



Universitat Autònoma de Barcelona

FACULTAT DE CIÈNCIES

DEPARTAMENT D'INFORMÀTICA

IDENTIFICACIÓ DE LES DIMENSIONS D'UN ESPAI
DE REPRESENTACIÓ DE TEXTURES BASAT EN UN
MODEL COMPUTACIONAL DE PERCEPCIÓ
PREATENTIVA

MEMÒRIA DE TESI PRESENTADA PER
MARIA VANRELL I MARTORELL A LA
UNIVERSITAT AUTÒNOMA DE BARCELONA
PER A OBTENIR EL TÍTOL DE
Doctora en Ciències (Informàtica)

OCTUBRE 1996

P: *Són iguals ?*

R: *No*

P: *Qué són?*

R: *Un gat i un gos*

A la meva família

RESUM

La textura és un estímul molt important dins les tasques visuals. La seva complexitat ha fet que no existeixi una representació general com la que existeix per altres estímuls com el color.

En aquest treball es proposa una representació general basada en un model computacional existent sobre la percepció preatentiva de textures. La representació que proposem ha estat estudiada des de diversos punts de vista. Primerament estudiem el seu cost computacional. En segon lloc establim la dimensionalitat de l'espai de representació que es defineix. Finalment donem una interpretació perceptual de les seves dimensions.

Per a tal de millorar el cost computacional del càlcul de la representació introduïm operacions no lineals de la morfologia matemàtica en els passos que representen les inhibicions no lineals del model. Per fer això proposem un mètode de descomposició d'elements estructurants isotròpics que permet reduir la complexitat de les operacions morfològiques.

Es demostra que la representació es pot reduir a un espai de quatre dimensions. Aquest fet estableix la seva viabilitat com a representació. L'exploració de la representació es fa a partir del mètode de l'escalament multidimensional i a partir de la definició d'un espai paràmetric de textures, que ens permet donar una interpretació dels eixos en termes dels paràmetres de contrast, escala i orientació dels elements texturals de la imatge.

Finalment es presenten els resultats d'aplicar aquesta representació a recerques basades en el contingut textural de la imatge en dues bases de dades conegudes.

Paraules clau: *visió per computador - percepció preatentiva de textures - descomposició d'elements estructurants - representació de textures de baixa dimensió - Accés a bases de dades pel contingut de la imatge*

ABSTRACT

Texture is an important cue in vision tasks. Its inherent complexity has derived in the absence of a general representation. As it exists for other important visual cues as color.

In this work we propose a general representation based on a computational model of preattentive texture perception. We have studied this representation from different points of view. Firstly, we study the computational cost. Secondly, we establish the underlying dimensionality of the representation space. Finally, we give a perceptual interpretation of each axis in this space.

In order to improve the cost in computing the representation, we introduce a non-linear operation from mathematical morphology to compute the steps of the model representing non-linear inhibitions. Here, we propose a method to decompose isotropic structuring elements, allowing to reduce complexity of morphological operations.

We demonstrate that this representation can be reduced to a four-dimension space. This fact implies its feasibility as a general texture representation. The exploration of this texture representation space is made using a multidimensional scaling approach on a parametric texture space that we have defined. It allows to give an interpretation of the space axis in terms of perceptual parameters, such as contrast, scale and orientation of the textural elements.

Finally, we present the results of applying this representation to searches by textural image content in large image databases.

keywords: *computer vision - preattentive texture perception - structuring elements decomposition - low-dimensional texture representation - query by image content*

AGRAÏMENTS

Primerament vull agrair les importants aportacions a tot aquest treball del meu director, Dr. Jordi Vitrià. Vull agrair-li molt sincerament el recolzament i coratge que m'ha donat en tot moment i sobretot la confiança que ha demostrat tenir sempre en mi, inclús en els moments en que no ho véiem tot de la mateixa manera.

També al Dr. Juan José Villanueva per haver-me donat l'oportunitat d'entrar en contacte en el món de la recerca i la docència universitària.

Vull fer un agraïment molt especial a en Xavier Roca, sense ell i la seva paciència aquesta tesi no hauria estat mai possible. El seu suport ha estat incondicional i ha abarcat tots els aspectes tant professional, afectiu com vital.

A en Xavier Binefa per les inacabables discussions en les que ho arreglàvem tot.

A en Xavier Sànchez per haver estat un gran company de despatx inclús en els moments en que li he envaït la seva taula.

A en Josep Lladós i en Felipe Lumbreras per ser com són, i perquè sempre han estat disposats a ajudar-me en tot allò que han pogut.

A en Ramon Baldrich, en Jordi Regincós i n'Enric Martí per haver-me creat l'entorn de treball més animat de tot el laboratori.

A en Joan Serrat vull donar-li les gràcies per haver-me introduït en el *món de les textures*, i haver-me ajudat molt en les meves primeres passes.

També a n'Antonio M. López, na Carme Molina i na Thais Jurado que han permés que la docència fos molt més fàcil per mi.

I a n'Agustín Collado i en Joan Masoliver per haver fet possible que el programari que he utilitzat estigués sempre a punt.

A tots els col·legues del grup vull agrair les sobretalles que han fet que tot fos molt agradable, concretament a n'Ernest Valveny, na Petia Radeva, en Xavier Varona, en J.R. Serra, na Gemma Sànchez, n'Esperança López, n'Aleix Martínez, n'Erundina Vázquez, en David Lloret, en Ricardo Toledo, en Craig Von Land, en Joan Carbó, na Mari Carmen Barbero, n'Anna Augé i n'Albert Pujol. També a n'F.X. Pérez, en J.A. Pérez, en Narcís Parés i en J.A. Rodríguez.

I a na Carme Ramirez, na Pilar Villa, na Carme Alfonso, na Cristina Minguet i en Toni Ventura vull agrair el seu suport pràctic.

També un agraïment especial al Dr. Jaime López-Krahe i a tot el grup del Laboratori d'imatges de Telecom Paris, per una inoblidable estada a Paris.

CONTINGUTS

1	El problema de la representació de textures	1
1.1	La textura	1
1.2	Necessitat d'una representació	4
1.3	Paradigmes sobre la representació en visió	6
1.4	Representació de similituds	8
1.5	Representacions basades en detectors de característiques	9
1.5.1	Eixos d'un espai de representació	10
1.5.2	Visió preatentiva	11
1.6	Objectius de la tesi	14
2	Representacions basades en la mesura de característiques	17
2.1	Mesura de característiques texturals	18
2.2	La Granulometria com a mesura de la granularitat	20
2.2.1	Morfologia matemàtica	21
2.2.2	Definició de granulometria	26
2.2.3	Granulometria sobre el volum	30
2.2.4	Obertures algebraiques	32
2.3	Aplicació a la classificació d'imatges radiogràfiques	40
2.3.1	La pneumoconiosi	40
2.3.2	Automatització de la classificació	42
2.3.3	Resultats	49
2.4	Valoració d'aquesta aproximació	67
3	Representació basada en detectors de característiques	69
3.1	Models multicanal de percepció de textures	69
3.2	El model computacional de J. Malik i P. Perona	71

3.2.1	Selecció dels filtres lineals	74
3.2.2	Introducció de mecanismes no lineals	78
3.2.3	Inhibició intracortical	83
3.2.4	Càlcul del gradient de textura	86
3.3	Definició d'una representació general de textures	87
3.3.1	Radi de la post-inhibició i densitat de textons	91
3.4	Problemes de la representació definida	92
4	Algorisme eficient pel càlcul de la representació	95
4.1	Descripció dels passos de l'algorisme	95
4.2	Introducció d'operacions morfològiques a l'algorisme	99
4.2.1	Millora de les operacions morfològiques basada en la descomposició d'elements estructurants	100
4.2.2	Descomposició general d'elements estructurants	101
4.2.3	Mètode proposat per a la descomposició de discs	106
4.2.4	Anàlisi dels resultats	116
4.2.5	Extensió de la descomposició a elements bàsics de majors dimensions	120
4.3	Algorisme proposat	121
5	Exploració de l'espai de representació	125
5.1	Escalament multidimensional	125
5.1.1	Escalament multidimensional clàssic	128
5.1.2	Escalament multidimensional no mètric	130
5.1.3	Dimensionalitat de les dades	132
5.2	Definició d'un espai paramètric de textures	135
5.2.1	Algorisme de segmentació basat en el model de barres i taques . . .	136
5.2.2	Parametrització de textures segons el model de taques i barres . . .	137
5.3	Interpretació de les dimensions de l'espai de representació	141
5.3.1	Variacions sobre un paràmetre	142
5.3.2	Interaccions entre paràmetres	144
5.3.3	Dimensionalitat de l'espai de representació	152
5.3.4	Confirmació de la interpretació amb imatges naturals	154
5.4	Conclusions sobre l'aproximació i els resultats	156

6	Identificació de les dimensions de l'espai de representació	161
6.1	Selecció de les imatges	162
6.2	Càlcul de la dimensió de l'espai de representació	162
6.3	Hipòtesi per a una interpretació basada en paràmetres perceptuals	165
6.3.1	Extracció de paràmetres a partir de la representació	165
6.3.2	Construcció de l'espai definit a la hipòtesi	168
6.4	Confirmació de la hipòtesi	170
6.4.1	Anàlisi de correlació	170
6.4.2	Anàlisi basat en una regressió lineal múltiple	173
6.4.3	Estudi d'altres variables	175
6.5	Conclusions	176
7	Aplicació de la representació per a l'accés a bases de dades	179
7.1	Accés a bases de dades pel contingut de la imatge	179
7.2	Resultats experimentals	180
7.2.1	Experiment 1	181
7.2.2	Experiment 2	181
7.2.3	Avaluació dels resultats	215
8	Conclusions i vies de continuació	219
A	Textures (I)	225
B	Textures (II)	227
	BIBLIOGRAFIA	228

CAPÍTOL 1

EL PROBLEMA DE LA REPRESENTACIÓ DE TEXTURES

En aquest capítol fem una introducció al concepte de textura com a estímul visual important dins les tasques de la visió. Plantegem la complexitat de la seva definició de la que se'n deriva una manca de representació acceptada. Per tant, la representació de textures es planteja com l'objectiu d'aquesta treball.

A partir d'aquest problema exposem els diferents paradigmes que han parlat sobre la representació en visió per computador i optem per seguir les idees de la representació de similituds. Amb aquest nou paradigma i els resultats d'alguns treballs sobre percepció preatentiva de textures plantegem les propietats que creiem que ha de complir una bona representació de textures. Aquestes propietats seran les que dirigiran tot el treball. Al final d'aquest capítol donem un resum de com s'ha organitzat tot el treball.

1.1 La textura

Textura és el terme usat normalment per a caracteritzar la superfície d'un objecte. Aquesta definició és perfecta si considerem que la textura és el resultat d'una sensació tàctil. Ara bé, una textura també és un estímul visual en el que hi intervenen més factors que els inherents a la pròpia superfície. Així doncs, la rugositat d'una paret pot ser percebuda visualment de manera molt diferent en funció de factors com la il.luminació, el punt de vista o la distància de l'observador; en canvi amb el tacte és sempre percebuda igual.

Amb aquesta comparació simplement es pretenia introduir el concepte de textura com una propietat lligada directament als sentits i remarcar la infinitat de textures que el sistema visual humà és capaç de percebre. Ens restringirem a les textures que hem anomenat visuals, que són el resultat d'una determinada projecció d'una superfície, segons certes condicions d'il.luminació, escala, etc.

La textura és un estímul important si tenim en compte que hi ha moltes tasques visuals en les que està involucrada. Les textures permeten segmentar imatges, classificar objectes o tipus de superfícies, i fins i tot la textura ens permet determinar l'orientació i la profunditat a la que es troben certes superfícies dins l'espai [122]. D'aquí ve el fet de que sigui justificat el seu estudi dins la visió per computador.

No existeix una definició ni un model general de textura que sigui acceptat per a tothom. Podem veure un conjunt de conceptes o propietats que apareixen lligats a les textures, alguns d'aquests han estat recollits per R.M. Haralick i L.G. Shapiro a [41]:

- L'existència d'una primitiva o patró que es repeteix.
- L'organització espacial de les primitives que pot ser aleatòria, o bé basada en una dependència entre elles.
- Les relacions entre les variacions dels nivells de grisos de tota la textura.
- La invariància per translació d'una textura.
- La resolució de la textura definida com la mida mínima de la finestra d'observació que deixa les mesures texturals invariants [30].
- L'escala de la textura.
- L'orientació que presenten certs tipus de textures.
- La inexistència d'un vocabulari que descriu les textures [29], que fa que ens haguem de basar en avaluacions qualitatives de certes propietats: gruix, granularitat, aleatorietat, suavitat, finor, irregularitat.

Tots aquests són conceptes o propietats de les textures que s'han de tenir en compte a l'hora de donar una definició. La majoria de treballs s'han basat només en alguna d'aquestes propietats o en certs tipus de textures, fet que ha suposat l'aparició de moltes definicions diferents en diversos treballs. A la taula 1.1 n'hem fet una petita recopilació.

Com a resum podem dir que la complexitat d'aquest estímul fa que no existeixi un vocabulari que el permeti definir de manera precisa.

La manca de definició ha fet que el problema de les textures hagi estat tractat de maneres molt diferents dins la visió computacional. Això ha provocat que la majoria d'aproximacions

(...) considerem una textura com un camp de la imatge 2-dimensional estocàstic i possiblement periòdic. [21]

(...) la noció de textura apareix com depenent de tres ingredients: (1) Un ordre local repetit sobre una regió, que és gran en comparació a la mida de l'ordre, (2) l'ordre consisteix en una col·locació no aleatòria de les parts elementals, i (3) les parts són entitats més o menys uniformes que tenen aproximadament les mateixes dimensions dins tota la regió textural. [42]

(...) una textura és una regió d'imatge en la que existeix una finestra de dimensió reduïda, de manera que una observació d'aquesta és traduïda per una impressió (o percepció) visual idèntica per totes les posicions en que es pot traslladar la finestra dins la regió considerada. [109]

(...) una textura en una imatge és descrita pel nombre i el tipus de primitives i per la seva organització espacial o disposició. [39]

(...) textures visuals són definides com agregats de molts elements petits. Els elements poden ser punts de cert colors o simples patrons. [52]

(...) una textura és una funció aleatòria f , construïda a partir d'un procés puntual sobre $E = R^n$, i d'una funció aleatòria f' sobre E , anomenada funció primària. [76]

TAULA 1.1: Algunes definicions de textura

al problema tinguin un caràcter molt parcial. G.L. Gimel'farb i A.K. Jain a [32] resumeixen aquest fet classificant les aproximacions fetes en els tres següents grups:

Mètodes estructurals basats en elements primitius i en les regles espacials per a la seva localització.

Mètodes probabilístics que assumeixen que la imatge de la textura és el resultat d'un procés aleatori.

Mètodes espectrals que consideren la textura com un senyal del que extreuen el contingut de les freqüències espacials.

Des d'un punt de vista metodològic la manera més elegant de tractar el problema ha estat a partir de la definició de models matemàtics en els que les textures venen donades per una determinada expressió matemàtica del model ([21, 46, 77, 1]).

Trobem resums molt interessants sobre tots els treballs fets en textures i els diversos enfocaments a [33, 39, 41, 121, 108].

1.2 Necessitat d'una representació

Els problemes que ens hem trobat a l'hora de definir el concepte de textura ens pot fer entendre el fet de que no existeixi una representació computacional de les textures acceptada per tothom. Representació que sí existeix per altres estímuls visuals com el color.

El color és un estímulo visual que ha estat caracteritzat amb diverses representacions en espais de tres dimensions. Així doncs, tenim la representació HLS on cada eix s'identifica amb una característica perceptual com el to, la il.luminació i la saturació. O la representació RGB que identifica el color com la composició de canals de vermell, verd i blau.

Aquestes representacions del color, en les que un determinat color s'identifica amb un punt de l'espai, han fet possible una estandardització d'aquest estímulo que ha permès un fàcil ús d'aquest en sistemes de visualització, en l'establiment de llenguatges de descripció i ha permès entendre millor aquest estímulo. Des d'un punt de vista metodològic aquestes representacions permeten un tractament més clar del color en els sistemes que tenen com a objectiu una determinada tasca visual.

En textures s'està encara lluny de definir una representació semblant a les que s'han definit pel color. Una manera d'arribar a aquesta representació podria venir donada a

partir de la construcció d'una taxonomia de textures que fos general. Vegem-ne alguns intents:

- Les textures poden ser de dos tipus [41]: *microtextures* formades per regions petites amb el mateix nivell de gris, i les relacions espacials entre aquestes regions es restringeix a zones molt petites, i *macrotextures* formades per regions amb una forma específica i una organització regular. Aquests dos tipus de textures també s'han anomenat estadístiques i estructurades [39].
- Les textures es poden classificar en tres tipus [79]: *desordenades* són les que no presenten ni repetició d'un element ni orientació; *ordenades dèbilment* són les que presenten un determinat grau d'orientació a cada punt de la textura; i *fortament ordenades* són textures que presenten una localització específica de certes primitives, o bé tenen una determinada distribució d'un classe d'elements.

Intentar definir un espai de representació a partir d'aquestes taxonomies resulta difícil tant per la parcialitat de la primera, com per l'ambigüitat de la segona.

Un dels resultats més clars d'anar cap a un espai de representació de textures ho trobem en els treballs de A.R. Rao i G.L. Lohse [81, 80]. Aquests treballs acaben amb una proposta d'espai de representació de textures de tres dimensions on els eixos venen donats per propietats perceptuals de les textures. A la figura 1.1 podem veure l'espai proposat per aquests autors. És el resultat d'un conjunt d'experiments a partir d'una sèrie de textures i amb un conjunt de persones que han establert les propietats de cada textura.

La importància d'aquest espai de representació ve donada per tres fets, primer de tot perquè s'ha reduït el nombre de dimensions de l'espai a un número petit, sense pèrdua de generalitat en el tipus de textures. En segon lloc perquè cada eix correspon a una característica perceptual que pot ser interpretada. I el tercer fet important és que representa directament un espai perceptual en el sentit de R.N. Shepard [98]. És a dir, les similituds entre les textures definides per aquest espai són les mateixes que les que defineix el sistema visual humà (SVH). Tornarem a tractar aquesta propietat més endavant.

Cal tenir en compte que la representació proposada no és una representació computacional sinó una representació que permet descriure textures des del punt de vista perceptual. Queda per definir el pas d'aquesta representació a una representació computacional que pugui ser calculada automàticament, aquest pas es comentarà al capítol següent.

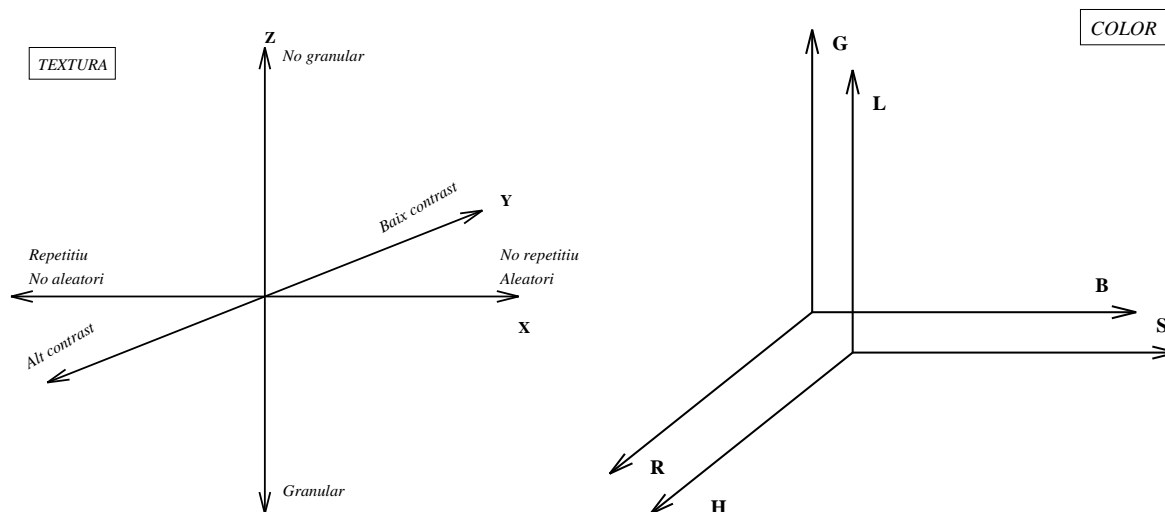


FIGURA 1.1: (a) Espai de representació de textures proposat a [81]. (b) Espais de representació de color.

Abans d'entrar a definir com pot ser una representació general de textures anem a veure els diferents paradigmes existents en visió per computador i el tractament que fan de la representació.

1.3 Paradigmes sobre la representació en visió

Abans de centrar-nos en el problema concret de com construir una representació de textures, anem a veure com han plantejat la representació els diferents paradigmes que s'han succeït en visió per computador.

El paradigma de Marr [68, 67] ha estat un dels més importants dins de la visió per computador, principalment pel fet de plantejar un sistema de visió com una conjunt de mòduls independents, cada un encarregat d'una determinada tasca visual. Aquest sistema visual des del punt de vista computacional ha d'estar dividit en tres nivells:

Nivell baix: els mòduls d'aquest nivell són els que operen directament sobre la imatge, amb l'objectiu de construir una representació que contingui tota la informació 2-dimensional de l'escena.

Nivell mig: en aquest nivell els mòduls fan una reconstrucció 3-dimensional de l'escena a

partir de la representació del nivell baix.

Nivell alt: aquí els mòduls utilitzen els resultats de la reconstrucció del nivell mig per a raonar sobre l'escena i actuar en funció d'aquesta.

Com podem veure en aquest paradigma la representació juga un paper molt important. Un sistema visual ha de primer reconstruir el món internament per aleshores poder extreure d'ell tota la informació necessària. Així doncs, aquest paradigma participa de la idea que la millor representació del món és aquella que és igual que el món i que per tant el reconstrueix totalment.

Aquest paradigma ha estat molt important dins la visió per computador degut a que va sentar les bases sobre quins eren els problemes de la visió i com es podien tractar de forma general des d'un punt de vista computacional. Ara bé, aquest model no ha donat bons resultats a la pràctica, això és degut principalment a la manera d'entendre la representació. La majoria dels problemes que intenten solucionar els mòduls del nivell baix i mig són problemes mal posats, que no estan completament condicionats i no tenen solució única. T. Poggio et al. [75] plantegen la teoria de la regularització per a solucionar aquest problema. De totes maneres el problema continua sent computacionalment intractable.

Com a resposta als problemes d'aquest paradigma en surten d'altres com el de la visió amb propòsit [2, 4], que es basa en la idea de que molts dels problemes de la visió es poden solucionar més fàcilment si l'observador és actiu i és dirigit per objectius concrets. Aquesta visió es contraposa a l'objectiu general del paradigma reconstruccionista en el que la representació 3-dimensional ens permetria solucionar-ho tot.

En aquest paradigma no es necessita una representació elaborada de l'escena, sinó que n'hi ha prou amb una representació que sigui capaç de contestar preguntes concretes sobre diferències entre objectes, estimacions de trajectòries, etc.

L'absència de resultats pràctics que s'ha derivat del paradigma de la reconstrucció ha portat també a paradigmes basats en la percepció directa, en els que es planteja que no hi ha cap necessitat d'una representació ja que la major part de tasques visuals es poden dur a terme directament des de la imatge [31, 78]. Aquest és un paradigma paral·lel al de sistemes basats en el comportament definits en el camp de la intel·ligència artificial [11].

Com a compromís entre el paradigma reconstruccionista i el paradigma de la visió sense representació, S. Edelman proposa el paradigma de la representació de similituds, que resumim en els següent apartat.

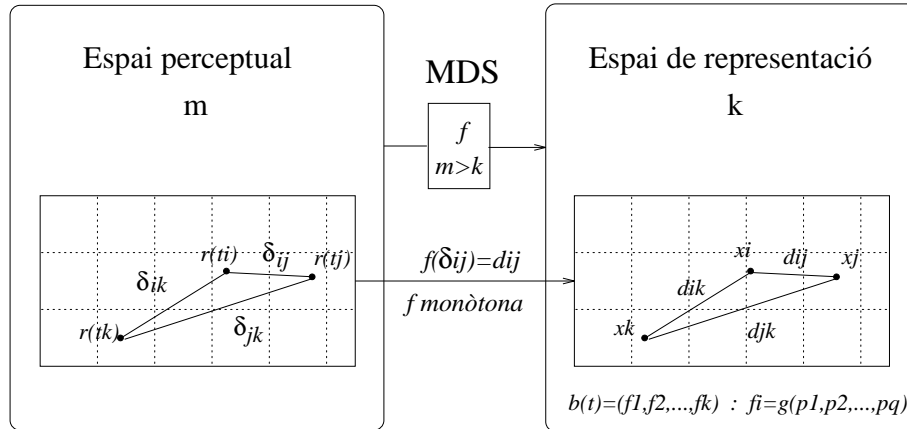


FIGURA 1.2: Esquema del mètode desenvolupat per R.N. Shepard que permet construir un espai que conservi la mètrica de les similituds respecte d'un altre.

1.4 Representació de similituds

La teoria de la representació que proposa S. Edelman [24] es basa en l'establiment d'isomorfismes de segon ordre entre la representació i allò que es representa. Aquesta noció d'isomorfisme de segon ordre com a principi de la representació va ser proposat per R.N. Shepard [96], i es pot resumir en els següents dos punts:

- (i) No volem que allò que es representa sigui semblant a la seva representació (Isomorfisme de 1r ordre).
- (ii) Volem que les similituds entre el que es representa es corresponguin amb les similituds entre les seves representacions (Isomorfisme de 2n ordre).

Amb aquesta idea R.N. Shepard i els seus col·legues [95, 59, 58] desenvoluparen un mètode, l'escalament multidimensional no mètric (MDS), que permet construir espais de representació que mantinguin la mètrica de les similituds en un sentit monòton. D'aquesta manera tenim una eina per a descobrir com s'organitzen les similituds en els espais perceptuals definits pel SVH. Un esquema d'aquest mètode es pot veure a la figura 1.2

Seguint aquestes idees com a principi, S. Edelman et al [22, 23] es plantegen la definició d'un espai de representació de la forma dels objectes. De manera que l'espai de representació de les formes es pugui codificar com un espai definit per característiques i de baixa dimensió, semblant a l'espai de representació del color.

Amb aquest objectiu proposa una representació que parteixi dels resultats d'aplicar un conjunt de detectors de característiques, i que els resultats de la sintonització d'aquests amb la imatge siguin la base de la representació.

Aleshores la representació seleccionada haurà de complir un isomorfisme de $2n$ ordre amb l'espai perceptual del SVH.

Partint d'aquesta breu introducció d'una complexa teoria com és la representació de similituds, ara ja podem definir quines són les propietats interessants que hauria de tenir una representació computacional de textures:

Perceptualment completa El seu comportament ha d'estar correlacionat amb les interpretacions del SVH. Per a que això sigui possible ha d'existir una mètrica de la similitud dins l'espai de representació, que es comporti igual, en el sentit de Shepard, que la mètrica de similituds de l'espai perceptual. És a dir, allò que sigui semblant dins l'espai perceptual ha de ser proper dins l'espai de representació.

Amb significat associat La representació ha de permetre interpretar la posició de la imatge dins l'espai, de manera que la posició dins l'espai ens permet deduir propietats perceptuals de la textura. Aquesta propietat permetria la utilització d'aquesta representació en molts sistemes artificials de visió en els que sigui necessari el tractament de textures.

Amb dimensionalitat baixa Aquesta propietat està molt lligada a les dues anteriors. L'estudi sobre la mètrica de l'espai de representació només serà possible si es pot veure en espais de poques dimensions, i per altra banda l'assignació de significat als eixos només és interessant si el número d'eixos és petit.

Aquestes propietats de la representació coincideixen en part amb els criteris definits per A. Pentland i els seus col·legues [74], sobre com ha de ser una representació d'imatges que permeti fer una bona recerca en una base de dades d'imatges.

1.5 Representacions basades en detectors de característiques

Abans de començar a definir una representació de textures anem a veure quins són els atributs de la textura que podrien representar els eixos de l'espai d'aquesta representació.

1.5.1 Eixos d'un espai de representació

A l'hora de definir els eixos de l'espai podem seguir diverses alternatives, sobretot si tenim en compte la diversitat d'aproximacions que s'han donat a aquest problema. Seguidament analitzarem tres possibilitats:

La mesura de característiques d'ordre superior es basa en mesures calculades directament sobre la imatge i que permeten representar quantitativament característiques amb una semàntica perceptual com la granularitat, la repetitivitat, la suavitat, el contrast, etc.

Extracció de formes i la seva organització espacial consisteix en l'aplicació d'algorismes que permetin extreure i reconèixer les formes que defineix la textura, seguit de l'estudi de la seva organització espacial. Aleshores la representació vindrà definida per la forma de l'objecte i les regles de localització.

Les respostes de detectors de característiques es basa en simular els camps receptius del SVH. Els diferents tipus de detectors es poden sintonitzar a determinades característiques de la imatge, obtenint un conjunt d'imatges resposta que permeten representar el tipus de característiques que s'han trobat a la imatge.

Per analitzar la viabilitat d'aquestes tres possibilitats a la taula 1.2 fem una comparació exposant les avantatges i els inconvenients que plantegen cada una d'elles.

Les representacions basades en la mesura de característiques presenten el problema de l'ambigüïtat en la mateixa definició sobre el que s'està mesurant. En aquesta línia estaria el treball de R. Rao et al [81] que ja hem comentat. Els problemes que es plantegen estan principalment en el fet de definir quines són les característiques que es mesuren i com es mesuren. És un via que planteja problemes importants però que és en la que més s'ha treballat i que per tant ja hi ha resultats molt interessants. En el capítol següent plantegem la resolució d'un problema fent servir aquesta aproximació. A la vista dels resultats comentarem més profundament la viabilitat d'una representació basada en aquestes mesures.

La segona aproximació està molt ben definida, consisteix en extreure un conjunt de formes de la imatge i examinar la seva organització [87, 128]. Aleshores es planteja un problema a nivell algorísmic sobre com extreure tota aquesta informació de la imatge, tenint en compte que es treballa normalment en un nivell simbòlic. El problema més important

Representació	Avantatges	Inconvenients
Mesura de característiques	Semàntica perceptual	Quines característiques? Com es mesuren? Ambigüitat
Extracció de formes i organització	Fàcil interpretació	Algorismes complexos No és aplicable a totes les textures
Respostes a detectors de característiques	Mecanisme general Validesa biològica No Ambigua	Dimensionalitat alta \Rightarrow Dificil interpretació Basada en models preatentius

TAULA 1.2: Comparació de les diferents aproximacions que es poden seguir en la construcció de la representació.

d'aquesta aproximació és que difícilment es aplicable a qualsevol textura, com seria el cas de textures aleatòries.

La darrera aproximació que és la que seguirem en aquest treball presenta un conjunt de propietats interessants, com és el fet de que computacionalment presenta un mecanisme general ben definit i es basa únicament amb característiques fotomètriques i geomètriques de determinades regions de la imatge, aquest fet fa que la seva interpretació no sigui ambigua.

Per altra banda presenta inconvenients importants com l'alta dimensió del resultat, ja que normalment aquest tipus d'aproximacions han d'aplicar conjunts molt grans de filtres detectors per poder mantenir la generalitat i per a que puguin ser aplicats a qualsevol tipus d'imatge. Aquesta alta dimensionalitat fa que aquest tipus d'aproximacions siguin difícils d'interpretar.

Com ja hem dit aquesta aproximació es basa en característiques fotomètriques i espacials, ja que és el resultat d'aplicar processos de percepció preatentiva. Per entendre aquest concepte fem una petita introducció al la teoria del textó de B. Julesz [49, 52, 48, 53] que és qui ha definit el concepte de procés preatentiu en la discriminació de textures.

1.5.2 Visió preatentiva

B. Julesz ha estat un dels autors més referenciats en el camp de la discriminació de textures. A partir dels seus treballs basats en experiments psicofísics ha desenvolupat la teoria del textó. Aquests experiments s'han fet amb textures sintètiques formades normalment per la

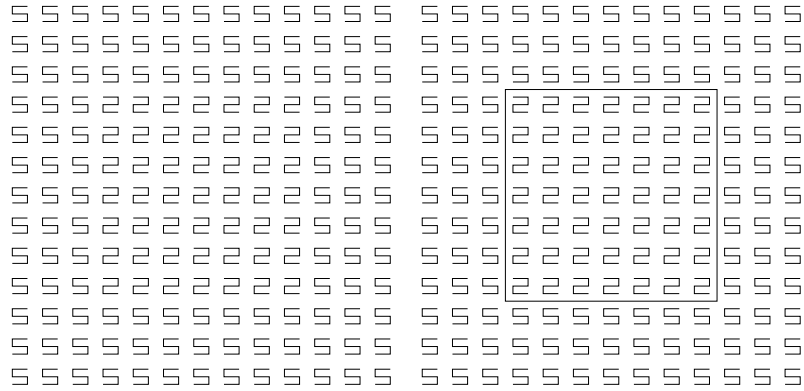


FIGURA 1.3: Exemple d'una imatge amb dues textures no discriminables preatentivament, però que si ens fixem atentivament en cada un dels elements que la formen podem veure dues regions diferents.

repetició d'un patró.

Els seus primers experiments es basaven principalment en la búsqueda de característiques permetien que la discriminació entre dues regions amb textures diferents. En un principi Julesz seguia la corrent *Gestaltista* de la psicologia que sostenia que la percepció visual és el resultat de càlculs estadístics globals de la imatge. Un estadístic de primer ordre representa la probabilitat de que punts llençats aleatòriament sobre la imatge caiguin sobre un determinat nivell de gris. De la mateixa manera un estadístic de segon ordre és la probabilitat de que parelles de punts separades una distància l i amb una orientació θ caiguin sobre una determinada parella de nivells de gris de la imatge. Denotarem l'estadístic global d'ordre n de la textura X com $E_n(X)$.

Anem a veure les principals hipòtesis que plantegen els seus resultats. L'existència de textures com les de la figura 1.3 fa que es pugui donar aquesta primera hipòtesi.

Hipòtesi 1.1 *Existeixen textures X i Y amb $E_2(X) = E_2(Y)$, tals que no són discriminables ràpidament però sí que ho són si ens fixem en cada un dels elements que la formen. Llavors es pot dir que en el sistema visual humà es donen dos modes d'operació:*

1. *Visió preatentiva. Es realitza de forma instantània, en paral·lel, i cubrint un ampli camp visual.*
2. *Visió atenta. Es produeix per una búsqueda en sèrie i gràcies a la focalització local.*

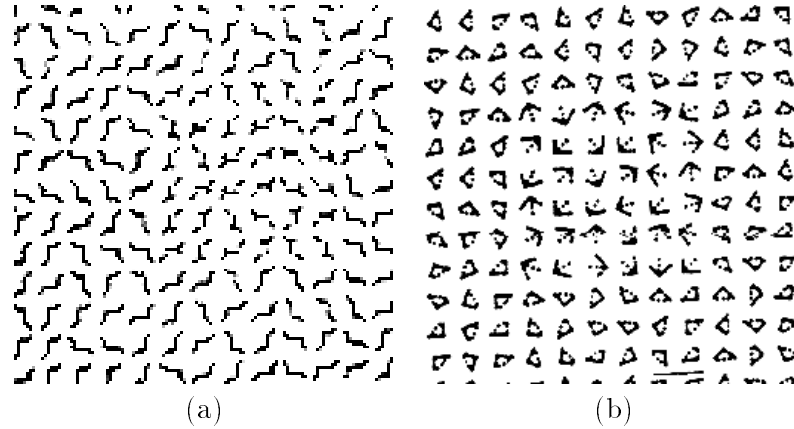


FIGURA 1.4: Parelles de textures amb estadístics de $2n$. ordre iguals, i que són fàcilment discriminables preatentivament (Aquestes imatges han estat preses dels treball de B. Julesz).

Per altra banda l'existència d'aquestes textures li permet establir una nova hipòtesi.

Hipòtesi 1.2 *Existeixen textures X i Y amb $E_2(X) = E_2(Y)$ i $E_3(X) \neq E_3(Y)$, i que no són discriminables preatentivament, aleshores podem concloure que el sistema preatentiu humà no pot calcular el estadístics d'ordre 3 o més grans.*

Després d'aquestes primeres hipòtesis s'arriba a la construcció de textures com les de la figura 1.4, a partir de les que es dedueix l'existència dels textons i es defineix una nova hipòtesi.

Hipòtesi 1.3 *Existeixen textures X i Y amb $E_2(X) = E_2(Y)$ i $E_3(X) \neq E_3(Y)$, i que són discriminables preatentivament, aleshores si tenim en compte la hipòtesi anterior podem dir que la discriminació és deguda al càlcul d'altres característiques locals, que anomenarem textons, i no a les diferències d'estadístics globals.*

Els estudis que segueixen es dirigeixen cap a definir quines característiques locals defineixen els textons. Experiments que el portaran a una darrera hipòtesi

Hipòtesi 1.4 *En percepció de textures el sistema visual preatentiu utilitza només característiques locals visibles, textons, i els textons no són aparellats entre ells. El sistema preatentiu utilitza els textons només globalment de la manera més simple possible, això és comptant el seu número (densitat).*

Els textons han estat definits per imatges en blanc i negre com:

- (i) Segments de línia definits per una longitud, una orientació, una amplada, etc.
- (ii) Acabaments o finals de línia.
- (iii) Creuements de segments de línia.
- (iv) Taques definides per segments de línia tancats, la seva forma no és important només els seus promitjos sobre la longitud, l'amplada i l'orientació.

La teoria del textó ha estat revisada a [49]. En aquest article es mantenen les idees que acabem de presentar però s'extenen amb resultats sobre com han de ser els passos de combinació global pel càlcul dels promitjos dels textons. Aquests resultats es comentaran més endavant quan es defineixi la representació.

Altres treballs posteriors han mostrat resultats que no són del tot consistents amb aquesta teoria del textó, però tot i així, la teoria del textó ha sentat les bases de la percepció preatentiva de textures.

1.6 Objectius de la tesi

Ara ja podem definir quins són els objectius concrets d'aquest treball i com es duran a terme. L'objectiu final és la construcció d'una representació de textures amb les següents propietats:

Generalitat com objectiu significa que volem una representació que se pugui calcular sobre qualsevol tipus d'imatge, sense que suposi cap aproximació específica.

Preatentiva Que no es basi en conceptes d'alt nivell per evitar l'ambigüitat i la falta de definició. Preferim una representació que es basi només en característiques fotomètriques i geomètriques locals, en el mateix sentit que els textons de la teoria de Julesz.

Amb una mètrica de la similitud com la del SVH volem que la representació defineixi un espai en el que la mètrica de les similituds entre les textures estigui d'acord en l'espai perceptual definit pel SVH. Per aquest motiu construirem una representació que es basi en un model computacional de la percepció humana. Tot i així, la construcció

de l'espai de representació haurà d'anar seguida d'un conjunt d'experiments psicofísics que confirmaran un comportament perceptual coherent.

De baixa dimensió per a que permeti una interpretació perceptual sobre el significat d'un punt a l'espai de representació, i per a que sigui possible analitzar la seva mètrica de similitud.

Amb una interpretació perceptual això significa que ens haurem de basar en paràmetres perceptuals que intervenen en la percepció preatentiva com l'escala, l'orientació, el contrast, etc.

Amb un cost computacional baix que permeti ser usada en aplicacions pràctiques.

Aquests objectius s'han dut a terme en els següents capítols seguint la següent organització. Al capítol 2 presentem els resultats de solucionar una aplicació concreta a partir d'una aproximació basada en la mesura d'una característica perceptual com la granularitat, a partir d'aquesta aplicació es treuen algunes conclusions sobre les limitacions d'aquesta aproximació per a una representació general de textures.

En el capítol 3 es fa una petita introducció als mètodes multicanal en els que es basarà la representació que es pretén definir. Aleshores s'introduirà el model computacional definit per J. Malik i P. Perona i a partir d'aquest es definirà una representació que serà la base d'aquest estudi.

Abans de passar a explorar la representació, en el capítol 4 es presenta un algorisme eficient per a calcular la representació. L'algorisme es basa en la introducció d'operacions morfològiques al model. Per a optimitzar aquestes operacions es proposa un mètode per a la descomposició d'elements estructurants rodons que representaran els veinatges del model.

En el capítol 5 es planteja com a objectiu trobar una interpretació perceptual de la representació aquesta interpretació es basarà en un espai de baixa dimensió que conservi la mètrica de similitud respecte de l'espai de representació. Amb aquest objectiu es defineix un espai paramètric de textures que permetrà la interpretació perceptual. El problema de l'ús d'aquest espai està en que els resultats només es poden donar per textures simples d'aquest espai.

En el capítol 6 s'extenen els resultats del capítol anterior al cas general, a partir de la definició d'un conjunt de textures naturals que assumim com a mostra general del conjunt de totes les textures. Primer de tot es confirma la baixa dimensionalitat de l'espai de

representació i aleshores es passa a demostrar una hipòtesi de combinació dels paràmetres perceptuals amb els que es vol interpretar la representació.

Finalment, en el capítol 7, donem els resultats d'aplicar la representació definida en una aplicació d'accés a bases de dades d'imatges. Es realitzen experiments sobre bases de dades conegudes i es dona una avaluació dels resultats.

CAPÍTOL 2

REPRESENTACIONS BASADES EN LA MESURA DE CARACTERÍSTIQUES

En aquest capítol presentem una de les aproximacions més usades en problemes de descripció de textures, que és la que es basa en la mesura de certes característiques texturals sobre la imatge. Aquestes característiques venen definides normalment per una propietat perceptual. En aquest capítol veurem els resultats de mesurar la propietat de la granularitat sobre les imatges d'un problema concret que es pretén automatitzar.

El càlcul de la granularitat com a característica textural es pot veure com una conseqüència de la definició del concepte de granulometria. La granulometria és un procés que ve de la ciència dels materials i que consisteix en la separació d'objectes a partir de tamisos de diferents mides. La morfologia matemàtica ha donat una definició matemàtica d'aquest concepte que permet ser reproduït per a l'anàlisi d'imatges. Abans de definir el concepte de granulometria fem una breu introducció a la morfologia matemàtica. El concepte d'obertura algebraica serà el que ens permetrà definir la granulometria sobre les imatges.

El problema sobre el que s'han aplicat aquestes mesures texturals és el de la classificació automàtica de radiografies de malalts de pneumoconiosi. El diagnòstic sobre l'estat de la malaltia es fa a partir de les plaques radiogràfiques i tenint en compte uns models estàndards. Els criteris pel diagnòstic són de mida i densitat de partícules a la imatge. Aquest procés es dut a terme pel radiòleg sense fer una mesura exhaustiva de les partícules, sinó que es basa normalment amb mesures aproximades sobre els patrons que formen les partícules. Aquest fet és el que ens ha portat a tractar el problema com un problema de classificació de textures.

Presentem els resultats d'aplicar el concepte de granulometria per a classificar aquestes imatges. Hem seguit diferents algorismes que anirem comparant i analitzant el seu rendiment.

Finalment i a la vista d'algunes limitacions que planteja el càlcul de mesures sobre la imatge, comentem alguns dels problemes que planteja la representació general de textures a partir d'aquesta aproximació.

2.1 Mesura de característiques texturals

Des del primer moment en que les imatges poden ser tractades com a vectors de dues dimensions amb un nivell de gris per a cada punt, es planteja el problema de la classificació o descripció de textures. Els primers treballs es plantegen el problema de definir quines són les característiques de la textura que la permeten descriure. Al mateix temps es comencen a plantejar les primeres mesures sobre les imatges que permetin representar quantitativament aquestes característiques.

En un d'aquests primers treballs R. M. Haralick et al. [40] presenten les matrius de co-ocurrència o matrius de dependència espacial dels nivells de gris. La definició d'aquestes matrius ve inspirada pel fet de que es pensava que la discriminació de textures era un procés visual global sobre la imatge basat en les diferències entre certs estadístics de segon ordre (distribucions de probabilitat conjunta) que havien estat definits per B. Julesz [47].

Aquestes matrius es defineixen per una distància d i un angle θ , de manera que el valor de la matriu p_{ij} representa la probabilitat de que es trobi un punt amb nivell de gris i a una distància d i amb una direcció definida per l'angle θ d'un punt amb nivell de gris j . A partir de les matrius de co-ocurrència es deriven un conjunt de mesures com el contrast, la correlació, la variància, etc.

El problema que es plantegen els autors, després de la definició d'aquestes mesures, és el següent:

Quin és el significat perceptual d'aquestes mesures definides?

Més endavant, H. Tamura i els seus col·legues [105] es plantegen el problema a l'inrevés, això és

Quines són les característiques texturals que són percebudes pel sistema visual humà?

Quines són les mesures computacionals que les permeten calcular?

A partir d'un conjunt de 16 textures de l'album de Brodatz [10], una sèrie d'experiments psicofísics amb persones i el següent conjunt de característiques perceptuals:

Gruix és la característica més fonamental i està relacionada amb la mida de les regions en que està dividida la textura.

Contrast és la característica que es basa en el rang de nivells de grisos de la textura.

Direccionalitat és una mesura global sobre la forma i la col·locació de les regions de la imatge. Dues regions que només es distingeixen per la seva orientació tenen el mateix grau de direccionalitat.

Linealitat és una característica relativa a la forma dels elements texturals i que fa referència a la isotropia dels elements.

Regularitat és una mesura sobre les variacions de les regles de col·locació dels elements dins de la textura.

Rugositat és una característica més pròpiament tàctil que visual, i que fa referència a la quantitat de canvis que es donen a la textura.

s'estableix la importància perceptual de cada una.

Primer de tot, cal remarcar que totes aquestes característiques no tenen un únic significat associat, i poden resultar ambigües, ja que persones diferents les tracten de manera diferent. A més a més moltes es poden correlacionar entre elles. Així doncs, el gruix està molt relacionat amb la rugositat. I la direccionalitat amb la linealitat.

Una de les conclusions a les que arriben amb l'experiment és que les característiques de gruix, contrast i direccionalitat són perceptualment molt importants. Un cop es definida la importància perceptual de cada una d'elles, es proposen les millors mesures de cada una d'aquestes característiques, en el sentit de que els resultats es correlacionin amb el comportament perceptual del sistema visual humà.

Després d'aquests dos treballs s'han definit moltes mesures diferents per mesurar aquestes mateixes característiques texturals. Al mateix temps que se n'han definit d'altres relacionades amb aquestes, com la granularitat o la suavitat, vegem-ne alguns exemples:

- El gruix de la textura s'ha mesurat amb la dimensió fractal de la superfície de la imatge [88].
- La rugositat a partir del promig de la curvatura de la superfície de la imatge [92].

- La direccionalitat a partir de l'estimació de l'orientació local d'una textura a partir del gradient orientat de la gaussiana [82].
- La borrositat, la granularitat, la discontinuitat, la linealitat i la curvatura són mesurades, a [14] a partir de la presència de canvis de nivell de gris suaus, presència de densitats altes de gradients no concatenats, nombre de línies discontinues, canvis de direcció sobtats a les línies d'una regió, densitat de línies rectes, densitat de línies no rectes, respectivament.
- La granularitat a partir d'implementar un procés granulomètric sobre la imatge. Un estudi sobre aquesta aproximació es veurà en els següents apartats ([110, 113, 115]).

Per tant, acabem d'introduir una aproximació habitual al problema de les textures que es basa en el càlcul de certes mesures sobre la imatge que es poden interpretar amb un significat perceptual. En els següents apartats veurem una mesura concreta, la corba granulomètrica, aplicada a un problema de classificació d'imatges mèdiques com és la pneumoconiosi en radiografies de tòrax i els resultats obtinguts. Seguidament donem una valoració sobre aquests tipus d'aproximacions i la seva validesa per a la definició de representacions de textures.

2.2 La Granulometria com a mesura de la granularitat

La granulometria és el concepte teòric per definir el mètode que prové de la Ciència dels Materials per a representar la distribució de mides de grans.

Si es pren com exemple l'anàlisi de materials en pols, un dels mètodes més usats per a mesurar les mides de les partícules és el de passar la pols per una sèrie de tamissos amb forats de mida cada vegada més petits (veure figura 2.1). Després de realitzar l'operació amb cada tamís quedaran les partícules de talla inferior a les que havien quedat en el tamís precedent, és a dir, mitjançant aquest procés obtenim una separació dels objectes en funció de la mida. Les mesures que realitzem sobre els objectes que resten a cada tamís ens donen una funció de distribució de certa mesura en relació a la talla de l'objecte. Aquesta funció pot ser un bon indicatiu de la talla mitja dels objectes de la imatge.

Fins ara hem vist la idea intuitiva del concepte de granulometria, ara la qüestió és: com podem utilitzar aquest concepte per determinar una mesura de la granularitat d'una textura?. Per a respondre aquesta pregunta cal solucionar els següents problemes:

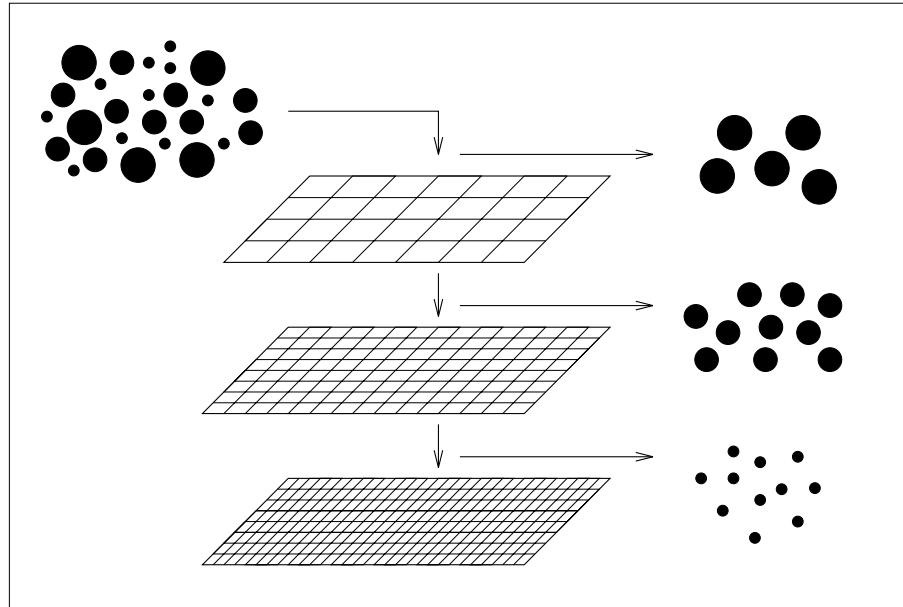


FIGURA 2.1: Esquema del procés de separació d'objectes de diferent talla, en funció de la mida del forat del tamís.

- Com podem implementar el concepte de tamís sobre les imatges?
- Quines mesures s'han de realitzar sobre els objectes que queden en el tamís?
- Com relacionar les mesures amb la talla de la transformació tamís?

Per a poder aplicar el concepte de granulometria a l'anàlisi d'imatges, anem ara a contestar cada una de les qüestions que ens hem plantejat. En els apartats següents veurem la formalització del concepte de granulometria per anàlisi d'imatges que s'ha fet des de la morfologia matemàtica, veurem el volum com una mesura interessant del que queda en el tamís, i finalment definirem la funció granulomètrica que ens permetrà mesurar la granularitat. Abans de tot, donem una petita introducció a la morfologia matemàtica.

2.2.1 Morfologia matemàtica

La morfologia matemàtica és un model matemàtic que ofereix un conjunt de mètodes aplicables a l'anàlisi d'imatges [91, 93]. Per a imatges amb nivells de gris la morfologia matemàtica es pot veure com un tipus de processament geomètric ja que considera la imatge com una superfície topogràfica sobre la que té definits una sèrie d'operadors.

Donarem les definicions bàsiques de la morfologia a l'espai dels conjunts, i mitjançant l'estructura de reticle complet ho extendrem a l'espai de funcions.

Tots els operadors morfològics es basen en la definició d'un element estructurant com element definidor de la connectivitat entre els punts de la imatge.

Definició 2.1 *Anomenem element estructurant a un conjunt B amb centre o que interectua amb un altre conjunt X . El conjunt B és més simple, en quan a talla i forma, que el conjunt X .*

La morfologia matemàtica utilitza dos operadors bàsics: l'erosió i la dilatació a partir dels quals es poden definir tots els altres operadors morfològics. Per a poder-los definir necessitarem abans el concepte de traslladat d'un conjunt respecte d'un altre.

Definició 2.2 *Definim el conjunt B_x com el conjunt B traslladat per x*

$$B_x = \{b + x; b \in B\} \quad (2.1)$$

Ara ja podem definir l'erosió.

Definició 2.3 *Definim la erosió ε d'un conjunt X per un element estructurant B com el conjunt de punts $x \in X$, tals que, el traslladat de B respecte x , B_x , estigui inclòs dins X .*

$$\varepsilon = \{x; B_x \subset X\} \quad (2.2)$$

Si tenim en compte l'expressió de la resta de Minkowski:

$$X \ominus B = \bigcap_{b \in B} X_b \quad (2.3)$$

aleshores podem expressar l'erosió en termes d'aquesta resta

$$\begin{aligned} \varepsilon &= \{x; B_x \subset X\} = \\ &= \bigcap_{b \in B} X_{-b} = \\ &= \bigcap_{-b \in B} X_b = \\ &= X \ominus \check{B} \end{aligned} \quad (2.4)$$

on \check{B} és el que anomenem conjunt transposat i el definim com:

$$\check{B} = \{-b; b \in B\} \quad (2.5)$$

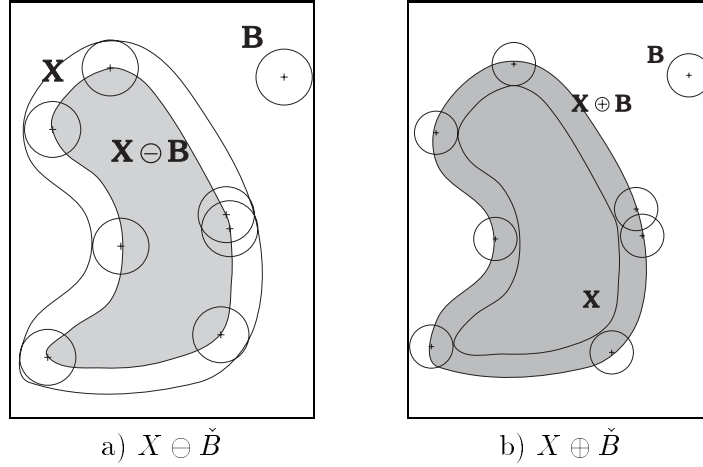


FIGURA 2.2: Exemples de l'erosió i de la dilatació del conjunt X per l'element estructurant \check{B}

L'altra operació bàsica de la morfologia és la dilatació, que compleix la dualitat respecte de l'erosió

Definició 2.4 *Definim la dilatació δ d'un conjunt X per un element estructurant B com el conjunt de punts x , tals que la intersecció dels traslladats d'aquests B_x amb X és no buida*

$$\delta = \{x; B_x \cap X \neq \emptyset\} \quad (2.6)$$

Es pot deduir de l'equació 2.4 que l'expressió analítica de la dilatació es pot escriure com:

$$X \oplus \check{B} = \bigcup_{b \in \check{B}} X_b \quad (2.7)$$

A partir d'aquestes expressions podem veure la dualitat de l'erosió i la dilatació:

$$X \ominus \check{B} = \bigcap_{-b \in B} X_b = \left[\bigcup_{-b \in B} X_b^c \right]^c = [X^c \oplus \check{B}]^c \quad (2.8)$$

Una vegada definides l'erosió i la dilatació anem a veure quines són les propietats algebraiques que compleixen:

1. L'erosió i la dilatació són transformacions creixents, $\forall X, Y \in E$, i $\forall B$ element estructurant

$$\text{Si } X \subset Y \implies \begin{aligned} X \ominus \check{B} &\subset Y \ominus \check{B} \\ X \oplus \check{B} &\subset Y \oplus \check{B} \end{aligned} \quad (2.9)$$

2. L'erosió és una transformació anti-extensiva i la dilatació és una transformació extensiva, $\forall X \in E$, i $\forall B$ element estructurant

$$X \ominus \check{B} \subset X \quad (2.10)$$

$$X \subset X \oplus \check{B} \quad (2.11)$$

3. De les propietats anteriors es poden deduir les següents propietats

$$\text{Si } B_1 \subset B_2 \implies \begin{aligned} X \ominus \check{B}_2 &\subset X \ominus \check{B}_1 \\ X \oplus \check{B}_1 &\subset X \oplus \check{B}_2 \end{aligned} \quad (2.12)$$

4. L'erosió i la dilatació no compleixen l'associativitat, però per altra banda compleixen les següents expressions:

$$(X \ominus \check{B}_1) \ominus \check{B}_2 = X \ominus (\check{B}_1 \oplus \check{B}_2) \quad (2.13)$$

$$(X \oplus \check{B}_1) \oplus \check{B}_2 = X \oplus (\check{B}_1 \oplus \check{B}_2) \quad (2.14)$$

aquesta propietat és la que permet calcular les operacions de dilatació i erosió amb la descomposició de l'element estructurant en elements més simples, problema que tractarem al capítol 4.

5. Ni l'erosió ni la dilatació són transformacions idempotents.

$$(X \ominus \check{B}) \ominus \check{B} = X \ominus (\check{B} \oplus \check{B}) \neq X \ominus \check{B} \quad (2.15)$$

$$(X \oplus \check{B}) \oplus \check{B} = X \oplus (\check{B} \oplus \check{B}) \neq X \oplus \check{B} \quad (2.16)$$

6. Distributivitat de l'erosió i de la dilatació

$$X \ominus (\check{B}_1 \cup \check{B}_2) = (X \ominus \check{B}_1) \cap (X \ominus \check{B}_2) \quad (2.17)$$

$$X \oplus (\check{B}_1 \cup \check{B}_2) = (X \oplus \check{B}_1) \cup (X \oplus \check{B}_2) \quad (2.18)$$

Totes les definicions vistes són nocions orientades als conjunts, per tant s'adapten molt bé per a l'estudi d'imatges binàries. Per a estendre aquestes definicions a altres tipus d'imatges podem fer servir la noció de reticle complet.

Definició 2.5 *Direm que un conjunt T té estructura de reticle complet si*

1. *Existeix un ordre parcial \geq sobre T .*
2. *Per a qualsevol família (A_i) de T existeix*
 - *un operador de màxim $\vee A_i$ anomenat sup.*
 - *un operador de mínim $\wedge A_i$ anomenat inf.*

Així doncs, els conjunts amb la inclusió com a relació d'ordre i amb les operacions d'unió i intersecció com *sup* i *inf* respectivament, formen el reticle dels conjunts o reticle booleà.

Si considerem una imatge amb nivells de grisos com una funció $f : E \rightarrow \mathbb{R}^+$ podem definir el reticle de les funcions semi-continues superiorment F mitjançant la següent relació d'ordre:

$$f \leq g \text{ si } \forall x \in E, f(x) \leq g(x) \quad (2.19)$$

i considerant com *sup* i *inf* els donats per les següents relacions:

$$f = \vee f_i \iff f(x) = \sup f_i(x), \forall x \in E \quad (2.20)$$

$$f' = \wedge f_i \iff f'(x) = \inf f_i(x), \forall x \in E \quad (2.21)$$

Sobre aquest reticle de funcions es poden definir anàlogament les mateixes operacions morfològiques d'erosió i dilatació que havíem definit pel reticle dels conjunts, conservant totes les propietats vistes [91, 93]. Les definicions són les següents:

Definició 2.6 *Siguin $F, G \subseteq E$, $f : F \rightarrow \mathbb{R}^+$ i $g : G \rightarrow \mathbb{R}^+$. L'erosió de f per g , denotada per $f \ominus g$ es defineix com:*

$$(f \ominus g)(x) = \max_{\substack{i \in F \\ i-x \in G}} \{f(i) - g(i-x)\} \quad (2.22)$$

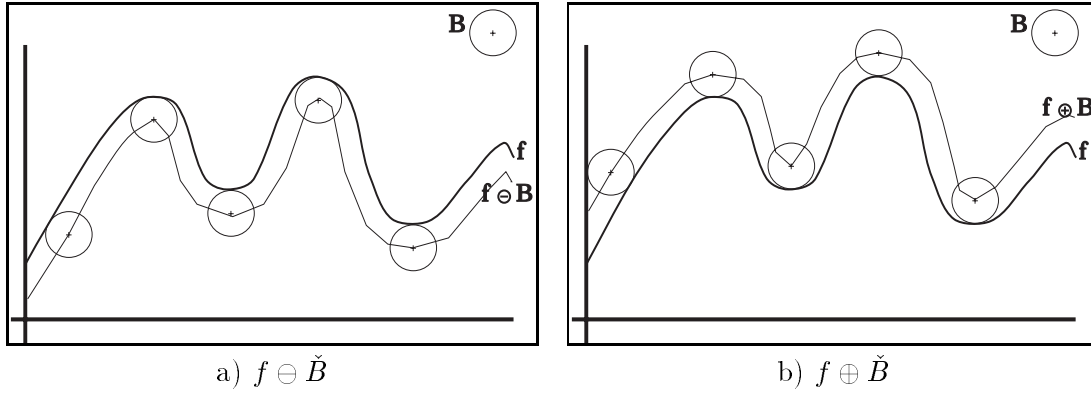


FIGURA 2.3: Exemples de l'erosió i de la dilatació de la funció f per l'element estructurant \check{B}

Per dualitat definirem dilatació com:

$$(f \oplus \check{g})(x) = \min_{\substack{i \in F \\ i-x \in G}} \{f(i) + g(i-x)\} \quad (2.23)$$

Molt sovint, per raons de simplicitat i per conservar importants propietats en l'aplicació dels algorismes, es treballa amb elements estructurants plans, $g : G \rightarrow \{0, 1\}$. En aquest cas particular, la definició d'erosió i dilatació, d'una funció f , per un element estructurant pla, B , s'expressa com:

$$\begin{aligned} (f \ominus \check{B})(x) &= \min_{i \in B} f(x-i) = \min_{i \in B_x} f(i) \\ (f \oplus \check{B})(x) &= \max_{i \in B} f(x-i) = \max_{i \in B_x} f(i) \end{aligned} \quad (2.24)$$

2.2.2 Definició de granulometria

La formalització del concepte de granulometria sobre una imatge és donada l'any 1975 per G. Matheron [69, 37]. En aquests treballs es defineixen els axiomes que ha de complir una transformació per ser una bona eina de càlcul de la distribució de la talla. La transformació depèn d'un paràmetre que representa la mida del forat del tamís.

Definició 2.7 Una família de transformacions paramètriques (ϕ_λ) on $\lambda > 0$ s'adapta bé al càlcul d'una granulometria si:

1. $\forall \lambda > 0, \phi_\lambda$ és una obertura algebraica.
2. Es compleix la propietat d'estabilitat respecte dels paràmetres, això vol dir que es compleix com a mínim una de les tres següents condicions:
 - $\phi_\lambda \circ \phi_\mu = \phi_\mu \circ \phi_\lambda = \phi_{\text{Sup}(\mu, \lambda)} \quad \forall \lambda, \mu \geq 0$
 - $\phi_\lambda \leq \phi_\mu \quad \forall \lambda, \mu \geq 0$
 - $B_\lambda \subseteq B_\mu \quad \forall \lambda, \mu \geq 0$

Per poder entendre aquest concepte cal que definim l'obertura algebraica:

Definició 2.8 Anomenem obertura algebraica a la transformació ϕ definida sobre X que compleix les següents propietats:

1. *Antiextensivitat*

$$\phi(X) \subset X \quad (2.25)$$

2. *Preservació de la inclusió (és a dir, ϕ és creixent)*

$$Y \subset X \implies \phi(Y) \subset \phi(X) \quad (2.26)$$

3. *Idempotència*

$$\phi[\phi(X)] = \phi(X) \quad (2.27)$$

Per tant, qualsevol família de transformacions paramètriques que compleixin les propietats d'antiextensivitat, creixement, idempotència i conservin l'estabilitat respecte dels paràmetres serà bona per a realitzar un procés de tamissatge sobre una imatge. L'aplicació d'aquest concepte de granulometria es presenta molt bé en el treball de M. Coster i J.L. Chermant [27].

En funció de com es mesurin els objectes que resten dins del tamís, és a dir, les imatges resultants d'aplicar les transformacions que acabem de definir, tindrem dos tipus de granulometries diferents

Granulometria en nombre es compta el nombre d'objectes que queden dins del tamís.

Granulometria en mesura s'aplica una mesura global sobre el conjunt dels objectes que queden al tamís.

Dimensió de l'espai	Funcionals de Minkowski d'ordre k			
	$k = 0$	$k = 1$	$k = 2$	$k = 3$
$n = 3$	Volum	Superfície	Norma	Connexitat
$n = 2$	Superfície	Perímetre	Connexitat	
$n = 1$	Perímetre	Connexitat		
$n = 0$	Connexitat			

TAULA 2.1: Funcionals de Minkowski W_k^n sobre $\mathcal{C}(\mathbb{R}^n)$.

Per a poder fer una granulometria en nombre caldrà definir dos conceptes:

- Com es presenten els objectes a la imatge.
- Quin és l'algorisme que ens permet detectar i posteriorment comptar els objectes.

Per altra banda en una granulometria en mesura caldrà elegir quina és la mesura que es fa sobre les imatges transformades. Al treball de P. Soille et al. ([101, 100]) es presenten resultats molt interessants de la geometria integral sobre quins són els funcionals bàsics (entenen com a funcional un paràmetre global associat a un conjunt) d'una mesura per a que aquesta sigui vàlida tan matemàtica com físicament, i com s'adapta tot això a les mesures d'imatges amb nivells de grisos.

La geometria integral ha aportat un interessant resultat en quan a la definició de mesures sobre l'espai $\mathcal{C}(\mathbb{R}^n)$ (classe dels conjunts compactes convexes no buits de \mathbb{R}^n). Es demostra que les mesures vàlides seran aquelles que siguin combinacions lineals dels funcionals de Minkowski (veure taula 2.1), ja que aquests compleixen les següents propietats de creixement, invariància per translació, homogeneïtat, additivitat i continuïtat. Per a qualsevol conjunt compacte i convexe de \mathbb{R}^n existeixen $n + 1$ funcionals de Minkowski. En el cas \mathbb{R}^3 aquests funcionals són el volum, l'àrea, la norma i el número d'Euler. Per tant per a una granulometria en mesura seran interessants totes aquelles mesures que es basin en un funcional d'aquests.

Després d'això podem dir que una granulometria en mesura presentarà un caràcter més general que una granulometria en nombre, ja que aquesta només pot donar informació que relacioni la talla i el número de partícules, en canvi una granulometria en mesura podrà donar informacions més generals sobre la imatge.

Fins ara hem vist que per a la construcció d'una granulometria cal seguir primerament dos passos:

- Donada una família de transformacions, (ϕ_λ) , calcular totes les imatges transformades per aquesta família.
- Per a cada una de les imatges resultants de la transformació s'ha de comptar el número de partícules, o bé determinar una mesura.

Després de tot això, ara només queda per definir la corba granulomètrica que permetrà relacionar tota la informació que s'extreu de la imatge en el procés de la granulometria. Anem a veure com es defineix depenent del tipus de granulometria.

Per a les granulometries en nombre la funció de distribució es defineix com

$$F(\lambda) = \frac{N(X) - N(\phi_\lambda(X))}{N(X)} \text{ on } 0 < F(\lambda) < 1 \quad (2.28)$$

on $N(X)$ representa el nombre de grans inicials i $N(\phi_\lambda(X))$ representa el nombre de grans després d'haver passat per un tamís de talla λ .

Per a les granulometries en mesura es defineix la funció de distribució com

$$G(\lambda) = \frac{Mes(X) - Mes(\phi_\lambda(X))}{Mes(X)} \text{ on } 0 < G(\lambda) < 1 \quad (2.29)$$

on Mes representa la mesura que es faci, que vindrà definida en termes d'algun dels funcionals de Minkowski.

Una vegada definida les funcions de distribució podem obtenir les funcions de densitat $f(\lambda)$ i $g(\lambda)$, com les derivades de $F(\lambda)$ i $G(\lambda)$

$$F(\lambda) = \int_0^\lambda f(\lambda)d\lambda \quad (2.30)$$

$$G(\lambda) = \int_0^\lambda g(\lambda)d\lambda \quad (2.31)$$

i que com a funcions de densitat de probabilitat han de complir que

$$\int_0^\lambda f(\lambda)d\lambda = 1 \quad (2.32)$$

$$\int_0^\lambda g(\lambda) d\lambda = 1 \quad (2.33)$$

En el cas de la granulometria en nombre, $f(\lambda)$ correspon al nombre d'objectes de talla λ , i que denotem com $n(X, \lambda)$, dividit pel nombre total d'objectes

$$f(\lambda) = \frac{n(X, \lambda)}{N(X)} \quad (2.34)$$

D'aquesta manera queda definit el concepte general de granulometria. Seguint el procés que s'ha presentat aquí podem definir diferents granulometries en funció del tipus d'informació que es vulgui extreure de la imatge. A l'hora de definir-la cal fer una bona elecció de quines són les obertures algebraiques que més bé s'adapten al problema i com s'ha de definir el paràmetre de les obertures perquè es pugui definir una bona corba granulomètrica. A més a més s'haurà d'escollir el tipus de granulometria i si és en mesura s'haurà de triar també la mesura.

2.2.3 Granulometria sobre el volum

En aquest apartat donem l'algorisme 2.1 que s'ha dissenyat pel càlcul de la granulometria en volum. La família de transformacions es denota amb (ϕ_{λ_i}) i la imatge f definida per $f : X \rightarrow Y$ on $X \subset \mathbb{N}^2$ i el conjunt Y ve donat normalment per l'interval $[0, 255]$. En aquest cas la seqüència $\{\lambda_{i_{min}}, \dots, \lambda_i, \dots, \lambda_{i_{max}}\}$ defineix la mida dels elements estructurants de les obertures algebraiques que s'aplicaran. La funció de distribució $G(\lambda)$ es calcula a partir de la diferència de volum que s'elimina després d'aplicar cada transformació.

Per a poder definir bé la funció de densitat de probabilitat $g(\lambda)$ cal que la seva derivada sigui igual a 1, com hem vist a l'expressió (2.33). És per això que hem introduït el factor η .

Des d'un punt de vista teòric la granulometria consisteix en un procés de filtratge on anem augmentant un paràmetre que va eliminant en cada pas una porció més gran de volum, fins que és eliminat tot el volum de la imatge. Ara bé, des d'un punt de vista més pràctic ens interessa la informació granulomètrica per a un conjunt més restringit de paràmetres, aleshores no cal invertir temps de càlcul en filtratge per paràmetres més grans dels que no ens interessa la informació estreta.

A l'aplicació que veurem hem seleccionat un valor λ_{max} a partir del qual no calcularem la resta d'obertures, d'aquesta manera no arribarem a una transformació que elimini tot el volum de la imatge. És per això que en un principi s'ha seleccionat com a factor de

Algorisme 2.1 *Granulometria en volum de la imatge $f : X \rightarrow Y$ pel conjunt de paràmetres $\{\lambda_{i_{min}}, \dots, \lambda_{i_{max}}\}$*

1. $i = i_{min}$
2. Mentre ($i \neq i_{max}$) fer
 - (a) Aplicació de la transformació de radi λ_i .
 - (b) Càlcul del volum $V(\phi_{\lambda_i}(f)) = \sum_{(x,y) \in X} f(x, y)$.
 - (c) Incrementar i .
3. Càlcul de $G(\lambda_i) = \frac{V(f) - V(\phi_{\lambda_i}(f))}{\eta} \quad \forall \lambda_i : i \in \{i_{min}, \dots, i_{max}\}$
4. Càlcul de la derivada $g(\lambda) = \frac{\delta G(\lambda)}{\delta \lambda}$.

normalització, η , el volum total d'imatge eliminat per l'obertura de paràmetre més gran,

$$\eta_1 = V(f) - V(\phi_{\lambda_{max}}(f)) \quad (2.35)$$

D'aquesta manera obtenim una corba de densitat $g(\lambda)$ amb integral igual a 1. Encara que, per altra banda ens apareixerà un problema a l'hora d'interpretar la corba granulomètrica.

Per interpretar un corba d'aquest estil ens interessa veure per quins valors es donen els màxims de la corba, és a dir, quin és el paràmetre que extreu més volum de la imatge. Aleshores resulta molt interessant el poder interpretar els salts que dona la corba entre cada paràmetre.

Si fem la normalització amb η_1 estem normalitzant respecte del volum que hem extret de la imatge en qüestió, aleshores quan comparem la corba d'imatges diferents estarem comparant variacions que no tenen la mateixa importància en una imatge que en l'altra.

Per evitar aquest problema proposem que el factor de normalització sigui el volum de tota la imatge

$$\eta_2 = V(f) \quad (2.36)$$

de manera que afegim un valor més al final de la corba que es suposa com una transformació final que elimina tot el volum que resta de la imatge. Amb això aconseguim fer una corba que permeti una interpretació més coherent entre les diferents imatges.

Tot i així, no podem dir que les corbes estiguin completament normalitzades, sempre que desconeguem l'anamorfoosi introduïda per l'aparell de captació de la imatge. D'aquest efecte en parlem a l'apartat 3.5 de [110] quan introduïm la propietat de la dimensionalitat del volum. Una total normalització de totes les corbes només es podria aconseguir amb un sistema de captació tancat, és a dir, un sistema en el que poguéssim conèixer els paràmetres de l'aparell en el moment de la captació de cada imatge. En els resultats que exposarem més endavant s'indicarà si la normalització ha estat feta amb el factor η_1 o η_2 .

2.2.4 Obertures algebraiques

Fins ara hem vist com es defineix una granulometria en general i com es construeixen les corbes granulomètriques. Però el pas previ a tot aquest procés passa sempre per definir quina és l'obertura algebraica i com és la relació entre els paràmetres d'aquesta família

d'obertures.

A l'apartat 2.2.2 hem vist la definició d'una obertura algebraica com tota transformació creixent, antiextensiva e idempotent. Podem trobar molts exemples de transformacions, que compleixen aquestes propietats de tamís, al treball de J. Serra i L. Vincent [94].

L'obertura més simple és la transformació d'umbralització fent servir un criteri creixent C . Direm que un criteri C és creixent si compleix les següents relacions

$$\text{Si } X \text{ satisfà } C \text{ i } Y \geq X \implies Y \text{ satisfà } C \quad (2.37)$$

$$\text{Si } X \text{ no satisfà } C \text{ i } Y \leq X \implies Y \text{ no satisfà } C \quad (2.38)$$

Aleshores a partir d'un criteri creixent podem definir una família d'obertures algebraiques amb transformacions d'umbral:

$$\phi_C(X) = \begin{cases} X & \text{si } X \text{ satisfà } C \\ 0 & \text{en altre cas} \end{cases} \quad (2.39)$$

per tant, la família de transformacions (ϕ_C) podria servir per a definir una granulometria.

Altres exemples d'obertures algebraiques que s'adapten molt bé per a una granulometria són

- Obertures morfològiques.
- Erosions amb reconstrucció.
- Suprem d'obertures morfològiques.

i que presentem en els següents apartats.

Obertura morfològica Donarem la definició d'obertura morfològica a partir de les definicions d'erosió i de dilatació donades per a imatges amb nivells de grisos.

Definició 2.9 *Una obertura morfològica és la transformació resultat de realitzar consecutivament una erosió seguida de la seva dilatació dual, es denota com f_B*

$$f_B = (f \ominus \check{B}) \oplus B \quad (2.40)$$

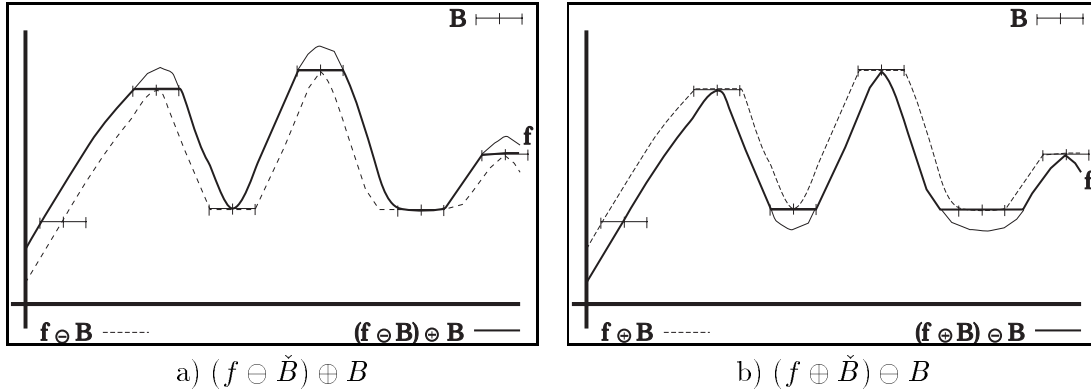


FIGURA 2.4: Exemples de l'obertura i del tancament de la funció f per l'element estructurant \check{B}

Per l'obertura també existeix una operació dual que és el tancament

Definició 2.10 *Un tancament morfològic és la transformació resultat de realitzar consecutivament una dilatació seguida de la seva erosió dual, es denota com f^B*

$$f^B = (f \oplus \check{B}) \ominus B \quad (2.41)$$

A la figura 2.4 podem veure el resultat d'una obertura i un tancament sobre una funció unidimensional. L'efecte d'una obertura és el d'eliminar de la funció tots aquells pics en els que l'element estructurant no hi cap, les altres parts de la funció queden igual. Podem veure doncs un efecte de tamís, en el que després de la transformació només quedaran aquells pics de la funció amb base més gran que l'element estructurant.

Després d'això és fàcil veure que l'obertura morfològica és una transformació que compleix les propietats d'una obertura algebraica: antiextensivitat, creixement i idempotència. Per altra banda el tancament no les compleix ja que és una transformació extensiva.

L'obertura morfològica és per tant una transformació apropiada pel càlcul d'una granulometria, en la que el paràmetre es correspondrà amb l'element estructurant dels operadors morfològics que la defineixen. Si construïm una família d'elements estructurants que permetin complir la condició d'estabilitat de la seqüència d'obertures, obtindrem una família de transformacions adaptables per a una granulometria. A la figura 2.5 podem veure els resultats d'una granulometria on hem canviat el radi de l'element estructurant.

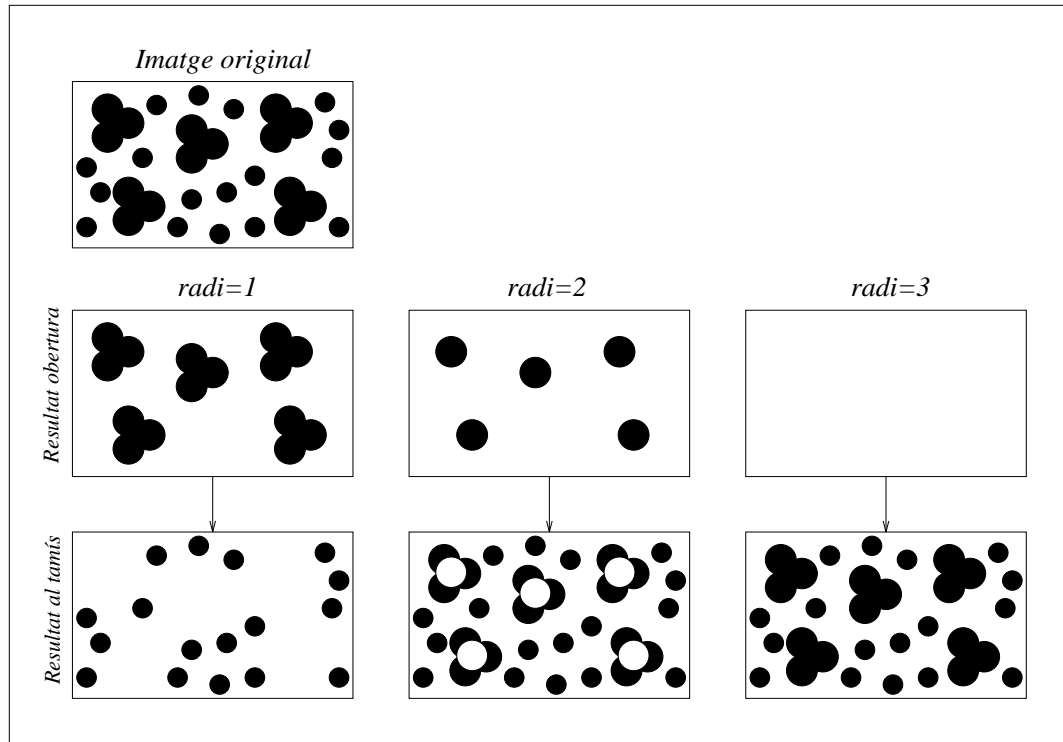


FIGURA 2.5: Seqüència d'imatges resultants d'aplicar obertures morfològiques amb diferents elements estructurants rodons, abaix donem les imatges que queden dins del tamís.

Erosió amb reconstrucció Una altra obertura algebraica que permet calcular una granulometria és l'erosió seguida de reconstrucció geodèsica. Abans de definir aquesta obertura donarem algunes definicions sobre transformacion geodèsiques. Primer de tot definirem la distància geodèsica

Definició 2.11 Donat el conjunt $X \subset \mathbb{N}^2$ es defineixi la distància geodèsica entre dos punts p_1 i p_2 de X com el camí de longitud mínima, si existeix, entre p_1 i p_2 , que està inclòs dins X

$$d_X(p_1, p_2) = \inf\{l(C_{p_1, p_2}), C_{p_1, p_2} \subset X\} \quad (2.42)$$

on C_{p_1, p_2} representa un camí entre els punts p_1 i p_2 .

També podem definir la bola geodèsica

Definició 2.12 Anomenem bola geodèsica de radi $n \in \mathbb{N}$ i de centre $p \in X$ al conjunt $B_X(p, n)$ definit per

$$B_X(p, n) = \{p' \in X, d_X(p', p) \leq n\} \quad (2.43)$$

I a partir d'aquestes podem definir l'erosió i la dilatació geodèsica

Definició 2.13 La dilatació geodèsica $\delta_X^{(n)}(Y)$ de mida n sobre el conjunt Y i dins del conjunt X ve donada per

$$\delta_X^{(n)}(Y) = \bigcup_{p \in Y} B_X(p, n) = \{p' \in Y, \exists p \in Y, d_X(p, p') \leq n\} \quad (2.44)$$

Definició 2.14 L'erosió geodèsica $\varepsilon_X^{(n)}(Y)$ de mida n sobre el conjunt Y i dins del conjunt X ve donada per

$$\varepsilon_X^{(n)}(Y) = \{p \in Y, B_X(p, n) \subseteq Y\} = \{p \in Y, \forall p' \in X/Y, d_X(p, p') \leq n\} \quad (2.45)$$

Una dilatació geodèsica es pot definir fàcilment a partir d'una dilatació, això és:

$$\delta_X^{(1)}(Y) = (Y \oplus H) \cap X \quad (2.46)$$

una dilatació geodèsica de mida n és obtinguda a partir de la següent expressió

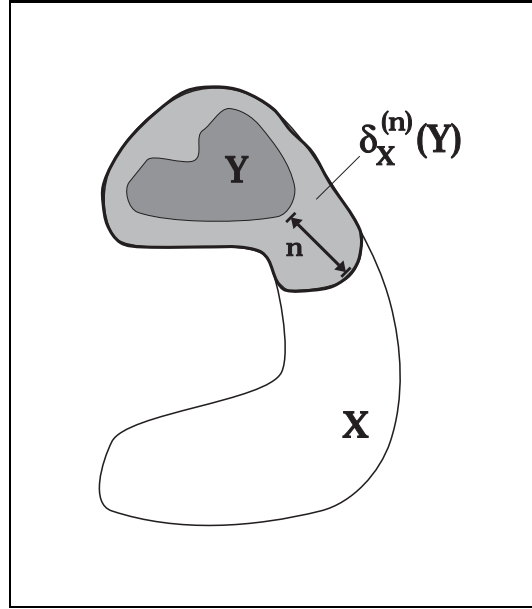


FIGURA 2.6: Exemples d'erosions i dilatacions geodèsiques del conjunt Y dins del conjunt X .

$$\delta_X^{(n)}(Y) = \delta_X^{(1)}(\delta_X^{(1)} \dots n \dots (\delta_X^{(1)}(Y))) \quad (2.47)$$

Si es realitzen successives dilatacions geodèsiques sobre un conjunt Y dins d'un conjunt X , és impossible connectar conjunts de Y que no estiguin continguts dins una component connexe de X . A més a més si realitzem una seqüència de dilatacions prou gran podem arribar a obtenir exactament els conjunts inicials X que contenen a Y .

Com que tractem amb conjunts finits, podem assegurar que existeix un valor n_0 tal que

$$\forall n > n_0, \delta_X^{(n)}(Y) = \delta_X^{(n_0)}(Y) \quad (2.48)$$

aleshores podem dir que en el pas n_0 ja hem reconstruït completament la component X . Aquesta operació és anomenada reconstrucció

Definició 2.15 *La reconstrucció $r_X(Y)$ del conjunt X finit, a partir del conjunt Y , ve donada per la següent fórmula:*

$$r_X(Y) = \lim_{n \rightarrow +\infty} \delta_X^{(n)}(Y) \quad (2.49)$$

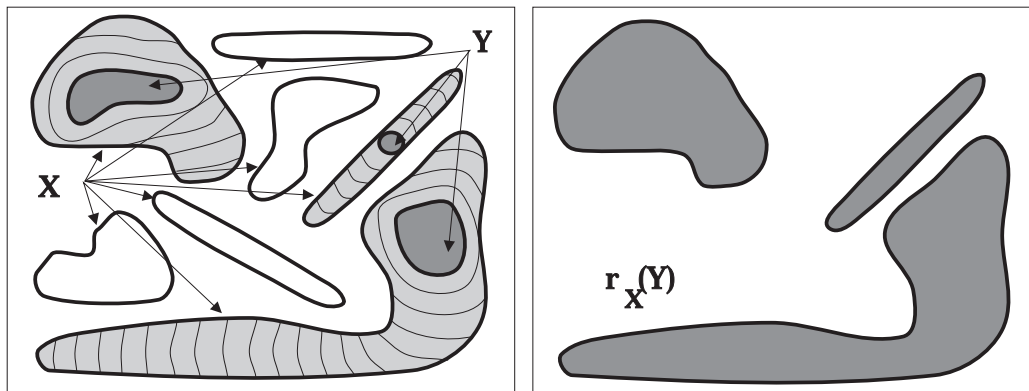


FIGURA 2.7: Exemple de la reconstrucció del conjunt Y dins de X .

Ara ja podem definir la nova obertura algebraica

Definició 2.16 *l'obertura algebraica basada en erosió seguida de reconstrucció geodèsica, sobre el conjunt X i amb l'element estructurant B_r , es defineix com:*

$$r_X(X \ominus B_r) \quad (2.50)$$

Passant totes aquestes definicions de conjunts a funcions podem definir l'algorisme pel càlcul de la granulometria, amb aquesta nova obertura substituint el punt 2.a de l'algorisme 2.1 per l'algorisme 2.2.

Suprem d'obertures morfològiques En el treball de J. Serra i L. Vincent [94] es donen els lligams entre les obertures algebraiques i les obertures morfològiques. Un resultat interessant és el següent teorema:

Teorema 2.1 *Una aplicació $\gamma : \mathcal{T} \rightarrow \mathcal{T}$ és una obertura si i només si aquesta és el suprem d'una família (γ_i) d'obertures morfològiques. A més a més, si una translació és definida sobre \mathcal{T} , γ és invariant per translació si i només si les γ_i són invariants per translació.*

Aquest teorema ens dóna la possibilitat de definir una nova obertura algebraica a partir d'altres obertures més simples.

Si agafem com a transformació el suprem de les obertures per elements lineals en totes les direccions, tenint en compte el teorema 2.1 podem assegurar que aquesta nova transformació és també una obertura algebraica

Algorisme 2.2 *Erosió amb reconstrucció*

1. Càlcul de l'erosió de la imatge amb un element estructurant rodó de radi r

$$g = f \ominus \check{B}_r \tag{2.51}$$

2. Calculem la dilatació geodèsica de g dins de f

$$\delta_g(f) = \min\{f, g \oplus H\} \tag{2.52}$$

on H correspon a un element estructurant quadrat i plà 3×3 .

3. La reconstrucció és el resultat d'aplicar consecutivament la dilatació geodèsica de g sobre ella mateixa, fins arribar a la idempotència

$$r_g^{n_0}(f) = \delta_g(\delta_g(\dots n_0 \dots (\delta_g(f)))) \tag{2.53}$$

on $\forall n > n_0, r_g^{n_0}(f) = r_g^n(f)$.

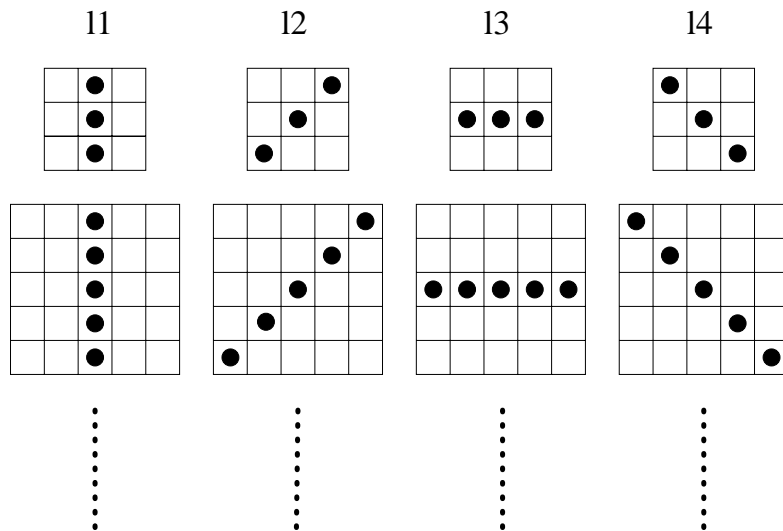


FIGURA 2.8: Elements estructurants lineals en totes les direccions

$$\gamma(f) = \sup\{f_{l_1}, f_{l_2}, f_{l_3}, f_{l_4}\} \quad (2.54)$$

Per a la construcció de la granulometria agafarem una família de transformacions $(\gamma_{\lambda_i}(f))$ en la que λ_i correspon a

$$\lambda_i = \frac{\text{longitud}(l_i)}{2} \quad \forall i \in \{1, 2, 3, 4\} \quad (2.55)$$

D'aquesta manera és fàcil veure que es compleix l'estabilitat respecte del paràmetre, ja que $\gamma_{\lambda_i}(f) \leq \gamma_{\lambda_j}(f)$ per qualsevol $i \geq j$.

2.3 Aplicació a la classificació d'imatges radiogràfiques

En aquest apartat mostrem els resultats que hem obtingut d'aplicar la granulometria com a mesura de la granularitat. Ha estat aplicada a l'automatització del procés de classificació d'imatges radiogràfiques. Més concretament a radiografies de tórax afectades de pneumoconiosi. Seguidament donem una petita introducció al problema de la pneumoconiosi així com l'aplicació de la granulometria i els resultats obtinguts.

2.3.1 La pneumoconiosi

La pneumoconiosi és la malaltia que és conseqüència d'alteracions broncopulmonars que són produïdes per la constant inhalació de pols mineral o metàl·lic ([8] capítol 15). Els símptomes de la pneumoconiosi es presenten sovint en treballadors de mines de silici, carbó, amiant, ferro, etc., o bé, en treballadors d'indústries que manipulen aquests productes o derivats d'ells.

La pel·lícula radiogràfica és la peça essencial per a la diagnòsi de pneumoconiosi, ja que en ella els malalts presenten signes molt característics, com és l'aparició de nòduls o petites opacitats, grans opacitats, lesions pleurals, etc.

Aquestes bronconeumopaties per inhalació de pols inorgànic, mineral o metàl·lica es desenvolupen de diferent manera, depenent del tipus de partícula inorgànica que l'ha provocada. Així doncs, tenim diferents tipus de pneumoconiosi, com ara la silicosi, l'antracosi, l'asbestosi, la beriliosi, la siderosi, entre d'altres.

Degut a la importància social de la pneumoconiosi en el camp del treball, s'han definit

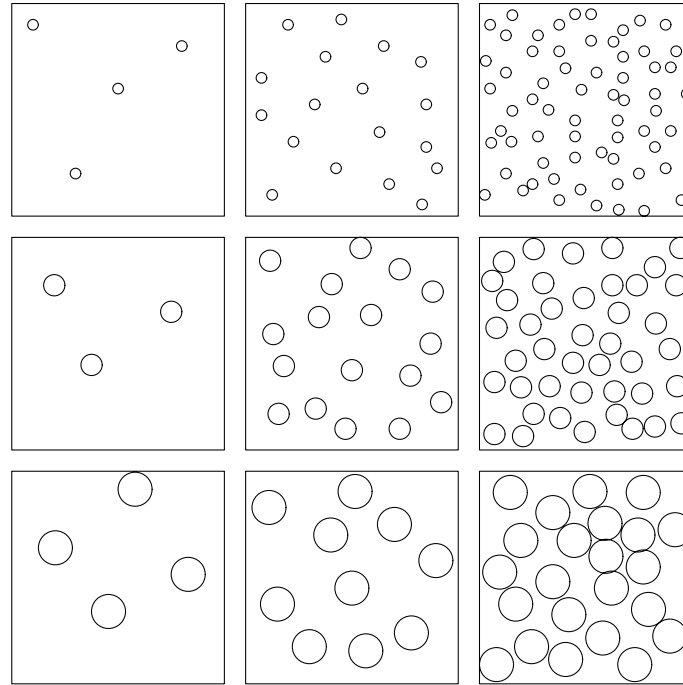


FIGURA 2.9: Esquema de com es la distribució d'opacitats en funció del tamany i la densitat.

classificacions radiològiques que faciliten el procés de classificació. S'han fet varies classificacions, i una de les més importants i potser de les més complicades, degut al grau de detall que presenta, és la classificació internacional de la ILO de 1980 [43].

Un dels símptomes essencials d'aquesta malaltia, és l'aparició d'engruiximents en el parènquima que apareixen a la radiografia com *petites opacitats rodones*. Aquests engruiximents de teixit es produeixen com a resposta a les partícules estranyes, que s'introdueixen en els pulmons, en el procés d'inhalació d'aire, en ambients amb determinades condicions. Aquestes situacions es presenten sovint en llocs on es troben moltes partícules inorgàniques o metàl·liques en suspensió a l'aire, com ara mines o indústries on es tracten certs tipus de materials. Una constant exposició a situacions com aquestes fa que es provoqui la malaltia, i per tant el nombre d'opacitats petites va augmentant.

Així doncs aquestes opacitats rodones seran un símptoma de pneumoconiosi. Es podran presentar en diversos graus de densitat, depenent de la concentració d'opacitats que es presenti, i diversos graus de mida, en funció del diàmetre de l'opacitat. Podem veure un exemple de com es distribueixen aquestes opacitats en funció d'aquests dos factors: densitat

Interval diàmetre (d)	Etiqueta	Pixels Imatges
$d < 1.5mm$	p	8.55
$1.5mm < d < 3mm$	q	17.1
$3mm < d < 10mm$	r	57.0

TAULA 2.2: Correspondència entre el diàmetre de l'opacitat i l'etiqueta assignada en els estàndards. També es dona la correspondència amb el número de pixels de cada diàmetre sobre la imatge. Veure el procés de discretització a l'apèndix sobre l'equipament

No Opacitats	0/0 (Example 1)			
Opacitats petites rodones	$\Delta densitat$			
	- \rightarrow +			
	-	1/1 p/p	2/2 p/p	3/3 p/p
	$\Delta mida \downarrow$	1/1 q/q	2/2 q/q	3/3 q/q
+	1/1 r/r	2/2 r/r	3/3 r/r	

TAULA 2.3: Imatges dels estàndards. Particularitats de les imatges: 0/0 (Example 1) Il·lustra bé el disseny vascular. 2/2 p/p és massa clara. 3/3 q/q, 1/1 r/r i 2/2 r/r són de qualitat insuficient

i mida, a la figura 2.9.

En els estàndards [43] la profusió o densitat d'opacitats petites ve donada en quatre grans grups 0, 1, 2, o 3 per comparació amb les imatges estàndards. Per altra banda la classificació del tamany es fa per la mesura del diàmetre de la opacitat, es classificaran com p, q o r segons la mida del diàmetre. La correspondència entre la mida del diàmetre i la classificació que rep es pot veure a la taula 2.2.

2.3.2 Automatització de la classificació

L' automatització en la classificació d'imatges radiogràfiques de tòrax afectades de pneumoconiosi és un problema molt complicat per esser afrontat d'un sol cop. Una classificació completa suposa la consideració d'un conjunt molt ampli d'aspectes radiogràfics. Aspectes molt diferents i que requereixen d'una gran quantitat de coneixement. En aquest treball es presentarà la possibilitat de realitzar un sistema de visió capaç de donar la classificació de la imatge considerant només les *petites opacitats rodones*, que són els símptomes essencials

per a la diagnosi.

Per aquest propòsit es tractarà principalment amb un conjunt de 10 imatges que correspondrà a la classificació abreujada. Les imatges seleccionades per a la comprovació del sistema seran les imatges dels estàndards, que venen etiquetades amb els noms que donem a la taula 2.3.

Aleshores podem dir que estem davant un problema de classificació automàtica d'imatges, dins d'unes certes classes que venen donades per unes imatges model. Per tant, la solució vindrà donada amb el dos següents passos:

1. Mesura sobre la imatge.
2. Classificació en funció de la mesura feta i de les imatges models.

En el primer pas s'han d'extreure mesures de certes característiques de la imatge. Les mesures que es triïn han de donar informació sobre:

1. La distribució i densitat de les opacitats.
2. La talla de les opacitats.

Ara que hem introduït el problema de forma general, anem a veure'l des del punt de vista de les imatges. Una radiografia toràcica és una imatge que presenta sempre uns elements molt concrets, les parets toràciques, que estan sobreposades al pulmó, i que presenten varis elements com les clavícules, la columna vertebral, les costelles i el diafragma; per altra banda tenim la pleura, el mediastí, el parènquima pulmonar (invisible en condicions normals) i tota la vascularització pulmonar. Aleshores les anomalies que hem de mesurar, les opacitats, es presentaran sobreposades en una imatge com la que acabem de descriure. Les opacitats es veuran principalment, a tota la zona parenquimal, és a dir, a les zones intercostals.

L'anàlisi que fa el radiòleg quan compara una placa radiogràfica d'un pacient amb les plaques dels estàndards, no consisteix en fer una correspondència exacta entre les dues imatges. I si tenim en compte que els radiòlegs dediquen molt poc temps a la classificació de cada una de les radiografies, això fa pensar que el que realitzen no és un procés de contacte o atencional, sinó més bé correpon a una classificació preatentiva de la imatge dels aspectes més significatius o anormals, com són les opacitats.

D'aquí que hagim decidit tractar el problema com un problema de classificació de textures. Ja que aquestes opacitats es distribueixen de forma aleatòria, i quan la densitat és

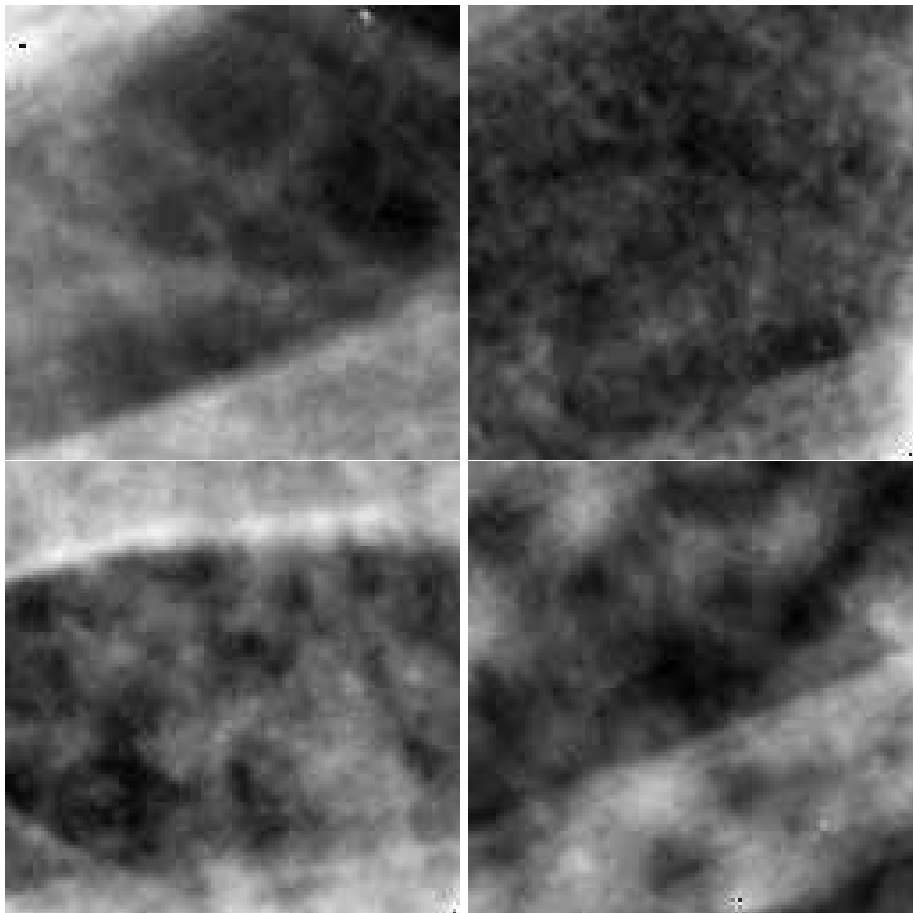


FIGURA 2.10: Exemples de les textures que es presenten en el parènquima pulmonar. D'esquerra a dreta i de dalt a baix corresponen a un pulmó sà en el que s'hi pot apreciar la vascularització pulmonar, un 3/3 p/p, 3/3 q/q i 3/3 r/r amb densitats altes d'opacitats i variacions en la talla d'aquestes.

més alta es localitzen totes juntes formant una textura. A la figura 2.10 es poden veure exemples d'aquestes textures.

Per tant si abans havíem definit el problema com la classificació d'imatges, ara que s'ha definit el contingut de les imatges, es pot especificar millor el problema dient que consisteix en aconseguir la classificació de les textures que apareixen a les imatges en qüestió. Aquesta classificació es durà a terme a partir de mesurar la granularitat de la imatge, extraient informació de la densitat i de la mida de les opacitats.

El primer pas consisteix en aïllar de forma automàtica les zones que ocupa el parènquima.

Aquest és un problema que està resolt en varis treballs [17], [13]. Una vegada aïllat el parènquima s'ha de procedir a analitzar les opacitats. Aquestes opacitats o textures anòmales que apareixen en el parènquima pulmonar estan superposades al que podríem anomenar un *fons*, format per parts de costelles o vasos sanguinis. Aquesta qüestió determina bastant la manera com s'aplicaran les mesures sobre les imatges.

Abans de passar a veure els resultats de calcular mesures granulomètriques sobre aquestes imatges veurem un resum d'alguns treballs anteriors sobre aquest problema.

Treballs anteriors Hi ha referències bibliogràfiques sobre aquest tema a partir de la publicació dels estàndards de la ILO de 1971. Podem trobar diferents aproximacions que pretenen automatitzar aquest mateix procés. Seguidament veurem aquests treballs que hem separat en dos grups: els que mesuren la textura del pulmó i els que busquen directament les opacitats com a parts de la imatge.

Mètodes que mesuren la textura En aquest apartat hem reunit tots els mètodes que mesuren la textura de certes zones de la imatge. Per fer això, en tots ells es treballa en zones parcials de la imatge. Zones on apareixi una textura més o manco homogènia, és a dir, sense influència del fons del que hem parlat a l'apartat anterior. A la taula 2.4 donem un resum de tots aquests treballs que ara anem a explicar.

En el treball presentat per R.P. Kruger [57] s'apliquen de forma complementària dues tècniques, fent una discriminació textural en els dominis espacial i freqüencial de la imatge, primer aplicant tècniques de mesura d'imatges digitals amb textures i després fent una aproximació usant mètodes òptics-digitals fent servir la transformada de Fourier òptica.

En el domini espacial fa una extracció de zones quadrades dels espais entre costelles, les dades de totes aquestes zones són combinades per formar un histograma de nivell de gris per tot l'espai intercostal d'on són agafats els quadrats. Aplicant un transformació de nivells de grisos aconseguix representar un espai intercostal amb 8 nivells de gris igualment probables. Sobre aquesta nova representació calcula matrius de dependència espacial de nivell de gris, cada matriu de 8 per 8, com resultat d'aplicar un determinat angle i una determinada distància entre punts. Sobre cada una de les matrius aplicarà després cinc mesures de textura: una mesura de l'autocorrelació com mesura de la vastitud de la textura, una altra de la disemblança sovint anomenada moment d'inèrcia, mesura de l'extensió, de l'homogeneïtat i una altra mesura de disemblança.

En el domini freqüencial fa una mesura global de la imatge incluint la projecció de les costelles. Usa un sistema de reconeixement, dispositiu utilitzat per extreure les mesures de la textura visual en l'espai de freqüències. Fa servir la idea de que un regió pulmonar amb opacitats pneumoconiotiques generarà més altes freqüències que una zona que no en té.

Per triar quines característiques contenen la informació més útil pel classificador, es defineixen uns valors per a cada característica mesurada. Els resultats obtinguts després de la classificació es poden considerar semblants als obtinguts per una sèrie de radiòlegs sobre el mateix conjunt d'imatges. Aquest treball és una primera aproximació que vol demostrar la factibilitat del projecte. Les característiques mesurades són utilitzades per la classificació de la densitat de les petites opacitats, donant a cada imatge una categoria de 1 a 4.

Un altre treball sobre el tema és el de E.L. Hall i altres [38], en aquest la classificació es basa, igual que en la referència anterior, en els estàndards publicats per la ILO el 1971. I també en aquest cas intentarà donar la classificació per a la densitat de les petites opacitats.

L'anàlisi de la imatge es considerarà com un problema de reconeixement de patrons, per tant seguirà un esquema de preprocessament, extracció de característiques i classificació. En el pas del preprocessament aïlla en dues zones als dos pulmons i aplica una umbralització per poder separar la imatge en dues on hi hagi les zones amb costelles per una banda, i les zones inter-costals per altra, per així poder fer un millor estudi de les característiques.

Les mesures que són calculades per la classificació consisteixen en 62 característiques per cada una de les 6 zones en que s'han dividit els dos pulmons. 6 de les mesures són fetes a partir de l'histograma de cada zona pulmonar, la màxima intensitat, el nombre de punts, la mitja de les intensitats, la desviació estàndard, el biaix i la kurtosi. També hi ha 9 mesures de moments espacials que són l'àrea, la massa total d'intensitat, les coordenades centroïd i els 5 moments de primer i segon ordre. Hi ha 10 moments espacials més, però calculats sobre el gradient horitzontal de la imatge. Les darreres 6 últimes mesures són moments de l'histograma calculats sobre la imatge gradient. Aquestes 31 operacions són calculades sobre les dues subimatges que hem descrit abans.

A l'experiment que es presenta en aquest treball es fa una classificació de normal o anormal a cada zona, primer se'n seleccionen un conjunt sobre les que el sistema aprendrà, i després es fa la classificació sobre la resta. La classificació obtinguda per zones es pot comparar a la dels radiòlegs, on es presenten més problemes és a l'hora de fer la classificació global de la imatge en la que es parteix del valor donat a cada una de les zones i per tant apareixen tota una sèrie de casos en que la decissió és difícil.

El darrer treball que presentem que tracta també mesures de la textura és el de J.R. Jagoe i K.A. Paton [45]. Es planteja com un mètode bastant més senzill que els dos anteriors, i segons l'autor dóna millors resultats que en els altres. En aquest cas es fa un estudi del gradient de la imatge. Separa la imatge en quadres proporcionals en alçada a l'amplitud inter-costal, cada un d'ells el divideix en 9 subparts, i per cada una de les subparts calcula el percentatge en l'increment de la densitat en les 8 possibles direccions. Agafant el màxim d'entre totes les direccions obté el gradient d'aquesta subpart. Per a una part defineix la diversitat que és el nombre de gradients de subparts que són diferents en aquella part. El grau de pneumoconiosi vindrà donat per la mitja de la distribució de freqüència de la diversitat. Per tant si calculem la mitja sobre les diversitats de totes les parts de la imatge obté el grau de pneumoconiosi i per tant la classificació.

Mètodes que detecten les opacitats En aquesta secció hem recollit tots aquells treballs que fan la classificació a partir de la detecció directe de les opacitats. Un cop s'han detectat les opacitats es pot passar a avaluar-ne el nombre o la densitat. La detecció de les opacitats es pot fer de forma global sobre tota la imatge.

En aquest grup de treballs s'hi troben els més actuals que hem trobat sobre el tema. Aquests han estat realitzats per X. Chen, J. Toriwaki i J. Hasegawa de dues universitats del Japó [15] [16]. Però abans de veure aquests articles més recents, anem a veure els estudis de dos grups que tamé han tractat el problema basant-se en la detecció directe de les opacitats.

En el treball de C.C. Li [62] es recopilen els resultats dels algorismes desenvolupats pel mateix autor i el seus col·legues en [89], [63], junt amb el treball de D.B.C Shu [99]. Tots els algorismes són aplicats sobre les radiografies estàndards de la ILO.

A l'article [89] es dóna un algorisme de creixement de regions per a la detecció de les opacitats. L'algorisme troba primer els píxels que són sospitosos de pertanyer a una opacitat; aquests seran les llavors que aniran creixent. Per aplicar aquest algorisme es necessita un criteri de creixement, d'aquí que es defineixi una mesura sobre l'aïllament d'un píxel *ICI* (*Isolation contrast integral*). Una vegada acabat d'aplicar el creixement, es farà una classificació per triar si una regió obtinguda correspon a una opacitat real. Per fer aquesta última classificació farà servir els paràmetres ICI/A i P/A , aquest darrer és el quocient entre el perímetre i l'àrea per a donar una mesura de l'arrodoniment de la regió. Les opacitats triades com bones són després classificades en les classes p , q o r segons el valor de A . El mètode com a tal no dóna uns resultats prou bons i necessita ser millorat

per augmentar el rendiment.

Com ja hem dit abans a [62] es junten els resultats obtinguts en tots els treballs d'aquest grup. Els passos que es segueixen és l'aplicació de l'algorisme de creixement de regions, després mesures de l'arrodoniment i de la irregularitat, i finalment la detecció d'objectes lineals per detectar un cert tipus d'opacitats irregulars que apareixen com objectes lineals petits amb un contrast relativament baix. A l'algorisme de creixement de regions fa una sèrie de modificacions que permeten detectar opacitats que abans no es trobaven. Un objecte serà acceptat com a opacitat si el contrast en profunditat supera un cert umbral, a més a més proposa mesures d'arrodoniment com p^2/A (on p és el perímetre i A l'àrea), o bé mitjançant l'aproximació elíptica del contorn de l'objecte a través de la transformada de Fourier. Per detectar opacitats lineals fa servir la transformada de Hough.

Després de l'aplicació dels algorismes sobre el conjunt dels estàndards més sobre un conjunt de radiografies no afectades amb pneumoconiosi, es pot veure que en el cas de radiografies normals detecta molts fals positius. En les altres radiografies es pot comprovar que a mesura que el grau de pneumoconiosi augmenta els resultats són més dolents, en canvi són bastant bons pels casos de primers símptomes de la malaltia.

L'altre estudi és el presentat per Guo Lei et al. [61] on es presenta un nou enfoc del problema. Es realitza una suavització de la superfície que defineixen els nivells de gris de la imatge. La suavització es fa aproximant la imatge per una corba B-Spline cúbica, obtenint superfícies més suaus i amb derivades de $2n$ ordre contínues, a més a més s'elimina el soroll. Aleshores l'àrea que conté una opacitat serà entesa com una superfície local convexa, és a dir amb les seves dues principals curvatures més petites que zero. Modelitzat d'aquesta manera el problema es redueix a calcular les principals curvatures per detectar les superfícies convexes i obtenir així àrees candidates a ser opacitats.

L'aproximació per B-Splines cúbiques s'obté mitjançant la convolució de la imatge per una sèrie de matrius, amb els resultats de les convolucions calcularem les principals curvatures. Per establir si els objectes detectats són realment opacitats s'usaran mesures d'arrodoniment, de concavitat, de densitat i valors mínims de les corbatures principals. Tenint tot això en compte es classificaran les imatges en 3 categories diferents: 0/0, 1/1(q/q), 2/2(q/q). L'algorisme és aplicat sobre molt poques imatges, el que fa que no es puguin extreure conclusions prou fiables, el que es pot dir és que els resultats milloren quan la densitat és gran, en canvi són bastant dolents en el cas de la radiografia 0/0. De totes maneres s'ha de destacar com un nou enfoc geomètric prou interessant.

Finalment queden per veure els treballs del grup de X. Chen [15] [16], com ja hem dit sembla que és l'únic grup que està treballant actualment en aquest tema concret. En ambdós treballs es presenten algorismes similars, però més perfeccionat en el segon. Es defineix el problema com la classificació de talla i de densitat. I es presenta un algorisme dividit en els passos següents:

Pas 1 Millora amb filtres de tipus p, q o r (depèn del tipus d'opacitat que es vol detectar).

Pas 2 Extracció dels màxims de la imatge com centres d'opacitats.

Pas 3 Càlcul del valor mig de la mida de les opacitats.

Pas 4 Si aquest valor és més gran que un cert umbral es classifica com p, q o r en funció del filtre que s'ha aplicat a (1).

Pas 5 Sinó tornem a (1) i fem el mateix pel següent filtre.

Els filtres anomenats p, q i r corresponen a filtres diferencials de $2n$. ordre. La diferència entre ells està en el radi del filtre, radis que s'adapten a les mides dels tres possibles tipus d'opacitats p,q i r. L'objectiu d'aquests filtres és el de fer que les opacitats que s'hi adapten en mida siguin suavitzades de tal manera que presentin un únic màxim en el seu centre. Ja que en general no sempre una opacitat es correspon amb un màxim. Els resultats d'aquest algorisme són bastant bons comparats amb tots els altres treballs.

Després de veure tots aquests treballs realitzats sobre el tema podem interpretar que encara no s'ha trobat una solució prou bona per a obtenir una classificació fiable. Tots els estudis fets s'han basat simplement en la detecció de les anomalies del parènquima, sense tenir en compte tots els altres símptomes de pneumoconiosi que potser podrien ajudar a una més completa classificació.

2.3.3 Resultats

En aquest apartat donem el conjunt de resultats de calcular les corbes granulomètriques sobre un conjunt d'imatges dels estàndards. Les corbes es calcularan a partir de diferents obertures algebraiques, que són les que s'han introduït a l'apartat 2.2.4, és a dir, s'aplicaran les obertures morfològiques amb elements estructurants rodons, les 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.

TREBALL	Caract. analitzada	Classificació donada	globals o zones	Tècniques usades
R.P.Kruger 1974 ILO-1971	Opacitats petites	0 1 2 3 (densitat)	zones intercostals i globals (espai freqüencial)	Mesures: autocorrelació vastitud moment d'inèrcia extensió homogeneïtat A l'espai freqüencial
E.L.Hall 1975 ILO-1971	Opacitats petites	0 1 2 3 (densitat)	zones intercostals i zones costals	Mesures: 6 sobre histograma 9 moments espacials 10 sobre imatge grad horitzontal 6 sobre histograma imatge gradient
J.R.Jagoe 1976 ILO-1971	Opacitats petites	0 1 2 3 (densitat)	zones intercostals	Estudi del gradient (mesura de diversitat)
C.C.Li 1980,82 ILO-1980	Opacitats petites rodones i lineals	0 1 2 3 p q r s t u (densitat i tamany)	global	Algorisme de creixement de regions Transformada de Hough Mesures d'arrodoniment
Guo Lei 1986 ILO-1980	Opacitats petites rodones	0 1 2 3 p q r (densitat i tamany)	global	Suavitització B-Splines cúbiques Càlcul de corbatures Detecció convexitats
X. Chen 1988,90 ILO-1980	Opacitats petites rodones	0 1 2 3 p q r	global	Filtratge Adaptatiu Extracció màxims Mesura de talla mitja Càlcul densitat

TAULA 2.4: Sumari dels diferents treballs realitzats

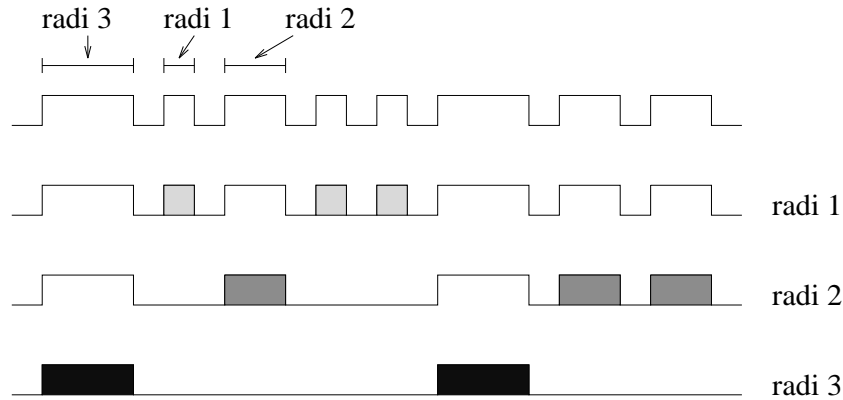


FIGURA 2.11: Seqüència de perfils d'opacitats, i com van desapareixent amb les obertures morfològiques.

Granulometria amb obertures morfològiques Les obertures morfològiques sobre les radiografies de tòrax haurien d'anar eliminant el volum que compona una opacitat en funció del radi de l'opacitat i del radi de la transformació.

El comportament hauria de ser com el que s'expressa a la figura 2.11. A mesura que anem augmentant el radi de l'obertura van desapareixent les opacitats de radi corresponent. D'aquesta manera es va afegint el volum d'imatge que desapareix en cada cas. Aleshores la corba granulomètrica hauria de presentar per a cada radi el volum que correspon al conjunt d'opacitats de radi corresponent.

A les figures 2.12, 2.13 i 2.14 es mostren alguns dels estats intermitjos que es donen quan es calcula la granulometria sobre algunes imatges dels estàndards que s'han triat com a mostra. Al costat de cada una de les imatges intermitges es dona la diferència entre la imatge original i la transformada. El volum d'aquesta imatge serà el que s'acumularà a la funció de distribució que estem definint.

Totes les imatges de la mostra tenen les opacitats d'una sola mida, de manera que hauríem de tenir una corba amb un màxim al radi de les opacitats de la imatge. A la figura 2.15 es mostren les corbes corresponents a un conjunt d'imatges especificades a l'apèndix E de [110]. Per a poder interpretar millor les corbes granulomètriques de les gràfiques hem marcat els diferents intervals, a partir dels quals es fa la classificació en els estàndards de la ILO. A partir de les corbes granulomètriques que hem indicat podem veure que en pràcticament tots els casos es faria una bona discriminació. Podríem assignar la classe corresponent a l'interval on es troba el pic de la corba.

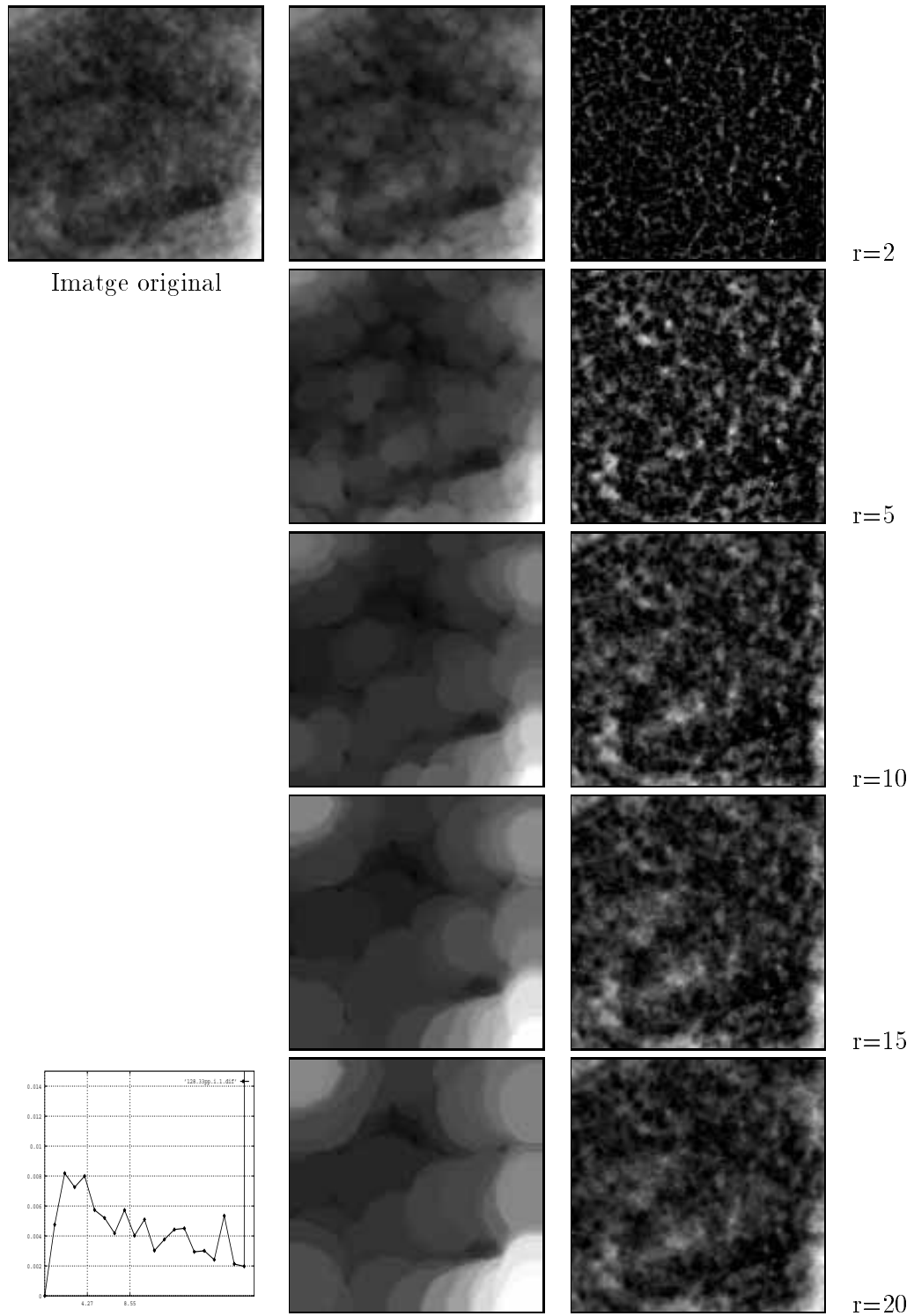


FIGURA 2.12: Imatge 3/3 p/p. Seqüència d'imatges intermitges en el procés de càlcul de la granulometria amb obertures morfològiques. Normalització η_2

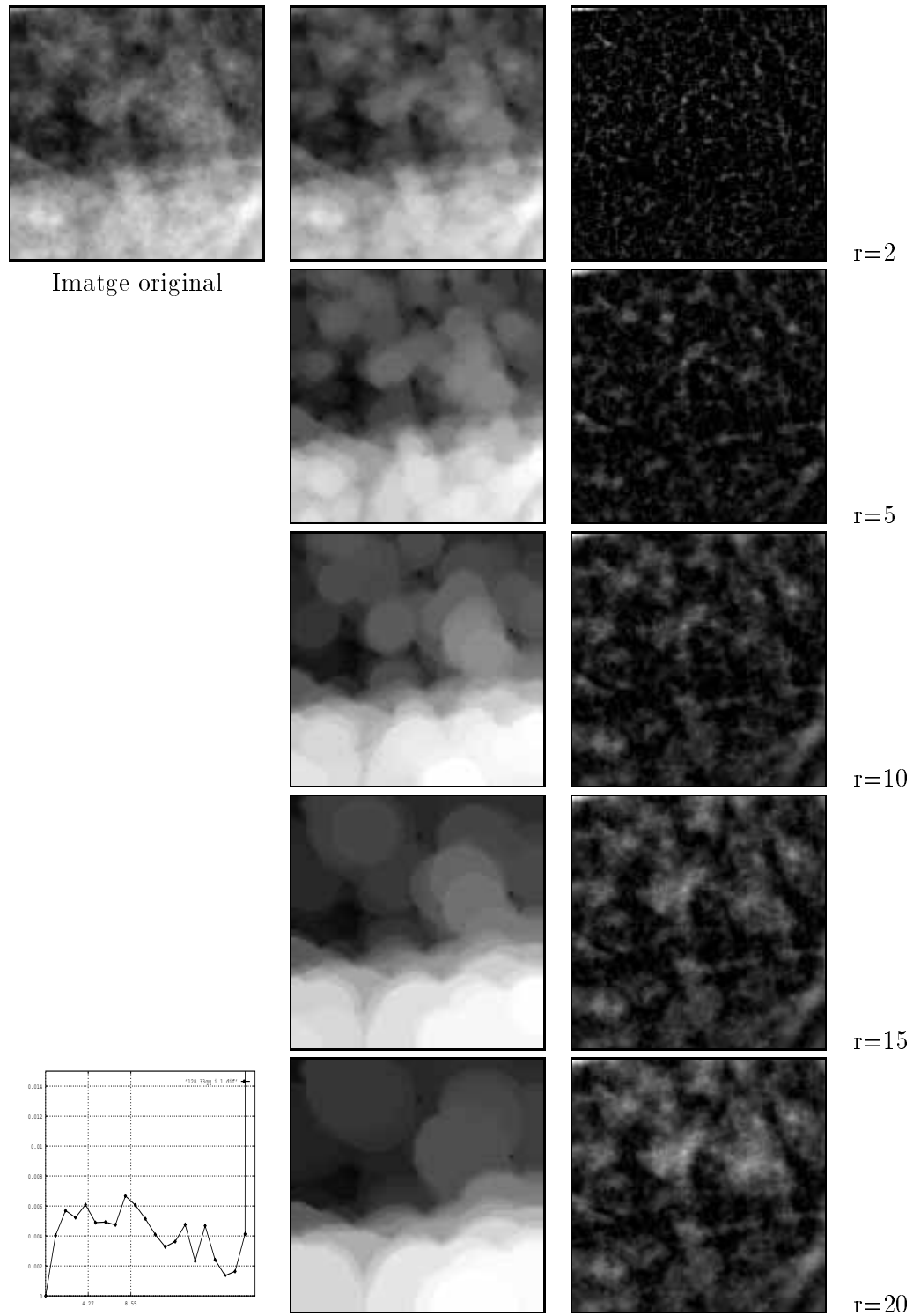


FIGURA 2.13: Imatge 3/3 q/q. Seqüència d'imatges intermitges en el procés de càlcul de la granulometria amb obertures morfològiques. Normalització η_2

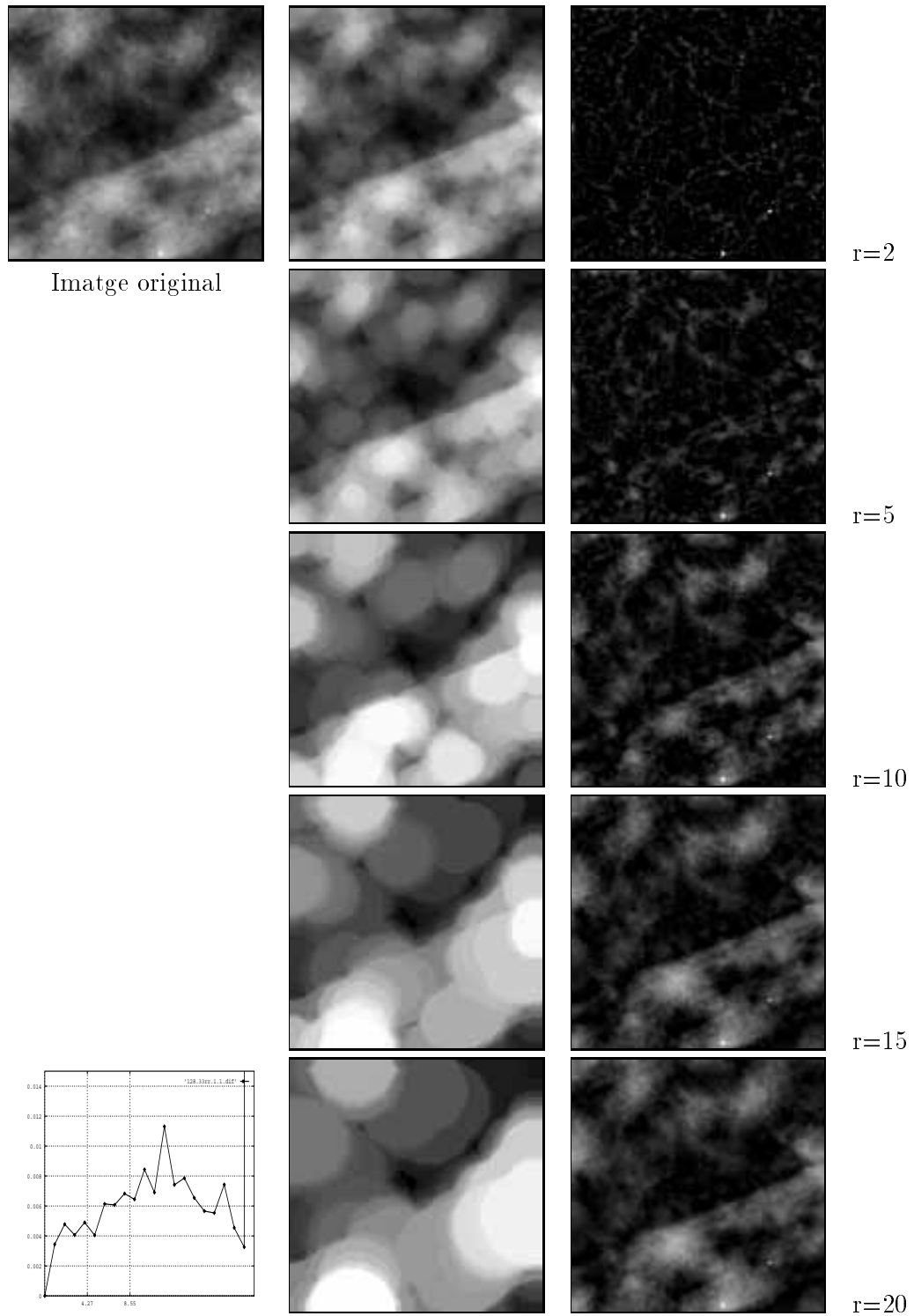


FIGURA 2.14: Imatge 3/3 r/r. Seqüència d'imatges intermitges en el procés de càlcul de la granulometria amb obertures morfològiques. Normalització η_2

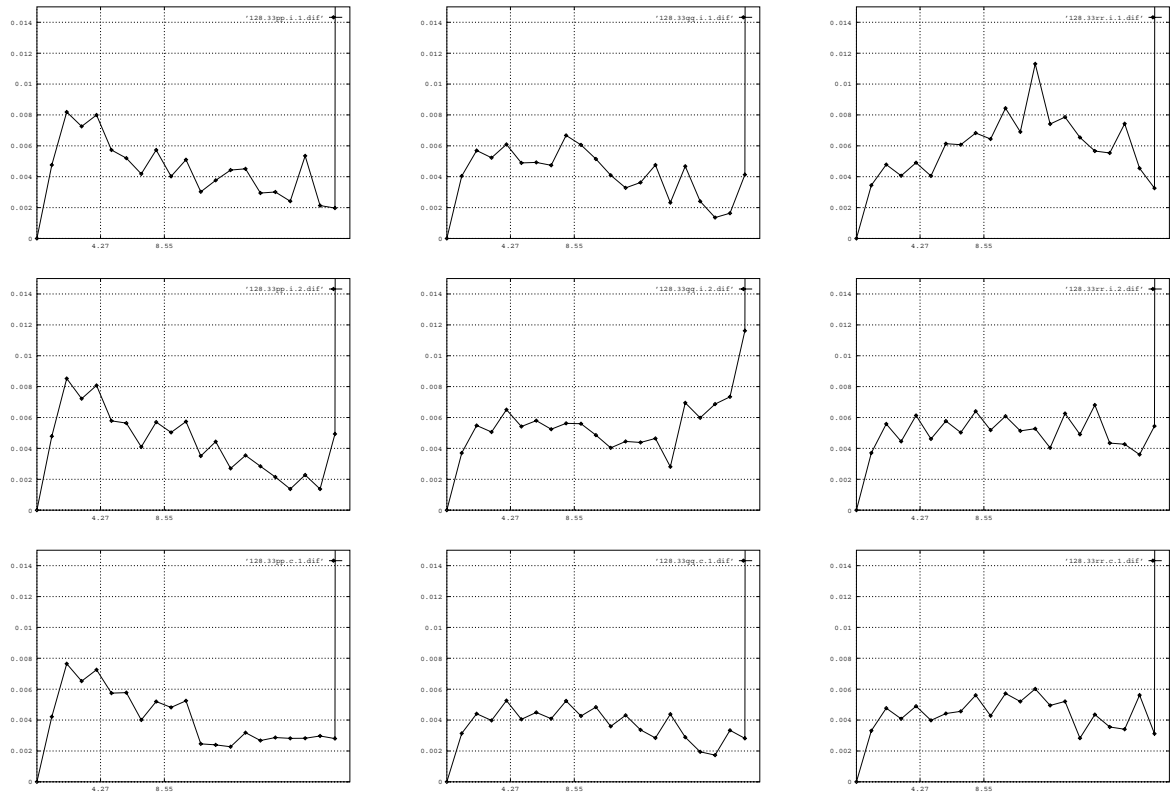


FIGURA 2.15: Corbes granulomètriques resultants d'aplicar una granulometria amb obertures morfològiques. D'esquerra a dreta els resultats corresponen a imatges classificades com 3/3 p/p, 3/3 q/q, 3/3 r/r (128x128). Normalització η_2 .

Ara bé si tenim en compte els resultats intermitjos en la construcció de la granulometria podem deduir que les opacitats no són com les que havíem dibuixat en el perfil de la figura 2.11, ja que no desapareixen totalment amb el seu radi corresponent. Sinó que van desapareixent petits troços de la imatge que es van fent cada cop més grans.

D'aquí podem dir que les opacitats són com petits turons, aleshores no van desapareixent així com podríem pensar inicialment. Les obertures morfològiques no eliminen en cada pas tota l'opacitat sinó que aquesta es va rompent a poc a poc.

Degut a això podem veure que per a les opacitats petites apareix un pic clar a la corba a l'interval corresponent. Però per a casos d'opacitats més grans els pics de la corba no són tant clars, ja que part del volum de les opacitats s'han perdut amb les obertures de radis inferiors.

A partir d'aquests resultats podem dir que l'obertura morfològica es comporta bé per a les imatges amb densitats altes d'opacitats. En els altres casos en els que no hi ha tantes opacitats, i tenint en compte que a totes les imatges hi ha molts elements com les costelles o venes que també augmenten els volums eliminats amb les obertures, resulta més difícil que la mesura sigui fàcilment interpretable.

Si tenim en compte els resultats intermitjos de les figures 2.12, 2.13 i 2.14, podem veure que l'obertura morfològica no és prou restrictiva amb la forma de les opacitats, es pot veure que a les imatges diferència no es distingeixen partícules rodones. Sinó que simplement s'elimina tot allò on l'element estructurant no hi cap. Per a intentar solucionar això canviarem l'obertura algebraica.

Granulometria amb obertura geodèsica La nova obertura algebraica que hem escollit per a solucionar el problema que hem explicat abans és l'erosió amb element estructurant rodó seguida d'una reconstrucció. D'aquesta manera es pot solucionar el problema ja que les opacitats que no siguin completament rodones no seran eliminades per l'erosió fins que l'element estructurant no càpiga dins de l'opacitat.

És a dir, l'opacitat conservarà la seva forma i per tant una major part del seu volum fins que sigui completament eliminada per l'erosió. A la figura 2.16 podem veure l'efecte d'aquesta obertura geodèsica sobre opacitats que no són rodones i la comparació amb l'obertura morfològica.

Els resultats intermitjos d'aplicar aquesta obertura es donen a les figures 2.17, 2.18 i 2.19. A la figura 2.20 donem les corbes granulomètriques pel conjunt d'exemples que ja

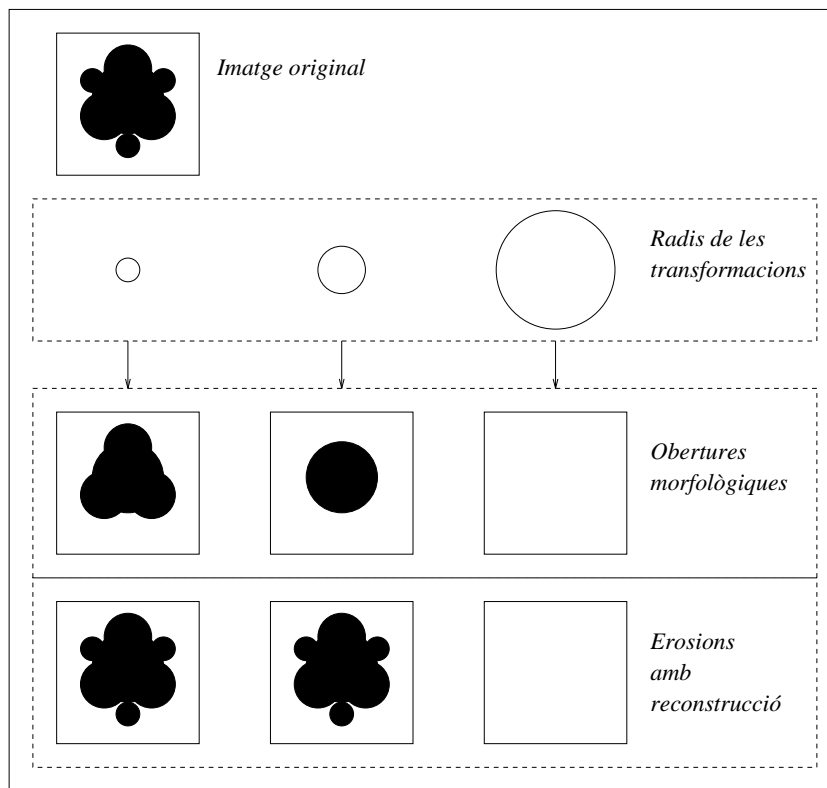


FIGURA 2.16: Exemple de la reconstrucció d'opacitats no rodones.

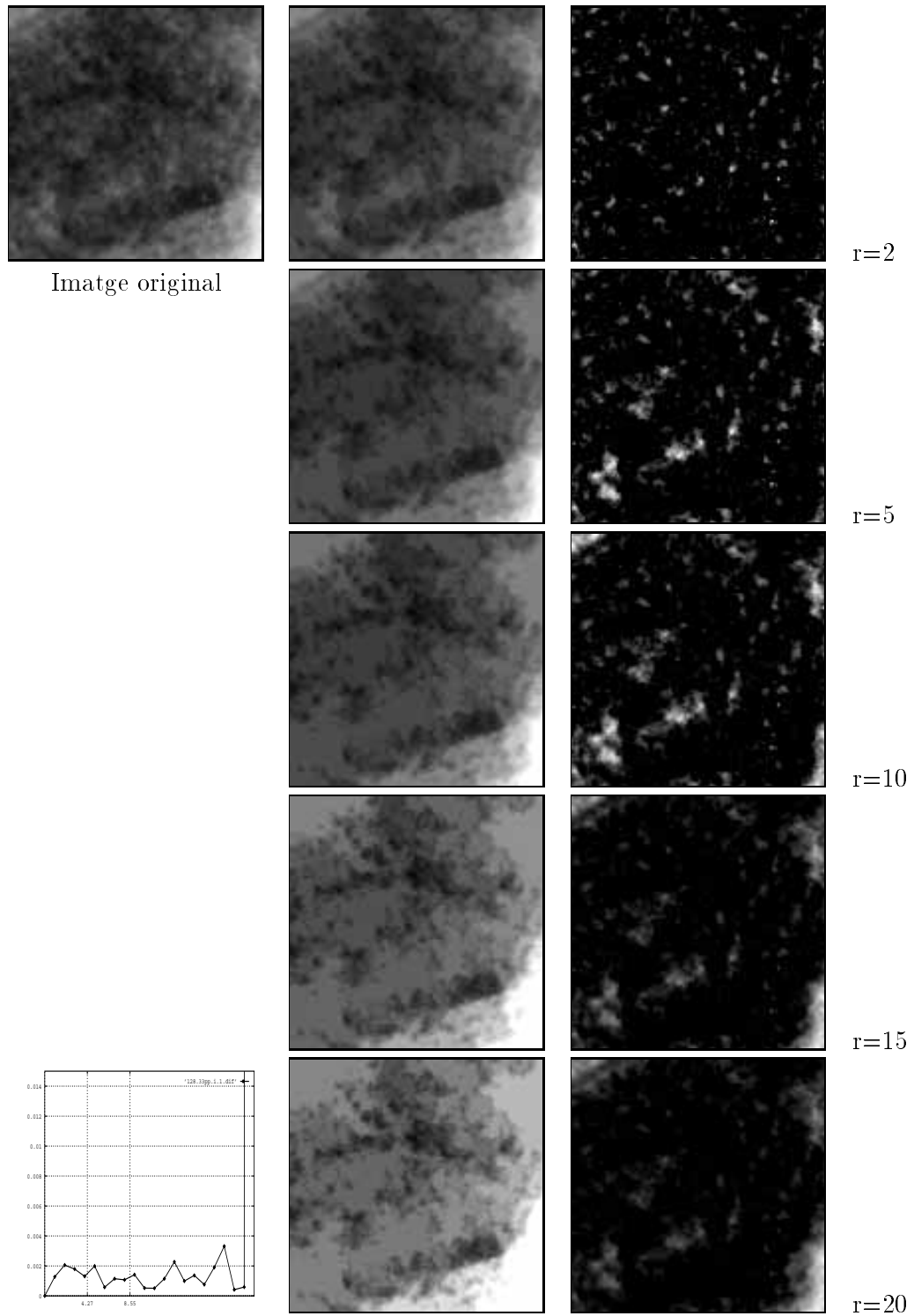


FIGURA 2.17: Imatge 3/3 p/p. Seqüència d'imatges intermitges en el procés de càlcul de la granulometria amb erosió més reconstrucció. Normalització η_2

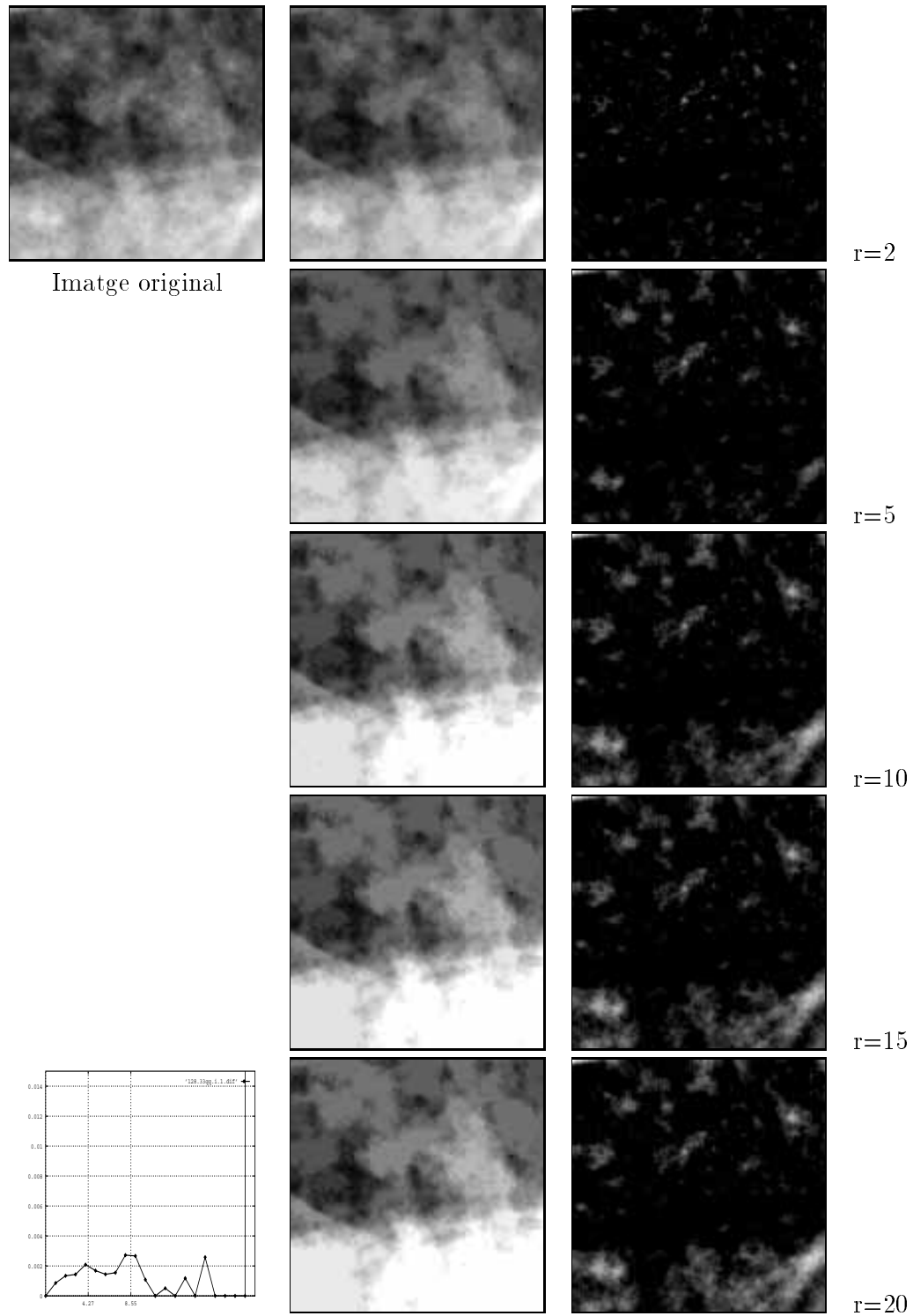


FIGURA 2.18: Imatge 3/3 q/q. Seqüència d'imatges intermitges en el procés de càlcul de la granulometria amb erosió més reconstrucció. Normalització η_2

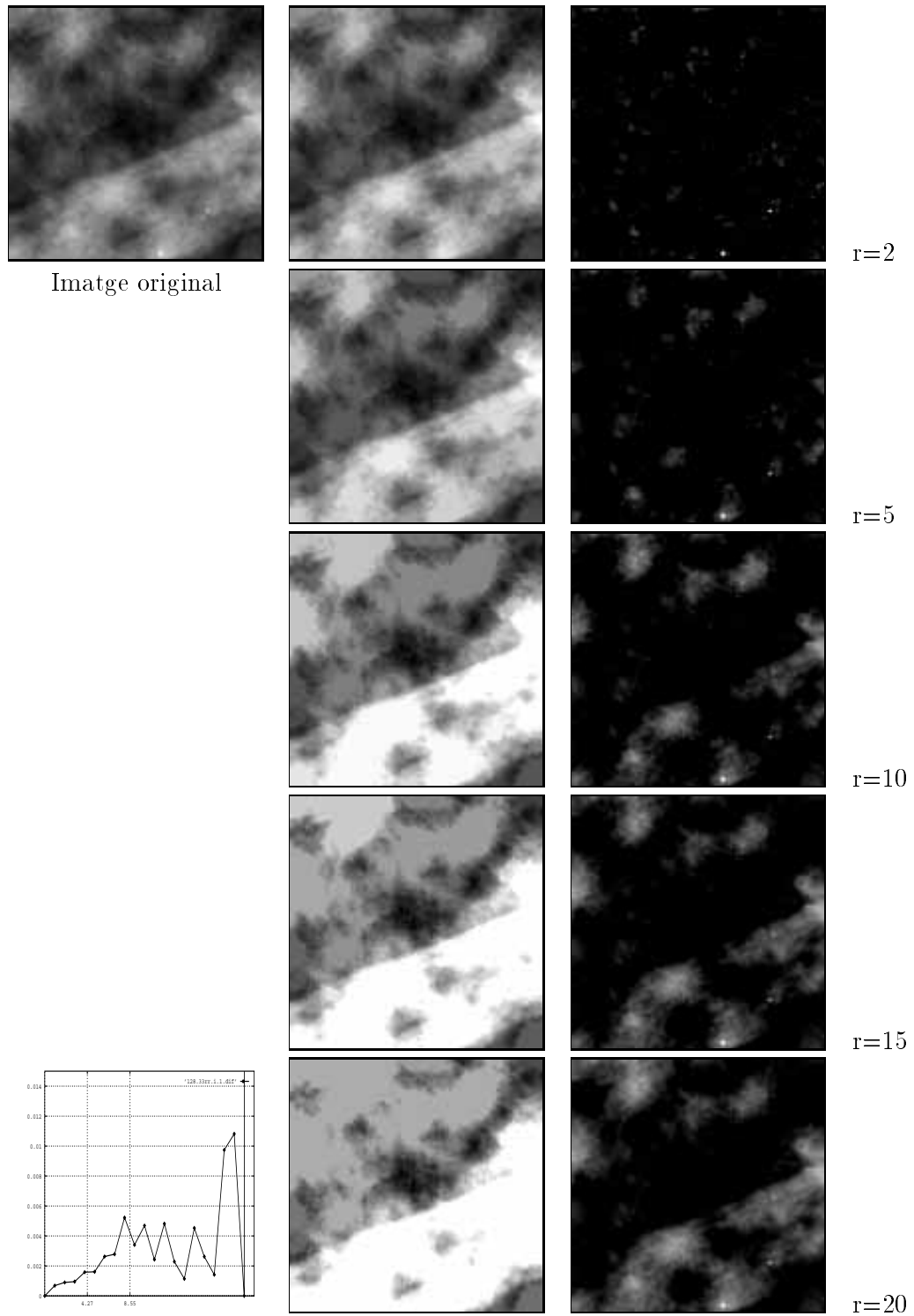


FIGURA 2.19: Imatge 3/3 r/r . Seqüència d'imatges intermitges en el procés de càlcul de la granulometria amb erosió més reconstrucció. Normalització η_2

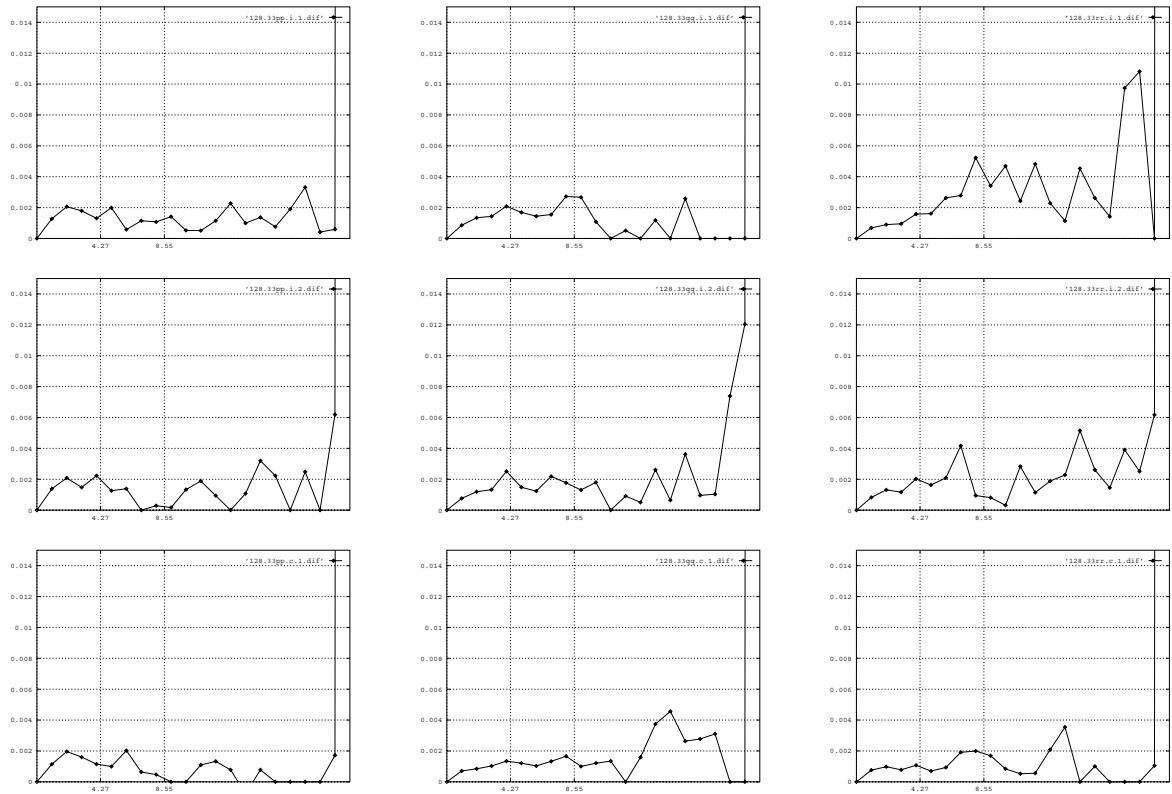


FIGURA 2.20: Corbes granulomètriques resultants d'aplicar una granulometria amb erosió seguida de reconstrucció. D'esquerra a dreta els resultats corresponen a imatges classificades com 3/3 p/p, 3/3 q/q, 3/3 r/r (128x128). Normalització η_2 .

hem especificat a l'apartat anterior.

Com en el cas de l'obertura morfològica els millors resultats s'han obtingut pels casos d'imatges amb densitat alta d'opacitats. Tot i això s'han millorat els resultats per a les opacitats de tipus r però s'han empiorat els resultats per a les opacitats de tipus p.

En els resultats intermitjos de la figura 2.19 es pot veure com s'ha millorat l'extracció de les opacitats. El volum que s'elimina correspon al volum de l'opacitat i no tant a altres elements que provocaven soroll amb l'obertura anterior.

En canvi per a la imatge amb opacitats petites el resultat ha estat empiorat. Tot i que l'obertura continua sent més restrictiva amb la forma, apareix un nou factor que és el fet de que les opacitats estan unides entre elles.

Degut a això les reconstruccions van inundant zones veïnes i les opacitats van desaparèixer molt més lentament fet que provoca que la corba granulomètrica sigui més plana.

És a dir, el fet de tenir una transformació més restrictiva fa que les transformacions tardin més en perdre el seu volum. Evidentment aquest efecte afavoreix els casos en que les opacitats són grans, i en canvi empitjora els casos en que les opacitats són petites i estan unides entre elles.

Granulometria amb suprem d'obertures Si tenim en compte els resultats que hem vist fins ara podem dir que les obertures morfològiques amb elements estructurants rodons no són prou restrictives amb la forma de les opacitats. Per altra banda la reconstrucció com a obertura és massa estricta i no permet que les opacitats estiguin unides, en canvi els elements estructurants rodons de l'erosió són un bon marcador.

A la vista d'aquests resultats es pot concloure que l'obertura morfològica és una bona transformació però cal que s'elegeixin uns elements estructurants que siguin més restrictius en quan a la forma. Per això hem elegit el suprem d'obertures morfològiques amb elements lineals com a nova obertura.

Aquesta obertura és més restrictiva que l'obertura simple amb element rodó, ja que només elimina aquells pics de la imatge en els que no hi capiguen cap dels elements lineals donats. A més a més, té en compte la forma rodona de les opacitats i no tenim el problema de que les opacitats estiguin unides com es donava amb la reconstrucció.

Com en els casos anteriors els millors resultats s'han obtingut per a imatges amb densitat alta d'opacitats.

Amb aquesta nova obertura s'aconsegueix eliminar els problemes de l'anterior cas. El

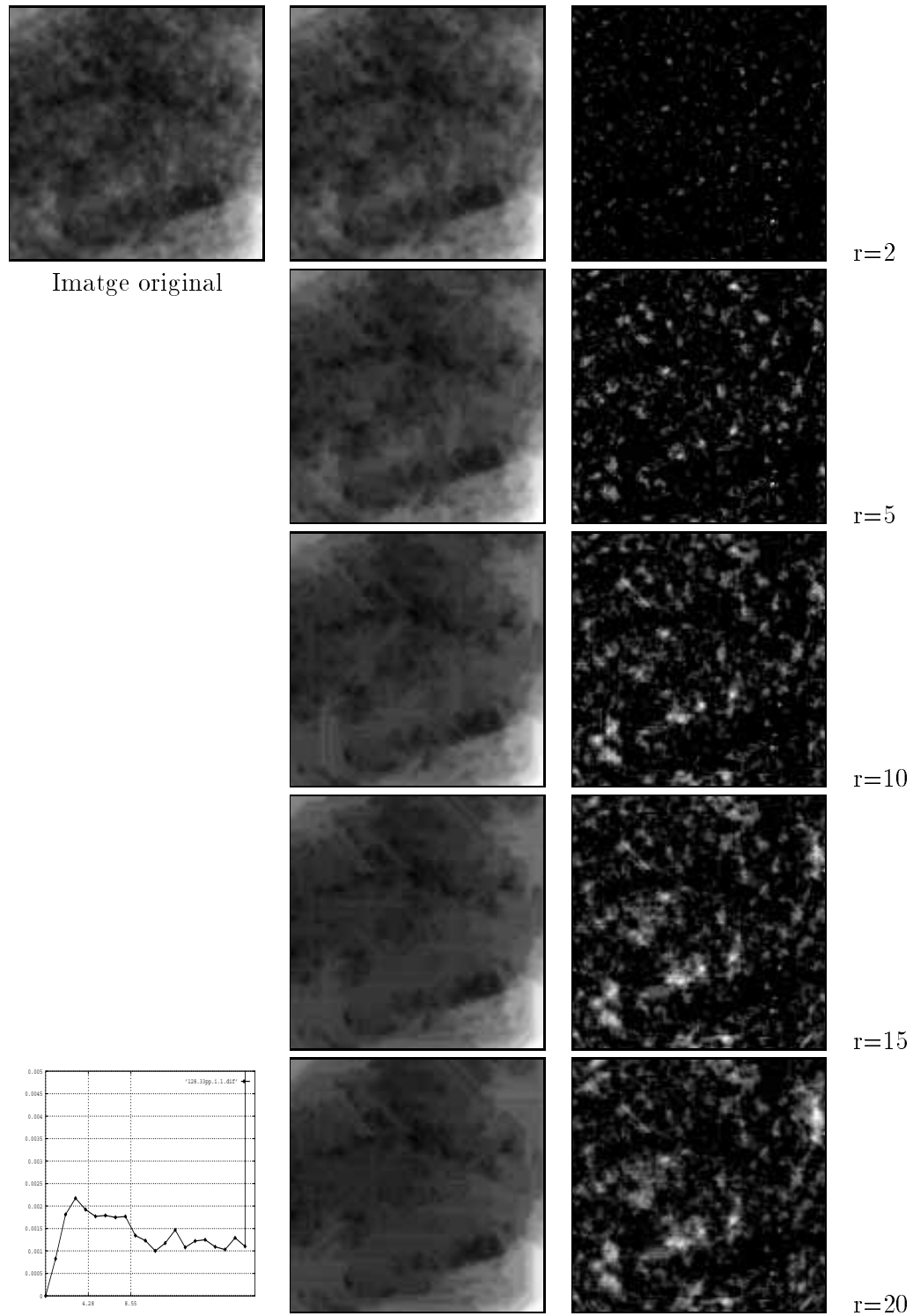


FIGURA 2.21: Imatge 3/3 p/p. Seqüència d'imatges intermitges en el procés de càlcul de la granulometria amb el suprem d'obertures amb elements lineals. Normalització η_2

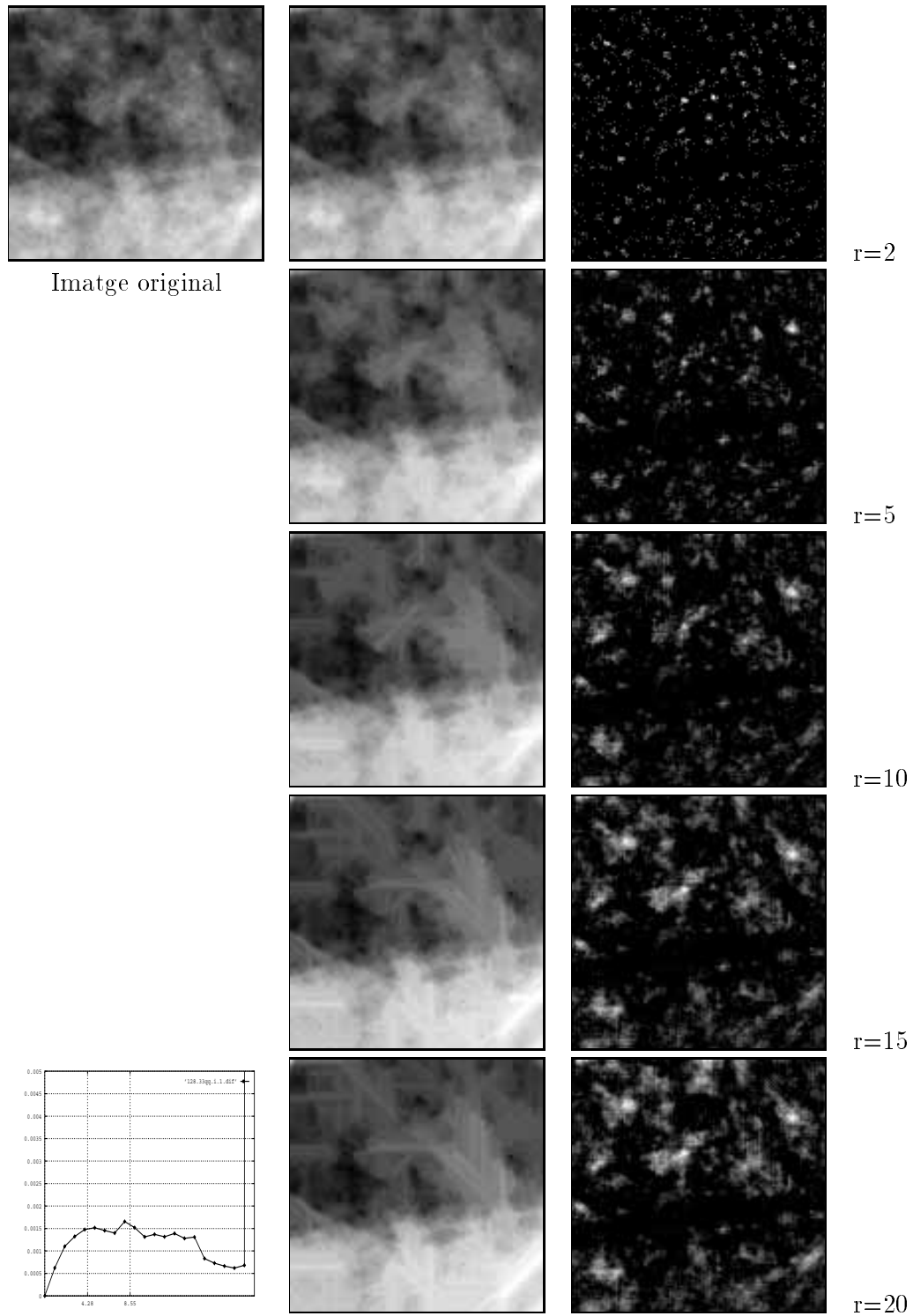


FIGURA 2.22: Imatge 3/3 q/q. Seqüència d'imatges intermitges en el procés de càlcul de la granulometria amb el suprem d'obertures amb elements lineals. Normalització η_2

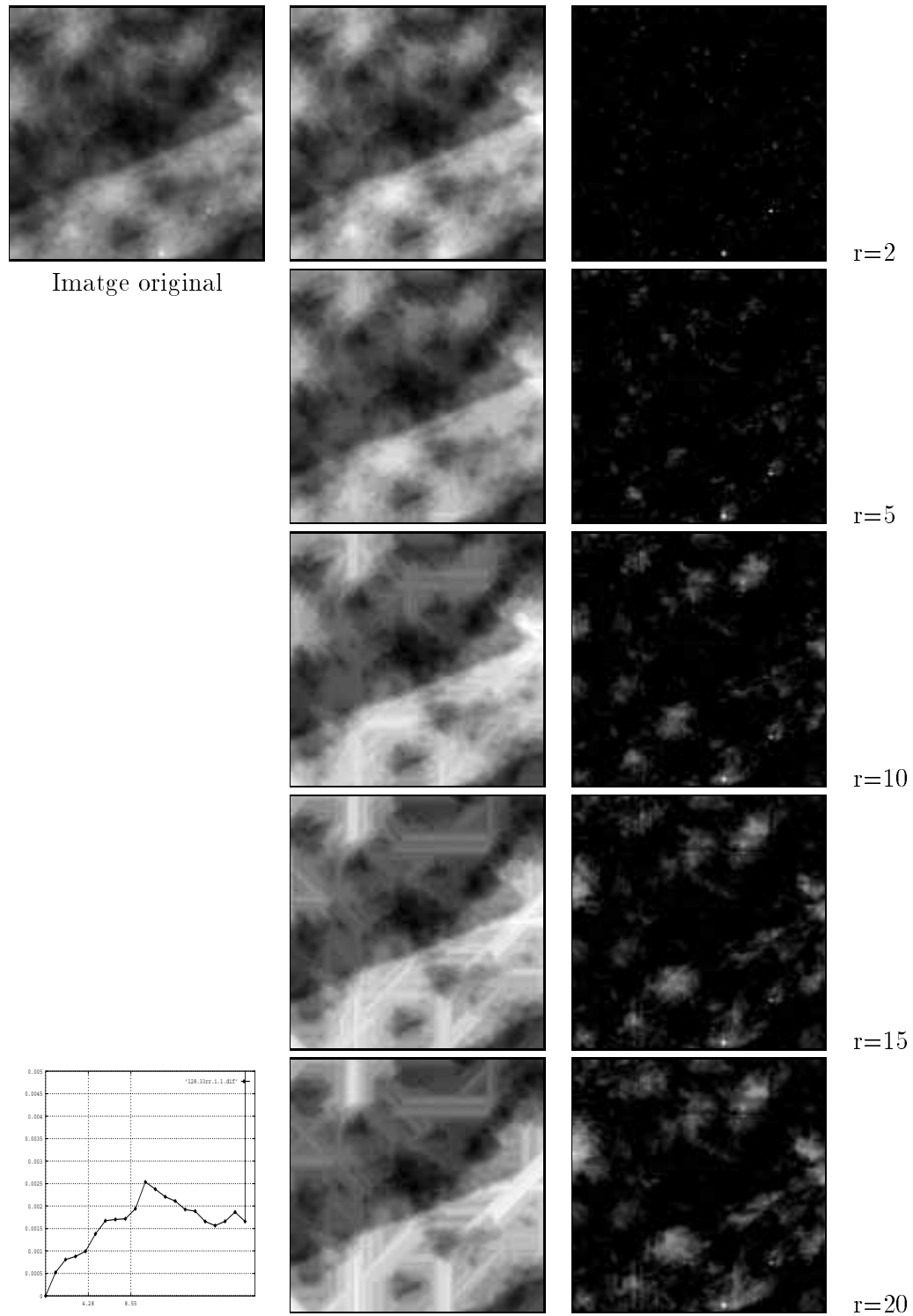


FIGURA 2.23: Imatge 3/3 r/r. Seqüència d'imatges intermitges en el procés de càlcul de la granulometria amb el suprem d'obertures amb elements lineals. Normalització η_2

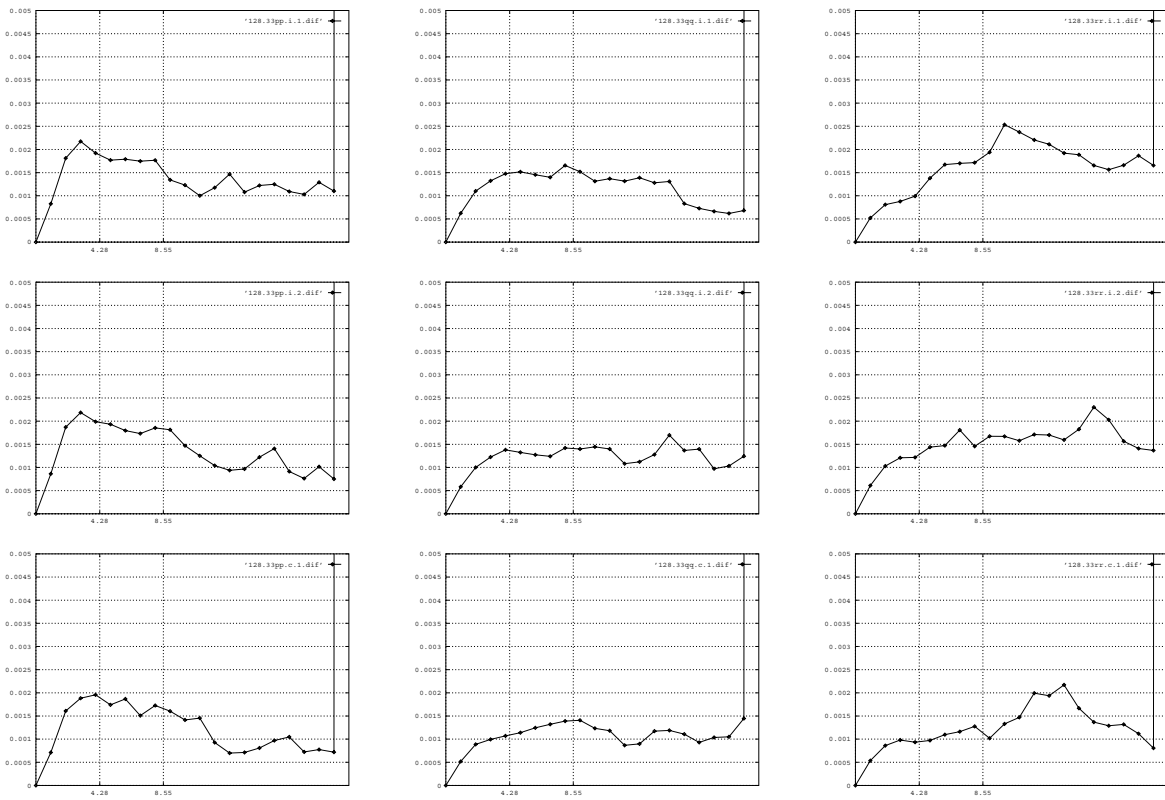


FIGURA 2.24: Imatges 3/3 p/p, 3/3 q/q, 3/3 r/r (128x128). Granulometria amb Suprem d'obertures amb elements lineals. Normalització η_2

suprem d'obertures lineals permet eliminar de la corba tots aquells elements com venes o costelles que suposaven un important soroll. Els elements lineals no deixen que aquests volums siguin considerats i per tant no afecten a la forma de la corba granulomètrica. S'a dir, totes aquelles parts de la imatge en les que hi cap un dels elements lineals no seran eliminades de la imatge i el seu volum no serà comptabilitzat com volum d'una opacitat.

A les imatges intermitges de la figura 2.21 podem veure com des de la primera transformació desapareix un volum considerable. Per altra banda a la figura 2.23 es pot apreciar com el volum es conserva quasi fins a la transformació de radi 15. Les corbes granulomètriques resultants d'aplicar aquesta obertura es donen a la figura 2.24

Si a tot això hi afegim el fet d'una disminució considerable de temps de càlcul de les transformacions respecte de les transformacions amb els elements rodons, tot ens porta a pensar que és la millor família pel càlcul de la granulometria en volum.

2.4 Valoració d'aquesta aproximació

Primer de tot, hem de dir que per al problema concret de l'automatització d'imatges radiogràfiques afectades de pneumoconiosi s'han aconseguit resultats vàlids des del punt de vista que la classificació, ja que aquesta sembla ser possible en la majoria dels casos (el conjunt de totes les mesures fetes sobre un ampli conjunt d'imatges es pot trobar a l'apèndix G de [110]). Des d'un punt de vista més teòric podem concloure:

- La granulometria amb obertures algebraiques ha demostrat ser una bona eina per a mesurar la granularitat com a característica textural.
- L'elecció de l'obertura algebraica és una decisió important a l'hora de calcular aquesta mesura. La validesa del resultat depèn en gran mesura del tipus d'imatge i del tipus d'obertura.

Per tant, podem concloure que la representació de textures basada en la mesura de característiques texturals com la granularitat, no sembla ser el millor pas cap a la generalització d'una representació de textures. Ja que com hem vist el mètode depèn molt de la imatge.

Aleshores podem deduir que per a la resolució o automatització de problemes concrets, la mesura de certes característiques texturals pot ser una solució molt interessant, ja que el fet de mesurar característiques molt concretes fa que s'estalviïn processos innecessaris, i es puguin donar temps de resposta petits. Una aproximació en aquest sentit ha estat feta

a [3], on es soluciona el problema a partir de de les mesures de dues propietats texturals concretes i que són suficients pel tipus d'imatges que s'estan tractant.

Ara bé, si ens plantegem la representació de textures de forma general amb aquestes mesures, abans haurem de contestar les següents preguntes:

- Quin conjunt de mesures és suficient fer per a que la representació sigui general?, una resposta podria ser la donada per R. Rao et al. a [81], on les mesures que s'han de fer venen donades pels eixos de l'espai de representació que es defineix, que són la granularitat, la repetitivitat, l'aleatorietat, el contrast o la direccionalitat.
- La mesura de característiques normalment es basa en característiques d'ordre superior, de manera que sempre porten de forma inherent una certa ambigüitat en el seu significat. Això es pot veure en els experiments de [81, 80], en els que una mateixa imatge és descrita de forma diferent per diferents persones. Com podem eliminar aquesta ambigüitat inherent a aquestes mesures?
- És evident que amb una mesura sola no en tenim prou, per tant la qüestió és, com les combinem entre elles per a una representació general?

Per tant, la representació basada en mesura de característiques sembla una via vàlida si es contesten les anteriors preguntes.

CAPÍTOL 3

REPRESENTACIÓ MULTIDIMENSIONAL BASADA EN DETECTORS DE CARACTERÍSTIQUES

En aquest capítol donem una petita introducció als models multicanal definits per a representar el problema de percepció de textures. Aquests models es basen en evidències biològiques sobre l'existència de detectors de característiques en el sistema visual primari, i han estat la base per desenvolupar models computacionals de percepció.

Ens centrarem en un d'aquests models computacionals que és el definit per J. Malik i P. Perona. Explicarem detalladament cada un dels passos del model i les justificacions d'aquests donades pels autors. A partir d'aquest model computacional de percepció definirem una representació de textures, basada en les respostes dels canals del model abans de passar pel nivell de combinació. Una funció d'aquestes respostes és la que s'agafarà com a representació i serà l'objectiu d'estudi de tot aquest treball.

Finalment, plantejarem els problemes que suposa la representació definida. Aquests problemes són els que guiaran els objectius dels següents capítols.

3.1 Models multicanal de percepció de textures

Els models multicanals són aquells que es basen en la representació d'una textura a partir d'un conjunt d'imatges, que són el resultat d'aplicar un conjunt considerable de filtres amb paràmetres específics.

L'ús d'aquests filtres ha estat recolzat per resultats psicofísics i fisiològics que evidencien l'existència d'aquests analitzadors al SVH, que es sintonitzen amb determinades freqüències espacials i determinades orientacions. Tenint en compte aquest enfocament J.R. Bergen i M.S. Landy [7] veuen la capacitat de discriminació preatentiva de textures com una conseqüència de la mateixa estructura del sistema visual primari localitzat a les àrees primàries

del córtex visual.

A partir d'aquestes idees s'han definit molts models computacionals diferents amb determinades capacitats per dur a terme la discriminació pretentiva de textures ([7, 35, 18, 6, 5, 107, 84, 28, 9, 65, 70, 19, 36]). El que és comú en tots aquests models és el fet de que comparteixen d'una representació genèrica de la imatge que és el resultat del filtratge d'aquesta per un conjunt important de filtres que es sintonitzen amb determinades característiques de les imatges i que normalment es relacionen amb freqüencials espacials i amb l'orientació. A la figura 3.1 podem veure l'esquema general del primer pas d'aquests nivells i a la figura 3.2 veiem un exemple d'una imatge amb una textura i els conjunt de canals resultants d'aplicar els filtres de la família de les derivades de la gaussiana [125].

En els treballs de M. Clark i A.C. Bovik [9, 19] s'insisteix en la consistència d'aquests models multicanal amb la teoria del textó de Julesz [52], i amb la línia de treballs que es basen en la detecció de barres i blobs com H. Voorhees i T. Poggio a [120]. Aquest fet és el que ens ha portat a agafar el model d'imatges que definim en el capítol 5.

En tots els models definits, el filtratge és seguit de mecanismes no lineals sense els que seria impossible la discriminació de textures, després es calculen promitjos espacials sobre els canals per a trobar després el gradient textural. Vegem alguns dels nivells que poden apareixer en els models basats en aquesta aproximació:

Filtratge lineal és el nivell que ja hem especificat. En aquest nivell la imatge és filtrada amb un conjunt important de filtres que es sintonitzen en determinades característiques, normalment relacionades amb l'orientació i la freqüència espacial.

Rectificació o operacions puntuals no lineals. Aquests mecanismes representen les úniques no linealitats incorporades per alguns models com el de D. Sagi [83, 84]. Poden ser mesures d'energia, rectificacions de mitja ona, etc.

Canals complexos. En alguns models [35, 36] s'han introduït canals complexos que poden substituir el filtratge lineal o apareixer conjuntament amb ell. Realitzen el filtratge lineal en dos passos entre els que s'hi posa una operació no lineal.

Xarxa d'inhibició és una no linealitat que ve donada amb una xarxa d'interaccions entre els diferents canals. Al següent apartat explicarem la xarxa d'inhibició del model J. Malik i P. Perona [65]. Aquesta no linealitat no apareix en tots els models. En un treball recent de N. Graham i A. Sutter [36] s'argumenta la consistència amb resultats psicofísics del comportament d'aquest nivell.

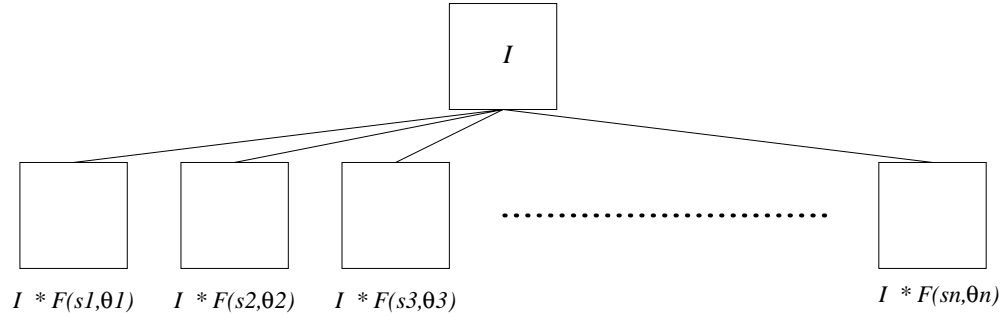


FIGURA 3.1: Esquema general de la representació d'una imatge a partir del resultat del filtratge d'aquesta amb filtres que difereixen en l'orientació i la freqüència espacial.

Càlcul del gradient textural . Aquest pas apareix en tots els models dirigits a la segmentació de textures. El gradient es calcula després d'un promig espacial sobre les imatges dels diferents canals. En alguns casos es calcula aplicant un filtre de simetria senar com el gradient de la funció gaussiana de suavització, $\nabla^2 G$, o bé amb un operador de laplaciana, ΔG .

Nivell de combinació en el que totes les respostes són combinades amb algun operador com la suma o el màxim. En aquest nivell es perd tota la informació que fa referència a la característica específica de cada canal.

Nivell de decisió en el que a partir del resultat dels nivells anteriors es dona una resposta sobre el resultat de la tasca visual.

D'entre tots aquests models hem seleccionat el model computacional de J. Malik i P. Perona, a partir del que volem definir una representació. La selecció d'aquest model ha estat motivada per ser un model que és el resultat d'intentar integrar les teories més conegudes sobre la discriminació preatentiva de textures en el camp de la visió per computador.

3.2 El model computacional de J. Malik i P. Perona

El model dissenyat per J. Malik i P. Perona [64, 65] és un model de segmentació preatentiva de textures que intenta agrupar els resultats dels treballs de la teoria del textó de Julesz [48], de J. Beck [5] i de A. Treisman [107]. Per a la construcció del model els autors es fixen tres criteris a seguir:

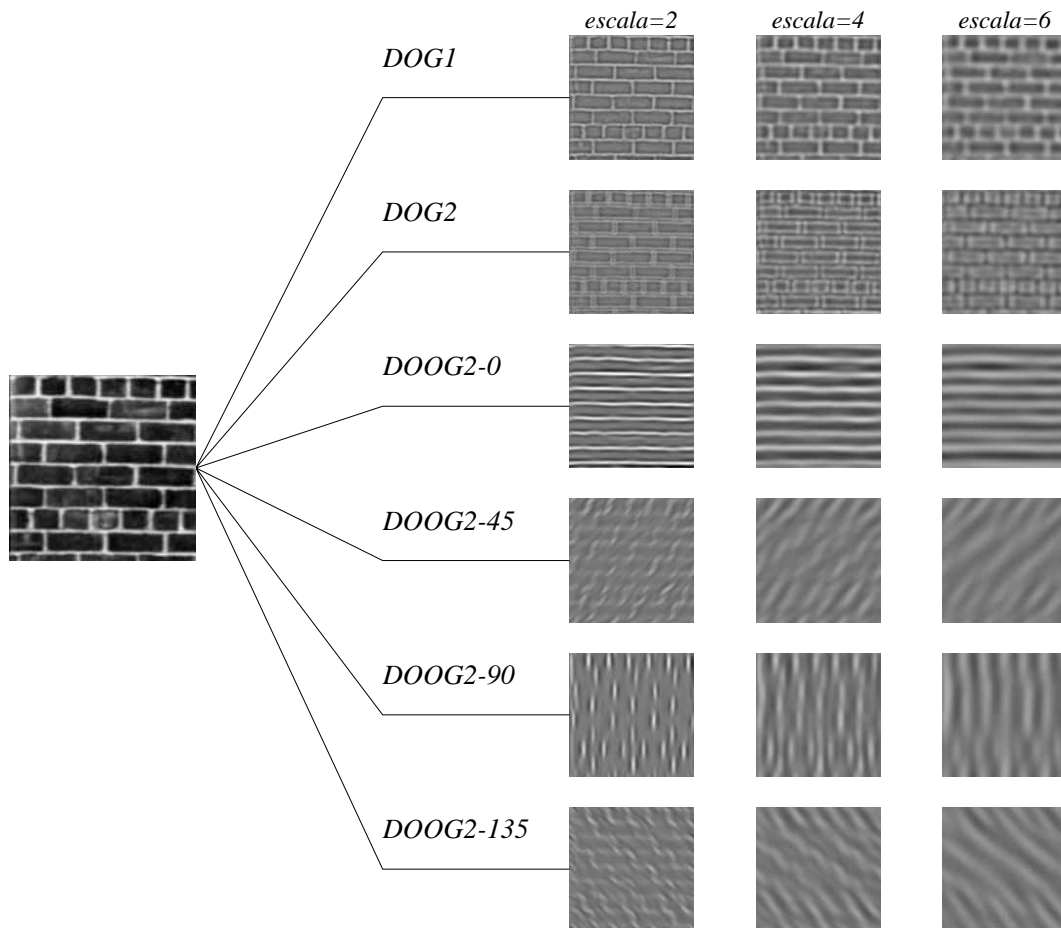


FIGURA 3.2: Exemple de la representació d'una imatge a partir d'un conjunt de filtres definits per certes escales i orientacions.

Plausibilitat biològica Els estats del model haurien de ser motivats i consistents amb mecanismes fisiològics coneguts de visió primària.

Generalitat El model hauria de ser prou general de manera que pugui ser aplicat a qualsevol imatge amb nivells de gris.

Correspondència quantitativa amb dades psicofísiques El model hauria de donar una predicció quantitativa sobre la importància d'una frontera entre dues regions amb textures diferents. Aquestes mesures haurien de coincidir amb les mesures psicofísicament.

L'esquema general del model ve donat en els següents tres estats:

- **Estat 1** Es modelitzen les sortides de cel·les simples. La imatge es convolucionada amb un banc de filtres lineals F_k seguit d'una rectificació de mitja ona. Aleshores tindrem un conjunt de respostes neuronals que representarem com $R_i(x, y)$, on i indica el canal amb el que s'ha convolucionat.

Distingirem la part positiva i negativa de la resposta de la següent manera:

$$R^+(x, y) = \max\{R(x, y), 0\} \quad (3.1)$$

$$R^-(x, y) = \max\{-R(x, y), 0\} \quad (3.2)$$

Aleshores el conjunt de respostes obtingudes serà:

$$R_{2k} = (I * F_k)^+(x, y) \quad (3.3)$$

$$R_{2k+1} = (I * F_k)^-(x, y) \quad (3.4)$$

Els filtres radialment simètrics modelitzen les cel·les simples no orientades, i els filtres orientats en una direcció modelitzen cel·les simples sensibles a barres.

- **Estat 2** Consisteix en una inhibició no lineal, que elimina respostes dèbils quan hi ha respostes més importants en posicions acostades. La postinhibició l'expressarem com

$$PIR_i(x_0, y_0) = \max_{x, y \in S_i(x_0, y_0)} \frac{1}{1 - \alpha_{ii}} (R_i(x, y) - T_i(x, y))^+ \quad (3.5)$$

on $S_i(x_0, y_0)$ és un veïnatge de mostreig, i T_i representa una transformació d'umbralització definida de la següent manera:

$$T_i(x_0, y_0) = \max_j \max_{(x,y) \in I_{ji}(x_0, y_0)} \alpha_{ji} R_j(x, y) \quad (3.6)$$

on I_{ji} és un veïnatge de (x_0, y_0) , en el que les neurones del canal j poden inhibir les neurones del canal i , i α_{ji} és una mesura de l'efectivitat de la inhibició.

- **Estat 3** Càlcul del gradient de la textura, que es defineix com

$$\max_i \|\nabla(PIR_i * G_{\sigma'})\|(x, y) \quad (3.7)$$

Com ja hem dit abans l'estructura general del model segueix els resultats de [48, 5, 107], per tant, fan la següent assumpció:

Assumpció 3.1 *En visió preatentiva les relacions posicionals entre textons no són importants, per tant tot es redueix a un problema de densitats de textons.*

Aleshores, dues textures T_1 i T_2 , són discriminables si tenen diferents promitjos espacials d'algunes respostes calculades localment, això és:

$$\int \int_{T_1} R(x, y) \neq \int \int_{T_2} R(x, y) \quad (3.8)$$

Per tant els mecanismes triats per a la construcció del model es basen en el càlcul de respostes neurals R , seguit d'un mecanisme pel càlcul del gradient. Anem a veure més detalladament cada un dels estats del model.

3.2.1 Selecció dels filtres lineals

Els autors argumenten que qualsevol de les següents famílies de filtres és igualment vàlida per aquest estat del model, que modelitza cel·les simples o subunitats de cel·les complexes:

- Funcions de Gabor
- Diferències de gaussianes desplaçades (*DOOG*)¹

¹Simplificació de l'anglès *Differences of offset gaussians*.

- Diferències de diferències de gaussianes desplaçades

Qualsevol d'aquestes famílies de funcions són igualment vàlides. En aquest model els autors han triat les diferències de gaussianes desplaçades basades en el model de derivades de gaussianes de R.A. Young [125]. La selecció d'aquests filtres és argumentada en el fet de que es comporten de forma molt semblant a mesures fisiològiques fetes i perquè computacionalment són molt simples.

Aquesta família de filtres es construeix a partir de la funció gaussiana següent:

$$G(x_0, y_0, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\left(\frac{x-x_0}{\sigma_x}\right)^2 + \left(\frac{y-y_0}{\sigma_y}\right)^2\right)} \quad (3.9)$$

a partir d'aquesta es defineixen les classes $DOG1(\sigma)$ i $DOG2(\sigma)$ de filtres radialment simètrics i que modelitzen cel.lles simples no orientades. La classe $DOG1(\sigma)$ s'obté com a combinació lineal de dues gaussianes concèntriques

$$DOG1(\sigma) = aG(0, 0, \sigma_i, \sigma_i) + bG(0, 0, \sigma_0, \sigma_0) \quad (3.10)$$

amb $\sigma_i : \sigma : \sigma_0$ en una relació de 0.71 : 1 : 1.14, i els coeficients $a : b$ en una relació de 1 : -1.

La classe $DOG2(\sigma)$ s'obté com a combinació lineal de tres gaussianes concèntriques

$$DOG2(\sigma) = aG(0, 0, \sigma_i, \sigma_i) + bG(0, 0, \sigma, \sigma) + cG(0, 0, \sigma_0, \sigma_0) \quad (3.11)$$

amb $\sigma_i : \sigma : \sigma_0$ en una relació de 0.62 : 1 : 1.6, i $a : b : c$ a -1 : 2 : -1.

I finalment tenim la classe de filtres $DOOG2$ que s'obtenen com a combinació lineal de tres funcions gaussianes idèntiques desplaçades

$$DOOG2(\sigma, r) = aG(0, y_a, \sigma_x, \sigma_y) + bG(0, y_b, \sigma_x, \sigma_y) + cG(0, y_c, \sigma_x, \sigma_y) \quad (3.12)$$

on $\sigma_y = \sigma$, $\sigma_x = r \cdot \sigma$, els autors suggereixen $r = 3$, els desplaçaments són $y_a = -y_c = \sigma$, $y_b = 0$, i els coeficients són $a : b : c$ com -1 : 2 : -1, pel filtre amb eix de simetria en la direcció x ($\theta = 0$). Els altres filtres d'aquesta classe $DOOG2$ s'obtenen per rotació sobre el centre de la gaussiana del mig.

Classe de filtres	σ	factor	desplaçament
$DOG1(\sigma) = aG(0, 0, \sigma_i, \sigma_i)$ + $bG(0, 0, \sigma_0, \sigma_0)$	$\sigma_i : \sigma : \sigma_0$ 0.71 : 1 : 1.14	$a : b$ 1 : -1	
$DOG2(\sigma) = aG(0, 0, \sigma_i, \sigma_i)$ + $bG(0, 0, \sigma, \sigma)$ + $cG(0, 0, \sigma_0, \sigma_0)$	$\sigma_i : \sigma : \sigma_0$ 0.62 : 1 : 1.6	$a : b : c$ -1 : 2 : -1	
$DOOG2(\sigma, r) = aG(0, y_a, \sigma_x, \sigma_y)$ + $bG(0, y_b, \sigma_x, \sigma_y)$ + $cG(0, y_c, \sigma_x, \sigma_y)$	$\sigma_y = \sigma$ $\sigma_x = r \cdot \sigma$ ($r = 3$)	$a : b : c$ -1 : 2 : -1	$y_a = -y_c = \sigma$ $y_b = 0$
$a_{DOG1} : a_{DOG2} : a_{DOOG2}$ 3 : 4.15 : 2			

TAULA 3.1: Expressions corresponents a la família de filtres *DOOG*, formada per les classes *DOG1*, *DOG2* i *DOOG2*.

$$G(x_0, y_0, \sigma_x, \sigma_y, \theta) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\left(\frac{(y-y_0)\cos(\theta)+(x-x_0)\sin(\theta)}{\sigma_x}\right)^2 - \left(\frac{(x-x_0)\cos(\theta)-(y-y_0)\sin(\theta)}{\sigma_y}\right)^2\right)} \quad (3.13)$$

Els coeficients d'escala $a_{DOG1} : a_{DOG2} : a_{DOOG2}$ segueixen la relació 3 : 4.15 : 2, designat així per equalitzar el rang dinàmic de les respectives respostes. Cal tenir en compte que tots aquests filtres són de mitja zero. A la figura 3.3.a podem veure les imatges corresponents als filtres definits, i a la figura 3.3.b els seus perfils.

El paràmetre σ de les tres classes de filtres usades es donen en freqüència espacial de cicles/graus, donat que es coneixi la distància a la vista i la mida de la imatge. Per mostrejar l'espai de freqüències els autors agafen un rang de tots els valors sencers entre 3 i 14 cicles/grau. Aquest mostreig fa que es tinguin 96 filtres diferents.

En el model *DOOG* s'assumeix que el perfil del camps receptius en la direcció perpendicular a l'eix d'orientació del filtre, són bé de simetria parell, és a dir

$$f(t) = f(-t) \quad (3.14)$$

o bé de simetria senar,

$$f(t) = -f(-t) \quad (3.15)$$

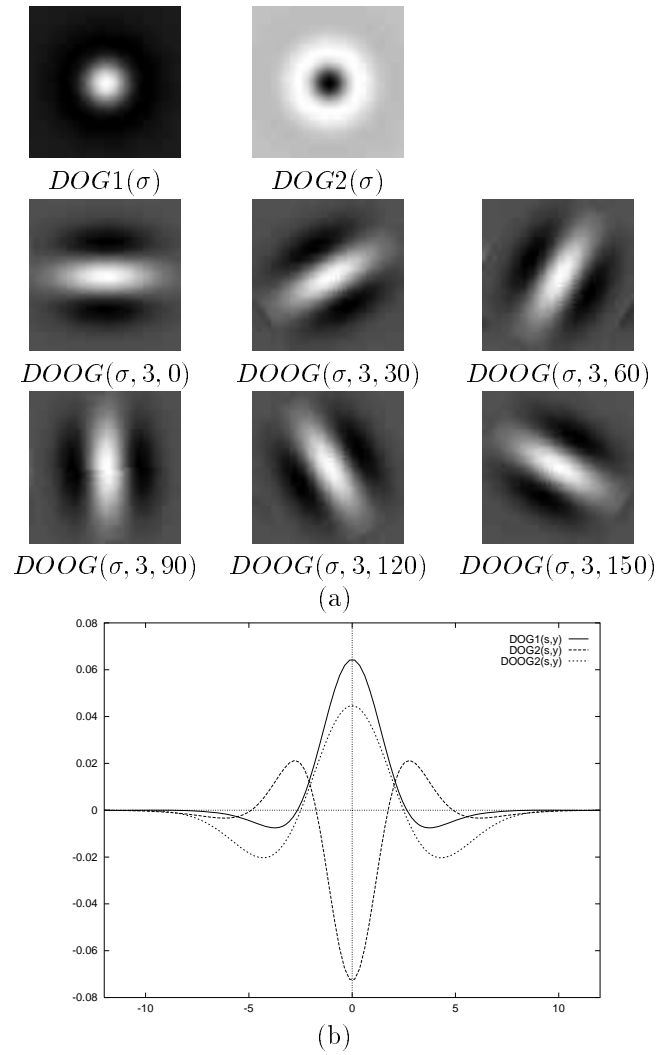


FIGURA 3.3: (a) Imatges corresponents als filtres $DOG1$, $DOG2$ i $DOOG2$. (b) Perfil d'aquest filtres segons l'eix y .

i no de qualsevol fase intermitja. Aquest model ha estat proposat a partir d'experiments psicofísics, que confirmen l'existència de detectors de barres i contorns [12]. A més a més els autors del model fan la següent assumpció:

Assumpció 3.2 *La informació que prové de mecanismes de simetria senar, no és usada en percepció de textures, i per tant s'exclou en aquest model. Els mecanismes de simetria senar responen òptimament a contorns orientats.*

Aquesta assumpció es basa en resultats experimentals que demostren que parells de textures formades per patrons que presenten simetria de mirall entre elles no són discriminables preatentivament. Mentre que parells de textures que no presenten aquesta simetria si que són discriminables. A la figura 3.4.a podem veure una imatge formada per dues textures, amb patrons P_1 i P_2 , tals que presenten simetria de mirall

$$P_1(x) = P_2(-x) \quad (3.16)$$

respecte de la línia que separa les dues textures, aquestes textures no són preatentivament discriminables. Per altra banda a la figura 3.4.b tenim un parell de textures que no presenten la simetria anterior sinó que compleixen que

$$P_1(x) = -P_2(-x) \quad (3.17)$$

i són fàcilment discriminables. Això fa pensar que les respostes als filtres de simetria parell i senar no són tractades de la mateixa manera. Es pot deduir que el signe de les respostes de filtres de simetria senar no són considerats, mentre que el signe de les respostes dels filtres de simetria parell si que ho són. D'aquí ve el fet de que en aquest model només es considerin filtres de simetria parell.

3.2.2 Introducció de mecanismes no lineals

Fins ara només s'ha usat l'operador de la convolució, per tant només hem vist mecanismes lineals. Ara es presenta la necessitat d'introduir mecanismes no lineals.

Suposem dues textures T_1 i T_2 amb promig d'intensitats igual. Si es convolucionen amb el filtre F obtenim les respostes

$$R_{T_1} = T_1 * F \quad (3.18)$$

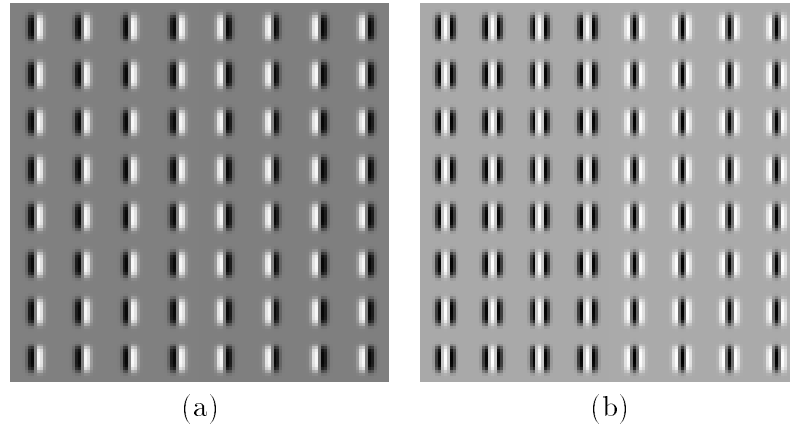


FIGURA 3.4: (a) Parell de textures formades per patrons que presenten simetria de mirall i que no són discriminables preatentivament. (b) Parell de textures formades per patrons sense simetria de mirall i que són preatentivament discriminables.

$$R_{T_2} = T_2 * F \quad (3.19)$$

que per les propietats de convolució continuen tenint promitjos espacials iguals. Per altra banda, es poden trobar exemples de textures amb promitjos espacials iguals que són discriminables preatentivament. Per tant, es necessita una no-linealitat en el sistema per a tenir un comportament d'acord amb el del SVH.

Per tant, en aquest model es fa una tercera assumpció:

Assumpció 3.3 *En un model computacional de discriminació de textures es necessita aplicar alguna no-linealitat sobre les respostes dels filtres convolucionats per a poder discriminar textures amb promitjos espacials iguals, i que l'ull humà és capaç de discriminar.*

Un exemple que justifica aquesta assumpció és la textura de la figura 3.4.b. Les dues textures d'aquesta imatge han estat formades a partir de funcions amb simetria parell i cada una d'elles presenta un promig espacial igual a 0 respecte del fons de la textura, l'única diferència ve donada pel signe del contrast. Per tant un promig espacial sobre les dues donarà el mateix valor. Il·lustrem aquest efecte a la figura 3.5.

La no-linealitat seleccionada en aquest model és la *rectificació de mitja ona*, que modelitza dues cel·les diferents per representar la part positiva i la part negativa de la resposta del filtre

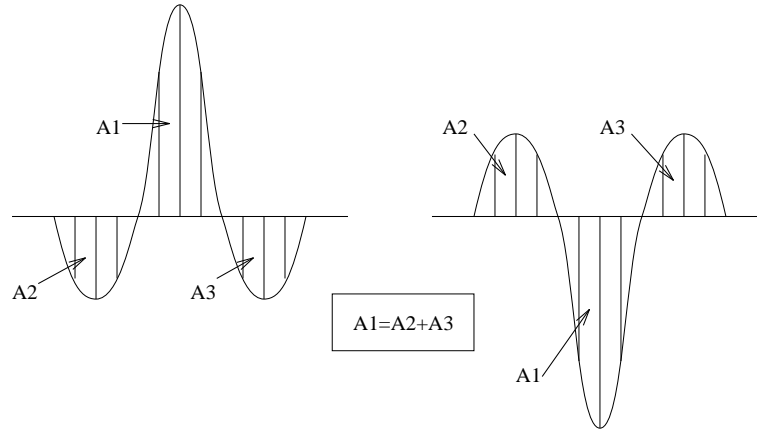


FIGURA 3.5: Exemple de patrons texturals que tenen el mateix promig en el seu entorn però que poden formar part de textures discriminables preatentivament.

$$\begin{aligned} R_{2k}^+(x, y) &= \max\{(I * F_k)(x, y), 0\} \\ R_{2k+1}^-(x, y) &= \max\{-(I * F_k)(x, y), 0\} \end{aligned} \quad \forall k \in \{0, \dots, 95\} \quad (3.20)$$

A altres models s'han fet servir altres mecanismes no lineals com:

1. *Rectificació completa* (J.R. Bergen i E.H. Adelson a [6])

$$R_k(x, y) = |(I * F_k)(x, y)| \quad (3.21)$$

2. *Càlcul de l'Energia* (I. Fogel i D. Sagi a [28])

$$R_k(x, y) = |(I * S_k)(x, y)|^2 + |(I * C_k)(x, y)|^2 \quad (3.22)$$

on S_k i C_k són una parella de filtres que estan en quadratura de fase.

Els autors donen dues raons per a l'elecció de la rectificació de mitja ona:

- És l'elecció més natural en el contexte de les evidències biològiques actuals, ja que és conegut que el filtratge lineal seguit de rectificació de mitja ona és una bona aproximació per modelitzar les respostes de cel·les simples i les subunitats de cel·les complexes. No s'ha demostrat que hi hagi cel·les que calculin l'energia de les respostes.

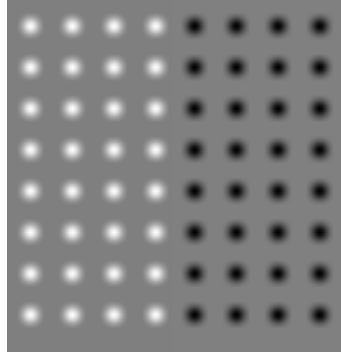


FIGURA 3.6: Exemple de textures en que els elements texturals només varien en el signe i són clarament discriminables.

- Tant en el càlcul de l'energia com amb la rectificació completa, es perd el signe de la resposta. Per tant dues textures que només variïn en el signe de les seves respostes no serien discriminables. Aquest seria el cas de dues textures una amb patró M i l'altra $-M$, si perdem el signe de les respostes aquestes són iguals i per tant no distingibles, i el nostre sistema visual és capaç de discriminar-les.

D'aquestes raons se'n deriva una consideració important, i és que el SVH és capaç de discriminar textures amb patrons invertits en nivell de gris. Aquest fet ens fa deduir que no es pot perdre el signe de les respostes dels filtres ja que sinó textures formades per patrons $(M, -M)$ com els de la figura 3.6, no serien discriminables.

En passos posteriors s'hauran d'introduir més mecanismes no lineals, ja que en algun moment s'han d'unir les respostes fent un promig espacial d'aquestes en àrees més grans que el tamany que ocupa un patró de la textura. Per tant, si no introduïm cap altra no-linealitat tindrem que textures formades amb patrons de mitja zero (com la imatge de la figura 3.4.b) el promig espacial de les seves respostes després de la rectificació tampoc no permet la discriminació.

Tenint en compte que els filtres triats són de mitja zero, això és

$$\iint F = 0 \quad (3.23)$$

aleshores per a una textura amb patró M de mitja zero, les sortides dels canals positiu i negatiu tenen la mateixa mitja

$$\int \int (M * F)^+ = \int \int (M * F)^- \quad (3.24)$$

i entre textures diferents tenim

$$(M * F)^- = (-M * F)^+ \quad (3.25)$$

per tant els promitjos espacials entre dues textures diferents compleix

$$\int \int (M * F)^+ = \int \int (M * F)^- = \int \int (-M * F)^+ = \int \int (-M * F)^- \quad (3.26)$$

Per tant la discriminació és impossible si no afegim alguna altra no linealitat abans de fer el promig espacial.

Aquesta segona no-linealitat és segons els autors del model, la principal mancança, ja que no ha estat demostrada pels resultats que hi ha actualment. Abans de triar aquesta nova no-linealitat es planteja la següent qüestió:

On s'aplica dins del model?, abans de la resposta?, o després?

Una possible no-linealitat primària bastant coneguda és l'**adaptació retinal**, s'ha vist que aquesta juga un paper important en discriminació de textures. Aquesta no-linealitat permet al sistema de respondre a contrastos locals amb canvis d'il·luminació que estan per sobre de 10 unitats logarítmiques.

La qüestió és si aquesta és suficient per a discriminar textures. Tenint en compte que la majoria d'informació sobre discriminació de textures ens ve d'imatges binàries i per tant amb contrast constant, aleshores en aquests casos l'adaptació retinal no pot tenir un paper massa important. Per tant sembla clar que farà falta una no-linealitat més tardana.

Existeixen dues eleccions possibles fisiològicament:

Funció no lineal de resposta al contrast Funció amb forma sigmoidal que es presenta a neurones que tenen un efecte d'umbral per contrastos baixos i de saturació per contrastos alts.

Inhibició intracortical Pot ocórrer a dins i entre diferents canals.

Aquesta darrera és la que ha estat elegida per aquest model, i s'explica en el següent

apartat.

3.2.3 Inhibició intracortical

Tècniques anatòmiques i fisiològiques han demostrat l'existència d'interaccions i connexions inhibidores entre les neurones del córtex visual primari. Els experiments fisiològics no han arribat a un model definitiu per aquesta inhibició, i és bastant possible que hi hagi varis circuits amb diferents papers. Tenint tot això en compte, en aquest model s'estableix el següent criteri:

Criteri *En aquest model s'aplica la inhibició intracortical per a la supressió de respostes de canals que no han estat ben sintonitzats amb l'estímul, i que anomenarem respostes falses.*

D'aquesta manera tindrem que les respostes més altes corresponen a canals ben sintonitzats amb l'estímul. Per evitar que els promitjos espacials de respostes sorolloses puguin tenir els mateixos valors que els promitjos de les respostes ben sintonitzades, el model proposa la definició d'una funció d'umbralització

$$T_i(x_0, y_0) = \max_j \max_{x,y \in I_{ji}(x_0, y_0)} \alpha_{ji} R_j(x, y) \quad (3.27)$$

on $T_i(x_0, y_0)$ representa umbralitzacions per neurones que pertanyen al canal i amb coordenades retinotòpiques (x_0, y_0) . El factor I_{ji} representa un veïnatge de (x_0, y_0) , en el que les neurones en el canal j poden inhibir neurones en el canal i , i α_{ji} és un coeficient de l'efectivitat d'aquesta inhibició.

Fent servir aquesta funció la resposta de la postinhibició ve donada per

$$PIR_i(x_0, y_0) = \max_{x,y \in S_i(x_0, y_0)} \frac{1}{1 - \alpha_{ii}} [R_i(x, y) - T_i(x, y)]^+ \quad (3.28)$$

que suposa una supressió de respostes que estan per sota l'umbral. La funció $S_i(x_0, y_0)$ és un veïnatge, a partir del que les respostes fortes del canal i són seleccionades per a un processament posterior.

El mecanisme aplicat pot ser vist dins de les xarxes neuronals com una xarxa amb realimentació on *el líder s'emporta la majoria*, una variant de les *guanyador s'ho emporta tot*.

α_{ji}		j		
		$DOG1(\sigma_j)$	$DOG2(\sigma_j)$	$DOOG2(\sigma_j, r_j, \theta_j)$
i	$DOG1(\sigma_i)$	0.20	0.45	0.15
	$DOG2(\sigma_i)$	0.45	0.25	0.20
	$DOOG2(\sigma_i, r_i, \theta_i)$	0.15	0.20	$0.65\delta(\theta_i, \theta_j)$

TAULA 3.2: Coeficients inhibitoris α_{ji} pel model A, on $\delta(\theta_i, \theta_j) = 1$ si $\theta_i = \theta_j$ i 0 en altre cas. Els coeficients inhibitoris pel model B, C i D són constants: $\alpha_{ji}^B \equiv 0.5$, $\alpha_{ji}^C \equiv 0$ i $\alpha_{ji}^D \equiv 0$.

I_{ji}		j		
		$DOG1(\sigma_j)$	$DOG2(\sigma_j)$	$DOOG2(\sigma_j, r_j, \theta_j)$
i	$DOG1(\sigma_i)$	$2\sigma_j$	$1.5\sigma_j$	$1.25\sigma_j$
	$DOG2(\sigma_i)$	$2\sigma_j$	$1.5\sigma_j$	$1.25\sigma_j$
	$DOOG2(\sigma_i, r_i, \theta_i)$	$2\sigma_j$	$1.5\sigma_j$	$1.25\sigma_j$

TAULA 3.3: Radis dels veinatges inhibitoris I_{ji} dels models A i B. Als models C i D els veinatges són irrelevants.

Els valors d' I_{ji} i d' α_{ji} són triats segons criteris que permetin eliminar respostes sorolloses. Aquests criteris s'han definit a partir de l'estudi de les respostes dels filtres sobre estímuls amb la mateixa forma que els filtres, de manera que per a cada cas s'ha definit quina era la resposta ben sintonitzada i quina era la resposta sorollosa depenent de l'estímul i del filtre (Aquestes respostes es poden veure a la figura 3.7).

Les diferents maneres de donar valors als factors α_{ij} , I_{ij} fa que s'hagin definit els tres següents models.

Model A És el model més complicat, però és també el que ha donat millors resultats. Els valors dels coeficients i dels veinatges es donen a les taules 3.2 i 3.3.

Model B Aquest dóna la mateixa efectivitat a totes les inhibicions $\alpha_{ji} = 0.5$, això significa no tenir grups específics d'inhibició de neurones, fet que fa que sigui més consistent amb evidències fisiològiques.

Model C En aquest model s'elimina el pas de la primera inhibició fent $\alpha_{ji} = 0$, però si que realitza la postinhibició sobre un veinatge S_i . Aquest model es pot relacionar amb

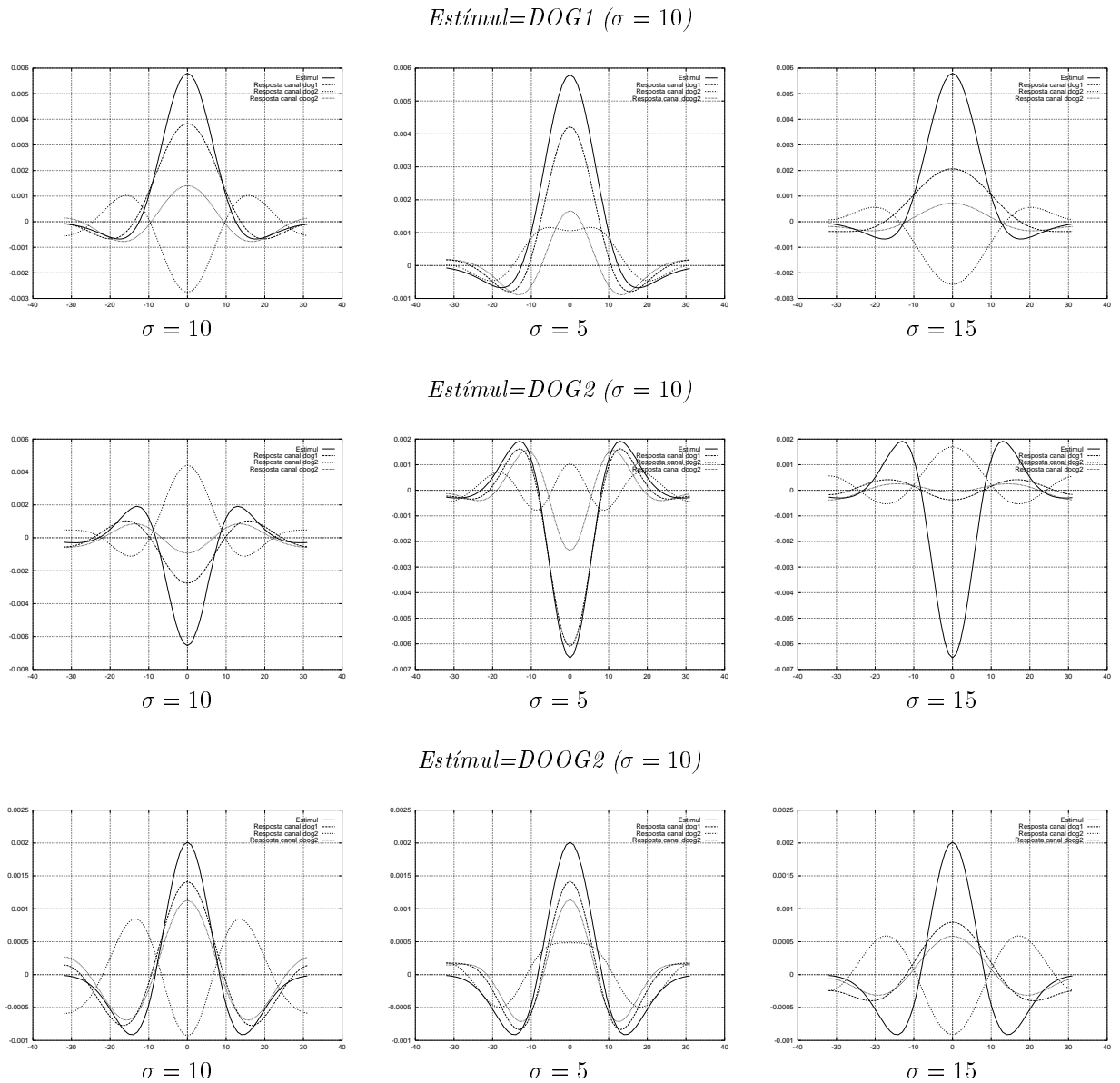


FIGURA 3.7: Perfils de les respostes dels filtres *DOOG* a estímuls de la mateixa forma que els filtres.

les sortides de cel·les complexes que no mostren modulació significativa a les seves respostes quan un estímul es mou a través del camp receptiu.

Model D Aquest model no realitza aquesta postinhibició, per tant $\alpha_{ji} = 0$, i $PIR_i = R_i$.

En el model A s'aplica el conegut agrupament entre neurones, això és, neurones sin-tonitzades sense cap orientació tendeixen a ocórrer juntes, i neurones que comparteixen la mateixa orientació tendeixen també a ocórrer juntes. Per tant, aquest agrupament ens porta a tenir 8 grups de canals, 2 radialment simètrics i 6 orientats. Els valors de I_{ji} i α_{ji} són idèntics per a tots els canals i dins un d'aquests grups. Els valors d'aquests apareixen a les taules 3.2 i 3.3.

3.2.4 Càlcul del gradient de textura

S'han donat dues característiques com les bàsiques que tot model de discriminació de textures ha de complir:

- La discriminació depèn de la densitat dels patrons de la imatge, que amb densitats altes porten a una discriminació més fàcil.
- La discriminació depèn de les diferències locals i no de les diferències globals.

Per tant, aquestes dues idees fan que es defineixi el càlcul del gradient de les respostes de la postinhibició suavitzades. Aquest gradient es defineix com:

$$\max_i \nabla (PIR_i * G_{\sigma'}) (x, y) \quad (3.29)$$

Biològicament, aquest gradient pot ser calculat usant mecanismes orientats, amb simetria senar, similars a les cel·les sensibles als contorns. Aquests mecanismes responsables de calcular el gradient de la textura tenen camps receptius grans (σ' donarà la mida d'aquests), i ocorren en àrees extraestriades.

Finalment, l'operació de màxim és una manera habitual de combinar les sortides de diferents canals. Les fronteres de les diferents textures correspondran als pics del màxim de la magnitud del gradient calculat.

A la figura 3.8 donem una representació gràfica de tots els nivells del model.

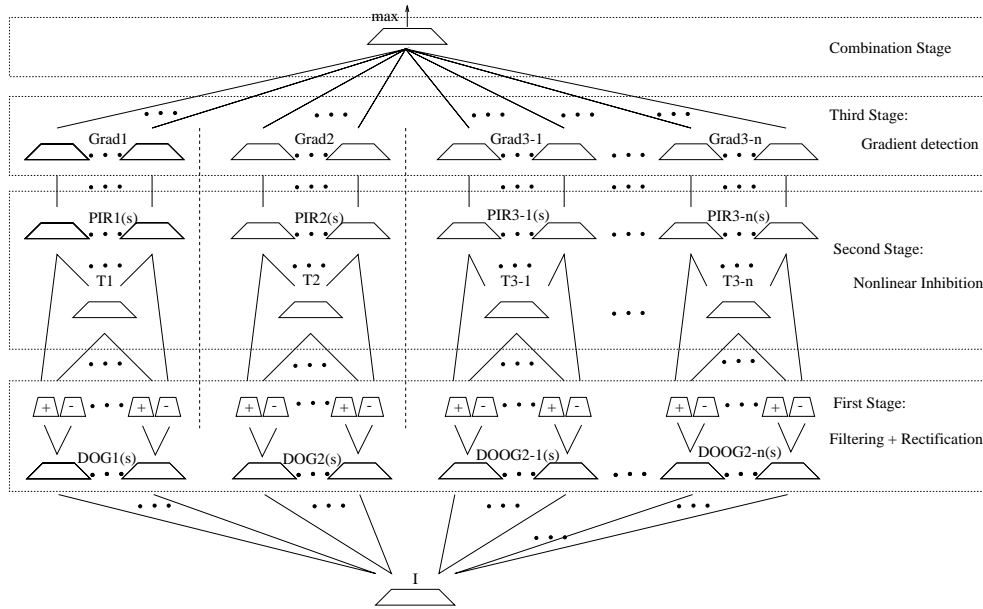


FIGURA 3.8: Diagrama dels nivells del model computacional de discriminació preatentiva de textures de J. Malik i P. Perona

3.3 Definició d'una representació general de textures

El model computacional de segmentació preatentiva de textures que acabem de presentar dona com a resultat el gradient textural de la imatge d'entrada. El gradient textural apareix entre dues textures si existeixen diferències detectades pel model. Aquest gradient és calculat sobre les respostes que resulten del segon estat del model, això és després de la postinhibició.

A cada canal intermig del model s'hi representa informació referent a un determinat tipus de característica. Així doncs, els canals resultants de convolucionar amb els filtres $DOG1$ i $DOG2$ es poden veure com la detecció de les taques isotròpiques de la textura, i els canals dels filtres $DOOG2$ com la detecció de barres orientades de la imatge. Cada tipus de característica té un canal per a cada escala.

Els canals que han detectat alguna característica d'acord amb la seva forma, orientació o escala presenten un conjunt de pics en tots aquells punts de la imatge on s'ha fet la detecció. El nivell de gris d'aquests pics depèn de la sintonització del filtre amb la imatge i del contrast del textó detectat. A la figura 3.9 veiem un exemple del comportament d'aquest

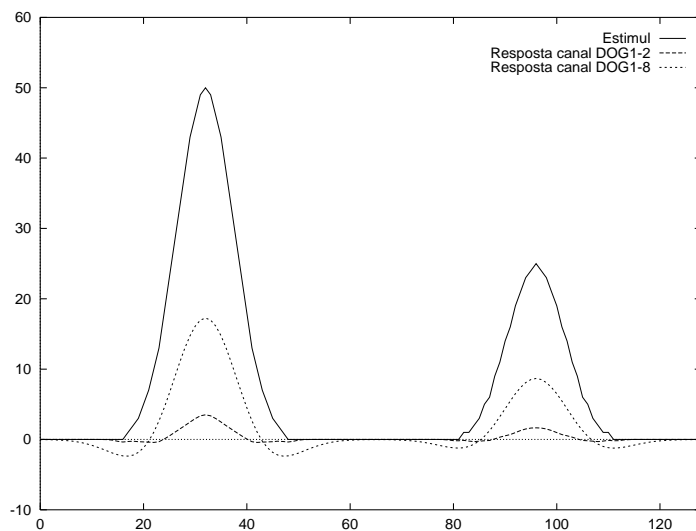


FIGURA 3.9: Perfil d'una línia de la imatge resultant de convolucionar amb els filtres *DOG1* amb escala 2 i 8. La línia contínua de la gràfica representa la imatge original formada per dues funcions gaussianes de mida 8 i amb diferent contrastos. Les línies puntejades representen les respostes dels filtres a la imatge.

filtre, el filtre de mida més petita que la taca de la imatge dóna una resposta petita, mentre que el de la mateixa mida que la taca dóna un màxim important. Per altra banda el valor del màxim podem veure que depèn del contrast de la taca.

La rectificació de mitja ona separa en canals diferents les respostes degudes a patrons amb contrastos positius o negatius respecte del fons de la imatge, això és, taques blanques o taques fosques respecte del fons en cada punt de la imatge. A la figura 3.10 es veu la necessitat de definir una rectificació que conservi el signe de la taca, si no fós així les respostes per taques blanques i negres amb el mateix contrast sempre serien iguals.

En el segon estat del model la inhibició s'encarrega d'eliminar totes les respostes sorolloses. Primer, restant la funció T_i calculada per a cada canal. Aquesta substracció s'encarrega d'eliminar les respostes negatives i les mal sintonitzades que ens han aparegut a les respostes de la figura 3.9. La postinhibició final s'encarrega d'expandir els pics importants, de manera que es donin regions amb un nivell de gris homogeni sempre que el radi de la postinhibició S_i sigui tant gran com la densitat de la característica que correspon a cada canal (veure apartat 3.3.1).

A la figura 3.11.a podem veure el resultat d'aplicar la post-inhibició sobre la resposta

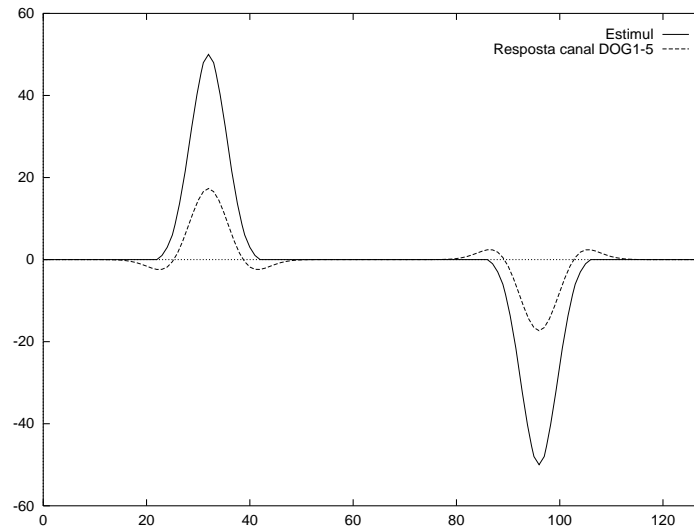


FIGURA 3.10: Perfil d'una línia de la imatge resultant de convolucionar amb el filtre *DOG1* amb escala 5. La línia contínua de la gràfica representa la imatge original formada per dues funcions gaussianes de diferent signe i de mida 5. Les línies puntejades representen les respostes del filtre a la imatge.

positiva d'un filtre amb un radi d'inhibició prou gran que fa que apareixin dues regions homogènies resultants dels dos tipus de taques, mentre que a la figura 3.11.b el radi és més petit i fa aparèixer diverses regions que crearien més d'un gradient textural, malgrat només existeixin dues textures diferents a la imatge.

Per a que existeixi un gradient textural entre dues textures d'una imatge hi ha d'haver algun canal previ al càlcul del gradient, que presenti dues zones amb nivell de gris homogeni dins cada regió i diferent entre ells. De manera que l'operador de gradient detecti la diferència en el nivell de gris.

Si apliquem els dos primers estats del model (filtratge, rectificació i postinhibició) sobre una imatge que presenta una sola textura, tindrem que els diferents canals estaran formats, en principi, per un únic nivell de gris homogeni per tota la imatge. Aquest conjunt de nivells de gris que es distribueixen en tots els canals són els que permetrien discriminar aquesta textura de totes les altres.

A partir d'aquesta interpretació podem dir que el conjunt de valors de nivell de gris distribuït en els diferents canals representa la textura des d'un punt de vista perceptual i preatentiu. Aleshores, en aquest treball proposem aquest vector de valors de nivell de gris

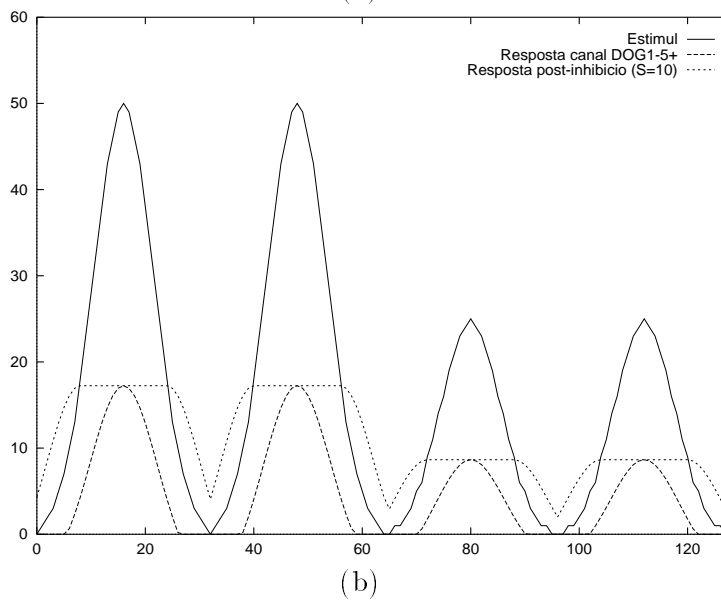
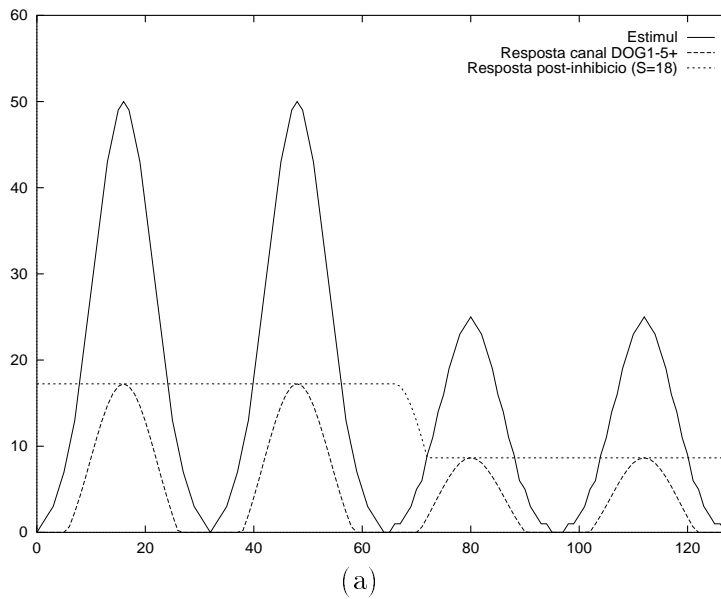


FIGURA 3.11: La línia contínua de la gràfica representa la imatge original formada per quatre funcions gaussianes de la mateixa mida i diferent contrast. Les línies puntejades representen la part positiva del filtre i la post-inhibició. (a) La post-inhibició ha estat feta amb $S = 18$. (b) La post-inhibició ha estat feta amb $S = 8$.

com una representació de la textura, i que ara definim més formalment:

Definició 3.1 *Donada una imatge de dimensions $N \times N$, que conté la textura t , definim la seva representació r , com*

$$r(t) = (v_1, \dots, v_n) \quad \text{on} \quad v_i = \sum_{(x,y) \in \{0, \dots, N\}^2} \frac{PIR_i(x, y)}{N^2} \quad (3.30)$$

on n depèn del nombre de canals del model, en aquest cas $n = 192$.

Aquesta representació ens permet considerar una textura com un punt a un espai de 192 dimensions.

Des d'aquest punt de vista la discriminabilitat entre dues textures pot venir donada per la proximitat en aquest espai de representació que acabem de definir. Per tant, la distància entre textures en aquest espai pot ser considerada com una mesura de dissimilaritat entre elles.

3.3.1 Radi de la post-inhibició i densitat de textons

En aquesta representació que acabem de definir queda un factor sense determinar i que és important en l'efecte que pot tenir sobre la representació. En el model no s'especifica cap valor pel paràmetre S_i , que representa el radi del veinatge de la postinhibició.

La funció de la post-inhibició és la d'expandir les respostes ben sintonitzades de manera que en l'expansió tota la imatge del canal prengui com a valor de nivell de gris el valor de les respostes. Per a que això sigui així el radi d'aquesta operació ha de ser prou gran com perquè sigui capaç de juntar respostes que estiguin relativament lluny entre elles.

Sembla doncs, que el radi d'aquesta post-inhibició està relacionat amb la densitat dels textons. A la teoria del textó es parla d'algunes distàncies crítiques que s'han de donar entre els elements texturals per a que la discriminació preatentiva es pugui dur a terme [48, 86, 70, 85]. Vegem-les:

\triangleleft -*neighborhood* és el veinatge de mida \triangleleft dins del que els elements d'una textura han de caure per a que aquesta pugui ser tractada preatentivament, ja que representa l'àrea dins la que els textons són comptats i el seu gradient determinat. Aquesta distància ha estat quantificada i es demostra que la discriminació preatentiva es dóna quan els elements adjacents estan més aprop que dues vegades la mida de l'element, és a dir,

el diàmetre de \triangleleft -*neighborhood* es proporcional a dues vegades la mida de l'element textural.

Obertura de l'atenció focal ve donada per la distància dins de la que els textons d'una mateixa textura són ajuntats. La mida d'aquesta obertura sembla estar també linealment relacionada amb el tamany del segment.

ϵ és la distància màxima a la que poden estar els elements texturals entre ells per a formar un únic agrupament dins de la textura. Sempre ha de complir la següent restricció $\epsilon \ll \triangleleft$.

L'escalatge que permetria calcular la mida d'aquests veinatges respecte de la mida dels elements texturals és l'única operació global realitzada pel sistema visual preatentiu.

Per a la nostra representació aquests veinatges i la seva forma semblen ser un bon indicatiu de com seleccionar els valors dels radis de post-inhibició. En una primera aproximació sembla que el radi de post-inhibició sembla coincidir amb la mida de l'obertura de l'atenció focal. Sembla doncs, que s'hauria de calcular la densitat entre els diferents elements texturals que pot canviar dins de cada canal. Aleshores aquest problema en ell mateix obre una via de continuació per acabar d'automatitzar la representació.

De totes maneres si agafem un radi de post-inhibició prou gran aleshores ja evitem qualsevol problema. En canvi si que s'ha de solucionar el problema per a l'aplicació del model en problemes de segmentació.

3.4 Problemes de la representació definida

Per a que aquesta representació sigui una bona representació de textures s'han de solucionar alguns problemes que ara plantegem.

Primer de tot, una representació interessant serà aquella que el seu càlcul no suposi un cost computacional molt gran, ja que en la majoria d'aplicacions sempre ens interessarà que es pugui representar qualsevol textura de manera eficient. En principi, el model presenta un cost molt alt, tant pel nombre d'imatges intermitges que s'han de calcular com pel nombre d'operacions que s'han de realitzar.

Per altra banda, tenim el problema que es deriva de l'alta dimensionalitat en que es dona la representació. Aquest fet impedeix de donar una interpretació d'aquesta i per tant és difícil de predir el seu comportament davant de qualsevol imatge. A més a més, es fa

molt difícil establir quins són els criteris que determinen les similituds entre les diferents textures.

El model ha estat construït a partir de criteris biològics i se li demana que es correspongui amb dades psicofísiques del problema. La comprovació d'aquesta correspondència amb el comportament del sistema visual humà es fa difícil si tenim en compte l'alta dimensió de la representació. La dificultat ve donada pel fet que es fa difícil generar tots els experiments, amb textures concretes, que ens permetin establir la correspondència entre el comportament de la representació i el comportament del SVH.

Per tant podem concretar els problemes en els següents tres punts:

- El càlcul de la representació suposa un cost computacional molt alt.
- L'espai de representació té una dimensió molt alta, fet que provoca una manca de significat del que està representant.
- Dificultat de comprovació del comportament de la representació respecte de dades psicofísiques del SVH, degut principalment a que no tenim una interpretació perceptual de la representació.

En els capítols que resten ens proposem de donar solució als problemes que ens acaba de presentar aquesta representació. Primer de tot, al capítol 4 donem un algorisme millorat per fer el càlcul d'aquesta representació. En el capítol 5 definim un model paràmetric de textures a partir del qual intentarem trobar l'estructura de l'espai de representació. Aquesta estructura és la que ens permetrà posar en correspondència l'espai de representació amb l'espai perceptual generat pel SVH. Finalment, en el capítol 6 extenem els resultats del capítol 5 al cas general.

CAPÍTOL 4

ALGORISME EFICIENT PEL CÀLCUL DE LA REPRESENTACIÓ

En el capítol anterior hem explicat un model general de discriminació preatentiva de textures, ens hem fixat en els diferents estats i en la seva justificació. En aquest primer apartat donarem una descripció d'aquest model que ens permeti construir un algorisme per a calcular la representació d'una manera eficient. Tenir un algorisme eficient que realitzi aquesta tasca és important donat que qualsevol aplicació d'aquest model requerirà temps de càlcul petits.

Per a millorar el cost computacional de càlcul de la representació hem introduït operacions morfològiques en els passos no lineals. Aquestes operacions s'encarregaran de calcular els passos del nivell d'inhibició sobre veinatges isotròpics.

Les operacions morfològiques han estat millorades fent servir la propietat de la descomposició dels elements estructurants. Proposem un mètode de descomposició d'elements estructurants circulars, que representen el veinatge de la inhibició. Aquest mètode permet reduir la complexitat del càlcul de les dilatacions, passant d'una complexitat d'ordre quadràtic a una complexitat lineal.

4.1 Descripció dels passos de l'algorisme

Per al disseny de l'algorisme separarem els diferents mòduls o passos en els que es divideix el model:

Pas 1 Filtratge lineal de la imatge amb tots els filtres seleccionats.

Pas 2 Rectificació de mitja ona de cada una de les imatges resultants del pas anterior.

Pas 3 Càlcul de la funció d'umbral que elimini respostes mal sintonitzades.

Pas 4 Substracció de la funció del pas anterior i post-inhibició.

En aquests passos hi ha alguns paràmetres per considerar, el conjunt de valors que agafa el paràmetre σ dels filtres, el conjunt d'orientacions, que s'apliquen sobre el filtre *DOOG2* i la mida del veinatge d'inhibició, que com ja hem dit és interessant que estigui d'acord amb la densitat del tipus de textó corresponent. Anem a analitzar cada un d'aquests passos.

El filtratge lineal es dur a terme a partir d'una família de tres filtres que denotem com $\{f_1, f_2, f_3\}$, l'interval de valors que pot agafar la σ dels filtres, $\{\sigma_1, \dots, \sigma_k\}$ i el conjunt d'orientacions que s'han d'aplicar al filtre f_3 , $\{\theta_1, \theta_2, \dots, \theta_n\}$. A partir d'això s'ha de generar el següent conjunt de $k \times (n + 2)$ respostes:

$$\begin{aligned} R_{1,\sigma_j}(x, y) &= I * f_{1,\sigma_j} \quad \forall j \in \{1, \dots, k\} \\ R_{2,\sigma_j}(x, y) &= I * f_{2,\sigma_j} \quad \forall j \in \{1, \dots, k\} \\ R_{3,\sigma_j,\theta_i}(x, y) &= I * f_{3,\sigma_j,\theta_i} \quad \forall j \in \{1, \dots, k\}, \forall i \in \{1, 2, \dots, n\} \end{aligned} \quad (4.1)$$

A les aplicacions que s'han fet de l'algorisme s'ha agafat $k = 12$ i $n = 6$, de manera que $\theta_i \in \{0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ\}$, d'aquí que el valor de $k \times (n + 2)$ sigui 96.

El filtratge lineal es pot calcular en el domini de Fourier, fent servir la propietat del producte de convolució

$$R_i(x, y) = \mathcal{F}^{-1}\{\mathcal{F}\{I\} \cdot \mathcal{F}\{f_i\}\} \quad (4.2)$$

donat que

$$\mathcal{F}\{I\}(w_x, w_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I(x, y) e^{2\pi i(xw_x + yw_y)} dx dy \quad (4.3)$$

i la inversa es calcularà aplicant la transformació següent:

$$\mathcal{F}^{-1}\{H\}(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} H(w_x, w_y) e^{-2\pi i(xw_x + yw_y)} dw_x dw_y \quad (4.4)$$

Per a poder calcular aquesta transformada de Fourier sobre una imatge discreta aplicarem la transformada de Fourier discreta que s'obté amb la següent expressió

$$\mathcal{F}(I)(k_x, k_y) \equiv \sum_{n_y=0}^{N-1} \sum_{n_x=0}^{N-1} I(n_x, n_y) e^{2\pi i n_x k_x / N} e^{2\pi i n_y k_y / N} \quad (4.5)$$

Per aplicar-la es pot utilitzar l'algorisme *FFT* (*Fast Fourier Transform*) que per a una imatge de dimensions $N \times N = n = 2^k$ presenta una complexitat de

$$O(n \log_2 n) = O(n \log_2 2^k) = O(nk) \quad (4.6)$$

assumint la periodicitat respecte dels límits que hem pres:

$$\forall x, y \in \left[-\frac{N}{2}, \frac{N}{2}\right] f(x, y) = f\left(x + \frac{N}{2}, y\right) = f\left(x, y + \frac{N}{2}\right) = f\left(x + \frac{N}{2}, y + \frac{N}{2}\right) \quad (4.7)$$

Per a seleccionar el conjunt de filtres a aplicar el model indica que s'ha de mostrear adequadament el rang central de freqüències espacials entorn al pic de la funció de sensibilitat del contrast, per tant s'agafen les freqüències del següent conjunt $\{3cicles/grau, \dots, 14cicles/grau\}$. En el nostre cas hem agafat el píxel com unitat i li hem donat valor sencers que van de l'1 al 12.

La família de filtres *DOOG* definits per R. Young [125] com ja hem vist es formen a partir de la funció gaussiana de l'expressió 3.9. Pels filtres DOG1 i DOG2 aquesta funció està centrada al $(0, 0)$ i $\sigma_x = \sigma_y = \sigma$. Pels filtres orientats tenim l'expressió 3.13.

En el segon pas d'aquest estat s'aplica la rectificació de mitja ona, que suposa aplicar una operació de màxim sobre cada píxel de totes les imatges que són resultat del pas anterior, això és

$$\begin{aligned} R_k^+ &= \max\{R_k, 0\} \\ R_k^- &= \max\{-R_k, 0\} \end{aligned} \quad (4.8)$$

en el que el nombre de comparacions de $N \times N \times 96 = n \times 96$. Per altra banda el nombre d'imatges per a representar cada canal ha estat doblada, per tant els recursos de memòria en aquest pas són doblats.

Per calcular l'umbral que inhibirà aquelles respostes que no han estat ben sintonitzades amb el filtre, cal calcular la següent funció:

$$T_i(x_0, y_0) = \max_j \max_{(x,y) \in I_{ji}(x_0, y_0)} \alpha_{ji} R_j(x, y) \quad (4.9)$$

on α_{ji} és una mesura de l'efectivitat de la inhibició, i I_{ji} representa un veïnatge de (x_0, y_0) . Els valors que han de tenir aquests factors per a cada funció T_i han estat donats a les taules 3.2 i 3.3.

i	j				
	$R_1 \dots R_{2n}$	$R_{2n+1} \dots R_{4n}$	$R_{4n+1} \dots R_{6n}$...	$R_{2n(p+1)+1} \dots R_{4n+2np}$
			θ_1	...	θ_p
T_1	$\alpha_{ji} = 0.2$	$\alpha_{ji} = 0.45$	$\alpha_{ji} = 0.15$		
...	$I_{ji} = 2\sigma_j$	$I_{ji} = 1.5\sigma_j$	$I_{ji} = 1.25\sigma_j$		
T_{2n}					
T_{2n+1}	$\alpha_{ji} = 0.45$	$\alpha_{ji} = 0.25$	$\alpha_{ji} = 0.20$		
...	$I_{ji} = 2\sigma_j$	$I_{ji} = 1.5\sigma_j$	$I_{ji} = 1.25\sigma_j$		
T_{4n}					
T_{4n+1}	$\alpha_{ji} = 0.15$	$\alpha_{ji} = 0.20$	$\alpha_{ji} = 0.65$...	$\alpha_{ji} = 0$
...	$I_{ji} = 2\sigma_j$	$I_{ji} = 1.5\sigma_j$	$I_{ji} = 1.25\sigma_j$...	
T_{6n}					
...	"	"	$\alpha_{ji} = 0$...	$\alpha_{ji} = 0$
$T_{2n(p+1)+1}$	$\alpha_{ji} = 0.15$	$\alpha_{ji} = 0.20$...	$\alpha_{ji} = 0.65$
...	$I_{ji} = 2\sigma_j$	$I_{ji} = 1.5\sigma_j$	$\alpha_{ji} = 0$...	$I_{ji} = 1.25\sigma_j$
T_{4n+2np}					

TAULA 4.1: Extensió de les taules dels paràmetres d'inhibició del model A

Suposem que el conjunt de respostes resultants d'aplicar el pas 1 s'ha fet tenint en compte els següents paràmetres, n , és el nombre d'escapes aplicades per a cada tipus de filtre, i suposem p que representa el nombre d'orientacions seleccionades per als filtres *DOOG2*, aleshores després de la rectificació tindrem un total de $2n$ respostes pel filtre *DOG1*, $2n$ respostes pel filtre *DOG2* i $2np$ respostes pels filtres orientats *DOOG2*. Tenint en compte aquest agrupament de respostes i la forma de l'expressió de la funció T_i podem estendre les taules 3.2 i 3.3 a la taula 4.1, on es mostren els factors en funció de l'agrupament que hem fet.

Tenint en compte aquesta taula podem deduir $2 + p$ funcions T_i diferents

$$T_i = \begin{cases} T_{DOG1} & \forall i \in \{1, \dots, 2n\} \\ T_{DOG2} & \forall i \in \{2n + 1, \dots, 4n\} \\ T_{DOOG2\theta_k} & \forall i \in \{4n + 2n(k-1) + 1, \dots, 4n + 2nk\}, \forall k \in \{1, \dots, p\} \end{cases} \quad (4.10)$$

on T_{DOG1} , T_{DOG2} i $T_{DOOG2\theta_k}$, es poden expressar com

$$T_{DOG1} = \max_j \{ \max_{(x,y) \in I_{j1}(x_0,y_0)} (\alpha_{j1} R_j(x,y)) \} \quad (4.11)$$

$$T_{DOG2} = \max_j \{ \max_{(x,y) \in I_{j(2n+1)}(x_0,y_0)} (\alpha_{j(2n+1)} R_j(x,y)) \} \quad (4.12)$$

$$T_{DOOG2\theta_k} = \max_j \{ \max_{(x,y) \in I_{j(4n+2n(k-1)+1)}(x_0,y_0)} (\alpha_{j(4n+2n(k-1)+1)} R_j(x,y)) \} \quad (4.13)$$

La substracció d'aquests umbrals a les respostes seguit de la postinhibició es presenta a la següent expressió, que s'ha de realitzar per a cada canal del procés.

$$PIR_i(x_0, y_0) = \max_{x,y \in S_i(x_0,y_0)} \frac{1}{1 - \alpha_{ii}} [R_i(x,y) - T_i(x,y)]^+ \quad (4.14)$$

on $S_i(x_0, y_0)$ és un veinatge de (x_0, y_0) que no ha estat especificat en el model però que per obtenir bons resultats hauria de funcionar d'acord amb la densitat de cada característica en cada canal.

En aquests dos darrers passos tenim que la complexitat de càlcul és molt gran, s'han de calcular $2 + p$ funcions T_i , una per a cada agrupament de canals. El càlcul de cada una suposa aplicar una operació de màxim local a cada punt de cada canal. Per altra banda, la postinhibició suposa una resta per a cada canal seguit d'una operació de màxim local amb un radi de veinatge considerable si tenim en compte que volem assegurar que sigui tan gran com la densitat del textó.

Donat un radi de veinatge de R píxels i unes dimensions de la imatge de $N \times N$ tindrem que el cost d'una operació de màxim suposa $N \times N \times \pi R^2$ comparacions. En el següent apartat donem una millora a aquest pas de l'algorisme.

4.2 Introducció d'operacions morfològiques a l'algorisme

El passos més costosos de l'algorisme són els que es basen en operacions de màxim en cada punt de la imatge. Per millorar aquests passos tindrem en compte la definició donada a 2.2.1 de dilatació morfològica per imatges amb nivells de grisos. Aquesta operació pel cas d'elements estructurants plans no és res més que una operació de màxim local a l'entorn definit per l'element estructurant. Per tant, es pot solucionar el pas del càlcul del màxim aplicant l'algorisme 4.1 de càlcul de la dilatació.

La complexitat de l'algorisme de dilatació, en general ve donada pel nombre de suprems que depén del nombre de punts que té l'element estructurant.

En aquest cas els autors parlen de radis de veinatge, per tant hem de definir dilatacions morfològiques amb elements estructurants isotròpics que amb els radis donats a la taula

Algorisme 4.1 *Dilatació morfològica de la imatge i amb l'element estructurant B amb centre c i radi r*

1. $i_r = i$
2. Per a (cada punt x de l'element B) fer
 - (a) $i_t =$ Translació de la imatge i_r amb el vector $c\vec{x}$
 - (b) $i_r =$ Suprem de la imatge i i la imatge i_t

que presenta el model A d'inhibició.

La complexitat d'aplicar operacions morfològiques amb elements estructurants rodons ve definida pel nombre de translacions que suposa l'element estructurant. Per tant serà de l'ordre de R^2 , si R representa el radi de l'element estructurant.

Per a millorar el cost computacional d'aquestes operacions morfològiques ens podem basar en la descomposició dels elements estructurants. Per a poder fer això hem definit un mètode que permet la descomposició d'elements estructurants circulars. En els següents apartats veurem el desenvolupament d'aquest mètode. Primer de tot introduïrem el problema de la descomposició general d'elements estructurants, i aleshores passarem al mètode que proposem per a la descomposició de discs [111, 116, 117]. Finalment farem un anàlisi dels resultats en els que veurem que amb el mètode desenvolupat es possible passar d'una complexitat d'ordre quadràtic a una complexitat lineal.

4.2.1 Millora de les operacions morfològiques basada en la descomposició d'elements estructurants

Com ja hem vist a les expressions 2.13 i 2.14, les operacions d'erosió i de dilatació compleixen la següent propietat:

$$A \ominus B = (((((A \ominus B^1) \ominus B^2) \dots) \ominus B^n) \dots) \ominus B^n \quad (4.15)$$

$$A \oplus B = (((((A \oplus B^1) \oplus B^2) \dots) \oplus B^n) \dots) \oplus B^n \quad (4.16)$$

$$\text{on } B = B^1 \oplus B^2 \oplus \dots \oplus B^n$$

Aquesta propietat permet calcular erosions i dilatacions amb elements estructurants grans a partir d'anar aplicant recursivament l'operació per elements estructurants més petits. Les avantatges d'aquesta descomposició són les següents:

- El cost que suposa calcular una dilatació (erosió) en quan a nombre de translacions es pot reduir fent servir una bona descomposició.
- La descomposició fa que l'operació es pugui paral·lelitzar més fàcilment [72].

Diverses arquitectures de màquines han dirigit els criteris òptims per a la descomposició:

Arquitectura seqüencial La complexitat d'una operació morfològica ve donada pel nombre de translacions que suposa fer l'operació morfològica sobre el total dels elements estructurants de la descomposició.

Arquitectura Pipeline En aquesta arquitectura paral·lela la complexitat en general ve definida pel nombre d'elements estructurats en que que es fa la descomposició, independentment del nombre de translacions que suposi cada element.

Arquitectura vectorial En general en aquesta arquitectura, igual que en màquines seqüencials, depèn del nombre total de translacions que requereix l'operació morfològica sobre tots els elements estructurants. Si els processadors estan connectats a quatre, aleshores la complexitat depèn de la suma de les distàncies de tots els punts de l'element estructurant en mètrica *city block*.

La descomposició d'elements estructurants ha estat estudiada en varis treballs [127],[126]. En alguns casos la descomposició ve guiada per la forma de l'element estructurant. En el treball de J. Xu [123] es considera el problema de la descomposició pel cas d'elements estructurants convexos. Una aproximació recent desenvolupada per H. Park and R.T. Chin [73], descomposa elements estructurants de qualsevol forma. En el següent apartat fem un resum d'aquest treball com a base per al desenvolupament de la nostra aproximació que té com a objectiu donar la definició de discs òptims descomponibles en elements 3×3 .

4.2.2 Descomposició general d'elements estructurants

El mètode per a la descomposició d'elements estructurants de qualsevol forma presentat a [73], permet la descomposició d'elements estructurants binaris simplement connectats en

elements 3×3 . Les fronteres de l'element es representen en codis de cadenes de Freeman. Aquests codis contenen tota la informació suficient de la forma de l'element per a obtenir una bona descomposició.

S'han de distingir dos tipus de fronteres: *convexes* i *còncaves*. Un element estructurant convex es defineix com

Definició 4.1 *Un element estructurant binari S és convex si és el resultat de la intersecció de semiplans orientats en direccions que són múltiples de 45° . El codi de cadena de la frontera convexa d'un element es representa com*

$$S = 0^{x_0}1^{x_1}2^{x_2}3^{x_3}4^{x_4}5^{x_5}6^{x_6}7^{x_7} = [i^{x_i}]_{i=0,\dots,7} \quad (4.17)$$

on $x_i > 0$ significa que el codi de cadena i es repeteix x_i vegades, i $x_i = 0$ quan la direcció i no apareix a la frontera de l'element. Un element estructurant convex és vàlid si compleix les següents condicions

$$x_7 + x_0 + x_1 = x_3 + x_4 + x_5 \quad (4.18)$$

$$x_1 + x_2 + x_3 = x_5 + x_6 + x_7 \quad (4.19)$$

del codi de cadena de l'element S també es denotarà amb el vector següent $(x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7)$

A la figura 4.1.a es pot veure un element estructurant convex.

Per altra banda, les fronteres d'un element estructurant còncav es divideixen en dos tipus separats de frontera, còncava i convexa. Existeixen infinits tipus diferents de fronteres còncaves. Ara bé, si tenim en compte que totes les fronteres còncaves de l'element estructurant original han d'estar contigudes per algun element dels descomposts, i que volem una descomposició en elements de dimensió 3×3 , podem assegurar que només existeixen 28 possibles fronteres còncaves. Aquestes fronteres es poden veure a la figura 2 de [73].

Les fronteres còncaves estan dividides en cinc tipus diferents, depenent de la forma de la frontera, que denotarem amb Q_{T_i} i on T indica el tipus que pot ser un d'aquests cinc U, J, L, V, r , que es mostren a la figura 4.3, i indica la direcció de començament del codi.

Anomenarem elements estructurants còncavs als elements que compleixen la següent definició:

Definició 4.2 *Un element estructurat binari i connectat S és còncav si les seves fronteres*

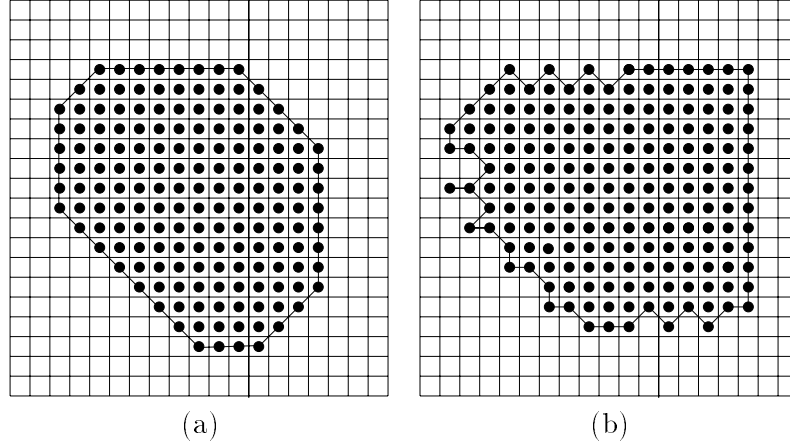


FIGURA 4.1: (a) Element estructurant convex amb codi de cadena $0^3 1^3 2^7 3^4 4^7 5^2 6^5 7^7$. (b) Element estructurant còncav amb codi de cadena $Q_{U_0} Q_{J_0} Q_{L_0} Q_{r_0} 0^2 Q_{V_1}^2 Q_{r_1} 2^{12} 4^6 Q_{V_5}^3 5^3 6$

es poden representar amb la següent expressió

$$S = [Q_{T_0}^{S_{T_0}}]_{T=U,J,L,r} 0^{S_0} [Q_{T_1}^{S_{T_1}}]_{T=J,V,r} 1^{S_1} \dots [Q_{T_7}^{S_{T_7}}]_{T=U,J,L,r} 7^{S_7} \quad (4.20)$$

on S_i denota la repetició de la direcció i a la fronteres convexes, i S_{T_i} denota la repetició de la frontera còncava Q_{T_i} . Assumim que els superíndexos han de complir les següents restriccions per a ser un codi de cadena vàlid

$$S_{U_i}(S_{i-1} + S_{i-2} + S_{i-3} + S_{r(i-1)} + S_{r(i-2)} + S_{V(i-1)}) = 0 \quad (4.21)$$

$$S_{J_i}(S_{i-1} + S_{i-2} + S_{r(i-1)}) = 0 \quad (4.22)$$

$$S_{L_i} S_{i-1} = 0 \quad (4.23)$$

$$S_{V_i} S_{i-1} = 0 \quad (4.24)$$

Si $S_{T_i} = 0 \forall T_i$, aleshores X és convex. En altre cas, S podrà ser o no descompost en elements 3×3 , depenent de si compleix les condicions de la proposició 4.2. A la figura 4.1.b es pot veure un element estructurant còncav.

Ara que hem definit els elements estructurants còncavs i convexos a partir de les seves fronteres, anem a presentar el problema de la descomposició. Abans però donarem una nova

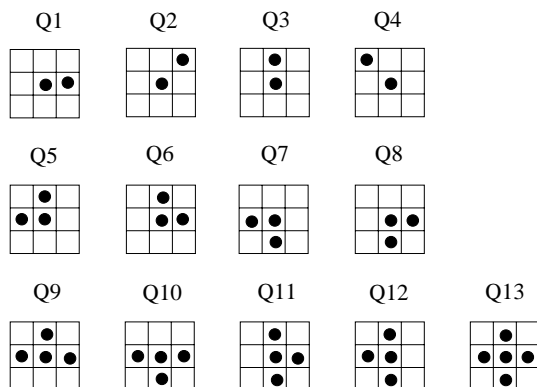


FIGURA 4.2: Conjunt d'elements estructurants que són factors primers de la descomposició d'elements convexes.

definició

Definició 4.3 *Un element A és un factor de S si existeix un element estructurant B , tal que $S = A \oplus B$. Direm que el factor A de S és un factor primer si i només si tot factor d' A és equivalent a A o és una imatge d'un sol píxel.*

Ara, introduïrem alguns criteris de descomposabilitat per elements estructurants convexos i còncaus. La descomposició d'elements convexos ha estat solucionat a [123]. Aquest treball es pot resumir amb la següent proposició

Proposició 4.1 *Tot element estructurant S és descomponible en una seqüència de dilatacions amb un conjunt de factors primers de la forma Q_1, Q_2, \dots, Q_{13} , que es donen a la figura 4.2.*

Una demostració constructiva d'aquesta proposició proporciona un algorisme de descomposició que es presenta a [123], i que ha estat millorat a [72] per a processadors vectorials connectats a 4. Aquest algorisme s'ha afegit com a darrer pas d'un algorisme més general per a descompondre elements estructurants còncaus.

Com hem dit abans, la definició 4.2 restringeix el tipus d'elements que poden ser descomposts en factors 3×3 . A la taula I de [73] s'especifica el conjunt de tots els possibles factors primers per a la descomposició de qualsevol element estructurant còncau. Aquest conjunt de factors primers és la base per a construir un algorisme general de descomposició. Abans

de donar una proposició general sobre la descomposabilitat, definirem algunes notacions prèvies.

Anomenem S a l'element còncau que volem descompondre, i $\{A^i\}$ al conjunt de tots els factors primers (de la taula I de [73]) que tenen fronteres còncaues que formen part de la frontera de S . Aleshores, si S és descomponible tindrem que

$$S = C^1 \oplus C^2 \oplus \dots \oplus C^m \oplus B \quad (4.25)$$

on $C^j \in \{A^i\}$ i B és un factor convex de S i per tant descomponible per la proposició 4.1.

A partir de la definició de frontera còncaua podem assegurar que la frontera de S està formada per

- (i) m fronteres còncaues distintes, V_1, V_2, \dots, V_m , on V_k és una de les Q_{T_i} 's tal que $S_{T_i} \neq 0$.
- (ii) l fronteres convexes distintes, d_1, d_2, \dots, d_l , on d_k representa una de les direccions i 's tal que $S_i \neq 0$.
- (iii) n factors primers còncaus (de la taula I de [73]) A^1, \dots, A^n que poden ser factors primers de S .

A partir d'aquestes components de la frontera podem construir les matrius Θ i Ω , de dimensions $m \times n$ i $l \times n$, respectivament.

$$[\Theta]_{ij} = \text{Nombre de } V_i \text{'s a } A^j \quad (4.26)$$

$$[\Omega]_{ij} = \text{Nombre de } d_i \text{'s a } A^j \quad (4.27)$$

Per altra banda es defineixen els vectors \mathbf{Y} i \mathbf{Z} de dimensions m i l respectivament

$$[\mathbf{Y}]_i = \text{Nombre de } V_i \text{'s a } S \quad (4.28)$$

$$[\mathbf{Z}]_i = \text{Nombre de } d_i \text{'s a } S \quad (4.29)$$

Finalment, definim el vector \mathbf{X} de m variables que poden tenir valors sencers no negatius. Seguidament donem una proposició que defineix la descomposabilitat de S

Proposició 4.2 *Un element estructurant S còncau és descomponible si existeix un \mathbf{X} tal que*

$$(i) \Theta \mathbf{X} = \mathbf{Y}$$

$$(ii) \Omega \mathbf{X} \leq \mathbf{Z}$$

(iii) $x_1 A^1 \oplus \dots \oplus x_n A^n \oplus B$ està simplement connectat i $x_i A^i$ indica x_i dilatacions amb A^i ,

(iv) B és un factor convex de S , definit pel codi de cadena següent (b_0, b_1, \dots, b_7) , i $b_i = [\mathbf{Z} - \Omega \mathbf{X}]_i, \forall i : 0 \leq i \leq 7$.

Per tant, si existeix un tal \mathbf{X} , la descomposició de S és

$$S = x_1 A^1 \oplus \dots \oplus x_n A^n \oplus B \quad (4.30)$$

Les definicions i proposicions que acabem de donar ens permeten definir un mètode per descompondre elements estructurants rodons en elements estructurants bàsics de dimensions 3×3 .

4.2.3 Mètode proposat per a la descomposició de discs

Com ja hem vist a l'anterior apartat tenim un mètode general per a la descomposició d'elements estructurants de qualsevol forma. El mètode es basa en representar la forma de l'element amb el seu codi de cadena. A partir d'aquesta expressió ja es pot aplicar la proposició 4.2.

Per tant, el primer problema que hem de solucionar ara és el de trobar el codi de cadena d'una aproximació d'un disc que ens permeti assegurar la seva descomposició en elements estructurants 3×3 . Un cop hagim definit la forma d'aquest disc definirem un algorisme que ens trobi la seva descomposició per a qualsevol radi.

Les aproximacions discretes de cercles ha estat estudiada en el camp dels gràfics per computador [25]. L'algorisme més eficient és el que genera iterativament el disc fent servir el criteri del punt mig. Aquest algorisme es basa en anar incrementant en cada pas una de les coordenades, l'altra coordenada s'incrementa o no depenent de la distància dels dos possibles següents punts al punt que està a la frontera del cercle continu de radi corresponent.

En general, els cercles generats per aquest algorisme presenten fronteres còncaues i convexes, que no poden ser contingudes pels elements estructurants bàsics de dimensions

3×3 que hem presentat a la secció anterior. Per tant, el cercle òptim no pot ser descompost en elements estructurants d'aquesta dimensió.

Per tant, per poder assegurar la descomposició haurem de definir la forma d'una aproximació discreta d'un cercle que permeti ser descompost. Si no considerem fronteres còncaues, aleshores poden garantir la descomposició (proposició 4.1). En aquest cas només podríem aproximar un disc per un polígon de 8 costats, en el que cada un dels costats aniria en una de les 8 direccions bàsiques d'una frontera convex. Ara bé, aquesta aproximació de disc no és gaire bona i pot ser millorada.

Anem doncs, a definir fronteres còncaues descomponibles que aproximïn discs. Com hem vist abans hi ha 28 fronteres còncaues possibles, que s'agrupen en cinc tipus. Les úniques fronteres còncaues que poden millorar l'aproximació del disc són les de tipus r (figura 4.3.a). La resta de fronteres (figura 4.3.b) introduirien importants errors en la forma del cercle.

Aleshores, podem assegurar que la frontera d'un element estructurant que només contingui fronteres còncaues de tipus r es pot expressar com

$$Q_{r0}^{x_0} 0^{x_1} Q_{r1}^{x_2} 1^{x_3} Q_{r2}^{x_4} 2^{x_5} Q_{r3}^{x_6} 3^{x_7} Q_{r4}^{x_8} 4^{x_9} Q_{r5}^{x_{10}} 5^{x_{11}} Q_{r6}^{x_{12}} 6^{x_{13}} Q_{r7}^{x_{14}} 7^{x_{15}} \quad (4.31)$$

que es pot considerar la frontera d'un polígon de 16 costats, sempre que $x_i \neq 0 \forall i$, vuit d'aquests costats venen donats per les direccions bàsiques, és a dir les direccions múltiples de 45° , i els altres vuit per les direccions de tipus r que denotem com $\{Q_{ri}^{x_j}\}$, $\forall i \in \{0, \dots, 7\}$ i que són aproximacions de segments amb pendents $-\frac{1}{2}, \frac{1}{2}, 2, -2, -\frac{1}{2}, \frac{1}{2}, 2, -2$, respectivament.

Per mantenir la isotropia d'un cercle imposarem la simetria de cada costat respecte del centre de l'element, això és

$$\begin{aligned} a &= x_0 = x_2 = x_4 = x_6 = x_8 = x_{10} = x_{12} = x_{14} \\ b &= x_1 = x_5 = x_9 = x_{13} \\ c &= x_3 = x_7 = x_{11} = x_{15} \end{aligned}$$

aleshores les fronteres del polígon de 16 costats que acabem de definir s'expressaran com

$$Q_{r0}^a 0^b Q_{r1}^a 1^c Q_{r2}^a 2^b Q_{r3}^a 3^c Q_{r4}^a 4^b Q_{r5}^a 5^c Q_{r6}^a 6^b Q_{r7}^a 7^c \quad (4.32)$$

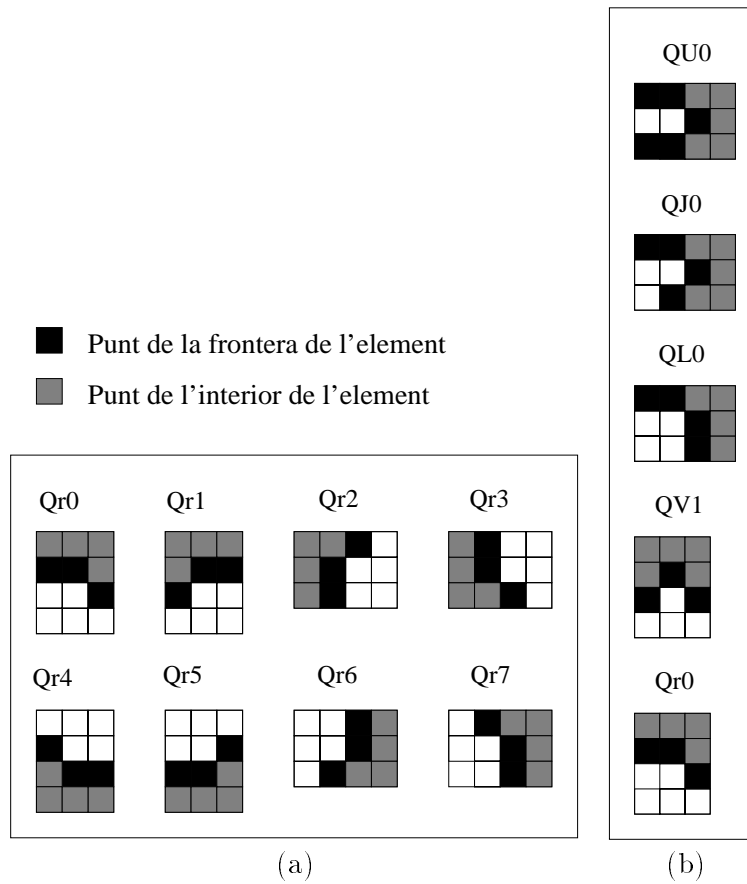


FIGURA 4.3: (a) Fronteres còncaves de tipus r , el conjunt de $\{Q_{ri}\}$ representa les 8 possibles direccions de començament en que es pot trobar la frontera. (b) Fronteres de tipus U , J , L , and V per una direcció de començament donada.

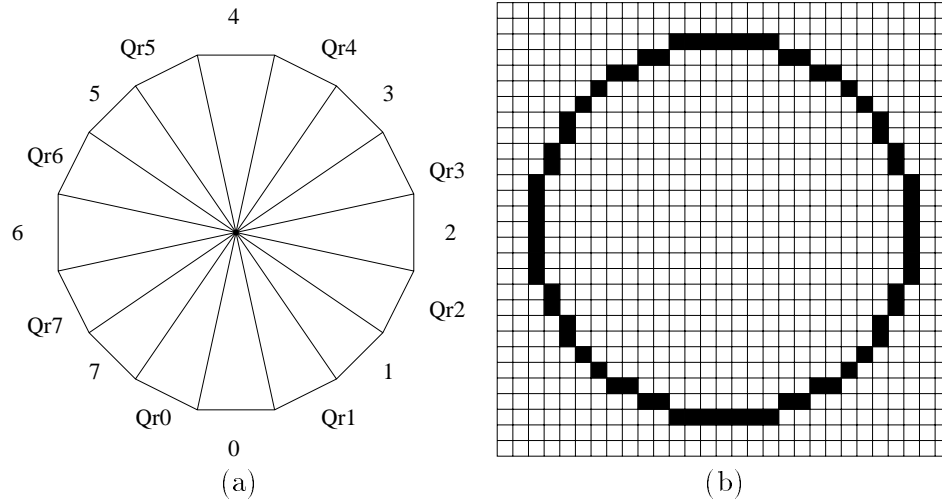


FIGURA 4.4: (a) Exemple d'un hexadecagon en el que indiquem les direccions i el tipus de frontera de cada costat. (b) Hexadecagon discret que aproxima un disc de radi 12. Els paràmetres del seu codi de cadena corresponent són $a = 2$, $b = 6$ i $c = 3$.

A la figura 4.4.a mostrem un hexadecagon com el que acabem de definir, per a cada costat indiquem la direcció corresponent expressada en el codi de cadena donat a 4.31. Per altra banda, a la figura 4.4.b donem una aproximació discreta d'un disc de radi 12 píxels que presenta les 16 direccions i la simetria que acabem de definir.

Per a un radi donat podem trobar diferents valors per a , b i c que defineixen polígons de 16 costats. Per poder obtenir discs que siguin òptims buscarem la combinació de valors per aquests paràmetres que minimitzi l'error d'aproximació entre el polígon i el cercle.

Discs descomponibles òptims En aquest apartat donem un algorisme per trobar els paràmetres que defineixen el polígon que aproxima un cercle amb mínim error quadràtic.

Primer de tot, definirem la funció d'error entre la frontera d'un disc de radi R i un polígon. Només considerarem la funció d'error per un quart del disc, donat que aquest presenta simetria respecte del centre. Així doncs, la funció d'error ve donada per

$$E(a, b, c) = \| \mathbf{F} - \hat{\mathbf{F}} \|^2 = \sum_{x=-(\frac{b}{2}+2a)}^{R-a-1} (F(x) - \hat{F}(x))^2 \quad (4.33)$$

on $F(x) = \sqrt{R^2 - x^2}$, i $\hat{F}(x)$ representa el valor de la coordenada y de la frontera

del polígon que volem construir. S'ha de tenir en compte que c pot ser expressat com $c = R - (\frac{b}{2} + 3a)$, i és per aixó que no apareix a l'expressió.

Aquesta funció E es pot dividir en tres parts diferents:

$$\begin{aligned} E(a, b, c) &= 2E_B(b') + 2E_A(a, b') + E_C(a, b') \\ &= 2 \sum_{x=1}^{b'} F(x)^2 + 2 \sum_{x=b'+1}^{b'+2a} (F(x) - k(x, b'))^2 + \sum_{x=b'+2a+1}^{R-a-1} (F(x) - (x - b' - a))^2 \end{aligned}$$

on $b' = \frac{b}{2}$, i

$$k(x, b') = \begin{cases} \frac{x-b'}{2} & \text{si } x - b' \text{ és parell} \\ \frac{x-b'+1}{2} & \text{si } x - b' \text{ és senar} \end{cases} \quad (4.34)$$

Per a obtenir els paràmetres a , b i c que minimitzen aquesta funció d'error s'ha construït un algorisme de propagació de restriccions. Les restriccions són les següents

$$\begin{aligned} b' &\in \{0, \dots, R\} \\ a &\in \{0, \dots, (R - b')/3\} \end{aligned}$$

L'algorisme es pot millorar si fem servir l'expressió recursiva dels errors parcials, per tant

$$E_B(b') = \sum_{x=1}^{b'} F(x)^2 = E_B(b' - 1) + F(b')^2 \quad (4.35)$$

$$\begin{aligned} E_A(a, b') &= \sum_{x=b'+1}^{b'+2a} (F(x) - k(x, b'))^2 \\ &= E_A(a - 1, b') + \end{aligned}$$

Algorisme 4.2 *Els valors de a , b i c que minimitzen la funció E es poden obtenir de la següent manera:*

Pas 1 Construir una taula que guardi els valors de $F(x)\forall x \in \{0, \dots, R\}$, on $F(x) = \sqrt{R^2 - x^2}$.

Construir una taula que guardi els valors de $E_B(b')\forall b' \in \{0, \dots, R\}$, calculats recursivament amb l'expressió 4.35.

Construir una taula que guardi els valors de $E_A(a, b')\forall (a, b') \in \{0, \dots, (R - b')/3\} \times \{0, \dots, R\}$, calculats recursivament amb l'expressió 4.36.

Pas 2 Per a cada $(a, b') \in \{0, \dots, (R - b')/3\} \times \{0, \dots, R\}$, calcular

$$E(a, b, c) = 2E_B(b') + 2E_A(a, b') + \sum_{x=b'+2a+1}^{R-a-1} (F(x) - (x - b' - a))^2 \quad (4.37)$$

usant els valors de les taules construïdes en els anteriors passos. Un parell és guardat si minimitza E per tot (a, b') .

Pas 3 Donat el parell (a, b') , que és el resultat del pas anterior, retornar a , $b = 2b'$ i $c = R - 3a - b$ com a solució.

$$\begin{aligned} & +(F(b' + 2a) - k(b' + 2a, b'))^2 + (F(b' + 2a - 1) - k(b' + 2a - 1, b'))^2 \\ = & E_A(a - 1, b') + (F(b' + 2a) - a)^2 + (F(b' + 2a - 1) - a)^2 \end{aligned} \quad (4.36)$$

d'aquesta manera definim l'algorisme 4.2.

A la taula 4.2 donem alguns resultats d'aplicar aquest algorisme. Els polígons que representen aquests resultats es mostren a la figura 4.6. A la mateixa taula donem l'error $E(a, b, c)$ per a un quadrant del polígon, i l'error mig \bar{E} , que representa l'error mig per punt. Finalment i també a la mateixa taula donem l'error mig per punt de l'aproximació discreta òptima d'un disc de radi donat generat per un algorisme basat en el criteri del punt mig.

A la gràfica 4.5.a podem veure la representació de la funció d'error pel disc òptim i pels

<i>Radius</i>	1	2	3	4	5	6	7	8	9	10	11	12	13
<i>a</i>	0	0	0	0	0	0	1	1	2	1	1	2	2
<i>b</i>	0	2	2	2	4	4	4	4	4	6	6	6	6
<i>c</i>	1	1	2	3	3	4	2	3	1	4	5	3	4
$E(a, b, c)$	0	0.14	0.11	0.66	0.36	0.61	0.53	0.63	1.22	0.81	1.31	1.00	1.44
\bar{E}_{Hex}	0	0.07	0.03	0.11	0.06	0.08	0.05	0.06	0.10	0.06	0.09	0.06	0.08
\bar{E}_{Mid}	0	0.07	0.03	0.11	0.06	0.08	0.05	0.06	0.10	0.06	0.09	0.06	0.07

<i>Radius</i>	14	15	16	17	18	19	20	21	22	23	24	25
<i>a</i>	2	2	3	3	3	3	3	4	4	3	4	4
<i>b</i>	8	8	8	8	8	10	10	10	10	12	12	12
<i>c</i>	4	5	3	4	5	5	6	4	5	8	6	7
$E(a, b, c)$	1.97	1.38	2.13	1.74	2.67	2.67	2.22	3.05	2.87	4.06	3.59	3.45
\bar{E}_{Hex}	0.10	0.06	0.09	0.07	0.11	0.10	0.08	0.10	0.09	0.13	0.10	0.10
\bar{E}_{Mid}	0.08	0.05	0.09	0.07	0.08	0.07	0.06	0.08	0.06	0.09	0.06	0.07

TAULA 4.2: Resultats de l'algorisme aplicat sobre un conjunt donat de radis. Per a cada radi donem els paràmetres a , b i c que retorna l'algorisme, i que defineixen el disc òptim descomponible amb fronteres còncaues 3×3 . $E(a, b, c)$ representa la funció total de l'error. L'error mig per punt de l'aproximació poligonal \bar{E}_{Hex} , i l'error mig per punt del cercle òptim generat per un algorisme basat en el criteri del punt mig \bar{E}_{Mid} .

polígons calculats pel nostre algorisme. Es pot veure que pels 25 primers radis l'aproximació poligonal es comporta pràcticament igual que les aproximacions òptimes. Evidentment, quan el radi augmenta, l'error també augmenta, ja que l'aproximació d'un cercle per un polígon de 16 costats sempre suposa errors, aquest efecte es pot veure a les figures 4.5.b i 4.5.c.

Encara que s'hagin introduït errors importants l'objectiu d'aquesta aproximació ha estat el de mantenir la propietat de ser descomponibles amb elements estructurants de dimensions 3×3 . Ara anem a demostrar que aquests discs són sempre descomponibles i calcularem la seva descomposició.

Descomposició de discs Com hem vist a la proposició 4.2, per a trobar la descomposició d'un element estructurant còncau hem de trobar la solució del sistema d'equacions que es planteja. Aleshores per a un disc donat S i definit per

$$S = Q_{r_0}^a 0^b Q_{r_1}^a 1^c Q_{r_2}^a 2^b Q_{r_3}^a 3^c Q_{r_4}^a 4^b Q_{r_5}^a 5^c Q_{r_6}^a 6^b Q_{r_7}^a 7^c \quad (4.38)$$

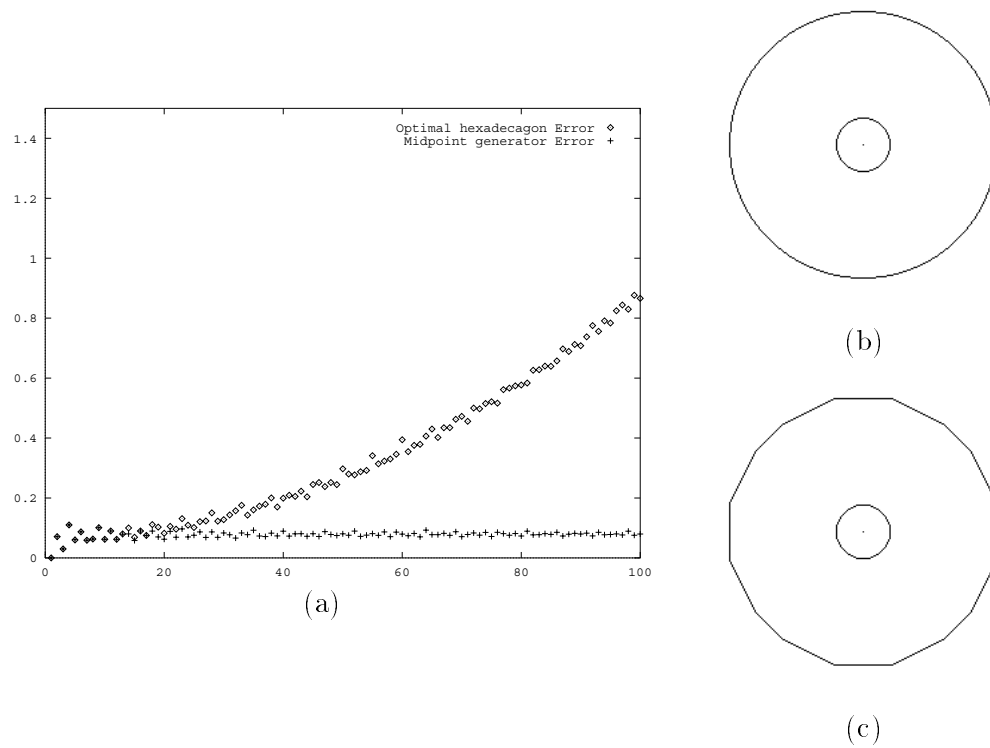


FIGURA 4.5: (a) Comparació gràfica entre l'error (en píxels) que presenten els polígons de 16 costats òptims i l'aproximació discreta òptima de cercles generats per un algorisme de punt mig. (b) Exemple de dos discs òptims generats per un algorisme de punt mig (radis de 20 i 100). (c) Exemple de dos hexadecagons òptims (radis de 20 i 100).

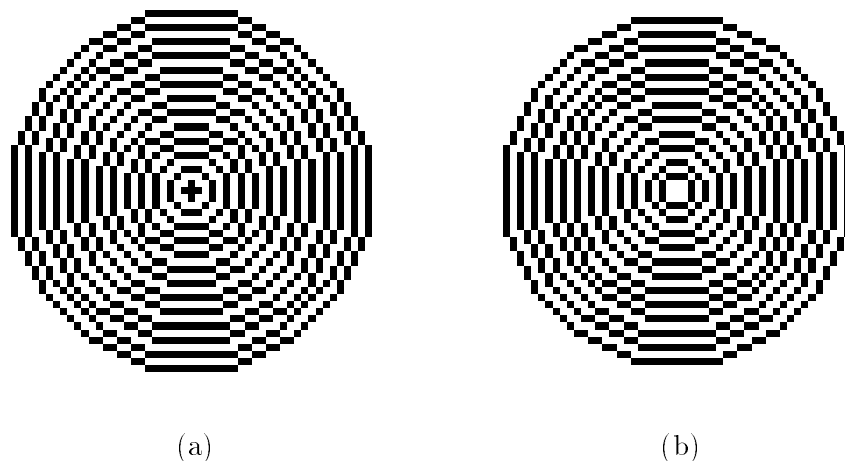


FIGURA 4.6: Hexadecagons que aproximen discs descomponibles amb mínim error quadràtic. (a) Discs amb radis de valor senar $R = 1, \dots, 25$. (b) Discs amb radis de valor parell $R = 2, \dots, 24$.

hem de construir el corresponent sistema d'equacions lineals. Aleshores, haurem de trobar la solució d'aquest sistema que ens asseguri que S és descomponible per a qualsevol radi.

Primer de tot, hem de considerar el conjunt de tots els possibles factors primers còncaus que poden permetre descompondre S . Aquests els agafarem de la taula I de [73], on seleccionarem el conjunt d'element estructurants que tenen fronteres que poden ser contingudes per a qualsevol hexadecagon S . Aquest conjunt d'elements seleccionats apareixen a la taula 4.7, on han estat agrupat per famílies de la mateixa forma.

Per tant, tots aquests elements representen tots els possibles factors primers que es poden usar per provar la descomposició dels discs definits. Els denotarem com $\{A_i\} \forall i \in \{1, \dots, 56\}$. A partir d'ells podem construir les matrius Θ and Ω , amb dimensions 56×8 , com s'especifica a la proposició 4.2.

També hem de construir els vectors \mathbf{Y} i \mathbf{Z} a partir de S , que vindran donats per

$$\mathbf{Y} = (a, a, a, a, a, a, a, a) \quad (4.39)$$

$$\mathbf{Z} = (b, c, b, c, b, c, b, c) \quad (4.40)$$

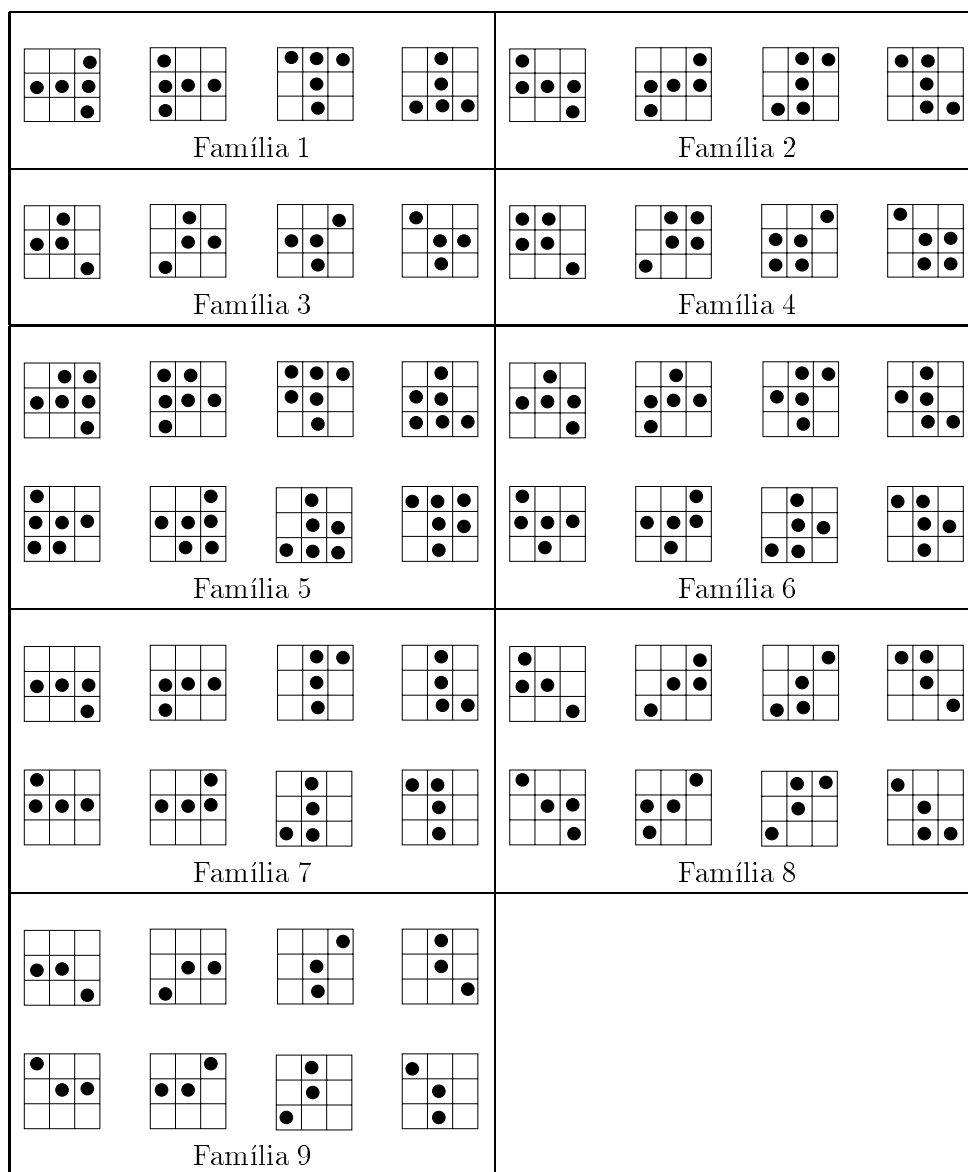


FIGURA 4.7: Conjunt d'elements estructurants còncaus amb fronteres contingudes a l'aproximació poligonal d'un disc. Cada família representa una forma d'element amb totes les possibles direccions de començament de la frontera

aleshores el sistema lineal a solucionar és el següent

$$\begin{aligned}
 \Theta \mathbf{X} &= \mathbf{Y} \\
 \mathbf{Z} - \mathbf{B} &= \Omega \mathbf{X} \\
 b_7 + b_0 + b_1 &= b_3 + b_4 + b_5 \\
 b_1 + b_2 + b_3 &= b_5 + b_6 + b_7
 \end{aligned}
 \tag{4.41}$$

on $\mathbf{X} = (x_1, \dots, x_{56})$ és el vector de variables. Els valors de la seva solució representaran el nombre de dilatacions per cada element estructurant que sigui factor primer còncau de S . El vector $\mathbf{B} = (b_0, b_1, b_2, b_3, b_4, b_5, b_6, b_7)$ és també un vector de variables que representarà el codi de cadena de l'element estructurant convex B . Les dues darreres equacions de 4.41 fan que la solució de les variables b_i representin un codi de cadena vàlid per a un element estructurant convex. Per tant, la solució d'aquest sistema donarà la següent descomposició

$$S = x_1 A_1 \oplus \dots \oplus x_{56} A_{56} \oplus B \tag{4.42}$$

El problema ve donat pel fet que 4.41 és un sistema indeterminat, amb 64 variables i 18 equacions i presentant per tant un gran nombre de possibles solucions.

Per evitar aquesta indeterminació hem considerat la descomposició per cada família per separat, construint nou matrius Θ i Ω , una per a cada família d'elements estructurants bàsics. Cada família per separat ens assegura que tota frontera còncaua de S serà continguda en un element estructurant bàsic de la família.

Les famílies d'elements estructurants 1, 2, 3 i 4 impliquen haver de solucionar sistemes d'equacions lineals amb 18 equacions i 12 variables. Les famílies 5, 6, 7, 8 i 9 construeixen sistemes formats per 18 equacions i 16 variables. En tots els casos el sistema és indeterminat, encara que per a cada un d'ells s'ha obtingut una única solució en termes dels paràmetres del polígon (la solució es pot veure a la taula 4.3).

4.2.4 Anàlisi dels resultats

Anem ara a analitzar les solucions que presentem a la taula 4.3. Primer de tot hem de veure que per totes les famílies es verifica que $x_i = a \forall i$, aleshores la descomposició tindrà sempre

1	$\mathbf{X} = (a, a, a, a)$	$\mathbf{B} = (b - 2a, c, b - 2a, c, b - 2a, c, b - 2a, c)$
2	$\mathbf{X} = (a, a, a, a)$	$\mathbf{B} = (b - 2a, c, b - 2a, c, b - 2a, c, b - 2a, c)$
3	$\mathbf{X} = (a, a, a, a)$	$\mathbf{B} = (b, c - a, b, c - a, b, c - a, b, c - a)$
4	$\mathbf{X} = (a, a, a, a)$	$\mathbf{B} = (b - 2a, c, b - 2a, c, b - 2a, c, b - 2a, c)$
5	$\mathbf{X} = (a, a, a, a, a, a, a, a)$	$\mathbf{B} = (b - 5a, c - 2a, b - 5a, c - 2a, b - 5a, c - 2a, b - 5a, c - 2a)$
6	$\mathbf{X} = (a, a, a, a, a, a, a, a)$	$\mathbf{B} = (b - 2a, c - 4a, b - 2a, c - 4a, b - 2a, c - 4a, b - 2a, c - 4a)$
7	$\mathbf{X} = (a, a, a, a, a, a, a, a)$	$\mathbf{B} = (b - 6a, c, b - 6a, c, b - 6a, c, b - 6a, c)$
8	$\mathbf{X} = (a, a, a, a, a, a, a, a)$	$\mathbf{B} = (b - 2a, c - 4a, b - 2a, c - 4a, b - 2a, c - 4a, b - 2a, c - 4a)$
9	$\mathbf{X} = (a, a, a, a, a, a, a, a)$	$\mathbf{B} = (b - 2a, c - 2a, b - 2a, c - 2a, b - 2a, c - 2a, b - 2a, c - 2a)$

TAULA 4.3: Resultats obtinguts de calcular la solució del sistema lineal que prova la descomposabilitat sobre cada família d'elements estructurants bàsics. La columna de l'esquerra identifica cada família. Els valors dels x_i 's representen el nombre de dilatacions amb cada element A_i de la família corresponent. Els valors dels b_i 's donen el codi de cadena de l'element estructurant convex resultant i del que tenim la descomposició assegurada.

la següent forma

$$S = aA_1 \oplus aA_2 \oplus \dots \oplus aA_n \oplus B \quad (4.43)$$

on els elements A_i depenen de quina família s'estigui fent servir per la descomposició (pot ser qualsevols de les que mostrem a la figura 4.7). Cal fer notar que hi ha famílies amb $n = 4$ i d'altres amb $n = 8$.

Per altra banda, hem de remarcar que cada solució verifica que $b_0 = b_4$, $b_1 = b_5$, $b_2 = b_6$ i $b_3 = b_7$. Per tant, d'acord amb l'algorisme 1 de [123] la descomposició de B és

$$B = b_0Q1 \oplus b_1Q2 \oplus b_0Q3 \oplus b_1Q4 \quad (4.44)$$

els elements Q_i 's es donen a la figura 4.8, i els valors dels b_i 's depenen de la família seleccionada. Aquests valors són sempre una funció dels paràmetres del disc a , b i c .

Per exemple, un disc de radi 16 es podria descompondre amb la família 1 com

$$S_{16} = 3A_1 \oplus 3A_2 \oplus 3A_3 \oplus 3A_4 \oplus 2Q1 \oplus 3Q2 \oplus 2Q3 \oplus 3Q4 \quad (4.45)$$

Per obtenir aquesta descomposició hem considerat el millor polígon de 16 costats que aproxima un cercle de radi 16. Aquests valors els trobem a la taula 4.2, on $a = 3$, $b = 8$ i $c = 3$, i els hem substituït a la solució donada per la família 1.

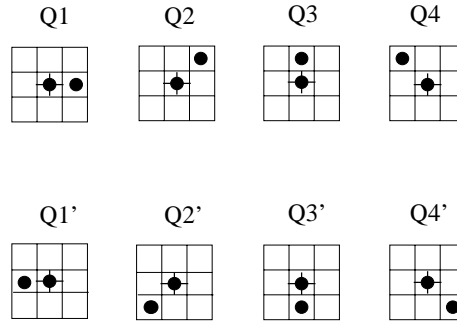


FIGURA 4.8: Conjunt de factors primers per descompondre elements estructurants convexos.

Aquesta descomposició pot introduir algunes translacions tal com ha estat exposat a [123]. Per evitar-les es pot canviar el centre dels elements estructurants bàsics i introduir translacions addicionals.

Si considerem les solucions obtingudes es pot veure que cada una verifica que $b_0 = b_2 = b_4 = b_6$ i $b_1 = b_3 = b_5 = b_7$. Per tant, es poden evitar translacions innecessàries usant els elements Q'_i 's donats a la figura 4.8 i substituir les següents dilatacions a l'expressió 4.44:

$$b_0 Q1 \oplus b_0 Q3 = k Q1 \oplus k Q1' \oplus k Q3 \oplus k Q3' \quad (4.46)$$

si $b_0 = 2k$, i

$$b_0 Q1 \oplus b_0 Q3 = (k + 1) Q1 \oplus k Q1' \oplus (k + 1) Q3 \oplus k Q3' \oplus \{(1, 0)\} \quad (4.47)$$

si $b_0 = 2k + 1$, i

$$b_1 Q2 \oplus b_1 Q4 = p Q2 \oplus p Q2' \oplus p Q4 \oplus p Q4' \quad (4.48)$$

si $b_1 = 2p$, i finalment

$$b_1 Q2 \oplus b_1 Q4 = (p + 1) Q2 \oplus p Q2' \oplus (p + 1) Q4 \oplus p Q4' \oplus \{(0, 1)\} \quad (4.49)$$

si $b_1 = 2p + 1$. $\{(x, y)\}$ indica la translació que s'ha d'aplicar.

Ara que ja hem exposat tots els passos per obtenir una descomposició de discs, anem a analitzar com és la descomposició en funció de la família que es faci servir. A partir

Família	Restriccions	λ
1, 2, 4	$b \geq 2a \quad c \geq 0$	$4(3a + b + c)$
3	$b \geq 0 \quad c \geq a$	$4(3a + b + c)$
5	$b \geq 5a \quad c \geq 2a$	$4(5a + b + c)$
6	$b \geq 2a \quad c \geq 4a$	$4(4a + b + c)$
7	$b \geq 6a \quad c \geq 0$	$4(2a + b + c)$
8	$b \geq 2a \quad c \geq 4a$	$4(2a + b + c)$
9	$b \geq 2a \quad c \geq 2a$	$4(2a + b + c)$

TAULA 4.4: Restriccions i cost computacional derivats de fer la descomposició amb cada una de les famílies.

d'aquesta anàlisi intentarem donar algunes preferències sobre quines famílies poden ser millors en algun sentit donat.

Per definició cada solució ha de complir que $x_i \geq 0$ and $b_i \geq 0$, per a tot i . Per tant, primer considerarem quines restriccions introdueixen aquestes condicions per a que sigui possible la descomposició amb cada família. Com es pot veure pels resultats de la taula 4.3, la condició es compleix per tot els x_i 's, en canvi per les solucions dels b_i 's s'han d'afegir algunes restriccions que exposem a la taula 4.4.

Per altra banda, també analitzarem la complexitat computacional de cada solució. Donem el número total de punts en tots els elements estructurants de la descomposició com una mesura d'avaluació del cost. Denotarem aquest número amb l'expressió λ i el seu valor es presenta també a la taula 4.4.

A partir de les restriccions i dels costos per a cada família, podem exposar les següents conclusions. El millors costos els donen les famílies 7, 8 i 9, però al mateix temps aquestes imposen restriccions molt fortes.

Les aproximacions de discs que s'han calculat i que es donen a la taula 4.2, no compleixen les restriccions $b \geq 5a$, $c \geq 4a$ i $c \geq 2a$. Per tant les famílies 3, 5, 6, 7, 8 i 9 no asseguren la capacitat de descompondre tots els discs.

Només les famílies 1, 2 i 4 imposen restriccions que són satisfetes per les aproximacions de discs que s'han fet. S'ha provat que aquestes restriccions es compleixen pels discs aproximats que van des de radi 1 fins a 500. Encara que no s'ha demostrat la capacitat de descomposició global, creiem que aquest resultat es pot considerar com bo ja que a la majoria d'aplicacions d'anàlisi d'imatges radis més grans no s'utilitzen freqüentment.

Algorisme 4.3 *Descomposició d'un disc de radi R*

Pas 1 Aplicar l'algorisme desenvolupat per calcular els valors de a , b i c que troben l'hexadecagon de mínim error quadràtic de radi R .

Pas 2 Trobar la descomposició de S com

$$S = aA_1 \oplus aA_2 \oplus aA_3 \oplus aA_4 \oplus (b - 2a)Q1 \oplus cQ2 \oplus (b - 2a)Q3 \oplus cQ4 \quad (4.50)$$

on els elements A_i 's corresponen a qualsevol de les famílies de la figura 4.7, i els Q_i 's a elements de la figura 4.8.

En resum, nosaltres optem per seleccionar les famílies 1, 2 i 4 per separat, com a conjunts de factors primers que ens donen la descomposició dels hexadecàgons òptims amb un cost computacional de $4(3a + b + c)$, si considerem el número de translacions necessàries. Aquest valor està linealment relacionat amb el valor del radi de l'element estructurant que volem descompondre, per tant hem passat a un complexitat lineal respecte del radi.

L'algorisme de descomposició de discs en basa en dos passos principals que es poden veure a l'algorisme 4.3.

4.2.5 Extensió de la descomposició a elements bàsics de majors dimensions

Depenent del tipus d'arquitectura en la que volguem implementar l'algorisme potser serà interessant la descomposició en elements estructurants bàsics de mides més grans, ja que en certes arquitectures una dilatació amb un element 3×3 pot tenir el mateix cost que amb un element 8×8 . Per tant, seria interessant estendre els resultats del treball anterior a dimensions més grans. Anem a veure els passos que s'haurien de fer per poder fer aquesta extensió.

Primer de tot hauríem de definir les fronteres còncaves que podria tenir l'element estructurant bàsic que descomposi un disc. Aquestes fronteres serien extensions de les fronteres 3×3 que hem usat. Podem veure una seqüència d'aquestes a la figura 4.9.

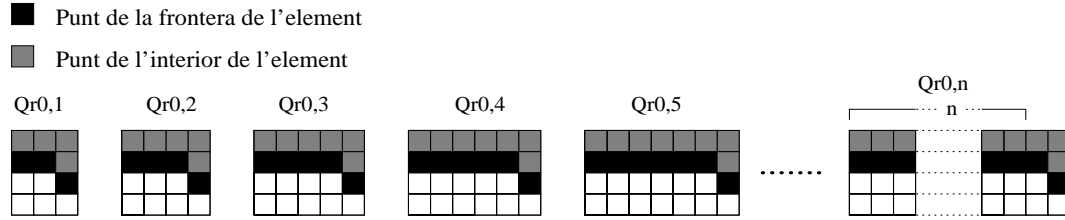


FIGURA 4.9: Fronteres còncaves de varies dimensions i que poden formar part de la frontera d'un disc.

Aleshores el codi de cadena del disc vindria donat per la següent expressió:

$$Q_{r0,1}^{m_1}, \dots, Q_{r0,n}^{m_n} 0^b Q_{r1,1}^{m_1}, \dots, Q_{r1,n}^{m_n} 1^c \dots 7^c \tag{4.51}$$

i que per conservar la simetria tindriem que la frontera

$$0^b Q_{r1,1}^{a_1}, \dots, Q_{r1,n}^{a_n} 1^c \tag{4.52}$$

es repetiria 8 vegades formant tota la frontera del disc. Així doncs trobar el disc òptim suposaria trobar els valors dels paràmetres

$$\{a_1, \dots, a_n, b, c\} \tag{4.53}$$

extenent l'algorisme 4.2.

Per altra banda a l'hora de fer la descomposició podríem estendre la família 1 d'elements estructurants bàsics, obtenint els elements de la figura 4.10. I caldria resoldre el sistema d'equacions 4.41 construït a partir del codi de cadena del disc i dels codis de cadena dels nous elements estructurants definits.

4.3 Algorisme proposat

L'algorisme 4.4 és l'algorisme que proposem pel càlcul de la representació, en el que hem introduït les operacions morfològiques de dilatació basades en la descomposició dels elements estructurants circulars que hem vist a l'apartat anterior.

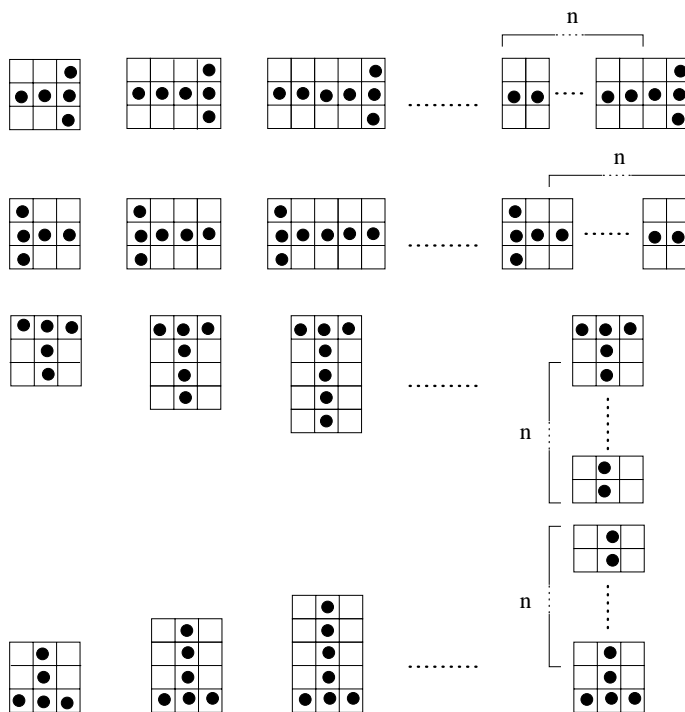


FIGURA 4.10: Elements estructurants bàsics de varies dimensions per a la descomposició de discs de qualsevol radi.

Algorisme 4.4 Algorisme pel càlcul de la representació r per la imatge t de dimensions $N \times N$. Està definit per 12 escales diferents $\{\sigma_1, \dots, \sigma_{12}\}$, i per 6 orientacions del filtre DOOG2, $\{\theta_1, \dots, \theta_6\}$:

Pas 1 Filtratge lineal

1. $Ft = FFT(t)$
2. Per $k = 1$ fins a $k = 12$ fer $R_k = FFT^{-1}(Ft \cdot FFT(DOG1_{\sigma_k}))$
3. Per $k = 1$ fins a $k = 12$ fer $R_{k+12} = FFT^{-1}(Ft \cdot FFT(DOG2_{\sigma_k}))$
4. Per $j = 1$ fins a $j = 6$ fer
 - (a) Per $k = 1$ fins a $k = 12$ fer $R_{kj+24} = FFT^{-1}(Ft \cdot FFT(DOOG2_{\sigma_k, \theta_j}))$

Pas 2 Rectificació de mitja ona

1. Per $i = 1$ fins a $i = 191$ fer (Pas = 2)
 - (a) $R_i^+ = \max\{R_i, 0\}$
 - (b) $R_{i+1}^- = \max\{R_i, 0\}$

Pas 3 Càlcul de l'umbral d'eliminació de respostes mal sintonitzades

1. Definim les taules $\alpha = (\alpha_{ij})$ i $I = (I_{ij})$ que mostrem a la taula 4.5.
2. Per $i = 1$ fins a $i = 192$ fer

$$T_{DOG1} = \max\{T_{DOG1}, (\alpha[1, j]R_i \oplus C_{I[1, i] \cdot \sigma_i})\}$$
3. Per $i = 1$ fins a $i = 192$ fer

$$T_{DOG2} = \max\{T_{DOG2}, (\alpha[2, j]R_i \oplus C_{I[2, i] \cdot \sigma_i})\}$$
4. Per $j = 1$ fins a $j = 6$ fer
 - (a) Per $i = 1$ fins a $i = 48$ fer

$$T_{DOOG2\theta_j} = \max\{T_{DOOG2\theta_j}, (\alpha[j, i] \cdot R_i \oplus C_{I[j, i] \cdot \sigma_i})\}$$
 - (b) Per $i = 48 + 24(j - 1) + 1$ fins a $i = 48 + 24(j - 1) + 24$ fer

$$T_{DOOG2\theta_j} = \max\{T_{DOOG2\theta_j}, (\alpha[j, i] \cdot R_i \oplus C_{I[j, i] \cdot \sigma_i})\}$$

Pas 4 Post-inhibició no lineal

1. Per $i = 1$ fins a $i = 24$ fer $PIR_i = S_i \oplus \frac{1}{1 - \alpha[1, 1]} \max\{0, R_i - T_{DOG1}\}$
2. Per $i = 25$ fins a $i = 48$ fer $PIR_i = S_i \oplus \frac{1}{1 - \alpha[2, 2]} \max\{0, R_i - T_{DOG2}\}$
3. Per $j = 1$ fins a $j = 6$ fer
 - (a) Per $i = 49 + 24(j - 1)$ fins a $i = 49 + 24j$ fer

$$PIR_i = S_i \oplus \frac{1}{1 - \alpha[j, j]} \max\{0, R_i - T_{DOOG2\theta_j}\}$$

Pas 5 Càlcul de la representació $r(t)$

1. Per $i = 1$ fins a $i = 192$ fer

$$r_i(t) = \sum_{(x, y) \in \{0, \dots, N\}^2} \frac{PIR_i(x, y)}{N^2}$$

on $C_{I[j, i]}$ i S_i representen elements estructurants circulars de radi $I[j, i]$ i indefinit respectivament. Les dilatacions amb aquests elements es poden fer tenint en compte la descomposició de l'algorisme 4.3.

$$\alpha = \begin{pmatrix} 0.2 & 0.45 & 0.15 & 0.15 & 0.15 & 0.15 & 0.15 & 0.15 \\ 0.45 & 0.25 & 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.15 & 0.2 & 0.65 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.15 & 0.2 & 0.0 & 0.65 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.15 & 0.2 & 0.0 & 0.0 & 0.65 & 0.0 & 0.0 & 0.0 \\ 0.15 & 0.2 & 0.0 & 0.0 & 0.0 & 0.65 & 0.0 & 0.0 \\ 0.15 & 0.2 & 0.0 & 0.0 & 0.0 & 0.0 & 0.65 & 0.0 \\ 0.15 & 0.2 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.65 \end{pmatrix}$$

$$I = \begin{pmatrix} 2 & 1.5 & 1.25 & 1.25 & 1.25 & 1.25 & 1.25 & 1.25 \\ 2 & 1.5 & 1.25 & 1.25 & 1.25 & 1.25 & 1.25 & 1.25 \\ 2 & 1.5 & 1.25 & - & - & - & - & - \\ 2 & 1.5 & - & 1.25 & - & - & - & - \\ 2 & 1.5 & - & - & 1.25 & - & - & - \\ 2 & 1.5 & - & - & - & 1.25 & - & - \\ 2 & 1.5 & - & - & - & - & 1.25 & - \\ 2 & 1.5 & - & - & - & - & - & 1.25 \end{pmatrix}$$

TAULA 4.5: Taules α i I de valors per la inhibició, contruïdes per l'algorisme 4.4 a partir de la taula 4.1.

CAPÍTOL 5

EXPLORACIÓ DE LA REPRESENTACIÓ SOBRE UN ESPAI PARAMÈTRIC DE TEXTURES

En el capítol anterior hem definit una representació de textures que hem derivat d'un model computacional basat en filtres detectors de característiques. La representació ens ha definit un espai de representació on les distàncies entre punts representen distàncies entre textures o dissimilituds. Aquesta representació presenta el problema de tenir una dimensió molt alta que fa que sigui difícil obtenir una interpretació. Amb l'objectiu d'explorar aquesta representació introduïm el mètode de l'escalament multidimensional que permet construir un espai de baixa dimensió conservant les dissimilituds entre punts de l'espai de representació de partida. Per a que l'espai resultant pugui ser analitzat en termes de paràmetres perceptualment interpretables s'ha definit un espai paramètric de generació d'imatges, de manera que a partir de les variacions sobre les imatges d'entrada es puguin identificar les dimensions de l'espai de representació.

5.1 Escalament multidimensional

Els mètodes d'escalament multidimensional (MDS)¹ han estat desenvolupats com eines d'anàlisi i visualització de dades ([124, 44, 66]). L'objectiu és que configuracions de punts en un espai de dimensió m es puguin veure a un espai de dimensió k on $k < m$, de manera que les distàncies entre punts es conservin en algun sentit. Una interessant valoració del mètode i dels programaris que l'implementen es pot trobar a [90].

Partint de la matriu de distàncies entre punts el MDS construeix la configuració de punts que compleixen les distàncies donades a la matriu. Aquesta configuració es calcula

¹Usarem la simplificació del nom en anglès MDS (*Multidimensional Scaling*) que és el nom habitual a la major part de bibliografia.

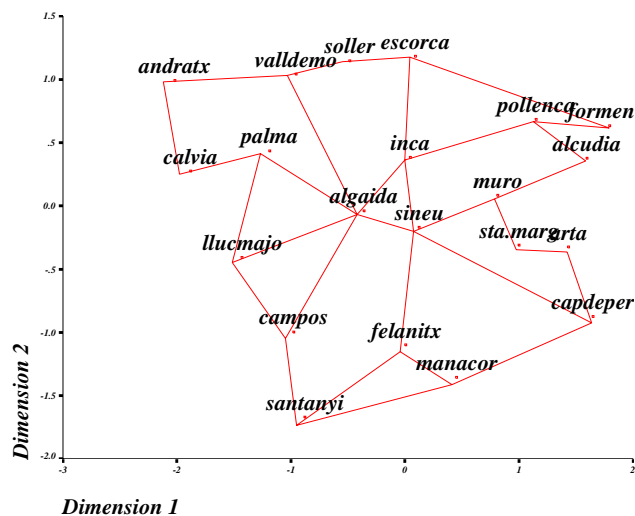


FIGURA 5.1: Configuració d'un conjunt de ciutats a partir de les seves distàncies en línia recta.

per a un espai de dimensions predeterminades. A la figura 5.1 es pot veure un exemple del resultat de reconstruir un mapa d'una illa mediterrània a partir de la seva matriu de distàncies entre ciutats. La configuració resultat de punts pot venir donada amb qualsevol rotació i desplaçament.

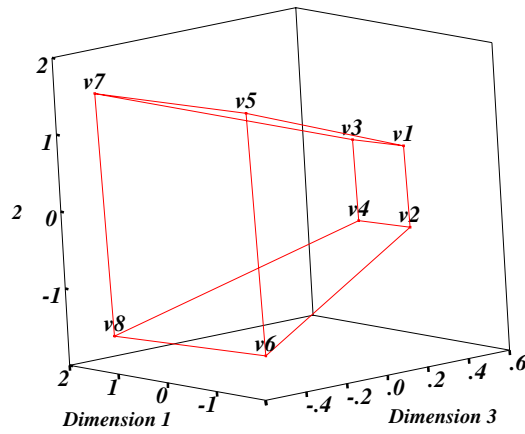
S'ha de fer notar que les distàncies pot ser que no representin una distància euclidiana real entre punts en un espai determinat, o que estiguin subjectes a determinats errors, aleshores el MDS ha de donar una configuració que sigui propera a la configuració real. A la figura 5.2.b podem veure el resultat d'aplicar el MDS sobre una matriu de distàncies entre vertèxos d'una figura en un espai de 3 dimensions i la seva configuració a l'espai de 2 dimensions.

Abans de passar a explicar els diferents mètodes que troben les solucions del MDS, anem a definir més detalladament l'objectiu del mètode.

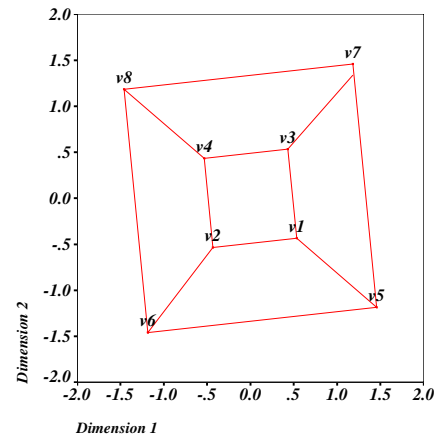
Encara que parlem de matrius de distàncies, aquestes no tenen perquè estar basades en una distància euclidiana sinó que poden representar qualsevol dissimilitud calculada.

Definició 5.1 *Direm que \mathbf{D} de dimensions $n \times n$ és una matriu de distàncies si és simètrica i compleix*

$$\mathbf{D} = \begin{pmatrix} 0 & - & - & - & - & - & - & - \\ 1 & 0 & - & - & - & - & - & - \\ 1 & 1.414 & 0 & - & - & - & - & - \\ 1.414 & 1 & 1 & 0 & - & - & - & - \\ 1.732 & 2.449 & 2.449 & 3 & 0 & - & - & - \\ 2.449 & 1.732 & 3 & 2.449 & 3 & 0 & - & - \\ 2.449 & 3 & 1.732 & 2.449 & 3 & 4.242 & 0 & - \\ 3 & 2.449 & 2.449 & 1.732 & 4.242 & 3 & 3 & 0 \end{pmatrix} \quad (5.1)$$



(a)



(b)

FIGURA 5.2: Matriu \mathbf{D} de distàncies entre 8 punts. (b) Configuració calculada pel MDS a partir de la matriu \mathbf{D} i escalat a un espai de 3 dimensions ($stress = 0.001$). (c) Configuració calculada pel MDS sobre la matriu \mathbf{D} i sobre un espai de 2 dimensions ($stress = 0.021$).

$$d_{ii} = 0 \forall i \in \{1, \dots, n\}, \quad d_{ij} \geq 0 \forall i, j \in \{1, \dots, n\} : i \neq j \quad (5.2)$$

A partir d'una matriu de distàncies \mathbf{D} que representen les distàncies entre punts p_1, \dots, p_n d'un espai de dimensió m i que no cal que siguin coneguts, l'objectiu del MDS és trobar els punts x_1, \dots, x_n d'un espai de k dimensions de manera que la distància entre aquests nous punts vingui donada per la matriu $\hat{\mathbf{D}}$ i que sigui semblant en algun sentit a la matriu \mathbf{D} .

La configuració de punts x_1, \dots, x_n resultant és indeterminada respecte de translacions, rotacions i reflexions, és a dir, per a qualsevol matriu ortogonal \mathbf{A} i qualsevol vector \mathbf{b} tals que

$$y_i = \mathbf{A}x_i + \mathbf{b} \quad i \in \{1, \dots, n\} \quad (5.3)$$

els punts y_1, \dots, y_n són també una solució del MDS.

Existeixen diversos mètodes per calcular el MDS que donen dos tipus de solucions. La solució clàssica parteix directament de les distàncies i intenta mantenir la similitud entre aquestes en els dos espais. I la solució no mètrica que parteix del rang d'ordre de les distàncies de la matriu

$$d_{i_1j_1} \leq d_{i_2j_2} \leq \dots \leq d_{i_tj_t} \quad \text{on } t = \frac{1}{2}n(n-1) \quad i \quad i_k \leq j_k \quad \forall (i_k, j_k) \in \{1, \dots, n\}^2 \quad (5.4)$$

i calculen els punts de manera que el rang d'ordre entre els nous punts es conservi sota una funció monòtona creixent f , és a dir

$$d_{i_1j_1} \leq d_{i_2j_2} \leq \dots \leq d_{i_tj_t} \iff f(d_{i_1j_1}) \leq f(d_{i_2j_2}) \leq \dots \leq f(d_{i_tj_t}) \quad (5.5)$$

Els mètodes que troben les primeres solucions que hem definit els anomenarem mètodes mètrics i als segons els anomenarem no mètrics.

5.1.1 Escalament multidimensional clàssic

El MDS clàssic és un mètode mètric i aplica l'algorisme que es deriva de suposar que tenim una matriu de distàncies \mathbf{D} que corresponen a n punts en un espai de p dimensions. Abans de donar l'algorisme cal veure alguns resultats previs

Definició 5.2 Una matriu de distàncies $\mathbf{D} = (d_{ij})$ és euclidiana si existeix una configuració de punts en un espai euclidià amb distàncies entre punts que venen donades per la matriu \mathbf{D} , això és, per un m donat existeixen punts $x_1, \dots, x_n \in \mathbb{R}^m$, tal que

$$d_{ij} = \sqrt{\sum_{a=1}^m (x_i^a - x_j^a)^2} \quad (5.6)$$

on $x_i = (x_i^1, \dots, x_i^m)$.

El problema que es planteja és el de com trobar la configuració de punts que tenen \mathbf{D} com a matriu de distàncies euclidianes. Abans necessitarem la següent notació. Per a qualsevol matriu de distàncies \mathbf{D} tenim que

$$\mathbf{A} = (a_{ij}) \quad : \quad a_{ij} = -\frac{1}{2}d_{rs}^2 \quad (5.7)$$

i definim

$$\mathbf{B} = \mathbf{H}\mathbf{A}\mathbf{H} \quad (5.8)$$

on \mathbf{H} és la matriu de concentració ($n \times n$).

Definició 5.3 La matriu $n \times n$ de concentració \mathbf{H} es defineix com $\mathbf{H} = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'$, on $\mathbf{1}$ representa el vector columna amb tots els elements igual a 1, i $\mathbf{1}'$ és el seu transposat. Aleshores \mathbf{H} compleix

- (i) $\mathbf{H}' = \mathbf{H}$, $\mathbf{H}^2 = \mathbf{H}$.
- (ii) $\mathbf{H}\mathbf{1} = \mathbf{0}$, $\mathbf{H}\mathbf{1}\mathbf{1}' = \mathbf{1}\mathbf{1}'\mathbf{H} = \mathbf{0}$.
- (iii) $\mathbf{H}x = x - \bar{x}\mathbf{1}$, on $\bar{x} = \frac{1}{n} \sum x_i$.
- (iv) $x'\mathbf{H}x = \frac{1}{n} \sum (x_i - \bar{x})^2$.

a partir d'aquestes definicions es pot donar el següent teorema:

Teorema 5.1 Sigui \mathbf{D} una matriu de distàncies, a partir d'aquesta definim \mathbf{B} com ho hem fet a l'expressió 5.1.1. Aleshores \mathbf{D} és euclidiana si i només si \mathbf{B} és positiva i semi-definida. Es donen els següents resultats en les dues direccions:

1. Si \mathbf{D} és la matriu de distàncies euclidianes entre punts, per la configuració $\mathbf{Z} = (z_1, \dots, z_n)'$, aleshores

$$b_{ij} = \sum_{r=1}^m (z_i^r - \bar{z}^r)(z_j^r - \bar{z}^r) \quad i, j \in \{1, \dots, n\} \quad (5.9)$$

que en forma the matriu es pot expressar com $\mathbf{B} = (\mathbf{HZ})(\mathbf{HZ})'$, per tant $\mathbf{B} \geq 0$. Cal notar que \mathbf{B} es pot interpretar com la matriu del producte escalar centrat?? per la configuració \mathbf{Z} .

2. En el sentit invers, si \mathbf{B} és positiva i semi-definida de rang k , aleshores es pot construir una configuració que correspon a \mathbf{B} en el següent sentit. Siguin $\lambda_1 > \dots > \lambda_k$ els valors propis de \mathbf{B} amb els corresponents vectors propis $\mathbf{X} = (x_1, \dots, x_k)$ normalitzats amb

$$x_i'x_i = \lambda_i \quad i \in \{1, \dots, k\} \quad (5.10)$$

Aleshores les distàncies entre els punts $x_i = (x_i^1, \dots, x_i^k)$ per tot $i \in \{1, \dots, k\}$ venen donades per la matriu \mathbf{D} . A més a més aquesta configuració de punts té com a centre de gravetat el vector $\mathbf{0}$, i \mathbf{B} representa la matriu del producte escalar per aquesta configuració.

La demostració i conseqüències d'aquest teorema es poden veure a [66]. L'ús d'aquest com a base pel MDS va ser presentada per Torgeson a [106] i ampliada a [34]. L'algorisme 5.1.1, que permet calcular el MDS mètric, és el resultat d'aquests treballs.

5.1.2 Escalament multidimensional no mètric

En el MDS mètric partíem d'una matriu de distàncies o dissimilituds δ_{ij} i a partir d'aquestes l'objectiu era el de trobar una configuració de punts de manera que les distàncies entre ells d_{ij} complissin que

$$d_{ij} = \delta_{ij} + e_{ij} \quad (5.11)$$

on e_{ij} representa l'error que pot ser degut al fet de que realment les dissimilituds δ_{ij} 's no puguin ser representades en un espai de dimensió k .

Algorisme 5.1 *Donada una matriu de distàncies \mathbf{D} per a la que suposem que existeix una configuració de punts a un espai de k dimensions. Es pot calcular aquesta configuració seguint els següents passos:*

1. *Construir la matriu $\mathbf{A} = (-\frac{1}{2}d_{ij}^2)$ a partir de $\mathbf{D} = (d_{ij})$*
2. *Obtenir la matriu \mathbf{B} a partir de \mathbf{A} .*
3. *Trobar els k valors propis $\lambda_1 > \dots > \lambda_k$ més grans de \mathbf{B} , amb els corresponents vectors propis $\mathbf{X} = (x_1, \dots, x_k)$ normalitzats (estem suposant que els k primers valors propis són tots positius).*
4. *Les coordenades dels punts que busquem venen donades per les columnes de la matriu \mathbf{X} .*

En alguns casos voler assegurar aquesta igualtat pot ser un condició massa restrictiva i possiblement sigui més realista intentar trobar una configuració de punts que les distàncies entre ells verifiquin

$$d_{ij} = f(\delta_{ij} + e_{ij}) \quad (5.12)$$

on f representi una funció desconeguda però que sigui monòtona creixent. D'aquesta manera s'està exigint que la configuració resultant mantingui només el rang d'ordre de les dissimilituds.

Les primeres aproximacions per construir configuracions basades en mantenir el rang d'ordre de les disimilituds fóren presentades per R. N. Shepard [95, 97]. El mateix autor a [98] proposa usar el MDS com un camí per descobrir l'estructura del que ell anomena l'espai perceptual. Aquest espai està format per les regions resultants que es comporten com les classes naturals. Amb aquest resultat un nou objecte és fàcilment reconegut o classificat a partir de la introducció a dins d'aquest espai. L'espai perceptual es construeix a partir d'un conjunt de dades de dissimilitud obtingudes a partir d'experiments psicològics.

Per aquesta aproximació partirem d'una matriu $\Delta = (\delta_{ij})$ de dimensions $n \times n$ on δ_{ij} representa una mesura de dissimilitud entre els punts p_i i p_j d'un espai m -dimensional. L'objectiu és trobar la configuració dels punts x_1, \dots, x_n a un espai k -dimensional on d_{ij}

representa les distàncies entre els punts x_i i x_j . Els valors de d_{ij} volem que estiguin monòtonament relacionats amb δ_{ij}

$$\delta_{ij} \leq \delta_{qp} \implies d_{ij} < d_{qp} \quad \forall (i, j) \in \{1, \dots, n\} \times \{1, \dots, n\} \quad (5.13)$$

Donada una configuració arbitrària de punts a \mathbb{R}^k amb distàncies entre ells \hat{d}_{ij} , el problema es soluciona aplicant un mètode iteratiu que minimitzi la següent expressió:

$$S = \sqrt{\frac{\sum_{i,j} (d_{ij} - \hat{d}_{ij})^2}{\sum_{i,j} \hat{d}_{ij}^2}} \quad (5.14)$$

els valors de d_{ij} que minimitzin el valor de S compliran la relació de variació monòtona del rang d'ordre amb δ_{ij} . L'algorisme de minimització d'aquesta funció va ser dissenyat per J.B. Kruskal, a [58], basant-se en un mètode de gradients. El valor final de S ens dona informació sobre com de bona és la configuració calculada. Aquest valor mínim l'anomenem l'*stress* i va ser definit a [59] com a mesura de la bondat de la configuració trobada. Així doncs, $S = 0$ representa una aproximació perfecta, és a dir, que la configuració trobada respecta exactament el rang d'ordre de les dissimilituds de partida.

Com en el cas del MDS mètric el valor de k és donat com a paràmetre de partida. Trobar el valor de k que més bé s'adapta a les dades és el problema de la dimensionalitat que ha estat tractat de manera molt interessant a [60].

5.1.3 Dimensionalitat de les dades

El problema de la dimensionalitat és el problema de trobar el número de dimensions o eixos que ha de tenir l'espai reconstruït pel MDS, valor que fins ara hem denotat amb la lletra k . En els mètodes que hem descrit fins ara sempre hem suposat que abans d'aplicar l'algorisme es dona el valor de k pel que volem reconstruir l'espai.

Abans d'entrar a tractar el problema s'han de distingir dos conceptes, un és el concepte de la *dimensionalitat real* o *correcta* que és la dimensió real de les dades en elles mateixes i per les que es podria demostrar estadísticament. L'altre concepte és el de *dimensionalitat apropiada*, que és aquella que permet l'anàlisi de les dades d'una manera més satisfactòria. Aquest darrer concepte és que fa que J.B. Kruskal [58] defineixi el problema de trobar la dimensionalitat com un problema de criteri científic.

Ara que hem definit el concepte de dimensionalitat correcta anem a veure els dos possibles camins que s'han seguit per determinar-la. Una aproximació estadística basada en la simulació computacional a partir de números aleatoris, és la que s'ha anomenat l'aproximació de Monte-Carlo. Per altra banda, s'ha proposat una aproximació intuïtiva que es basa amb els valors d'*stress* com a guia.

En qualsevol de les aproximacions cal tenir en compte una variable important que intervé en la decisió de la dimensionalitat, k , i que és el número de punts, I , de partida. Qualsevol interpretació d'*stress* pel càlcul de la dimensionalitat no serà sensible a la variable I si $I > 4k$.

S'han donat algunes regles basades en l'experiència per a l'anàlisi del valor d'*stress*. A partir de la funció S donada a 5.1.2 denotem

$$S_k = \min_{\hat{d}_{ij}} S(\hat{d}_{ij}) \quad (5.15)$$

que representa el valor d'*stress* mínim per a una configuració en k dimensions. Aleshores per calcular la dimensionalitat es pot calcular la seqüència de valors S_1, S_2, \dots , fins que el valor de S_k sigui baix, el valor de k corresponent serà la dimensionalitat correcta. A la següent taula es donen alguns valors d'*stress* i la seva interpretació verbal donada per J.B. Kruskal a [58]

<i>stress</i>	Interpretació
0.200	"pobre"
0.100	"just"
0.050	"bo"
0.025	"excel.lent"
0.000	"perfecte"

L'aproximació estadística es basa en valors d'*stress* ja calculats. Al treball de I. Spence i J. Graef [102, 103] es dona la taula de valors d'*stress* pel cas de 36 punts de partida i que reproduïm a la taula 5.1.2. Aquesta aproximació es basa en partir de configuracions conegudes de punts, aleshores aquestes es modifiquen per la introducció d'errors controlats i es calculen les seves matrius de dissimilituds. A partir d'aquestes matrius es calculen les solucions resultants d'aplicar el MDS per un conjunt de dimensions, 1, 2, 3, 4 i 5. És el que hem anomenat les dimensions escalades.

Usant aquests resultats la dimensionalitat d'un determinat conjunt de dades es pot

Dimensió generada m	Dimensió escalada k	Nivell d'error				
		0.0000	0.0625	0.1225	0.2500	∞
1	1	0.000	0.096	0.192	0.331	0.529
	2	0.000	0.089	0.164	0.257	0.349
	3	0.000	0.081	0.140	0.203	0.259
	4	0.000	0.074	0.121	0.167	0.205
	5	0.000	0.069	0.107	0.142	0.169
2	1	0.334	0.342	0.360	0.438	0.529
	2	0.000	0.078	0.144	0.255	0.349
	3	0.000	0.071	0.123	0.200	0.259
	4	0.000	0.064	0.106	0.162	0.205
	5	0.000	0.061	0.096	0.138	0.169
3	1	0.413	0.403	0.428	0.470	0.529
	2	0.197	0.196	0.228	0.289	0.349
	3	0.000	0.063	0.121	0.203	0.259
	4	0.000	0.057	0.104	0.163	0.205
	5	0.000	0.053	0.091	0.137	0.169
4	1	0.387	0.447	0.467	0.502	0.529
	2	0.204	0.233	0.255	0.312	0.349
	3	0.097	0.125	0.160	0.221	0.259
	4	0.000	0.060	0.111	0.172	0.205
	5	0.000	0.055	0.097	0.143	0.169

TAULA 5.1: Valors d'*stress* del MDS a partir de dades generades en un espai de dimensió m , amb diferents nivells d'error i reconstruïdes a un espai de dimensió k . Extretes del treball de I. Spence et al-74 pel cas de 36 punts de partida.

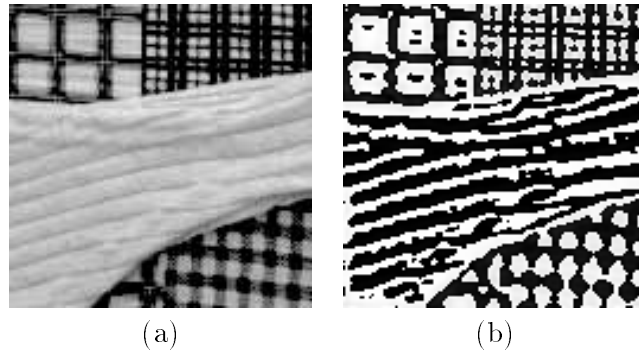


FIGURA 5.3: (a) Imatge natural. (b) Interpretació de la imatge en termes de barres i taques. Aquesta imatge correspon a un pas intermig de l'algorisme de H. Voorhees i T. Poggio.

calcular a partir de trobar una corba d'*stress* per a les pròpies dades que sigui similar a les corbes que es donen a la taula 5.1, evidentment tenint en compte els paràmetres de número de punts, error i dimensió escalada.

5.2 Definició d'un espai paramètric de textures

Fins ara hem definit el mètode que utilitzarem per explorar l'espai. Ara anem a definir un espai paramètric de textures que permetrà introduir determinats paràmetres a la interpretació d'aquest.

Per a la definició d'aquest espai paramètric de textures hem fet la següent assumpció:

Assumpció 5.1 *Tota textura t és una imatge formada per la repetició d'elements bàsics que bé són isotròpics, que els anomenem taques, o bé són allargats i els anomenem barres. Cada element taca o barra té associat un conjunt d'atributs, com l'escala, el contrast o l'orientació.*

El model d'imatge que acabem d'assumir va ser definit per D. Marr [67, 68] a la seva representació els elements bàsics s'anomenen *blobs* i tenen característiques associades de contrast, mida i orientació, formant el *primal sketch* de la imatge. Aquesta assumpció també està d'acord, en general, amb els resultats de la teoria del textó de Julesz [52, 48, 49].

Per tant, per dur a terme l'estudi de les dimensions treballarem sobre un espai paramètric

de textures formades per l'alineament de barres i taques² seguint l'assumpció feta. Abans de definir l'espai paramètric anem a veure un treball previ assumint aquest model d'imatge.

Un dels únics intents de definició d'un algorisme pel tractament de textures segons aquest model de barres i taques amb nivell de gris ha estat fet per H. Voorhees i T. Poggio a [118, 120, 119], i que resumim en el següent apartat.

5.2.1 Algorisme de segmentació basat en el model de barres i taques

En aquest treball es persegueix la segmentació de textures a partir de la detecció de totes les barres i les taques de la textura, seguit del càlcul dels seus atributs i finalment d'un procés de segmentació basat en la definició d'un estadístic de 1r. ordre.

Anem a veure primer de tot la detecció de les taques i de les barres que es basa en la següent definició:

Definició 5.4 *Les barres i les taques són les duals dels contorns. Si considerem que els contorns es detecten calculant els creuaments per zero del laplacà de la imatge, $I * \Delta G$, aleshores les taques es detectaran com les zones de $I * \Delta G$ que no canvien de signe.*

Així doncs, la detecció de les taques i les barres es realitza fent una umbralització sobre $I * \Delta G$, si els contorns són els punts on es canvia de signe, aleshores obtindrem les taques negres fent una umbralització positiva:

$$I * \Delta G > l_0 \quad (5.16)$$

i les taques blanques fent una umbralització negativa:

$$I * \Delta G < l_0 \quad (5.17)$$

Aquest detector presenta el problema de quina ha de ser la mida de la desviació estàndard, σ , de la funció G per a que la detecció sigui correcta. Per a solucionar-ho els autors plantegen trobar l'escala a partir d'un procés d'estimació de soroll, en el que s'assumeix la presència de soroll gaussià i s'estima la seva mida a partir de l'aproximació de la derivada de l'histograma per una funció de Rayleigh.

²És la traducció que utilitzarem per anomenar els *blobs* i *bars* definits per Marr, i que també anomenarem *textons* seguint els treballs de Julesz.

Un cop s'han detectades les taques es segueixen processos diferents per a la segmentació de les taques i les barres.

Per una banda les taques isotròpiques que són detectades per la umbralització ja estan pràcticament segmentades, els únics problemes que es plantegen venen donats pel fet de que dues taques puguin estar juntes quan per la forma que presenten haurien d'estar completament separades. Per a solucionar aquest problema es defineix el següent criteri d'estabilitat:

Criteri d'estabilitat Una regió connectada serà separada en dues si apareix com regions separades amb un rang de nivells de gris més gran o igual que el nombre de nivells en els que apareix com una regió connectada

Les regions que compleixin el criteri d'estabilitat donat hauran de ser separades en dues regions en el nivell inferior. Per tant, s'ha d'aplicar un algorisme que faci créixer les regions separades en nivells superiors sobre les regions del nivell inferior creant la frontera corresponent en el nivell inferior. A la figura 5.4 es pot veure com es comporta aquest criteri d'estabilitat i com s'aplica a un cas concret de taques.

Per altra banda les barres es segmenten a partir del l'esquelet de la regió, aleshores l'esquelet és segmentat per aquells punts on hi ha junctures de dues línies de l'esquelet i pels punts terminals. A més a més, les línies d'esquelet molt curtes són eliminades i les que són molt llargues són partides en vàries. A partir de la reconstrucció de aquest esquelet modificat es reconstrueixen els segments allargats o barres.

Una vegada s'han segmentat totes les taques i totes les barres de la imatge, per a cada una d'elles es calculen els seus atributs: contrast, escala i orientació.

A partir de la imatge representada per les taques i les barres i tots els seus atributs els autors defineixen un estadístic que s'aplicarà sobre diferents veinatges de la imatge amb l'objectiu d'obtenir la segmentació.

5.2.2 Parametrització de textures segons el model de taques i barres

Inspirats amb aquesta manera d'entendre una imatge amb una textura, anem a definir un espai paramètric de generació i representació de textures basat en aquest model de barres i taques.

A partir de les consideracions que acabem de fer podem definir una textura com un punt en un espai paramètric de quatre dimensions. Cada dimensió està associada a un

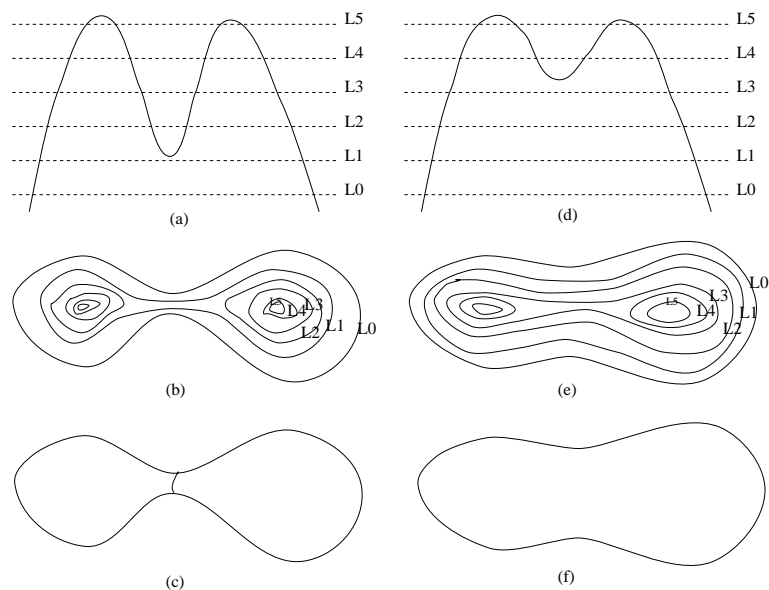


FIGURA 5.4: (a) Perfil de dues taques que seran separades pel criteri. (b) Superfícies de nivell que poden correspondre a la imatge del perfil (a). (c) Segmentació final de la imatge (a). (d) Perfil de taques que no seran separades pel criteri. (e) Superfícies de nivell que poden correspondre a la imatge del perfil (d). (f) Segmentació final de la imatge (d).

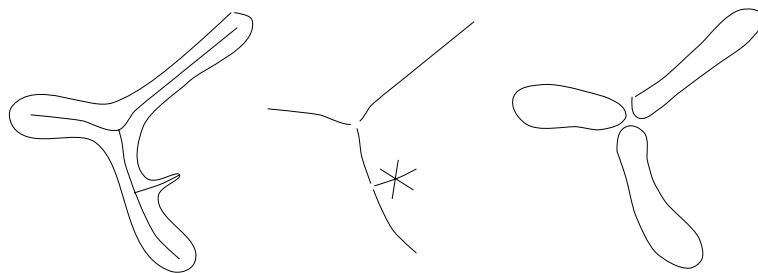


FIGURA 5.5: Exemple de la segmentació de barres a partir de la poda de l'esquelet de les taques allargades.

atribut dels textons. Com a paràmetres hem seleccionat els següents: la mida, la longitud l'orientació i el contrast.

En aquest primer pas de l'estudi estem suposant textures amb un sol tipus de textó que es repeteix a tota la imatge. Aleshores la textura vindrà definida per un punt en aquest espai paramètric (a la figura 5.6 podem veure un conjunt d'imatges d'aquest espai). Una variació en un eix representa una variació de l'atribut corresponent al textó de la imatge.

A totes les imatges introduïm una densitat constant que haurà de coincidir amb el radi d'inhibició de l'algorisme que calcula la representació. Aquesta decisió no afecta a la representació donat que com ja hem dit aquesta no depèn de la posició del textó però sí de la densitat d'aquests. També en totes elles hem introduït un nivell de gris fixe pel fons.

Aquest espai ens permetrà treballar amb conjunts d'imatges de les que podrem aïllar i controlar certes característiques de manera que es puguin preparar experiments que ajudin a fer la interpretació de l'espai.

D'aquesta manera hem definit un model d'imatge simplificat en el que una textura, t , és un punt en un espai paramètric de quatre dimensions, tal que

$$t = i(p_1, p_2, p_3, p_4) = i(s, l, \theta, c) \quad (5.18)$$

on els paràmetres s , l , θ i c representen la mida, la longitud, l'orientació i el contrast, respectivament, del textó de la imatge.

Cal fer notar que el paràmetre de la longitud s'ha introduït per donar coherència al model paramètric, de manera que es pugui passar de forma continua d'una taca a una barra, malgrat l'algorisme que calcula la representació no faci en principi cap detecció específica de la longitud.

En tots els experiments que s'han fet els paràmetres han estat extrets d'un subespai de l'espai $[1, 10] \times [1, 10] \times [0^\circ, 180^\circ] \times [-128, 128]$. El subespai ve donat per tots els punts que compleixin que $s \leq l$, en altre cas tindriem punts diferents que representarien la mateixa imatge.

Les taques es generen a partir d'una funció Gaussiana de dues dimensions com la de l'expressió 3.9, i les barres a partir de l'expressió 3.13. En ambdós casos les mides venen donades per les desviacions estàndards σ_x , σ_y .

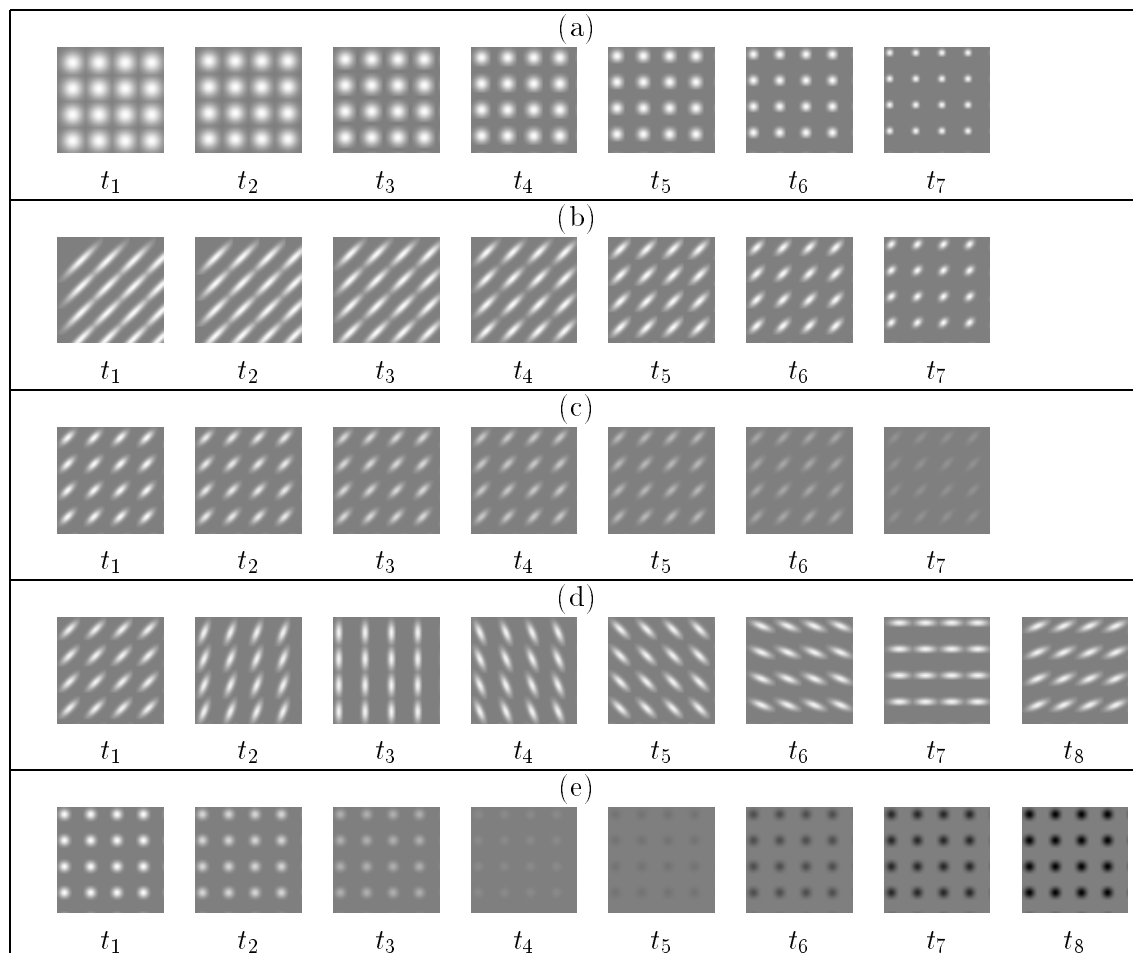


FIGURA 5.6: Grups d'imatges que presenten la variació d'alguns paràmetres. (a) Variació de mida. (b) Variació de longitud. (c) Variació de contrast. (d) Variació d'orientació. (e) Variació de contrast amb creuament per zero.

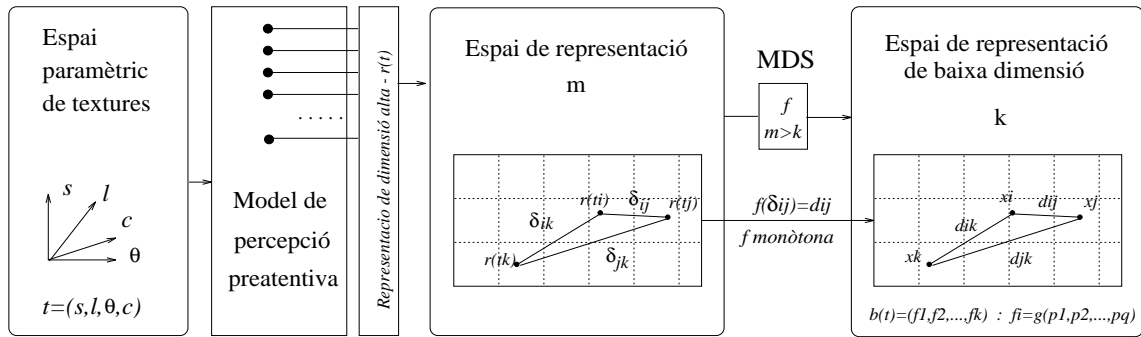


FIGURA 5.7: Diagrama dels passos que seguirem per a l'exploració de l'espai de representació. Els paràmetres perceptuals en els que volem basar la interpretació seran introduïts a l'espai a partir de les imatges d'entrada generades amb el model paramètric d'imatge que hem construït.

5.3 Interpretació de les dimensions de l'espai de representació

La representació de textures que hem definit a l'apartat 3.3 ha creat un espai de representació de dimensió alta (192 dimensions). Ara ens plantegem el problema d'analitzar com es distribueixen les textures representades dins d'aquest espai, és a dir, ens plantegem el problema d'identificar quines són les dimensions d'aquest espai de representació en termes de paràmetres amb un significat perceptual.

Per a obtenir aquesta interpretació amb significació hem introduït els paràmetres perceptuals a les imatges d'entrada, de manera que les diferents variacions d'aquests paràmetres a les entrades ens permetin identificar el comportament d'aquestes dins de l'espai de representació.

El que proposem per aquesta exploració de l'espai de representació es presenta al gràfic de la figura 5.7. Com podem veure en aquest gràfic utilitzarem el MDS per a la reducció de dimensió de l'espai. Ens basarem en un MDS no mètric per disminuir les restriccions en el pas de l'escalament, ja que el nombre de dimensions és molt alt i seria interessant passar a un espai de dimensió molt més baixa.

Així doncs, l'exploració de l'espai que volem dur a terme es pot resumir en els següents passos:

Pas 1 Seleccionem un conjunt de n imatges a l'espai paramètric de textures. Agafarem conjunts de punts a l'espai d'entrada que mantinguin una relació d'ordre entre ells.

Pas 2 Apliquem l'algorisme basat en el model de percepció preatentiva de textures i obtenim la representació, r , de cada imatge.

Pas 3 Construïm la matriu $n \times n$ de dissimilituds entre totes les representacions calculades al pas anterior.

Pas 4 Seleccionem la dimensió k de l'espai al que volem escalar la representació i apliquem el MDS. La desviació o manteniment de l'estructura de l'espai d'entrada dins de l'espai de representació ens permetrà definir aspectes sobre els eixos de l'espai de representació.

En els següents apartats farem servir aquesta metodologia per a fer l'exploració de l'espai de representació. Primer de tot ho farem a partir de variacions d'un sol paràmetre de l'espai paramètric i després passarem a les interaccions entre ells. Cal dir, que aquesta exploració s'ha fet per molts conjunts d'imatges, encara que en els següents apartats només mostrem un exemple de cada un. No hem cregut interessant posar tots els exemples degut a que les configuracions de l'espai de representació escalades sempre presentaven el mateix tipus de deformació.

Un cop hagim identificat cada una de les dimensions de l'espai de representació escalat, confirmarem els resultats sobre un conjunt d'imatges naturals.

5.3.1 Variacions sobre un paràmetre

Primer de tot aplicarem el MDS per explorar el comportament de cada un dels paràmetres de l'espai d'entrada a l'espai de representació. Ho farem a partir de conjunts d'imatges que variïn monòtonament respecte d'un eix, aleshores analitzarem com es comporten les relacions d'ordre entre les representacions d'aquestes imatges.

Variacions en els eixos de mida i longitud a l'espai d'entrada (veure figura 5.6.a i 5.6.b) han estat escalades a variacions monòtones a espais d'1 dimensió ($stress = 0.001, 0.008$ respectivament). Els resultats de l'escalament es mostren a la figura 5.8.a. En aquest moment podem dir que aquest dos paràmetres podrien ser dos eixos a l'espai de representació.

$$(f_1(s), f_2(l)) \tag{5.19}$$

on f_1 i f_2 representen funcions monòtones.

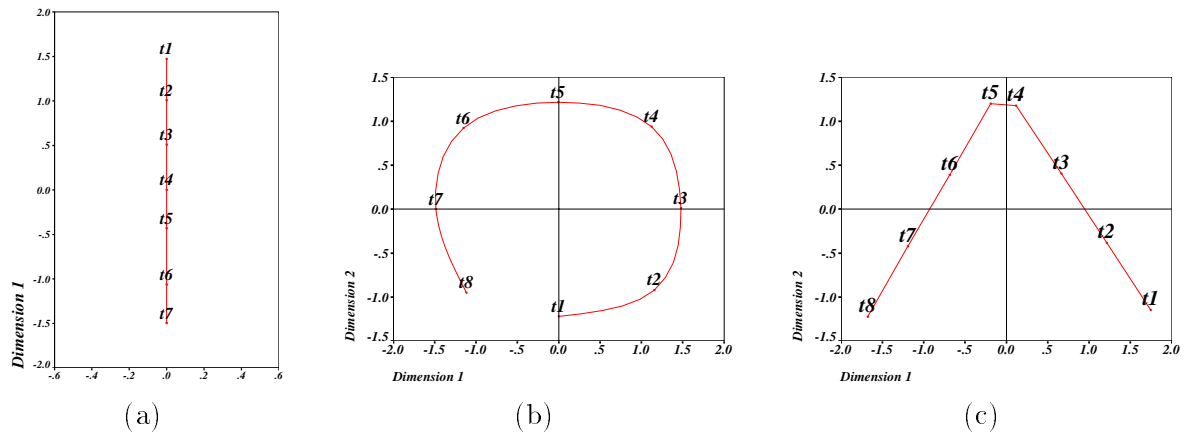


FIGURA 5.8: Configuracions resultants d'aplicar el MDS sobre configuracions de punts dins de l'espai de representació. (a) t_1, \dots, t_7 representen les representacions de les imatges de la figura 5.6.a . (b) t_1, \dots, t_8 representen les representacions de les imatges de la figura 5.6.d . (c) t_1, \dots, t_8 representen les representacions de les imatges de la figura 5.6.e.

Per altra banda les variacions al llarg de l'eix d'orientació implica una configuració no monòtona en un espai d'1 dimensió i amb un valor d'*stress* molt alt ($stress = 0.26$). Ara bé, el valor d'aquest decreix de forma important quan es fa l'escalament a un espai de 2 dimensions, presentant la configuració de la figura 5.8.b ($stress = 0.0002$). La configuració circular apareix indistintament per a qualsevol mida, longitud o contrast.

A partir d'aquest resultat podem assegurar que l'algorisme presenta distàncies màximes sempre que les orientacions difereixen en 90° i que no existeix un origen de coordenades per l'eix de l'orientació. Aquestes observacions ens fan fer una nova hipòtesi sobre dos nous eixos de l'espai de representació:

$$(f_3(\cos 2\theta), f_4(\sin 2\theta)) \quad (5.20)$$

per a representar variacions a l'orientació. En aquestes expressions f_3 i f_4 determinen el radi de la configuració circular. En aquest moment no podem dir res encara sobre la forma d'aquestes funcions.

De la mateixa manera, les variacions de contrast al llarg d'un creuent per zero impliquen una variació no monòtona de la representació en un espai d'1 dimensió ($stress = 0.2$), mentre que l'escalament per a un espai de 2 dimensions dóna un *stress* de 0.00004,

la configuració d'aquest escalament es pot veure a la figura 5.8.c. En aquest experiment les imatges d'entrada presentaven un increment de contrast constant des de t_8 fins a t_1 , amb un creuament per zero entre les imatges t_4 i t_5 . D'aquest resultat podem deduir també que existeix un comportament simètric entre els textons blancs i negres³, i una diferència creixent entre ells a mesura que el valor absolut del contrast creix. Per tant, podem fer la hipòtesi de com són els eixos que descriuen les textures amb diferent signe de contrast:

$$(sgn(c) \cdot f_5(|c|), f_6(|c|)) \quad (5.21)$$

on la funció sgn representa el signe de l'argument i f_5 , f_6 semblen ser funcions monòtones.

Fins aquí ja podem donar una primera aproximació sobre l'estructura de l'espai de representació, en termes dels paràmetres s , l , θ i c . Per tant, amb aquest resultat anem des d'un espai de representació de 192 dimensions a una aproximació d'aquest en 6 dimensions. A aquest nou espai l'anomenarem b i ve donat per

$$b(t) = (f_1, \dots, f_k) \simeq (f_1, f_2, f_3, f_4, f_5, f_6) \quad (5.22)$$

on k representa la teòrica dimensionalitat correcta i f_1 , f_2 , f_5 i f_6 són funcions mòno- tones, definint el següent espai

$$b(t) \simeq (f_1(s), f_2(l), f_3(\cos 2\theta), f_4(\sin 2\theta), sgn(c) \cdot f_5(c), f_6(|c|)) \quad (5.23)$$

5.3.2 Interaccions entre paràmetres

Per a poder refinar la hipòtesi sobre les dimensions de l'espai de representació, anem a explorar com interaccionen les dimensions entre elles. Aquesta exploració es pot fer construint conjunts d'imatges en les que les variacions es produeixin en més d'un paràmetre alhora. L'objectiu estarà en veure si la forma de les configuracions de dues dimensions a l'espai paramètric són mantingudes a l'espai de representació. Construirem configuracions en forma de triangle o quadrat i veurem si es conserva o si és deformada.

A partir de la configuració de la figura 5.9.a de l'espai de textures podem veure la reconstrucció de la configuració a l'espai de representació en dues dimensions a la figura

³Entenem textons blancs com aquells amb un nivell de gris alt i textons negres com aquells que tenen un nivell de gris negatiu respecte del nivell del fons.

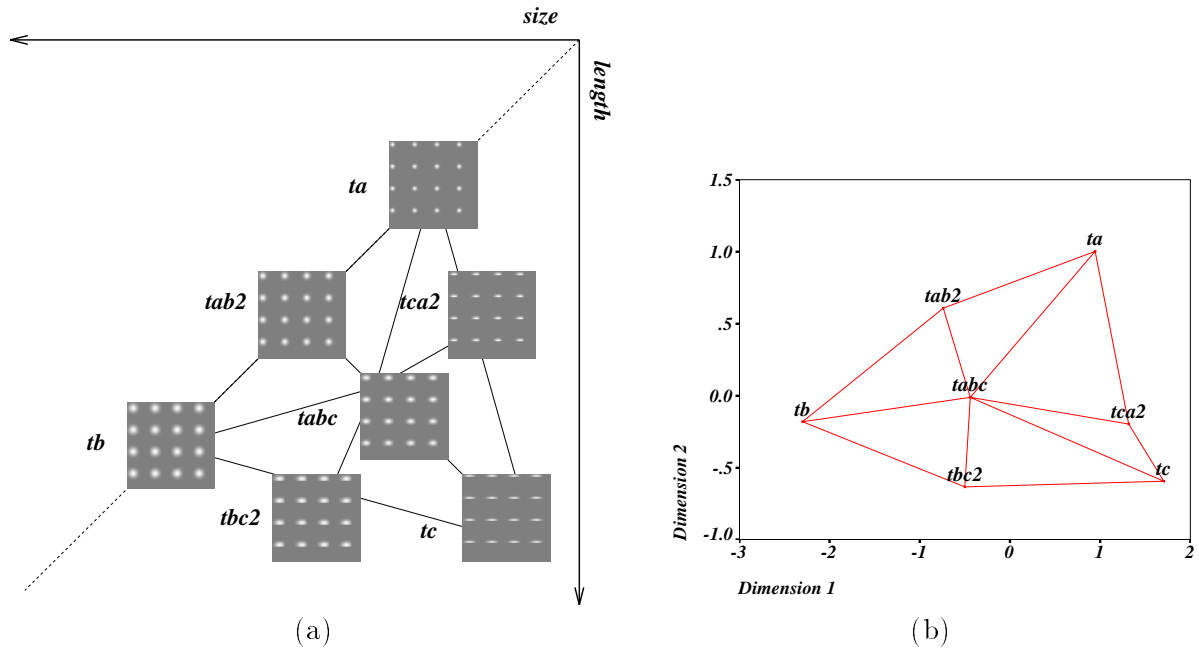


FIGURA 5.9: (a) Punts a l'espai paramètric de textures presentant una configuració en forma de triangle. Les imatges t_a , t_b i t_c estan situades als vèrtexs d'un triangle equilàter, t_{ij2} és el punt mig entre t_i i t_j , i t_{abc} és el punt central del triangle. Les variacions es donen sobre els paràmetres de mida i longitud (assumint $s \leq l$). (b) Resultat d'aplicar el MDS sobre les representacions de les imatges d'(a) sobre una espai de 2 dimensions ($stress = 0.003$).

5.9.b. Podem deduir que l'estructura és preservada per la representació de forma monòtona. Existeix una desviació deguda a la imatge t_c , però no es pot deduir cap interacció clara entre els dos eixos. Altres configuracions de triangles que han estat escalades presenten també certes desviacions per algunes imatges, com en aquest cas, però no permeten deduir cap relació específica.

La variació entre el contrast i la mida es pot veure a la figura 5.10.a. La configuració reconstruïda pel MDS sobre les representacions es pot veure a la figura 5.10.b. Podem interpretar que les distàncies entre textures del mateix contrast creixen a mesura que el contrast creix, això és

$$f_1 = g_1(s, c) \quad (5.24)$$

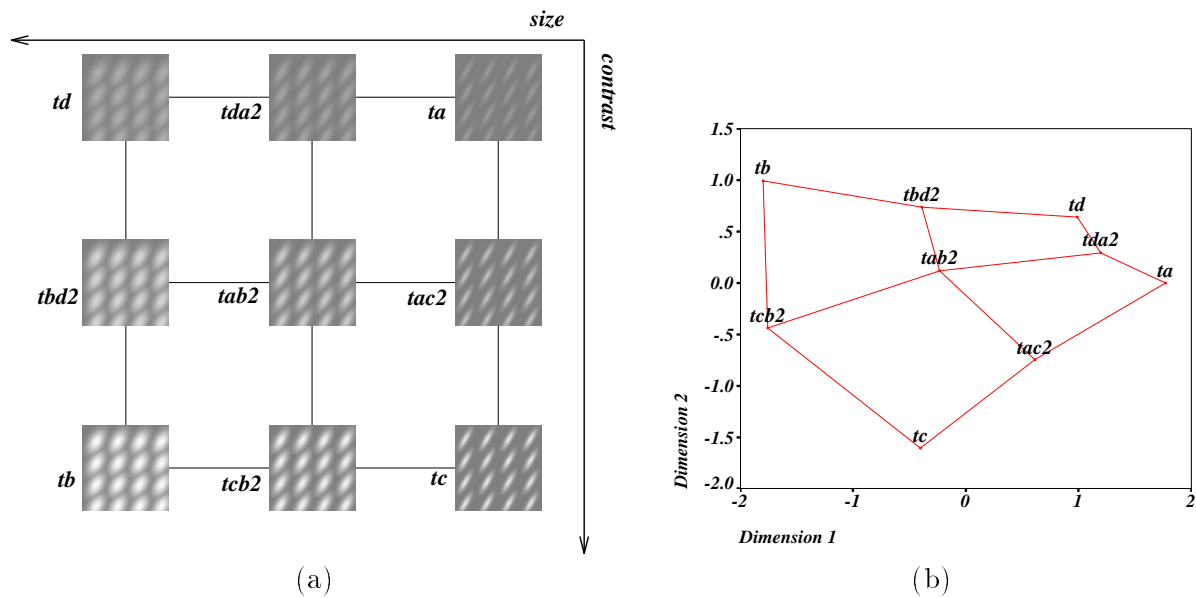


FIGURA 5.10: (a) Configuració de quadrat amb punts de l'espai paramètric, les variacions són de mida i contrast. (b) Solució en 2 dimensions del MDS sobre les representacions de les imatges d'(a) ($stress = 0.0006$).

on g_1 és una funció monòtona de s i de c . Mentre que les distàncies entre textures de diferent contrast sembla dependre només de la diferència de contrast, siguin quins siguin la resta de paràmetres.

$$f_6 = g_6(|c|) \quad (5.25)$$

on g_6 és una funció monòtona que depèn només de $|c|$.

Aquestes mateixes conclusions es poden deduir a partir de la configuració quadrada amb imatges que presenten un contrast negatiu. Els quadrats són deformats de la mateixa manera a partir de la configuració donada, fet que confirma les expressions (5.24) i (5.25).

Interpretacions semblants es donen a partir de configuracions en les que varien la longitud i el contrast (veure figura 5.11) i configuracions que varien el contrast respecte de l'eix d'igual longitud i mida (veure figura 5.12). Aleshores podem deduir també que

$$f_2 = g_2(l, c) \quad (5.26)$$

on g_2 és una funció monòtona.

Interaccions entre el paràmetre d'orientació i la resta de paràmetres també han estat estudiades. A la figura 5.13.a hem seleccionat un conjunt d'imatges que varien en contrast i orientació, i a la figura 5.13.b podem veure la configuració d'aquestes imatges a l'espai de representació. De la configuració de l'espai escalat podem dir de forma aproximada que la dimensió 1 correspon a variacions del contrast i les dimensions 2 i 3 corresponen a les variacions de l'orientació en el sentit que ha estat expressat amb les funcions f_3 i f_4 . Podem veure que les imatges que varien en orientació i que tenen el mateix contrast es troben sobre el mateix pla en una configuració circular. A més a més podem observar que per una orientació donada la representació es troba en el mateix pla mantenint la mateixa configuració respecte del contrast que hem vist a la figura 5.8.c. Una observació final ve donada pel fet de que el radi del cercle és independent del signe del contrast, aleshores

$$f_3 = g(|c|) \cos(2\theta) \quad (5.27)$$

$$f_4 = g(|c|) \sin(2\theta) \quad (5.28)$$

A partir d'aquesta darrera deducció hem volgut comprovar si el radi depèn només del contrast o també depèn de la resta de paràmetres. Les gràfiques de la figura 5.14 donen

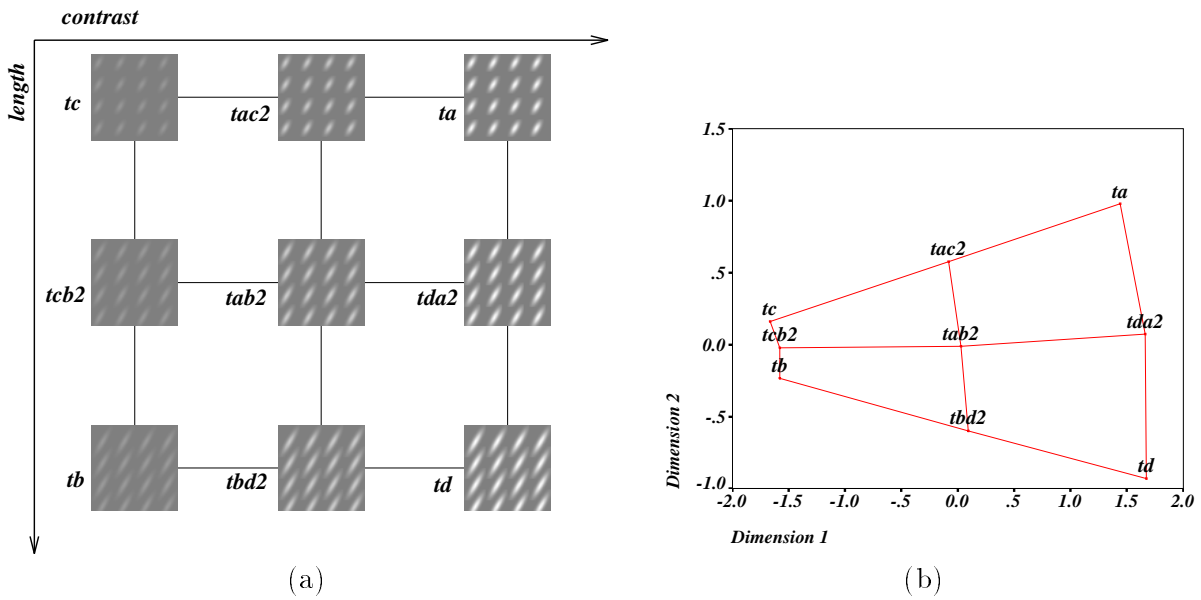


FIGURA 5.11: (a) Configuració de quadrat amb punts de l'espai paramètric, les variacions són de longitud i contrast.(b) Solució en 2 dimensions del MDS sobre les representacions de les imatges d'(a) ($stress = 0.0002$).

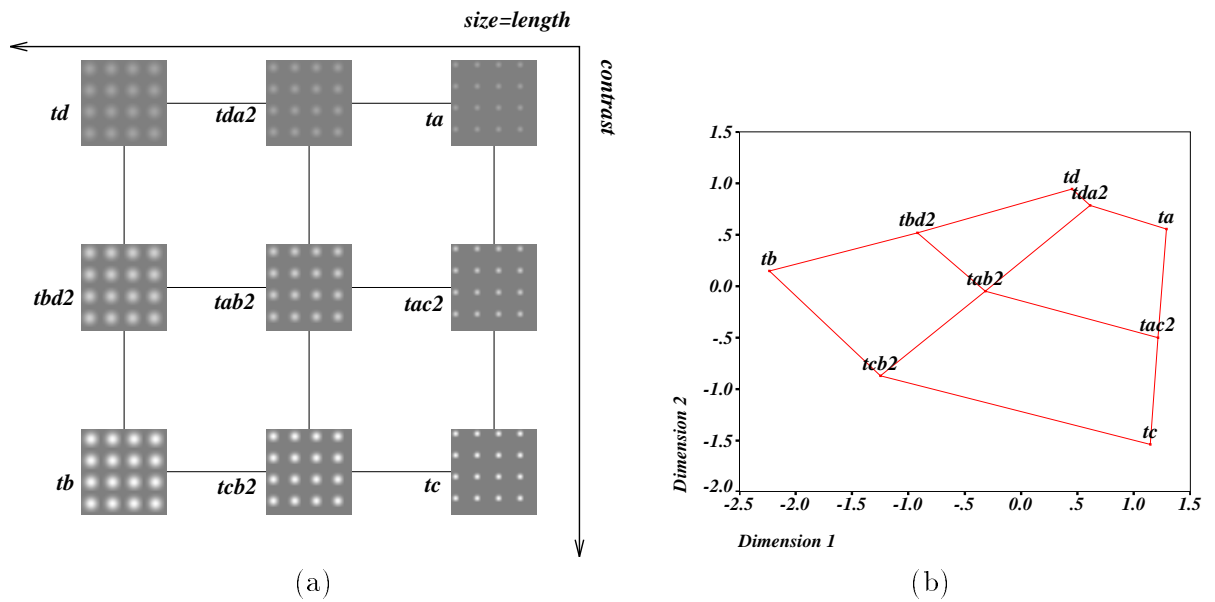
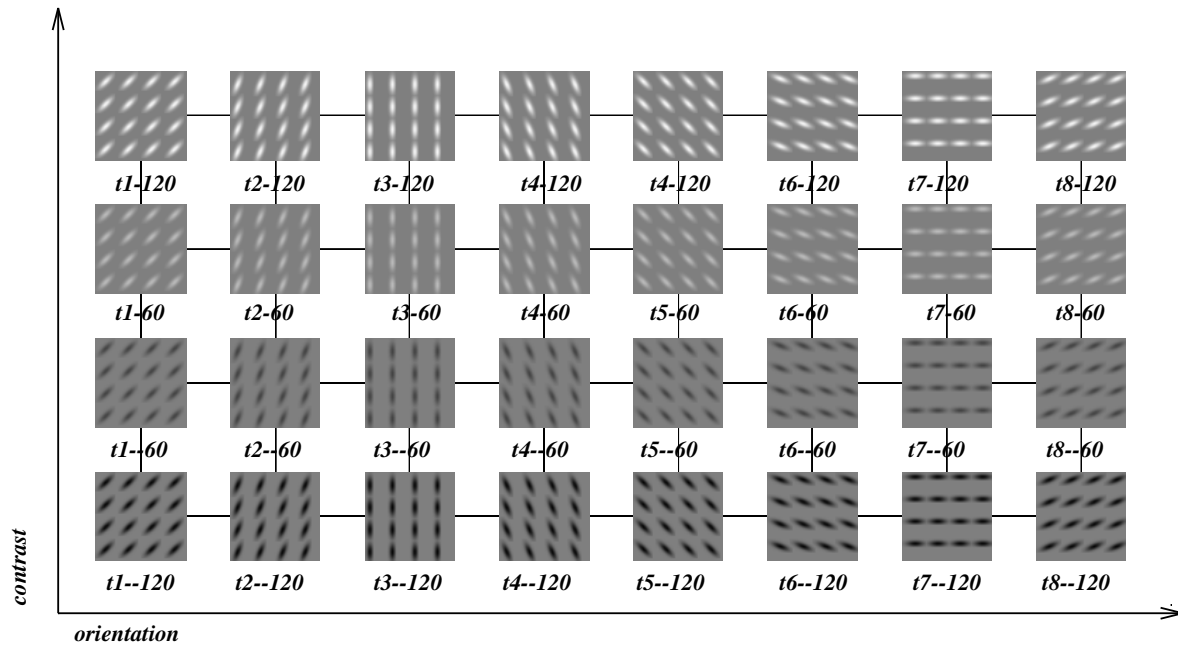
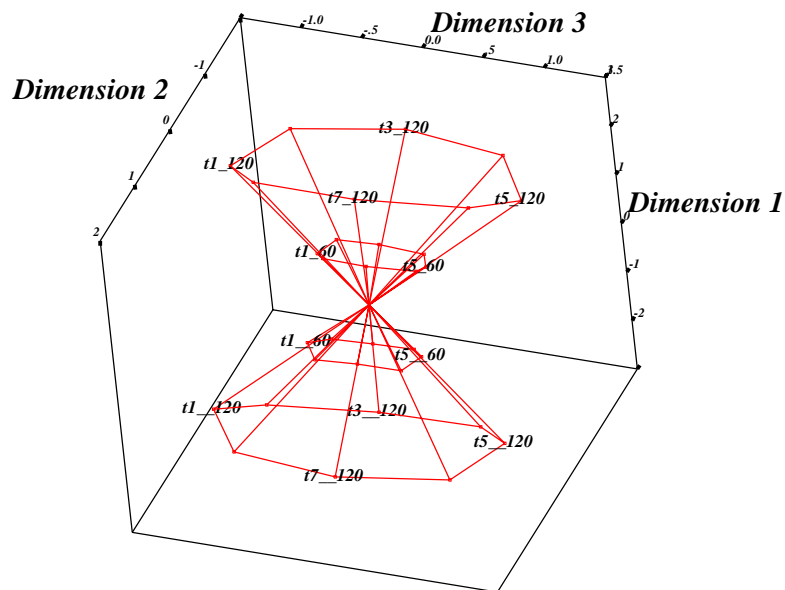


FIGURA 5.12: (a) Configuració de quadrat amb punts de l'espai paramètric, les variacions són de contrast respecte de l'eix que forma la línia d'igual mida i longitud. (b) Solució en 2 dimensions del MDS sobre les representacions de les imatges d'(a) ($stress = 0.005$).



(a)



(b)

FIGURA 5.13: (a) Trama de 8×4 punts a l'espai paramètric de textures, amb variacions en orientació (diferències d'angle= 23°) i contrast a través d'un creuament per zero. (b) Solució en 3 dimensions del MDS de les representacions corresponents de les imatges de (a) ($stress = 0.038$).

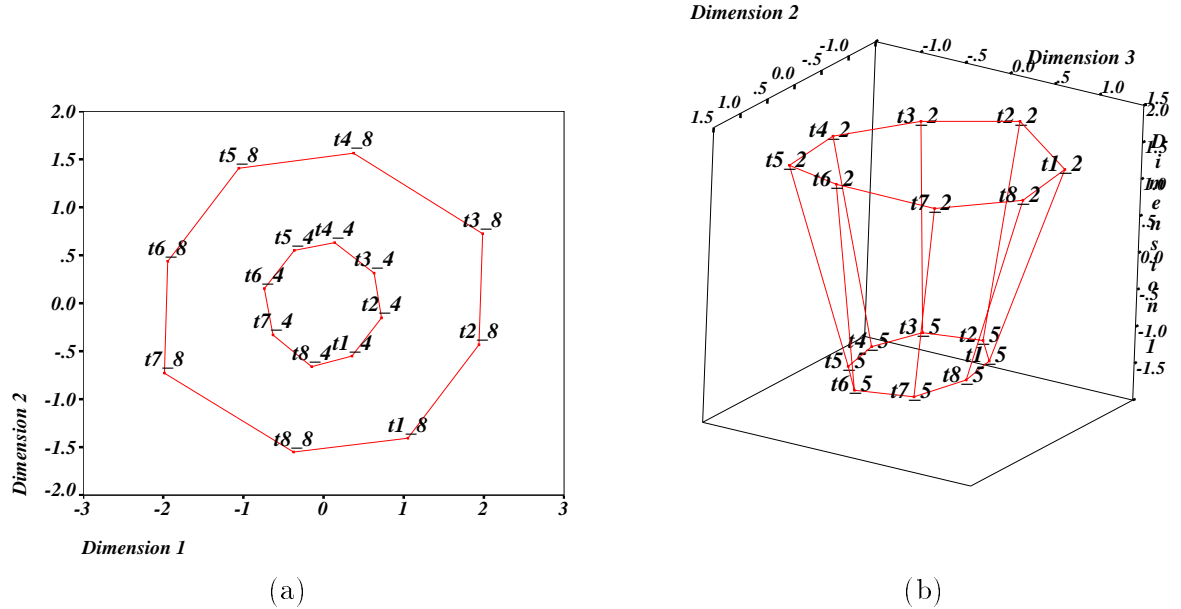


FIGURA 5.14: (a) Solució en 2 dimensions del MDS de les representacions de dos conjunts d'imatges representades pels punts $\{t1_4, \dots, t8_4\}$ i $\{t1_8, \dots, t8_8\}$, on tk_i representa la textura amb orientació θ_k i longitud i ($stress = 0.016$). (b) Solució en 3 dimensions del MDS de les representacions de dos conjunts de punts $\{t1_2, \dots, t8_2\}$ i $\{t1_5, \dots, t8_5\}$, on tk_j representa la textura amb orientació θ_k i longitud j ($stress = 0.002$). En ambdós casos les orientacions són separades consecutivament per intervals de 23° .

suport a aquesta hipòtesi, ja que com podem observar el radi de la configuració varia per diferents longituds de barres (figura 5.14.a), i per diferents mides (figura 5.14.b). A partir d'aquesta observació podem deduir les següents expressions

$$f_3 = g(s, l, |c|) \cos(2\theta) \quad (5.29)$$

$$f_4 = g(s, l, |c|) \sin(2\theta) \quad (5.30)$$

que significa que el radi depèn de variacions monòtones de mida, longitud i del valor absolut del contrast. Aquest fet ha estat expressat amb la funció g , ja que la configuració circular no presenta cap distorsió en cap de les seves dimensions.

Finalment hem estudiat les representacions entre imatges que es distingeixen pel signe del contrast i varien dependent d'un altre paràmetre que pot ser la longitud, la mida o les

dues alhora. Podem veure els resultats a la figura 5.15. Les distàncies corresponents entre les representacions ens porten a fer la hipòtesi de que aquests paràmetres no afecten les distàncies. Aleshores sembla que només depèn del valor absolut del contrast, per tant

$$f_5 = \text{sgn}(c) \cdot g_5(|c|) \quad (5.31)$$

on g_5 representa una funció monòtona.

Per acabar podem resumir totes les observacions en la següent expressió

$$b(t) \simeq (g_1(s, c), g_2(l, c), g(s, l, c) \cos(2\theta), g(s, l, |c|) \sin(2\theta), \text{sgn}(c) \cdot g_5(|c|), g_6(|c|)) \quad (5.32)$$

que representa la forma dels eixos de l'espai de representació en un espai de 6 dimensions.

5.3.3 Dimensionalitat de l'espai de representació

La darrera qüestió que ens queda per respondre és sobre la dimensionalitat correcta d'aquest espai de representació. Per a poder respondre a aquesta pregunta hem calculat els valors d'*stress* resultants d'aplicar el MDS sobre les representacions del conjunt de 36 imatges que hem fet servir en tots els experiments dels apartats anteriors. A la figura 5.16.b donem els valors d'*stress* per a cada dimensió reconstruïda en front dels valors d'*stress* de les dades simulades amb l'aproximació de Monte-Carlo i que hem donat a la taula 5.1.

A partir de la gràfica construïda i dels valors que podem comparar, podem deduir que les dades presenten una dimensionalitat correcta amb valor 4. Tenint tot això en compte, donarem un darrera hipòtesi sobre la forma de les dimensions de l'espai

$$b(t) \simeq (g(s, l, c) \cos(2\theta), g(s, l, |c|) \sin(2\theta), \text{sgn}(c) \cdot g_5(|c|), g_6(|c|)) \quad (5.33)$$

on hem eliminat dues dimensions que havíem introduït abans i que podien ser considerades redundants si tenim en compte que la mida, la longitud i el contrast ja apareixien a les altres dimensions. A més a més, aquesta nova expressió manté totes les restriccions que els experiments ens han fet introduir al llarg de tot l'estudi.

Cal fer notar que pels casos de textures amb textons no orientats hem de tenir en compte el següent

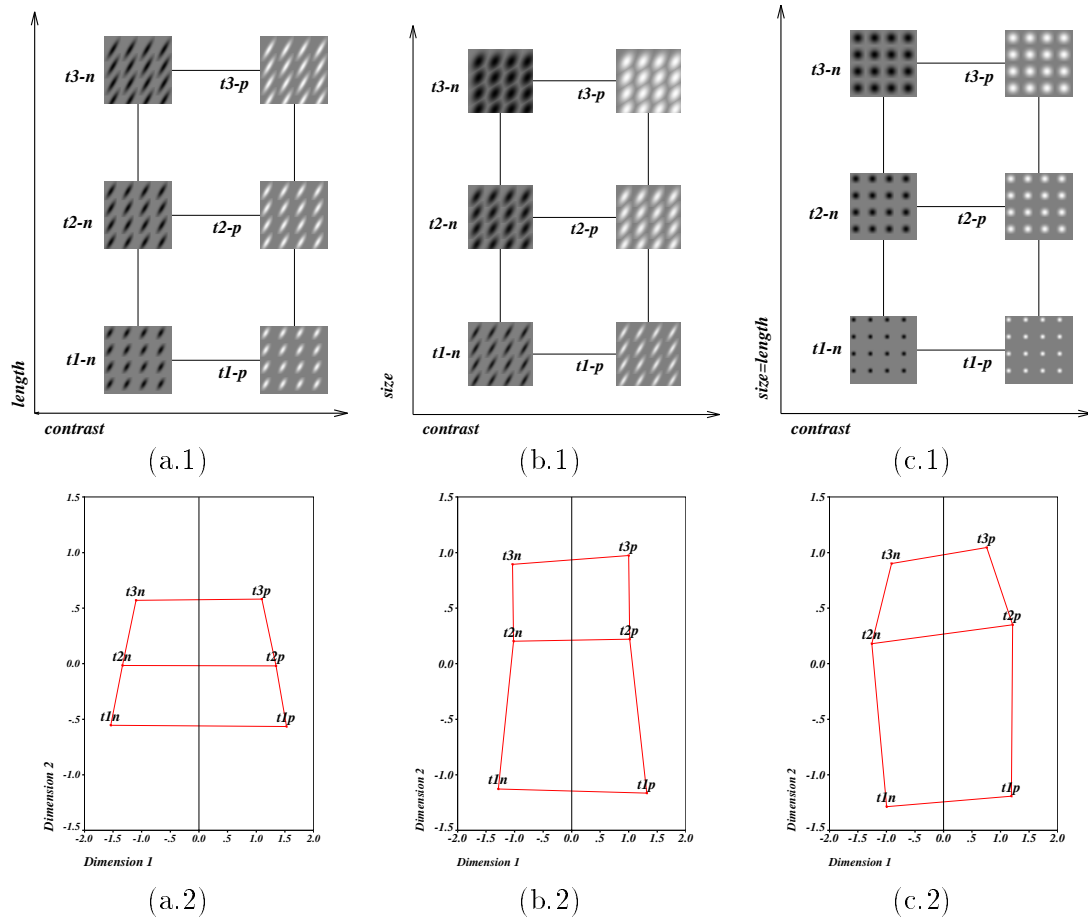
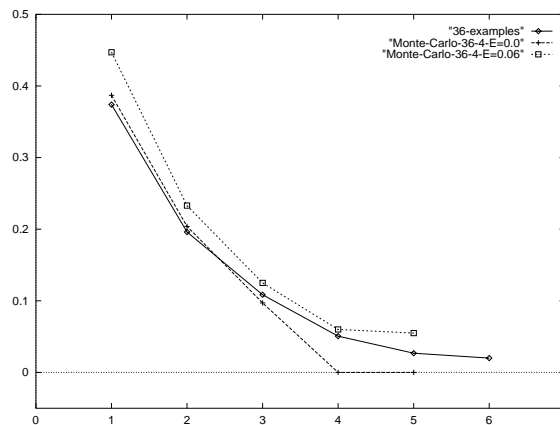


FIGURA 5.15: (a.1) Conjunt d'imatges que varien en longitud i en el signe del contrast. (a.2) Solució en 2 dimensions del MDS sobre les representacions de les imatges de (a.1) ($stress = 0.001$). (b.1) Conjunt d'imatges que varien en mida i en el signe del contrast. (b.2) Solució en 2 dimensions del MDS sobre les representacions de les imatges de (b.1) ($stress = 0.001$). (c.1) Conjunt d'imatges que varien en mida i longitud alhora i en el signe del contrast. (c.2) Solució en 2 dimensions del MDS sobre les representacions de les imatges de (c.1) ($stress = 0.001$).

k	Stress (dades)	Simulació de Monte-Carlo	
		$e=.0000$	$e=.0625$
1	0.37402	0.387	0.447
2	0.19609	0.204	0.233
3	0.10843	0.097	0.125
4	0.05063	0.000	0.060
5	0.02685	0.000	0.055
6	0.02000	-	-

(a)



(b)

FIGURA 5.16: (a) Valors d'*stress* per conjunts de dades escalades a diverses dimensions, denotades amb k . La primera columna correspon a les solucions del MDS sobre les representacions de 36 imatges de l'espai paramètric. Les altres columnes corresponen a dades simulades provinents de la taula 5.1. (b) Representació gràfica dels valors d'*stress* de la taula (a).

$$b(t) \simeq \begin{cases} (g(s, l, c) \cos(2\theta), g(s, l, |c|) \sin(2\theta), \operatorname{sgn}(c) \cdot g_5(|c|), g_6(|c|)) & \text{Si } t \text{ formada per barres} \\ (g(s, l, c), g(s, l, |c|), \operatorname{sgn}(c) \cdot g_5(|c|), g_6(|c|)) & \text{Si } t \text{ formada per taques} \end{cases} \quad (5.34)$$

Com a resum podem dir que la conclusió més relevant ve donada pel fet que hem aconseguit una reducció important en el nombre de dimensions i que il·lustrem a la figura 5.17, al mateix temps que coneixem com interaccionen els diversos paràmetres que és poden donar a les imatges d'entrada, i per tant s'ha donat una interpretació a un model de dimensionalitat alta.

5.3.4 Confirmació de la interpretació amb imatges naturals

Per a donar suport a la interpretació de l'espai de representació que hem donat a l'apartat anterior, ara provarem si el podem confirmar sobre un conjunt d'imatges naturals. Hem seleccionat un conjunt d'imatges de l'àlbum de Brodatz [10], que en un cert sentit es poden interpretar com punts dins de l'espai paramètric definit.

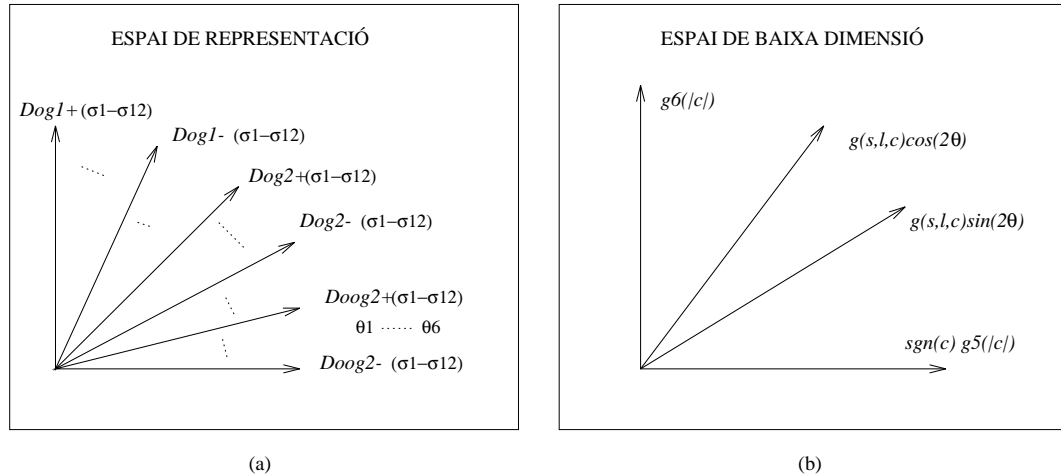


FIGURA 5.17: Il·lustració de la reducció de dimensions aconseguida a partir de l'estudi de les dissimilituds entre les representacions. (a) Espai de representació donat pel model computacional. (b) Espai deduït després de l'escalament a la dimensió correcta.

Primer de tot, hem aplicat el mateix procés sobre el conjunt d'imatges de la figura 5.18. Aquestes imatges presenten barres amb una direcció predominant i que hem manipulat per aconseguir diferents contrastos entre elles. La solució en 2 dimensions d'aplicar el MDS sobre les representacions d'aquestes imatges es poden veure a la figura 5.18(baix). Com podem veure la configuració confirma la forma que havíem predit a la interpretació de l'espai de representació. Les diferències degudes a l'orientació es configuren formant un cercle amb el radi dependent del contrast.

Per altra banda, hem provat el comportament d'una imatge formada per taques sense cap orientació que es pugui percebre (veure figura 5.19.a imatge *b-120*). En una primera aproximació podem suposar que és una imatge formada per taques negres sobre un fons clar, aquest seria el resultat d'una interpretació d'alt nivell sobre la imatge.

Aleshores hem variat el contrast de la imatge amb transformacions fotomètriques, obtenint les imatges *b120*, *b60* i *b-60* (figura 5.19.a), intentant simular el mateix efecte que les variacions de contrast sobre les textures paramètriques de la figura 5.19.c. Aplicant el MDS sobre les representacions d'aquests dos conjunts d'imatges obtenim la configuració de la figura 5.19.d, aquest resultat ens permet interpretar com es comporta la representació. Podem veure que la textura *b120* està més aprop de *t-120* en lloc d'estar més aprop de *t120* com havíem predit per la interpretació que en feiem de la textura. El mateix efecte és obtingut

per la resta de textures. La interpretació que se'n pot fer és que la imatge és representada per les taques que forma el fons, en lloc de les taques que nosaltres havíem suposat, a partir d'una interpretació basada en el coneixement de l'escena.

Per confirmar aquesta interpretació hem construït el conjunt d'imatges de la figura 5.19.b, per les que hem repetit les mateixes transformacions fotomètriques que abans però sobre el fons de la imatge en lloc de sobre les taques negres. La solució de 2 dimensions del MDS confirma la hipòtesi que hem fet. Ara la imatge *bk120* és més propera a *t120*.

De la mateixa manera hem volgut testejar el paràmetre de la mida. Ho hem fet a partir de la reducció de la resolució de la imatge. Hem construït una configuració quadrada de textures variant en contrast i en resolució (veure figura 5.20.a), com si estiguéssim variant els paràmetres de contrast i mida a l'espai de paràmetres. La configuració en 2 dimensions resultant del MDS preserva la configuració quadrada de la mateixa manera que hem predit per les textures de l'espai paramètric (veure figura 5.20.b).

5.4 Conclusions sobre l'aproximació i els resultats

De l'exploració que acabem de fer de l'espai de representació volem extreure algunes conclusions sobre la metodologia i sobre els resultats.

La metodologia seguida ha vingut inspirada per treballs recents en intel·ligència artificial [20] que tendeixen a l'ús de mètodes empírics per a la caracterització de programes, entorns, tasques i comportaments de sistemes computacionals. Aquests mètodes empírics donen una estratègia de generalització en cinc passos:

- (i) Construcció d'un programa que tingui un comportament interessant.
- (ii) Identificació de característiques específiques del programa, de les tasques i dels entorns que influencien el comportament objectiu.
- (iii) Desenvolupar i testejar un model causal de com aquestes característiques influeixen el comportament objectiu.
- (iv) Generalitzar les característiques de manera que altres programes, tasques i entorns siguin abarcats pel model causal definit.
- (v) Testejar si el model general prediu el comportament d'aquest conjunt de programes, tasques i entorns.

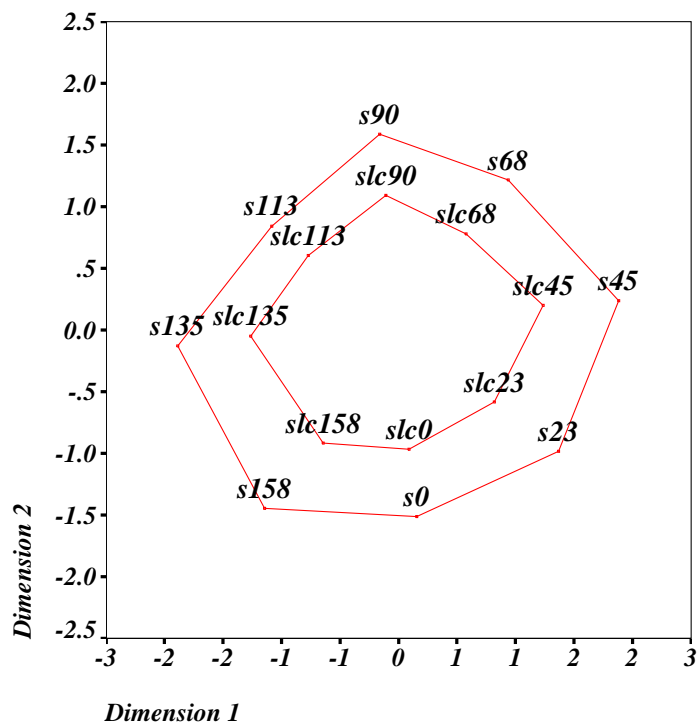
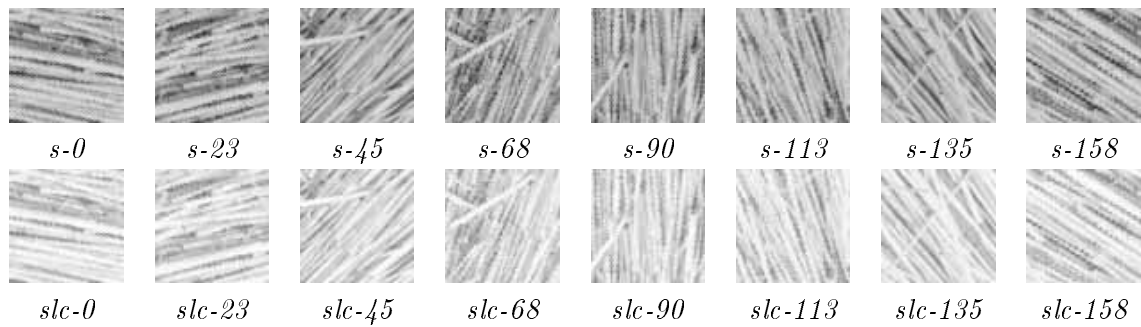


FIGURA 5.18: Conjunt de textures amb una orientació predominant per a les barres. Les textures $\{s-0, \dots, s-158\}$ presenten transformacions fotomètriques dins del rang de nivells de gris que van de 0 a 255. Les textures $\{slc-0, \dots, slc-158\}$ presenten una transformació fotomètric en el rang de 128 a 255. Abaix es dona la configuració de 2 dimensions resultant del MDS sobre les representacions d'aquestes imatges ($stress = 0.05$).

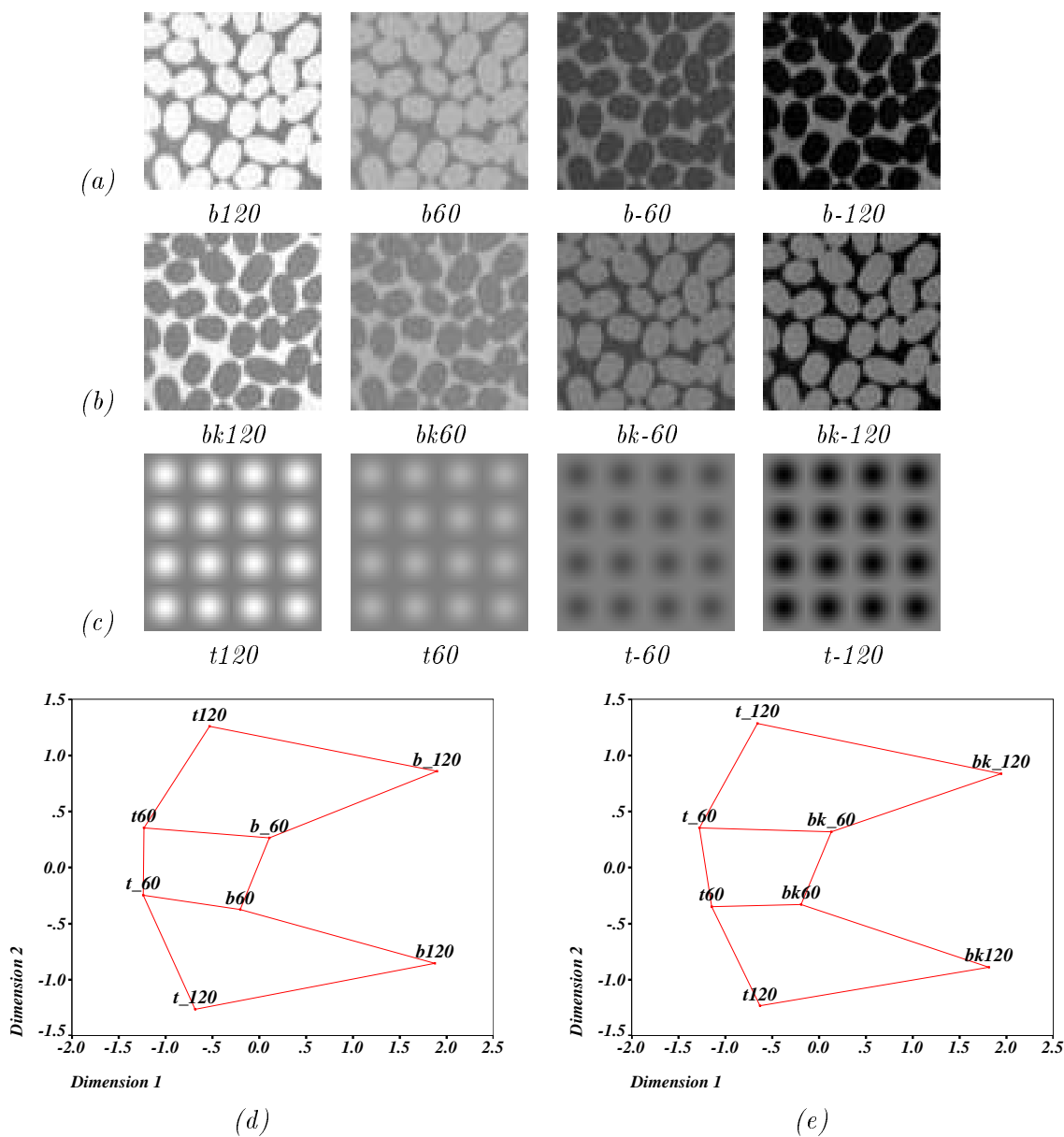


FIGURA 5.19: (a) Variacions fotomètriques d'una textura, el nivell de gris del fons s'ha fixat a 128, i el nivell de gris de les taques varia en diferents intervals respecte del fons, aquests són $[128, 256]$, $[128, 182]$, $[64, 128]$, $[0, 128]$. (b) Variacions fotomètriques d'una textura, el nivell de gris de les taques és fixat a 128 i el nivell de gris del fons ha estat variat en els mateixos intervals de nivell de gris que (a). (c) Quatre textures de l'espai paramètric formades per taques amb igual mida i longitud i un nivell de gris fixe pel fons, les variacions de contrast de les taques passen al llarg d'un creuament per zero. (d) Solució de 2 dimensions de les representacions de les imatges de (a) i (c) ($stress = 0.03$). (e) Solució de 2 dimensions de les representacions de les imatges de (b) i (c) ($stress = 0.03$).

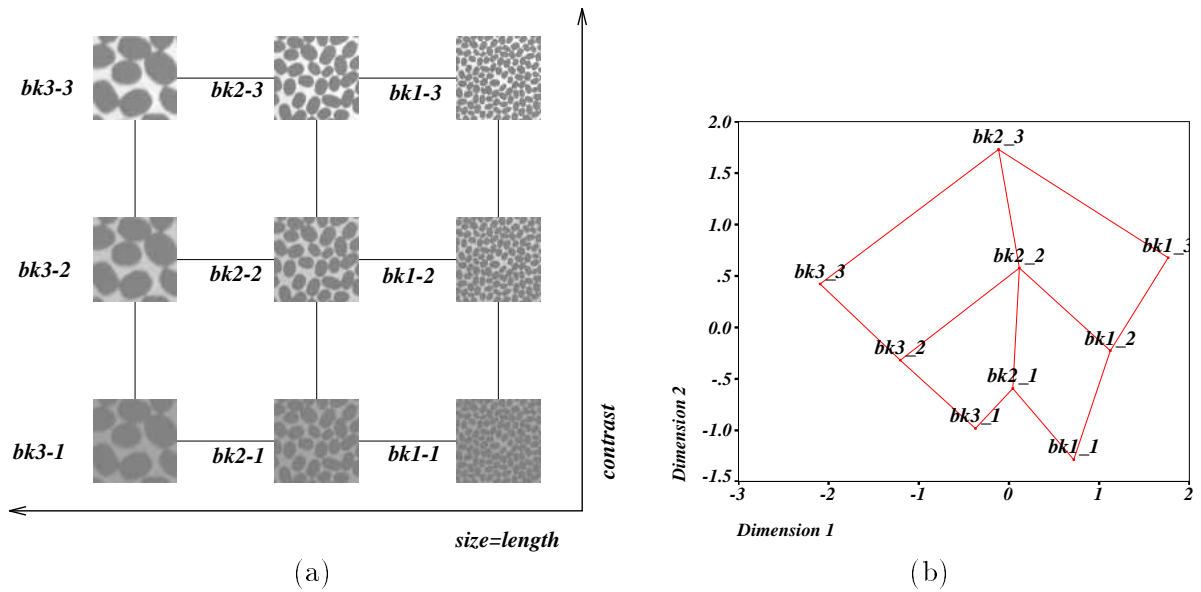


FIGURA 5.20: (a) Imatges resultants de transformar la resolució i el nivell de gris del fons. Les variacions de resolució s'han posicionat al llarg de l'eix d'igual mida i longitud. Les variacions de contrast sobre el fons estan a l'eix de contrast. (b) Configuració en 2 dimensions resultant d'aplicar el MDS sobre les representacions de les imatges de (a) ($stress = 0.004$)

En aquest capítol ens hem centrat en el segon pas d'aquesta estratègia. A partir del tipus de filtres que utilitza el programa, del tipus d'imatge i tenint en compte que l'objectiu és el de donar una interpretació a la representació, hem definit un conjunt de paràmetres. Aquests ens han permès generar imatges representant l'entorn del programa i ens ha permès veure com influeixen aquests paràmetres en el comportament de la representació.

La conservació de les relacions de similitud en un i altre espai ens estaran validant l'ús de l'algorisme per a la tasca corresponent. Aquestes idees han estat recollides a [114, 112] com a metodologia d'exploració del comportament d'un algorisme.

Per tant, en aquest capítol hem explorat l'espai de representació definit i hem trobat una determinada estructura definida pels paràmetres que hem especificat. Valorar la validesa perceptual del comportament de l'algorisme vol dir construir l'espai perceptual corresponent i veure si es comporta amb la mateixa estructura. Aquesta és una de les vies obertes que queden per acabar de completar aquest treball.

Ara que hem extret algunes conclusions sobre la metodologia seguida, i tenint en compte que l'estructura ha estat deduida a partir d'un model simplificat de les imatges d'entrada, anem a veure com podríem generalitzar aquests resultats.

El que és evident és que tota textura no es pot definir per una orientació, una mida o un contrast. Si volem estendre el resultat podríem intentar repetir el procés per textures que continguin més d'un tipus de textó. Això significa que una textura ja no vindria definida per quatre paràmetres sinó que necessitaríem quatre paràmetres per a cada tipus de textó. Per exemple, una imatge amb n tipus de textons diferents la representariem com

$$t = i(s_1, l_1, \theta_1, c_1, s_2, l_2, \theta_2, c_2, \dots, s_n, l_n, \theta_n, c_n) \quad (5.35)$$

Aquesta parametrització ens portaria a un espai paramètric de $n \times 4$ dimensions, per al que repetir l'exploració que hem fet pel de 4 implicaria la interacció entre molts paràmetres. Aquest procediment resultaria inviable i de poc interès. Al següent capítol proposem una solució a aquest problema a partir de la construcció de la parametrització directament sobre la representació.

CAPÍTOL 6

IDENTIFICACIÓ DE LES DIMENSIONS DE L'ESPAI DE REPRESENTACIÓ

En aquest capítol es pretén estendre els resultats obtinguts al capítol anterior per a qualsevol textura. Per tant, l'objectiu que ens plantejem és doble, primer de tot volem veure si la reducció de dimensions de l'espai de representació es dona no només per les textures de l'espai paramètric sinó que la dimensionalitat real de l'espai de representació és 4 independentment de les textures que fem servir en l'exploració. I en segon lloc volem veure si la interpretació de les dimensions que hem fet i que es resumeix a la figura 5.17.b es pot estendre per a qualsevol textura.

Per poder assolir aquests dos objectius hauríem d'estendre l'espai paramètric vist en el cas general. Com ja hem comentat, una parametrització general de l'espai d'entrada ens portaria a un nou espai d'alta dimensió. Per solucionar aquest problema hem seguit els mateixos passos que R. Rao i G.L. Lohse en els seus experiments [81, 80]. Hem seleccionat un conjunt d'imatges en el que hi ha imatges prou diferents entre elles i per a les que fem l'assumpció de que representen un conjunt general de totes les textures¹.

Per dur a terme l'estudi canviarem la metodologia respecte de l'anterior capítol. Els paràmetres en els que hem basat tota la interpretació no es coneixen a la imatge de partida sinó que s'obtenen a partir de la mateixa representació. Per altra banda, la identificació de les dimensions es farà a partir d'una anàlisi de correlació entre la parametrització de la representació de la textura i l'espai de representació escalat.

¹Els autors de [81] argumenten aquest fet així: *"(...) We made this assumption in order to avoid a circular problem: how do we generate a fair sample from a space whose dimensions we are trying to uncover? (...)"*

Dimensió escalada	<i>stress</i>
1	0.35341
2	0.15683
3	0.08905
4	0.04881
5	0.03614
6	0.02911

TAULA 6.1: Valors d'*stress* resultants d'aplicar el MDS per diverses dimensions sobre les representacions de les imatges seleccionades a l'apartat 6.1

6.1 Selecció de les imatges

Hem seleccionat un conjunt d'imatges de l'album de Brodatz [10]. La decissió d'agafar aquestes imatges ve donada pel fet de que són imatges que han estat usades en molts treballs de recerca en el camp de la visió per computador, i perquè es poden obtenir en moltes bases de dades d'imatges. De les 112 imatges que té l'album hem intentat agafar imatges prou diverses entre elles. Una gran part coincideixen amb les usades als experiments que ja hem referenciat anteriorment. A les figures 6.1 i 6.2 donem les 40 imatges seleccionades.

6.2 Càlcul de la dimensió de l'espai de representació

A cada una de les imatges introduïdes a l'apartat anterior hem calculat la seva representació i hem aplicat el MDS per a calcular la dimensió correcta de l'espai. Els valors de l'*stress* que ens ha donat el MDS per diferents dimensions d'escalament són a la taula 6.1. No tenim els resultats de l'aproximació de Monte-Carlo pel cas de 40 punts, però si considerem els corresponents a 36 punts i que donem a la taula 5.16.a, podem veure que podrien representar conjunts semblants de dades. La dada més important és que l'escalament a dimensió 4 dóna un valor d'*stress* que és aproximadament de 0.05 i que per tant fa que poguem considerar que l'escalament a dimensió 4 és bo.

Aquest resultat està d'acord amb el resultat obtingut per les representacions de les textures de l'espai paramètric, per a les que havíem considerat totes les possibles variacions entre els paràmetres. Sembla doncs, que l'espai de representació que hem definit podrà ser interpretat com un espai general de representació de textures de 4 dimensions.

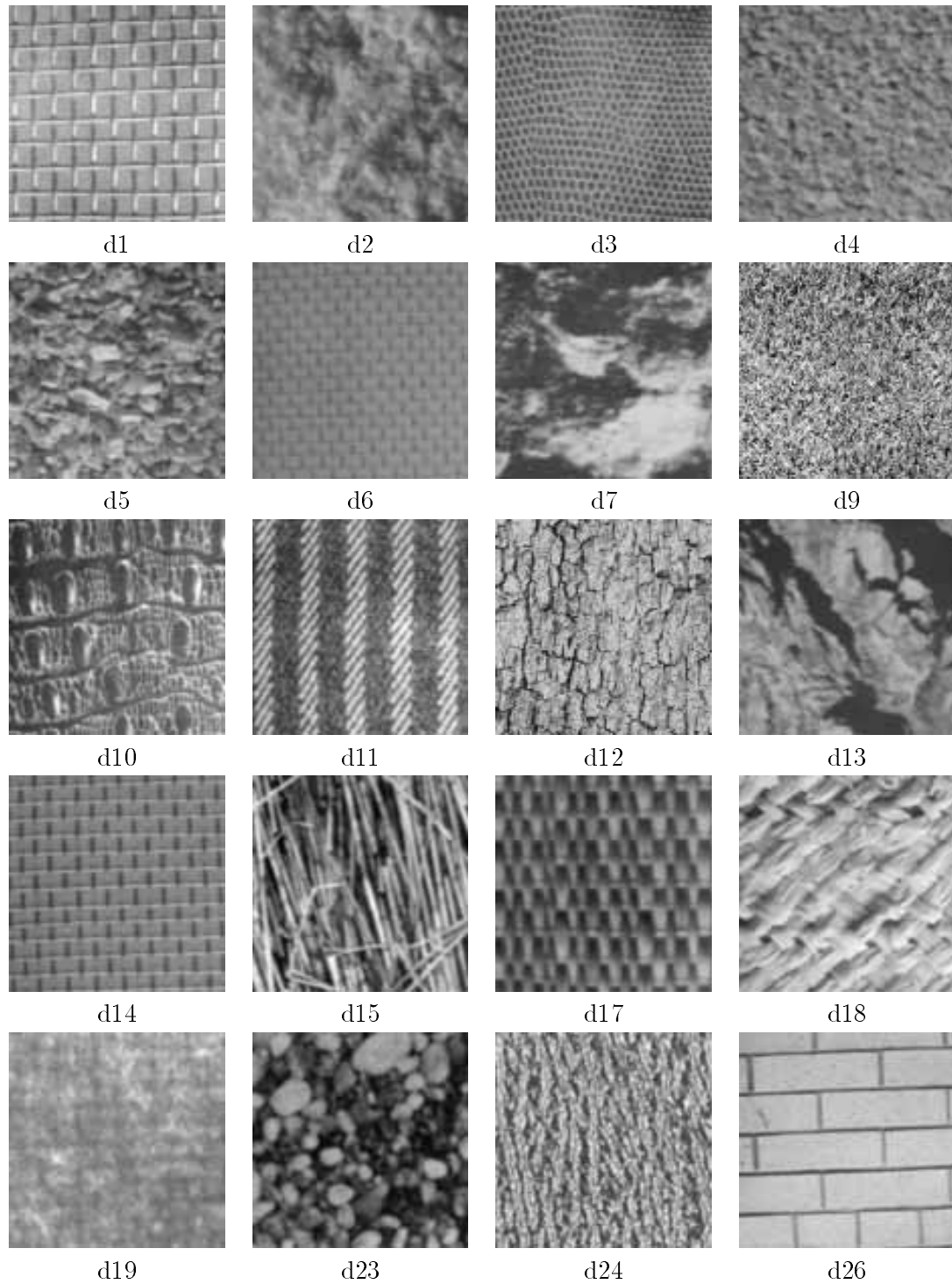


FIGURA 6.1: Conjunt d'imatges seleccionades de l'album de Brodatz (continua).

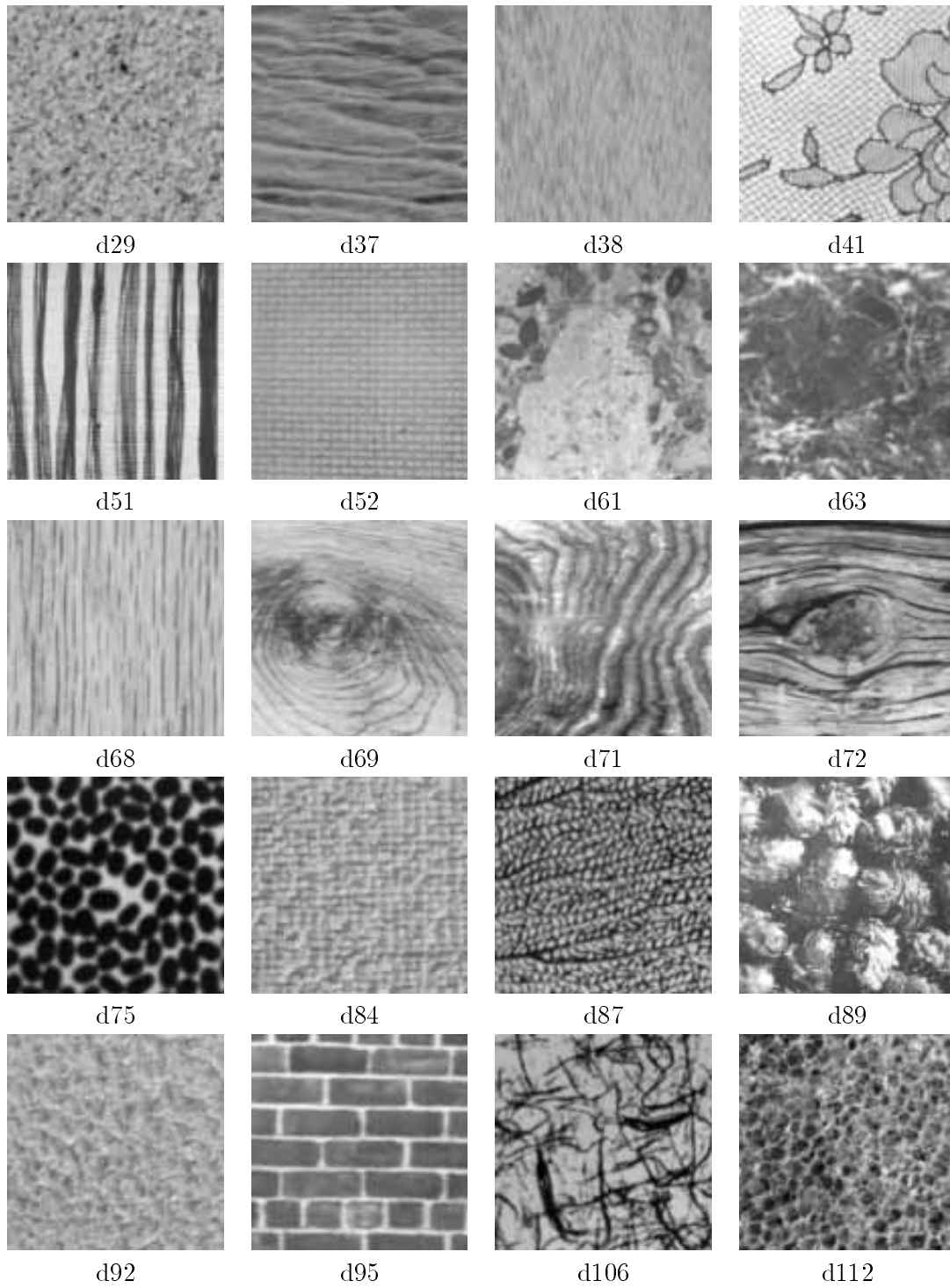


FIGURA 6.2: Conjunt d'imatges seleccionades de l'album de Brodatz.

D'aquesta manera hem demostrat el primer objectiu que ens havíem proposat. Hem vist que l'espai de representació pot ser reduït a un espai de dimensió 4.

6.3 Hipòtesi per a una interpretació basada en paràmetres perceptuals

Com ja hem argumentat és difícil parametritzar totes les textures a priori. Per a poder fer una parametrització que ens permeti una exploració sistemàtica hauríem d'aplicar un algorisme semblant al que es defineix al treball de H. Voorhees i T. Poggio [120] (apartat 5.2.1). Aquest algorisme ens permetria detectar de forma automàtica barres i taques, per a les que calcularíem tots els seus atributs. Aleshores podríem parametritzar les textures abans de calcular la seva representació construint la parametrització en termes de l'expressió 5.4. Ara bé, ja hem argumentat que tot i tenir aquesta parametrització l'exploració de la manera que l'hem feta abans seria complexa i pràcticament inviable.

Una solució més adient per a poder obtenir una representació paramètrica general per a tots els tipus de textons de la imatge, es basa en extreure els paràmetres de la pròpia representació calculada. Per això, podríem interpretar el significat de cada un dels canals en els que es basa la representació i a partir d'aquesta interpretació extreure la parametrització.

Vegem un exemple concret de com farem aquesta interpretació: respostes altes o ben sintonitzades en un canal que correspon a un filtre del tipus $DOOG2(\sigma, \theta)$ suposarem que ens està detectant textons que són barres orientades en la direcció θ i que tenen una amplitud de mida σ , a més a més, el valor de la resposta en aquesta coordenada de la representació representarà el valor de contrast d'aquest textó.

Com ja hem dit abans els detectors que s'apliquen en el càlcul de la representació són independents de la longitud de les barres. Per tant eliminarem el paràmetre de longitud de la interpretació de l'espai de baixa dimensió.

Una vegada obtinguem aquesta representació paramètrica per a cada textura del conjunt mostra seleccionat podrem fer la interpretació de l'espai a partir d'un anàlisi de correlació. Aquesta anàlisi es veurà més endavant, abans definirem la parametrització de la representació.

6.3.1 Extracció de paràmetres a partir de la representació

Donada la textura t la seva representació ve donada per $r(t)$ tal com hem definit a 3.30. El vector $r(t)$ és un vector de 192 valors. Cada un d'aquests valors correspon a un canal

del model i com ja hem dit representa la sintonització d'un filtre determinat a una escala concreta, seguit d'una rectificació de mitja ona que fa que la resposta correspongui bé a la resposta positiva o bé a la resposta negativa.

Inicialment tenim 8 tipus de filtres diferents:

$(DOG1, DOG2, DOOG2-0^\circ, DOOG2-30^\circ, DOOG2-60^\circ, DOOG2-90^\circ, DOOG2-120^\circ, DOOG2-150^\circ)$

Per a cada un d'aquests hem derivat un canal per a la part positiva (+) i un per a la negativa (-). I per a cada tipus de filtre tenim 12 escales diferents. Per tant, els 192 canals es poden dividir en 16 grups, cada un d'ells amb 12 valors, si reordenem la representació de la següent manera:

$$r(t) = (v_1, \dots, v_{192}) \longrightarrow r'(t) = (v_1^1, \dots, v_{12}^1, \dots, v_1^{16}, \dots, v_{12}^{16}) \quad (6.1)$$

on

$$v^k = f(R_k^+) \quad \forall k : 1 \leq k \leq 8 \quad (6.2)$$

$$v^k = f(R_k^-) \quad \forall k : 8 < k \leq 16 \quad (6.3)$$

aleshores $r'(t)$ representa les 12 respostes que corresponen a cada un dels filtres, seguint l'ordre següent:

$(DOG1^+, DOG2^+, DOOG2-0^\circ+, DOOG2-30^\circ+, DOOG2-60^\circ+, DOOG2-90^\circ+, DOOG2-120^\circ+, DOOG2-150^\circ+, DOG1^-, DOG2^-, DOOG2-0^\circ-, DOOG2-30^\circ-, DOOG2-60^\circ-, DOOG2-90^\circ-, DOOG2-120^\circ-, DOOG2-150^\circ-)$

A partir de $r'(t)$ extreurem els paràmetres perceptuals de contrast, escala i orientació que ens han de permetre interpretar l'espai de representació de la mateixa manera que ho fem per imatges de l'espai paramètric.

Per a cada un dels 16 grups de $r'(t)$, que venen donats pel subconjunt (v_1^k, \dots, v_{12}^k) , definim la seva parametrització

$$(c_k, s_k, \theta_k) : \forall k \in \{1, \dots, 16\} \quad (6.4)$$

on

$$c_k = \delta(k) \max_{i=1 \dots 12} v_i^k \quad : \quad \delta(k) = \begin{cases} 1 & 1 \leq k \leq 8 \\ -1 & 9 \leq k \leq 16 \end{cases} \quad (6.5)$$

$$s_k = s \quad : \quad v_s^k = \max_{i=1 \dots 12} v_i^k \quad (6.6)$$

$$\theta_k = \begin{cases} \emptyset & \text{Si } k \in \{1, 2, 9, 10\} \\ 0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ & \text{Si } k \in \{3, 11\}, \{4, 12\}, \{5, 13\}, \{6, 14\}, \\ & \{7, 15\}, \{8, 16\} \text{ respectivament} \end{cases} \quad (6.7)$$

$$(6.8)$$

Per tant la parametrització de la textura a partir de la seva representació vindrà donada per

$$p(t) = i(p_1^1, p_1^2, p_1^3, \dots, p_{16}^1, p_{16}^2, p_{16}^3) = i(c_1, s_1, \theta_1, \dots, c_{16}, s_{16}, \theta_{16}) \quad (6.9)$$

que ve a ser una aproximació de la parametrització que havíem intuït a l'expressió 5.4. Aquesta nova representació forma un espai paramètric de 16×3 dimensions.

Com que l'exploració d'aquest espai paramètric continua sent impossible com ho hem fet abans, anem a definir una altra manera d'extendre la interpretació de l'espai al cas general.

El que proposem és definir una hipòtesi sobre com es poden combinar els diferents paràmetres d'una textura. Aleshores a partir d'aquesta hipòtesi podrem construir un nou espai pel que haurem de veure si es correlaciona amb la reducció de l'espai de representació.

Vegem doncs, la hipòtesi de combinació dels paràmetres perceptuals dins la representació:

Hipòtesi 6.1 *Una textura t és el resultat de sumar els punts que representen cada una de les textures simples $\{t_1, \dots, t_q\}$ que la formen. Cada textura simple, t_i , ve donada per un valor de contrast, escala i orientació, que denotem com (c_i, s_i, θ_i) . El punt a l'espai de representació de 4 dimensions que representa la textura t_i és el punt*

$$b(t_i) = (g(s_i, c_i) \cos(2\theta_i), g(s_i, |c_i|) \sin(2\theta_i), \text{sgn}(c_i) \cdot g_5(|c_i|), g_6(|c_i|)) \quad (6.10)$$

si $\theta_i \neq \emptyset$, en altre cas el punt ve donat per

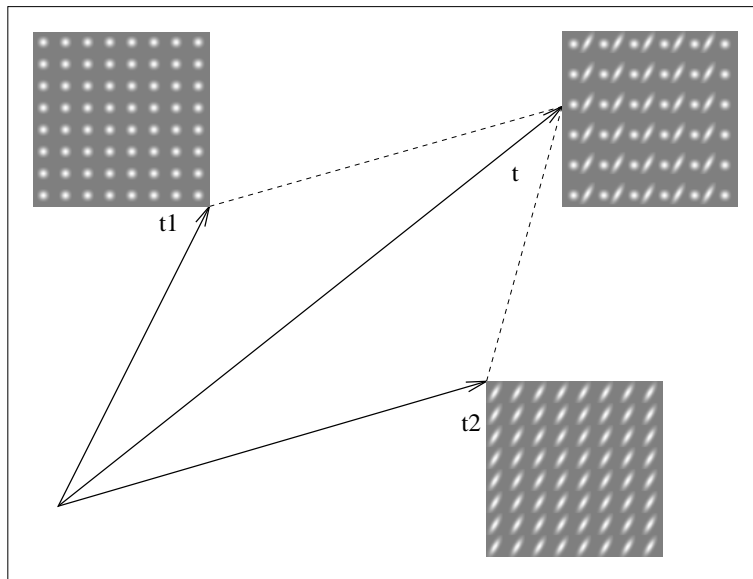


FIGURA 6.3: Il·lustració de la hipòtesi 6.1

$$b(t_i) = (g(s_i, c_i), g(s_i, |c_i|), \text{sgn}(c_i) \cdot g_5(|c_i|), g_6(|c_i|)) \quad (6.11)$$

aleshores la textura t ve donada a l'espai de representació pel punt

$$b(t) = \sum_{i=1}^q b(t_i) \quad (6.12)$$

Per tant amb aquesta hipòtesi estem dient que la representació de tota textura és el resultat de sumar les representacions de les textures que la formen. Cada una d'aquestes textures ve definida per un únic tipus de textó que ve determinat pels seus atributs d'escala, contrast i orientació.

6.3.2 Construcció de l'espai definit a la hipòtesi

Ara que tenim una parametrització de cada una de les textures que hem seleccionat, i tenint en compte la hipòtesi de combinació d'aquestes a l'espai de representació per a formar una textura general, podem calcular per a cada representació d'una textura les seves dimensions dins d'aquest espai de 4 dimensions.

Els passos que seguirem per a construir l'espai de 4 dimensions segons la hipòtesi són els següents:

Pas 1 Per a una textura, t , calcular la seva representació $r(t)$.

Pas 2 Donat el vector $r(t)$ ordenar-lo per obtenir el vector $r'(t)$.

Pas 3 Parametritzar cada un dels 16 subgrups de $r'(t)$, seguint les expressions de 6.8 de manera que obtenim el vector $p(t)$.

Pas 4 Per a cada subconjunt de la forma (c_i, s_i, θ_i) de $p(t)$ que representa a la subtextura t_i , calculem el seu punt corresponent a l'espai de 4 dimensions segons les expressions 6.10 i 6.11, per tant obtenim $b(t_i) = (b_1(t_i), b_2(t_i), b_3(t_i), b_4(t_i))$.

Pas 5 Les dimensions de t a l'espai de 4 dimensions vindran donades per

$$b(t) = \sum_{i=1}^{16} (b_1(t_i), b_2(t_i), b_3(t_i), b_4(t_i)) \quad (6.13)$$

Les dimensions de l'espai venen donades per les següents expressions

$$(g(s_i, c_i) \cos(2\theta_i), g(s_i, |c_i|) \sin(2\theta_i), \operatorname{sgn}(c_i) \cdot g_5(|c_i|), g_6(|c_i|)) \quad (6.14)$$

on sabem que g , g_5 i g_6 són funcions monòtones, encara que desconexem la seva forma exacta. Per tant, per a poder constuir l'espai de la hipòtesi haurem d'assumir una forma per elles que mantingui la propietat de la monotoneïtat. Farem doncs, la següent assumpció:

$$g(x, y) = x \cdot y \quad (6.15)$$

$$g_5(x) = g_6(x) = x \quad (6.16)$$

Si la hipòtesi és certa (a la figura 6.3 podem veure una il.lustració d'aquesta), les posicions, $b(t)$, de les textures no haurien d'estar gaire lluny de les que resulta de reduir amb MDS l'espai de representació a un espai de 4 dimensions. Anem a veure l'anàlisi de correlació entre aquests dos espais.

6.4 Confirmació de la hipòtesi

Per a confirmar la hipòtesi treballarem amb el conjunt d'imatges que hem introduït a l'apartat 6.1 i que denotarem com $\{t^1, \dots, t^{40}\}$. Per a cada una d'aquestes calcularem la seva representació, $r(t^i)$. Aplicant el MDS sobre el conjunt de distàncies definides per totes aquestes representacions escalem l'espai de representació a un espai de 4 dimensions obtenint les coordenades, (x_1, x_2, x_3, x_4) , per a cada textura t^i .

Per altra banda, construïm l'espai de la hipòtesi seguint els passos que hem definit a l'apartat anterior. Per tant, per a cada textura de $\{t^1, \dots, t^{40}\}$ obtenim les seves coordenades a l'espai de la hipòtesi, aquestes coordenades les denotarem com (b_1, b_2, b_3, b_4) .

Per veure la correspondència entre aquests dos espais farem primer una anàlisi de correlació i a la vista dels resultats plantejarem un estudi basat en un anàlisi de regressió lineal múltiple.

6.4.1 Anàlisi de correlació

Per a analitzar la correlació entre aquests dos espais veurem quins són els coeficients de correlació de *Pearson* entre les coordenades de cada espai sobre el conjunt de les 40 imatges analitzades. Aquest coeficient, R , és la mesura més usada normalment per quantificar l'associació lineal entre dues variables, X i Y , ve donada per la següent expressió:

$$R = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{(N - 1)S_X S_Y} \quad (6.17)$$

on N és el nombre de casos, en el nostre cas 40, S_X i S_Y són les desviacions estàndards de les dues variables. Aleshores, el valor absolut de R indicarà la importància d'aquesta relació lineal. A la figura 6.4 podem veure alguns exemples del comportament d'aquest coeficient. Quan l'associació lineal és perfecte aleshores $R = 1$ i la pendent és positiva, quan l'associació és perfecte però invertida, aleshores $R = -1$ i la pendent de la recta és negativa. A les figures 6.4.c i 6.4.d podem veure dos exemples en els que l'associació lineal és impossible, aleshores $R = 0$.

Els resultats de calcular aquest coeficient de correlació entre les coordenades de les imatges a l'espai de representació escalat a 4 dimensions, (x_1, x_2, x_3, x_4) , i les coordenades d'aquestes imatges a l'espai de la hipòtesi, (b_1, b_2, b_3, b_4) , es poden veure a la taula 6.2.

Els coeficients que relacionen els dos espais són els de la part superior dreta de la taula. Els valors de correlació més importants venen donats entre les següents variables:

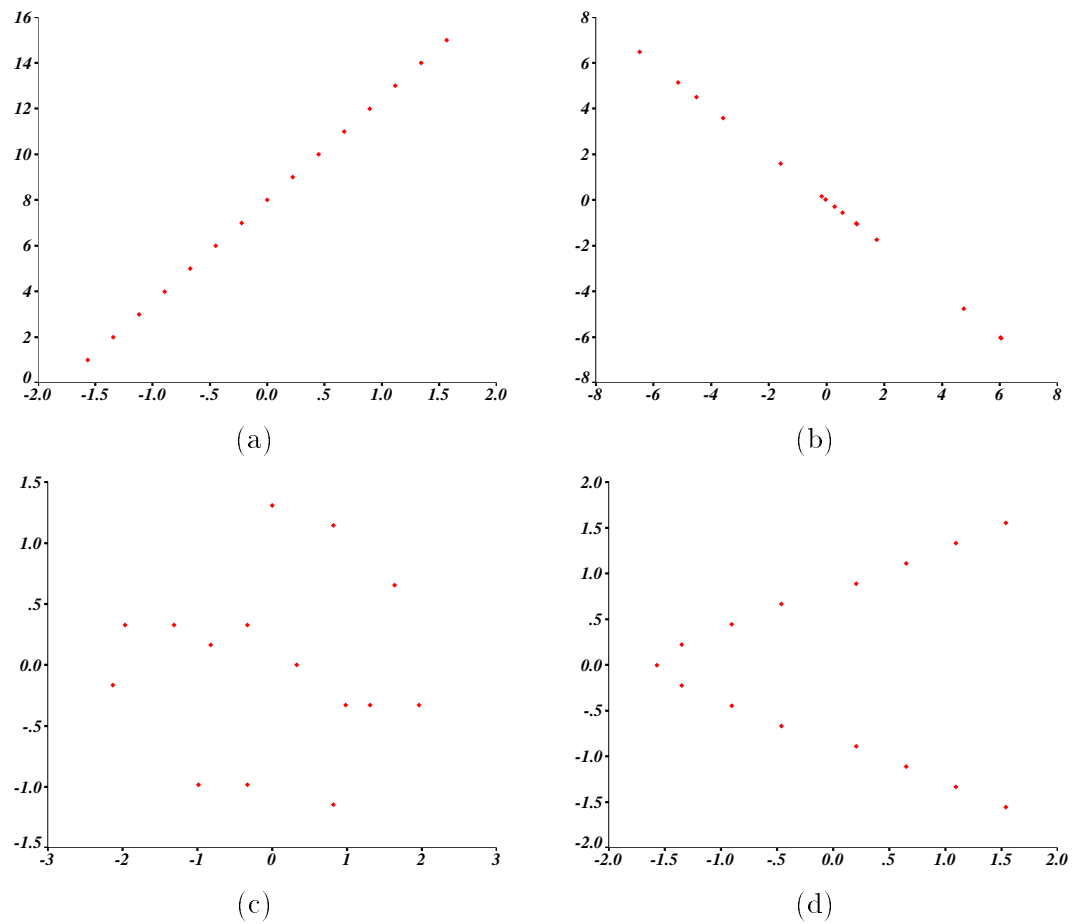


FIGURA 6.4: Exemples d'associacions lineals amb diferents coeficients de correlació de *Pearson*. (a) $R = 1$. (b) $R = -1$. (c) $R = 0$. (d) $R = 0$.

	x_1	x_2	x_3	x_4	b_1	b_2	b_3	b_4
x_1	1.0000 P= .	-.0239 P= .883	-.0484 P= .767	.0145 P= .929	-.4437 P= .004	-.7618 P= .000	.1845 P= .255	-.6949 P= .000
x_2	-.0239 P= .883	1.0000 P= .	.0120 P= .942	-.0570 P= .727	.0448 P= .784	.2111 P= .191	-.1230 P= .449	-.6009 P= .000
x_3	-.0484 P= .767	.0120 P= .942	1.0000 P= .	-.0013 P= .993	-.0449 P= .783	.2212 P= .170	.5003 P= .001	-.0267 P= .870
x_4	.0145 P= .929	-.0570 P= .727	-.0013 P= .993	1.0000 P= .	.7418 P= .000	-.1040 P= .523	.3644 P= .021	-.0912 P= .576
b_1	-.4437 P= .004	.0448 P= .784	-.0449 P= .783	.7418 P= .000	1.0000 P= .	.3549 P= .025	.0653 P= .689	.2197 P= .173
b_2	-.7618 P= .000	.2111 P= .191	.2212 P= .170	-.1040 P= .523	.3549 P= .025	1.0000 P= .	-.2739 P= .087	.5220 P= .001
b_3	.1845 P= .255	-.1230 P= .449	.5003 P= .001	.3644 P= .021	.0653 P= .689	-.2739 P= .087	1.0000 P= .	-.1176 P= .470
b_4	-.6949 P= .000	-.6009 P= .000	-.0267 P= .870	-.0912 P= .576	.2197 P= .173	.5220 P= .001	-.1176 P= .470	1.0000 P= .

TAULA 6.2: Coeficients de correlació de *Pearson* entre les diferents dimensions de l'espai de representació escalat a 4 dimensions i l'espai de la hipòtesi. El valor de P representa la probabilitat de que el coeficient de correlació canviï.

	b_1	b_2	b_3	b_4
x_1	-	-0.7618	-	-0.6949
x_2	-	-	-	-0.6009
x_3	-	-	0.5003	-
x_4	0.7418	-	-	-

A la vista d'aquests resultats podem dir que sembla haver-hi correlació entre certes variables. Tenint en compte que les coordenades dels punts que dona el MDS poden haver sofert qualsevol transformació a l'espai les dimensions d'un i altre espai no tenen perquè coincidir. Això es veu en el fet de que la coordenada x_1 està pràcticament igual correlacionada amb b_2 i b_4 . Per evitar aquest problema hem fet un anàlisi basat en una regressió lineal múltiple de l'espai de la hipòtesi respecte del conjunt de variables de l'espai de representació.

6.4.2 Anàlisi basat en una regressió lineal múltiple

Per a l'estudi dels eixos d'interés d'un espai escalat pel MDS, a [60] i a [90] es planteja el problema com el de calcular una regressió lineal múltiple de la variable que es vol identificar com eix de l'espai, respecte de les variables independents que formen els eixos coneguts de l'espai.

Per tant la regressió lineal múltiple busca la recta de l'espai que permetria correlacionar-se més amb els valors d'una determinada variable. Donada una variable b_i que representa una propietat, i donat el conjunt de variables $\{x_1, \dots, x_n\}$ que representen l'espai escalat, volem veure si existeix un eix dins l'espai $\{x_1, \dots, x_n\}$ que es correlacioni amb la propietat b_i . Per fer això es construeix la recta

$$b_i \cong \bar{b}_i = B_0 + B_1x_1 + \dots + B_nx_n \quad (6.18)$$

tenint en compte que b_i i $\{x_1, \dots, x_n\}$ representen variables que poden agafar diferents valors depenent del nombre de casos que es tinguin a l'espai.

En el nostre cas tenim que les rectes de regressió resultants d'aquest procés són les següents, per a cada una de les dimensions que volem trobar b_1 , b_2 , b_3 i b_4 :

$$b_1 \cong \bar{b}_1 = 166.61 - 52.30x_1 + 11.48x_2 - 12.66x_3 + 150.01x_4 \quad (6.19)$$

$$b_2 \cong \bar{b}_2 = 196.60 - 81.05x_1 + 26.01x_2 + 32.66x_3 - 15.47x_4 \quad (6.20)$$

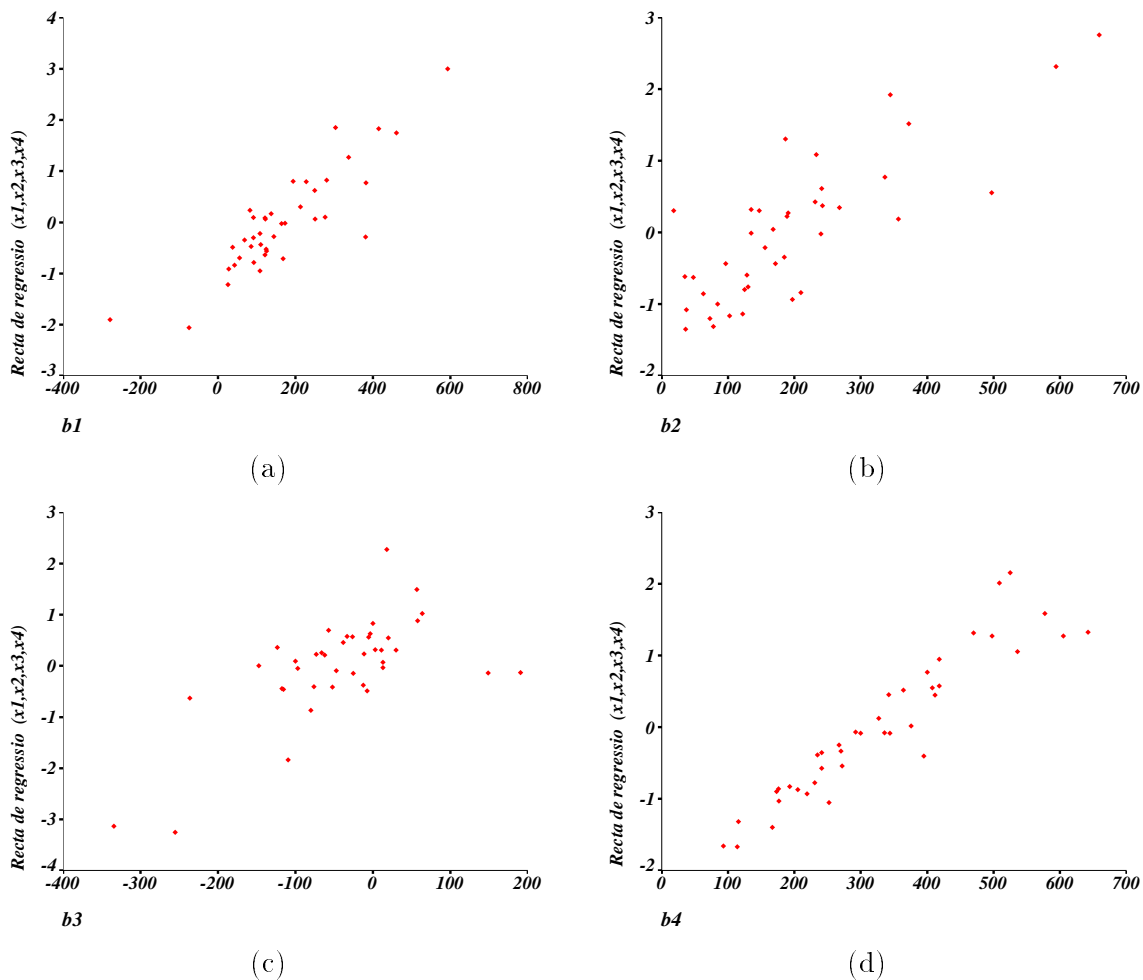


FIGURA 6.5: Associació lineal entre la dimensió b_i de l'espai i la recta de regressió calculada \bar{b}_i .

$$b_3 \cong \bar{b}_3 = -41.97 + 14.64x_1 - 9.72x_2 + 61.25x_3 + 44.87x_4 \quad (6.21)$$

$$b_4 \cong \bar{b}_4 = 328.40 - 75.47x_1 - 85.26x_2 - 9.41x_3 - 21.46x_4 \quad (6.22)$$

A la figura 6.5 podem veure les gràfiques de l'associació lineal entre els valors de b_i i els valors de les rectes \bar{b}_i . I els valors dels coeficients de correlació de *Pearson*, R , donats per aquesta relació són:

b_1	b_2	b_3	b_4
0.875	0.810	0.660	0.938

D'aquesta manera millorem els resultats de l'anàlisi de correlació que havíem obtingut a

l'apartat anterior. I per tant podem donar per confirmada la hipòtesi, tot i que hi ha certs valors de correlació com els de la variable b_4 que s'haurien de millorar. Aquesta millora es podria intentar dur a terme canviant la forma de la funció g_5 .

6.4.3 Estudi d'altres variables

Un altre argument que pot ajudar a corroborar la hipòtesi ve donat pel fet de que no hem trobat variables definides sobre la parametrització de la representació que s'hagin pogut correlacionar amb algun eix de l'espai de representació escalat.

Hem definit variables que representéssin certes característiques relacionades amb els paràmetres definits, i aleshores hem intentat correlacionar aquestes amb les coordenades de l'espai de representació $\{x_1, x_2, x_3, x_4\}$. No n'hem trobat cap que es correlacionés millor que el que ens defineix l'espai de la hipòtesi.

Vegem algunes de les variables que hem estudiat sense obtenir bons resultats:

- Escala del canal que conté el valor de màxima resposta pels filtres isotròpics.
- Escala del canal que conté el valor de màxima resposta pels filtres orientats.
- Quocient entre el màxim contrast del filtres orientats i el segon màxim d'aquests.
- Desviació estàndard dels valors màxims de contrast dels filtres orientats.
- Sinus de l'angle (multiplicat per 2) que representa el filtre de màxima resposta.
- Cosinus de l'angle (multiplicat per 2) que representa el filtre de màxima resposta.
- La diferència entre els mòduls de les representacions dels canals que corresponen a rectificacions positives i els que corresponen a rectificacions negatives.
- La diferència entre el valor del màxim contrast de canals positius i el valor del màxim contrast de canals negatius.

La dimensió del contrast, com l'hem definit, no deixa cap marge de dubte sobre el fet de que representa un eix important de l'espai de representació. Per altra banda s'ha intentat veure si l'escala per ella mateixa podia ser un eix de l'espai obtenint valors de correlació molt baixos.

On s'ha trobat més dificultat ha estat a l'hora de trobar variables que relacionessin el paràmetre de l'orientació, fet que fa valorar més els valors obtinguts per les dimensions representades amb les variables b_1 i b_2 de l'espai de la hipòtesi.

Finalment, no hem trobat cap mesura millor que intentés reflectir el comportament de la representació respecte del signe de les taques i les barres, que és el que pretenien les dues darreres variables definides a la llista. Aquest eix continua sent el que deixa més dubtes sobre la interpretació feta.

6.5 Conclusions

Quan hem introduït aquest treball i hem plantejat la necessitat d'una representació per a les textures, hem comparat aquest fet amb un altre estímul visual com el color i amb l'existència de representacions basades en espais de poques dimensions amb una interpretació perceptual associada.

La proposta de representació que fem en aquest treball podria comparar-se a aquest espai del color. Una representació de 4 dimensions on cada eix representa una funció d'una característica interpretable perceptualment.

Els resultats obtinguts ens han permès establir una certa estructura respecte d'uns paràmetres dins l'espai de representació. Aquesta estructura defineix l'espai psicomètric definit a [55, 56]

Espai psicomètric és l'espai resultant d'aplicar un algorisme que faci la correspondència entre una imatge d'entrada i un punt dins d'aquest espai.

L'ús d'aquest espai psicomètric és el de veure la correspondència aproximada entre aquest i el corresponent espai perceptual

Espai perceptual ha estat definit per R.N. Shepard [98] i és l'espai generat a partir dels judicis de similitud fets per persones davant d'un conjunt d'estímuls.

Per tant, si suposem un conjunt d'imatges com estímuls, un conjunt de persones i un algorisme que pretén dur a terme el mateix objectiu que les persones davant dels estímuls es poden construir els corresponents espais perceptuals i psicomètrics.

Això deixa veure una línia de continuació en aquest sentit que ens permetés respondre a la qüestió sobre la correspondència entre aquest espai de representació de textures amb l'espai perceptual del SVH.

L'aportació més interessant d'aquest resultat no és el resultat en si, sinó que és el plantejament d'unes idees que obren un camí per a treballar cap a una representació general o com a mínim útil de les textures.

Amb aquest resultat volem demostrar la viabilitat d'un espai de representació general de textures segons els següents criteris:

Basada en models perceptuals fet que suposa que haurà de partir de la representació que dona un filtratge multicanal que simuli els detectors de característiques del SVH, i que per tant es basarà en processos de visió primària o preatentius.

Espai de dimensió baixa es confirma la possibilitat de que la representació de textures es pugui fer en poques dimensions com ja havíem fet R. Rao i G.L. Lohse a la seva aproximació.

En el capítol següent donem alguns resultats sobre el comportament d'aquesta representació sobre dues bases de dades d'imatges. Tot i que els resultats que obtenim són prou bons el que caldria fer ara es provar la validesa perceptual d'aquests. Aquesta validació podria descobrir les mancances d'aquesta, ja que aquesta representació hauria de ser millorada en diversos aspectes:

- Búsqueda d'una funció que sigui capaç de representar com es combina el signe de les taques d'una mateixa textura.
- Extendre evidentment aquesta representació per a que pugui incloure el color de la textura.
- Treballar sobre el problema que planteja el radi de la inhibició, intentant introduir en el mètode alguns dels resultats de la teoria del textó.

CAPÍTOL 7

APLICACIÓ DE LA REPRESENTACIÓ PER A L'ACCÉS A BASES DE DADES

Amb l'objectiu de comprovar el comportament de la representació definida sobre imatges naturals, en aquest capítol plantegem dos experiments sobre grans bases de dades d'imatges accessibles per tothom. Donem els resultats de tots aquests experiments i fem una avaluació del rendiment de la representació respecte d'altres treballs fets sobre aquest mateix tema. Finalment fem una valoració d'aquesta avaluació i fem notar alguns aspectes que s'haurien de millorar per a obtenir un millor anàlisi del rendiment.

7.1 Accés a bases de dades pel contingut de la imatge

Accedir a grans bases de dades d'imatges per a trobar imatges amb determinats continguts és un problema important en els sistemes actuals d'emmagatzemament. Degut a la importància d'aquest fet s'han desenvolupat interessants projectes amb aquests objectius:

QBIC [26] (*Query by Image Content*) Aquest sistema permet recuperar imatges i vídeos. És capaç de calcular característiques que representin informació sobre el color, la textura, la forma, la posició o el moviment, i trobar imatges a partir d'exemples de manera que siguin perceptualment similars.

Photobook [74] En aquest sistema es planteja una compressió de la imatge que preservi tot el seu significat, depenent del tipus d'imatge l'algorisme de compressió s'hauria de comportar de forma diferent. En cap cas no es pretén una representació simbòlica de tot el contingut de la imatge que pugui ser demanat.

Chabot [71] Integra l'ús de text sobre la imatge amb la capacitat de recerca basada amb l'anàlisi del color de la imatge.

Piction [104] És un sistema especialitzat en trobar cares en fotografies dels diaris i fer la seva identificació a partir dels peus que acompanyen les fotografies.

Una propietat comú a a molts d'aquests sistemes de recerca d'imatges, és la capacitat de dirigir la recerca a partir de les propietats texturals de la imatge.

A. Pentland i els seus col·legues defineixen a [74] quines són les propietats que ha de tenir una representació d'imatges per a que sigui vàlida per aquest tipus d'aproximacions. Aquestes propietats han estat comentades al capítol 1 d'aquest treball.

7.2 Resultats experimentals

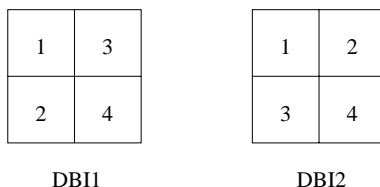
Per a investigar la capacitat de la representació dissenyada, hem fet dos experiments d'accés a bases de dades per proximitat entre la representació. Per dissenyar els experiments ens hem basat en dues bases de dades d'imatges (BDI), de textures més concretament, que especifiquem als apèndixs. Per a cada imatge hem mantingut el seu nom original en anglès.

Per realitzar els experiments hem calculat les representacions de cada una de les textures de la BDI. Aleshores en cada prova ens plantejarem quines són les textures més properes, dins de l'espai de representació, d'una textura determinada que indicarem amb (?).

La primera BDI és la que referenciem com BDI1 i es dona a l'apèndix A. Consisteix en un conjunt de 82 imatges dividides en tres grups organitzats segons la semàntica del contingut de les imatges. Són imatges de dimensions 512×512 , que després de fer una reducció de dimensió per promig a 256×256 han estat partides en 4 parts no solapades, formant imatges de 128×128 . Per tant hem construït un conjunt de $4 \times 82 = 328$ imatges.

Per altra banda, la segona base de dades és la que referenciem com BDI2 i que es pot veure a l'apèndix B. Està formada per un conjunt de 167 imatges de dimensions 512×512 . Com abans, hem fet una reducció de dimensió per promig a 256×256 , i aquestes han estat partides en 4 parts no solapades, formant imatges de 128×128 . Per tant hem construït un conjunt de $4 \times 167 = 668$ imatges.

Cal tenir en compte que la partició d'imatges de 256×256 a 4 imatges de 128×128 s'ha fet de diferent manera per a cada BDI:



Aquestes imatges es referenciaran a partir del nom donat als apèndixs seguit del número que indica la partició a la que correspon.

7.2.1 Experiment 1

Aquest experiment ha estat realitzat a partir de les imatges de la base de dades BDI1. Els resultats es donen a les taules 7.1, 7.2 i 7.3. En aquestes taules indiquem el nom de la imatge qüestionada, i per a cada part d'aquesta 1, 2, 3 i 4 donem respectivament les imatges més properes a elles que s'han trobat a la BDI1. A la taula donem la primera, la segona i tercera textura més propera.

A les figures 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 7.10, 7.11, 7.12, 7.13, 7.14, donem alguns exemples d'aquests resultats.

En aquests exemples es pot veure el concepte de preatentivitat de la representació que fem. Podem veure textures semblants si tenim en compte les propietats relacionades amb les taques de la imatge, contrast, escala, orientació, sense fer cas del contingut de la imatge.

Un exemple que il·lustra aquest fet és el de la figura 7.10, en els que entre les imatges més properes a la qüestionades tenim plantes i fruites alhora, en ambdós casos les textures que formen es poden veure preatentivament molt semblants.

7.2.2 Experiment 2

Aquest experiment ha estat realitzat a partir de les imatges de la base de dades BDI2. Els resultats es donen a les taules 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 7.10, 7.11, 7.12, 7.13, 7.14, 7.15, 7.16, 7.17, 7.18, 7.19, 7.20, 7.21, 7.22. Els noms de les imatges s'han substituït per les inicials. Com a l'anterior experiment, en aquestes taules indiquem el nom de la imatge qüestionada, i per a cada part d'aquesta 1, 2, 3 i 4 donem respectivament les imatges més properes a elles que s'han trobat a la BDI1. A la taula donem la primera, la segona i tercera textura més propera.

A les figures 7.15, 7.16, 7.17, 7.18, 7.19, 7.20, donem alguns exemples d'aquests resultats.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
f1(1,2,3,4)	f9,f15,f1,f1	f9,f1,f15,f1	f15,f9,f1,f9
f2(1,2,3,4)	f2,f2,p21,f2	f2,f2,f18,f2	p21,p21,f2,f2
f3(1,2,3,4)	f3,f3,f3,f3	f3,f3,f10,f3	f3,f3,f3,f3
f4(1,2,3,4)	f4,f4,f4,f8	f4,f4,p19,f15	f7,f10,f4,f8
f5(1,2,3,4)	f5,f5,f5,f5	f5,f5,f5,f5	f5,f5,f5,f5
f6(1,2,3,4)	f6,f6,f6,f6	f6,f6,f6,f6	f6,f6,p7,f6
f7(1,2,3,4)	f7,f4,f13,f13	f13,f17,f7,f7	f7,f7,f12,f7
f8(1,2,3,4)	f15,f8,f5,f8	f17,f15,f8,f15	f11,f8,f15,f8
f9(1,2,3,4)	f9,f9,f9,f9	f9,f8,f1,f15	f3,f15,f15,f15
f10(1,2,3,4)	f10,f10,f3,f10	f10,f15,f3,f17	f17,f10,p21,f8
f11(1,2,3,4)	f11,f18,f11,f5	f11,f11,f11,f11	f11,f11,f5,f11
f12(1,2,3,4)	f12,f12,f12,f12	f12,f13,f13,f12	m17,f1,f12,f13
f13(1,2,3,4)	f13,f7,f13,f13	f12,f13,f13,f1	f13,f12,f7,f13
f14(1,2,3,4)	f14,f14,f14,f16	f16,f16,f14,f14	f16,f14,f16,f16
f15(1,2,3,4)	f15,f8,f9,f17	f10,f11,f15,f17	f9,f8,f1,f15
f16(1,2,3,4)	f16,f16,f16,f16	f16,f16,f16,f16	f16,f14,f16,f16
f17(1,2,3,4)	f17,f17,f17,f17	f8,f15,f10,f8	f10,f15,p21,f15
f18(1,2,3,4)	p21,f18,f11,f11	p21,f18,f18,f5	f18,f11,f18,f5

TAULA 7.1: Resultats sobre BDI1 a partir de preguntes sobre les imatges de tipus f.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
p1(1,2,3,4)	p1,p1,p1,p1	p1,p1,p1,p1	p1,p1,p1,p1
p2(1,2,3,4)	p3,p2,p2,p2	p3,p2,p2,p13	p2,p5,p13,m13
p3(1,2,3,4)	p3,p3,p3,p3	p2,p3,p3,p3	p3,p3,p2,p3
p4(1,2,3,4)	p4,p4,p4,p4	m18,m26,m18,p4	m18,p4,m18,m26
p5(1,2,3,4)	p23,p23,p5,p13	p5,p15,p23,p5	p13,p15,m19,p5
p6(1,2,3,4)	p6,p6,p6,p6	p6,p6,p6,p6	p9,p6,p6,p6
p7(1,2,3,4)	p7,p7,p12,m23	p26,p26,p12,p7	p26,p11,m23,p12
p8(1,2,3,4)	p8,p8,p8,p8	p25,m12,p29,p8	p25,m12,p29,p25
p9(1,2,3,4)	p9,p9,p9,p9	p9,p9,p9,p27	p9,p27,p11,p11
p10(1,2,3,4)	p10,p10,p10,p10	p10,p10,p10,p10	p10,p10,p10,p10
p11(1,2,3,4)	p11,p7,p11,p11	p7,p26,p7,p11	p11,p9,p27,p7
p12(1,2,3,4)	p16,p12,p12,p7	m19,p7,p7,p12	p12,p27,p12,p26
p13(1,2,3,4)	p13,p13,p13,p13	p13,p5,p13,p5	p13,p13,p13,p23
p14(1,2,3,4)	p14,p14,p14,p14	p14,p14,p14,p14	p14,p14,p14,p14
p15(1,2,3,4)	p15,p15,p16,p15	p17,p16,p16,p5	p16,p16,p15,p22
p16(1,2,3,4)	p15,p16,p16,p16	p16,p15,p15,p5	p15,p5,p18,p15
p17(1,2,3,4)	p17,p15,p17,p17	p17,p17,p17,p16	p23,p17,p23,p16
p18(1,2,3,4)	p25,p18,p18,p18	p29,p18,p18,p18	p18,p16,p18,p16
p19(1,2,3,4)	p21,f10,p19,p19	p21,p19,p21,p19	f18,p21,p19,p21
p20(1,2,3,4)	p20,p19,f7,p19	m17,f16,f7,f17	m17,p19,f13,f17
p21(1,2,3,4)	p21,p21,p21,p21	p21,f2,p21,f18	p21,p21,f18,f10
p22(1,2,3,4)	p22,p22,p22,p22	p22,p22,p22,p22	p22,p22,m19,p22
p23(1,2,3,4)	p13,p5,m19,m19	p23,p15,p8,p5	p17,p15,p5,m19
p24(1,2,3,4)	p24,p24,p24,p24	p24,p24,p24,p24	p24,p15,p8,p22
p25(1,2,3,4)	p25,p25,p25,p25	p25,p25,p25,p25	p25,p25,p25,p25
p26(1,2,3,4)	p26,p26,p26,p7	m23,p12,m23,p7	p12,p11,p26,p26
p27(1,2,3,4)	p27,p27,p27,p27	p27,p27,p27,p7	p11,p27,p27,p30
p28(1,2,3,4)	p28,p28,p22,p28	p15,p28,p15,p23	p28,p26,p15,p5
p29(1,2,3,4)	p29,p29,p29,p29	p29,p29,p25,p29	p29,p29,p29,p31
p30(1,2,3,4)	p30,p30,p30,p30	p30,p30,p30,p30	p27,p30,p30,p30
p31(1,2,3,4)	m25,p29,m3,p31	m10,p31,m3,m3	p29,p29,m3,m3

TAULA 7.2: Resultats sobre BDI1 a partir de preguntes sobre les imatges de tipus p.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
m1(1,2,3,4)	m1,m1,m1,m1	m1,m1,m1,m1	m1,m1,m1,m1
m2(1,2,3,4)	m2,m2,m2,m2	m2,m2,m2,m2	m2,m2,m2,m2
m3(1,2,3,4)	m3,m3,m3,m3	m3,m3,m3,m3	m3,m3,m3,m3
m4(1,2,3,4)	m4,m29,m4,m4	m4,m29,m4,m4	m11,m24,m11,m11
m5(1,2,3,4)	m32,m5,m5,m5	m5,m5,m32,m5	m32,m32,m32,m32
m6(1,2,3,4)	m6,m6,m6,m6	m6,m6,m6,m6	m6,m6,m6,m6
m7(1,2,3,4)	m7,m7,m7,m7	m7,m7,m7,m7	m7,m7,m7,m7
m8(1,2,3,4)	m8,m8,m8,m8	m8,m8,m8,m8	m8,m8,m8,m8
m9(1,2,3,4)	m9,m9,m9,m9	m9,m9,m9,m9	m9,m9,m9,m9
m10(1,2,3,4)	m10,m10,m10,m10	m10,m10,m11,m11	m11,m11,m11,m10
m11(1,2,3,4)	m11,m11,m11,m11	m11,m11,m11,m10	m11,m10,m10,m11
m12(1,2,3,4)	m12,m12,m12,m12	m12,m12,m12,m12	m12,m12,m12,m12
m13(1,2,3,4)	m13,m13,m13,m13	m13,m13,f14,m13	m13,p2,f16,p2
m14(1,2,3,4)	m14,p2,m14,m14	m14,p2,m14,m13	p7,p2,p2,p2
m15(1,2,3,4)	m15,m15,m15,m15	m15,m15,m15,m15	m15,m15,m15,m15
m16(1,2,3,4)	m16,m16,m16,m16	m20,m16,m20,p29	m16,m20,m20,m20
m17(1,2,3,4)	m17,m17,m17,m17	m17,m17,m17,m17	f12,m17,f12,m17
m18(1,2,3,4)	m18,m18,m18,m18	m18,m18,m18,m18	m18,m18,m18,m18
m19(1,2,3,4)	m19,m19,m19,m19	p23,m19,m19,m19	m19,m19,p23,m19
m20(1,2,3,4)	m20,m20,m20,m20	m20,p3,m20,m19	m16,m16,m16,m16
m21(1,2,3,4)	m21,m21,m21,m21	m21,m21,m21,m21	m21,m21,m21,f12
m22(1,2,3,4)	m22,m22,m22,m22	m22,m22,m22,m22	m22,m32,m22,m32
m23(1,2,3,4)	m23,m23,m23,m23	m23,m23,m23,m23	p26,m23,m23,m23
m24(1,2,3,4)	m24,m24,m24,m24	m24,m24,m24,m24	m24,m24,m24,m24
m25(1,2,3,4)	m25,m25,m25,m25	m25,m25,m25,m11	m25,m25,m25,m25
m26(1,2,3,4)	m26,m26,m26,m26	m26,m26,m26,m26	m26,m18,m26,m5
m27(1,2,3,4)	m27,m27,m27,m27	m27,m27,m27,m27	m27,m27,m27,m27
m28(1,2,3,4)	m28,m28,m28,m28	m28,m28,m28,m28	m28,m28,m28,m28
m29(1,2,3,4)	m29,m29,m29,m29	m29,m29,m4,m4	m29,m29,m24,m24
m30(1,2,3,4)	m30,m30,m30,m30	m30,m30,m30,m30	m30,m30,m30,m30
m31(1,2,3,4)	m31,m31,m31,m31	m31,m31,m31,m31	m31,m31,m31,m31
m32(1,2,3,4)	m32,m32,m32,m5	m5,m32,m32,m32	m5,m32,m5,m5
m33(1,2,3,4)	m33,m33,m33,m33	m33,m33,m33,m7	m33,m7,m33,m32

TAULA 7.3: Resultats sobre BDI1 a partir de preguntes sobre les imatges de tipus m.

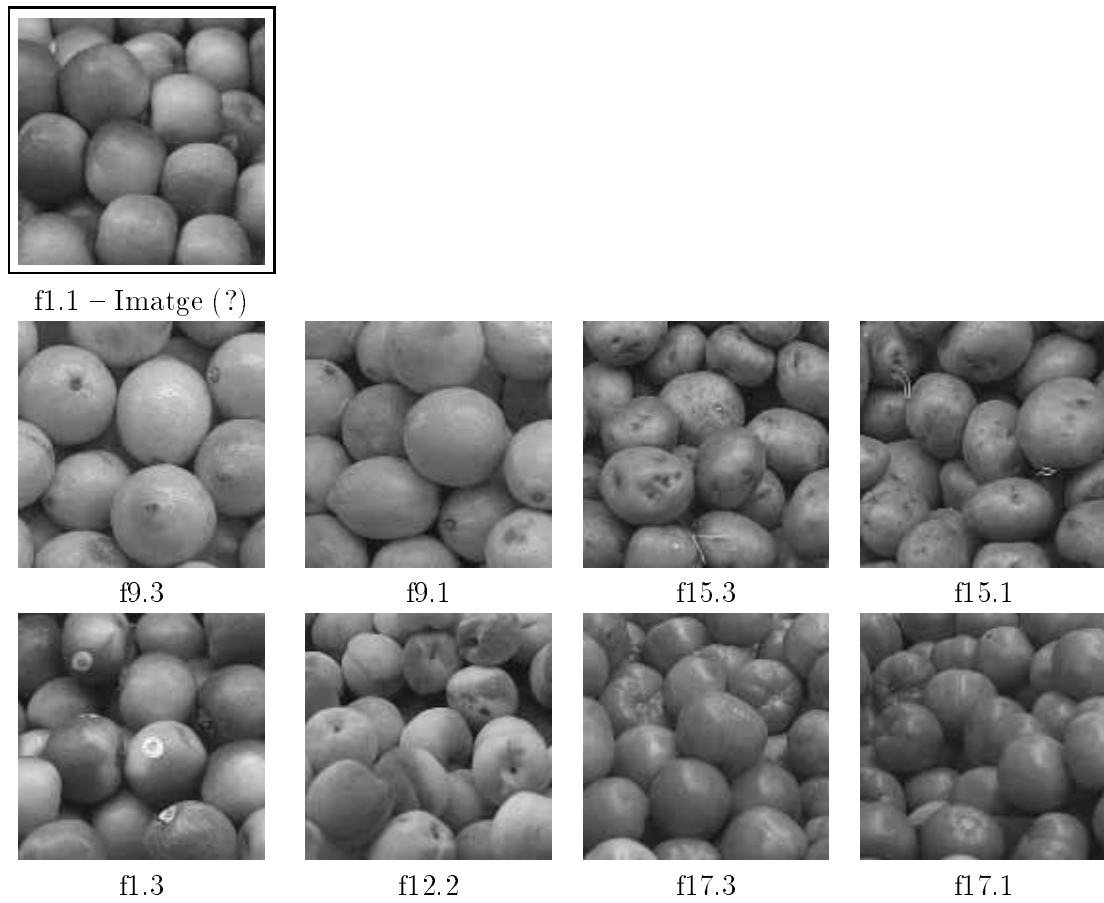


FIGURA 7.1: Experiment sobre DBI1.

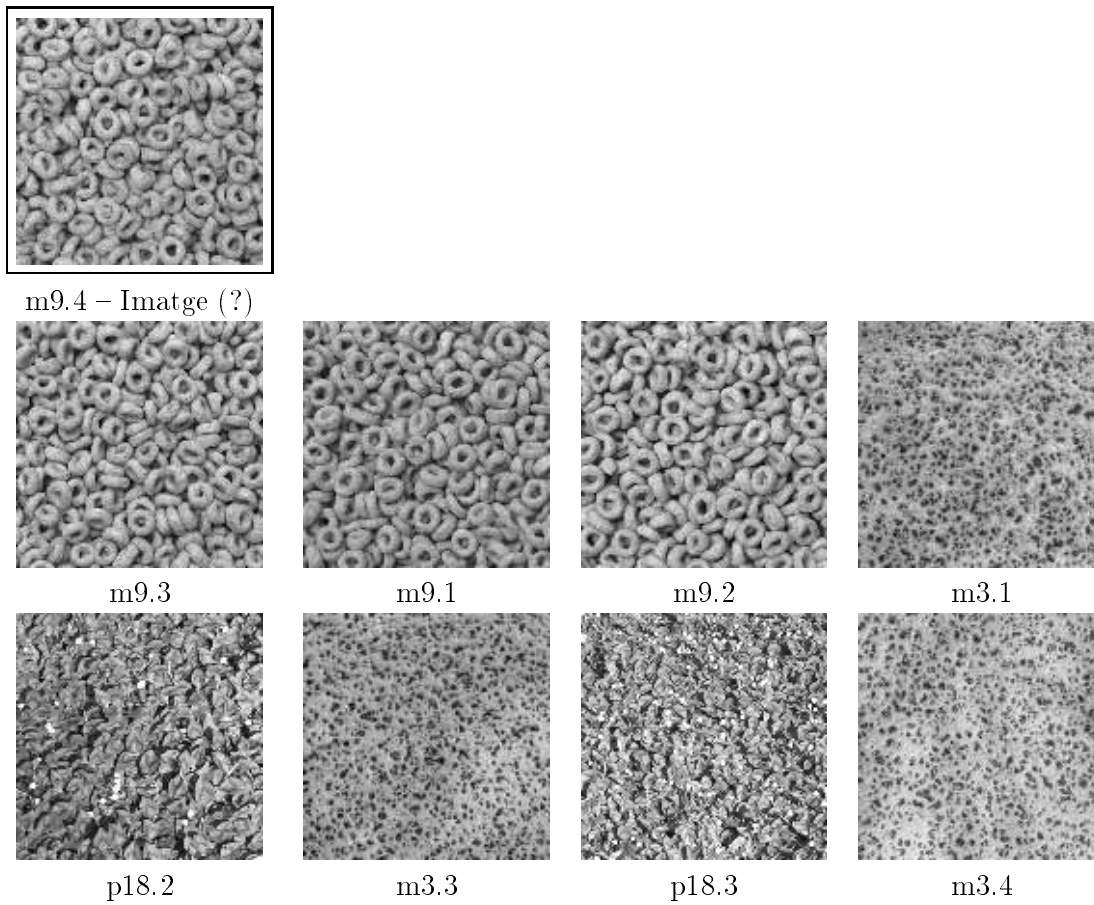


FIGURA 7.2: Experiment sobre DBI1.

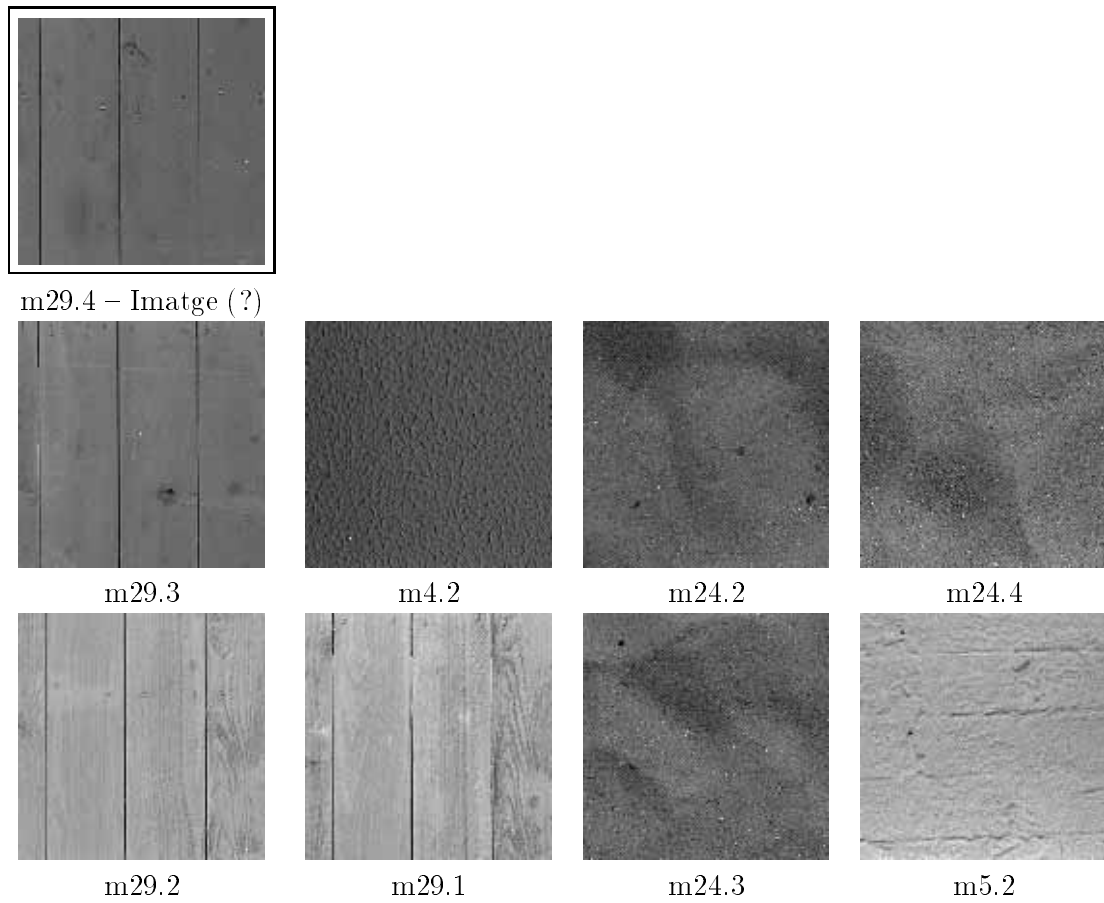


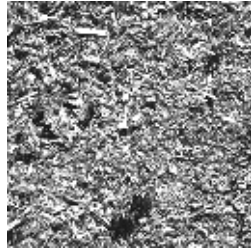
FIGURA 7.3: Experiment sobre DBI1.



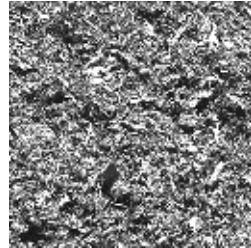
p8.2 - Imatge (?)



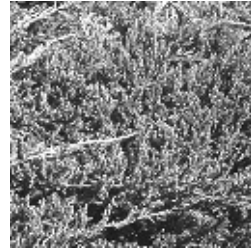
p8.1



m12.2



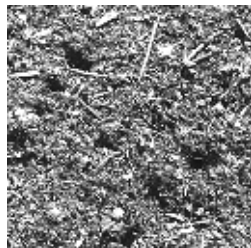
m12.1



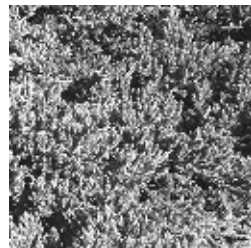
p23.3



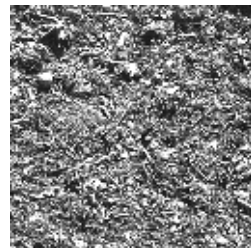
p5.3



m12.4



p23.4



m12.3

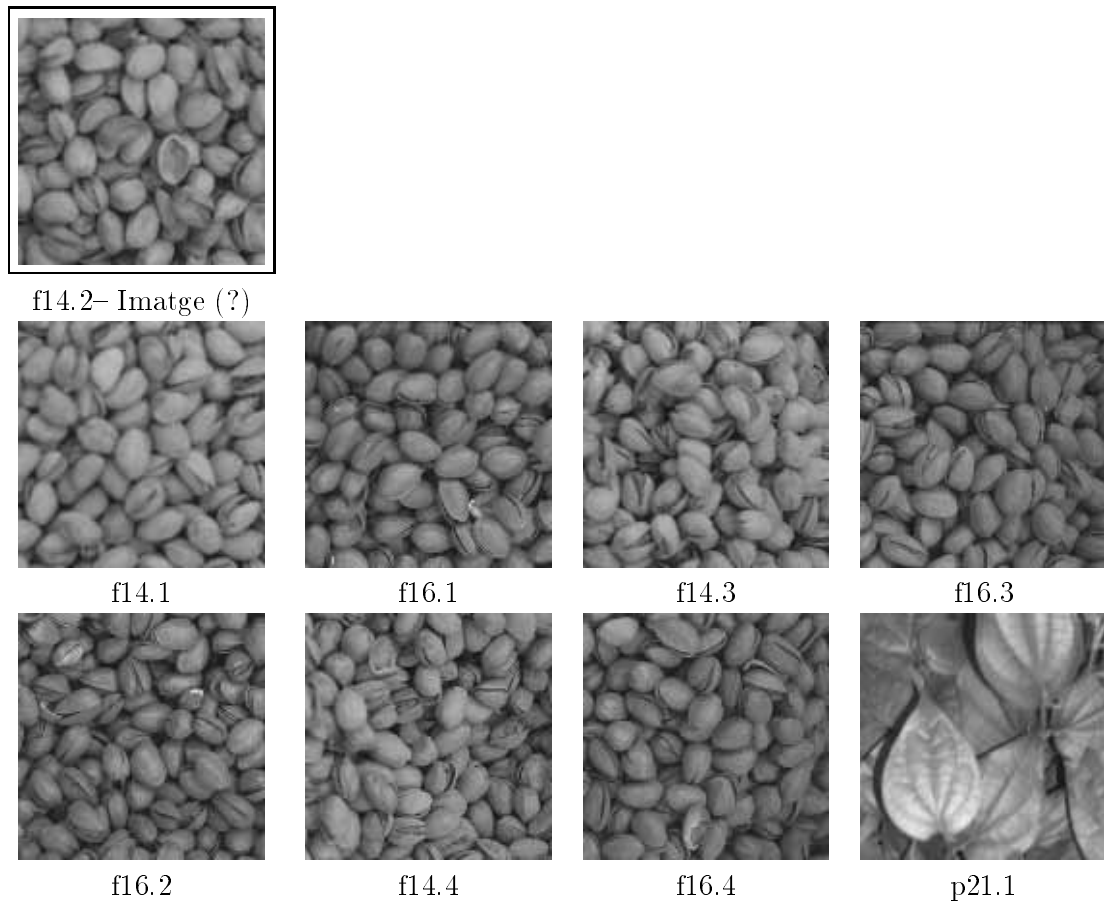


FIGURA 7.4: Experiment sobre DBI1.



p11.2- Imatge (?)



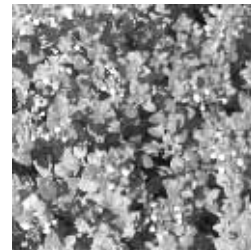
p7.1



p26.2



p9.3



p12.4



p11.1



p26.1



p7.2



p7.4

FIGURA 7.5: Experiment sobre DBI1.

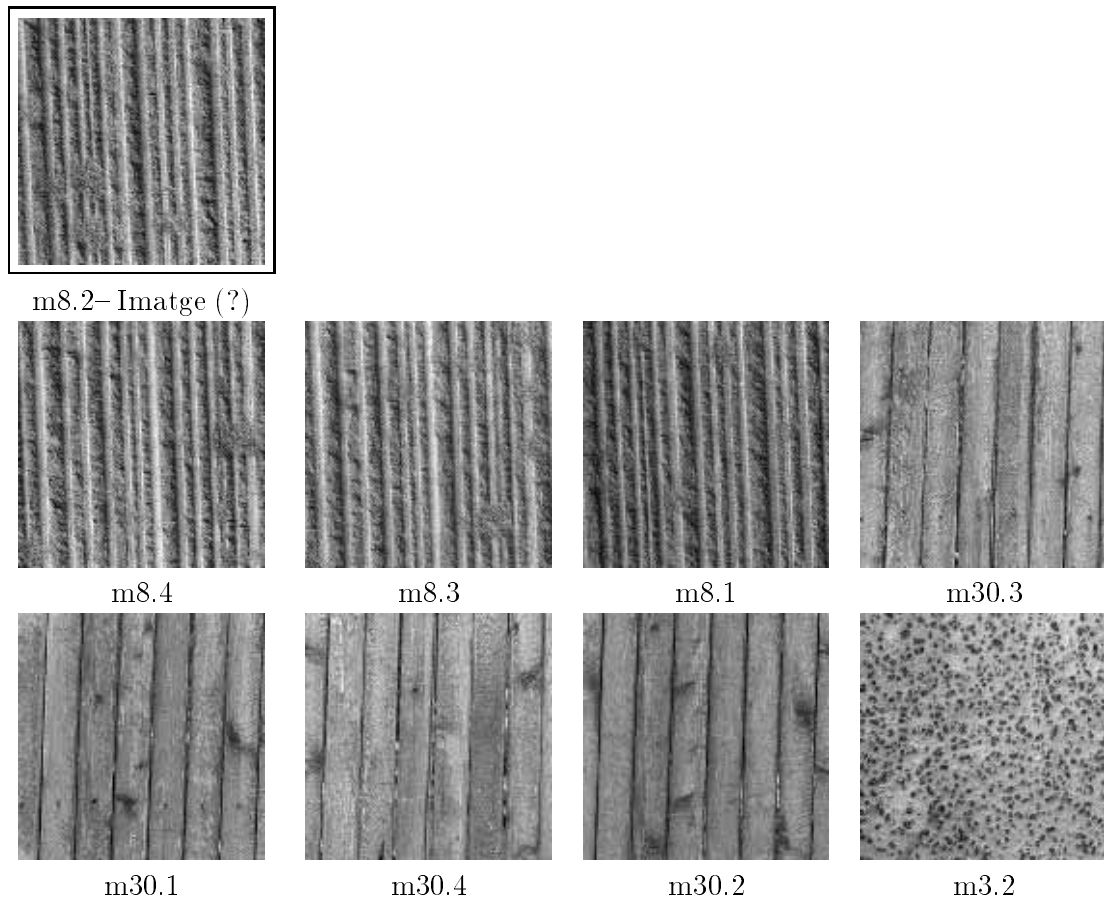
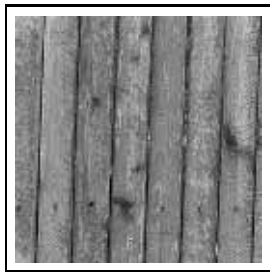
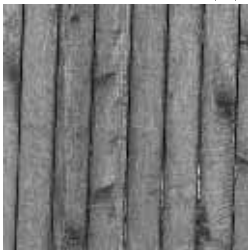


FIGURA 7.6: Experiment sobre DBI1.



m30.1- Imatge (?)



m30.2



m30.3



m30.4



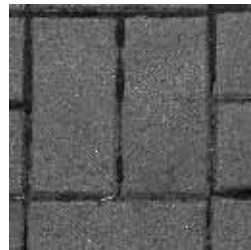
m29.1



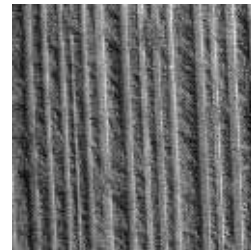
m31.1



m31.2



m6.4



m8.1

FIGURA 7.7: Experiment sobre DBI1.

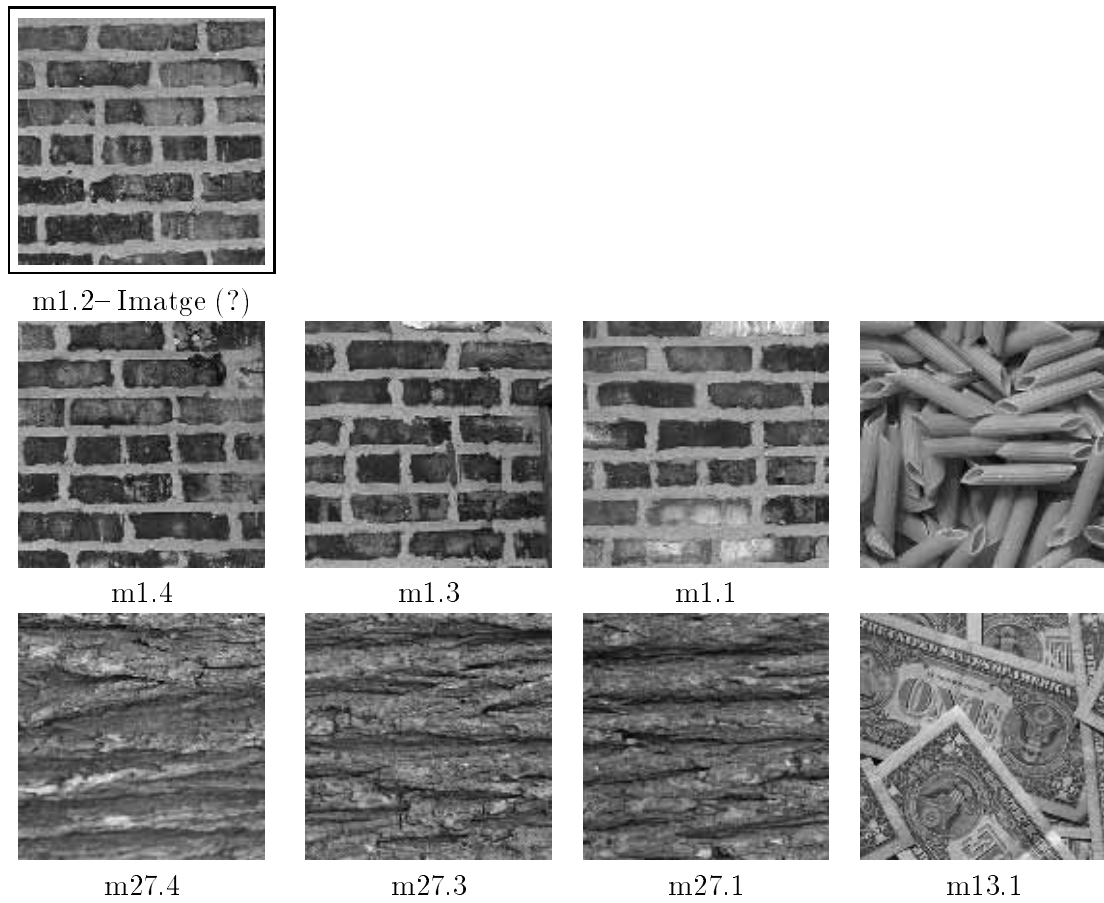


FIGURA 7.8: Experiment sobre DBI1.

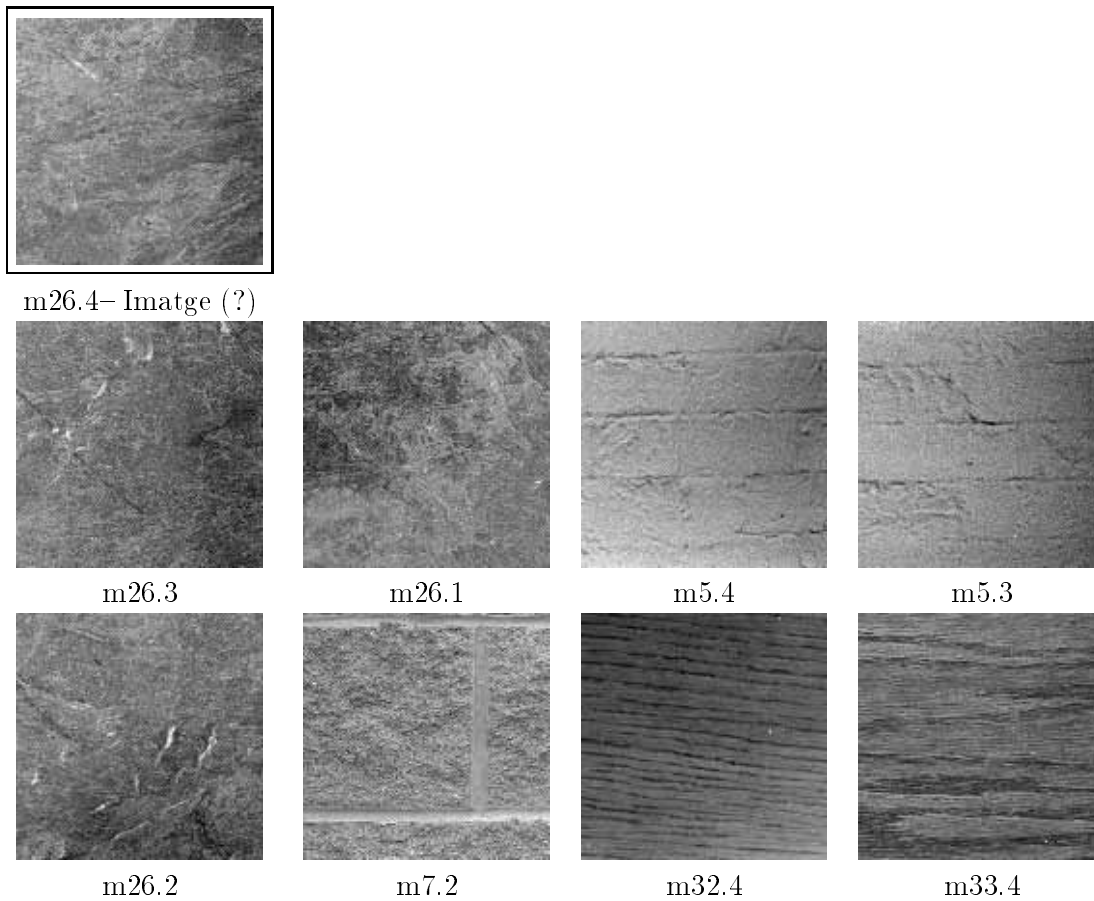


FIGURA 7.9: Experiment sobre DBI1.

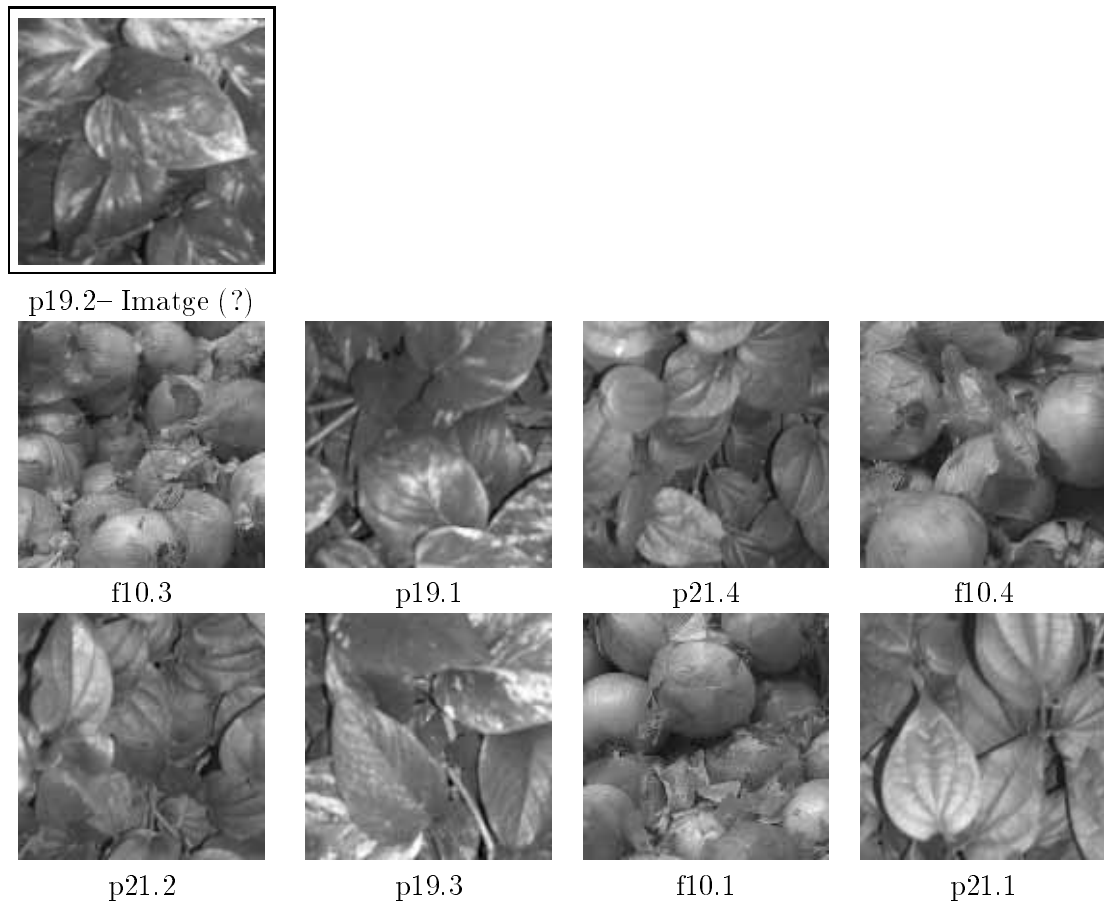


FIGURA 7.10: Experiment sobre DBI1.

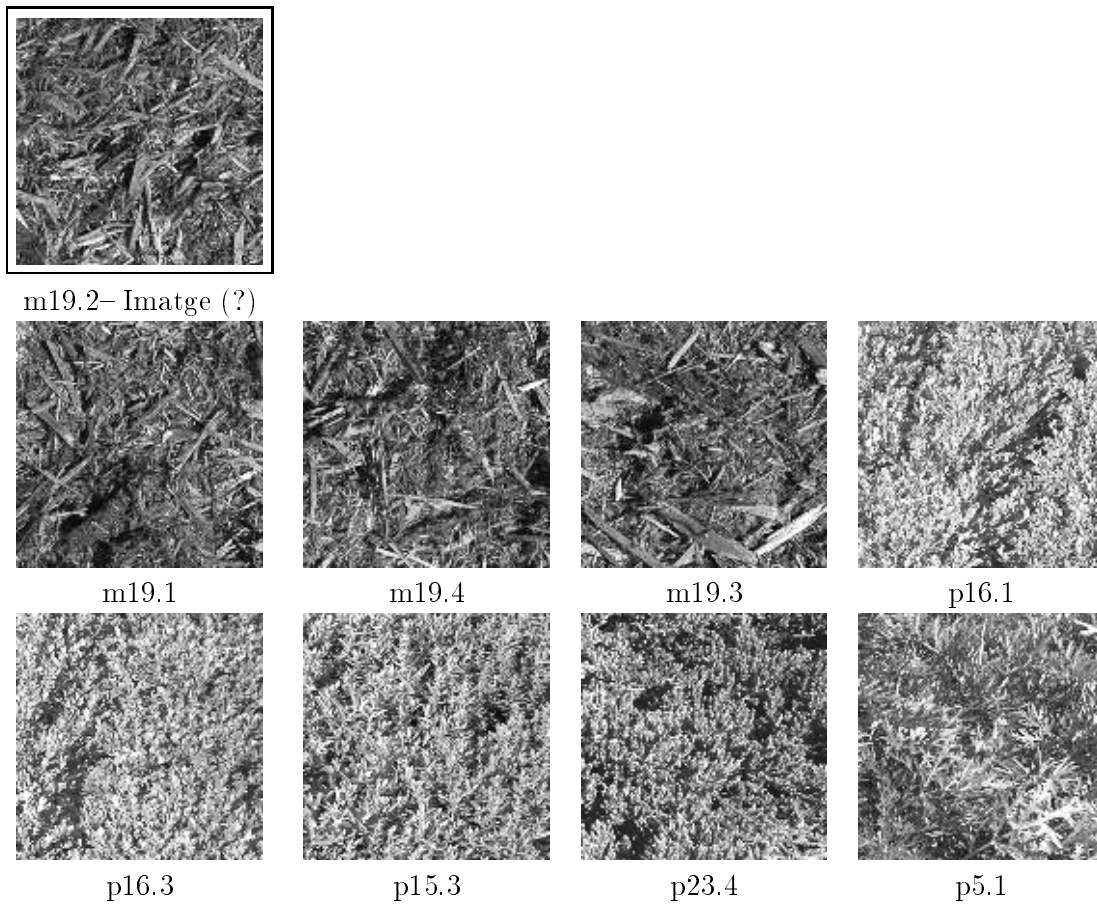


FIGURA 7.11: Experiment sobre DBI1.

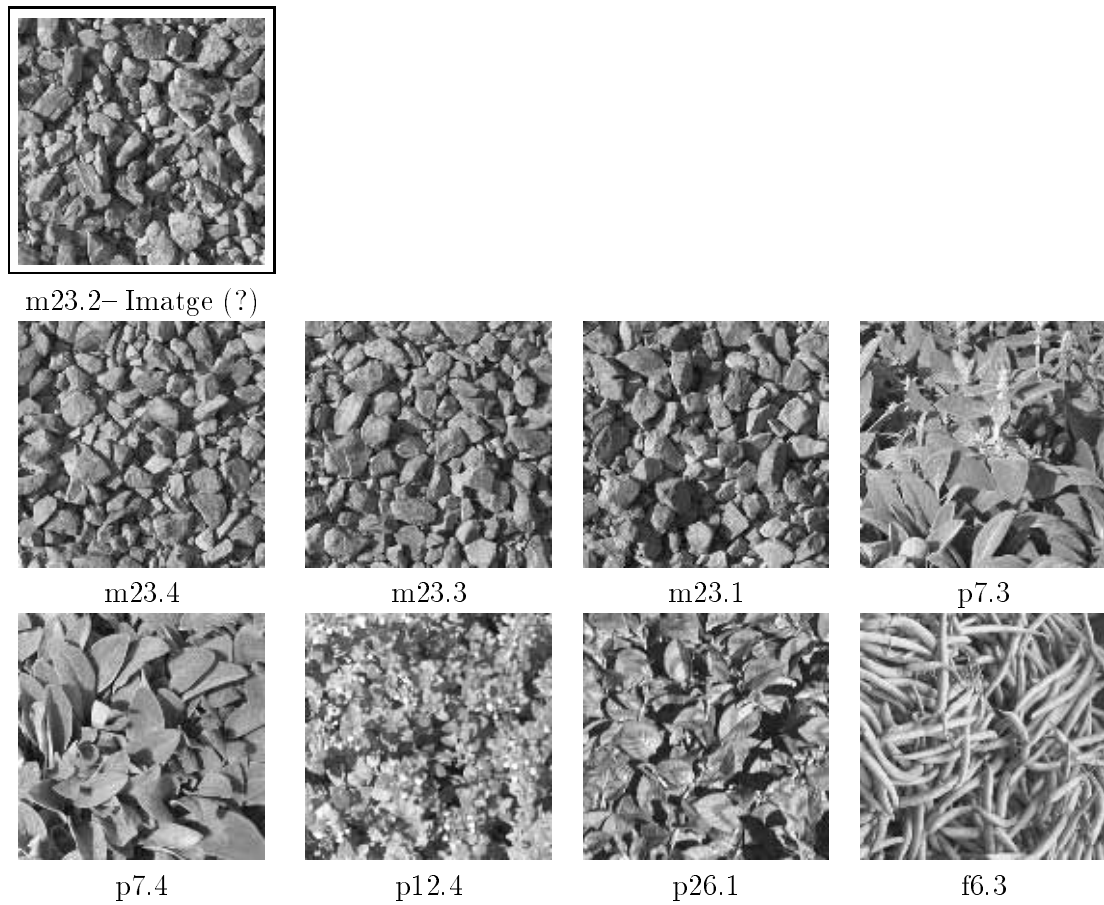
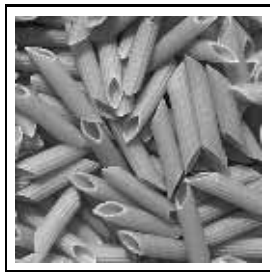


FIGURA 7.12: Experiment sobre DBI1.



m17.4- Imatge (?)



m17.1



m17.2



m17.3



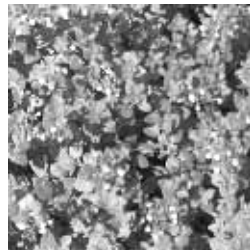
m13.2



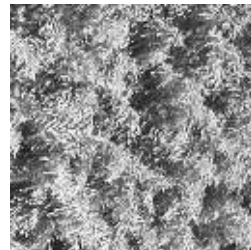
p7.3



m14.4



p12.4



p27.4

FIGURA 7.13: Experiment sobre DBI1.

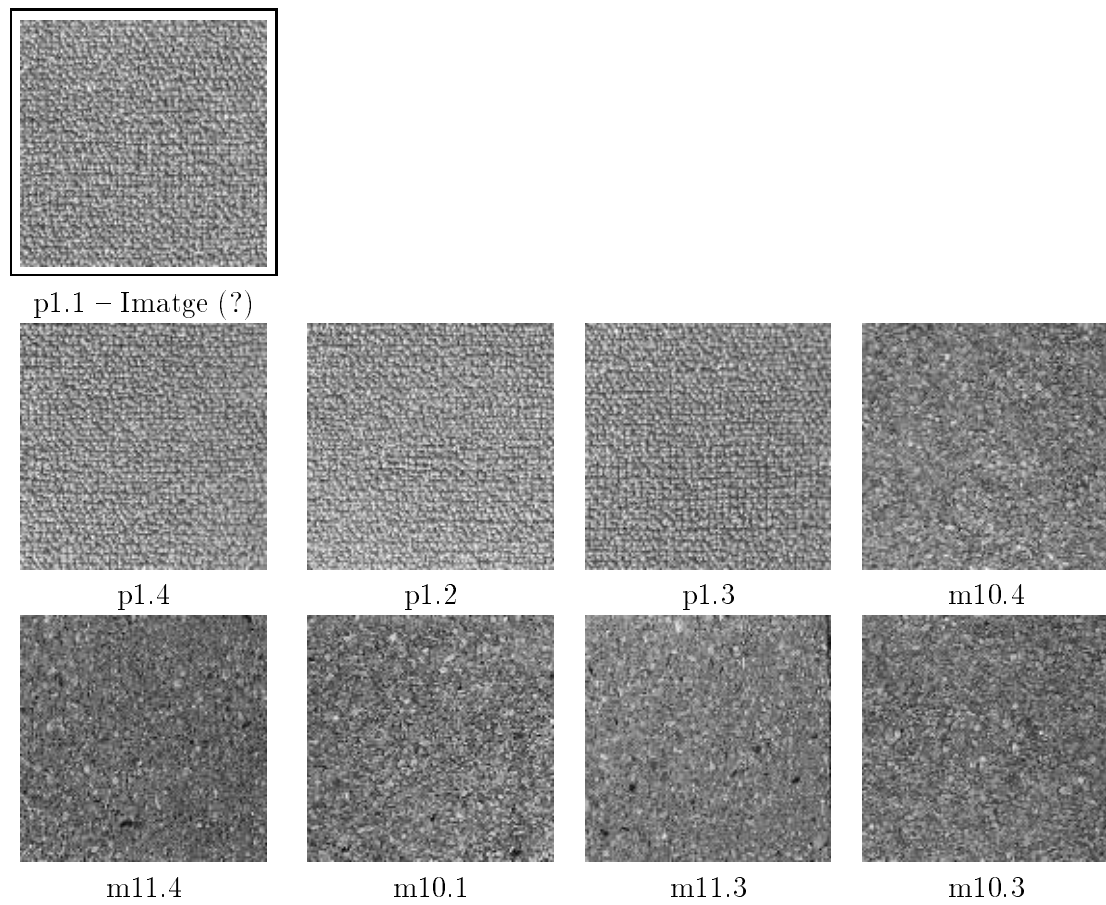


FIGURA 7.14: Experiment sobre DBI1.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Ba.0(1,2,3,4)	Ba.0,Ba.0,Ba.0,P.21.0	Fo.7,P.21.1,P.1.1,Ba.0	Fo.6,Ba.2,P.1.1,P.31.1
Ba.1(1,2,3,4)	Ba.2,Ba.2,Ba.1,Ba.2	Ba.1,Ba.1,Fl.1,Ba.2	Ba.1,St.3,Ba.2,Ba.1
Ba.2(1,2,3,4)	Ba.2,Ba.1,Ba.1,Ba.1	Ba.1,Ba.1,Ba.2,Ba.1	Fl.3,Ba.1,Ba.1,Ba.2
Ba.3(1,2,3,4)	Ba.3,Ba.3,Ba.3,Ba.3	Ba.3,Ba.3,Ba.3,Ba.3	Fo.4,Ba.3,Ba.3,Ba.3
Ba.4(1,2,3,4)	Ba.4,Ba.4,Ba.4,Ba.4	Ba.4,Ba.4,Ba.4,P.31.1	Bu.5,Bu.5,Ba.4,Ba.4
Ba.5(1,2,3,4)	Ba.5,Fo.10,Ba.5,P.21.1	Fo.10,Fo.10,P.21.1,P.21.0	Fo.10,Ba.5,Fo.10,P.21.0
Ba.6(1,2,3,4)	Ba.6,Ba.6,Ba.6,Ba.6	Ba.6,Ba.6,Ba.6,Ba.6	Te.0,Ba.6,Ba.6,Ba.6
Ba.7(1,2,3,4)	Ba.7,Ba.7,Ba.7,Ba.7	P.11.1,Ba.12,P.11.1,Ba.9	P.11.1,Ba.11,P.11.1,L.14
Ba.8(1,2,3,4)	Ba.8,Ba.8,Ba.8,Ba.8	Ba.8,Ba.8,Ba.8,Ba.8	Ba.8,P.11.4,Ba.8,L.10
Ba.9(1,2,3,4)	Ba.9,P.11.0,Ba.9,Ba.10	Ba.10,P.11.4,Ba.11,Ba.11	Ba.11,P.11.1,Ba.11,Ba.12
Ba.10(1,2,3,4)	Ba.10,Ba.10,Ba.9,Ba.10	Ba.10,Ba.10,Ba.12,Ba.10	Ba.10,Ba.9,Ba.11,Ba.9
Ba.11(1,2,3,4)	Ba.11,Fo.2,Ba.12,Ba.12	Ba.9,Ba.11,Ba.12,Ba.11	Ba.12,Fo.2,Ba.11,Ba.12
Ba.12(1,2,3,4)	Ba.12,Ba.12,Ba.11,Ba.11	Ba.11,Ba.11,Ba.12,Ba.11	Ba.11,Ba.10,L.15,Ba.12

TAULA 7.4: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Bark*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Br.0(1,2,3,4)	Br.0,Br.1,Br.0,Br.1	Br.0,Br.1,Br.1,Br.0	Br.1,Br.1,Br.0,Br.0
Br.1(1,2,3,4)	Br.0,Br.0,Br.0,Br.1	Br.1,Br.1,Br.0,Br.1	Br.1,Br.1,Br.1,Br.0
Br.2(1,2,3,4)	Br.2,Br.2,Br.2,Br.2	Br.2,Br.2,Br.2,Br.2	Br.2,Br.2,Br.2,Br.2
Br.3(1,2,3,4)	Br.3,Br.3,Br.3,Br.3	Br.3,Br.3,Br.3,Br.3	Wa.7,Br.3,Wa.7,Wa.7
Br.4(1,2,3,4)	Br.4,Br.4,Br.4,Br.4	Br.4,Br.4,Br.4,Br.4	Mi.3,Br.4,Br.4,Br.4
Br.5(1,2,3,4)	Br.5,Br.5,Br.5,Br.5	Br.5,Br.5,Br.5,Br.5	Br.5,Br.5,Br.5,Br.5
Br.6(1,2,3,4)	Br.6,St.2,Ba.1,Ti.6	Ti.6,Br.6,Ba.2,Br.6	Ba.1,St.2,Fl.3,Ti.3
Br.7(1,2,3,4)	Br.7,G.0,Br.7,Br.7	Mi.0,St.0,P.11.3,Br.7	P.31.0,Bu.4,P.31.0,P.31.0
Br.8(1,2,3,4)	Ti.3,Te.1,Wh.1,L.7	Fl.0,P.21.1,L.6,Br.7	Fl.0,P.21.0,L.5,Fl.2

TAULA 7.5: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Brick*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Bu.0(1,2,3,4)	Bu.2,P.11.1,Bu.4,P.41.1	Br.6,Ba.1,P.31.0,P.11.0	Bu.1,P.11.0,Ba.4,P.41.1
Bu.1(1,2,3,4)	Bu.2,Bu.2,Bu.2,Bu.2	Bu.3,Bu.0,Br.7,Bu.2	Ba.9,L.15,Bu.1,Bu.1
Bu.2(1,2,3,4)	Bu.1,Bu.0,Bu.1,Bu.1	Bu.3,Bu.1,Bu.1,Bu.1	Bu.5,Br.6,Br.7,Bu.2
Bu.3(1,2,3,4)	Bu.1,Bu.2,Bu.3,Bu.3	Bu.2,Bu.1,Bu.1,Ba.0	L.14,Ba.7,Br.8,Bu.2
Bu.4(1,2,3,4)	Bu.4,Bu.4,G.0,Fa.4	Bu.1,Bu.5,Ba.4,Bu.5	Ba.0,Bu.1,Bu.0,L.6
Bu.5(1,2,3,4)	L.15,Ba.7,Ba.4,Ba.4	Bu.6,Ba.11,Ba.4,G.0	St.0,Ba.9,Bu.0,Ba.4
Bu.6(1,2,3,4)	Bu.5,P.41.1,Wa.3,Bu.6	Ba.11,P.41.1,Wa.5,Fa.6	Ba.12,Bu.0,Wa.3,Wa.3
Bu.7(1,2,3,4)	Bu.0,Fl.4,St.2,Bu.7	P.41.1,Fl.4,Wa.5,Bu.0	Br.8,P.31.0,St.2,St.2
Bu.8(1,2,3,4)	Bu.10,Bu.8,Bu.8,Bu.8	Bu.10,Bu.10,Bu.10,Bu.10	Bu.8,Bu.10,Bu.10,Bu.10
Bu.9(1,2,3,4)	Bu.9,Bu.9,Bu.9,Bu.9	Bu.9,Bu.9,Bu.9,Bu.9	Bu.9,Bu.9,Bu.9,Bu.9
Bu.10(1,2,3,4)	Bu.10,Bu.10,Bu.10,Bu.10	Bu.8,Bu.8,Bu.8,Bu.8	Bu.8,Bu.8,Bu.8,Bu.8

TAULA 7.6: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Buildings*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
C.0(1,2,3,4)	C.0,C.0,C.1,C.1	C.0,C.0,C.0,C.0	C.1,C.1,C.0,C.1
C.1(1,2,3,4)	Wo.2,C.1,C.0,C.1	C.1,C.1,C.0,C.0	Wo.2,C.0,C.0,C.0

TAULA 7.7: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Clouds*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Fa.0(1,2,3,4)	Fa.1,Fa.0,Fa.1,Fa.1	Fa.1,Fa.1,Fa.0,Fa.0	Fa.0,Fa.0,Fa.1,Fa.0
Fa.1(1,2,3,4)	Fa.1,Fa.0,Fa.1,Fa.0	Fa.1,Fa.1,Fa.1,Fa.0	Fa.0,Fa.0,Fa.0,Fa.1
Fa.2(1,2,3,4)	Fa.2,Fa.2,Fa.2,Fa.2	Fa.2,Fa.2,Fa.16,Fa.2	Fa.2,Fa.3,Fa.16,Fa.3
Fa.3(1,2,3,4)	Fa.3,Fa.3,Fa.3,Fa.3	Fa.3,Fa.3,Fa.3,Fa.3	Fa.3,Fa.2,Fa.2,Fa.3
Fa.4(1,2,3,4)	Fa.5,Fa.4,Fa.5,Fa.4	Fa.4,Fa.5,Fa.5,Fa.5	Fa.5,Fa.4,Fa.4,Fa.4
Fa.5(1,2,3,4)	Fa.5,Fa.5,Fa.6,Fa.5	Fa.6,Fa.4,Fa.6,Fa.4	Fa.6,Fa.4,Fa.5,Wa.0
Fa.6(1,2,3,4)	St.5,Fa.6,Fa.6,Fa.6	Fa.6,Fa.6,Fa.5,Fa.5	St.4,Fa.5,St.5,Fa.6
Fa.7(1,2,3,4)	Fa.7,Fa.7,Fa.7,Fa.7	Fa.7,Fa.7,Fa.7,Fa.7	Fa.7,Fa.7,Fa.7,Fa.7
Fa.8(1,2,3,4)	Fa.8,Fa.8,Fa.10,Fa.10	Fa.8,Fa.8,Fa.10,Fa.8	Fa.10,Fa.10,Fa.8,Fa.8
Fa.9(1,2,3,4)	Fa.9,Fa.9,Fa.9,Fa.9	Fa.9,Fa.9,Fa.9,Fa.9	Fa.9,Fa.9,Fa.9,Fa.9
Fa.10(1,2,3,4)	Fa.10,Fa.10,Fa.10,Fa.10	Fa.10,Fa.10,Fa.10,Fa.10	Fa.10,Fa.10,Fa.10,Fa.10
Fa.11(1,2,3,4)	Fa.11,Fa.11,Fa.11,Fa.11	Fa.11,Fa.11,Fa.11,Fa.11	Fa.11,Fa.11,Fa.11,Fa.11
Fa.12(1,2,3,4)	Fa.12,Fa.12,Fa.12,Fa.12	Fa.12,Fa.12,Fa.12,Fa.12	Fa.12,Fa.12,Fa.12,Fa.12
Fa.13(1,2,3,4)	Fa.13,Fa.13,Fa.13,Fa.13	Fa.13,Fa.13,Fa.13,Fa.13	Fa.13,Fa.13,Fa.13,Fa.13
Fa.14(1,2,3,4)	Fa.14,Fa.14,Fa.14,Fa.14	Fa.14,Fa.14,Fa.14,Fa.14	Fa.14,Fa.14,Fa.14,Fa.14
Fa.15(1,2,3,4)	Fa.15,Fa.16,Fa.16,Fa.16	Fa.16,Fa.15,Fa.15,Fa.15	Fa.15,Fa.15,Fa.16,Fa.16
Fa.16(1,2,3,4)	Fa.16,Fa.15,Fa.15,Fa.16	Fa.16,Fa.16,Fa.16,Fa.16	Fa.16,Fa.16,Fa.16,Fa.16
Fa.17(1,2,3,4)	Fa.17,Fa.17,Fa.17,Fa.17	Fa.17,Fa.17,Fa.17,Fa.17	Fa.17,Fa.17,Fa.17,Fa.17
Fa.18(1,2,3,4)	Fa.18,Fa.18,Fa.18,Fa.18	Fa.18,Fa.18,Fa.18,Fa.18	Fa.19,Fa.18,Fa.18,Fa.18
Fa.19(1,2,3,4)	Fa.18,Fa.19,Fa.18,Fa.19	Fa.19,Fa.19,Fa.18,Fa.19	Fa.19,Fa.18,Fa.18,Fa.18

TAULA 7.8: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Fabric*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Fl.0(1,2,3,4)	Fl.1,Fl.0,Fl.0,P.21.1	Ba.2,Fl.0,Fl.0,Fl.0	Fl.0,P.21.1,Fl.0,Fl.0
Fl.1(1,2,3,4)	Te.8,Fl.1,Fl.1,Fl.1	Te.8,Te.8,Fl.1,Fl.1	Fl.1,St.3,Ba.1,P.21.0
Fl.2(1,2,3,4)	Fl.2,Fl.5,Fl.2,Fl.2	P.1.1,Fl.5,Fl.2,Fl.2	Fl.5,Fl.5,Wh.0,Fl.2
Fl.3(1,2,3,4)	Fl.3,Fl.3,Fl.3,Fl.3	Fl.3,Fl.3,Fl.3,Fl.3	Fl.6,Fl.3,Ba.2,Fl.6
Fl.4(1,2,3,4)	Fl.4,Fl.4,Fl.4,Fl.4	Fl.4,Fl.4,Fl.4,Fl.5	L.10,L.10,L.10,Fl.2
Fl.5(1,2,3,4)	Fl.5,Fl.5,Fl.5,Fl.5	Fl.2,Fl.5,Fl.2,Fl.2	Fl.5,Fl.2,Fl.5,Fl.5
Fl.6(1,2,3,4)	Fl.6,Fl.6,Fl.6,Fl.6	Fl.3,Fl.3,Fl.6,Fl.6	Fl.6,Fl.6,Fl.6,P.21.0
Fl.7(1,2,3,4)	Fl.7,Fl.7,Ba.2,Fl.7	Fl.7,Fl.7,Ba.2,Fl.7	Ba.2,Ba.2,Fl.7,Fl.7

TAULA 7.9: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Flowers*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Fo.0(1,2,3,4)	Fo.0,Fo.0,Fo.0,Fo.0	Fo.0,Fo.0,Fo.0,Fo.0	Fo.0,Fo.0,Fo.0,Fo.0
Fo.1(1,2,3,4)	Fo.1,Fo.1,Fo.1,Fo.1	Fo.1,Fo.1,Fo.1,Fo.1	Fo.1,Fo.1,Fo.1,Fo.1
Fo.2(1,2,3,4)	Fo.2,Fo.2,Fo.2,Fo.2	Fo.2,Fo.2,Fo.2,Fo.2	Fo.2,Fo.2,Fo.2,Fo.2
Fo.3(1,2,3,4)	Fo.3,Fo.3,Fo.3,Fo.3	Fo.3,Fo.3,Fo.3,Fo.3	Fo.3,Fo.3,Fo.3,Fo.3
Fo.4(1,2,3,4)	Fo.4,Fo.4,Fo.4,Fo.4	Fo.4,Fo.4,Wh.2,Fo.4	Fo.4,L.9,L.16,Wh.2
Fo.5(1,2,3,4)	Fo.5,Fo.5,Fo.5,Fo.5	Fo.5,Fo.5,Fo.5,Fo.5	Fo.5,Fo.5,Fo.5,Fo.5
Fo.6(1,2,3,4)	Fo.7,Fo.7,Fo.7,Fo.7	Fo.7,Fo.7,Fo.7,Fo.7	Fo.7,Fo.7,Fo.7,Fo.6
Fo.7(1,2,3,4)	Fo.6,Fo.6,Fo.6,Fo.6	Fo.7,Fo.6,Fo.7,Fo.7	Fo.6,Fo.7,Fo.7,Fo.6
Fo.9(1,2,3,4)	Fo.8,Fo.8,Fo.8,Fo.8	Fo.8,Fo.8,Fo.8,Fo.8	Fo.9,Fo.9,Fo.9,Fo.9
Fo.9(1,2,3,4)	Fo.8,Fo.8,Fo.8,Fo.8	Fo.8,Fo.8,Fo.8,Fo.8	Fo.9,Fo.9,Fo.9,Fo.9
Fo.10(1,2,3,4)	Fo.10,Fo.10,Fo.10,Fo.10	Fo.10,Ba.5,Fo.11,Fo.10	Ba.5,Fo.10,Fo.11,Fo.11
Fo.11(1,2,3,4)	Fo.11,Fo.11,Fo.11,Fo.11	Fo.10,Fo.11,Fo.11,C.0	Fo.11,Fo.11,Fo.10,C.0

TAULA 7.10: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Food*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
G.0(1,2,3,4)	G.0,G.0,Bu.5,Wa.7	P.31.1,G.0,Ba.4,G.0	G.0,Ba.4,G.0,Te.10
G.1(1,2,3,4)	G.1,G.2,G.1,G.1	G.2,G.1,G.1,G.1	G.2,G.2,G.2,G.2
G.2(1,2,3,4)	G.2,G.1,G.2,G.2	G.2,G.1,G.1,G.1	G.1,G.2,G.1,G.1

TAULA 7.11: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Grass*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
L.0(1,2,3,4)	P.21.0,P.21.0,L.0,Wa.7	L.0,P.21.0,Ba.1,Fl.0	Fl.3,P.21.0,L.0,L.0
L.1(1,2,3,4)	Te.6,Te.6,Te.1,Te.2	Wa.6,Te.5,Te.4,Te.2	Wa.6,Te.10,Te.0,Te.6
L.2(1,2,3,4)	L.2,L.2,L.2,L.2	L.2,L.2,Mi.2,Fa.15	Mi.2,Bu.8,Mi.2,Fa.15
L.3(1,2,3,4)	L.3,L.3,L.3,L.3	L.3,L.3,L.3,L.3	L.3,L.3,L.3,L.3
L.4(1,2,3,4)	L.4,L.4,L.4,L.4	L.4,L.4,L.4,L.4	L.4,L.4,L.4,Fo.9
L.5(1,2,3,4)	L.6,L.4,Wh.0,Wh.0	L.4,L.4,Mi.0,Wh.0	L.11,L.4,L.7,L.11
L.6(1,2,3,4)	L.6,Wh.0,L.7,P.31.0	L.6,L.9,L.6,P.31.0	Wh.0,L.6,Wh.0,P.11.3
L.7(1,2,3,4)	L.7,L.7,L.6,P.31.0	L.7,Wh.0,P.31.1,L.5	Fl.2,L.6,L.7,Br.7
L.8(1,2,3,4)	L.8,L.8,L.8,L.8	L.8,L.8,Wh.0,L.8	L.8,L.3,L.8,L.8
L.9(1,2,3,4)	L.9,L.9,L.16,L.9	L.16,L.16,L.9,L.9	L.9,L.9,L.9,L.16
L.10(1,2,3,4)	L.10,L.10,L.10,L.10	L.10,Mi.0,Mi.0,Mi.0	Fl.4,Mi.0,L.10,Mi.0
L.11(1,2,3,4)	Wh.0,L.11,L.11,L.11	L.11,L.11,L.11,L.11	Wh.1,L.11,Wh.1,L.11
L.12(1,2,3,4)	L.12,L.12,L.12,L.12	L.12,L.12,L.12,L.13	L.8,Fl.4,L.13,Mi.0
L.13(1,2,3,4)	L.13,L.13,L.13,L.13	L.12,L.12,L.13,L.12	Mi.0,L.12,L.13,L.13
L.14(1,2,3,4)	L.14,L.14,L.14,L.14	L.14,L.14,L.14,St.1	Ba.7,L.14,L.14,St.1
L.15(1,2,3,4)	P.11.4,L.15,P.11.1,L.15	P.11.4,L.15,St.0,P.11.1	L.15,P.11.0,P.11.0,Ba.12
L.16(1,2,3,4)	L.16,L.16,L.16,L.9	L.16,Mi.0,L.16,L.9	Wh.0,Mi.1,L.9,L.9

TAULA 7.12: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Leaves*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Me.0(1,2,3,4)	Me.0,Me.0,Me.0,Me.0	Me.0,Me.0,Me.0,Me.0	Me.0,Fa.19,Me.0,Fa.18
Me.1(1,2,3,4)	Me.1,Me.1,Me.1,Me.1	Me.1,Me.1,Me.1,Me.1	Me.0,Me.1,Me.1,Me.0
Me.2(1,2,3,4)	Me.3,Me.2,Me.3,Me.2	Me.2,Me.2,Me.2,Me.3	Me.3,Me.3,Me.2,Me.2
Me.3(1,2,3,4)	Me.3,Me.2,Me.3,Me.2	Me.2,Me.3,Me.3,Me.3	Me.3,Me.2,Me.2,Me.2
Me.4(1,2,3,4)	Me.5,Me.5,Me.5,Me.5	Me.4,Me.4,Me.5,Me.5	Me.5,Me.5,Me.4,Me.4
Me.5(1,2,3,4)	Me.4,Me.4,Me.4,Me.4	Me.4,Me.4,Me.5,Me.4	Me.4,Me.5,Me.4,Me.5

TAULA 7.13: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Metal*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Mi.0(1,2,3,4)	Mi.0,Mi.0,Mi.0,P.31.0	Mi.1,Mi.0,Mi.0,Mi.0	Mi.0,Mi.0,Mi.0,L.10
Mi.1(1,2,3,4)	Mi.1,Mi.1,Mi.1,Mi.1	Mi.1,Mi.1,Mi.1,Mi.1	Mi.1,Mi.0,Mi.1,Mi.1
Mi.2(1,2,3,4)	Mi.2,Mi.2,Mi.2,Mi.2	Mi.2,Mi.2,Mi.2,Mi.2	Mi.2,Sa.4,Mi.2,Sa.0
Mi.3(1,2,3,4)	Mi.3,Mi.3,Mi.3,Mi.3	Mi.3,Mi.3,Mi.3,Mi.3	Mi.3,Mi.3,Mi.3,Mi.2

TAULA 7.14: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Misc.*

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
P.1.0(1,2,3,4)	Te.10,Br.5,St.3,P.11.0	Te.10,Br.5,St.0,P.11.1	Mi.3,Br.5,Ba.1,Fa.16
P.1.1(1,2,3,4)	P.1.1,P.1.1,Wh.2,P.1.1	Fl.2,Fl.2,P.1.1,P.1.1	P.11.2,P.11.3,P.1.1,Fl.2
P.11.0(1,2,3,4)	P.11.0,P.11.0,P.11.2,P.11.1	P.1.0,P.11.1,P.11.4,Te.10	P.11.1,P.11.1,P.11.1,Te.1
P.11.1(1,2,3,4)	P.11.0,Te.10,P.1.0,P.11.2	L.15,St.0,Fa.16,P.11.1	P.11.0,P.11.2,Fa.15,P.11.1
P.11.2(1,2,3,4)	P.11.1,Te.10,P.1.1,P.11.0	P.11.1,Te.10,P.1.1,St.3	P.11.1,Te.0,P.1.1,P.11.4
P.11.3(1,2,3,4)	P.11.3,P.11.3,P.11.3,P.11.3	P.11.3,P.11.3,P.11.3,P.11.3	P.11.3,P.11.3,Fa.16,P.11.3
P.11.4(1,2,3,4)	P.11.4,P.11.4,Fa.16,P.11.4	P.11.4,P.11.4,Fa.16,P.11.4	St.0,P.41.0,P.31.0,Mi.3
P.21.0(1,2,3,4)	L.0,P.21.0,P.21.0,Te.1	Ba.1,Fl.6,P.21.1,P.21.0	P.21.0,Te.8,Ba.5,P.21.1
P.21.1(1,2,3,4)	P.21.1,Fl.0,P.21.1,P.21.0	P.21.1,P.21.1,St.0,P.21.0	P.11.1,Fl.0,P.11.1,Te.1
P.31.0(1,2,3,4)	P.31.0,P.31.0,Mi.0,P.11.3	P.31.0,St.0,P.31.0,P.31.0	P.31.0,Ba.4,L.10,P.31.0
P.31.1(1,2,3,4)	P.31.1,P.31.1,L.7,P.11.3	Ba.0,Fl.0,P.1.1,Ba.4	P.31.1,Fl.0,P.31.1,Ba.4
P.41.0(1,2,3,4)	P.41.0,P.41.0,P.41.0,P.41.0	P.41.0,P.41.0,P.41.0,P.41.1	P.11.4,P.11.4,P.41.1,P.11.4
P.41.1(1,2,3,4)	P.41.1,P.11.4,P.41.1,P.41.1	P.41.0,P.11.4,Bu.0,P.11.4	P.41.1,P.41.1,P.11.0,P.41.0

TAULA 7.15: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Paintings.*

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Sa.0(1,2,3,4)	Sa.0,Sa.0,Sa.0,Sa.0	Sa.0,Sa.0,Sa.0,Sa.4	Sa.0,Sa.2,Sa.0,Sa.0
Sa.1(1,2,3,4)	Sa.1,Sa.1,Sa.1,Sa.1	Sa.1,Sa.1,Sa.1,Sa.1	Sa.1,Sa.1,Sa.1,Sa.3
Sa.2(1,2,3,4)	Sa.2,Sa.2,Sa.2,Sa.2	Sa.2,Sa.2,Sa.2,Sa.2	Sa.4,Sa.2,Sa.0,Sa.4
Sa.3(1,2,3,4)	Sa.3,Sa.4,Sa.4,Sa.4	Sa.4,Sa.0,Mi.3,Sa.4	Sa.4,Sa.3,Mi.3,Sa.3
Sa.4(1,2,3,4)	Sa.3,Sa.4,Sa.3,Sa.0	Sa.3,Sa.2,Sa.4,Sa.2	Sa.0,Sa.0,Sa.3,Sa.4
Sa.5(1,2,3,4)	Sa.5,Sa.5,Sa.5,Sa.5	Sa.5,Sa.5,L.12,Sa.5	Ba.9,Sa.6,Ba.9,Sa.6
Sa.6(1,2,3,4)	Sa.6,Sa.6,Sa.6,Sa.6	Sa.5,Sa.6,Sa.6,Sa.6	Sa.6,Sa.6,Sa.5,Sa.6

TAULA 7.16: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Sand.*

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
St.0(1,2,3,4)	P.1.0,L.15,Te.10,P.1.0	P.11.4,St.0,P.11.1,St.3	St.0,P.11.1,P.11.2,St.0
St.1(1,2,3,4)	St.1,St.1,St.1,St.1	St.1,St.1,St.1,St.1	L.14,L.14,St.1,L.14
St.2(1,2,3,4)	St.2,St.2,St.2,St.2	St.2,St.2,St.2,St.2	St.2,St.2,Br.1,St.2
St.3(1,2,3,4)	P.1.0,St.3,St.3,St.3	Ba.1,St.3,Ba.6,St.3	St.3,St.3,St.3,St.3
St.4(1,2,3,4)	St.4,St.4,St.4,St.4	St.4,St.5,St.4,St.5	St.5,St.4,St.4,St.5
St.5(1,2,3,4)	St.4,St.5,St.5,St.5	St.4,St.4,St.4,St.4	St.5,St.4,St.4,St.5

TAULA 7.17: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Stone*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Te.0(1,2,3,4)	Te.5,Te.3,Te.1,Te.1	Te.5,Te.3,Te.4,Te.0	Te.1,Te.3,Te.3,Te.3
Te.1(1,2,3,4)	Te.4,Te.9,Te.0,Te.2	Te.3,Te.6,Te.0,Te.2	Te.3,Te.0,Te.1,Te.9
Te.2(1,2,3,4)	Te.2,Te.6,Te.2,Te.1	Te.1,Te.2,Te.9,Te.2	Te.9,Te.1,Te.1,Te.2
Te.3(1,2,3,4)	Te.3,Te.3,Te.5,Te.3	Te.4,Te.3,Te.0,Te.0	Te.3,Te.4,Te.8,Te.3
Te.4(1,2,3,4)	Te.3,Te.5,Te.3,Te.5	Te.3,Te.0,Te.4,Te.1	Te.3,Te.3,Te.3,Te.5
Te.5(1,2,3,4)	Te.4,Te.5,Te.5,Te.6	Te.5,Te.3,Te.0,Te.5	Te.0,Te.9,Te.4,Te.5
Te.6(1,2,3,4)	Te.6,Te.9,Te.5,Te.1	Te.10,Te.8,Te.6,Te.2	L.1,Te.7,Te.5,Te.1
Te.7(1,2,3,4)	Te.7,Te.7,Te.8,Te.8	Te.6,Te.6,Te.8,Te.6	Te.9,Te.5,Te.6,Te.9
Te.8(1,2,3,4)	Te.7,Te.7,Fl.1,Fl.1	Te.6,Te.3,Te.8,Te.8	Te.8,Te.6,Ti.5,Fl.1
Te.9(1,2,3,4)	Te.1,Te.0,P.11.2,Te.2	Te.6,Te.1,Te.10,Te.1	Te.7,Te.3,Te.10,Te.2
Te.10(1,2,3,4)	Te.4,Te.6,P.11.2,P.11.2	Te.9,Te.5,Te.10,Te.10	Te.3,L.1,P.11.1,Te.9

TAULA 7.18: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Terrain*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Ti.0(1,2,3,4)	Ti.0,Ti.0,Ti.0,Ti.0	Ti.0,Ti.0,Ti.0,Ti.0	Ti.0,Ti.0,Ti.0,Ti.0
Ti.1(1,2,3,4)	Ti.1,Ti.1,Ti.1,Ti.1	P.1.0,Br.4,Ba.2,Ba.2	Ba.2,P.1.0,Ba.2,Fl.3
Ti.2(1,2,3,4)	Ti.2,Ti.2,Ti.2,Ti.2	Ba.6,Te.3,Br.8,Ti.6	Ba.6,Te.4,Ti.3,Ti.3
Ti.3(1,2,3,4)	Ti.3,Ti.3,Ti.3,Ti.3	Ti.3,Ti.3,Ti.3,Ti.3	Ti.3,Fl.1,Br.8,Ti.3
Ti.4(1,2,3,4)	Ti.4,Ti.4,Ti.4,Ti.4	Ti.4,Ti.4,Ti.4,Ti.4	Ti.4,Ti.4,Ti.4,Ti.4
Ti.5(1,2,3,4)	Ti.5,Te.8,P.11.4,Ti.5	P.11.4,G.0,P.11.1,Bu.5	P.11.4,Te.8,P.31.0,Ba.4
Ti.6(1,2,3,4)	Ti.5,Br.6,St.3,Ti.5	P.41.1,St.2,Te.1,Ti.3	Te.8,Br.8,Ba.6,Br.8
Ti.7(1,2,3,4)	Ti.7,Ti.7,Ti.7,Ti.7	Ti.7,Ti.7,Ti.9,Ti.7	Ti.9,Ti.9,Ti.9,Ti.9
Ti.8(1,2,3,4)	Ti.8,Ti.8,Ti.8,Ti.8	Ti.8,Ti.8,Ti.8,Ti.8	Ti.8,Ti.8,Ti.8,Ti.8
Ti.9(1,2,3,4)	Ti.9,Ti.9,Ti.9,Ti.9	Ti.9,Ti.9,Ti.9,Ti.7	Ti.10,Ti.9,Ti.7,Ti.10
Ti.10(1,2,3,4)	Ti.10,Ti.10,Ti.10,Ti.10	Ti.10,Ti.10,Ti.10,Ti.10	Ti.10,Ti.10,Ti.10,Ti.10

TAULA 7.19: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Tile*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Wa.0(1,2,3,4)	Wa.0, Wa.0, Wa.0, Wa.0	Fa.5, Fa.5, Wa.4, Ba.4	Fa.5, Fa.5, Ba.4, Wa.4
Wa.1(1,2,3,4)	Wa.1, Wa.1, Wa.1, Wo.2	Wa.6, Wa.6, Wa.1, Wo.2	Wa.1, Wa.6, Wa.1, Wo.2
Wa.2(1,2,3,4)	Wa.2, Te.2, Wa.7, Wa.2	Wa.2, Te.2, Wa.2, Wa.7	Sa.4, Wa.5, Wa.2, Te.1
Wa.3(1,2,3,4)	Wa.3, Wa.3, Wa.6, Wa.3	Wa.6, Wa.6, Wa.6, Wa.6	Wa.6, Wa.6, Wa.3, Wa.5
Wa.4(1,2,3,4)	Wa.4, Wa.4, Wa.4, Wa.4	Wa.0, Wa.0, L.7, P.31.1	Wa.0, Wa.0, P.31.1, Ba.4
Wa.5(1,2,3,4)	Wa.5, Wa.5, Wa.5, Wa.5	Wa.3, Wa.3, Wa.3, Wa.3	Wa.3, Wa.3, Wa.3, Wa.3
Wa.6(1,2,3,4)	Wa.6, Wa.6, Wa.6, Wa.6	Wa.3, Wa.3, Wa.3, Wa.3	Wa.3, Wa.3, Wa.6, Wa.6
Wa.7(1,2,3,4)	Wa.7, Wa.7, Fl.1, Wa.7	P.1.0, Wa.2, Ba.6, Ti.5	Te.10, Wa.2, Ba.6, G.0

TAULA 7.20: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Water*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Wh.0(1,2,3,4)	L.11,L.5,Wh.2,Fl.4	Wh.0,Wh.0,L.8,Wh.1	Wh.0,Wh.0,L.7,Fl.2
Wh.1(1,2,3,4)	Wh.1,Wh.1,Wh.1,Wh.1	Fo.7,Wh.2,Wh.1,Wh.2	Fo.6,Wh.2,Wh.1,Wh.0
Wh.2(1,2,3,4)	Fo.6,Wh.2,Wh.1,Wh.2	Fo.7,Fo.7,Wh.2,Wh.1	Wh.2,Wh.2,Wh.1,Wh.0

TAULA 7.21: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Whereswaldo*.

Imatge (?)	Primera elecció	Segona elecció	Tercera elecció
Wo.1(1,2,3,4)	Wo.0,Wo.0,Wo.0,Wo.0	P.11.3,Wo.0,P.1.0,Wo.0	P.11.3,Wo.1,Fa.16,Wo.1
Wo.1(1,2,3,4)	Wo.0,Wo.0,Wo.0,Wo.0	P.11.3,Wo.0,P.1.0,Wo.0	P.11.3,Wo.1,Fa.16,Wo.1
Wo.2(1,2,3,4)	Wo.2,Wo.2,Wo.2,Wo.2	Wo.2,Wo.2,Wo.2,Wo.2	Wo.2,Wo.2,Wo.2,Wo.2

TAULA 7.22: Resultats sobre BDI2 a partir de preguntes sobre les imatges de tipus *Wood*.

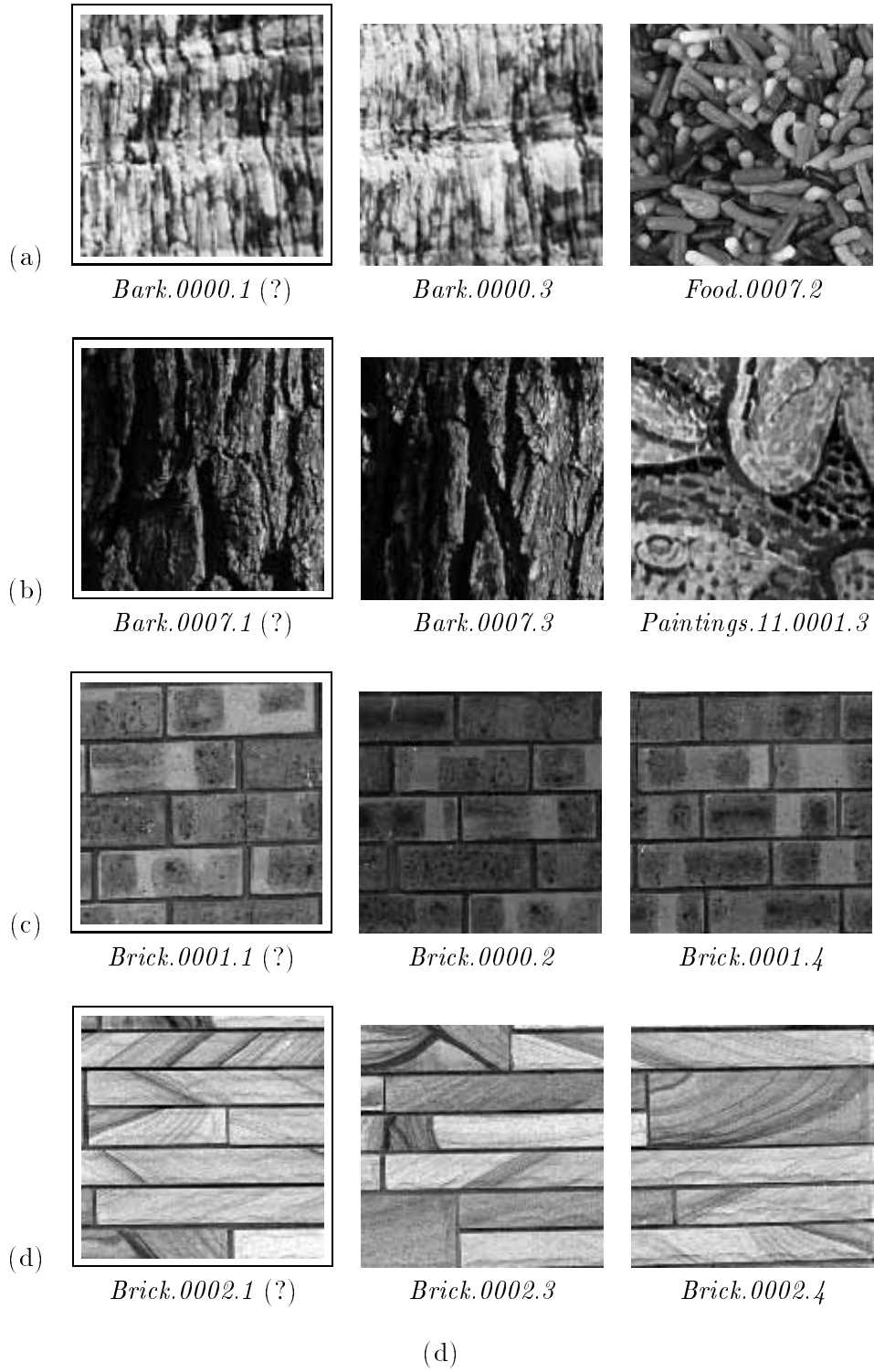


FIGURA 7.15: Experiments sobre DBI2

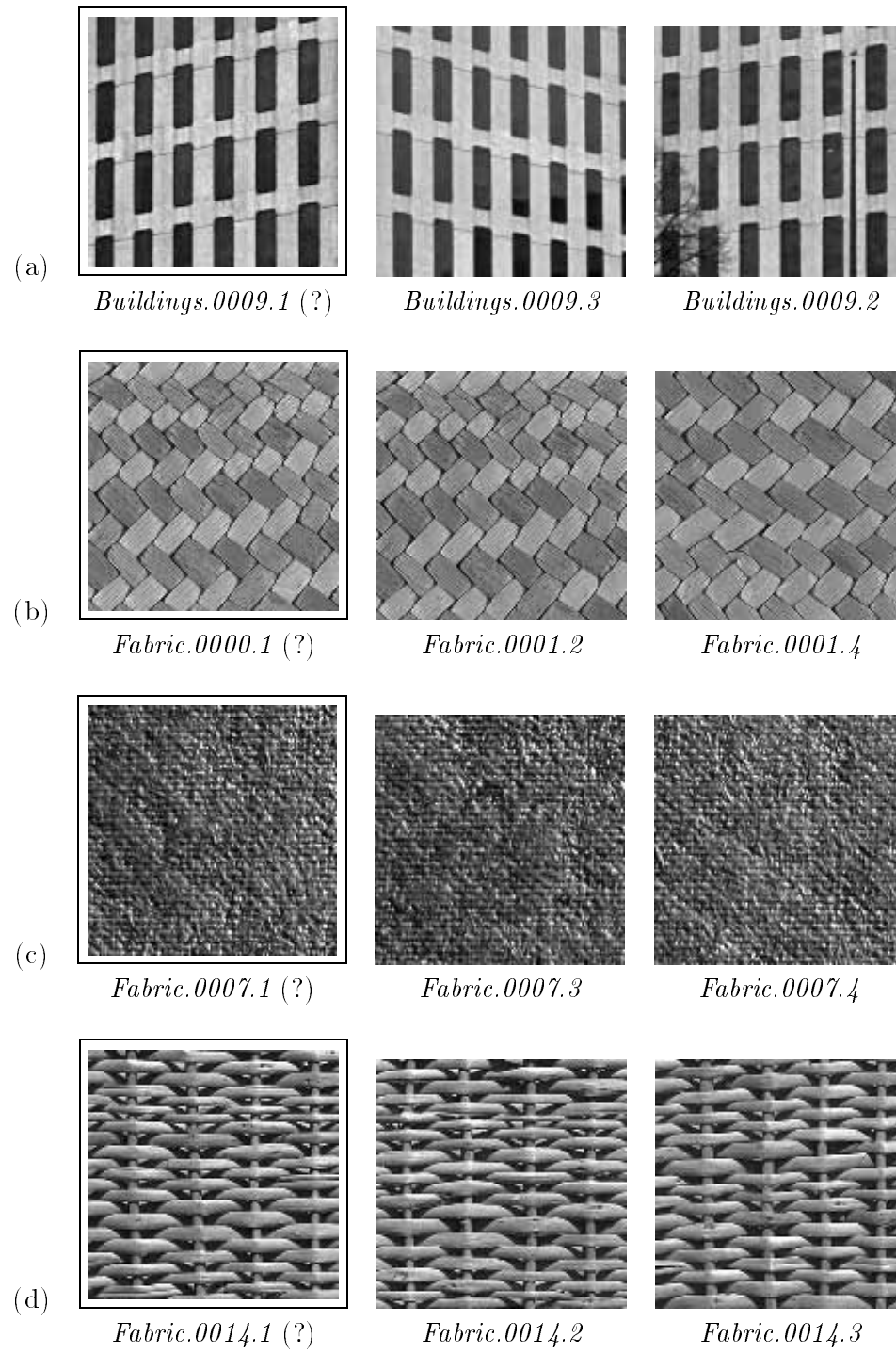


FIGURA 7.16: Experiments sobre DBI2

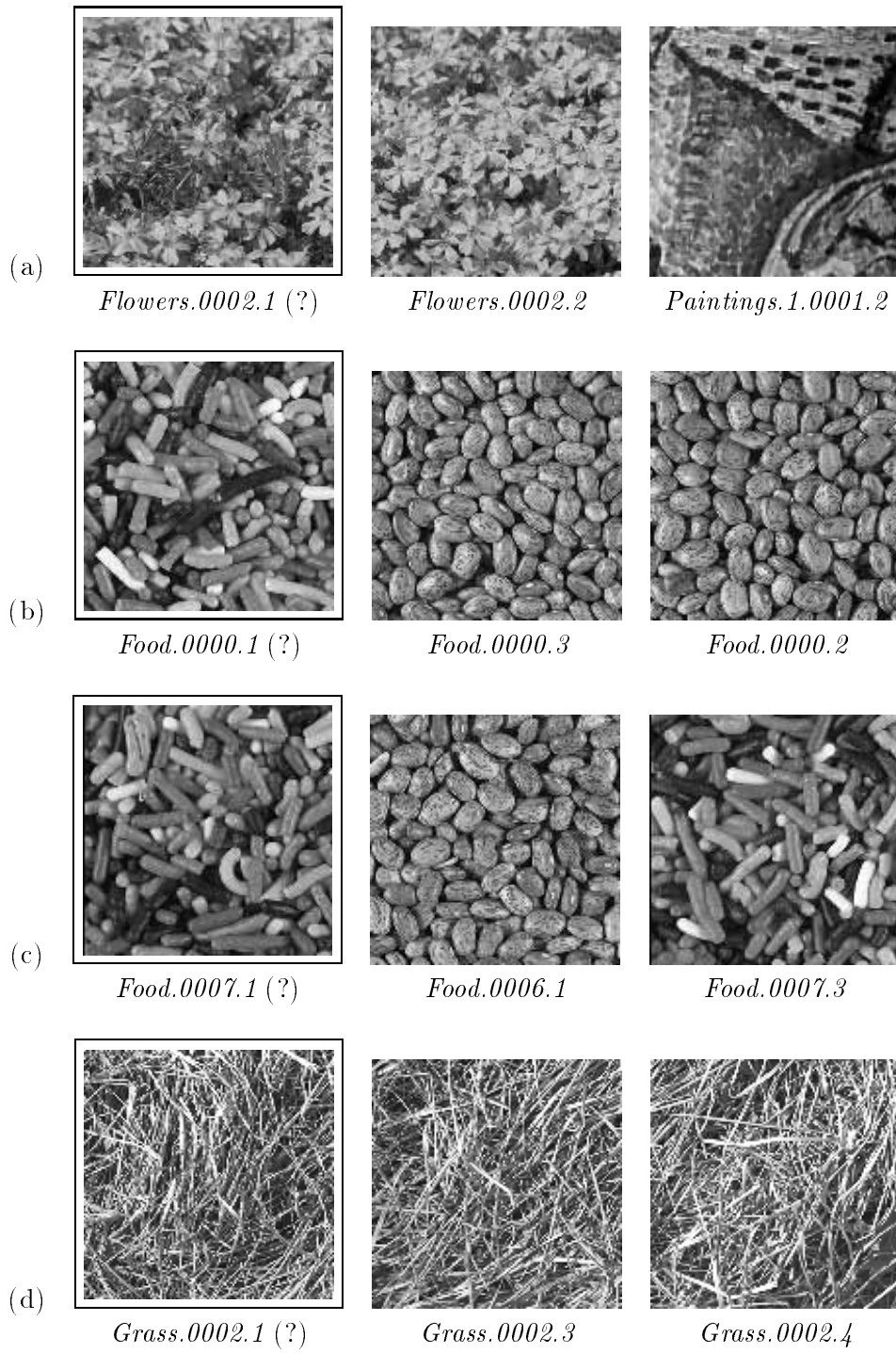


FIGURA 7.17: Experiments sobre DBI2

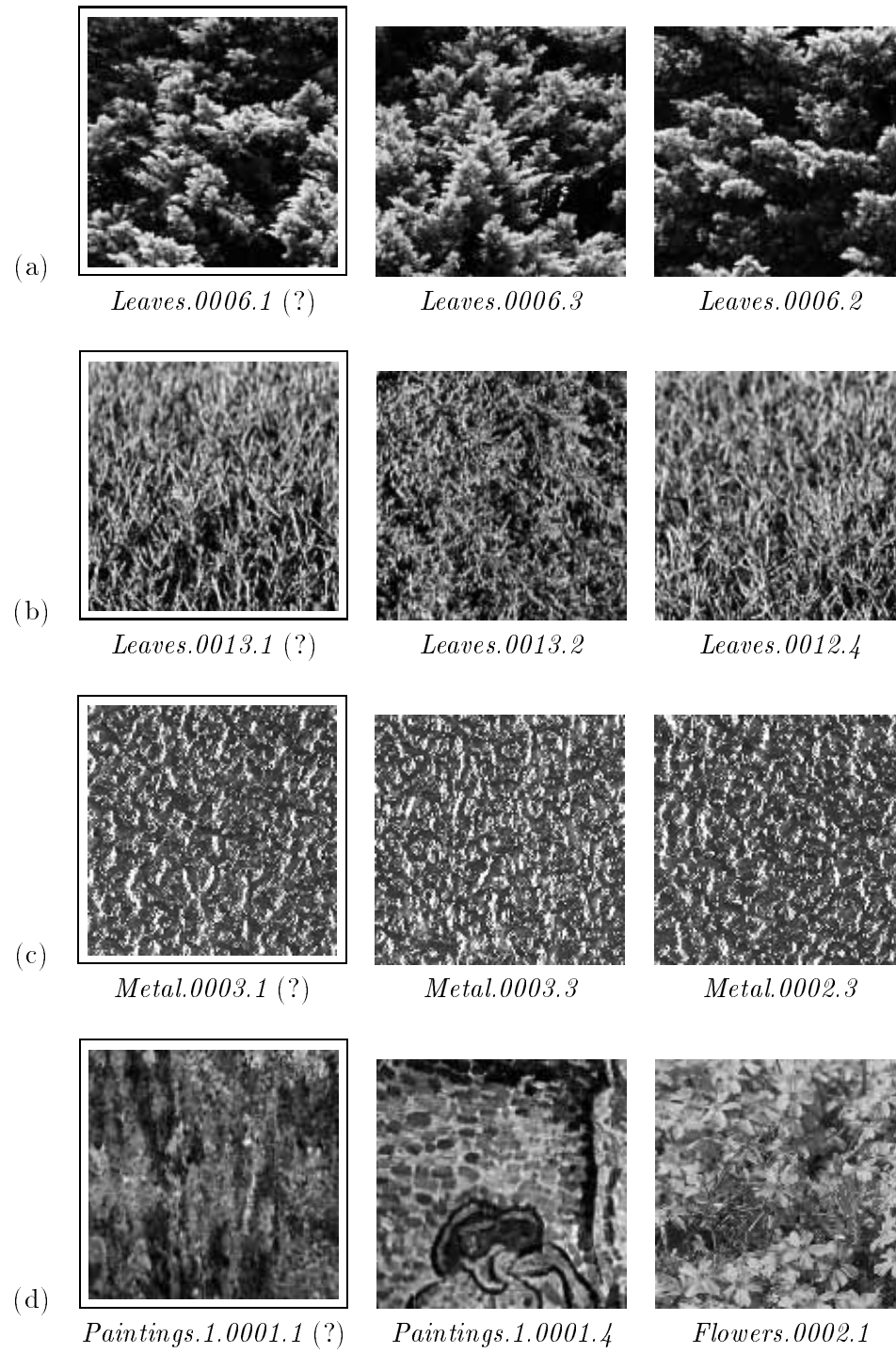


FIGURA 7.18: Experiments sobre DBI2

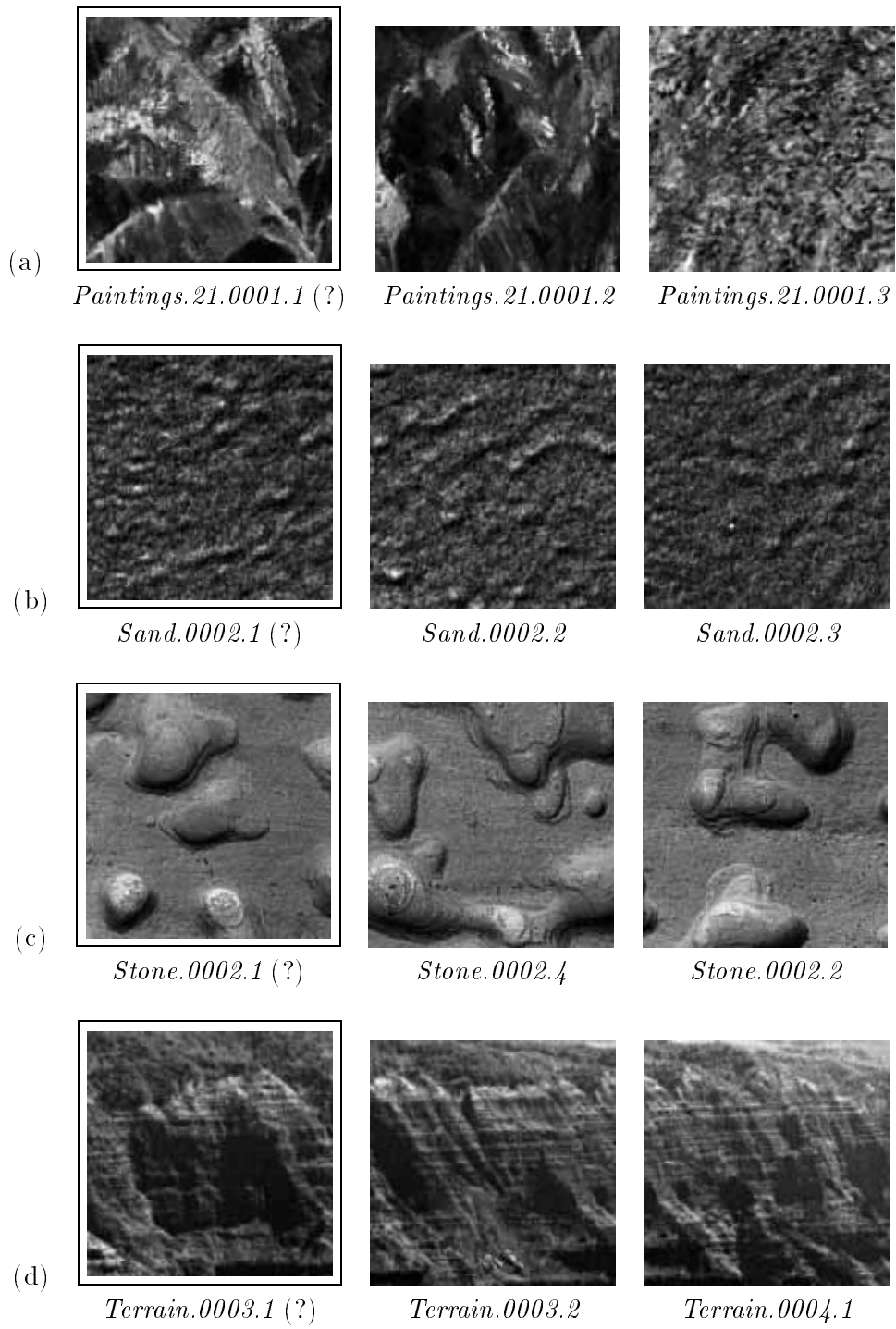


FIGURA 7.19: Experiments sobre DBI2

7.2.3 Avaluació dels resultats

En el treball de G.L. Gimelfarb i A.K. Jain [32] es presenta una aproximació semblant a aquesta en la que es defineix una mesura de la distància entre textures. A partir d'aquesta mesura es fan un conjunt d'experiments semblants als que hem fet nosaltres.

Per a l'avaluació dels resultats es plantegen uns criteris que permetran donar una valoració d'aquests:

Criteri 1 Quantes vegades la mateixa imatge ha estat seleccionada a la *primera selecció*.

Criteri 2 Quantes vegades la mateixa imatge ha estat seleccionada a la *segona selecció*.

Criteri 3 Quantes vegades la mateixa imatge ha estat seleccionada a la *tercera selecció*.

Quan diem la mateixa imatge es té en compte el fet de que les imatges han estat dividides en parts més petites. Aleshores aquestes mesures es basen en el fet de que una recerca és bona si selecciona com a imatge més propera una imatge que provingui de la mateixa imatge d'on ve la que s'està demanant.

Des del nostre punt de vista aquests criteris no són prou objectius per varis motius:

- (i) Primer de tot perquè les imatges no sempre han estat captades amb les mateixes condicions d'il·luminació a tota l'escena, i perquè algunes d'elles representen escenes en els que certes parts no representen cap textura en si mateixa.
- (ii) Existeixen conjunts d'imatges que són semblants entre elles. Per tant, la imatge que es selecciona no té perquè provenir de la mateixa imatge original sinó que pot venir de qualsevol imatge del conjunt i pot continuar sent un resultat vàlid.

D'aquí que en un dels experiments els autors hagin afegit un quart criteri:

Criteri 4 Quantes vegades la mateixa imatge o imatges que s'han definit de conjunts semblants, han estat seleccionades dins de la *primera selecció*, o bé, dins de la primera i la *segona selecció* alhora.

que els hi fa millorar de forma notable l'avaluació.

Els experiments que han realitzat en el seu treball ha estat fet sobre dues BDI's. La primera formada per imatges de l'album de Brodatz [10], són imatges regulars i en condicions de captació homogènies per tota la imatge, els resultats que obtenen sobre un conjunt de 240 preguntes fetes entre els que hi ha 8 imatges de cada tipus, són:

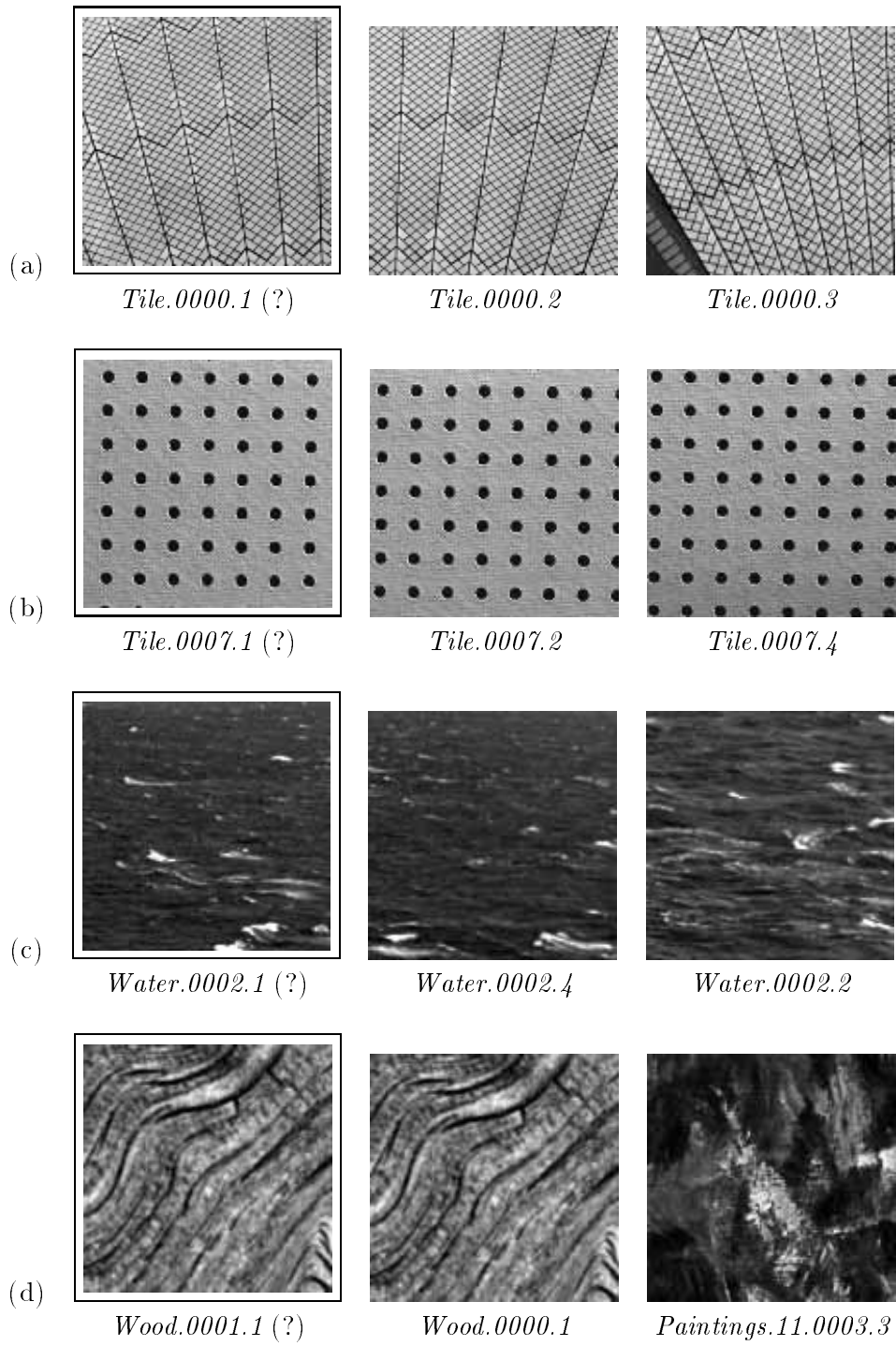


FIGURA 7.20: Experiments sobre DBI2

Criteri 1	Criteri 2	Criteris 1 i 4	Criteris 2 i 4
88.8%	96.7%	93.6%	97.9%

La segona BDI està formada per imatges de la mateixa base de dades del nostre segon experiment. Els resultats obtinguts sobre 360 preguntes, en una base de dades on hi ha 4 imatges d'una mateixa imatge, són:

Criteri 1	Criteri 2	Criteri 3
66.6%	73.3%	77.8%

Tot i que creiem que aquesta avaluació no és la més adient, hem fet les mateixes avaluacions. Per la BDI1 i per 328 preguntes hem obtingut els següents percentatges d'encert:

Criteri 1	Criteri 2	Criteri 3
84.4%	90.8%	93.5%

Sobre la BDI2 i per 668 imatges hem obtingut:

Criteri 1	Criteri 2	Criteri 3
64.9%	75.5%	83.5%

La diferència d'encerts entre la primera i la segona base de dades és degut principalment que la segona base de dades conté moltes imatges on realment no hi ha cap textura, i que per tant el seu resultat no és predictable.

Per altra banda, el que caldria fer seria veure els resultats per un criteri com el 4, però la seva aplicació implica definir un conjunt de similituds entre grups que hem cregut millor no fer, ja que haguéssim pogut introduir certs dubtes a l'experiment. Aquesta definició de similituds s'hauria de fer amb persones no experimentades i amb experiments que tinguessin en compte la preatentivitat.

Si ho comparem amb els resultats del treball dels altres autors que hem introduït anteriorment, podem veure que per la segona base de dades els resultats són semblants. Cal tenir en compte que a la BDI2 hi havia més imatges escollides sense cap criteri.

CAPÍTOL 8

CONCLUSIONS I VIES DE CONTINUACIÓ

Per extreure les conclusions de tot aquest treball hem de tenir en compte els objectius que ens havíem fixat. L'objectiu central era el de definir una representació de textures que complís les propietats de ser general, preatentiva, amb un cost computacional baix, que el seu espai presentés una mètrica de similituds semblant a la del SVH, que fos de baixa dimensionalitat i en les que els eixos es poguéssin interpretar perceptualment.

Abans de definir la representació ens hem plantejat diferents alternatives. Si teníem en compte els treballs anteriors realitzats, la primera idea seria plantejar una representació que es basés en la mesura sobre les imatges de certes propietats amb un significat d'ordre superior associat. Tenint en compte aquest tipus d'aproximació hem presentat els resultats d'aplicar aquestes idees sobre una aplicació concreta.

L'aplicació consisteix en una classificació d'imatges radiogràfiques afectades amb pneumoniosi. Per fer la classificació ens hem basat en una mesura concreta: la granularitat.

L'anàlisi dels resultats d'aquesta aplicació ens ha portat a les següents conclusions: la definició de la mesura influeix molt en el resultat, i el tipus d'imatge fa que els resultats d'una mateixa mesura es puguin interpretar de forma diferent. En resum, aquest tipus d'aproximació pot presentar cert grau d'ambigüitat i manca de generalitat en el tipus d'imatge que pot representar.

Per evitar aquests dos problemes ens hem plantejat buscar una representació basada en característiques que permeten una discriminació preatentiva en el sentit de la teoria del textó de Julesz [52]. En aquest cas, la textura ha de ser representada a partir de característiques locals de la imatge com taques i barres, i dels seus atributs, independentment de la posició que aquestes ocupen dins de la imatge, és a dir, basant-se només en criteris de densitat d'aquestes característiques.

Una representació en el sentit que acabem de definir podria basar-se en un conjunt

d'algorismes que detectéssin les taques i les barres i que aleshores calculéssin els seus atributs. Aquests tipus d'aproximacions [120] presenten gran complexitat en la definició d'algorismes prou robustos i problemes en la combinació de tota la informació de forma global. Un opció més senzilla des del punt de vista metodològic consisteix en agafar un model multicanal de percepció preatentiva de textures, en els que tant la detecció com la combinació és més automàtica i més general davant de qualsevol tipus d'imatge. A més a més, aquests models estan basats en evidències biològiques i en resultats psicofísics que fan que puguin presentar un comportament d'acord amb el del SVH, que és una de les propietats que ens havíem proposat per aquesta representació.

A partir de la selecció d'un model d'aquest tipus [65], hem definit una representació, que ha presentat els problemes habituals en aquest tipus d'aproximacions:

- Cost computacional de càlcul elevat.
- La dimensionalitat de l'espai de representació molt gran.
- La manca d'una interpretació perceptual.

El cost computacional en el càlcul de la representació ha estat rebaixat de forma important a més a més amb la capacitat de ser paral·lelitzable. S'han introduït operacions morfològiques en els passos d'inhibició no lineal del model. Aquestes operacions aporten una millora important si tenim en compte que existeixen algorismes molt eficients per calcular-les. En l'intent de millorar aquests passos hem aportat un mètode per a la descomposició d'elements estructurants isotròpics. Aquest tipus d'elements s'han utilitzat pel càlcul de les dilatacions morfològiques que hem introduït a l'algorisme del model. El mètode desenvolupat ens permet passar d'una complexitat d'ordre quadràtic a una complexitat d'ordre lineal en el càlcul d'aquestes operacions.

Una altra aportació interessant és que hem demostrat que la representació que presentava una dimensionalitat molt alta, es pot reduir a una espai de quatre dimensions. Característica molt important per a qualsevol representació. Fixem-nos que l'estímul visual del color ve representat físicament en un espai d'infinites dimensions mentre que les seves representacions normalment es fan en espais de tres dimensions.

Finalment, hem explorat l'espai de representació amb l'objectiu de donar-li una interpretació segons certs paràmetres perceptuals objectius. La interpretació dels eixos de l'espai reduït s'ha fet en dos passos (en podem veure un esquema a la figura 8.1).

En el primer pas s'han identificat les possibles dimensions a partir d'un espai paramètric de textures. Aquest espai paramètric d'imatges, o espai d'entrada, ha permès controlar determinats paràmetres a l'entrada de manera que s'ha pogut veure l'efecte de les seves variacions a l'espai de representació. Així doncs, s'ha aconseguit una interpretació basada en els paràmetres d'escala, de contrast i d'orientació de les taques i les barres que apareixen a la imatge.

En el segon pas ens hem plantejat la generalització d'aquesta interpretació pel cas general de qualsevol textura, i no només per a textures paramètriques. Fer servir la mateixa metodologia pel cas general és impossible ja que no tenim l'espai paramètric de totes les textures possibles. El que hem fet ha estat definir una parametrització de la representació i una hipòtesi sobre la combinació de paràmetres d'una textura dins de l'espai.

La hipòtesi ha estat confirmada a partir d'una anàlisi de correlació, sobre les representacions d'un conjunt d'imatges naturals, un conjunt semblant al que altres autors [81] havien agafat com a mostra de l'espai de totes les textures.

La confirmació d'aquesta hipòtesi ens permet dir que una textura es localitza dins l'espai de representació a partir de la suma de les posicions de totes les subtextures que la formen. Considerarem una subtextura d'aquesta al conjunt de característiques de la imatge, taques o barres que presenten el mateix tipus d'atribut, aquests atributs són el contrast i el seu signe, l'escala i l'orientació, combinats entre ells segons certes funcions que ja hem especificat.

Com a línies de continuació d'aquest treball podem dir que caldria millorar la correlació d'un dels eixos de la representació. Concretament l'eix que representa el signe dels elements que formen la textura.

Aquesta millora aniria acompanyada d'un estudi més precís sobre l'estructura de l'espai de representació. De manera que poguéssim determinar l'origen, la relació entre els diferents eixos, la continuïtat dins de l'espai, etc. Aquest estudi es podria fer seguint la metodologia que estableix l'escalament multidimensional partint d'experiments orientats en aquesta direcció. En quan a aquest fet hem de dir que hem preferit testejar primer la viabilitat de la generalització de la interpretació abans d'entrar més a fons a definir l'estructura.

Una vegada hagim definit l'estructura de l'espai caldria validar-la amb l'espai perceptual corresponent. S'haurien de dur a terme un conjunt d'experiments amb persones que permetéssin establir l'estructura del nostre espai perceptual, per poder veure si existeix una correspondència. Aquest fet confirmaria la validesa de la mètrica de similituds de l'espai de representació definit.

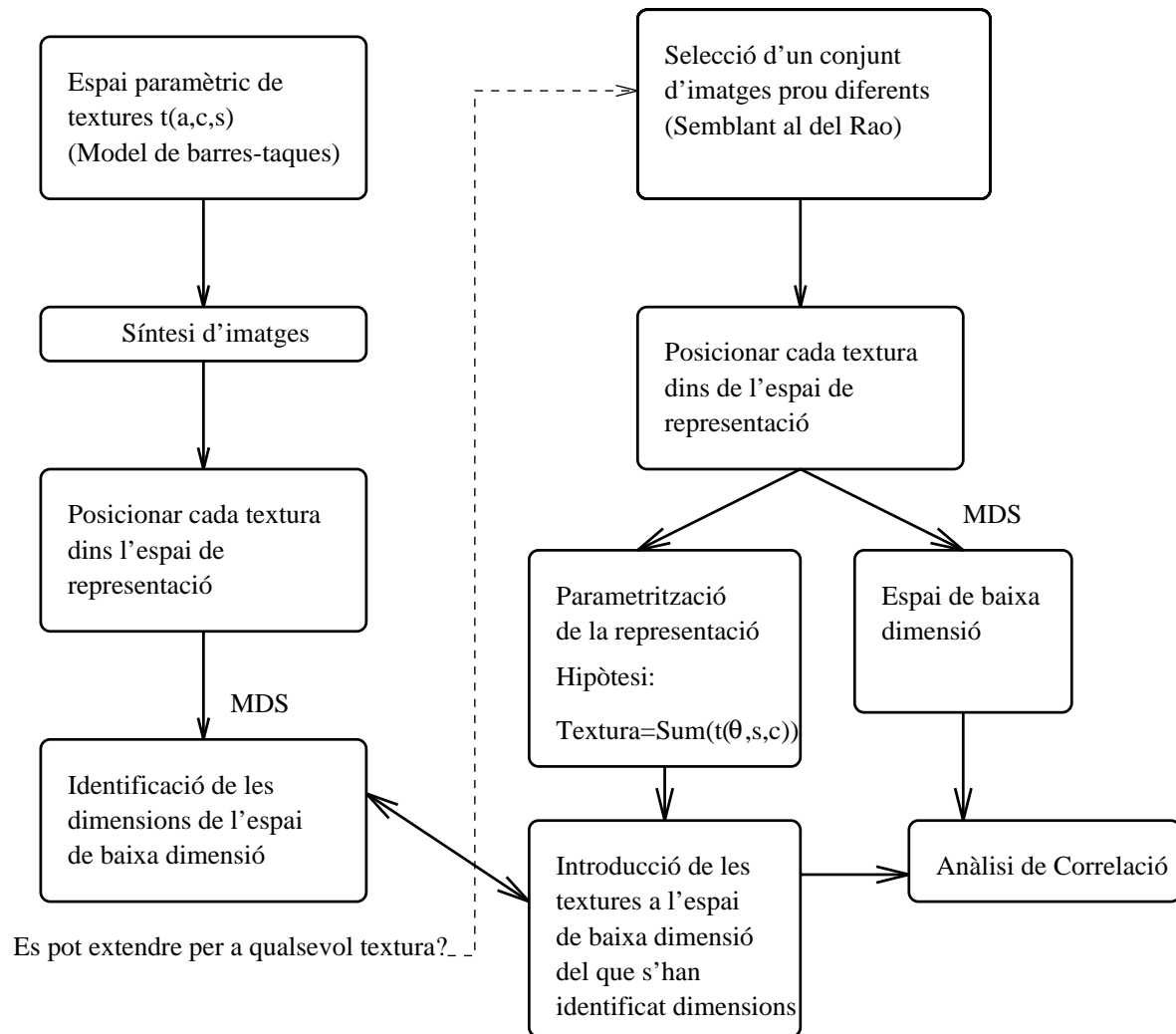


FIGURA 8.1: Esquema del procés seguit per a la interpretació dels eixos de l'espai de representació.

Per a que la generalització fos real caldria estendre tot el treball a imatges en color.

En quan al model utilitzat i a la representació definida caldria introduir algunes modificacions respecte de la sintonització del radi de la inhibició. En aquest sentit ja hi ha alguns resultats dins la teoria del textó que podrien dirigir aquesta recerca.

L'aportació feta amb el mètode de descomposició d'elements estructurants isotròpics podria ser generalitzada per a qualsevol dimensió de l'element bàsic. Aquesta generalització és important si tenim en compte que els *hardware's* que ens permetrien implementar aquest algorisme de forma eficient ofereixen la possibilitat de fer operacions amb elements estructurants més grans amb el mateix cost que elements petits.

Finalment, i referent al capítol d'aplicació de la representació a l'accés a bases de dades d'imatges, caldria una metodologia de validació millor, ja que la validació feta no reflecteix ben bé la capacitat de la representació.

APÈNDIX A

TEXTURES (I)

APÈNDIX B

TEXTURES (II)

BIBLIOGRAFIA

- [1] N. Ahuja and Azriel Rosenfeld. Mosaic models for textures. *IEEE*, PAMI-3:225–235, 1981.
- [2] Y. Aloimonos. *Active Perception*. Lawrence Erlbaum Associates, 1993.
- [3] J. Amat and A. Casals. *Computer Vision: Craft, Engineering and Science. Workshop Proceedings*, chapter Processors for Road Images Segmentation from Textures, pages 151–164. Springer-Verlag, 1994.
- [4] R. Bajcsy. Active perception. *Proceedings IEEE*, 76(8):996–1005, 1988.
- [5] J. Beck, A. Sutter, and R. Ivry. Spatial frequency channels and perceptual grouping in texture segregation. *Computer Vision, Graphics, and Image Processing*, 37:299–325, 1987.
- [6] J.R. Bergen and E.H. Adelson. Early vision and texture perception. *Nature*, 333:363–364, 1988.
- [7] J.R. Bergen and M.S. Landy. *Computational Models of Visual Processing*, chapter Computational Modeling of Visual Texture Segregation, pages 253–271. The MIT Press, 1991.
- [8] I. Blajot. *Radiología clínica del tórax*. Ediciones TORAY, 1977.
- [9] A.C. Bovik, M. Clark, and W.S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12:55–73, 1990.
- [10] P. Brodatz. *Textures*. New York: Dover, 1966.
- [11] Rodney A. Brooks. Intelligence without representation. *Artificial Intelligence*, 47:139–159, 1991.

- [12] D. Burr, C. Morrone, and D. Spinelli. Evidence of edge and bars detectors in human vision. *Vision Research*, 29:419–431, 1989.
- [13] M. José Carreira. Interpretación automática de imágenes de tórax basada en conocimiento. Master's thesis, Universidad de Santiago de Compostela, 1989.
- [14] A. Casals, J. Amat, and A. Grau. *Lecture Notes in Computer Science. Computer Vision - ECCV'92*, chapter Texture parametrization method for image segmentation, pages 160–164. Springer-Verlag, 1992.
- [15] X. Chen, J. Hasegawa, and J. Toriwaki. Quantitative diagnosis of pneumoconiosis based on recognition of small rounded opacities in chest x-ray images. In *9th International Conference on Pattern Recognition*, pages 462–464, 1988.
- [16] X. Chen, J. Toriwaki, and J. Hasegawa. Automated classification of pneumoconiosis radiographs based on recognition of small rounded opacities. *Systems and Computers in Japan*, 21(12):33–44, 1990.
- [17] D. Cheng and M. Goldberg. An algorithm for segmenting chest radiographs. In *Visual Communications and Image Processing*, pages 261–268. SPIE, 1988.
- [18] C. Chubb and M.S. Landy. *Computational Models of Visual Processing*, chapter Orthogonal Distribution Analysis: A New Approach to the Study of Texture Perception, pages 291–301. 1991.
- [19] M. Clark and A.C. Bovik. Experiments in segmenting texton patterns using localized spatial filters. *Pattern Recognition*, 22:707–717, 1989.
- [20] P. R. Cohen. *Empirical methods for Artificial Intelligence*. The MIT Press, 1995.
- [21] G.R. Cross and A.K. Jain. Markov random field texture models. *IEEE PAMI*, 5:61–75, 1983.
- [22] F. Cutzu and S. Edelman. Exploration of shape space. Technical report, Weizmann Institute, 1995.
- [23] S. Duvdevani-Bar and S. Edelman. On similarity to prototypes in 3d object representation. Technical report, Weizmann Institute, 1995.

- [24] S. Edelman. Representation is representation of similarities. Technical report, Weizmann Institute, 1996.
- [25] J.D. Foley et al. *Computer Graphics, Principles and Practice*. Addison-Wesley Publishing Company, 1990.
- [26] M. Flickner et al. Query by image and video content: The qbic system. *Computer*, 28 (9):23–32, 1995.
- [27] M. Coster et J. Chermant. *Précis d'Analyse d'Images*, chapter Granulométrie, pages 120–161. Editions du Centre National de la Recherche Scientifique, 1985.
- [28] I. Fogel and D. Sagi. Gabor filters as texture discriminator. *Biological Cybernetics*, 61:103–113, 1989.
- [29] H. Freeman, editor. *Machine Vision for Semiconductor Wafer Inspection*, chapter Machine Vision for Inspection and Measurement, pages 283–315. Academic Press, 1989.
- [30] A. Gagalowicz. *Vers un modèle de textures*. PhD thesis, Université Pierre et Marie Curie. Paris VI, 1983.
- [31] J.J. Gibson. *The ecological approach to visual perception*. Lawrence Erlbaum Associates, 1986.
- [32] G.L. Gimel'farb and A.K. Jain. On retrieving textured images from an image database. *Pattern Recognititon*, 29:1461–1483, 1996.
- [33] L.van Gool, P. Dewaele, and A. Oosterlinck. Survey: Texture analysis anno 1983. *Computer Vision, Graphics and Image Processing*, 29:336–357, 1985.
- [34] J.C. Gower. Some distance properties of latent root and vector methods in multivariate analysis. *Biometrika*, 53:315–328, 1966.
- [35] N. Graham. *Computational Models of Visual Processing*, chapter Complex Channels, Early Local Nonlinearities, and Normalization in Texture Segregation, pages 274–289. 1991.

- [36] N. Graham and A. Sutter. Effect of spatial scale and background luminance on the intensive and spatial nonlinearities in texture segregation. *Vision Research*, 36:1371–1390, 1996.
- [37] A. Haas, G.Matheron, and J.Serra. Morphologie mathématique et granulometries en place. *Annales des Mines*, (XI):735–782, 1967.
- [38] E.L. Hall, W.O. Crawford, and Jr.F.E. Roberts. Computer classification of pneumoconiosis from radiographs of coal workers. *IEEE Transactions on Biomedical Engineering*, BME-22(6):518–527, 1975.
- [39] R.M. Haralick. Statistical and structural approaches to texture. In *Proceedings of de IEEE*, pages 304–322, May 1979.
- [40] R.M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Trans. on System, Man and Cybernetics*, SMC-3:610–621, 1973.
- [41] R.M. Haralick and L.G. Shapiro. *Computer and Robot Vision*, chapter Texture, pages 453–507. Addison-Wesley Publishing Company, 1992.
- [42] J.K. Hawkins. *Picture Processing and Psychopictorics*, chapter Texture, pages 347–370. Academic Press, 1970.
- [43] Instructions pour l’utilisation de la classification internationale du bit des radiographies de pneumoconioses. Série Sécurité, Hygiène et Médecine du travail 22, Bureau International du Travail, 1980.
- [44] M.J. Norusis/ SPSS Inc., editor. *SPSS Professional Statistics 6.1*. 1994.
- [45] J. Roger Jagoe and Keith A. Paton. Measurements of pneumoconiosis by computer. *IEEE Transactions on Computers*, pages 95–97, 1976.
- [46] D. Jeulin. Morphological modeling of images by sequential random functions. *Signal Processing*, 16:403–431, 1989.
- [47] B. Julesz. Experiments in the visual perception of texture. *Scientific American*, pages 34–44, 1978.
- [48] B. Julesz. Textons, the elements of texture perception, and their interactions. *Nature*, 290:91–97, 1981.

- [49] B. Julesz. Texton gradients: The texton theory revisited. *Biological Cybernetics*, 54:245–251, 1986.
- [50] B. Julesz. Early vision and focal attention. *Review on Modern Physics*, 63(3):735–772, 1991.
- [51] B. Julesz. *Dialogues on perception*. The MIT Press, 1995.
- [52] B. Julesz and J.R. Bergen. Textons, the fundamental elements in preattentive vision and perception of textures. *Bell Systems Technological Journal*, 62:1619–1645, 1983.
- [53] B. Julesz, E.N. Gilbert, and J.D. Victor. Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics*, 31:137–140, 1978.
- [54] B. Julesz and B. Krose. Features and spatial filters. *Nature*, 333:302–303, 1988.
- [55] V. Kayargadde and J. Martens. Perceptual characterization of images degraded by blur and noise: experiments. *Journal of the Optical Society of America*, 13(6):1166–1177, 1996.
- [56] V. Kayargadde and J. Martens. Perceptual characterization of images degraded by blur and noise: model. *Journal of the Optical Society of America*, 13(6):1178–1188, 1996.
- [57] R.P. Kruger, W.B. Thompson, and A.F. Turner. Computer diagnosis of pneumoconiosis. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-4(1):40–49, 1974.
- [58] J.B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29:1–27, 1964.
- [59] J.B. Kruskal. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29:115–129, 1964.
- [60] J.B. Kruskal and M. Wish. *Multidimensional Scaling*. Sage University Paper, 1978.
- [61] G. Lei, W. Chengke, and L. Xinru. A new method for computer recognition of small rounded pneumoconiosis opacities in chest x-rays. In *8th International Conference on Pattern Recognition*, pages 475–477, 1986.

- [62] C.C. Li, David B.C. Shu, et al. Computer analysis of ilo standard chest radiographs of pneumoconiosis. In *International Symposium on Medical Imaging and Image interpretation*, pages 424–428, 1982.
- [63] C.C. Li, H.T. Tai, et al. Computer-aided recognition of linear opacities in chest radiographs. In *Proceedings of the IEEE Conf. Cybern. and Soc.*, pages 754–759, 1980.
- [64] J. Malik and P. Perona. A computational model of texture segmentation. In *International Conference on Computer Vision and Pattern Recognition*, pages 326–332, 1989.
- [65] J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America*, 7:923–932, 1990.
- [66] K.V. Mardia, J.T. Kent, and J.M. Bibby. *Multivariate Analysis*, chapter Multidimensional Scaling, pages 394–423. Academic Press, 1995.
- [67] D. Marr. Analyzing natural images: A computational theory of texture vision. *Philosophical Transactions of the Royal Society of London*, 275:483–524, 1976.
- [68] D. Marr. *Vision*. W. H. Freeman and Company, 1982.
- [69] G. Matheron. *Random Sets and Integral Geometry*. John Wiley and Sons, 1975.
- [70] H.C. Nothdurft. Sensitivity for structure gradient in texture discrimination tasks. *Vision Research*, 25(12):1957–1968, 1985.
- [71] V.E. Ogle and M. Stonebraker. Chabot: Retrieval from a relational database of images. *Computer*, 28 (9):40–48, 1995.
- [72] H. Park and R.T. Chin. Optimal decomposition of convex morphological structuring elements for 4-connected parallel array processors. *IEEE Trans. on PAMI*, 16(3):304–313, 1994.
- [73] H. Park and R.T. Chin. Decomposition of arbitrarily shaped morphological structuring elements. *IEEE Trans. on PAMI*, 17(1):2–15, 1995.
- [74] A. Pentland, R.W. Picard, and S. Sclaroff. Photobook: Content-based manipulation of images databases. *International Journal of Computer Vision*, 18 (3):233–254, 1996.

- [75] T. Poggio, V. Torre, and C. Koch. Visual perception and regularization theory. *Nature*, 138:645–647, 1985.
- [76] F. Preteux. *Description et Interprétation des images par la Morphologie Mathématique. Application à l’Imagerie Médicale*. PhD thesis, Université Pierre et Marie Curie-Paris 6, 1987.
- [77] F. Preteux and M. Schmitt. *Image Analysis and Mathematical Morphology*, chapter Boolean Texture Analysis and Synthesis, pages 377–401. Academic Press, 1988.
- [78] V.S. Ramachandran. *AI and the Eye*, chapter Visual Perception in People and Machines, pages 21–77. John Wiley and Sons, 1990.
- [79] A.R. Rao. *A Taxonomy for Texture Description and Identification*. Springer-Verlag, 1990.
- [80] A.R. Rao and Gerald L. Lohse. Identifying high level features of texture perception. *CVGIP: Graphical Models and Image Processing*, 55:218–233, 1993.
- [81] A.R. Rao and Gerald L. Lohse. Towards a texture naming system: Identifying relevant dimensions of texture. *Vision Research*, 36:1649–1669, 1996.
- [82] A.R. Rao and Brian G. Schunck. Computing oriented texture fields. *CVGIP: Graphical Models and Image Processing*, 53:157–185, 1991.
- [83] D. Sagi. *Channels in the visual nervous system: Neurophysiology, Psychophysics and Models*, chapter Spatial filters in texture segmentation tasks, pages 397–424. London: Freund Publishing House, 1991.
- [84] D. Sagi. *Early Vision and Beyond*, chapter The Psychophysics of Texture Segmentation, pages 69–78. The MIT Press, 1991.
- [85] D. Sagi and B. Julesz. ”where” and ”what” in vision. *Science*, (228):1217–1219, 1985.
- [86] D. Sagi and B. Julesz. Short-range limitations on detection of feature differences. *Spatial Vision*, 2:39–49, 1987.
- [87] G. Sànchez. Segmentació i anàlisi de textures lineals en documents gràfics. Master’s thesis, Departament d’Informàtica. Universitat Autònoma de Barcelona, 1996.

- [88] N. Sarkar and B.B. Chaudhuri. An efficient approach to estimate fractal dimension of textural images. *Pattern Recognition*, 25(9):1035–1041, 1992.
- [89] A.M. Savol, C.C. Li, and R.J. Hoy. Computer-aided recognition of small rounded pneumoconiosis opacities in chest x-rays. *IEEE, PAMI-2*(5):479–482, 1980.
- [90] S.S. Schiffman, M.L. Reynolds, and F.W. Young. *Introduction to Multidimensional Scaling. Theory, Methods and Applications*. Academic Press, 1981.
- [91] J. Serra. *Image analysis and mathematical morphology*. Academic Press, 1982.
- [92] J. Serra. Descriptions of flatness and roughness. *Journal of Microscopy*, 134(3):227–243, 1984.
- [93] J. Serra. *Image analysis and mathematical morphology II. Theoretical Advances*. Academic Press, 1988.
- [94] J. Serra and L. Vincent. Lecture notes on morphological filtering. Les Cahiers du Centre de Morphologie Mathematique 8, Ecole Nationale Superieure des Mines de Paris, 1989.
- [95] R.N. Shepard. The analysis of proximities: multidimensional scaling with an unknown distance function. i and ii. *Psychometrika*, 27:125–140,219–246, 1962.
- [96] R.N. Shepard. Cognitive psychology: A review of the book by u. neisser. *Ameriacan Journal of Pshychology*, 81:285–289, 1968.
- [97] R.N. Shepard. Multidimensional scaling, tree-fitting, and clustering. *Science*, 210:390–210, 1980.
- [98] R.N. Shepard. Toward a universal law of generalization for psychological science. *Science*, 237:1317–1323, 1987.
- [99] D.B.C. Shu, C.C. Lin, et al. Piecewise linear detection of small thin objects in radiographs. In *IEEE Conf. Cybern. and Society*, pages 357–363, 1981.
- [100] P. Soille. *Morphologie Mathematique: du Relief a la Dimensionalite. Algorithmes et Methodes*. PhD thesis, Université Catholique de Louvain, 1992.

- [101] P. Soille, J. Serra, and J.F. Rivest. Dimensional measurements and operators in mathematical morphology. In *Nonlinear Image Processing III*, pages 127–138. SPIE, 1992.
- [102] I. Spence and J. Graef. The determination of the underlying dimensionality of an empirically obtained matrix of proximities. *Multivariate Behavioral Research*, 9:331–342, 1974.
- [103] I. Spence and J.C. Oglivie. A table of expected stress values for random rankings in nonmetric multidimensional scaling. *Multivariate Behavioral Research*, 8:511–517, 1973.
- [104] R.K. Srihari. Automatic indexing and content-based retrieval of captioned images. *Computer*, 28 (9):49–56, 1995.
- [105] H. Tamura, S. Mori, and T. Yamawaki. Textural features corresponding to visual perception. *IEEE Trans. on System, Man and Cybernetics*, SMC-8:460–473, 1978.
- [106] W.S. Torgerson. *Theory and Methods of Scaling*. Wiley, New York, 1958.
- [107] A. Treisman. Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing*, 31:156–177, 1985.
- [108] M. Tuceryan and A.K. Jain. *Handbook of Pattern Recognition and Computer Vision*, chapter Texture Analysis, pages 235–276. World Scientific, 1993.
- [109] M. Unser. *Description Statistique de Textures: Application l'Inspection Automatique*. PhD thesis, Ecole Polytechnique fédérale de Lausanne, 1984.
- [110] M. Vanrell. Anàlisi granulomètrica de textures: Aplicació a la classificació de la pneumoconiosi. Master's thesis, Departament d'Informàtica. Universitat Autònoma de Barcelona, 1992.
- [111] M. Vanrell. Descomposició general d'elements estructurants. aplicació a la descomposició de discs. Technical Report 1, CVC, 1995.
- [112] M. Vanrell. Exploring the behavior of a texture perception algorithm. Technical Report 12, CVC, 1996.

- [113] M. Vanrell and X. Roca et al. A general morphological framework for perceptual texture discrimination based on granulometries. In *Mathematical Morphology and its Applications to Signal Processing*, pages 112–118, 1993.
- [114] M. Vanrell, X. Roca, and J. Vitrià. A multidimensional scaling approach to explore the behavior of a texture perception algorithm. *Machine Vision and Applications*, pages 262–271, 1997.
- [115] M. Vanrell and J. Vitrià. Mathematical morphology, texture perception and granulometries. In *Image Algebra and Morphological Image Processing IV*, pages 152–161. SPIE's Int. Symp. on Optical Instrumentation and Applied Science, 1993.
- [116] M. Vanrell and J. Vitrià. 3×3 decomposition of circular structuring elements. In *International Conference on Image Processing*, volume 3, pages 5–8. IEEE, 1996.
- [117] M. Vanrell and J. Vitrià. Optimal 3×3 decomposable disks for morphological transformations. *Image and Vision Computing*, 1997. In Press.
- [118] H. Voorhees. Finding texture boundaries in images. Technical Report 968, MIT Artificial Intelligence Laboratory, 1987.
- [119] H. Voorhees and T. Poggio. Detecting textons and texture boundaries in natural images. In *First International Conference on Computer Vision*, pages 250–258, 1987.
- [120] H. Voorhees and T. Poggio. Computing texture boundaries from images. *Nature*, 333:364–367, 1988.
- [121] H. Wechsler. Texture analysis - a survey. *Signal Processing*, 2:271–282, 1980.
- [122] A.P. Witkin. Recovering surface shape and orientation from texture. *Artificial Intelligence*, 17:17–45, 1981.
- [123] J. Xu. Decomposition of convex polygonal morphological structuring elements into neighborhood subsets. *IEEE Trans. on PAMI*, 13(2):153–162, 1991.
- [124] F.W. Young, R.A. Faldowski, and M.M. McFarlane. *Handbook on Statistics 9 - Computational Statistics*, chapter Multivariate statistical visualization, pages 959–998. North Holland, 1993.

- [125] R.A. Young. Simulation of human retinal function with the gaussian derivative model. In *International Conference on Computer Vision and Pattern Recognition*, pages 564–569, 1986.
- [126] X. Zhuang. Decomposition of morphological structuring elements. *Journal of Mathematical Imaging and Vision*, 4(1):5–18, 1994.
- [127] X. Zhuang and R.M. Haralick. Morphological structuring element decomposition. *Computer Vision, Graphics and Image Processing*, 35:370–382, 1986.
- [128] S.W. Zucker. Toward a model of texture. *Computer Graphics and Image Processing*, 5:190–202, 1975.