(n1,n2)$  és funció de la imatge original  $I(n1,n2)$  més la suma del soroll aleatori  $v(n1,n2)$ . La taula 4.5 resumeix els resultats obtinguts.

Imatge Fig. 4.19	Variància $\sigma_w$	Nombre de tèxels mal classificats (sobre 1024)	% d'encerts en la classificació
a)	Sense soroll	0	100%
c)	100	45	95.60%
e)	128	80	92.19%

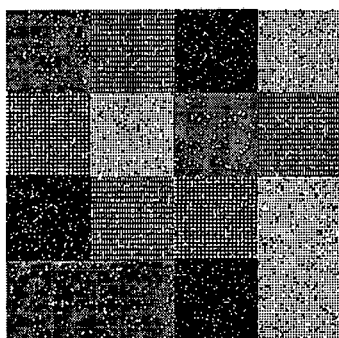
Taula 4.5. Resultats de la classificació en imatges sorolloses.



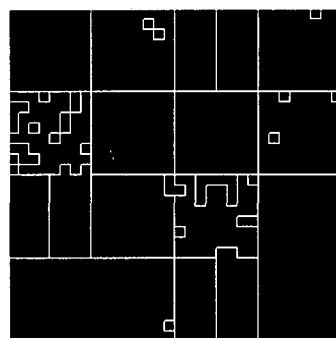
a)



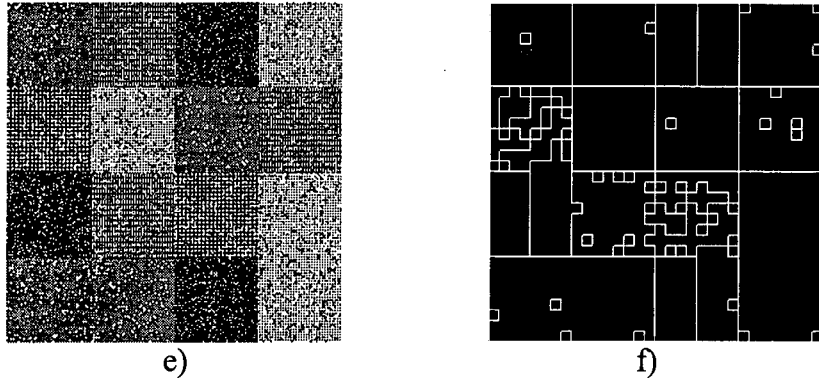
b)



c)



d)



**Figura 4.19.** Qualitat en la classificació respecte al soroll afegit. Imatges d'entrada i resultats en la seva segmentació. a) sense soroll,  $\sigma_w = 0$ ; c)  $\sigma_w = 100$ ; e)  $\sigma_w = 128$ .

Tal com s'aprecia en els resultats de la segmentació en la figura 4.19, el procés d'extracció de característiques de textura i el procés de la classificació conjuntament es comporten molt bé davant la introducció d'una funció aleatòria amb una variància tant elevada, per la qual cosa podem dir que el mètode que es presenta i s'avalua és robust al soroll extern.