

**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 2. ANTECEDENTS.

Durant molts anys, l'anàlisi i classificació de textures ha estat un punt d'estudi en el qual molts autors han investigat i escrit els seus treballs. En aquest capítol es vol fer un recull dels mètodes més rellevants que han aportat quelcom de nou i interessant al problema de l'anàlisi i segmentació per textures. Els enfocaments principals d'anàlisi de textures que hem vist en el capítol anterior seran tractats per separat. Mentre que els enfocaments estadístics són útils quan les primitives de textura corresponen a àrees o regions petites de la imatge, *microtextura*, esdevé necessari un enfocament estructural quan apareixen *macrot textures*, on les seves primitives corresponen a regions grans de la imatge i, a més, cal determinar-ne les regles d'emplaçament.

Després de fer un viatge en el temps recorrent les diverses i més variades tècniques per a la classificació de textures, situarem el mètode presentat en aquest treball dins l'àrea oportuna, on diversos investigadors segueixen línies semblants a la que es presenta aquí. Això ens encoratja a seguir la investigació, pensant que no estem sols dins aquest mar tant extens com és el tractament i anàlisi de textures. Després de veure l'evolució al llarg del temps dels diferents treballs, som conscients que amb aquest treball no hem arribat a cap final de recorregut sinó que solament hem arribat al final d'una etapa i que encara queda un llarg trajecte dins aquesta línia d'investigació.

Hem dividit aquest capítol en tres seccions. La primera secció ens acosta als mètodes que pertanyen a l'enfocament estadístic. La segona secció explica resumidament l'enfocament estructural i alguns dels mètodes més interessants. S'ha afegit una tercera secció que descriu esquemàticament tècniques per a la síntesi de textures que seran útils per trobar i provar tècniques de classificació de textures.

### 2.1. Enfocament estadístic.

L'enfocament estadístic per a la caracterització de textures intenta trobar i definir unes característiques globals de la textura. Com a descriptors de textura s'utilitzen propietats estadístiques de la distribució espacial dels nivells de gris de la imatge. La característica clau d'aquest enfocament és la dependència exclusiva dels descriptors en les propietats d'un punt o píxel, sense l'ús explícit d'elements o subregions.

La majoria dels models de textura existents estan basats en aquest enfocament. Diversos autors han realitzat classificacions segons la tècnica usada per a l'extracció de característiques de textura i s'han creat grans grups de mètodes dins l'enfocament estadístic.

Aquests grups són:

- models de camps aleatoris i autoregressius
- *run length* de nivells de gris

- transformades d'energia de textura
- funcions d'autocorrelació
- probabilitats de co-ocurrència espacial de nivells de gris
- transformades digitals
- extracció de característiques paramètriques
- altres mètodes: mosaics aleatoris, fractals, ...

El treball que aquí es presentarà està ubicat dins el grup d'extracció de característiques paramètriques. Diversos autors han optat per aquest via a l'hora de caracteritzar una imatge de textura.

A continuació, es recull els diferents mètodes més importants amb els treballs més rellevants per a cadascun d'aquests grups.

### 2.1.1. Models dels camps aleatoris i autoregressió.

Els models dels camps aleatoris han estat àmpliament utilitzats en l'anàlisi d'imatges per a la seva modelització, compressió o altres aplicacions. Els camps aleatoris de Markov (MRF) no són només descriptius sinó també generatius i s'usen per a la síntesi i classificació de textures i segmentació, restauració i compressió d'imatges. L'interès d'aquests mètodes es centra en el reconeixement de textures usant paràmetres de models MRF com a característiques. El model de camp aleatori de Markov assumeix que la textura és estocàstica, estacionària (invariant a translacions) i satisfà un criteri d'independència condicional.

Sigui  $\mathbf{R} \times \mathbf{C}$  el domini espacial de la imatge i per a cada  $(x,y) \in \mathbf{R} \times \mathbf{C}$ ,  $N(x,y)$  indica els veïns de  $(x,y)$ . A causa de la estacionarietat del camp,  $(a,b) \in N(x,y)$  si i només si  $(a+i, b+j) \in N(x+i, y+j)$ . Això significa que la configuració espacial del veïnatge és la mateixa per a tota la imatge. Hi ha una dificultat òbvia per a que aquesta condició es compleixi en els límits de la imatge. La manera habitual de tractar el problema teòricament és considerar la imatge com a un objecte de revolució (esquema toroïdal). En aquest cas, el veïnatge canònic espacial pot venir representat per  $N(0,0)$ .

El criteri d'independència condicional és que la probabilitat condicional d'un píxel  $(x,y)$  donats tota la resta de píxels en la imatge és igual a la probabilitat condicional d'aquest píxel donats només els píxels del seu veïnatge. Això és

$$P(I(x,y) | I(i,j) : (i,j) \in \mathbf{R} \times \mathbf{C}, (i,j) \neq (x,y)) = P(I(x,y) | I(i,j) : (i,j) \in N(x,y))$$

Aquest criteri s'anomena *markovianietat*.

Un camp aleatori és *estacionari*, en termes de moviment euclidià, si existeix invariància sota translacions. Un camp aleatori és *isotròpic* si la seva funció de covariància és invariant sota qualsevol rotació al voltant d'un punt fix. Llavors, un camp aleatori és *homogeni* si és estacionari i isotròpic. L'homogeneïtat implica invariància sota qualsevol moviment euclidià rígid. Després d'aquestes definicions, els models de variacions espacials són camps aleatoris homogenis escalars reals discrets.

L'aplicació més habitual d'aquest mètode és crear un model de camp aleatori de Markov per a la generació i anàlisi d'imatge de textura. En aquests models es produeixen sistemes d'anàlisi i síntesi de textura on es pren com a entrada una textura, s'analitzen els seus paràmetres d'acord amb el model MRF i es genera com a resultat imatges textura que s'assemblen visualment a la imatge d'entrada i coincideixen molt acuradament des del punt de vista estadístic. [CRO83] aplica un sistema semblant creant textures sintètiques que els humans no poden diferenciar de les textures reals d'entrada al sistema.

[ONU91] modelitza fractals de textura a partir de camps aleatoris de Markov en cadascun dels seus nivells de resolució. La textura es controla segons els paràmetres del model MRF.

Un altre model a considerar és el camp aleatori de Gauss-Markov (GMRF), ja que representa classes de textura més àmplies. En aquest model, els nivells de gris d'una textura segueixen una distribució Gaussiana, la mitjana de la qual és una combinació lineal dels nivells de gris veïns i la seva variància és una constant pròpia de la textura [SCH80a].

*Definició de camp aleatori gaussià:* suposem que tenim una regió de textura  $A \subseteq E^2$  d'àrea unitat. Si  $G(x)$  denota el nivell de gris del punt  $x \in A$ , aleshores un camp real  $\{G(x) \mid x \in A\}$  és un camp aleatori gaussià si per a cada conjunt finit de punts  $x_i \in A$ , els corresponents nivell de gris  $G(x_i)$  són variables aleatòries de Gauss conjuntament. Un camp aleatori gaussià homogeni està completament caracteritzat per la seva mitjana i la seva covariància.

El problema dels GMRF és l'estimació dels paràmetres del model i la segmentació de les textures basades en els valors estimats. Una forma de fer aquesta estimació seria utilitzant un model autoregressiu (AR). Aquest model permet l'extracció d'un estadístic de textura per a cada píxel, el qual pot ser ponderat al llarg de regions arbitràries de píxels per oferir un estadístic agregat a la regió i és, d'altra banda, adaptat per a l'ús en tècniques de múltiple resolució. L'avantatge d'un model predictiu de textura sobre un model que usa només la mitjana i la variància és particularment important en l'estimació del nombre de textures diferents quan aquestes tenen una alta correlació espacial.

Es poden obtenir models d'ordre superior i GMRF, però pràcticament tots els autors usen models de segon ordre, [MAN91], [BOU91], [KAN82], [OHA92], ja que no compensa l'esforç en el càlcul de models d'ordre superior amb els resultats que s'en obtenen. [COH91] aplica un model GMRF per a la inspecció tèxtil usant propietats visuals de textura.

Darrerament s'estan usant models més complexos i més desenvolupats com ara la tècnica del *hidden Markov model* (HMM). Aquesta tècnica estadística s'usa per a classificar textures [POV95]. El model HMM assumeix que una imatge pot ser modelitzada per un procés estadístic els estats del qual són no directament observables i que cada estat té una certa distribució estadística d'una quantitat observable, com ara el nivell de gris. El potencial avantatge del mètode HMM sobre altres mètodes de

classificació de textures és que un HMM intenta discernir una estructura fonamental subjacent en una imatge que no pot ser directament observable. La determinació d'aquesta estructura fonamental, representada pels estats dels píxels, pot donar una caracterització a la imatge potencialment més robusta.

### 2.1.1.1. Model autoregressiu.

El model discret del camp aleatori de Markov és una generalització dels models en sèries de temps autoregressius. A partir de models autoregressius es pot atacar el problema de la caracterització de textures i, amb aquesta caracterització, realitzar una partició de la imatge basada en la textura. Donada una imatge, [DEG78] presenta un enfocament basat en l'estimació del nivell de gris d'un píxel a partir d'una suma ponderada (amb pesos) dels nivells de gris dels píxels veïns. Els coeficients dels pesos són determinats minimitzant l'estimació de l'error quadràtic mitjà. L'estimador òptim bi-dimensional s'usa per a la caracterització de la textura i per a la partició de la imatge. En aquest segon pas, una imatge es parteix en regions amb característiques de textura uniformes.

Suposem una imatge com una matriu d' $R \times C$  elements  $(x,y)$ . Suposem que aquesta imatge és la sortida d'un sistema lineal bi-dimensional amb una entrada de soroll blanc i que la textura de cada imatge, generada per aquest sistema lineal, és caracteritzada per la seva estructura. Més específicament, es pot dir que el nivell de gris d'un píxel  $(x,y)$  ve donat per

$$(x,y) = \sum_{\substack{p=-M \\ (p,q) \neq (0,0)}}^M \sum_{q=-N}^N a_{p,q} (x-p, y-q) + n_{x,y} \quad \text{per a tot } x,y$$

on  $n_{x,y}$  és el soroll blanc de l'entrada,  $a_{p,q}$  són els coeficients dels pesos dins una finestra de mida  $(2N+1)(2M+1)$ . Aquests pesos caracteritzen la textura.

Quan s'estima el nivell de gris d'un píxel a partir de la suma lineal amb pesos del nivell de gris dels píxels veïns, es produeix un error entre aquest valor i el valor real del nivell de gris del píxel. Aquest error s'anomena  $e_{x,y}$  i el valor esperat de l'error quadràtic ve donat per

$$S = \sum_x \sum_y E[e_{x,y}^2]$$

L'objectiu serà minimitzar la funció  $S$ . Per trobar el mínim absolut d' $S$ , cal trobar la mida òptima de finestra.

En treballs més recents, [LU95] usa models autoregressius bi-dimensionals per a la descripció de textures. Els coeficients del model AR, com a paràmetres, poden usar-se per identificar textures. El conjunt de paràmetres del model autoregressiu caracteritza la dependència d'un píxel amb els seus veïns. Aquests paràmetres són usats com a

vector de característiques per a discriminar tipus de textures. En aquell treball s'adoptaven dues tècniques per a l'estimació dels paràmetres: LSE i MSE, és a dir, error mínim quadràtic i mètode d'estimació de màxima probabilitat. La implementació del càlcul dels paràmetres es feia amb una xarxa neuronal, donada la seva natura d'execució en paral·lel i la seva facilitat per a l'aprenentatge adaptatiu.

Els models autoregressius són variants a les rotacions. Aquest efecte tindria valor quan es volgués discriminar entre textures amb orientacions diferents. Hi ha hagut alguns intents per a derivar característiques invariants a les rotacions en l'anàlisi de textura. [KAS86a] suggereixen un model invariant a rotacions anomenat model autoregressiu circular (CAR). En aquest model, la mitjana dels nivells de grisos del píxel tractat i els seus vuit veïns defineixen una nova variable aleatòria per a cada píxel. [MAO92] amplia el model circular i construeix un model invariant a rotacions multivariante (RISAR).

Existeixen dos problemes principals associats a qualsevol model autoregressiu. Un és el problema d'escollir la mida del veïnatge on es mira la dependència del píxel central observat en aquell moment. L'altre problema és seleccionar la mida apropiada de finestra en la qual es mira la textura esperant-ne la seva homogeneïtat i sobre la qual s'estimaràn els paràmetres del model autoregressiu. Molts estudis usen mida fixa pel veïnat i mida fixa per la finestra, les quals es calculen empíricament. El problema de treballar amb mida fixa és la no-adaptabilitat. [LU95] treballa amb finestra de veïns no quadrada. [MAO92] treballa amb una gran finestra (25x25) i usa només la resolució que necessita per a cada tipus de textura.

### 2.1.2. Run Lengths.

Per a l'anàlisi de textures, una característica útil són els *run length* de nivell de gris. Una de les autores que fa un dels estudis més exhaustius i profunds és Galloway a [GAL75].

Un *run* de nivell de gris és un conjunt consecutiu i colineal de punts d'una imatge que tenen el mateix valor de nivell de gris. La longitud del *run* és el nombre de punts en la imatge dins el *run*. Per a una imatge donada, es pot calcular una matriu de *run length* de nivells de gris per *runs* que tinguin una direcció donada. L'element  $(i,j)$  de la matriu especifica el nombre de vegades que la imatge conté un *run* de longitud  $j$ , en una direcció donada, que està format per punts amb un nivell de gris  $i$ .

El següent exemple mostra una imatge de 4x4 píxels amb 4 nivells de gris (0..3) cadascun i es mostren les matrius de *run length* de nivells de gris resultants en les quatre direccions principals.

0	1	2	3
0	2	3	3
2	1	1	1
3	0	3	0

a)

		Run Length			
		1	2	3	4
Nivell de gris	0°	4	0	0	0
	1	1	0	1	0
	2	3	0	0	0
	3	3	1	0	0

b)

		Run Length			
		1	2	3	4
Nivell de gris	45°	4	0	0	0
	1	4	0	0	0
	2	0	0	1	0
	3	3	1	0	0

c)

		Run Length			
		1	2	3	4
Nivell de gris	90°	2	1	0	0
	1	4	0	0	0
	2	3	0	0	0
	3	3	1	0	0

d)

		Run Length			
		1	2	3	4
Nivell de gris	135°	4	0	0	0
	1	4	0	0	0
	2	3	0	0	0
	3	5	0	0	0

e)

**Figura 2.1.** a) Imatge de 4x4 elements. Matrius de *run length* en les quatre direccions principals: b) 0°; c) 45°; d) 90°; e) 135°.

El càlcul d'aquestes matrius és realment senzill i el nombre d'operacions a realitzar és directament proporcional al nombre de punts de la imatge, la qual cosa resulta d'un alt cost computacional per a imatges amb un nombre mitjà de nivells de gris (64, 128, 256) i mida també mitjana (per exemple 256x256).

Per a obtenir mesures de textura numèriques a partir de les matrius, es poden calcular funcions anàlogues a les utilitzades per [HAR73] en les matrius de co-ocurrència de nivells de grisos.

Partint de les següents definicions, les diferents mesures que [GAL75] va calcular venen detallades a continuació. Sigui

- $p(i,j)$  la posició  $i,j$ -èssima de la matriu de *run length*
- $N_g$  el nombre de nivells de gris de la imatge
- $N_r$  el nombre de diferents *run lengths* que hi ha (per tant la matriu serà de dimensió  $N_g \times N_r$ )
- $P$  el nombre de punts de la imatge.

1. Èmfasis de *runs* curts.

$$RF1 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{p(i,j)}{j^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j)}$$

Aquesta funció divideix cada valor de *run length* per la longitud del *run* al quadrat. Això tendeix a emfatitzar els *runs* curts. El denominador és el nombre total de *runs* de la imatge i serveix com a factor normalitzador.

2. Èmfasis de *runs* llargs.

$$RF2 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j^2 p(i,j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j)}$$

Aquesta funció multiplica cada valor de *run length* per la longitud del *run* al quadrat. Emfatitza els *runs* llargs. El denominador és per a normalitzar.

3. No-uniformitat de nivell de gris.

$$RF3 = \sum_{i=1}^{N_g} \left( \sum_{j=1}^{N_r} p(i, j) \right)^2 / \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j)$$

Aquesta funció eleva al quadrat el nombre de *run lengths* per a cada nivell de gris. La suma dels quadrats es normalitza respecte al nombre total de *run* de la imatge. Això mesura la no-uniformitat de nivells de gris de la imatge. Quan els *runs* estan distribuïts equitativament entre tots els nivells de gris, la funció pren el mínim valor.

4. No-uniformitat del *run length*.

$$RF4 = \sum_{j=1}^{N_r} \left( \sum_{i=1}^{N_g} p(i, j) \right)^2 / \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j)$$

Es normalitza la suma dels quadrats del nombre de *runs* per a cada longitud. Aquesta funció mesura la no-uniformitat dels *run lengths*. Si els *runs* estan distribuïts equitativament per les longituds, la funció tindrà un valor baix.

5. Percentatge de *run*.

$$RF5 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j) / P$$

Aquesta funció és una relació del nombre total de *runs* respecte el nombre total de possibles *runs* si tots ells tinguessin longitud 1. Tindrà valor mínim per a imatges amb una estructura lineal màxima.

Amb totes aquestes característiques, [GAL75] va aconseguir un 83% d'identificació de textures en 6 grups d'imatges: llacs, vies de trens, aiguamolls, horts, bosc baix i edificacions. De totes maneres, l'autora remarca que s'haurien d'utilitzar les funcions de *run length* juntament amb altres tècniques de classificació de patrons.

En l'extracció de característiques de textura, [WES76] utilitza els *run length* de nivells de gris per calcular 4 paràmetres de [GAL75], a més d'altres paràmetres de transformada digital i estadístics de primer i segon ordre. Després, la segmentació de la imatge la realitza amb una classificació mitjançant la tècnica del discriminant lineal de Fisher.



### 2.1.3. Transformades d'energia de textura.

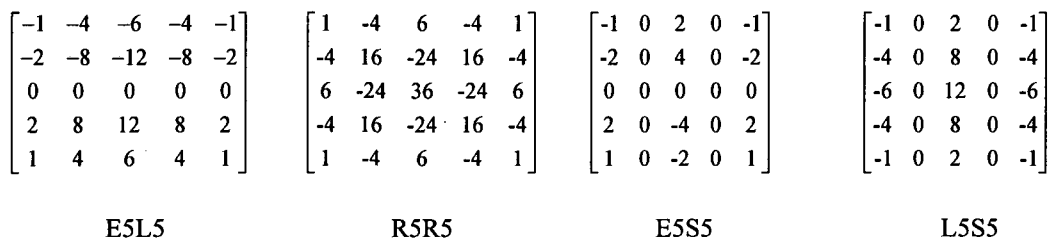
Els treballs de Laws en [LAW80] presenten un conjunt de transformades d'energia de textura que donen mesures de textura per a cada píxel d'una imatge monocroma. Aquestes mesures de la textura pertanyen al domini espacial. La caracterització de la textura es realitza en dos passos: en primer lloc, la imatge es convoluciona amb un conjunt de filtres de petita àrea; en segon lloc es creen estadístics a partir les convolucions anteriors. Aquests filtres usats en la convolució de la imatge s'anomenen màscares de microtextura i les seves sortides són les característiques de microtextura. Aquestes màscares contenen coeficients o pesos que es necessitaran en el procés de la convolució bi-dimensional.

La convolució bi-dimensional d'una imatge  $f(i,j)$  i una màscara  $h(i,j)$  de mida  $2a+1$  per  $2a+1$  ve donada per la següent relació

$$x(i,j) = (h * f)(i,j) \\ = \sum_{k=-a}^a \sum_{l=-a}^a h(k,l) f(i+k, j+l) \quad \text{per } i=0,1,\dots,N-1 \text{ i } j=0,1,\dots,N-1$$

on  $*$  és l'operació convolució bi-dimensional.

Les màscares de microtextura de Laws estan dissenyades per actuar com a filtres de comparació (*matched filters*) per a cert tipus de variacions quasi-periòdiques que sovint es troben en imatges de textura. Típicament, aquestes màscares són de mida  $7 \times 7$  o més petites. La suma dels elements de les màscares és zero, amb la qual cosa els resultats de la imatge de sortida tenen una mitjana zero. Aquestes màscares de convolució intenten ser sensibles a estructures com ara arestes, ondulacions i taques. La següent figura mostra els coeficients de quatre màscares de microtextura que tenen el millor poder de discriminació pel mosaic de textura que usa Laws.



**Figura 2.2.** Quatre màscares de microtextura de Laws.

En un segon pas, amb els resultats de les convolucions es creen diversos estadístics. Per a màscares de suma zero, les mesures de textura local es determinen com variàncies estadístiques de la imatge filtrada sobre els valors amb signe de la imatge. Des del punt de vista de l'eficiència computacional, Laws proposa 3 alternatives al càlcul de la desviació estàndard: una ponderació amb els valors absoluts de la convolució (ABS SAVE), amb els valors positius (POS SAVE) i amb els valors negatius (NEG SAVE). La que millor resultat li dona és l'ABS SAVE. Per a àrees amb mitjana

igual a zero, aquesta característica es pot considerar una aproximació ràpida de la desviació estàndard. L'ABSVE  $s(i,j)$  resulta

$$s(i,j) = \frac{1}{(2n+1)^2} \sum_{k=i-n}^{i+n} \sum_{l=j-n}^{j+n} |x(k,l)|$$

Aquestes característiques, com el ABSVE de cada imatge filtrada de microtextura, s'anomenen mesures d'energia de textura. Una vegada s'han extret les característiques d'energia de textura, hi ha un pas opcional per a reduir la dimensionalitat en el procés de classificació. Hi ha dos mètodes de selecció i extracció de característiques, depenent de si es disposa o no d'etiquetes per a les classes i els seus estadístics. Per a reduir la dimensionalitat de l'espai de característiques es pot usar el mètode de l'anàlisi múltiple del discriminant si es coneixen les etiquetes; aquest mètode és una generalització del discriminant lineal de Fisher per al problema de múltiples classes. Llavors, s'usa un classificador de mínima distància als centres de massa per assignar l'etiqueta del test d'aprenentatge. Si no es disposa de les etiquetes inicials, Laws usa el mètode de l'anàlisi de components principals.

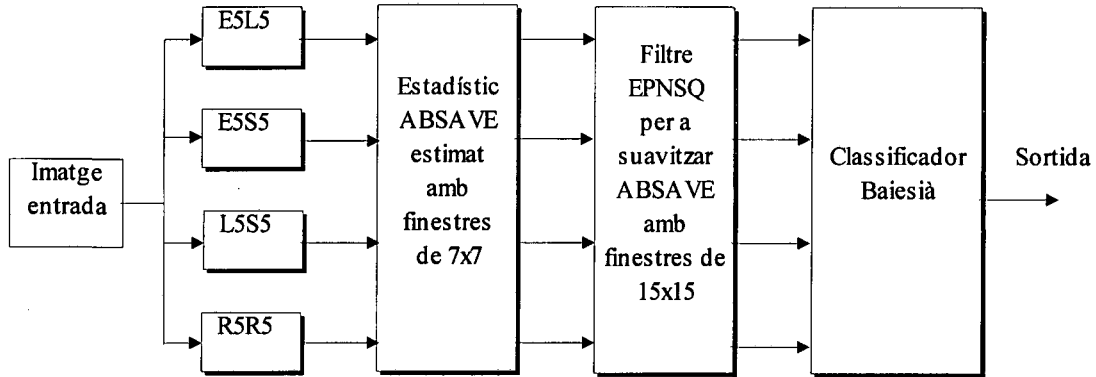
#### *Problemes potencials.*

El mètode desenvolupat per Laws dona bons resultats per regions d'àrea petita. Quan s'apliquen les característiques d'energia de textura al problema de la segmentació, hi ha errors considerables en els interiors de regions àmplies i també a prop de les fronteres entre regions. Aquest últim tipus d'error és degut al mètode d'estimació de l'energia local de textura. Recordant, les característiques macroestadístiques s'obtenen processant els plans de microtextura amb un filtre no-lineal d'energia local de textura. Aquest filtre no-lineal s'usava per estimar la desviació estàndard local de la imatge filtrada (o un estadístic semblant aproximat per una ponderació dels valors absoluts de la finestra). Tot i que aquestes operacions de finestra mòbil són ràpides i senzilles, introdueixen errors significatius al llarg de les fronteres entre regions. La raó és que la finestra usada per a calcular la característica macroestadística sovint solapa més d'una textura a la frontera de la regió. Quan això passa, els estadístics resultants són una barreja de dos conjunts d'estadístics i això fa que es pensi que la textura pertanyi a una altra classe. En aquest cas, els píxels de la frontera (o píxels solapats) es classifiquen en una tercera classe, diferent a les dues classes a les quals pertanyen les regions avaluades.

Els treballs de [HSI89] intenten millorar el problema mencionat anteriorment. Una manera d'evitar-lo seria estimar més acuradament la desviació estàndard de les imatges filtrades de microtextura. Hsiao *et al.* es plantegen el problema com el de suavitzar una imatge sorollosa: com suavitzar el soroll sense difuminar els contorns. Hi ha diverses maneres d'enfocar el problema i aquests autors donen algunes solucions: utilitzar l'estimador localment lineal de l'error quadràtic mitjà mínim (LLMMSE), que té la propietat de suavitzar el soroll en les regions planes i deixar invariant els veïns dels contorns; una altra solució és l'ús d'un filtre que aplica el mètode de preservació de contorns en el suavitzat del soroll (EPNSQ) el qual, en lloc de suavitzar indiscriminadament cada punt, suavitza la funció  $s(i,j)$  per a determinats

veïns que tenen major probabilitat de pertànyer a la mateixa regió que el punt donat  $(i,j)$ .

La seqüència proposada per [HSI89] per una imatge d'entrada és la següent:



**Figura 2.3.** Diagrama de blocs del segmentador amb el classificador de Bayes.

El filtre EPNSQ és aplicat a cadascuna de les quatre imatges que representen la sortida dels plans de característiques de les quatre microtextures originals i selecciona el veïnatge que és més probable que estigui dins una regió uniforme de la imatge. La sortida del filtre EPNSQ (vectors de característiques  $x$ ) és enviada a un classificador baiesià, assumint que les probabilitats *a priori* són equivalents per a cada classe. Llavors, la classificació es pot aconseguir usant funcions discriminants.

Un altre autor que usa les màscares de Laws per a la detecció i extracció de regions de textura compactes és [SHE89]. Per implementar aquests processos, utilitza una piràmide cel·lular que requereix d'un algorisme d'ordre  $\log N$ , sent  $N \times N$  la dimensió de la imatge a analitzar.

Les màscares de Laws, amb coeficients constants, no poden donar característiques invariants a canvis com ara rotacions i escalat, tant sovint en el món real. Una manera d'evitar aquest problema és reemplaçar les màscares fixes per màscares adaptatives per a ser invariants a aquests canvis. A [YOU93] es donen diversos criteris per a generar màscares adaptatives. En aquests algorismes es porta a terme un aprenentatge amb moltes mostres de textures pures, en les quals s'ajusten els coeficients per tal d'optimitzar un criteri. Els coeficients de les màscares continuen sumant zero.

#### **2.1.4. Funció d'autocorrelació.**

Des d'un cert punt de vista, la textura està relacionada amb la mida espacial de les primitives tonals de la imatge. Aquestes primitives tonals poden tenir una mida ampla o una mida estreta indicant així textures gruixudes o textures fines, respectivament. La funció d'autocorrelació genera una característica que ens indica la mida de les primitives tonals.

[HAR79] explica l'autocorrelació amb un exemple molt clar. Suposem dues transparències, una és una còpia exacta de l'altra. Posem una transparència sobre de l'altra i, amb una font lluminosa uniforme, mesurem la llum transmesa a través de la doble transparència. Ara, es mou una transparència respecte l'altra i mesurem la llum transmesa a través de la porció d'imatge on una transparència solapa l'altra. La funció d'autocorrelació bi-dimensional de la transparència es representa com un graf d'aquestes mesures, que és funció de les posicions traslladades  $(x,y)$ , normalitzada respecte l'origen de la translació  $(0,0)$ .

Sigui  $I(u,v)$  un punt qualsevol de la imatge sobre la qual volem trobar les primitives tonals. Suposem que fora dels límits d'una regió rectangular, definida per  $0 \leq u \leq L_x$  i  $0 \leq v \leq L_y$ , la imatge val 0. Sigui  $(x,y)$  la translació respecte els dos eixos. La funció d'autocorrelació per a la transparència es defineix formalment com a:

$$\rho(x,y) = \frac{\frac{1}{(L_x - |x|)(L_y - |y|)} \iint_{\infty} I(u,v)I(u+x,v+y)dudv}{\frac{1}{L_x L_y} \iint_{\infty} I^2(u,v)dudv}$$

on  $|x| < L_x$  i  $|y| < L_y$

Discretitzant, la funció d'autocorrelació també es pot definir en forma de sumatoris de la forma:

$$C(i,j) = \frac{\sum_{r,c} I(r,c)I(r+i,c+j)}{\sum_{r,c} I^2(r,c)}$$

on  $r$  i  $c$  són els índexs de fila i columna dins la imatge. Els índexs de desplaçament  $i$  i  $j$  poden prendre valors negatius, sent la funció simètrica respecte l'origen.

Si les primitives tonals de la imatge són relativament grans, llavors el resultat de la funció d'autocorrelació disminuirà lentament amb la distància. Si les primitives tonals són petites, llavors la funció disminuirà ràpidament amb la distància, marcada per  $x$  i  $y$  o per  $i$  i  $j$  en cada cas. Si les primitives tonals són espacialment periòdiques, la funció d'autocorrelació disminuirà i s'incrementarà de nou seguint una forma periòdica.

La primitiva tonal en el model d'autocorrelació és el nivell de gris. L'organització espacial és caracteritzada pel coeficient de correlació, que és una mesura de la dependència lineal d'un píxel respecte un altre. La relació entre la funció d'autocorrelació i la densitat espectral és la transformada de Fourier una de l'altra.

[KAI55] va dur a terme un experiment per veure si la funció d'autocorrelació tenia alguna relació amb la textura que els interpretadors de fotografies veien en les imatges. S'usava una sèrie de set fotografies aèries d'una regió àrtica i va determinar la funció d'autocorrelació de les imatges. Se suposa que la funció d'autocorrelació és

circularment simètrica i calculada només com a una funció de distància radial. Per a cada imatge, va trobar la distància  $d$  per la que la funció d'autocorrelació  $\rho$  prenia el valor

$$\rho(d) = 1/e$$

L'investigador va preguntar a 20 individus que ordenessin les set imatges en una escala de gruixudesa (de detalls més fins a més gruixuts). Va correlar les imatges ordenades amb les distàncies corresponents al valor  $1/e$  de la funció d'autocorrelació. Amb això va trobar un coeficient de correlació de 0.99. Amb aquests resultats va establir que, al menys amb el seu conjunt de dades, la funció d'autocorrelació i els individus varen mesurar la mateixa classe de característiques de textura.

Després d'haver generat textures estocàstiques, [FAU80] comprova que un conjunt de característiques de textura suficientment universal ha de donar la informació necessària per a discriminar camps gaussians que tinguin diferent funció d'autocorrelació, així com discriminar camps no gaussians que tinguin idèntica funció d'autocorrelació. Aleshores, a partir de mesures d'autocorrelació espacial i d'un estadístic de primer ordre calcula quatre moments: mitjana, desviació, biaix i kurtosis. Per segmentar la imatge en diferents textures usa un classificador baiesià amb l'anterior conjunt de característiques de textura.

### **2.1.5. Co-ocurrència: Dependència espacial de nivells de gris.**

L'enfocament de la dependència espacial dels nivells de gris caracteritza la textura per la co-ocurrència dels seus nivells de gris. En les textures bastes o gruixudes la distribució canvia només de forma suau amb la distància; en canvi, en les textures fines la distribució canvia ràpidament amb la distància.

La potència de l'enfocament co-ocurrent es basa en caracteritzar les interrelacions espacials dels nivells de gris en un patró textural i això es pot fer de manera que sigui invariant sota transformacions monotòniques de nivell de gris. El desavantatge d'aquest enfocament és que no té en compte els aspectes de forma de les primitives tonals. Per això, no és possible treballar bé amb textures compostes per primitives que ocupin una gran àrea de la imatge.

Un aspecte de la textura fa referència a la distribució espacial i la dependència espacial entre els nivells de gris en una àrea local. [JUL62] va ser el primer en usar pels experiments en discriminació de textures els estadístics de co-ocurrència en la dependència espacial dels nivells de gris. [DAR68] van usar estadístics obtinguts de matrius de transició de veí més proper de nivell de gris per a mesurar aquesta dependència en imatges de satèl·lit de núvols i van identificar tipus de núvols a partir de la seva textura. [BAR69] i [WEI70] van utilitzar co-ocurrència uni-dimensional en aplicacions mèdiques. [ROS70] i [HAR71] proposaren dependències espacials bi-dimensionals dels nivells de gris en matrius de co-ocurrència per a cada distància fixada i/o relacions espacials angulars. [HAR73] va usar estadístics basats en matrius de co-ocurrència com a mesures de textura en imatges per satèl·lits, imatges aèries i

imatges microscòpiques. En [CHI74] es mostra l'aplicació de la co-ocurrència de nivells de gris per automatitzar l'anàlisi de raigs X al pit. [CON76] treballa a partir d'imatges radiogràfiques en aplicacions mèdiques. [PRE76] mostra l'aplicació per a la discriminació de cèl·lules cervicals. [CHE79] usaren la co-ocurrència juntament amb el procediment d'*split and merge* per a segmentar una imatge a partir de la seva textura. Tots els estudis donaren resultats raonables per a diferents textures usant la co-ocurrència de nivells de gris. A continuació s'explica resumidament com s'extreuen les característiques de co-ocurrència.

### 2.1.5.1. Matrius de co-ocurrència.

Suposem que l'àrea a analitzar per textura sigui rectangular i tingui  $N_c$  elements en direcció horitzontal,  $N_r$  elements en direcció vertical i el nivell de gris, per a cada element, es quantifiqui en  $N_g$  nivells.

Sigui:

$L_c = \{1, 2, \dots, N_c\}$  el domini de l'espai horitzontal  
 $L_r = \{1, 2, \dots, N_r\}$  el domini de l'espai vertical  
 $G = \{1, 2, \dots, N_g\}$  el conjunt d' $N_g$  nivells de gris.

El conjunt  $L_r \times L_c$  és el conjunt d'elements de la imatge ordenats per les seves designacions fila-columna. La imatge  $I$  pot ser representada com una funció que assigna algun nivell de gris de  $G$  a cada element o parell de coordenades del conjunt  $L_r \times L_c$ :

$$I: L_r \times L_c \rightarrow G$$

La co-ocurrència de nivells de gris es pot especificar com una matriu de freqüències relatives  $P_{ij}$  en la qual dos elements veïns separats una distància  $d$  estiguin en la imatge, un amb el nivell de gris  $i$  i l'altre amb el nivell de gris  $j$ . En general, s'obtenen una sèrie de matrius  $P_\theta(i, j, d)$ , on  $\theta$  és la direcció d'avaluació amb distància  $d$ . En lloc d'usar un únic desplaçament, s'utilitza un conjunt de desplaçaments per obtenir la propietat desitjada a la qual correspongui la característica de textura. Aquestes matrius de freqüències de dependència espacial de nivell de gris són simètriques i són funció de la relació angular entre elements veïns així com funció de la distància entre ells.

Les matrius de co-ocurrència de nivells de gris (GLCM) no s'usen generalment com a un vector de característiques. Cal transformar-les en característiques secundàries. [HAR73] calcula les següents característiques, a partir de les matrius de co-ocurrència, i realitza un nombre d'experiments d'identificació:

- Uniformitat de l'energia  $\sum_{i,j} P_{ij}^2$

- Entropia 
$$-\sum_{i,j} P_{ij} \log P_{ij}$$
- Màxima probabilitat 
$$\max_{i,j} P_{ij}$$
- Contrast 
$$\sum_{i,j} (i-j)^2 (P_{ij})$$
- Correlació 
$$\sum_{i,j} \frac{ijP_{ij} - \mu_x \mu_y}{\sigma_x \sigma_y}$$

on  $\mu_x$  i  $\sigma_x$  són la mitjana i desviació estàndard de la suma de les files de la matriu, mentre que  $\mu_y$  i  $\sigma_y$  ho són de les columnes.

- Moment de diferència inversa 
$$\sum_{\substack{i,j \\ i \neq j}} \frac{(P_{ij})^l}{|i-j|^k}$$
- Probabilitat d'un *run length*  $n$  per un nivell de gris  $i$  
$$\frac{(P_i - P_{ii})^2 (P_{ii})^{n-1}}{P_i^n}$$
  
on  $P_i = \sum_j P_{ij}$

La transformació en característiques secundàries condueix a una pèrdua significativa d'informació. En alguns casos, es desitja aquesta pèrdua d'informació, ja que les noves característiques redueixen la informació continguda en la imatge fins a propietats de textura. Hi ha autors, com ara [LOH94], que, tot i extreure característiques secundàries, continuen utilitzant la informació completa de les matrius de co-ocurrència. El nombre de matrius a calcular per a obtenir unes bones característiques de textura depèn de la relació entre l'angle i la distància i aquest nombre pot ser, potencialment, enorme. [TOU77] discuteixen un enfocament d'extracció de característiques basat en els vectors propis de la matriu per tal d'ajudar a solucionar aquest problema.

Els experiments de [WES76] fan pensar que les característiques espàcio-freqüencials  $i$ , per tant, la característica d'autocorrelació no són tant bones mesures de textura com les característiques de co-ocurrència. Es pensa que la raó per la qual les probabilitats de co-ocurrència tenen molta més informació que la funció d'autocorrelació és que tendeix a haver-hi restriccions naturals entre les probabilitats de co-ocurrència a una distància i les característiques d'autocorrelació a una altra distància. Per aquestes relacions, molta informació a una certa distància pot determinar la poca quantitat d'informació en la funció d'autocorrelació a diverses distàncies.

L'entropia o transformació conspícua va estar proposada també per [HAR75]. Aquesta és una manera de generar un pla de textura sense calcular les matrius de co-ocurrència per a cada punt. Les matrius es calculen per píxels en una gran àrea, a ser possible per la imatge sencera. La probabilitat de pertànyer un píxel a la mateixa regió de textura

que els seus veïns es calcula buscant el seu nivell de gris i els dels seus veïns dins les matrius. Les probabilitats, o alguna funció derivada, poden usar-se per a la segmentació de les textures. Els píxels comuns es treuen com a un sol segment i els estadístics de co-ocurrència es recalculen per els píxels restants. Els segments són identificats després sense necessitat de classificar píxels com a tipus de textura, com ho fa [OHL78]. Aquestes mesures de probabilitat són semblants a les transformades conspicues de [WIN78] i a les tècniques de predicció lineal de [DEG78].

[HEP94] calcula un vector de característiques de textura que representa la matriu de co-ocurrència d'ordre  $n$ -èssim. Per a cada píxel de la imatge es convolucionen quatre màscares amb diverses orientacions i segons els resultats, els ordena per ordre de major a menor, 4 elements donen 24 combinacions. Llavors per a cada píxel de la imatge obté un vector de 24 components. Això forma un espai de característiques sobre el qual es pot realitzar una classificació. En aquest cas, usa un algorisme de NNC (classificació per veí més proper).

Per a generar vectors de paràmetres a partir de matrius de co-ocurrència d'ordre  $n$ -èssim no cal obtenir les dades d'imatges multinivell. En particular, [PAY94] parteix d'imatges binàries per a trobar co-ocurrències i generar un vector d'estat per a cada píxel binari de la imatge. Després obté les regions homogènies de textura classificant aquests vectors. [HE87] proposa característiques de textura que involucren només un petit subconjunt de la matriu de co-ocurrència. El fet d'usar senzilles operacions aritmètiques sobre aquestes característiques permet que aquestes siguin trobades més ràpidament. Defineix tres característiques de textura: moment diagonal, moment d'alt nivell i moment de baix nivell.

#### 2.1.5.2. Matrius de co-ocurrència generalitzades.

Una altra tècnica és la que proposa [DAV79a], revisada i millorada en [DAV81b]. Aquest nou enfocament està basat en les matrius de co-ocurrència generalitzades (GCM). Per tal d'analitzar una textura, reemplaça la textura per una altra imatge que indica les posicions d'unes certes propietats locals en la imatge original, com ara els contorns. Normalment, aquesta segona imatge s'obté convolucionant la imatge textura amb un reduït conjunt de màscares i assignant descripcions de característiques locals als pics, o màxims locals, de les superfícies obtingudes després d'aplicar la convolució. Cada punt corresponent a un màxim local té alguna descripció associada amb ell. Per exemple, si les màscares fossin operadors de contorns, llavors la descripció podria incloure la magnitud i l'orientació del contorn. La matriu de co-ocurrència GLCM, definida normalment, no serveix com a eina per a descriure la distribució d'aquests màxims locals, perquè la majoria de màxims locals no tindran cap altre màxim a certa distància relativa.

[DAV79a] suggereix una generalització de la matriu de co-ocurrència per extreure algunes de les propietats espacials importants de la distribució de màxims locals i defineix aquesta matriu com el conjunt

$$P = \{p_i = (x_i, y_i, m_i)\}_{i=1}^n$$



on  $(x_i, y_i)$  és una posició de la imatge i  $m_i$  la descripció del  $i$ -èssim màxim local  $p_i$ . Se suposa que aquesta descripció ha estat quantificada d'alguna manera (normalment a partir d'operadors de convolució).

Sigui  $F$  un predicat de restricció espacial de la següent forma

$$F(p_i, p_j, k) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \leq k$$

Aquest predicat serà cert si i només si la distància entre  $p_i$  i  $p_j$  és menor o igual que  $k$ . Llavors, la posició  $(m, m')$  de la matriu,  $\mathbf{GF}$ , de co-ocurrència generalitzada (GMC) és el nombre de vegades que un parell  $(p_k, p_l)$  compleixen el predicat  $F$  i també,  $m$  i  $m'$  les descripcions de  $p_k$  i  $p_l$ , respectivament. Es poden definir diversos predicats de restricció espacial els quals incrementen la potència de les GMC.

De la mateixa manera que [HAR73] va descriure tota una varietat d'estadístics per a les GLCM, també es poden aplicar una sèrie d'estadístics diferents per a les GCM. Aquests estadístics són:

- Contrast generalitzat
- Uniformitat
- Entropia
- Correlació
- Relacions de característiques derivades de diferents GCM (aplicant diferents predicats de restricció espacial).

Aquestes característiques es poden aplicar en:

1. Detecció de grups o classes a partir de les GCM: [ZUC76] examina els histogrames de màxims locals per tal de segmentar una imatge composta de diverses textures però en regions individuals. D'aquesta mateixa manera, es podria buscar, en les GCM, pics o màxims locals per a segmentar una imatge. Aquesta segona manera seria molt més precisa que la segmentació derivada de les anàlisis dels histogrames.
2. Detecció de contorns de textura: aquest punt és el més interessant i es pot realitzar a dos nivells. Un és comparant directament les GCM amb alguna mètrica definida per a matrius; de fet, [CHE79] proposa quelcom semblant per a les GLCM per tal de segmentar imatges que contenen regions de textura usant una estructura de dades piramidal. L'altre nivell seria comparar les característiques derivades de les GCM.

[DAV79a] aplica tres mètodes diferents per a la segmentació d'imatges a partir de les característiques que obté de les GCM. Usa primer un algorisme d'agrupament clàssic, el veí més proper; després classifica les textures amb funcions de discriminants lineals i finalment, usa la mesura de distància de Battacharyya. Com es veu a continuació, aquest autor usa aquesta mesura de distància entre descriptors basant-se en diferents característiques.

### 2.1.5.3. Polarograms.

Per a l'anàlisi de textures en imatges, [DAV81a] va desenvolupar una eina anomenada *polarogram*. Un polarogram és una representació polar d'un estadístic específic d'orientació d'una textura i va ser àmpliament utilitzat per a obtenir informació sobre la gruixudesa i direccionalitat d'una textura. Es calcula a partir d'una característica obtinguda d'una matriu de co-ocurrència. En concret, es fan dos experiments.

1. A partir d'un conjunt d'imatges, es calculen tres polarograms del descriptor de contrast de les matrius de co-ocurrència de nivell de gris. Es calculen les matrius per a les vuit direccions principals i per a distàncies entre punts d'1, 3 i 5. Es calculen també la matriu de co-ocurrència generalitzada (GCM), basada en tots els parells de 8 punts adjacents. Tant per les GLCM com per a les GCM, es troben quatre estadístics: contrast, uniformitat, entropia i correlació (definites anteriorment). Per a classificar les textures, a partir de tots els descriptors calculats, aplica una funció lineal discriminant multivariant.
2. El segon experiment es basa en la mesura de distància de Battacharyya per a la separabilitat de classes de textura. És un mètode proposat per [FAU80] per a comparar descriptors de textura i té diversos avantatges en front dels mètodes de classificació comparativa, incloent els pocs requeriments de dades a processar per aconseguir resultats fiables. Els resultats obtinguts a partir de la distància de Battacharyya entre parelles de classes de textura, donen millors resultats per als polarograms que per a les matrius de co-ocurrència de nivells de gris i per a les matrius de co-ocurrència generalitzades.

### 2.1.6. Mètodes de transformada digital de textura.

En el mètode de transformada digital per a l'anàlisi de textures, la imatge digital és, normalment, dividida en un conjunt de petites subimatges quadrades no solapades. Imaginem que la mida de la subimatge és de  $n \times n$  elements, llavors els  $n^2$  nivells de gris de la subimatge es poden considerar com les  $n^2$  components d'un vector  $n^2$ -dimensional. El conjunt de subimatges constitueixen llavors un conjunt de vectors  $n^2$ -dimensional. En la tècnica de transformació, cadascun d'aquests vectors es reescriu en un nou sistema de coordenades. La clau de la transformació és que la base del nou sistema de coordenades té una interpretació que es relaciona amb la freqüència espacial o seqüència  $i$ , a l'estar la freqüència estretament lligada a la textura, aquestes transformades poden ser útils per a l'anàlisi de textures.

Una imatge pot ser transformada per una o més de les següents raons:

- per permetre un emmagatzematge i representació més eficient (compressió d'imatges),
- per accentuar i remarcar les característiques per a la classificació,
- per permetre algun tipus de filtratge que es pot aconseguir més eficientment en l'espai transformat. En aquest cas, la imatge és transformada, filtrada i llavors es torna a transformar a l'espai original.

Per seleccionar la transformada apropiada, és important trobar el compromís entre la bondat de la representació de les dades en el nou espai i el temps de càlcul requerit.

De totes les transformades digitals, les més destacables que es poden aplicar a l'anàlisi de textura són:

- Transformada de Fourier
- Transformades *Wavelets*.

A continuació, veiem amb més detall cadascuna d'aquestes transformades i la seva relació amb la descripció i tractament de textures.

### 2.1.6.1. Transformada de Fourier.

Una imatge digitalitzada pot ser considerada com una funció real de dues variables,  $g(x,y)$ , on  $x$  i  $y$  són nombres naturals que indiquen la mida de la imatge ( $x$  i  $y$  són les coordenades cartesianes d'un punt de la imatge). La transformada discreta de Fourier bi-dimensional de  $g(x,y)$  és

$$F(u,v) = \frac{1}{p^2} \sum_{x=0}^{p-1} \sum_{y=0}^{p-1} g(x,y) \exp[-2\pi i(xu + yv)/p]$$

on  $p$  és la dimensió de la matriu imatge que suposem quadrada.

L'espectre de potència de  $F(u,v)$  és la magnitud d'aquesta funció complexa i ve donada per

$$p(u,v) = [F_{\text{Re}}^2(u,v) + F_{\text{Im}}^2(u,v)]^{1/2}$$

La fase de l'espectre és

$$\psi(u,v) = \arctan[F_{\text{Im}}(u,v)/F_{\text{Re}}(u,v)]$$

L'operador de Fourier té propietats que poden ser utilitzades per a l'anàlisi de textures. La transformada de Fourier és simètrica respecte a l'origen del pla complex si la funció original és real. L'espectre de potència és invariant a translacions en el domini espacial però no és invariant a rotacions. Això implica que la direccionalitat d'un patró en la imatge es conserva en  $p(u,v)$ . La transformada de Fourier comprimeix dades si són periòdiques o quasi periòdiques, característica pròpia de la textura. La fase conté informació de la posició dins la finestra de la imatge i tota la informació que ens podria donar en el processat de textures no es recull en l'espectre de potència, ja que la informació de la fase es perd.

En el treball de [BAJ76], anteriorment descrit per [BAJ73b], es desenvolupa una tècnica en la qual es generen descriptors de textura, basats en finestres locals, a partir

de les característiques de l'espectre de potència de la transformada de Fourier de la imatge.

Un altre conjunt de característiques varen ser extretes per [WES76] a partir de la transformada discreta de Fourier; de fet, [LEN70] va ser el primer investigador en postular el conjunt de característiques que més tard utilitzen i milloren tant [WES76] com [CON80]. La distribució radial de  $p$ , l'espectre de potència, és sensible a textures gruixudes en la imatge. Una textura gruixuda tindrà alts valors de  $p$  concentrats prop de l'origen, mentre que en una textura fina els valors de l'espectre estaran més desplegats. Llavors, per a analitzar el gruix d'una textura (que en definitiva no és més que la periodicitat), un conjunt de característiques que serien d'utilitat són les mitjanes de l'espectre de potència al llarg de regions en forma d'anell centrades a l'origen de la forma

$$\phi_{r_1 r_2} = \sum_{\substack{r_1^2 \leq u^2 + v^2 < r_2^2 \\ 0 \leq u, v \leq p-1}} |p(u, v)|^2$$

per a diversos valors de  $r_1$  i  $r_2$ , cercle interior i exterior, respectivament. De manera semblant, la distribució angular dels valors de l'espectre de potència és sensible a la direccionalitat de la textura en la imatge original. Una textura amb molts contorns o línies, en una direcció donada  $\theta$ , tindrà valors alts de  $p$  concentrats al voltant de la direcció perpendicular  $\theta + (\pi/2)$ , mentre que en una textura no-direccional l'espectre també serà no-direccional. Llavors, un bon conjunt de característiques per analitzar la direccionalitat d'una textura seran les mitjanes de l'espectre de potència al llarg de regions en forma de falca acabades a l'origen de la forma

$$\phi_{\theta_1 \theta_2} = \sum_{\substack{\theta_1 \leq \tan^{-1}(v/u) < \theta_2 \\ 0 < u, v \leq p-1}} |p(u, v)|^2$$

per a diversos valors de  $\theta$  que marquen la inclinació de la falca. Cal remarcar que la component de contínua,  $(u, v) = (0, 0)$ , s'ha omès ja que és comuna a totes les falques. Aquestes característiques són sensibles només a la mida (freqüència espacial), o només a l'orientació, però no a totes dues alhora.

[CON80] utilitza una característica més que [WES76]. A part de calcular les mitjanes en regions anulars i regions en forma de falca, també ho fa en regions en forma de talls paral·lels, però a partir de l'espectre de potència en coordenades cartesianes i no polars, tal com s'han extret les altres dues característiques. Aquest operador és útil per a distingir la regularitat d'una textura. [CON80] aplica aquests operadors a la discriminació de textures a partir d'imatges aèries, juntament amb d'altres operadors calculats per diferents mètodes (mètode de *run length* de nivells de gris, ...). Els resultats són semblants als obtinguts per [WES76], o el que és el mateix, resultats pobres comparats amb d'altres mètodes.

En un treball recent, [MCL93] usa quatre característiques extretes a partir de la potència local. Aquest conjunt dona bona discriminació de textures fins i tot en presència de soroll. En primer lloc es normalitza l'espectre de potència i s'obté  $p(u, v)$ , el qual té una certa distribució de probabilitat. Les quatre característiques són les següents:

1. Energia de pic:

$$f_1 = [p(u_1, v_1) \times 100]$$

on  $u_1, v_1$  són les coordenades en l'espai freqüencial del màxim pic de l'espectre de potència.

2. Percentatge d'energia, quadrant 1

$$f_2 = \sum_{u>0} \sum_{v>0} p(u, v)$$

3. Percentatge d'energia, quadrant 2

$$f_3 = \sum_{u<0} \sum_{v>0} p(u, v)$$

4. Entropia relativa

$$f_4 = \frac{\left[ - \sum_{u=-u_m/4}^{u_m/4} \sum_{v=-v_m/4}^{v_m/4} p_1(u, v) \log p_1(u, v) \right]}{\log \left[ \frac{u_m}{2} \frac{v_m}{2} \right]}$$

on  $u_m$  i  $v_m$  són el nombre màxim de freqüències en la imatge transformada. Per a normalitzar aquesta característica s'usa una funció diferent a la funció que normalitza les altres característiques.

Les característiques extretes amb la transformada de Fourier es discuteixen més àmpliament en [LIU90]. Segons els autors, l'enfocament de freqüència espacial és molt més senzill que el mètode GLCM (co-ocurrències de nivell de gris) usant només una simple transformada i quatre característiques calculades a partir de l'espectre de potència. L'ús de característiques de textura basades en l'espectre de potència de Fourier per a textures naturals donen resultats inferiors a característiques modelitzades estadísticament en el domini espacial.

#### 2.1.6.2. Transformada Wavelet.

Tal com s'ha vist anteriorment, un dels primers mètodes de descomposició de senyals és la transformada de Fourier que usa complexos sinusoides per a representar-los. Això s'aconsegueix per les propietats ortogonals de la família de les funcions base obtingudes a partir de les funcions sinus i cosinus amb duració infinita. Però el funcionament de les transformades de Fourier es deteriora quan el senyal conté transicions en la forma o components transitòries, les quals es consideren localitzades en el domini espacial. Per solucionar aquest problema, s'usa la STFT (*short-time Fourier Transform*) la qual implica multiplicacions d'una funció finestra pel senyal

original. En aquesta transformada la mida de la finestra és fixa. La transformada *wavelet*, a l'igual que la STFT, dona localització tant en el domini espacial com en el freqüencial. No obstant, la transformada *wavelet* dona un avantatge molt important respecte a la STFT: la mida de la finestra varia segons el contingut freqüencial de la imatge. Una funció arbitrària  $f$  pot ser representada com a una superposició de funcions bàsiques anomenades *wavelets*. Al contrari que les funcions sinus i cosinus de la transformada de Fourier, els *wavelets* tenen una durada finita la qual cosa dona localització espacial i freqüencial. Per explicar l'impacte dels *wavelets* en la visió per computador a baix nivell, cal fixar-se en tres grans aplicacions: processat multiresolució, detecció de contorn multiescala i discriminació de textures.

- Els algorismes multiresolució modifiquen la resolució de la imatge per processar les dades a diversos nivells de resolució per a una tasca en particular. Per dur a terme el ventall des de l'anàlisi gruixuda (*coarse*) fins a l'anàlisi fina (*fine*), es processa la imatge a baixa resolució i es va ampliant selectivament la zona d'informació a tractar. Aquests algorismes tenen clara aplicació a l'estèreo-visió i a les mesures de flux òptic.
- Sovint, els contrastos locals en la imatge contenen més informació que els valors d'intensitat lluminosa. Una transformada *wavelet* mesura variacions en el nivell de gris de la imatge a diferents escales. Els contorns de les estructures de la imatge corresponen a contrastos de formes i poden ser detectats a partir dels màxims locals de la transformada. Això obre la porta a algorismes de reconstrucció els quals poden reconstruir imatges a partir de contorns multiescala.
- Un dels problemes complexos en visió per computador és la discriminació de textures. Malgrat que les textures són ràpidament discriminades per l'observador humà, no existeix encara un model apropiat per les textures [MAL96]. La percepció de les textures depèn de propietats locals però no existeix una mida predefinida dels veïns sobre els quals la textura pot ser analitzada. Això motiva l'ús de les transformades *wavelets*, on es mesuren les propietats de la imatge sobre dominis de mida variant. Per a la discriminació de textures semblen importants les mesures de freqüència local derivades d'una transformada *wavelet* direccional [SUP95][JAI91][POR89]. El treball de [LEE92] minimitza funcionals d'energia en vectors amb components formats per *wavelets* de Gabor de quadratura. Encara no existeix cap teoria comprensiva que guiï les segmentacions de textura a partir dels coeficients dels *wavelets*.

Quan s'estudia l'aplicació de *wavelets* a la visió per computador, les majors dificultats es troben en la interfície entre els algorismes de baix nivell i els models visuals d'alt nivell. Les estratègies de cerques multiresolució han de dependre d'un coneixement *a priori* del món. De manera semblant, la detecció de contorns no es pot restringir a un processat orientat a punts tal com ens semblaria des del nostre punt de vista de contorn. Finalment, la discriminació de textures també requereix l'elaboració de models *a priori* els quals han de guiar els procediments d'agrupament per a les segmentacions d'imatges. Aquesta nova tècnica d'anàlisi de dades està creixent espectacularment i ja existeixen molts treballs en aquesta via, [UNS95][SAL95].

També molts paquets de programes estan incorporant aquesta tècnica en el seu programari (MatLab, S-Plus, Mathematica, PV-Wave) [SCW96].

Cal remarcar que són possibles altres transformacions digitals de la imatge. Els únics requeriments són que han de representar una imatge eficientment i que han de ser ràpides de calcular. La utilitat d'una imatge transformada depèn de quant bé representa les característiques importants de la imatge i quant eficientment es calcula. [AND71] compara les transformades de Fourier, Walsh/Hadamard, [HAD93], i Haar, [HAA10], en un experiment de classificació. Per a la comparació, selecciona les característiques en base a la variància decreixent. Les transformades de Fourier i Walsh/Hadamard són comparables (en quant a eficiència en la classificació, no en càlcul) i considerablement millors que la transformada de Haar.

### 2.1.7. Característiques perceptives de textura.

En l'anàlisi de textures de forma paramètrica, es pretén obtenir un conjunt de característiques amb les quals es mesurin tots els patrons d'entrada i amb els quals es donin resultats ben distribuïts. [TAM78] proposa un conjunt de 6 característiques extremes de textura que corresponen a la percepció visual humana. Tal com es diu en el treball de Tamura *et al.*, aquest conjunt perceptiu és molt útil per a la selecció òptima de característiques i per al disseny d'analitzadors de textures. Els sis paràmetres que proposa són:

- Gruixària: El gruix d'una textura és una de les seves característiques fonamentals i ja ha estat estudiat des de fa bastant temps ([HAY74]).
- Contrast: Quan dos patrons es diferencien només en la distribució del nivell de gris, es pot mesurar aleshores el seu contrast. S'assumeixen els següents quatre factors per a mesurar el contrast:
  1. rang dinàmic dels nivells de gris,
  2. polarització de la distribució del blanc i negre en l'histograma ; o la relació de les àrees blanques i negres,
  3. agudeses en les arestes, i
  4. període de la repetició de patrons
- Direccionalitat: Aquesta és una propietat global sobre una regió donada. La direccionalitat involucra tant la forma de l'element com el seu emplaçament. [BAJ73a] divideix la direccionalitat en dos grups, la monodireccionalitat i la bidireccionalitat. Aquí l'orientació del patró no és rellevant.
- Rectilinietat: aquest concepte només fa referència a la forma de l'element de textura on es valorarà si està format per línies.
- Regularitat: és una propietat per a mesurar les variacions d'una regla d'emplaçament.
- Asprositat: aquesta descripció era utilitzada originalment per a descriure textures tàctils, no per a visuals. Quan s'observen textures naturals, som capaços de comparar-les en termes d'asprositat i suavitat.

Tamura *et al.* compara les mesures extretes computacionalment sobre textures bàsiques naturals amb els resultats experimentals realitzant un test sobre un conjunt d'individus. Busca la correspondència entre l'experiment psicològic i el resultat computacional per confirmar l'aspecte perceptiu de les característiques de textura.

Per a descriure una àrea local de textura, [BAJ73a] usa una llista de descriptors com ara la direccionalitat, la mida de la separació entre dos elements de textura, l'amplada d'un element de textura  $i$ , en el cas de ser direccional, la seva orientació. Tots aquests descriptors s'usaran més endavant per formar regions amb descriptors semblants. La formació de regions, en el seu primer pas, és un procés a baix nivell que Bajcsy anomena *topological sheaves* o model matemàtic per formalitzar el procés de creació de regions.

Un element de textura s'aproxima per una regió amb nivells de gris homogenis. A partir d'aquesta definició, [TOM82] extreu unes propietats per a cada element de textura: il·luminació, àrea, mida (2 vegades l'àrea pel perímetre), elongació i curvatura. Si existeixen grups en la distribució de les propietats dels elements, aleshores aquests grups representen diferents classes d'elements. Els elements es classifiquen per tècniques de llindar recursiu sobre els histogrames de les propietats [TOM73] i [TSU73].

La línia d'investigació d'aquesta tesi es considera un mètode més dins aquest apartat de caracterització paramètrica de textura [CAS92b]. El conjunt de característiques de textura extret de la imatge original entra dins les característiques perceptives i s'obtenen els següents paràmetres:

- linealitat,
- abruptitat,
- granulositat,
- difuminat i
- discontinuïtat.
- curvitat.

Aquests paràmetres seran definits i àmpliament detallats en seccions posteriors. L'algorisme de classificació crea grups a partir de les mostres multiparamètriques que s'introdueixen al classificador. L'algorisme usa la mètrica euclidiana de distància.

### **2.1.8. Altres mètodes.**

Per obtenir un espai de característiques de textura, no cal partir de paràmetres perceptius. Un dels estudis més clàssics d'extracció de característiques de textura és el de Haralick, [HAR73], on s'arriben a proposar fins a 28 característiques de textura a partir de cada matriu de dependència espacial dels nivells de gris, com ara el segon moment angular, el contrast, la correlació, la variància, el moment de diferència inversa, la suma de mitjanes, la suma de variàncies, l'entropia, el coeficient de màxima correlació i altres. Segons el tipus de textures que es volien classificar, s'escollien unes característiques o unes altres per formar el vector de característiques de textura. Per



exemple, en el reconeixement d'imatges aèries, usava 11 característiques extretes de matrius de dependències de nivells de gris en quatre direccions; això fa un total de 44 components per a cada vector. L'algorisme de classificació era en aquest cas la regla de decisió *Min-Max*, sobre els valors màxim i mínim de la distribució uniforme de les mostres.

Recentment ha aparegut un nou model d'anàlisi de textura basat en el concepte de la unitat de textura, on una imatge pot ser caracteritzada pel seu espectre de textura [WAN90]. El concepte bàsic és que una textura pot ser descomposta en un conjunt d'unitats essencials petites, que s'anomenen unitats de textura. Cada unitat es representa amb 8 elements, on cadascun pot prendre tres possibles valors (0, 1 i 2) obtinguts amb un veïnatge de 3x3 píxels. En la següent figura (fig. 2.4) es pot veure la transformació d'un veïnatge en una unitat de textura.

Veïnatge		
63	28	45
88	40	35
67	40	21

Unitat de textura		
2	0	2
2		0
2	1	0

**Figura 2.4.** Exemple de creació d'una unitat de textura.

Es fixen dos llindars per crear els tres valors de la unitat de textura a partir de l'element central del veïnatge. Això condueix a  $3^8=6561$  unitats de textura estàndards. Com la unitat de textura representa l'aspecte local de la textura, els estadístics que s'en derivin sobre una imatge revelaran la seva informació de textura. La distribució d'ocurrència de les unitats de textura s'anomena espectre de textura on les abscesses indiquen el tipus d'unitat de textura i les ordenades representen la freqüència d'ocurrència. L'autor també proposa un mètode de classificació basat en l'espectre de textura.

Basant-se en aquest treball, ha aparegut una versió binivell proposada per [OJA94] i que ell mateix anomena patrons locals binaris (PLB). En aquesta versió només hi ha  $2^8=256$  unitats de textura possibles i el valor de l'únic llindar es fixa amb el valor central del veïnatge. Aquests mètodes són invariants a l'escala de nivells de gris i poden ser fàcilment combinables amb d'altres característiques, com ara el contrast o la covariància, per donar-li més potència alhora de discriminar textures.

Existeixen altres mètodes que basen l'extracció de característiques de textura en la dimensió fractal de la imatge. La dimensió de Hausdorff-Besicovitch (HB) o dimensió fractal d'un conjunt afitat  $A$  en  $\mathbb{R}^n$  és un nombre real que s'usa per a caracteritzar la complexitat geomètrica d' $A$ . Llavors, un conjunt es diu que és un conjunt fractal si la seva dimensió HB és estrictament superior a la seva dimensió topològica. [CHA95] proposa sis característiques basades en la dimensió fractal de: 1) la imatge original; 2) els valors de gris alts de la imatge; 3) els valors de gris baixos de la imatge; 4) la imatge suavitzada horitzontalment; 5) la imatge suavitzada verticalment i 6) la dimensió multifractal de la imatge.

[PEL84] extreu la signatura fractal d'una textura i compara diverses signatures per a classificar textures. La signatura fractal d'una textura diu que és el canvi en la dimensió fractal en una àrea variant l'escala.

Hi ha diversos autors que han inclòs el tractament del color per al reconeixement de textures. En particular, [GAG86b] intenta trobar models eficients per a textures en color. Els paràmetres d'aquests models es calculen a partir de camps de textures en color usant un algorisme d'agrupament dinàmic que substitueix les coordenades (R,G,B) de cada píxel per un simple codi. Aleshores, la mida dels models obtinguts és la mateixa que per textures en tons de gris. El problema és que l'agrupament per píxels propers dins el cub RGB no determina realment semblança de colors. Per una altra banda, [PAN95] introdueix models de camp aleatori de Markov per a textures en colors. Aquests models caracteritzen una textura en termes d'interacció espacial entre cada pla de color i d'interacció entre diferents plans de color. Per a la segmentació, usen un algorisme no supervisat d'agrupament jeràrquic aglomeratiu dirigit pel criteri basat en la probabilitat de la imatge. Per aquest motiu, consideren mètodes de màxima probabilitat per a l'estimació dels paràmetres. Un problema important és la selecció dels veïns durant el disseny dels models de color dels camps aleatoris, donat que poden aparèixer un gran nombre de possibles paràmetres per definir interaccions dins i entre les bandes de color.

Un mètode per crear models de textura són els mosaics aleatoris. Aquests models es construeixen en dos passos: un primer pas dona els mitjans per a la *tesselació* d'un pla en cel·les i el segon pas assigna un determinat valor a cada cel·la. Aquest mosaic representarà una regió de textura. Existeixen tota una sèrie de processos per a generar els mosaics aleatoris [YOU86]. En el model de la línia de Poisson, el pla es divideix per un conjunt aleatori de línies. Cada cel·la és, aleshores, una regió connectada i on els seus límits són segments que pertanyen al conjunt aleatori. En el model d'ocupació, es produeix una divisió per un procés aleatori en el qual es planten punts en el pla, llavors cadascun d'aquests punts determina una cel·la consistent en tots els punts del pla més propers a ell que no pas a qualsevol altre dels punts plantats. En el model de Delaunay, entre cada parell de punts plantats es dibuixa una línia, les seves cel·les corresponents en el model d'ocupació comparteixen un segment de frontera comú. D'altres serien el model de Johnson-Mehl, el model de tauler d'escacs rotat o els models de bomba. [SCH78] deriva propietats estadístiques per aquests models aleatoris de mosaic. [AHU81] compara propietats de textures generades sintèticament amb els seus valors teòrics. [SCH80a] analitza com les característiques de textura tenen relació amb els variograms (gràfiques entre la diferència al quadrat esperada de nivell de gris en funció de la distància) i amb la funció de correlació. [MOD81] calcula les funció de densitat de potència espectral per un pla *tesselat* per un procés aleatori de línies i en el qual els nivells de gris d'una cel·la tenen una dependència de Markov amb els nivells de gris de les cel·les veïnes. [THE83] usa un model autoregressiu per a cada cel·la i, a l'igual que Modestino *et al.* en [MOD81], imposa un model de camp aleatori de Markov per a descriure les transicions entre les cel·les.

Dins la morfologia matemàtica, [MAT67] i [SER73] proposen l'ús d'un element estructural i un filtratge per a textures en imatges binàries. La seva idea bàsica era definir un element estructural com a un conjunt de píxels constituint una forma

específica com ara una línia o un quadrat i generar una nova imatge binària movent l'element estructural a través de la imatge i erosionant les figures formades per píxels que tinguin el valor 1. Les característiques de textura s'obtenen ara de la nova imatge comptant el nombre de cel·les que tenen el valor 1. Aquest enfocament va trobar una àmplia aplicació en l'anàlisi qualitatiu de microestructures de material i biologia.

Un altre treball interessant de morfologia matemàtica és el presentat en [VAN96], on es calculen corbes granulomètriques sobre un conjunt d'imatges de radiografies de tòrax. Les corbes es calculen a partir de diferents obertures algebraïques amb elements estructurants rodons, erosions amb elements rodons seguides de reconstrucció i finalment l'obertura resultat de fer el suprem d'obertures morfològiques amb elements lineals orientats en quatre direccions.

## **2.2. Enfocament estructural.**

Amb l'enfocament estructural s'entén la textura com una disposició en l'espai d'un conjunt de subpatrons d'acord amb unes certes regles d'emplaçament. Els subpatrons, a la vegada, estant formats generalment per subpatrons més petits, posicionats també amb unes certes regles, ja siguin iguals o diferents. Aquesta natura recursiva de l'enfocament permet obtenir una estructura jeràrquica de les escenes naturals. Tant els subpatrons com les seves regles d'emplaçament poden ser caracteritzats estadísticament.

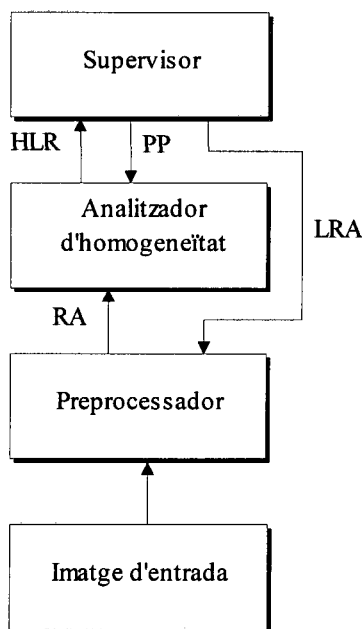
A continuació es fa un breu resum d'alguns mètodes de caracterització estructural de textures.

### **2.2.1. Recull de mètodes**

Un dels primers treballs amb mètodes estructurals fou presentat per [TSU73] *et al.* En aquest treball es proposa un mètode per a la partició d'escenes en regions on es troben algunes propietats repartides uniformement. En aquest procediment, s'examina, a partir de l'anàlisi estructural, la uniformitat o homogeneïtat de propietats complexes de la imatge com ara la forma, la mida o l'arranjament de subpatrons en les imatges. El procés d'anàlisi és supervisat i es controla l'interès de cada propietat en la imatge. Es realitzen successives particions de la imatge amb l'ànim de trobar una partició interpretable des del punt de vista humà.

En la figura 2.5 es mostra l'analitzador estructural d'escenes. El preprocessador busca regions atòmiques en les imatges digitalitzades. Una regió atòmica es defineix com un conjunt connectat d'elements de la imatge que tenen, al menys, el mateix nivell de gris. S'assumeix que cada regió atòmica té un contorn clar. En aquest treball no es dóna gaire importància a trobar les regions en sí mateixes, sinó que es centra l'interès en detectar els contorns d'àrees connectades en les quals els nivells de gris de cada punt siguin menors que el valor del llindar. Aquest valor del llindar es decideix a partir de l'histograma dels nivells de gris. Un cop les regions atòmiques han estat localitzades pel preprocessador, es passa a l'analitzador d'homogeneïtat tota la

informació corresponent a aquestes regions. Per a cada regió atòmica, el preprocessador passa a l'analitzador d'homogeneïtat la mitjana de nivell de gris de cada regió i informació dels límits de la regió. El primer que s'analitza són els descriptors de la imatge i són calculats per a cada regió atòmica. Es mesuren la forma, la mida i la posició de les regions. L'àrea i el perímetre venen especificats per la mida del descriptor i el descriptor de posició dona les coordenades bi-dimensionals del centre de gravetat de cada regió.



**Figura 2.5.** Diagrama de l'analitzador d'escenes.

Les abreviatures tenen el següent significat:

- RA és la llista de regions atòmiques
- HLR és l'histograma de la llista de regions
- PP és la propietat de la imatge seleccionada
- LRA és el llindar per trobar les regions atòmiques

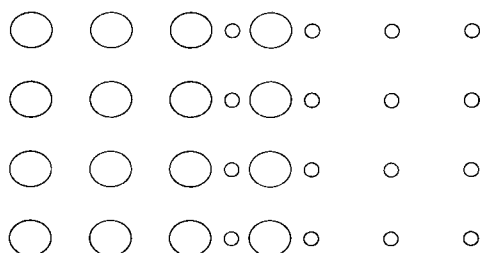
A causa de les limitacions en la resolució en les imatges d'entrada, no es tenen formes acurades de les regions atòmiques. Conseqüentment, el descriptor de la forma ha de ser seleccionat de manera que doni informació gruixuda sobre la forma però que no sigui sensible al soroll afegit en l'obtenció de la imatge. Per a satisfer aquest criteri, s'adopten una combinació de moments d'ordre  $i,j$ -èssims,  $M_{i,j}$ . Els moments utilitzats són  $M_{2,0}$ ,  $M_{0,2}$ ,  $M_{2,1}$  i  $M_{1,1}$ . S'han escollit aquests moments en particular perquè van ser adequats en una sèrie d'experiments per a la detecció de diferències de formes en petits objectes.

Els moments, al voltant del centre de gravetat  $(x_c, y_c)$  d'una regió R, es defineixen com

$$M_{i,j} = \sum_{x \in R} \sum_{y \in R} (x - x_c)^i (y - y_c)^j$$

L'ús de moments d'ordre  $i, j$ -èssims per a descriure informació sobre la forma no és nou. La generalització d'aquest mètode i l'ús del conjunt de moments invariants és descrit per [HU62] i [DUD73]. Una motivació per usar moments invariants per descriure formes ja apareix inicialment en [DEU55].

La següent figura és un exemple en el qual hi ha dues classes de regions atòmiques.



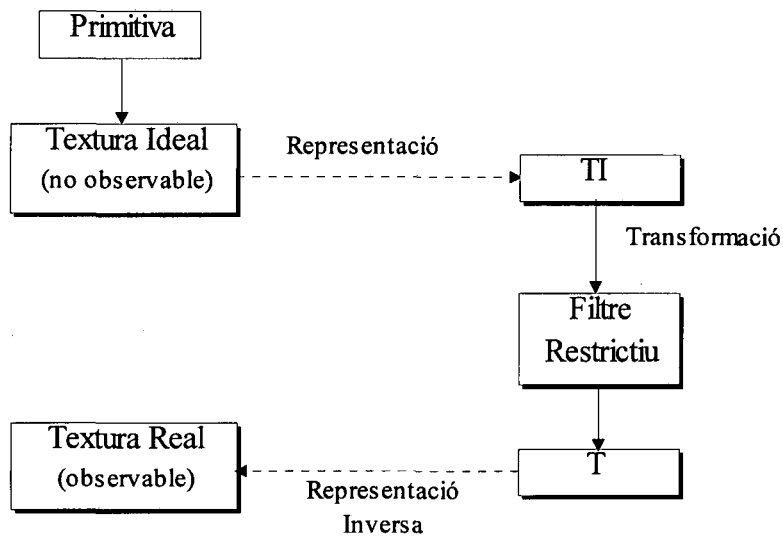
**Figura 2.6.** Exemple de tres regions construïdes amb dos conjunts.

Sigui  $R_1$  la regió coberta per punts petits i  $R_2$  la regió coberta per punts grossos. Al tenir una mida considerable l'àrea comuna  $R_1 \cap R_2$  de les dues regions, sembla raonable per l'analitzador considerar tres regions en lloc de dues. Les tres regions estan definides per  $R_1 \cap R_2$ ,  $R_1 \cap R'_2$  i  $R'_1 \cap R_2$ . En altres paraules, el sistema determina quan dues regions tenen parts solapades. Seguidament, es passa un test d'afitament per si alguna regió atòmica, en el grup de les regions aïllades, està situada entre dues regions o envoltada per alguna regió. Si només una regió toca alguna regió de les aïllades, aquestes s'ajunten per a formar una única regió. No obstant, si dues o més regions toquen una regió atòmica del grup de les aïllades, aquesta es queda en el grup de les aïllades. Això es fa perquè no existeix un mètode senzill per a decidir a quina de les regions col·lindants s'hauria d'integrar la regió aïllada.

S'envien al supervisor les regions descrites pels seus afitaments i els seus elements. El supervisor posa línies rectes als límits de les regions per obtenir una representació visual. Si alguna part de les línies no concorda amb el model, el supervisor considera que el resultat ha estat insatisfactori i caldrà realitzar una nova partició. El procés iteratiu de partició s'acaba quan alguna part de la regió es troba que concorda amb l'estructura del model o quan no es possible més crear més particions.

Aquest mètode d'anàlisi estructural dona resultats satisfactoris per a escenes d'entrada simples. No obstant, calen certes modificacions abans no s'analitzin correctament escenes complexes.

Un altre dels treballs pioners en aquest tipus d'anàlisi és el de [ZUC76]. En el seu treball desenvolupa un model de textura que té tres components principals: primitives, textures ideals i textures observables. En la figura 2.7 es poden veure les relacions entre elles.



**Figura 2.7.** Principals components del model i les seves interrelacions ([ZUC76]).

Bàsicament, les primitives locals es combinen d'una manera perfectament ordenada (concatenades) per tal de construir una textura ideal global. Aquesta textura no és observable, és altament estructurada i amb un patró perfecte, o sigui, és una imatge imaginària que no pot ser representada directament. Per aquest motiu, es desenvolupa una forma de representació per a textures ideals: un graf regular en el qual cada vèrtex es connecta als seus veïns d'una manera idèntica. Aquesta forma és la principal característica en el nivell de textura ideal. Aleshores, es defineixen transformacions per passar d'aquesta representació a la representació de textura real (o superfície). Aquestes transformacions faran observables els patrons no observables.

La classe de transformacions definides en una TI donada (Textura Ideal) generen moltes textures, cadascuna de les quals es consideren derivades de la mateixa TI. Per tal de limitar el poder generatiu del nivell transformacional, les transformacions no han de violar un conjunt de restriccions. Aquestes restriccions estan derivades a partir de les propietats matemàtiques del model, així com de dades psicofísiques extremes en la percepció de textures. Aquestes transformacions s'han d'entendre com a generació d'un conjunt de textures les quals són vistes com a versions distorsionades d'una textura ideal donada. També algunes textures són derivades a partir de més d'una textura ideal. Les restriccions són realment un conjunt de regles, tot i que en la figura anterior apareguin esquemàticament com a un filtre.

Cal dir que, posteriorment, Zucker relata les excel·lències dels mètodes estadístics treballant amb matrius de co-ocurrència en [ZUC81].

[LU78] proposa un model sintàctic per a la generació de textures estructurades així com per a la discriminació de textures. En primer lloc, es divideix els patrons de textura en finestres de mida fixa. Les finestres que pertanyen al mateix patró de textura es caracteritzen amb un arbre gramatical. Aquest arbre s'usa tant per a la síntesi com per a la discriminació de textures. Per a la discriminació, es poden aplicar automats per a la correcció d'errors construïts segons les gramàtiques creades a partir dels patrons de textura. També s'han suggerit gramàtiques estocàstiques en forma

d'arbre per al modelitzat de patrons de textura sorollosos i distorsionats [LU79][FU80].

D'altres autors, com ara [MAT82], proposen també analitzadors estructurals. En particular, aquest autor proposa un mètode per a l'anàlisi estructural de textures arranjades regularment. A partir d'una imatge de textura, primer extreu els elements de textura amb un procediment de creixement de regions (*region growing*), sabent que una textura conté moltes regions amb propietats semblants corresponent a elements de textura. En el procés de descripció dels arranjaments espacials entre els elements de textura extrets, primer s'extreuen els vectors de regularitat a partir d'un conjunt de vectors de posició relativa entre els elements de textura. Un vector de regularitat indica un vector relatiu el qual apareix sovint en l'arranjament dels elements de textura. Amb l'ús d'aquests vectors, es poden estimar localitzacions d'elements perduts així com també es poden eliminar elements erronis que apareixen en el processament de la imatge. Finalment, es busca la descripció més simple de l'arranjament dels elements de textura en termes de vectors de regularitat.

### **2.2.2. Problemes associats a l'enfocament estructural.**

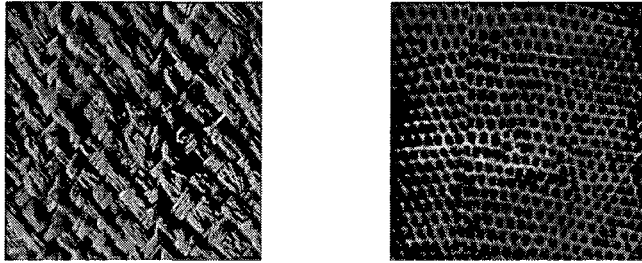
Com s'ha vist anteriorment, en aquest enfocament, una textura es defineix d'acord amb unes regles d'emplaçament ben definides i per subpatrons que apareixen repetidament dins tot el patró. Els exemples anteriors il·lustren com aquesta definició per textura és utilitzada per formular algorismes en l'enfocament estructural. Malauradament, s'associen un cert nombre de problemes en la formulació d'algorismes basats en aquesta definició, alguns d'aquests han estat enunciats resumidament per [HAW70] :

"Primer, la noció de que la textura depèn de la repetició de patrons locals en la imatge s'ha d'aplicar amb cautela. En la majoria dels casos pràctics, el patró que s'està repetint no és idèntic d'un lloc a un altre, com el paper de paret o el linoli. Millor dit, el patró local té propietats similars d'un lloc a un altre; usualment dos patrons dins la textura no són idèntics. El que es replica, per tant, és una classe de patró dins la textura en la qual tots els exemples de patró són contemplats com equivalents. A més, la periodicitat espacial del procés de replicació és rarament perfecta. Patrons locals, individualment, poden patir desplaçaments de fase respecte a tota la matriu. O la matriu que conté la periodicitat de la textura pot ser allargada o recargolada. O pot ser que no hi hagi tendència a la periodicitat a l'estar produïda per la interacció de forces naturals complexes, com ara les imatges aèries terrestres."

Per tots aquests problemes, semblaria que cal un algorisme de propòsit general per l'enfocament estructural consistent en els següents passos:

1. Extracció i identificació dels patrons unitaris.
2. Col·locació de cada patró unitari en la seva classe apropiada.
3. Determinació de les regles d'emplaçament per als patrons unitaris.

Cadascun d'aquests passos és un problema molt complex per ell mateix.



**Figura 2.8.** Ràfia i rèptil (de l'àlbum de Brodatz [BRO66]).

Per exemple, identificar i aïllar cada patró unitari en la figura 2.8, apareix com un problema molt complex que requereix un gran temps de càlcul. El problema de la localització de patrons pot ser, a més, complicat pel fet que aquests patrons unitaris, per ells mateixos, poden estar fet de elements estructurats o, més generalment, de subpatrons. I a la seva vegada, aquests subpatrons poden estar fets encara d'elements més petits, etc.

### **2.3. Síntesi de textures.**

La síntesi de textures consisteix en generar textures artificials que tinguin unes característiques predeterminades. La selecció de característiques apropiades és la part més crítica en qualsevol procés de classificació. En l'àrea de la classificació de textures, una manera de determinar si les característiques escollides són correctes és sintetitzar una imatge que tingui les característiques de textura donades. Si la imatge sintètica té una semblança molt propera a la imatge real, llavors poden assegurar amb bastant certesa que les característiques escollides són també apropiades.

L'anàlisi per síntesi és, de fet, una manera d'avaluar l'aptitud o capacitat absoluta d'un analitzador de textura.

Hi ha diversos mètodes de síntesi de textures, però més que donar una descripció detallada, es pretén fer una breu guia d'alguns dels treballs més significatius. [LOH94] descriu un algorisme (*Metropolis*) basat en matrius de co-ocurrència, anys abans [CRO83][GEM84] havien descrit una primera versió d'aquest algorisme. [TOU76a] utilitza un model de sèries de temps per a la síntesi de textures. A [YOK78], Yokohama i Haralick usen un model estructurat de creixement per a sintetitzar una imatge de textura altament complexa. En [PRA78a] i [PRA78b] es desenvolupa un conjunt de tècniques per a la generació de textures amb mitjanes, variàncies i funcions d'autocorrelació idèntiques però amb moments d'ordre superior diferents. En [GAG81] es dona una tècnica per a la generació de camps de textura binaris amb imposició d'estadístics de segon ordre. Chellappa descriu en [CHE85] una tècnica per a la generació d'imatges donat un camp aleatori de Gauss-Markov, així com [KHO87] i [KAS86a] basen la generació de la textura en models autoregressius simultanis (SAR). Amb mètodes estructurals, [TOM90] proposa un algorisme per a distribuir els diferents grups d'elements de textura existents segons les seves pròpies regles d'emplaçament.



Hi ha molts d'altres mètodes de síntesi de textures, però no menys interessants, com ara [SCH78], que proposa un model de mosaic aleatori per a generar i analitzar textures cel·lulars. En el procés de generar els mosaics hi intervenen diferents models com ara models d'ocupació, el model Johnson-Mehl, model de Poisson, model de tauler rotat, model de bombeig, ... [MAN82] proposa els fractals, que són un conjunt de funcions caracteritzades per la seva estructura infinitament recursiva, on es poden descriure patrons del món natural (terrenys, núvols, arbres, ...). [PEN84] aplica la teoria fractal per analitzar textures i demostra que hi ha una correlació entre la rugositat perceptible i la dimensió fractal de les textures.