



UNIVERSITAT POLITÈCNICA DE CATALUNYA

DEPARTAMENT D'ÒPTICA I OPTOMETRIA

Reconeixement òptic de patrons  
amb discriminació variable  
per mitjà de la correlació no-lineal dual.  
Aplicació al reconeixement d'objectes en color.

ELISABET PÉREZ CABRÉ

Memòria presentada per optar al grau de Doctora en Ciències Físiques

Terrassa, desembre 1998

Na Dra. María Sagrario Millán García-Varela, Catedràtica d'Escola Universitària de la Universitat Politècnica de Catalunya

CERTIFICA

que N'Elisabet Pérez Cabré, llicenciada en Ciències Físiques, ha realitzat sota la seva direcció i en el departament d'Òptica i Optometria de la Universitat Politècnica de Catalunya, el treball "Reconeixement òptic de patrons amb discriminació variable per mitjà de la correlació no-lineal dual. Aplicació al reconeixement d'objectes en color", que es recull en aquesta memòria per optar al grau de Doctora en Ciències Físiques.

I per a què consti d'acord amb la legislació vigent, signa aquest certificat

Dra. María Sagrario Millán García-Varela

Terrassa, 1 de desembre de 1998

Per a la realització d'aquest treball s'ha comptat amb el finançament i els recursos de diverses institucions. A totes elles, voldria expressar el meu agraïment. Especialment,

- A la Direcció General de Recerca de la Generalitat de Catalunya, per la concessió d'una beca predoctoral de formació d'investigadors, i per la borsa de viatge concedida per a la realització d'una estada a la Universitat de Varsòvia. Així mateix, pels ajuts obtinguts dins el marc de la Xarxa Temàtica en Processament Òptic d'Imatges, projectes XT94-39 i XT97-33.
- A la Comisión Interministerial de Ciencia y Tecnología (CICYT), pels projectes dins del *Plan Nacional de Tecnologías Avanzadas de la Producción*, nº TAP93-0667-CO3-02, TAP96-1015-CO3-02 i TAP97-1089.
- Al Departament d'Òptica i Optometria, pels recursos humans i materials amb els quals he comptat durant la realització de la tesi doctoral, i pel finançament parcial d'una segona estada curta a la Universitat de Varsòvia.
- A l'Institut de Geofísica, per la seva acollida i finançament parcial durant les dues estades realitzades a la Universitat de Varsòvia.
- A l'Instituto Valenciano de Investigaciones Agrarias (IVIA), principalment al Sr. Rafael Bono, per facilitar-me material fotogràfic de varietats de cítrics.

# Índex

<b>Capítol 1. Introducció i propòsit .....</b>	<b>1</b>
1.1 Reconeixement òptic de patrons .....	1
1.2 Reconeixement d'objectes en color .....	10
1.3 Propòsit d'aquesta memòria .....	17
1.4 Esquema del treball .....	22
<b>Capítol 2. Fonaments del reconeixement òptic d'objectes. Mètodes de filtratge i correladors optoelectrònics .....</b>	<b>25</b>
2.1 Funció correlació .....	26
2.2 Correlació òptica .....	27
2.2.1 Correlador clàssic de Vander Lugt o $4f$ .....	27
2.2.2 Correlador convergent .....	30
2.2.3 Correlador de transformades conjuntes .....	32
2.2.4 Comparació del correlador de Vander Lugt (VLC) i del correlador de transformades conjuntes (JTC) .....	35
2.3 Filtratge adaptat .....	38
2.3.1 Filtres de potències fraccionàries .....	39
2.3.2 Filtratge no-lineal .....	41
2.3.3 Avaluació del comportament dels filtres .....	43
2.4 Correlació multicanal en el reconeixement d'objectes en color .....	45
2.4.1 Canals $\lambda_n$ .....	45
2.4.2 Descomposició i correlació multicanal RGB .....	48
2.4.3 Llindars i criteris de decisió .....	50

2.5 Característiques dels dispositius optoelectrònics (càmeres CCD i moduladors espacials de llum) .....	52
2.5.1 Càmeres d'estat sòlid amb sensors CCD .....	53
2.5.2 Moduladors espacials de llum (SLM) .....	56

### **Capítol 3. Selecció de canals quasimonocromàtics per a la millora de la discriminació cromàtica en el reconeixement d'objectes en color.....59**

3.1 Introducció .....	59
3.2 Correlació multicanal adaptada. Selecció d' $n$ canals quasimonocromàtics .....	62
3.3 Resposta dels filtres de potències fraccionàries (FPF) per objectes similars en forma i de color uniforme .....	65
3.4 Aplicació del mètode de selecció adaptada de canals .....	68
3.4.1 Selecció de canals adaptada a les característiques de les escenes test .....	68
3.4.2 Descomposició multicanal experimental .....	74
3.4.2a Sistema de descomposició RGB .....	75
3.4.2b Sistema de descomposició adaptada en $n$ canals quasimonocromàtics .....	76
3.4.3 Correlacions simulades .....	78
3.4.3a Gamma magenta .....	78
3.4.3b Gamma taronja .....	81
3.4.3c Escena d'objectes naturals (cítrics) .....	84
3.5 Resultats experimentals de reconeixement per correlació òptica .....	86
3.5.1 Correlador òptic convergent .....	87
3.5.2 Resultats de reconeixement experimentals .....	90
3.5.2a Gamma magenta .....	90
3.5.2b Gamma taronja .....	95
3.5.2c Escena d'objectes naturals (cítrics) .....	98
3.6 Reconeixement d'objectes policromàtics .....	101

### **Capítol 4. Correlació no-lineal dual (DNC) .....107**

4.1 Definició de la DNC .....	109
4.2 Propietats de la DNC .....	115
4.2.1 Invariància a translacions .....	115
4.2.2 Variació de la DNC amb la intensitat del feix de llum que il·lumina l'escena .....	116
4.2.3 Pic d'autocorrelació no-lineal dual .....	117
4.2.4 Escenes multiobjecte .....	118
4.3 Modificació de la regió de suport mitjançant la funció $T(\vec{v})$ i els seus efectes en la discriminació d'objectes texturats .....	120
4.3.1 Cas de reconeixement I .....	125
4.3.2 Cas de reconeixement II .....	131
4.3.3 Cas de reconeixement III .....	134
4.4 Realització òptico-digital de la DNC .....	136
4.4.1 Dispositiu i processament simplificat .....	141

<b>Capítol 5. Anàlisi numèrica de les condicions experimentals per a la realització de la DNC .....</b>	<b>147</b>
5.1 Introducció .....	147
5.2 Saturació de la càmera CCD .....	153
5.3 Quantització del sistema de captació .....	160
5.4 Nivell de zero en el càlcul de la no-linealitat (paràmetre $\epsilon$ ) .....	167
5.5 Quantització del modulador espacial de llum (SLM) .....	173

**Capítol 6. Sistema de reconeixement amb capacitat de discriminació variable .....** **179**

6.1 Introducció .....	179
6.2 Variacions de forma entre els objectes a analitzar .....	182
6.2.1 Descripció del test 1 .....	183
6.2.2 Mètodes de filtratge convencionals (CMF, POF, IF, PPC) .....	185
6.2.3 Aplicació de la DNC .....	189
6.3 Variacions d'intensitat (nivell de gris) entre els objectes a analitzar .....	196
6.3.1 Descripció del test 2 .....	196
6.3.2 Mètodes de filtratge convencionals (CMF, POF, IF, PPC) .....	197
6.3.3 Aplicació de la DNC .....	200
6.4 Conclusió .....	209

**Capítol 7. Realització experimental de la DNC .....** **211**

7.1 Introducció .....	211
7.2 Descripció del dispositiu experimental .....	212
7.3 Calibració del sistema de captació .....	217
7.4 Caracterització del modulador espacial de llum .....	221
7.4.1 Característiques generals de les LCTV del videoprojector Epson model VP-100PS. Paràmetres i comportament espacial .....	222
7.4.2 Determinació del règim de modulació d'amplitud de la pantalla LCTV .....	229
7.5 Posada en funcionament del correlador no-lineal dual .....	232
7.6 Resultats experimentals de la correlació no-lineal dual .....	237
7.6.1 Variació de la capacitat de discriminació del correlador no-lineal dual davant variacions de forma .....	238
7.6.2 Variació de la capacitat de discriminació del correlador no-lineal dual davant variacions d'intensitat (nivell de gris) .....	253

## **Capítol 8. Aplicació de la DNC al reconeixement d'objectes en color ..... 267**

8.1 Variacions de forma i color entre els objectes a analitzar .....	269
8.1.1 Descripció del test 3 .....	271
8.1.2 Mètodes de filtratge convencionals (CMF, POF, IF, PPC) .....	274
8.1.3 Aplicació de la DNC .....	278
8.1.3a Discriminació a canvis de forma i tolerància a canvis en el color dels objectes analitzats (zona A) .....	281
8.1.3b Tolerància a canvis en la forma i discriminació a canvis en el color dels objectes analitzats (zona B) .....	284
8.1.3c Tolerància a canvis en la forma i en el color dels objectes analitzats (zona C) .....	286
8.1.3d Discriminació elevada respecte a la forma i al color dels objectes analitzats (zona D) .....	288
8.2 Descomposició adaptada en canals quasimonocromàtics aplicada al reconeixement d'objectes en color mitjançant la DNC .....	291
8.2.1 Tolerància a canvis de forma i discriminació a canvis en el color dels objectes analitzats .....	292
8.2.2 Discriminació elevada respecte a la de forma i al color dels objectes analitzats .....	293

## **Capítol 9. Conclusions ..... 295**

## **Annex I. Escenes i corbes de reflectància espectral ..... 305**

I.1 Gamma marró .....	306
I.2 Gamma blava .....	308
I.3 Gamma verda .....	310

## **Annex II. Publicacions relacionades amb el treball ..... 313**

## **Bibliografia ..... 317**

# Capítol 1

## Introducció i propòsit

### 1.1 Reconeixement òptic de patrons

El reconeixement òptic de patrons constitueix una de les línies de recerca bàsiques en els camps de la percepció artificial, la robòtica, la teledetecció, etc. Les contribucions en aquest camp, són nombroses i segueixen augmentant, fets indicatius de l'interès per aquesta matèria i la seva evolució constant i progressiva. Aquest important desenvolupament es deu a la possibilitat – i a l'hora al repte – de reproduir mitjançant sistemes de visió artificial, les mateixes tasques que actualment realitzen les persones, de manera que sigui possible lliurar-les de treballs perillosos, feixucs, repetitius i on la subjectivitat pugui afectar els resultats de la feina. Aquestes tasques les podem trobar en molts àmbits laborals, des del manteniment d'instal·lacions on es tracten matèries perilloses, la vigilància de llocs de difícil accessibilitat, el muntatge de components per a la fabricació d'un producte, el control de qualitat dels productes finals de qualsevol empresa, a l'anàlisi i interpretació d'imatges mèdiques, de satèl·lit, etc.



Per comparació amb els éssers humans, una gran part de la informació que procedeix de l'exterior, es rep a partir del sistema visual en forma d'imatges. El processament d'aquesta informació comporta, doncs, el tractament d'una quantitat molt important de dades. El principal objectiu dels sistemes de visió artificial radica en l'extracció de dades de l'entorn pel seu processament i, finalment, per la seva interpretació a fi i efecte de poder prendre decisions. D'aquí, es pot entendre que el processament d'imatges en general i, dins d'ell, el reconeixement de patrons, constitueixen camps molt rellevants dins el marc general de la percepció artificial i la robòtica.

El reconeixement òptic es basa fonamentalment, en la correlació entre l'escena que s'analitza i l'objecte buscat. L'operació de correlació és una mesura del grau de semblança de les dues funcions que es correlacionen. Els sistemes òptics capaços de realitzar la correlació entre dues funcions s'anomenen correladors i, en ells, cal situar, en els plans adjacents, l'escena a analitzar i un filtre adaptat a l'objecte a reconèixer. L'operació de correlació ens proporciona una funció  $c(\vec{x})$ , el valor de la qual en cada punt del pla  $\vec{x} = (x, y)$ , és proporcional al volum en comú sota les dues funcions bidimensionals (imatges)  $s(\vec{x})$  i  $r(\vec{x})$  que es correlacionen, a mesura que es desplacen una sobre l'altra. Matemàticament, la correlació s'expressa

$$c(\vec{x}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\vec{\xi}) r^*(\vec{\xi} - \vec{x}) d\vec{\xi}, \quad (1.1)$$

on  $*$  denota la funció complexa-conjugada.

Si una de les funcions correspon a l'objecte buscat,  $r(\vec{x})$ , i aquest objecte és present a l'escena d'entrada,  $s(\vec{x})$ , en el pla de correlació, llavors, la presència de l'objecte es detecta a través d'un màxim d'intensitat localitzat en la mateixa posició que ocupa l'objecte en l'escena. Això és degut a què és en aquesta posició on es produeix la màxima coincidència entre les dues funcions.

Podem situar els inicis del reconeixement òptic de patrons, l'any 1964, quan Vander Lugt va proposar un mètode de filtratge espacial complex [1]. Des de llavors, les

tècniques, les arquitectures i els algorismes s'han desenvolupat i millorat per construir correladors òptics eficients per a les aplicacions de reconeixement de patrons.

Els sistemes purament òptics proporcionen als correladors l'avantatge de processar en paral·lel i en temps real una gran quantitat d'informació, però presenten l'inconvenient que, mitjançant mètodes totalment òptics, no és possible programar-los, ni tampoc poden prendre decisions per ells mateixos. Aquests inconvenients són, però, els punts fort de l'electrònica i, sembla natural, per tant, combinar ambdues disciplines per obtenir sistemes híbrids òptico-digitals que gaudeixin dels avantatges d'ambdós camps. De fet, ha estat amb aquestes combinacions com més s'ha avançat cap a sistemes de reconeixement versàtils, ràpids i eficaços.

Existeixen, fonamentalment, dues arquitectures de correladors òptics. El primer a ésser introduït, va ser el correlador clàssic de Vander Lugt (VLC, de l'anglès *Vander Lugt Correlator*) o, altrament anomenat  $4f$  per la disposició dels seus elements [1]. Dos anys més tard, en el 1966, es va introduir el correlador de transformades conjuntes (JTC, de l'anglès *Joint Transform Correlator*) proposat simultàniament i de manera independent per Weaver i Goodman [2] i per Rau [3] l'any 1966.

En el correlador de Vander Lugt és necessària la utilització d'un filtre adaptat a l'objecte a reconèixer, que consisteix en una transparència, la transmitància de la qual conté la transformada de Fourier complexa-conjugada de l'objecte que es vol detectar. El correlador de transformades conjuntes, en canvi, requereix situar en un mateix pla les funcions que descriuen l'escena a analitzar i l'objecte a detectar separades una distància suficient per evitar solapaments entre el terme que proporciona la correlació i altres termes que apareixen en el pla de sortida del sistema òptic. En el JTC, l'espectre de potència conjunt (JPS, de l'anglès *Joint Power Spectrum*), que conté la distribució en intensitat de la transformada de Fourier (TF) del pla d'entrada del correlador, juga un paper tan rellevant com el filtre adaptat en el cas del VLC.

Inicialment, quan les funcions a reproduir i els plans de correlació s'enregistraven mitjançant tècniques de fotografia o holografia, el VLC va tenir un gran desenvolupament ja que el temps de processament era més curt, en comparació del JTC. Un cop es van introduir elements optoelectrònics que substituïen el procés fotogràfic i hologràfic, el JTC permetia realitzar una operació de correlació en temps real, sense necessitat de filtre adaptat i, a més, es podia modificar la informació de l'escena a analitzar i de l'objecte a reconèixer de manera immediata. Aquests dispositius optoelectrònics consistien bàsicament en càmeres amb detectors CCD (de l'anglès *Charge coupled-device*) i moduladors espacials de llum o SLM (de l'anglès *Spatial Light Modulator*). En aquest moment, el JTC obtenia els millors avantatges en quant a temps de processament i va adquirir, per tant, un nou impuls que va permetre el seu desenvolupament molt ràpid.

Per altra banda, el fet de poder obtenir hologrames generats per ordinador per a la realització de filtres adaptats generalitzats amb la possibilitat de reproduir-los en dispositius d'amplitud com els moduladors, ha permès reduir molt el temps de processament en els correladors òptics tipus VLC. Actualment, les dues arquitectures conviuen i, juntament amb altres dissenys basats en les configuracions bàsiques del VLC i del JTC, són les arquitectures constituents dels dispositius optoelectrònics més emprats en el reconeixement òptic de patrons.

A partir de les configuracions clàssiques del VLC i del JTC, el següent pas va consistir en la introducció de nous algorismes que permetien un processat de la informació a correlacionar. Segons el problema a resoldre, es potencia un tipus d'informació determinada de les funcions que es correlacionen. Així per exemple, és possible detectar un objecte per una característica determinada (contorns, informació interna, color, etc.), discriminar-lo millor d'altres objectes, reconèixer-lo fins i tot en presència d'algun tipus de distorsió (canvi d'escala, rotacions, etc.) o degradació del senyal (per exemple, presència de soroll), etc. Per tant, a partir d'aquest moment, es desenvolupa de manera molt important el disseny de filtres adaptats en el VLC i, el processament de l'espectre de potències conjunt (JPS) en el JTC. D'aquesta manera

s'aconsegueix diversificar i perfeccionar el procés de reconeixement d'objectes per correlació òptica.

Recollim a continuació, algunes de les propostes més significatives de modificacions dels filtres adaptats del correlador de Vander Lugt i del JPS del correlador de transformades conjuntes, per a la millora del reconeixement òptic d'objectes. Farem esment de distintes línies de desenvolupament com les que milloren la detectabilitat i qualitat del pic de correlació, l'eficiència del sistema, l'augment de la capacitat discriminativa, el reconeixement simultani de diversos objectes, el reconeixement invariant davant de distorsions (produïdes per girs, canvis d'escala, etc.) i el reconeixement estadístic per a la classificació.

Des del filtre clàssic de Vander Lugt [1], s'han proposat mètodes de processament que persegueixen, en general, augmentar la intensitat i la qualitat del pic de correlació per facilitar la detecció de l'objecte i la seva localització, si aquest és present a l'escena analitzada. Paral·lelament, també s'ha buscat augmentar la capacitat de discriminació del processador. En aquest sentit, trobem el filtre només de fase (POF, de l'anglès *Phase-Only Filter*) [4], o el filtre invers (IF, de l'anglès *Invers Filter*) [5] que posteriorment s'engloben en la definició dels filtres de potències fraccionàries (FPF, de l'anglès *Fractional Power Filters*) [6], els filtres només de fase binaris [7] i els filtres ternaris [8] com a millores en la capacitat discriminativa dels anteriors, o també el filtre de mínima energia mitja de correlació (MACE, de l'anglès *Minimum Average Correlation Energy*) [9]. Una anàlisi comparativa del comportament d'aquests filtres en el reconeixement òptic d'objectes, es pot trobar en [10]. El filtre invers i, de manera més general, els FPF van ser introduïts inicialment com a filtres adaptats per a un correlador de Vander Lugt, però posteriorment es va proposar el mètode per a ser realitzats en un correlador de transformades conjuntes [11,12]. També existeix una proposta, la dels filtres localment no-lineals, on es reuneixen, en una única definició de filtre adaptat i variant en freqüència, les característiques dels filtres CMF, POF i IF [13].

En quant al JTC, l'objectiu d'augmentar la capacitat de discriminació del sistema i millorar el pic de correlació, s'ha portat a terme mitjançant la introducció d'un

processat intermedi no-lineal del JPS. La primera proposta en aquest sentit, va ser formulada per Javidi i Kuo [14] i va consistir en la binarització del JPS aprofitant la característica binària de l'SLM que empraven, obtenint un increment en la intensitat del pic de correlació, una reducció de l'amplada del pic, i una millora de la sensibilitat discriminativa del sistema. El valor de tall per portar a terme la binarització del JPS s'ha escollit de diferents maneres. En l'anàlisi d'escenes multiobjectes, Javidi i col·laboradors han proposat varis mètodes, un basat en la mitjana del JPS [15], un segon mètode de tall bidimensional per a maximitzar l'eficiència del sistema [16], i posteriorment, va proposar un llindar dependent de les freqüències espacials i a partir de la mitjana d'un subconjunt [17]. Altres mètodes de binarització es poden trobar en [18] aplicats a la detecció i classificació d'objectes en imatges d'infraroig. En el cas d'escenes lliures de soroll, s'ha avaluat el valor del llindar més adient com a compromís entre l'eficiència i la discriminació del sistema [19]. S'han aplicat també mètodes de difusió de l'error per determinar el llindar [20].

Un altre exemple de no-linealitat aplicada al JTC correspon a l'anomenat JTC ajustat a les franges (*Fringe-adjusted JTC*) [21], que es caracteritza per tenir un temps de computació menor que JTC binari, i per obtenir una millor realització de la correlació sota canvis en la il·luminació del pla d'entrada [22].

Tots els treballs mencionats fins aquí en relació al correlador de transformades conjuntes, tenen en comú el processament de l'espectre de potències conjunt (JPS) un cop enregistrat mitjançant un detector d'intensitat. En [23] els autors proposen modificar, no l'amplitud d'aquesta funció, sinó, prèviament al seu enregistrament, la fase de l'espectre de Fourier corresponent al pla d'entrada del correlador. L'operació portada a terme, rep el nom de binarització de la fase, i es realitza a partir de la distribució de fase de l'espectre de l'objecte de referència. Els autors proposen una realització òptica en temps real d'aquesta binarització de la fase emprant un modulador treballant en règim de fase i un polaritzador, disminuint així el temps de processament.

En [24] s'afegeix una nova proposta de filtratge no-lineal en el JTC, de manera que el filtratge es porta a terme mitjançant un pre-processament de la informació de

l'objecte de referència abans d'ésser introduïda en el pla d'entrada del correlador. En aquest treball, es comparen els resultats de correlació obtinguts havent realitzat un filtratge en l'espai objecte (és a dir, a la informació de la referència), i els obtinguts amb el més tradicional filtratge en el pla de Fourier (per tant, processant el JPS). Una altra idea de pre-processament de la informació d'entrada del processador, és la proposada en [25], on la imatge corresponent al pla d'entrada del JTC (escena + referència) es reproduceix en un modulador en règim de modulació de fase. D'aquesta manera, una imatge habitualment en amplitud es transforma en una imatge només de fase. Es demostra que aquest sistema obté una major discriminació i eficiència lluminosa comparat amb el JTC convencional. Una proposta similar es troba en [26]. Combinant la representació del pla d'entrada del JTC codificat en fase, amb la binarització del JPS i la posterior codificació en fase, s'obtenen bons resultats de correlació, fins i tot treballant amb nivells molt baixos d'il·luminació [27]. En aquest mateix treball, Ledesma *et al.* proposen altres modificacions del JPS, com ara la codificació en fase del JPS, i també l'extracció de la fase de l'espectre de Fourier conjunt.

Paral·lelament, la codificació en fase de les escenes en el cas d'un correlador de Vander Lugt, es presenta com un procediment per a la millora de l'eficiència del sistema i un increment en la capacitat discriminativa del processador [28]. Seguint amb els correladors basats en el mètode de filtratge adaptat (VLC i derivats), la proposta d'introduir un filtratge no-lineal en el procés de correlació va ser realitzada per Ersoy i Zeng [29]. Aquests autors van proposar una generalització dels filtres adaptats POF i POF binari, de manera que s'aplicava la mateixa operació sobre la TF de l'escena i de la referència abans de ser multiplicades en l'espai de freqüències. D'aquesta manera va sorgir el filtre simètric només de fase, i el simètric només de fase binari. És a partir d'aquest estudi que es parla de filtratge no-lineal en el VLC. A partir d'aquesta proposta, es desenvolupa per diferents autors el que s'anomena l'extracció de la fase [30,31], i la correlació només de fase [32,33].

La introducció d'operacions no-lineals comporta, en moltes ocasions, l'aparició d'efectes no desitjats, com ara la presència de falses alarmes generades per termes

harmònics que s'amplifiquen, i l'increment del valor màxim de l'ordre central en el pla de sortida del correlador, que provoca una disminució en l'eficiència del pla de correlació. Nombrosos estudis s'han portat a terme per determinar aquests efectes negatius introduïts amb les operacions no-lineals [34-38]. S'han buscat alternatives a la binarització per evitar els efectes d'intermodulació en el reconeixement multiobjecte [39-41], per eliminar l'ordre zero [42-46] i els termes de correlació redundants entre els objectes de la mateixa escena [47,48].

Réfrégier [49-51] va proposar els anomenats filtres de compromís (*trade-off filters*) pel correlador de Vander Lugt que optimitzen varis criteris de qualitat, aconseguint un balanç òptim entre la discriminació i la robustesa al soroll en el pla d'entrada del correlador. Posteriorment, Réfrégier *et al.* [52,53] dissenyen un processador òptim en els mateixos termes. Aquest processador òptim consisteix en l'aplicació d'una operació no-lineal que es pot realitzar en l'arquitectura del JTC.

Una línia molt important de recerca en el camp del disseny de filtres adaptats i de processaments del JPS per al reconeixement òptic de patrons, és la del reconeixement invariant a distorsions per canvis d'escala i rotacions dels objectes. Les aportacions més destacades es centren en el disseny d'un únic filtre adaptat amb la informació de l'objecte a reconèixer transformada a una base que permeti la invariància. Una transformació de coordenades aplicada al filtre només de fase (POF) permet obtenir un sistema de reconeixement invariant a desplaçament i a canvis d'escala [54,55]. En quant a distorsions degudes a rotacions de l'objecte, s'introdueix l'expansió en harmònics circulars per a la resolució d'aquesta tasca de reconeixement [56-60]. La mateixa expansió en harmònics circulars s'aplica per a l'obtenció d'un reconeixement amb invariància a rotacions i al contrast [61].

Les rotacions fora de pla dels objectes, amb els conseqüents canvis de perspectiva, constitueixen una altra font de distorsió que degraden el pic de correlació en un sistema de reconeixement i, en el cas d'objectes plans, es pot considerar més senzillament com un canvi d'escala [62]. Per aconseguir un reconeixement invariant a aquests canvis de perspectiva, s'ha introduït l'expansió en harmònics logarítmics [63].

Els filtres composts van ser introduïts per Caulfield i Maloney en [64,65], i posteriorment desenvolupats per Casasent i col·laboradors [66,67]. Vijaya Kumar presenta en [68] un treball de recopilació dels diferents tipus de filtratge adaptat inclosos en el disseny dels filtres composts per a ser emprats en correladors del tipus VLC i derivats d'ell. Recentment (1997), Javidi *et al.* [69] estudien els filtres adaptats no-lineals composts per a la seva aplicació en el reconeixement d'objectes en diferents orientacions degut a girs fora de pla, i amb la presència de diferents tipus de soroll en el pla d'entrada. També en el 1997, trobem una nova aplicació dels filtres composts en el tractament d'imatges on hi ha objectes distorsionats per rotacions fora de pla [70]. El conjunt de filtres utilitzats en aquest treball, anomenats filtres composts híbrids, proporcionen resultats complexos de correlació, l'amplitud dels quals indica la posició de l'objecte a detectar, i la fase, convenientment tractada, informa de l'orientació particular de l'objecte detectat. En [71] es presenta una realització experimental dels filtres composts en un JTC. I seguint amb les aportacions destinades a ésser implementades en un JTC, trobem el treball de Tang i Javidi [72] que estudien el comportament del JTC no-lineal en presència d'objectes amb diferent escala i orientació en el pla d'entrada del correlador.

S'han emprat eines típiques del processament digital d'imatges, com ara la morfologia matemàtica, per a la realització òptica d'operacions aplicades en el tractament d'imatges [73] i, posteriorment, s'han introduït en el reconeixement òptic d'objectes [74]. La correlació morfològica no-lineal, realitzada experimentalment en [74], aconsegueix detectar un objecte d'intensitat menor que altres objectes presents en una escena, a partir d'un pic d'autocorrelació, que s'obté com el pic més intens.

Propostes recents de variacions de la configuració clàssica del JTC que incorporen nous materials i combinacions amb altres sistemes òptics, les podem trobar en [75-77]. En el primer treball, Khoury *et al.* [75] proposen l'ús d'un material fotorefractiu per portar a terme processos d'holografia en temps real. Amb el dispositiu proposat, obtenen la correlació entre dues funcions a partir d'un processador només de fase. Un altre material, el bacteriorhodopsin, és utilitzat per a l'obtenció d'un SLM adreçat òpticament amb l'avantatge de tenir una gran resolució espacial, possibilitant així, la



construcció d'un JTC de dimensions molt compactes [76]. En [77] es descriu un sistema que combina un interferòmetre Mach-Zehnder amb una configuració de JTC sense l'ús de lents. Aquest sistema és apropiat per determinar, amb alta resolució espacial, la posició relativa de dues fonts puntuals (*pinholes*).

Com qualsevol dispositiu òptic, el correlador pot veure afectat els seus resultats de correlació en més o menys mesura en funció de la presència de diferents tipus d'aberracions. En [78] es determina analíticament l'efecte de les aberracions geomètriques de primer ordre sobre el JPS en el primer pas del correlador de transformades conjuntes, i s'ha estimat numèricament la seva influència sobre el pic de correlació.

## 1.2 Reconeixement d'objectes en color

La introducció de la informació cromàtica dels objectes en el reconeixement de patrons suposa una millora en la capacitat discriminativa dels sistemes de reconeixement al permetre la identificació d'objectes no només per la seva forma sinó també per la seva distribució de color, i a més, significa un acostament del sistema de percepció artificial al sistema visual humà, molt més sensible a variacions de color que no pas d'intensitat.

Per comprendre la dimensió del problema, si es planteja la identificació d'objectes en escenes policromàtiques com un problema de reconeixement monocromàtic d'un sol canal, on només s'utilitza una longitud d'ona com a il·luminant, pot ésser impossible la detecció de l'objecte buscat si els contrastos en la imatge no són prou importants, o bé es poden produir falses alarmes si objectes de colors diferents apareixen com a semblants en el canal monocromàtic. És per aquest motiu que cal trobar maneres d'introduir la informació del color en el procés de reconeixement. El fet que la descripció del color sigui trivariant [79] suggereix, com a primera aproximació, la utilització de tres canals primaris centrats en les regions vermella, verda i blava (RGB, de l'anglès *red, green, blue*) de l'espectre visible. En aquests tres canals es porta a terme, el que s'anomena correlació multicanal.

La correlació multicanal és, doncs, la que s'efectua mitjançant una descomposició prèvia dels senyals de l'objecte a reconèixer i de l'escena en components, que són les que es correlacionen en paral·lel en canals d'informació independents. Els resultats obtinguts en cada canal es combinen seguint criteris de decisió o són les entrades d'alguna fórmula preestablerta. Al final s'obté un únic resultat de reconeixement pel qual s'afirma o es rebutja la presència de l'objecte buscat en l'escena.

Previ a la correlació multicanal, és necessari definir els canals i el mecanisme de descomposició dels senyals. La correlació multicanal no és, però, un procediment exclusiu del reconeixement d'objectes en color. Per exemple, fora d'aquest camp, citem dues propostes molt recents. En [80-82] es proposa l'ús d'un esquema de filtres de Gabor multiescala i multiorientació per a definir els canals d'informació i obtenir la descomposició dels senyals. En [74] es proposa una descomposició en làmines (*slices*) dels senyals en nivells de gris i la correlació en paral·lel de les esmentades components, que pot ser també considerada com una altra forma de correlació multicanal.

Tot i els dos exemples comentats, es pot dir que un dels casos més representatius de descomposició multicanal és el que s'aplica a les imatges en color. Així, prenent la descomposició RGB abans esmentada com a punt de partida, s'han proposat diferents arquitectures per realitzar-la òpticament, s'han aplicat mètodes, tant òptics com digitals, i s'han desenvolupat altres sistemes de canals, que a continuació descriurem, amb l'objectiu de millorar la capacitat de discriminació del sistema.

Badiqué *et al.* proposen una formulació per a la correlació digital d'imatges en color [83-85]. Un vector tridimensional (R, G, B) de la imatge en color es transforma en un vector imatge bidimensional per projecció en un pla de color. Amb l'obtenció d'un pla de color generalitzat òptim [84,85] per realitzar la correlació, aconsegueixen un filtre de reconeixement espacial-espectral que té una capacitat de discriminació més alta que el filtre que s'obté de la distribució d'intensitat. També han estudiat filtres invariants a la rotació i al canvi d'escala aplicats al reconeixement en color [86]. En [87] Pei i Hsu proposen una representació del logaritme complex de la informació de

l'objecte, basada en l'estructura retiniana humana, per a la millora del sistema de reconeixement invariament a rotacions i canvis d'escala dels objectes en color.

En el processament òptic de senyals, són destacables les contribucions de Yu en la utilització de llum blanca i de llum parcialment coherent [88,89]. En el reconeixement òptic en color, Yu proposa una tècnica de filtratge adaptat a objectes en color emprant xarxes de difracció [90]. Aquestes es posen en contacte amb la transparència d'entrada en la correlació i en la generació del filtre. Per a la realització hologràfica dels filtres, el dispositiu s'il·lumina amb tres fonts de llum coherents (vermella, verda i blava) i, mitjançant la xarxa de difracció, les tres transformades de l'escena apareixen separades en el pla de Fourier. Posteriorment, Yu i Hsu [91] demostren que la il·luminació del correlador en una banda ampla es pot fer servir pel reconeixement d'objectes en color. Yu i Javidi [92] introdueixen els moduladors òptics per aconseguir un reconeixement policromàtic en temps real. En el pla de correlació s'obtenen uns pics de color en la posició on hi ha l'objecte i el color indica la informació cromàtica d'aquest. En [93] posen de manifest els avantatges de la il·luminació parcialment coherent per reduir els efectes de distorsió en la fase introduïts pels moduladors espacials de llum.

Warde *et al.* [94] proposen un mètode de filtratge adaptat espacial-espectral en temps real. Utilitzant una xarxa codificadora per a tres colors i un modulador espacial de llum, l'objecte de l'entrada es descompon en un conjunt d'imatges coherents monocromàtiques. Un correlador de transformades conjuntes permet que la correlació es faci en temps real. Una simplificació d'aquest sistema la proposen Ludman *et al.* [95] on només cal un feix làser, i posteriorment Javidi *et al.*, [96] introdueixen els moduladors binaris magneto-òptics en els quals reproduïxen la informació de color de l'escena una vegada codificada amb una xarxa de difracció i binaritzada. El següent pas ha estat la introducció de les pantalles de televisió de cristall líquid (LCTV, de l'anglès *liquid-crystal television*) amb reproducció en color de les imatges. Yu *et al.* van introduir aquests dispositius pel reconeixement d'objectes en color tant en els correladors del tipus  $4f$  amb filtres adaptats [97], com en els de transformades conjuntes [98]. Més recentment, Wang *et al.* han proposat

L'ús d'aquestes pantalles combinat amb la utilització d'una tècnica incoherent amb filtres hologràfics de Fresnel multiespectrals (*multiwavelength*) en el procés de correlació [99]. Té l'avantatge d'un requeriment d'alineació dels components més relaxat, d'una major eficiència de difracció, i de la no aparició de termes de correlació creuada entre les diferents components cromàtiques, donat que els filtres hologràfics es separen espacialment per components, i es situen just al seu darrera uns filtres de colors primaris amb la component corresponent.

En el camp del reconeixement estadístic d'imatges, l'objectiu del qual és la identificació i classificació dels objectes, també s'han proposat mètodes per introduir la informació cromàtica de les mostres avaluades. En aquesta línia, Gu *et al.* [100] presenten dos mètodes de processament híbrid basats en la correlació òptica amb il·luminació incoherent i en el disseny de filtres estadístics adaptats a la informació espectral i espacial dels objectes a reconèixer. El primer mètode consisteix en la correlació seqüencial en els tres canals quasimonocromàtics RGB dels tubs de raigs catòdics, i en el segon mètode utilitza simultàniament la informació en paral·lel dels mateixos canals RGB, però proporcionats per un monitor de TV.

Per objectes policromàtics, Millán *et al.* [101] varen proposar un mètode de reconeixement per correlació multicanal. Es tracta de descompondre la informació d'amplitud i de color de l'escena en tres components cromàtiques, la vermella (R), la verda (G) i la blava (B). Mitjançant mètodes hologràfics i utilitzant feixos làsers de diferent longitud d'ona, es pot obtenir un filtre adaptat a l'objecte a reconèixer per cada un dels canals RGB. L'escena, enregistrada sobre diapositiva en color, es situa en el pla d'entrada d'un correlador. Donat que la seva transmitància dependrà de la longitud d'ona de l'il·luminant emprat, es pot realitzar una correlació monocromàtica per separat en cada canal si s'utilitza el mateix feix làser que el que ha generat el filtre adaptat situat en el correlador. Si només es considera la informació obtinguda en un sol canal, sovint apareixen contrastos poc accentuats entre els objectes de color que, en molts casos, es tradueixen en falsos reconeixements, anomenats falses alarmes. No obstant, tenint en compte els resultats de les correlacions monocromàtiques de tots els canals conjuntament, s'aconsegueix la detecció d'un

objecte tant per la seva forma com per la seva distribució de color. En [102] es presenten resultats experimentals obtinguts amb aquest mètode de correlació multicanal, i en [103] s'afegeixen els filtres d'harmònics circulars per aconseguir un reconeixement d'objectes en color invariant a rotacions.

De vegades, com en el reconeixement de caràcters en color, pot interessar reconèixer un aspecte parcial de l'objecte, per exemple la forma del caràcter independentment del color. En [104], Millán, *et al.* plantegen diferents estratègies a seguir segons es vulgui identificar un caràcter alfanumèric d'una determinada combinació de color caràcter-fons, o bé si es pretén reconèixer un caràcter independentment de la seva combinació cromàtica amb el fons. En el primer cas, cal aplicar l'operació lògica AND als resultats de les correlacions de cada objecte amb el patró obtinguts en cada un dels canals. Pel contrari, caldrà modificar el senyal al qual s'adapta el filtre i el criteri de decisió si es vol detectar la presència d'un determinat caràcter sense que la seva combinació de color sigui una informació rellevant.

Els mecanismes emprats per a la introducció de l'escena en el correlador han estat estudiats en el cas de la diapositiva en color [105] i de la càmera de vídeo [106]. L'escena captada i digitalitzada, s'emmagatzema en tres matrius corresponents a les components RGB. La diferència entre aquestes i les components cromàtiques obtingudes en el sistema de canals monocromàtics [101-104] radica en què l'ús de la il·luminació làser proporciona informació de la transmitància de l'escena en uns canals molt estrets de l'espectre, mentre que mitjançant l'ús de la càmera CCD, la informació de cada component prové d'una banda ampla de l'espectre on la sensibilitat del sensor és alta, amb l'efecte afegit de la discretització del sistema d'adquisició. En [106] s'analitza de quina manera el procés de reconeixement està influenciat per paràmetres com el *gain* i l'*offset*, controlables a través del software que governa la targeta digitalitzadora en el procés de captació. És necessari, per tant, establir en un primer moment, quins valors d'aquests dos paràmetres són més adients de manera que el senyal enregistrat es vegi alterat mínimament. També s'ha analitzat com reduir les distorsions cromàtiques introduïdes en el procés d'adquisició mitjançant càmeres en color [107].

Continuant amb la idea del reconeixement multicanal, García *et al.* proposen en [108] la utilització d'un únic filtre adaptat del tipus MACE (de l'anglès *minimum average correlation energy*) per aconseguir un reconeixement d'objectes tenint en compte llur informació de color. Es proposa la realització de la correlació multicanal en un sol pas, reproduint l'escena en una transparència o modulador en color, i il·luminant el sistema amb llum policromàtica. Amb l'ús d'elements difractius és possible obtenir les correlacions dels diferents canals cromàtics en una mateixa posició, independentment del contingut espectral de l'il·luminant [109], de manera que enregistrant un únic pla de sortida del sistema en un medi monocrom s'obté òpticament, el resultat final del reconeixement a partir de la suma de les intensitats dels pics de correlació dels diferents canals sobreposats en el mateix pla. Afegint el fet que el filtre MACE es generalitza a una expansió d'harmònics circulars, el procés de reconeixement portat a terme és invariant a rotacions de l'objecte buscat [110]. En [111], trobem una altra proposta de realització de la correlació multicanal en un sol pas del correlador, i en el qual s'obté un únic pla de sortida del processador.

Actualment també s'estudia la cerca d'objectes en una posició determinada de l'escena a analitzar (per exemple, la localització d'un segell a la cantonada superior d'una carta), per la qual cosa no es vol mantenir la propietat d'invariància a translacions pròpia de la correlació. Per aquesta aplicació s'ha introduït el que s'anomena correlació fraccionada, i la seva aplicació en el reconeixement en color la trobem en [112].

Les tècniques descrites fins aquí tenen en comú la descomposició de la informació del color en els canals cromàtics convencionals RGB, en els quals es realitza un procés multicanal ja sigui de manera seqüencial o en paral·lel. El resultat final d'aquest procés de reconeixement, s'obté després d'aplicar funcions llindar i operacions lògiques als resultats obtinguts de manera independent en cada un dels canals, o bé, es pot obtenir un únic pla de sortida del processador, si mitjançant tècniques òptiques es sobreposen en una mateixa posició els plans de correlació obtinguts en els diferents canals cromàtics processats en paral·lel.

Una manera diferent d'introduir la informació del color en el procés de reconeixement, parteix de la realització d'un pre-processament de les imatges en color a analitzar i de l'objecte a reconèixer, perquè d'aquesta manera es porti a terme un procés de correlació òptica en un processador monocromàtic convencional. Dins les tècniques de pre-processament de la informació del color, trobem aportacions amb l'objectiu d'aconseguir una millora en la capacitat de discriminació del sistema de reconeixement. Les tècniques que esmentarem, tenen en comú que parteixen de les components convencionals RGB, i mitjançant càlculs numèrics entre aquestes, s'obtenen una o varies components noves en les quals es porta a terme el reconeixement.

En la línia d'apropar els sistemes de reconeixement artificial al sistema visual humà, trobem propostes de pre-processaments on s'introdueixen els models de percepció humana del color en els processos de reconeixement d'objectes en color [113,114]. En ambdós mètodes s'apliquen transformacions lineals sobre les components RGB proporcionades per una càmera, segons diferents models de visió basats en la teoria de colors oponents. En [113] s'utilitza el model de Hurvich-Jameson [115] pel reconeixement de caràcters de colors diferents a partir de la informació obtinguda del procés d'aprenentatge d'una placa neuronal. Per la seva banda, Millán *et al.* [114] introdueixen els models de Guth *et al.*, [116] i de Boynton [117] de manera que aconseguen augmentar la capacitat discriminativa del sistema de reconeixement respecte al sistema que utilitza els canals RGB, i al mateix temps, és possible reduir el nombre de canals en els quals es processa la informació. En la memòria de tesi doctoral presentada per Corbalán [118] es presenten altres mètodes de pre-processament de la informació de les imatges en color que consideren la influència de possibles canvis de l'il·luminant i tenen en compte transformacions colorimètriques.

Altres transformacions de les components RGB que prenen com a referència el codi de color de Munsell, han estat proposades pel reconeixement d'objectes en color [119,120].

Moreno *et al.* [121] proposen una transformació de les components RGB mitjançant una operació de blanquejat al llarg de l'eix de color, que en [122] es amplia amb la introducció de noves modificacions respecte la transformació original. Una altra proposta de codificació s'introdueix en [123] on la codificació de la informació del color es porta a terme mitjançant xarxes de difracció, variant la seva orientació, freqüència i amplitud. Les components cromàtiques RGB són codificades en funció de l'orientació de la xarxa, i la transmitància dels objectes policromàtics en cada una de les components es codifica en un cas amb la freqüència de la xarxa, i en un segon exemple, mitjançant l'amplitud de la mateixa. En [124] es proposen diferents pre-processaments que obtenen imatges amb valors reals bipolars. En concret, s'indiquen tres mètodes diferents. En dos d'ells, els que representen el valor absolut del senyal d'entrada i el que es codifica la informació només en fase, s'aconsegueix una única matriu bipolar que constitueix el pla d'entrada d'un correlador monocanal. En canvi, en el tercer mètode de pre-processament es descomposa l'escena en tres imatges positives, de manera que cal realitzar tres experiments de correlació i analitzar els resultats de cada un d'ells conjuntament per completar el procés de reconeixement. Amb els tres mètodes de pre-processaments esmentats, s'aconsegueix identificar un objecte policromàtic, i discriminar-lo d'altres de similars mentre que el sistema de reconeixement RGB no és suficient per resoldre el reconeixement.

Els avantatges dels sistemes aquí exposats, tracten d'ampliar la capacitat discriminativa dels sistemes de reconeixement, camp d'investigació que continua obert en l'actualitat i en el qual hem intentat realitzar algunes aportacions (annex II).

### **1.3 Propòsit d'aquesta memòria**

En aquest treball, ens plantegem obtenir un mètode general de reconeixement de patrons mitjançant la correlació òptica, que tingui una capacitat de discriminació variable, de manera que sigui possible adaptar-lo a diverses tasques de reconeixement amb diferents graus de tolerància o discriminació selectiva. És a dir, en un mateix sistema, sense necessitat de modificar la informació de l'escena d'entrada, ni la de l'objecte a detectar que s'introdueix en el correlador, puguem



modificar el grau de discriminació del sistema, o el que és equivalent, el seu grau de tolerància.

Respecte als mètodes de reconeixement emprats en l'actualitat, la capacitat discriminativa dels quals ve fixada per les característiques pròpies de cada mètode, la novetat de la nostra proposta consisteix en avançar cap a un control variable de la capacitat discriminativa del sistema de reconeixement, dins d'un mateix mètode o marc global que estigui governat, de manera senzilla, per uns pocs paràmetres. La introducció de la tolerància no es farà variant els llindars de decisió que s'apliquen en el pla de correlació per a reconèixer o discriminar un objecte d'un altre. És clar que una mesura tan immediata com aquesta, generaria un augment de falses alarmes, a més de ser indiscriminada, no permetent l'establiment d'una tolerància selectiva a determinats aspectes dels objectes en anàlisi (contorns, intensitat, color, textura, soroll, etc.). D'acord amb el nostre propòsit, dos objectes semblants es reconeixeran com el mateix o seran discriminats en funció de les exigències de discriminació del problema plantejat a les quals s'haurà d'adaptar el mètode de reconeixement. Les variacions d'una adaptació a una altra, al fer front a diferents tasques de reconeixement, no ha d'implicar modificacions substancials del mètode o sistema emprats, la qual cosa garantirà l'eficàcia i versatilitat del model global que es pretén.

Considerarem diferents canvis possibles en els objectes a analitzar, respecte als quals volem fer variable la capacitat de discriminació del sistema de reconeixement. Així doncs, estudiarem el sistema davant de canvis en la forma, en la intensitat, en el color, i en la textura dels objectes. Entre tots ells, dedicarem una especial atenció a les variacions de color, dins una mateixa gamma. En la línia de millora de la precisió d'un sistema, ens proposem augmentar la capacitat discriminativa quan s'analitzen objectes semblants, dins una mateixa gamma de tonalitat. Com a primera proposta per aconseguir aquest objectiu, considerarem una flexibilització del mètode convencional de descomposició multicanal RGB i, plantejarem una selecció adaptada de canals quasimonocromàtics per a la descomposició de la informació de l'escena i de l'objecte. L'adaptació dels nous canals es farà en funció de la gamma concreta a la

que pertanyen els colors dels objectes. Per a l'anàlisi d'objectes amb variacions de textura, forma o intensitat, considerarem el marc global de la correlació no-lineal dual (DNC, de l'anglès *Dual Non-linear Correlator*). Aquest mètode general, introduït en [42], abraça una gran varietat de mètodes de correlació lineals i no-lineals. En general, permet l'aplicació d'un operador no-lineal a la informació de l'escena diferent al que s'aplica a la informació de l'objecte a reconèixer. Les no-linealitats es poden modificar de manera gradual i controlada per dos paràmetres. El model de la DNC inclou també una funció que delimita la regió de suport o domini en el pla de Fourier on es poden aplicar les no-linealitats. Considerem que el mètode de la DNC ens pot oferir el marc global que busquem per a obtenir un control variable de la capacitat discriminativa del sistema a variacions de textura, forma i intensitat a través del control adequat dels paràmetres que regeixen les no-linealitats i de la definició de la regió de suport sobre la qual s'apliquen.

L'arquitectura base de la DNC és el disseny del correlador de transformades conjuntes (JTC), del qual s'obté l'avantatge de no necessitar filtres adaptats a l'objecte a reconèixer i la relaxació en quant a l'alineació dels seus components. En un únic dispositiu optoelectrònic, s'aconsegueixen reproduir diferents mètodes de correlació lineals i no-lineals dels descrits en la literatura i, a més s'afegeixen noves possibilitats – fins i tot intermèdies – de processament de la informació. D'aquesta manera, la realització de la correlació seguint diferents mètodes de filtratge dins la DNC s'aconseguirà variant uns paràmetres involucrats en el processament intermedi de l'espectre de potència conjunt (JPS). El fet d'englobar amb un sol sistema, diferents mètodes de filtratge lineals i no-lineals, farà possible la seva comparació i, per tant, també serà possible escollir el mètode més idoni per a resoldre un problema de reconeixement donat.

Dins dels tipus de canvis en els objectes pels quals es pretén obtenir una discriminació variable en el marc de la DNC, les variacions en la intensitat (o nivell de gris) dels objectes acromàtics se'ns presenta com un cas especialment interessant. Apart de l'interès que pugui tenir en si mateix el desenvolupament d'un sistema de reconeixement amb un control variable en la seva capacitat de discriminació respecte

a variacions en la intensitat dels objectes (acromàtics), és immediat donar un pas més enllà i aplicar els resultats que s'obtinguin en el reconeixement d'objectes en color. Al cap i a la fi, el reconeixement d'objectes en color es realitza mitjançant la correlació multicanal i, en cada canal, les components de l'escena i de l'objecte són distribucions d'intensitats (o nivells de gris) que es processen separatament respecte als altres canals.

Aquesta extensió del model de la DNC cap a una nova aplicació en el reconeixement d'objectes en color ens portarà a enriquir la recerca considerant situacions mixtes en el control de la discriminació dels objectes per la seva forma i color. Així, ens pot interessar que el sistema de reconeixement tingui un grau de discriminació més elevat respecte a un determinat aspecte dels objectes (per exemple, la forma) i, al mateix temps, presenti un nivell de tolerància major a canvis segons un aspecte diferent (per exemple, el color), o viceversa.

Per ser la DNC un mètode que es formula a partir de les funcions escena i objecte, independentment de possibles descomposicions multicanals prèvies (RGB, o d'altre tipus), resulta que l'aplicació que puguem desenvolupar en el marc de la DNC pel reconeixement d'objectes en color no és incompatible, sinó que podem considerar complementària, amb el mètode de descomposició multicanal inicialment proposat basat en una selecció adaptada de canals quasimonocromàtics. D'aquesta manera es veurien reforçades ambdues tècniques, per la possibilitat que ambdues presenten de poder combinar-se entre si. El sistema de reconeixement resultant de la combinació hauria de tenir una capacitat de discriminació variable i, al mateix temps, quan fos necessari, una alta discriminació dels objectes segons el seu color. En aquesta memòria analitzarem aquesta possibilitat.

Amb els objectius plantejats fins aquí, és clar que si s'aconsegueix un sistema amb el grau de discriminació adaptable al problema plantejat, podem aconseguir un mètode de reconeixement molt flexible i versàtil. Es tracta a continuació d'obtenir un dispositiu experimental que, sense modificar la seva arquitectura, ni el marc de processament establert per al reconeixement, sigui capaç de modificar la seva

capacitat de discriminació, podent-la fer molt sensible a petits canvis dels objectes, o tot el contrari, ser més tolerant a aquestes mateixes modificacions.

Ens plantejem doncs, portar a terme la realització experimental tant del sistema DNC amb capacitat de discriminació variable, com la correlació multicanal amb una descomposició de la informació cromàtica prèvia en uns nous canals quasimonocromàtics seleccionats. Per a la realització experimental d'aquestes dues propostes, emprarem dues arquitectures de correladors diferents.

En la realització experimental de la proposta que inclou una descomposició cromàtica en canals seleccionats per a l'anàlisi d'objectes similars en forma i amb colors propers, emprarem el correlador convergent [125], que constitueix una modificació del correlador de Vander Lugt. L'ús d'aquesta arquitectura ve justificada pel fet d'analitzar escenes multiobjecte. Els resultats de correlació en el cas del VLC i dels seus derivats no es veuen modificats pel nombre d'objectes presents en l'escena d'entrada, mentre que un JTC veu disminuir la intensitat dels pics de correlació a mesura que el nombre d'objectes es va incrementant [126]. Per aquest motiu, i tenint en compte que en el reconeixement d'objectes en color és necessari realitzar un sistema de correlació multicanal, és més convenient l'ús de l'arquitectura del VLC, o en el nostre cas, del correlador convergent, per tal de no augmentar innecessàriament el nombre d'operacions a realitzar.

En el cas de la DNC, com ja s'ha comentat, utilitzarem el correlador de transformades conjuntes, per treure avantatge de la no necessitat de generar filtres adaptats a l'objecte a reconèixer. D'aquesta manera, les diferents aplicacions de processat de la informació les portarem a terme canviant, de manera molt senzilla, uns paràmetres involucrats en el càlcul digital intermedi que es realitza sobre l'espectre de potència conjunt. La facilitat en l'alineació dels components del muntatge, juntament amb el ràpid temps de processament, permetran obtenir un sistema versàtil per diferents tasques de reconeixement. Com una anàlisi prèvia, realitzarem un estudi detallat d'alguns condicionaments experimentals que poden limitar o afectar la realització experimental de la DNC. Aquests condicionaments experimentals venen donats, bàsicament, per les característiques tècniques dels

dispositius optoelectrònics (càmera CCD i SLM) que emprarem com a elements del processador. Amb aquest estudi numèric volem posar de manifest la viabilitat de la proposta i, si és així, determinar quines són les condicions de treball dels dispositius en les quals s'obtenen uns millors resultats de reconeixement. També volem conèixer les limitacions reals que poden sorgir al portar a la pràctica, amb els dispositius que tenim a l'abast, el model teòric de la DNC.

## **1.4 Esquema del treball**

En el capítol 2 farem una breu descripció dels fonaments en el reconeixement òptic de patrons, de les arquitectures bàsiques i de les seves variacions més importants que seran emprades posteriorment en les diferents realitzacions experimentals. Detallarem també, els mètodes més importants de filtratge de la informació a processar, i mencionarem les possibilitats d'aplicacions dels diferents mètodes a les arquitectures que s'hauran detallat. Donarem alguns criteris d'avaluació en la realització de la correlació òptica mitjançant diferents mètodes de processament o filtratge. Introduïrem també, la correlació multicanal com a eina essencial en el reconeixement d'objectes en color. Descriurem els criteris de decisió que adoptarem per l'anàlisi final dels resultats de correlació, i els adaptarem en el cas de la correlació multicanal. Finalment, descriurem les característiques fonamentals dels dispositius optoelectrònics emprats en els correladors híbrids òptico-digitals.

En el capítol 3, proposarem la selecció adaptada de canals per a la millora de la capacitat de discriminació cromàtica quan el sistema de reconeixement requereixi distingir entre objectes molt semblants en forma i amb colors pertanyents a una mateixa gamma de tonalitat. Descriurem el procediment experimental de descomposició en les components cromàtiques seleccionades i, a partir d'elles, en un primer pas, realitzarem la correlació multicanal de manera simulada per conèixer l'abast del mètode. A continuació, realitzarem la correlació òptica de manera experimental. Presentarem resultats experimentals en el reconeixement d'objectes en color, iguals en forma i de color uniforme dins una mateixa gamma de tonalitat. Aplicarem aquest mètode a objectes procedents tant d'un test de prova com objectes

naturals. En tots els casos, els resultats obtinguts es contrastaran amb els proporcionats pel mètode de descomposició convencional RGB. Finalment, estudiarem la implicació de la proposta en l'anàlisi d'escenes que continguin objectes policromàtics.

Iniciarem el capítol 4, amb la descripció teòrica de la correlació no-lineal dual (DNC) [42]. A continuació, estudiarem les possibilitats de la DNC, en el reconeixement d'objectes texturats amb patrons senzills, mitjançant la variació de la funció que defineix la regió de suport involucrada en el processament digital. Descriurem el dispositiu optoelectrònic inicialment proposat en [42] i, proposarem simplificacions en el muntatge òptic i en l'algorisme digital per reduir-ne el cost i el temps de computació. Estudiarem l'equivalència d'ambdós dispositius, l'original i el simplificat.

L'estudi d'alguns condicionaments experimentals, imposats bàsicament pels dispositius optoelectrònics del correlador dissenyat, es portarà a terme en el capítol 5. En concret, analitzarem l'efecte sobre els resultats de la DNC de la saturació dels sensors de la càmera CCD, de la quantització del sistema d'adquisició, de l'efecte d'un paràmetre involucrat en el càlcul digital de la DNC equiparable a l'*offset* del sistema de captació i, per últim, de la quantització introduïda pel modulador espacial de llum. Analitzarem la viabilitat de la proposta teòrica de la DNC, i establirem els valors més adients dels paràmetres per a la posterior realització experimental.

En el capítol 6, s'analitza el sistema DNC per a l'obtenció d'una capacitat de discriminació variable, davant de canvis en la forma i en la intensitat dels objectes, aspectes que s'analitzen de manera separada. S'estudiarà la proposta tant de manera teòrica com tenint en compte els condicionament experimentals, analitzats en el capítol anterior. Es compararan els resultats obtinguts amb els proporcionats pels mètodes de filtratge convencionals.

En el capítol 7, es presentaran resultats experimentals de reconeixement amb tolerància variable. Es descriurà el sistema optoelectrònic emprat per a la obtenció dels resultats en el laboratori. Seguint el mateix esquema que en el capítol anterior,

s'analitzarà per separat, la DNC davant de canvis en la forma i en la intensitat dels objectes.

Continuarem en el capítol 8, utilitzant les conclusions estretes de l'anàlisi prèvia de la DNC, per a la seva aplicació en el reconeixement d'objectes en color. Considerarem un sistema de reconeixement amb un cert nivell de tolerància a canvis simultanis en la forma i en el color dels objectes, podent variar la capacitat de discriminació del sistema de manera separada respecte la informació de la forma i a la informació cromàtica. Donarem un pas més amb la combinació de la DNC amb el sistema de selecció adaptada de canals quasimonocromàtics.

Finalment, farem una breu revisió i exposarem les conclusions del treball realitzat en el capítol 9.

## **Capítol 2**

# **Fonaments del reconeixement òptic d'objectes. Mètodes de filtratge i correladors optoelectrònics**

Aquest capítol resumeix les idees bàsiques del reconeixement òptic d'objectes i mostra els esquemes dels correladors més habituals amb els quals portar a terme l'operació de reconeixement. Els fonaments exposats aquí constitueixen el punt de partida sobre el qual es desenvolupa aquesta tesi. Alguns dels dispositius i components descrits en aquest capítol seran eines emprades en els capítols posteriors.



## 2.1 Funció correlació

Una de les tècniques més esteses en el reconeixement de patrons es basa en la correlació entre l'escena que es vol analitzar i l'objecte que es vol detectar. Si les funcions  $s(\vec{x})$  i  $r(\vec{x})$ , en general complexes, representen respectivament l'escena i l'objecte a reconèixer, la correlació entre elles ve donada per

$$c(\vec{x}) = s(\vec{x}) \otimes r(\vec{x}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\vec{\xi}) r^*(\vec{\xi} - \vec{x}) d\vec{\xi}, \quad (2.1)$$

on el símbol  $*$  denota la funció complexa-conjugada i  $\otimes$  simbolitza la correlació. Aquesta funció mesura el grau de semblança entre dos objectes [127,128].

La correlació es pot expressar en termes de les transformades de Fourier (TF) de les funcions que es correlacionen. Siguin  $S(\vec{\nu})$  i  $R(\vec{\nu})$  les transformades de Fourier de  $s(\vec{x})$  i  $r(\vec{x})$ , respectivament, on la transformada de Fourier d'una funció bidimensional està definida per l'expressió

$$S(\vec{\nu}) = TF\{s(\vec{\xi})\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\vec{\xi}) \exp[-i2\pi(\vec{\nu} \cdot \vec{\xi})] d\vec{\xi}. \quad (2.2)$$

La correlació de les funcions  $s(\vec{x})$  i  $r(\vec{x})$  satisfà la propietat

$$c(\vec{x}) = s(\vec{x}) \otimes r(\vec{x}) = TF^{-1}\{S(\vec{\nu}) R^*(\vec{\nu})\}, \quad (2.3)$$

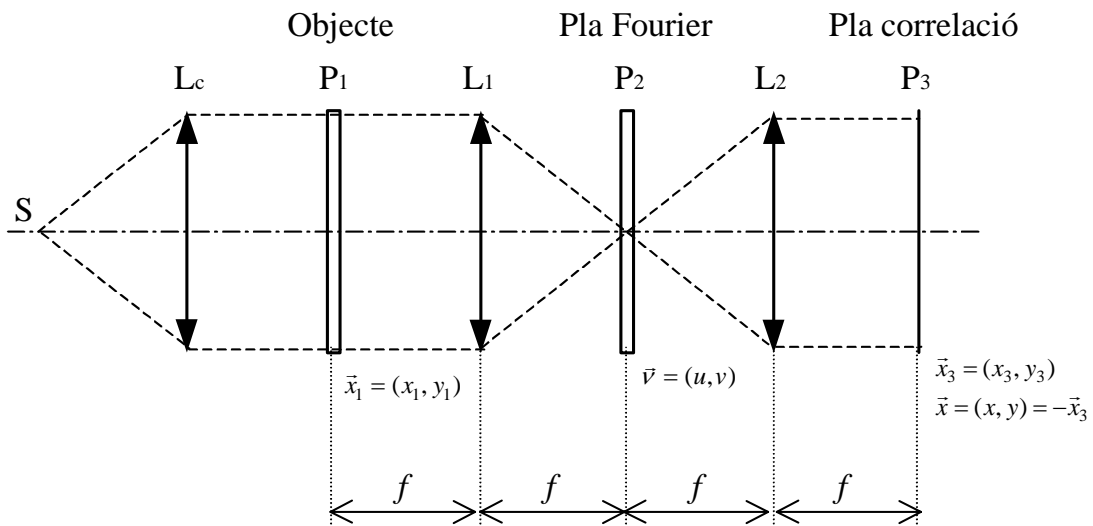
relació que ens indica que la correlació entre dues funcions es pot obtenir també, de manera molt senzilla, si realitzem el producte de les respectives transformades de Fourier en l'espai de freqüències (una d'elles complexa-conjugada), i realitzem una transformació de Fourier inversa del producte resultant.

## 2.2 Correlació òptica

Les tècniques òptiques permeten realitzar operacions complexes, com poden ser transformacions de Fourier o productes de funcions, processant en paral·lel i en temps real una gran quantitat d'informació. Amb aquestes dues operacions matemàtiques, la transformació de Fourier i el producte de funcions, és possible obtenir òpticament la correlació entre dues funcions tal com indica l'Eq. (2.3). Els sistemes capaços de realitzar-la, s'anomenen correladors òptics.

### 2.2.1 Correlador clàssic de Vander Lugt o $4f$

Considerem a continuació el dispositiu òptic de la Figura 2.1 que rep el nom de correlador clàssic de Vander Lugt (VLC, de l'anglès *Vander Lugt Correlator*) o  $4f$  per la posició relativa dels seus elements [1]. Aquest dispositiu permet realitzar òpticament l'operació de correlació descrita per l'equació (2.1) i va constituir el primer correlador òptic (1964). A continuació el descrivim.



**Figura 2.1.** Correlador clàssic de Vander Lugt o  $4f$ .

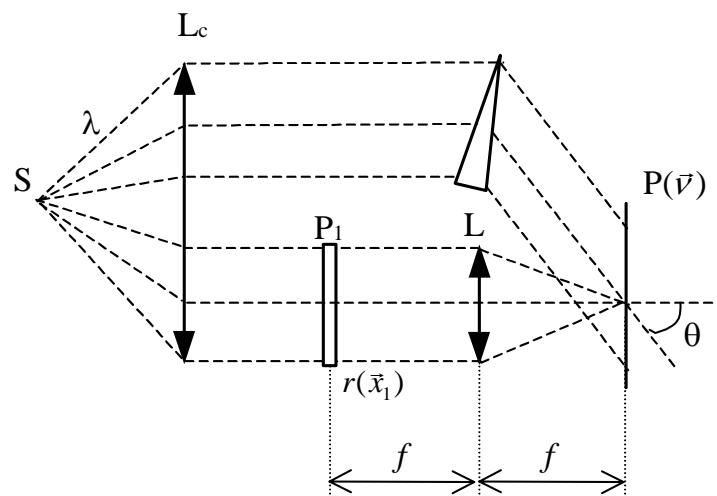
La llum provenint de la font puntual i monocromàtica  $S$  és col·limada per la lent  $L_c$ . L'ona plana incideix sobre el pla d'entrada  $P_1$  on es situa l'escena  $s(\vec{x}_1)$  a una distància  $f$  de la lent  $L_1$ . En els inicis, la presentació de l'escena en el correlador es feia utilitzant pel·lícula fotogràfica. Modernament i, en particular, degut a la necessitat d'obtenir un reconeixement en temps real en escenes dinàmiques, es fan servir moduladors espacials de llum que reproduïxen l'escena captada per una càmera CCD. Darrera la lent  $L_1$  s'obté la TF de l'escena,  $S(\vec{\nu})$ , en el seu pla focal imatge,  $P_2$ . En aquest mateix pla es situa el que s'anomena filtre adaptat a l'objecte a reconèixer, amb la informació necessària del patró que es vol detectar. El primer desenvolupament d'un filtre adaptat a un senyal  $r(\vec{x}_1)$  correspon a una transparència o placa, la transmitància de la qual,  $t(\vec{\nu})$ , conté la funció  $R^*(\vec{\nu})$ , és a dir, la transformada de Fourier complexa conjugada del senyal. Just darrera del filtre adaptat hi ha el producte  $S(\vec{\nu}) \cdot t(\vec{\nu})$ . El pla  $P_2$  coincideix amb el pla focal objecte d' $L_2$  i,  $P_3$  és el pla focal imatge de la lent  $L_2$ . La relació entre la distribució d'amplitud darrera el filtre en  $P_2$  i la distribució d'amplitud en el pla  $P_3$ , és la d'una transformació de Fourier, amb la qual s'obté la correlació buscada de les funcions  $s(\vec{x}_1)$  i  $r(\vec{x}_1)$  però invertida, és a dir, s'obté  $s(-\vec{x}_3) \otimes r(-\vec{x}_3)$ . Aquesta inversió és totalment irrellevant i és freqüent compensar-la, si més no analíticament, mitjançant una redefinició dels eixos de coordenades ( $\vec{x} = -\vec{x}_3$ ) al pla  $P_3$  [127-129]. Aquesta distribució d'amplitud es troba en el pla de sortida del correlador,  $P_3$  en la Figura 2.1, també anomenat pla de correlació.

La proposta de Vander Lugt inclou una tècnica per a l'obtenció del filtre adaptat [1] mitjançant la qual es pot aconseguir enregistrar la informació d'amplitud i de fase de la funció  $R^*(\vec{\nu})$  realitzant una experiència d'interferències amb el muntatge hologràfic de la Figura 2.2. En ell, la placa hologràfica  $P$  s'exposa al patró d'interferències produït per la superposició de dues ones coherents: una correspon a  $R(\vec{\nu})$ , la TF de l'objecte buscat  $r(\vec{x}_1)$  i, l'altra correspon al feix de referència, una ona plana inclinada que representem per  $Ae^{i2\pi f \nu \sin\theta}$ .

Suposant que l'exposició rebuda per la pel·lícula estigui dins de la seva zona de resposta lineal podem considerar que la transmitància de la placa fotogràfica revelada,  $t(\vec{v})$ , serà proporcional mitjançant una constant  $\beta$ , a la distribució d'intensitat produïda per la superposició de dues ones coherents segons l'expressió

$$\begin{aligned} t(\vec{v}) &= \beta \left| A e^{i2\pi f v \sin\theta} + R(\vec{v}) \right|^2 = \\ &= \beta A^2 + \beta |R(\vec{v})|^2 + \beta A e^{-i2\pi f v \sin\theta} R(\vec{v}) + \beta A e^{i2\pi f v \sin\theta} R^*(\vec{v}). \end{aligned} \quad (2.4)$$

Observem que l'expressió (2.4) conté, en el seu darrer sumand, la funció  $R^*(\vec{v})$  que es vol incloure en el filtre.



**Figura 2.2.** Enregistrament del filtre clàssic de Vander Lugt.

Emprant aquest tipus de filtre adaptat (anomenat filtre adaptat clàssic), la distribució d'amplitud que s'obté en el pla de correlació  $P_3$  (Figura 2.1) ve donada per l'expressió

$$\begin{aligned}
TF\{S(\vec{v}) \cdot t(\vec{v})\} &= \beta A^2 s(-\bar{x}_3) + \beta r(-\bar{x}_3) * r^*(\bar{x}_3) * s(-\bar{x}_3) + \\
&+ \beta A \delta(x_3, y_3 + f \sin \theta) * r(-\bar{x}_3) * s(-\bar{x}_3) + \\
&+ \beta A \delta(x_3, y_3 - f \sin \theta) * r^*(-\bar{x}_3) \otimes s(-\bar{x}_3).
\end{aligned} \tag{2.5}$$

En l'expressió (2.5), el darrer terme correspon a la correlació creuada entre l'escena  $s$  invertida i l'objecte  $r$  també invertit. La correlació apareix desplaçada de l'origen de coordenades una distància  $-f \sin \theta$ , i és el terme que considerarem posteriorment en el reconeixement. Els altres termes que trobem en l'Eq. (2.5) consisteixen en dos primers sumands centrats a l'origen de coordenades i sense importància en aquesta aplicació. En el tercer terme, obtenim la convolució de les funcions  $s$  i  $r$ , desplaçada també de l'origen de coordenades una distància  $f \sin \theta$  simètrica respecte al quart terme de la correlació creuada.

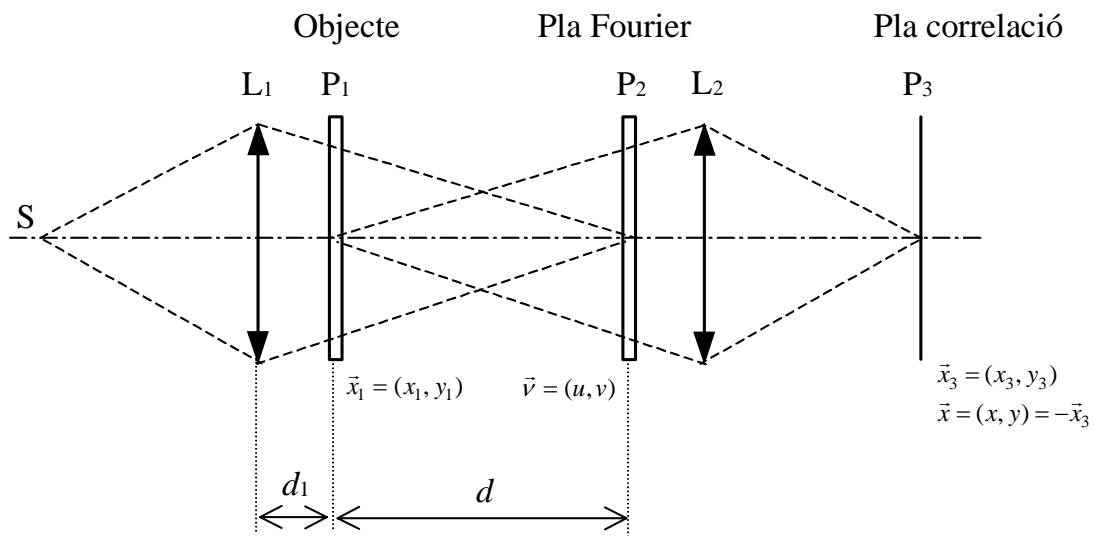
Acabem de descriure el correlador  $4f$ , que és el dispositiu òptic d'arquitectura més senzilla que ens permet realitzar l'operació de correlació eliminant els factors de fase quadràtics (Figura 2.1). A part d'aquest, existeixen altres arquitectures de sistemes òptics en els quals és possible obtenir la correlació entre dues funcions. Aquests correladors s'han introduït posteriorment en el reconeixement de formes, obtenint alguns avantatges respecte del correlador  $4f$ .

### 2.2.2 Correlador convergent

Una de les limitacions que presenta el correlador  $4f$  és la de ser molt crític respecte a la no coincidència en les escales de les transformades  $S(\vec{v})$  i  $R^*(\vec{v})$ , el producte de les quals es realitza en el pla de Fourier. Aquest inconvenient es pot esmenar utilitzant una arquitectura distinta, anomenada correlador convergent [130,125] (Figura 2.3), que és comunament acceptada com una variant del correlador VLC.

És ben coneguda la possibilitat de modificar l'escala de la TF d'un objecte situant-lo darrera la lent que s'utilitza per realitzar aquesta transformació [vegin-se, per

exemple, 131 i 132]. Per a obtenir una bona correlació cal fer coincidir exactament les escales de la TF òptica de l'escena amb la TF continguda en el filtre adaptat. Quan el filtre adaptat és un holograma generat per ordinador aquest procés d'ajust d'escales és sempre ineludible. Per aquest motiu haurem d'ajustar l'escala de la TF de l'escena, situant-la darrera la lent, variant la distància fins que coincideix amb les dimensions del filtre.



**Figura 2.3.** Correlador convergent.

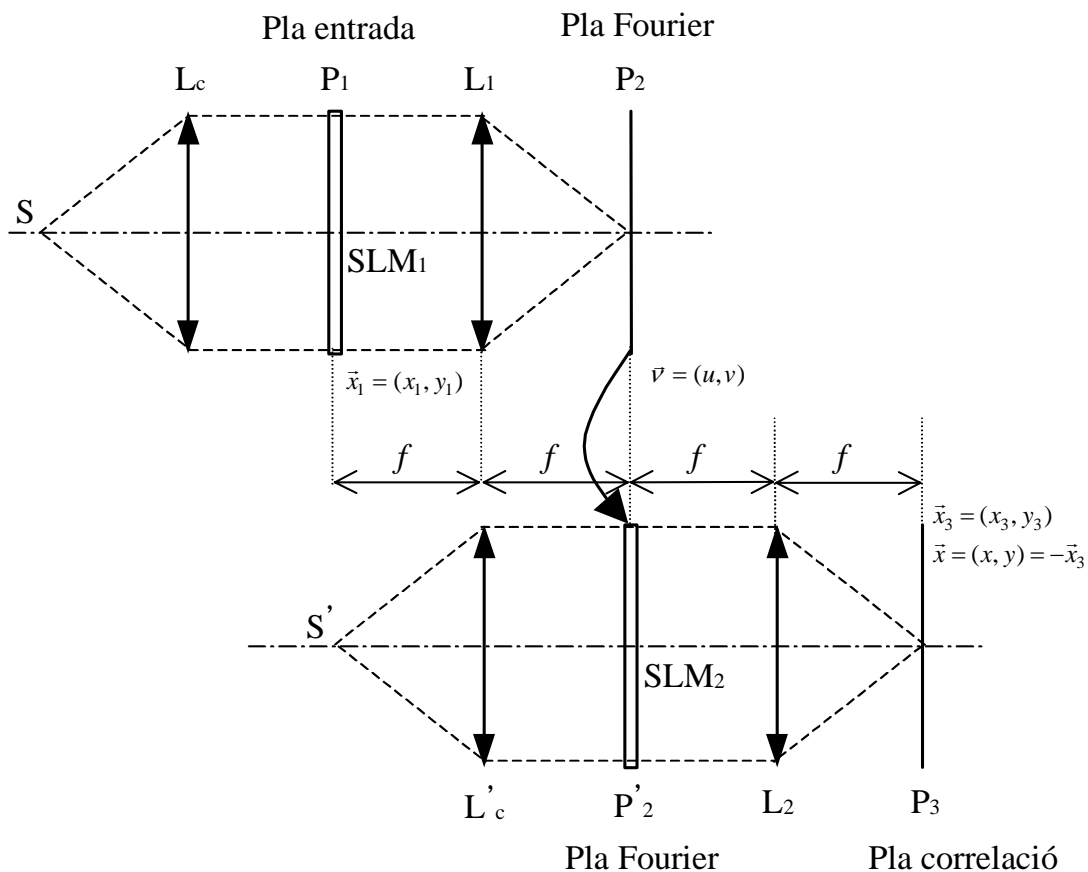
Considerem el muntatge de la Figura 2.3 corresponent al correlador convergent. A través de la lent  $L_1$ , la font puntual  $S$  té el seu conjugat imatge en el pla de freqüències  $P_2$ . En  $P_1$ , darrera la lent  $L_1$  a una distància  $d_1$ , s'ha interposat l'escena  $s(\vec{x}_1)$ , de manera que en  $P_2$  s'obté la seva TF de dimensions variables depenent de la distància  $d$  entre l'objecte i el pla de Fourier  $P_2$  [132]. El filtre  $R^*(\vec{v})$  està situat en el pla  $P_2$  i, just darrera del filtre s'obté el producte  $SR^*$ . Una segona lent,  $L_2$ , permet obtenir, en el pla conjugat de la transparència objecte,  $P_3$ , la transformada de Fourier del producte de les funcions  $SR^*$  i, per tant, la correlació  $s \otimes r$ . En el pla de correlació  $P_3$  s'obté una distribució d'intensitat similar a la que s'obté en el correlador  $4f$  [130], amb l'avantatge que es pot modificar la distància  $d$  per adaptar l'escala de la TF

d'  $s(\bar{x}_1)$  perquè coincideixi exactament amb l'escala de la TF d'  $r(\bar{x}_1)$  en el pla de freqüències  $P_2$ .

### 2.2.3 Correlador de transformades conjuntes

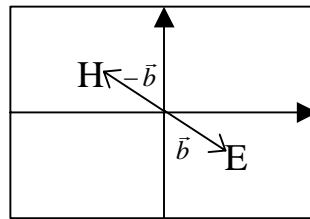
Una altra arquitectura de correlador és l'anomenat correlador de transformades conjuntes (JTC, de l'anglès *Joint transform correlator*), que té com a diferència més destacada en relació al correlador de Vander Lugt i derivats, la no necessitat d'una transparència amb el filtre adaptat a l'objecte a reconèixer. Aquest correlador va ser proposat simultàniament per dos grups de manera independent l'any 1996 [2,3].

Un esquema de l'arquitectura d'aquest correlador es mostra en la Figura 2.4.



**Figura 2.4.** Correlador de transformades conjuntes.

En la Figura 2.4 observem la font de llum S, puntual i monocromàtica de longitud d'ona  $\lambda$ , que es situa en el focus objecte de la lent  $L_c$  per a proporcionar el feix col·limat que il·lumina el pla d'entrada del JTC,  $P_1$ . En aquest pla, l'escena a analitzar  $s(\bar{x}_1)$  i l'objecte de referència  $r(\bar{x}_1)$  es situen de manera contigua en una transparència o modulador espacial de llum de transmitància  $s(\bar{x}_1 + \bar{b}) + r(\bar{x}_1 - \bar{b})$ , donat que ambdues funcions es desplacen una distància  $\bar{b}$  respecte l'eix òptic en direccions oposades (Figura 2.5).



**Figura 2.5.** Pla d'entrada del JTC on s'observa el desplaçament relatiu entre l'escena (esquerra) i la referència (dreta).

El pla d'entrada del JTC es troba situat per davant de la lent  $L_1$  a una distància  $f$ , de manera que en pla focal imatge de la lent,  $P_2$ , s'obté la seva transformada de Fourier que s'enregistra, en general, mitjançant un detector d'intensitat. En els inicis del JTC, aquest detector consistia en una placa fotogràfica o hologràfica i, més recentment, s'han introduït càmeres CCD o modulador espacials de llum de lectura i escriptura òptica. En qualsevol cas, tots aquests detectors enregistren la informació proporcional a la distribució d'intensitat que, en  $P_2$ , ve donada per l'expressió

$$\begin{aligned}
 I(\bar{\nu}) &= \beta \left| TF \left\{ s(\bar{x}_1 + \bar{b}) + r(\bar{x}_1 - \bar{b}) \right\} \right|^2 = \\
 &= \beta \left[ |S(\bar{\nu})|^2 + |R(\bar{\nu})|^2 + S(\bar{\nu}) \cdot R^*(\bar{\nu}) \cdot e^{i4\pi\bar{\nu} \cdot \bar{b}} + S^*(\bar{\nu}) \cdot R(\bar{\nu}) \cdot e^{-i4\pi\bar{\nu} \cdot \bar{b}} \right],
 \end{aligned} \tag{2.6}$$

on  $\beta$  denota la constant de proporcionalitat dependent del tipus de detector emprat.



La distribució de l'Eq. (2.6) rep el nom d'espectre conjunt d'intensitat (JPS, de l'anglès *Joint Power Spectrum*), i té un paper fonamental en el procés de correlació amb l'arquitectura del correlador de transformades conjuntes.

A continuació, el JPS es reproduïx en el modulador SLM<sub>2</sub> (o bé, en la placa fotogràfica on s'havia enregistrat prèviament) situat en el pla P'<sub>2</sub>, de manera que la lent L<sub>2</sub> proporciona en el pla P<sub>3</sub> la transformada de Fourier de l'expressió (2.6),

$$\begin{aligned} TF\{I(\bar{v})\} = & \beta [s(-\bar{x}_3) \otimes s(-\bar{x}_3)] + \beta [r(-\bar{x}_3) \otimes r(-\bar{x}_3)] + \\ & + \beta [s(-\bar{x}_3) \otimes r(-\bar{x}_3) * \delta(\bar{x}_3 + 2\bar{b})] + \beta [r(-\bar{x}_3) \otimes s(-\bar{x}_3) * \delta(\bar{x}_3 - 2\bar{b})]. \end{aligned} \quad (2.7)$$

Els dos primers sumands d'aquesta expressió corresponen a les autocorrelacions de l'escena a analitzar i de l'objecte a reconèixer, respectivament, i es troben centrats a l'eix òptic. El tercer sumand proporciona la correlació creuada entre l'escena i la referència desplaçada una distància  $2\bar{b}$  respecte l'eix. Finalment, el darrer sumand, simètric al tercer, proporciona la correlació entre les mateixes funcions, però intercanviant els seus papers d'escena i de referència.

L'arquitectura del JTC pot ser simplificada si tant sols s'utilitza la meitat del dispositiu que acabem de descriure, és a dir, amb un únic modulador espacial de llum [133,134]. El JPS s'enregistra amb una càmera CCD situada en el pla P<sub>2</sub> (Figura 2.4), en un primer pas del procés de correlació i, tot seguit s'envia aquesta distribució novament al primer modulador, SLM<sub>1</sub>, en el pla de Fourier que anomenarem P'<sub>2</sub> (= P<sub>1</sub>), i obtenim la TF del pla P<sub>3</sub> (= P<sub>2</sub>). La distribució aquí obtinguda correspon al pla de sortida del JTC, que conté els pics de correlació de l'escena i de la referència fora d'eix i, és enregistrat per la càmera CCD.

El fet de reduir el nombre de components òptics i, sobretot optoelectrònics (càmeres i moduladors), permet immediatament abaratir el cost del dispositiu. En contrapartida, suposa doblar el temps de processament del JTC, donat que el procés complert s'ha de realitzar en dos passos consecutius [135].

En [136] s'ha proposat un JTC d'escala variable combinant l'arquitectura del JTC amb un únic SLM i la configuració de correlació d'escala variable [137,125]. L'avantatge d'aquesta configuració radica en la possibilitat de modificar l'escala del JPS, que s'enregistra amb la càmera CCD. Donat que aquesta distribució consisteix en un patró de franges d'interferència, és necessari que l'amplada d'aquestes franges ocupi un nombre suficient de píxels en la càmera CCD per a poder ésser enregistrada mantenint la informació més rellevant.

De la mateixa manera que la resolució espacial de la CCD pot afectar el procés de reconeixement en un JTC, hi ha altres fonts de degradació del pic de correlació, relacionades directament amb el sistema òptic i les característiques físiques i de posicionament dels components que el constitueixen, que també han estat motiu d'estudi.

Per exemple, una bona resolució espacial del modulador pot influir positivament en els resultats del JTC donat que la reproducció de l'escena, de la referència i, posteriorment, del JPS, es farà de manera més fiable. Si no disposem d'una resolució suficient, aquest fet provocarà una degradació dels resultats de correlació [138].

L'alineació de l'SLM en el sistema òptic també ha estat objecte d'estudi en [138]. Concretament s'han estudiat desalineacions en sentit transversals (perpendiculars a l'eix òptic) i longitudinals (en la direcció de l'eix òptic). S'ha demostrat que l'efecte dels primers sobre el pic de correlació és nul, donat que el JPS no requereix un alineament estricte en el muntatge, però en canvi, el pic es degrada amb els desplaçament longitudinals. En aquest treball es quantifica el seu efecte i es compara amb el correlador clàssic de Vander Lugt.

#### **2.2.4 Comparació del correlador de Vander Lugt (VLC) i del correlador de transformades conjuntes (JTC)**

Dos anys més tard del desenvolupament del VLC per Vander Lugt (1964), va ser demostrat per primera vegada el correlador de transformades conjuntes (JTC), independentment per Weaver i Goodman [2] i per Rau [3]. No obstant, aquesta

arquitectura va romandre estancada fins els anys 80 degut a la manca de dispositius adequats que facilitessin l'emmagatzemat i la transmissió en temps curts de la informació processada. Així, per exemple, en el JTC era necessari enregistrar l'espectre de transformades conjuntes (JPS) en plaques fotogràfiques o hologràfiques per a cada escena a analitzar, tot i emprar el mateix objecte de referència. En aquest sentit, el VLC presentava una reducció important en el temps de processament donat que només s'enregistrava una placa hologràfica per a cada objecte patró que es volgués detectar. Aquesta placa s'emprava com a filtre adaptat per les diferents escenes a analitzar. El què era un avantatge pel VLC, va canviar amb la introducció de dispositius com les càmeres CCD i els SLM, que possibilitaven la realització en temps real de tot el procés de reconeixement amb l'arquitectura del JTC.

Va ser en el 1984 quan Yu i Lu introduïren per primera vegada l'SLM en un correlador per aconseguir un JTC programable [139]. Posteriorment, Yu *et al.*, [140] van afegir al JTC la utilització d'una càmera CCD per enregistrar el JPS i el pla de correlació final. Des d'aquest moment, el JTC ha adquirit un paper important en el reconeixement de patrons [141,142].

Lu *et al.* [141] van proposar una generalització del sistema òptic de manera que amb un sol dispositiu es pogués descriure conjuntament el correlador de Vander Lugt (VLC) i el de transformades conjuntes (JTC). La formulació del correlador generalitzat permet als autors obtenir expressions analítiques que descriuen els efectes sobre la qualitat del pic de correlació, del soroll introduït per l'SLM, tant en amplitud de les distribucions degut a les no-uniformitats de la transmitància del dispositiu, com en la fase a causa de les variacions de la superfície de l'SLM.

La majoria dels correladors elaborats per desenvolupar la recerca en el camp del reconeixement per correlació òptica, es basen en gran mesura en l'arquitectura del VLC i la del JTC, i llurs variacions. Donada la importància del VLC i del JTC, hi ha hagut molts treballs de comparació d'aquests correladors per establir-ne les seves limitacions i apreciar-ne els avantatges d'una arquitectura en front de l'altra [126,135,138,141]. En aquest apartat, volem presentar de manera resumida, algunes de les conclusions més importants que s'extreuen d'aquests treballs.

En [141] s'estudia analíticament la degradació del pic de correlació degut a variacions d'escala i a rotacions de l'escena respecte de la referència. En quant a la tolerància al soroll i a canvis d'escala i de rotacions, ambdues arquitectures presenten un comportament similar en l'anàlisi d'escenes monobjectes. Respecte al pla d'entrada del correlador, el JTC necessita el doble d'amplada de banda espacial en comparació al VLC, donat que el mateix pla d'entrada és compartit per l'escena i la referència i, en conseqüència, el JTC té un requeriment més estricte de coherència espacial. De l'avaluació qualitativa d'altres aspectes, s'extreu una millor realització de la correlació de funcions mitjançant l'arquitectura del JTC. Aquesta conclusió es deriva, principalment, de la realització en temps real i de la relaxació en l'exigència d'alineació dels seus components. Comentem una mica més les característiques del VLC i del JTC en relació a ambdós aspectes.

En relació al processament en temps real, actualment no existeix molta diferència entre ambdues arquitectures donat que en el VLC els filtres adaptats a l'objecte a reconèixer poden ser generats per ordinador i, tot i incrementar el temps de computació [143], és possible la implementació en temps real del procés de correlació. A més, en [142], s'introdueix la possibilitat d'aplicar operacions no-lineals al JPS abans d'ésser reproduït en el segon pas del JTC d'un sol braç. Amb aquesta operació els temps de càlcul d'una i altra arquitectura s'equiparen, però s'augmenta el nombre d'aplicacions de reconeixement que poden portar-se a terme pel sistema JTC. En el treball de Yu i Gregory [135], s'afegeix l'ús del JTC en sistemes de seguiment (*tracking*). Amb la introducció de nous materials com els discs òptics o materials fotorefractius amb una capacitat d'emmagatzemat massiva, una alta resolució i un enregistrament en temps real, es pot disposar d'un pla d'entrada dinàmic amb actualització constant de la referència.

En un treball recent de Purwosumarto i Yu [138], trobem una avaluació numèrica en quant als requeriments d'alineació dels components en ambdues arquitectures. S'estudia la precisió necessària en el posicionament de l'SLM en el pla de Fourier d'ambdós correladors. S'estableix una millor realització del JTC davant del VLC donat que els petits desplaçaments transversals no afecten al JPS, mentre que si es

necessita un posicionament molt precís en el filtre adaptat del VLC per obtenir la màxima coincidència de les transformades de Fourier de les funcions que es correlacionen. En relació a l'alineament longitudinal, el JTC és més estricte que el VLC, però amb tot i això, els requeriments són de l'ordre d'1 mil·límetre, i aquesta distància és factible a la pràctica amb els mecanismes actuals de posicionament.

El correlador clàssic de Vander Lugt, però, permet obtenir millors resultats en la detecció de múltiples objectes tant quan l'escena és lliure de soroll, com quan aquesta conté un cert nivell de soroll additiu [126]. Les conclusions extreïdes en la referència [126], posen de manifest que la intensitat del pic de correlació només és independent del nombre d'objectes i del nivell de soroll introduït pel VLC, mentre que ambdós factors fan disminuir la intensitat del pic a mesura que el nombre d'objectes o el nivell de soroll s'incrementen en el JTC.

En els capítols posteriors d'aquest treball de tesi es presentaran resultats experimentals de reconeixement amb ambdues arquitectures. En els casos d'anàlisi d'escenes multiobjecte en color s'ha utilitzat l'arquitectura de VLC, mentre que per a la realització de la correlació no-lineal dual és necessària l'arquitectura del JTC.

### 2.3 Filtratge adaptat

Tenint en compte la nomenclatura de [141], podem parlar de manera general de les operacions de filtratge en correlació, entenent que parlem dels diferents tipus de filtres adaptats a l'objecte a reconèixer en el cas del VLC, i del processat del JPS realitzat en un pas intermedi del JTC. En el reconeixement d'objectes és molt important la informació introduïda en el filtre de correlació per resoldre de manera satisfactòria un problema plantejat. S'han proposat diverses possibilitats per introduir la informació total, parcial o modificada, tant de la funció complexa  $R^*(\vec{v})$  com del JPS, en el filtratge de correlació. Aquest camp es troba en constant evolució donat que en funció de la informació codificada en el filtre o del tipus de modificació del JPS, és possible detectar un objecte per una característica determinada, discriminar-

lo millor d'altres, reconèixer-lo fins i tot en presència d'algun tipus de distorsió o degradació del senyal, etc.

En aquest apartat revisarem alguns dels filtres de correlació, àmpliament utilitzats en el reconeixement de patrons, fent especial esment en aquells que tenen una relació més estreta amb les propostes elaborades en aquesta tesi. Indicarem l'arquitectura per la qual s'han dissenyat els filtres, i la possibilitat d'aplicació a diferents correladors, especialment el VLC i el JTC.

### 2.3.1 Filtres de potències fraccionàries

Els filtres de potències fraccionàries (FPF) van ser introduïts per Kumar i Hassebrook [6], l'any 1990, com a generalització d'alguns filtres adaptats del VLC existents en la literatura. Un filtre adaptat pertanyent als FPF es defineix per l'expressió

$$H_{FPF}(\vec{\nu}) = \begin{cases} |R(\vec{\nu})|^p e^{-i\theta(\vec{\nu})} & \text{si } |R(\vec{\nu})| \neq 0 \\ 0 & \text{si } |R(\vec{\nu})| = 0, \end{cases} \quad (2.8)$$

on  $R(\vec{\nu}) = |R(\vec{\nu})| e^{i\theta(\vec{\nu})}$  és la TF de l'objecte a reconèixer  $r(\vec{x})$  i  $p$  és un nombre real.

Són casos particulars dels FPF, el filtre adaptat clàssic, el filtre només de fase i el filtre invers, que detallem a continuació.

El filtre adaptat clàssic (CMF, de l'anglès *Classical Matched Filter*), ja descrit en l'apartat 2.2.1, pren el valor de  $p = 1$  en la formulació dels filtres FPF (Eq. 2.8). En ell s'enregistra la funció TF complexa-conjugada de l'objecte a reconèixer, és a dir,

$$H_{CMF}(\vec{\nu}) = R^*(\vec{\nu}). \quad (2.9)$$

Aquest filtre presenta l'inconvenient de tenir poca discriminació entre objectes amb detalls i contorns semblants, dóna pics de correlació molt amples i és relativament sensible a deformacions dels objectes, ja siguin canvis de forma, d'escala o rotacions [144-146]. Per contra és molt resistent al soroll [147].

El filtre només de fase (POF, de l'anglès *Phase-only filter*), està inclòs en la formulació dels FPF considerant  $p = 0$  (Eq. 2.8). Aquest filtre millora la capacitat discriminativa del CMF i és més eficient energèticament. Va ser introduït per Horner i Gianino [4]. La transmitància d'aquest filtre ve descrita per l'expressió

$$H_{POF}(\vec{v}) = \frac{R^*(\vec{v})}{|R(\vec{v})|} = e^{-i\theta(\vec{v})}. \quad (2.10)$$

D'aquesta manera es té informació de les freqüències més altes afavorint la discriminació entre objectes semblants de distribució d'amplitud similar, i donant uns pics de correlació més estrets i amb uns lòbuls laterals molt reduïts [148]. La sensibilitat a les rotacions i als canvis d'escala és molt més gran en el filtre POF que en el CMF, però al mateix temps és menys resistent al soroll.

En [7] i [8] es proposen algunes millores dels filtres POF que augmenten la discriminació en la correlació. En [7] s'introdueixen els filtres POF binaritzats i en [8] es formulen filtres trivaluats (-1, 0, 1) basats en el quocient de les energies espectrals de l'objecte buscat amb un altre no buscat. Els filtres ternaris poden contemplar-se com una extensió dels filtres només de fase binaris, i milloren la capacitat discriminativa, en particular quan el senyal d'entrada està afectat per una gran quantitat de soroll. Els filtres dissenyats en [7] i [8] incideixen únicament en la informació d'amplitud de la funció  $R^*(\vec{v})$  introduïda en el filtre adaptat. No obstant, la distribució de fase juga un paper molt important en la capacitat discriminativa del sistema de reconeixement. Zhou *et al.* [149] han proposat una tècnica basada en el preblanquejat de les diferències de fase per a la síntesi d'un filtre compost (multiobjecte) que millora la discriminació d'objectes de classes diferents.

Un altre mètode que optimitza la capacitat discriminativa del POF, recolzant-se en la informació de fase, ha estat proposat per Ahouzi *et al.* [150]. En aquest treball, s'introdueix la noció de l'histograma de les diferències de fase per al disseny del filtre adaptat mitjançant el bloqueig d'unes freqüències determinades.

El filtre invers (IF, de l'anglès *Invers Filter*) és un cas particular dels FPF quan prenem  $p = -1$ . Aquest filtre proporciona pics de correlació encara més estrets que els filtres CMF i POF, i es defineix com [5]

$$H_{IF}(\vec{\nu}) = \frac{R^*(\vec{\nu})}{|R(\vec{\nu})|^2} = \frac{e^{-i\theta(\vec{\nu})}}{|R(\vec{\nu})|}. \quad (2.11)$$

En el cas en què l'objecte a reconèixer sigui present a l'escena, el producte  $S(\vec{\nu}) \cdot H_{IF}(\vec{\nu})$  es converteix en un valor constant, la TF del qual inclourà una funció autocorrelació consistent en una funció delta de Dirac. Aquest filtre potencia, encara més que el POF, la informació continguda en les freqüències altes per augmentar la capacitat discriminativa del sistema. Proporciona pics més estrets i amb lòbuls laterals més reduïts, i ha estat utilitzat per al reconeixement d'imatges desenfocades [151]. Awwal *et al.* [152] proposen una nova definició del filtre IF modificada per uns paràmetres. Controlant aquests paràmetres s'eliminen els punts d'indeterminació del filtre invers, i poden ser utilitzats, a més, per eliminar el soroll i per limitar la banda passant del filtre.

Els filtres CMF, POF, IF i, en general els FPF, van ser inicialment proposats com a filtres adaptats pel seu ús en un correlador de Vander Lugt, però també existeixen les corresponents versions pel JTC [11,12].

### 2.3.2 Filtratge no-lineal

Un cas ben conegut de mètode de filtratge no-lineal és la correlació només de fase [30-33]. En la correlació només de fase (PPC, de l'anglès *Pure Phase Correlation*)



només es té en compte la distribució de fase tant de l'escena com de la referència. La funció del filtre és

$$H_{PPC} = \frac{1}{|S(\vec{v})|} \cdot \frac{R^*(\vec{v})}{|R(\vec{v})|}. \quad (2.12)$$

En el cas que  $s(\vec{x}) = r(\vec{x})$ , s'obté  $S(\vec{v}) \cdot H_{PPC} = constant$ . En aquest sentit, el filtratge del PPC té una certa similitud amb el filtre invers (2.11), però només en el cas en què escena i referència són la mateixa funció.

La correlació només de fase dóna més pes a les altes freqüències i, en conseqüència, és un mètode molt sensible a distorsions introduïdes en l'objecte a detectar. Al mateix temps, s'aconsegueix un increment en la capacitat discriminativa del sistema [29,38]. S'han proposat diferents realitzacions optoelectròniques de la PPC, algunes basades en mètodes interferomètrics [30-31,153-155], i altres en l'arquitectura del JTC [156,17,157,27].

Una altra operació no-lineal important en el camp del reconeixement òptic de patrons, és l'anomenat correlador no-lineal. En [158], Javidi proposa per primera vegada el correlador de transformades conjuntes no-lineal (NJTC, de l'anglès *Nonlinear Joint Transform Correlator*). En ell, s'aplica una operació llindar en el JPS depenent d'un paràmetre  $k$  ( $0 \leq k \leq 1$ ) que indica la severitat de la no-linealitat. Prenent  $k = 1$ , correspon al cas lineal en què no es modifica la informació del JPS i, en l'altre extrem,  $k = 0$ , s'aconsegueix la binarització de la distribució. Casos intermedis corresponen, de manera general, a la no-linealitat d'ordre  $k$ -èssim, que es pot entendre com l'aplicació d'un filtre de banda passant, el pendent de la zona de llindar del qual varia amb l'ordre  $k$  de la no-linealitat. Variant la severitat de la no-linealitat mitjançant el paràmetre  $k$  es poden obtenir diferents tipus de filtres no-lineals i diferents senyals de correlació. Cal remarcar, que donat que la no-linealitat s'aplica a la distribució JPS en el pla de Fourier, la no-linealitat afecta per igual a la informació corresponent a l'escena a analitzada i a l'objecte de referència. En aquest

sentit, per tant, podem parlar d'un correlador no-lineal simètric. A partir dels resultats de simulació presentats en [158] s'observa que mitjançant aquest filtratge no-lineal s'aconsegueixen lòbuls laterals més reduïts i pics de correlació més estrets, produint, en conseqüència, pics de correlació més ben definits i localitzats. Resultats experimentals es recullen en [159].

En [160,161], Javidi generalitza la proposta del JTC no-lineal al cas dels filtres adaptats del VLC, i en [162] Javidi i Zhang en presenten resultats experimentals. La comparació del JTC no-lineal i del VLC amb filtres adaptats no-lineals, en relació ambdós al JTC lineal, es realitza en [163]. S'observa que, en el cas d'escenes amb soroll, ambdós correladors no-lineals obtenen millors resultats de reconeixement que el correlador JTC convencional lineal. Donat que la no-linealitat aplicada en el JTC afecta per igual a les amplituds de les TF de l'escena i de l'objecte de referència, s'obté una millor realització de la correlació pel JTC no-lineal comparat amb el correlador basat en el mètode de filtratge adaptat no-lineal. Les diferents propostes de Javidi i col·laboradors en quant al processament no-lineal del JTC i dels correladors del tipus VLC amb utilització de filtres adaptats a l'objecte a reconèixer, es troben recollides en [164,165].

Javidi *et al.* recullen en [166] diferents aportacions de tècniques no-lineals aplicades en el reconeixement òptic de patrons, ja sigui per l'ús de dispositius no-lineals, o bé mitjançant l'aplicació d'algorismes.

### 2.3.3 Avaluació del comportament dels filtres

Per avaluar el comportament dels filtres, s'han introduït un seguit de paràmetres [6]. A tall d'exemple, esmentem alguns d'aquests paràmetres que seran mencionats sovint en aquesta memòria. En general, per tots ells, es suposa que l'escena a analitzar conté un sol objecte.

La relació entre el pic i l'energia de correlació (PCE, de l'anglès *Peak-to-correlation energy*), està definida com

$$PCE = \frac{|c(\vec{0})|^2}{\iint |c(\vec{x})|^2 d\vec{x}}, \quad (2.13)$$

essent  $|c(\vec{0})|^2$  el valor a l'origen de la intensitat de la funció correlació, i el denominador l'energia total del pla de correlació. Aquest paràmetre està relacionat amb la detectabilitat de l'objecte. Indica si els pics de correlació són aguts i sobresurten del fons per tal de poder detectar un objecte i conèixer la seva posició sense dificultat. Una informació similar s'obté si es considera la relació entre el pic de correlació i el valor mig de tota la resta en el pla de correlació.

Un paràmetre important és l'eficiència, introduït per Horner [167]. Avaluat la quantitat de llum que arriba al pla de correlació en relació a la llum del pla d'entrada

$$\eta = \frac{\iint |c(\vec{x})|^2 d\vec{x}}{\iint |f(\vec{x})|^2 d\vec{x}}. \quad (2.14)$$

Aquest paràmetre prendrà valors entre 0 i 1 depenent de la informació introduïda en el filtre, de la codificació de la informació i del dispositiu utilitzat per la seva producció.

La relació senyal-soroll (SRN, de l'anglès *Signal-to-noise ratio*) és útil per avaluar la tolerància al soroll del sistema. Es poden introduir successives entrades afectades amb diferents nivells de soroll, i fer la mesura del quocient

$$SNR = \frac{|E\{c(\vec{0})\}|^2}{\text{var}\{c(\vec{0})\}}, \quad (2.15)$$

on el numerador és el quadrat del valor mig dels valors centrals del pla de correlació i, el denominador és la variància d'aquests valors. La localització del pic i, per tant,

la de l'objecte a l'escena, pot dependre del soroll de l'escena i sobre tot de l'agudesia del pic de correlació.

Un altre paràmetre és la capacitat de discriminació (DC, de l'anglès *Discrimination Capability*) que es defineix a partir de la relació entre el valor màxim del pic de correlació creuada ( $cc$ ) entre dos objectes diferents i el valor màxim del pic d'autocorrelació ( $ac$ ) d'un objecte amb si mateix [150],

$$DC = 1 - \frac{cc}{ac}. \quad (2.16)$$

Un valor alt d'aquesta mesura, proper a la unitat, indicarà una bona capacitat de discriminació amb pics d'autocorrelació molt més alts que els de correlació creuada. L'Eq. (2.16) correspon a la definició inicial de la  $DC$  però, posteriorment, aquesta s'ha modificat en el cas del reconeixement d'objectes en color [168]. Aquest punt serà ampliat en l'apartat 2.4.3, d'aquest mateix capítol.

## 2.4 Correlació multicanal en el reconeixement d'objectes en color

### 2.4.1 Canals $\lambda_n$

En els mètodes habituals de reconeixement d'objectes en color per correlació òptica, es porta a terme un primer pas de descomposició multicanal de la informació cromàtica. La descomposició clàssica és l'anomenada RGB, on es té la informació de l'escena en els canals vermell (R), verd (G) i blau (B). Però de manera més general, podem considerar que una descomposició cromàtica qualsevol es realitza en  $n$  canals, de banda ampla o estreta, caracteritzats cada un d'ells per una longitud d'ona principal  $\lambda_n$ . Amb aquesta descomposició de la informació cromàtica en diferents canals, és possible estendre fàcilment el mètode de reconeixement de formes per correlació monocromàtica (Secció 2.2) a escenes policromàtiques. La correlació es

realitza en cada un dels  $n$  canals per separat, de manera que l'anàlisi conjunta dels resultats de tots els canals permet tenir en compte la informació del color de l'escena.

Suposem que una escena policromàtica il·luminada per la longitud d'ona  $\lambda$  es descriu per la funció  $f_\lambda(\vec{x})$ . Aquesta escena pot estar formada per  $M$  objectes diferents,  $o_i$ , tal com indica l'expressió

$$f_\lambda(\vec{x}) = \sum_{i=1}^M o_{i\lambda}(\vec{x} - \vec{x}_i), \quad (2.17)$$

on el subíndex  $\lambda$  posa de manifest que la transmitància de l'escena i dels seus objectes depèn de la longitud d'ona. D'aquesta manera, per cada escena i objecte haurem de considerar  $n$  funcions diferents per les longituds d'ona  $\lambda_n$ . Per aquesta anàlisi considerem l'escena lliure de soroll i sense solapaments entre els diferents objectes, és a dir,

$$D\{o_{i\lambda}(\vec{x})\} \cap D\{o_{j\lambda}(\vec{x})\} = 0 \text{ per } i \neq j, \quad i \quad \forall \lambda \in \{\lambda_n\}. \quad (2.18)$$

Ens proposem reconèixer el primer objecte de l'escena,  $o_{1\lambda}(\vec{x})$ , per la qual cosa haurem d'introduir la informació d'aquesta funció en el sistema de reconeixement.

Segons l'arquitectura de correlador escollida, hem vist que la informació de l'objecte buscat s'introdueix de manera diferent. Per exemple, en un JTC, la funció  $o_{1\lambda}(\vec{x})$  hauria de ser introduïda en el pla d'entrada del JTC, distant de l'escena a analitzar una certa distància. En canvi, en els correladors  $4f$  o derivats, cal obtenir prèviament el conjugat de la TF de la referència que pot ser modificat de diferents maneres segons el tipus de filtrat. Els correladors proporcionen en el seu pla de sortida la correlació entre una funció  $f_\lambda(\vec{x})$ , que correspon a l'escena a analitzar, i  $o_{1\lambda}(\vec{x})$ , funció que representa l'objecte a reconèixer. Tot i que els passos intermedis difereixen segons l'arquitectura del sistema òptic, el resultat final de correlació el

podem entendre com a general per qualsevol correlador. Per tant, sense perdre generalitat en l'anàlisi, podem considerar que emprarem un correlador  $4f$  o derivat d'ell.

Suposem doncs, que hem d'adaptar un filtre en cada un dels canals. Descriurem el filtre per  $\hat{O}_{1\lambda}(\vec{v})$ . En el canal  $\lambda$ , la presència de l'objecte 1 podrà ser detectada mitjançant la correlació entre l'escena  $f_\lambda(\vec{x})$  i l'objecte  $o_{1\lambda}(\vec{x})$ ,

$$\begin{aligned} c_\lambda(\vec{x}) &= f_\lambda(\vec{x}) \otimes o_{1\lambda}(\vec{x}) = \\ &= o_{1\lambda}(\vec{x}) \otimes o_{1\lambda}(\vec{x}) * \delta(\vec{x} - \vec{x}_1) + \sum_{i=2}^M o_{i\lambda}(\vec{x}) \otimes o_{1\lambda}(\vec{x}) * \delta(\vec{x} - \vec{x}_i). \end{aligned} \quad (2.19)$$

El primer terme correspon a l'autocorrelació de l'objecte 1 en el canal  $\lambda$ . El pic d'intensitat està situat en la mateixa posició  $\vec{x}_1$  que ocupa l'objecte a l'escena. Els altres sumands són deguts a les correlacions creuades del senyal  $o_{1\lambda}(\vec{x})$  amb la resta d'objectes de l'escena.

Quan l'objecte està present a l'escena en la posició  $\vec{x}_1$ , repetint l'operació de correlació per cada canal  $\lambda_n$ , s'obtenen  $n$  màxims d'autocorrelació, un en cada canal, centrats tots ells en el punt definit per  $\vec{x}_1$ . Si un mateix objecte es reconeix independentment en tots els canals es dedueix que l'objecte coincideix amb el buscat. És important que en cada canal es correlacionin les components de l'objecte i de l'escena de l'esmentat canal. Imaginem que, per error o altra circumstància, realitzem la correlació de la component  $A$  de l'escena,  $f_A(\vec{x})$ , amb la component  $B$  de l'objecte 1,  $o_{1B}(\vec{x})$ , al qual hem adaptat el filtre. Encara que l'objecte estigui inclòs en l'escena no obtindrem l'autocorrelació del senyal, ja que els termes que apareixeran en el pla de correlació seran:

$$\begin{aligned}
 f_A(\vec{x}) \otimes o_{1B}(\vec{x}) &= o_{1A}(\vec{x}) \otimes o_{1B}(\vec{x}) * \delta(\vec{x} - \vec{x}_1) + \\
 &+ \sum_{i=2}^M o_{iA}(\vec{x}) \otimes o_{1B}(\vec{x}) * \delta(\vec{x} - \vec{x}_i).
 \end{aligned}
 \tag{2.20}$$

El primer sumand, per un objecte policromàtic, correspon a la correlació entre dues funcions espacialment diferents i, per tant, no coincideix amb l'autocorrelació. Molt probablement aquest senyal de correlació creuada no assolirà un pic d'intensitat suficient per permetre el reconeixement.

#### 2.4.2 Descomposició i correlació multicanal RGB

La descomposició multicanal RGB es pot obtenir per diferents mètodes. Un d'ells pren com a punt de partida una diapositiva en color amb l'escena a analitzar. L'objecte que es vol reconèixer,  $o_1(\vec{x})$ , es pot extreure de l'escena o bé ser enregistrat per separat també en diapositiva en color. Il·luminant successivament l'objecte amb feixos de llum làser de longituds d'ona centrades en les regions espectrals R, G i B [per exemple,  $\lambda(R) = 633$  nm del làser d'He-Ne,  $\lambda(G) = 514$  nm i  $\lambda(B) = 457$  nm d'un làser d'Ar], s'obtenen les components,  $o_{1\lambda}(\vec{x})$ , de l'objecte que es pretén reconèixer de l'escena. Mitjançant, per exemple, mètodes d'holografia clàssica, es pot registrar la informació de l'objecte en tres filtres adaptats diferents, un per cada canal,  $\hat{O}_{1\lambda}(\vec{v})$  ( $\lambda = R, G$  i  $B$ ) [90,101]. L'escena i un dels filtres es situen en els plans corresponents d'un correlador òptic i s'il·lumina el sistema amb la longitud d'ona utilitzada en la generació del filtre. A continuació es repeteix el procés de correlació amb successius feixos coherents de diferent longitud d'ona i els corresponents filtres [101].

Òbviament, si il·luminéssim el correlador amb una longitud d'ona diferent a la utilitzada per registrar el filtre adaptat, no obtindríem l'autocorrelació del senyal a detectar, ja que estaríem en el cas de l'Eq. (2.20).

Un altre mètode que permet una descomposició RGB diferent és partint de la informació enregistrada per una càmera en color. Alguns estudis realitzats en càmeres de sensors de dos tipus, 1 CCD i 3 CCD, posen de manifest la conveniència d'utilitzar càmeres en color de 3 CCD, en reconeixement de formes [106]. Al captar una escena, s'obtenen tres matrius de punts (o píxels) per a cada imatge que corresponen a les seves components RGB ( $f_R$ ,  $f_G$  i  $f_B$ ). La descomposició s'obté mitjançant filtres dicroics situats a l'interior de la càmera. A més, davant de cada detector CCD hi ha col·locat un filtre cromàtic de banda passant centrat en la zona dels vermells, dels verds i dels blaus, respectivament. Un cop digitalitzades les matrius es pot extreure la informació de l'objecte a reconèixer de cada una d'elles, de manera que a les components de l'objecte  $o_{1R}(\vec{x})$ ,  $o_{1G}(\vec{x})$  i  $o_{1B}(\vec{x})$  se'ls adaptarà separatament els corresponents filtres,  $\hat{O}_{1R}(\vec{v})$ ,  $\hat{O}_{1G}(\vec{v})$  i  $\hat{O}_{1B}(\vec{v})$ .

En aquesta descomposició cromàtica, no només caldrà disposar del filtre adaptat a l'objecte a detectar per cada canal, sinó que també s'haurà de reproduir separatament cada component de l'escena. Al situar l'escena i el filtre corresponents a un mateix canal en un correlador òptic, es podrà il·luminar el sistema amb un sol feix coherent de longitud d'ona qualsevol, donat que la informació cromàtica de l'escena i del filtre no vindrà donada ara per la  $\lambda$  del feix il·luminant, sinó per l'enregistrament previ de les diferents components cromàtiques per separat. El procés de la correlació òptica es repeteix seqüencialment per cada un dels canals RGB.

En aquest darrer mètode, la informació cromàtica que s'obté en els canals RGB no correspon exactament als canals monocromàtics centrats a les longituds d'ona dels feixos làsers que s'utilitzaven en el primer cas. Una càmera en color de tres CCD té un conjunt de filtres de color de banda ampla davant dels detectors CCD i, la informació enregistrada en cada canal es pot entendre com una integració sobre tota la banda de l'espectre on la sensibilitat del sensor és alta.

En quant a la introducció de l'escena i del filtre en el correlador òptic, podem dir que els mètodes d'enregistrament d'ambdues funcions no són únics, de manera que si cal recórrer a un processament en temps real de la informació es poden utilitzar



moduladors espacials de llum (SLM), que permeten modificar la informació de l'escena o del filtre ràpidament, i sense necessitat d'alinejar cada vegada aquest component òptic. Aquests dispositius van ser introduïts en el reconeixement en color en la referència [92]. Després, el seu ús ha esdevingut més generalitzat. Una altra possibilitat, quan el sistema no requereix operar en temps real, és la reproducció sobre transparència, per exemple, sobre fotolit en blanc i negre, codificant els diferents nivells de gris de la imatge amb un nombre determinat de punts per cel·la. Aquest procés aconsegueix una bona resolució espacial si es realitza per impressió d'alta qualitat, utilitzant els mitjans propis de les arts gràfiques.

### 2.4.3 Llindars i criteris de decisió

En el cas d'escenes acromàtiques d'objectes isoenergètics, sempre es compleix que el pic de més intensitat en el pla de correlació correspon a l'autocorrelació de l'objecte a detectar, en el cas de ser-hi present [169]. En el reconeixement acromàtic de patrons, es pren aquest pic com a referència i, com a proposta més senzilla, s'estableix un criteri de decisió basat en un llindar. Un objecte qualsevol de l'escena serà reconegut si el seu pic de correlació creuada amb el filtre supera el llindar establert. Els pics que no el superin seran considerats com a objectes diferents al patró i per tant no seran reconeguts com a tal.

No obstant, en el reconeixement d'objectes en color, la normalització de l'energia per a tots els objectes de l'escena no és possible si es requereix conservar la informació de color a més de la informació de forma. Si s'il·lumina un objecte de color amb successives longituds d'ona provenint d'una font de llum visible equienergètica, l'energia que presenta l'objecte, ja sigui per reflexió o transmissió, variarà per cada longitud d'ona. Això vol dir que objectes equienergètics de color diferent presentaran una energia igual o diferent en funció de la longitud d'ona que els il·lumina. Per altra banda, objectes de colors diferents poden tenir una intensitat similar quan s'il·lumina amb un longitud d'ona determinada. Pot succeir que, en un o més canals de reconeixement, hi hagi objectes de l'escena que pel seu color presentin una intensitat major que la de l'objecte que es vol reconèixer. Això implica

que el pic d'autocorrelació pot ser d'intensitat inferior als altres pics de correlació creuada entre objectes diferents.

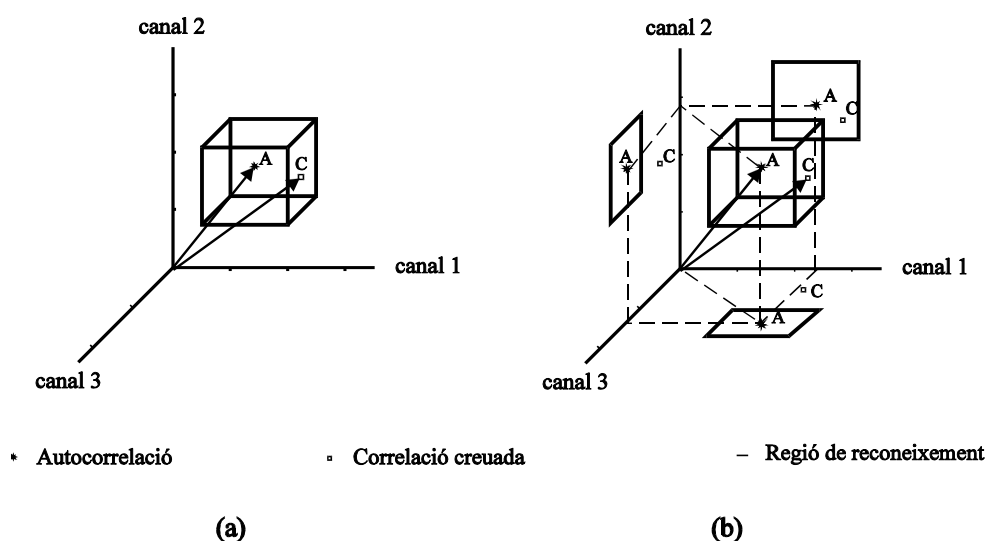
Per tant, en reconeixement en color no prendrem com a referència el pic d'intensitat màxima del pla de correlació, sinó que prendrem com a referència el valor de l'autocorrelació de l'objecte que es vol detectar. Això comporta determinar el valor d'autocorrelació del senyal en cada canal prèviament a cada prova de reconeixement. A més, no només es situarà un llindar que tingui en compte els valors inferiors al pic d'autocorrelació sinó que, per descartar valors molt superiors al pic d'autocorrelació, es situarà un doble llindar [168]. És a dir, el criteri de decisió que considerarem és que un objecte s'identifica en un canal determinat com el que es vol reconèixer, quan el seu valor màxim de correlació es situa dins el doble llindar al voltant de l'autocorrelació. L'operació del doble llindar es realitza en cada un dels canals que es considerin.

Aquest criteri de decisió, s'expressa ara en funció del valor absolut de la capacitat de discriminació que es va definir en l'Eq. (2.16) [150], quedant

$$DC = \left| 1 - \frac{cc}{ac} \right|; \quad DC \begin{cases} \leq u & \text{(reconeixement)} \\ > u & \text{(discriminació).} \end{cases} \quad (2.21)$$

El valor  $u$  és el llindar. Si el valor de  $DC$  calculat per a un objecte segons l'Eq. (2.21) no supera el valor  $u$ , llavors aquest objecte es reconeix com al patró de referència. En cas contrari, si  $DC > u$ , l'objecte és discriminat. El criteri de decisió expressat en l'Eq. (2.21) s'ha d'aplicar en cada un dels canals. Es pot representar gràficament en l'espai de les intensitats de correlació, la dimensió del qual coincidirà amb el nombre de canals ( $n$ ). En la Figura 2.6 es mostra un diagrama de la representació corresponent a un procés de reconeixement en tres canals. El vector  $A$  representa el vector d'autocorrelació en l'espai de les intensitats de correlació i, cada una de les seves components correspon al valor d'autocorrelació en cada un dels canals. Qualsevol dels objectes presents en l'escena a analitzar obtindrà el corresponent vector de correlació creuada en l'espai de les intensitats de correlació (per exemple,

el vector  $C$  en la Figura 2.6), Si escollim un llindar  $u$  i apliquem el criteri de decisió (Eq. 2.21) a cada un dels canals, es construeix una superfície de decisió formada per  $2n$  hiperplans definits per l'operació del doble llindar (amb la forma de cub en la Figura 2.6). Tots els objectes, els vectors representants dels quals estiguin dins la superfície de decisió, seran reconeguts com a similars al patró buscat. La resta d'objectes seran rebutjats i, per tant, diferenciats de l'objecte de referència. Per facilitar la visualització dels resultats en l'espai de les intensitats de correlació, representem les projeccions dels vectors i de la superfície de decisió de cada hiperplà en l'esmentat espai (Figura 2.6b).



**Figura 2.6.** (a) Superfície de decisió en l'espai de les intensitats de correlació. (b) Projeccions de la superfície de decisió i dels vectors de correlació.

## 2.5 Característiques dels dispositius optoelectrònics (càmeres CCD i moduladors espacials de llum)

Com ja s'ha fet esment en aquesta memòria, el gran desenvolupament dels sistemes de reconeixement per correlació òptica, es pot situar a partir de la introducció de les càmeres CCD i de les pantalles de cristall líquid actuant de moduladors espacials de llum (SLM). Ambdós dispositius optoelectrònics fan possible el reconeixement en

temps real, amb l'anàlisi d'imatges dinàmiques. A més, permeten la modificació digital de la informació a processar, de manera que s'assoleixen tasques de reconeixement on la capacitat de discriminació del sistema arriba a ésser molt elevada, en comparació als primers sistemes proposats. Els sistemes més moderns porten a terme identificacions d'objectes per característiques molt diverses o davant de distorsions importants en la informació a analitzar.

Donada la importància de les càmeres i dels moduladors, i tenint en compte també el seu ús en els dispositius emprats en aquest treball de recerca, volem esmentar, a manera de resum, algunes de les seves característiques més destacades.

### **2.5.1 Càmeres d'estat sòlid amb sensors CCD**

Els dispositius CCD (de l'anglès *Charge-Coupled Device*) van ser inventats per Boyle i Smith [170] l'any 1970. Una de les aplicacions més importants d'aquests sensors és la d'enregistrament d'imatges, formant part d'un conjunt anomenat popularment, càmera CCD. Per què tingui sentit com a tal, una càmera CCD, s'ha de considerar com el conjunt de l'òptica que forma la imatge sobre el sensor CCD, i l'electrònica que governa el funcionament d'aquest dispositiu.

D'aquesta manera, el dispositiu captador d'imatges CCD pot ser considerat com una superfície fotosensible composta d'un elevat nombre de fotodiodes (píxels) en formació matricial situada sobre una base, en la que les càrregues elèctriques originades per la incidència dels fotons de llum de la imatge són, primerament, emmagatzemades i, posteriorment, desplaçades per, finalment, ser convertides en senyals de tensió mesurables com a sortida de vídeo.

Trobem càmeres CCD tant en blanc i negre com en color, depenent de si la imatge enregistrada guarda la informació cromàtica de l'escena o no. A més, existeixen dues varietats de càmera en color, que consisteixen en les càmeres amb un sol sensor CCD, i les de 3 CCD. En el primer cas, s'utilitzen filtres òptics cromàtics situats en forma de mosaic davant de la superfície de captació per tal de diferenciar els colors. Un sistema divisor del conjunt d'elements fotosensibles en l'interior del sensor,

permet obtenir els senyals unitaris corresponents. Les càmeres de 3 CCD, en canvi, estan dissenyades amb tres sensors per captar de manera independent les tres components RGB de l'escena. Habitualment, s'utilitzen filtres dicroics per descompondre la llum de l'escena, i dirigir cada component de l'escena al sensor CCD corresponent.

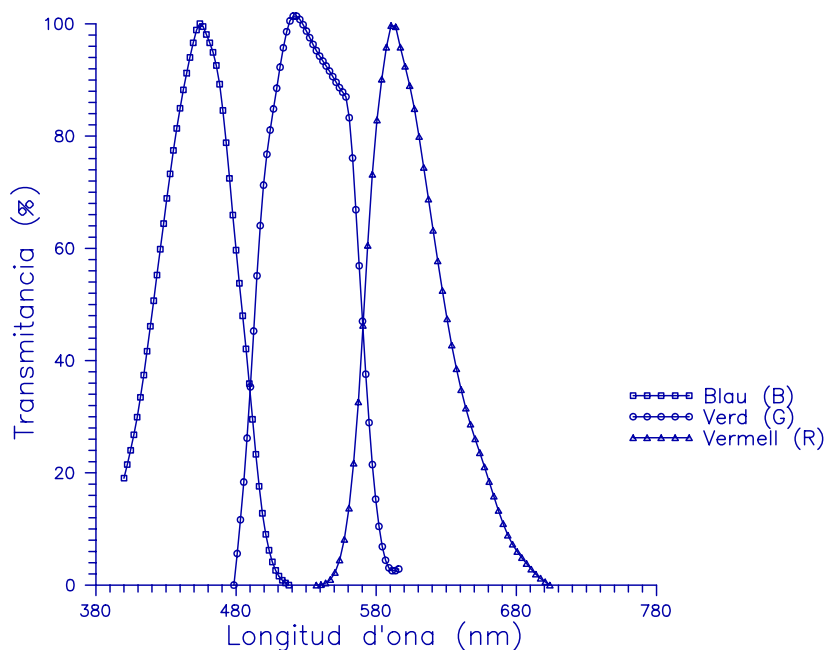
El tamany del detector, el format de la matriu de píxels, la resolució espacial, el rang dinàmic, la sensibilitat, la correcció de  $\gamma$  o el nivell de soroll, són algunes de les característiques que es tenen en compte, especialment en aplicacions d'enregistrament i processat d'imatges. Aquestes i altres característiques generals de les càmeres es poden estudiar amb detall a [171].

Una calibració prèvia del sistema de captació permet conèixer el rang dinàmic de la càmera, i a més, la linealitat de la resposta del sensor. Les càmeres en blanc i negre emprades en aquest treball corresponen ambdues a càmeres CCD Pulnix model TM-760. La diferència entre elles radica en la correcció de la gamma. En una d'elles, la sortida de la càmera és no-lineal a propòsit per tal d'optimitzar la reproducció de les imatges captades en un monitor de TV de tecnologia CTR. En els capítols corresponents, s'indicarà les característiques de la càmera emprada i es detallarà la calibració realitzada.

La resposta d'una càmera depèn tant de la resposta espectral del sensor com de la temperatura de color de la font d'il·luminació emprada. En [172] s'estudia l'efecte dels canvis d'il·luminant en les components cromàtiques enregistrades per una càmera en color, i es proposa un mètode per compensar els canvis en la informació enregistrada. De totes maneres, és moltes vegades necessari realitzar una calibració del sistema de captació en les mateixes condicions d'il·luminació en què s'enregistraran les imatges.

En la Figura 2.7 es presenten les corbes de sensibilitat espectral dels sensors RGB de la càmera en color de 3 CCD Sony, model DXC-930P, utilitzada en aquest treball. En tots els casos, l'amplada de banda dels sensors és superior als 70 nm i, per tant, la informació enregistrada correspon a la integració en tota la zona d'alta sensibilitat

dels sensors. Prèviament a l'enregistrament de cap imatge, cal situar les fonts d'il·luminació en la mateixa posició i orientació en que s'empraran, i realitzar els anomenats balanços de blanc i de negre. El balanç de blanc és necessari en els captadors d'estat sòlid per tal de prendre una referència de llum. A més, en alguns models és possible escollir entre diferents temperatures de color, i això permet adequar millor la referència de llum de la càmera amb les condicions en les que es realitzaran les captacions. El balanç de negre, per la seva part, fixa el nivell de zero en el sistema de captació, i marca el límit de les senyals de vídeo i sincronisme.



**Figura 2.7.** Sensibilitat espectral dels sensors de la càmera en color de 3 CCD Sony, model DXC-930P.

En l'enregistrament mitjançant càmera també cal tenir en compte paràmetres com el guany (*gain*) i la compensació (*offset*) del sistema de captació que, en general, es poden controlar mitjançant el software que governa la targeta d'adquisició que digitalitza el senyal rebut pel sensor de la càmera. El *gain* controla el pendent de la corba de resposta de la càmera així com el valor màxim de digitalització, i l'*offset* regula el nivell de zero, és a dir, el valor mínim d'energia incident que és necessari perquè la resposta de la càmera no sigui nul·la. En el cas d'una càmera en color, cal

determinar, prèviament a la captació, els valors de *gain* i *offset* adequats perquè la informació cromàtica de l'escena captada s'alteri el mínim possible [106].

### **2.5.2 Moduladors espacials de llum (SLM)**

Els moduladors espacials de llum [173,174] són dispositius capaços de transformar senyals electrònics o òptics incoherents, en senyals modulats espacialment que admeten un processament òptic coherent.

Existeixen diferents tecnologies en el desenvolupament dels SLM. Podem citar per exemple, els cristalls líquids, els magneto-òptics, els miralls deformables, els de múltiples pous quàntics, i les cèl·lules de Bragg acusto-òptiques. Nosaltres en centrarem en els cristalls líquids nemàtics, concretament d'estructura helicoidal.

Els cristalls líquids nemàtics es caracteritzen perquè les seves molècules són de forma allargada i es disposen orientades segons un eix de simetria que es coneix com a director del cristall. La fluïdesa característica del cristall líquid permet reorientar la direcció del director mitjançant l'aplicació d'un camp elèctric extern. La forma allargada de les molècules fa que el material actuï òpticament com un cristall birrefringent uniàxic, l'eix òptic del qual coincideix amb el director del cristall. Depenent de l'alineament del director, els canvis d'orientació de les molècules poden proporcionar ja sigui una modulació de fase, un retard de fase entre les components del camp elèctric que es propaga a l'interior del cristall líquid, o bé una rotació del pla de vibració de la llum. L'elecció d'una d'aquestes configuracions, conjuntament amb el posicionament adequat d'altres elements òptics com són els polaritzadors, fa possible treballar amb aquestes pantalles de cristall líquid, en règim de modulació quasi de fase, quasi d'amplitud, o d'acoblament dels dos règims de treball, simultàniament.

Les pantalles de cristall líquid emprades en aquesta memòria pertanyen a les pantalles estretes del videoprojector comercial Epson, model VP-100PS, que han estat àmpliament utilitzades en correladors òptics donat el seu baix cost i prestacions.

Existeixen diferents treballs que estudien les seves característiques, així com els règims de treball en modulació d'amplitud i de fase [175-184]. Per a treballar en un banc òptic, les pantalles de cristall líquid s'extreuen del videoprojector, però es mantenen connectats a ell per tal d'enviar la informació a través de l'electrònica del videoprojector. Les imatges es visualitzen mitjançant 8 bits, que equivalen a 256 nivells de gris per a cada píxel. L'aparell consta de tres potenciòmetres anomenats *brillo*, *contrast* i *color*, que controlen la visualització de la imatge mitjançant canvis en el senyal electrònic enviat a les pantalles. Donada una posició d'aquests potenciòmetres, la resposta en modulació d'amplitud i de fase és diferent per a cada pantalla. L'acció dels potenciòmetres sobre els senyals és també diferent. El *brillo* actua com un voltatge afegit a tota la pantalla, independent del nivell de gris enviat a cada píxel. El *contrast* actua com un guany en el voltatge aplicat a cada píxel. El *color* modifica la relació de voltatges aplicats sobre cada una de les tres pantalles del videoprojector. Per al seu ús en correladors òptics, és necessària la calibració prèvia de les pantalles emprades, de manera que es conegui la modulació d'amplitud i de fase que s'aconsegueix per a cada posició dels potenciòmetres de l'aparell, així com per diferents valors dels paràmetres *gain* i *offset* de la targeta digitalitzadora que transmet la informació digital cap al videoprojector.





## **Capítol 3**

# **Selecció de canals quasimonocromàtics per a la millora de la discriminació cromàtica en el reconeixement d'objectes en color**

### **3.1 Introducció**

Els mètodes actuals de descomposició multicanal per al reconeixement en color segueixen, en general, el mateix esquema RGB de partida, bé sigui amb canals monocromàtics o de banda ampla, bé sigui emprant els canals RGB directament o a partir de transformacions lineals o no-lineals aplicades sobre els seus valors.

En les primeres experiències descrites en la literatura [90,101,119], els colors dels objectes eren en general molt diferents, de manera que es va iniciar una recerca per millorar la discriminació cromàtica dels sistemes de reconeixement que encara està oberta en el present [114,118,121,122,168,183,185-187].

En aquest capítol presentem un mètode per aconseguir un augment de la capacitat discriminativa del sistema de reconeixement tenint en compte no només la forma, sinó especialment el color dels objectes. La proposta es basa en una flexibilització de l'esquema de canals cromàtics, adaptant la descomposició multicanal a la necessitat de discriminació en color del problema plantejat a través de la selecció de nous canals que seran quasimonocromàtics. Per tractar-se d'un mètode que s'aplica sobre la descomposició multicanal de la informació d'entrada, afecta exclusivament a les distribucions d'amplitud de l'objecte de referència i de l'escena a analitzar que s'utilitzaran en el procés de reconeixement. Per això, el mètode de canals seleccionats és compatible amb altres mètodes de millora de la capacitat discriminativa respecte a la forma o a les variacions d'intensitat dels objectes. Així, més endavant, en el capítol 8 d'aquesta mateixa memòria, s'apuntaran noves aplicacions en reconeixement d'objectes en color que permeten combinar el mètode de canals seleccionats amb altres mètodes desenvolupats per millorar la capacitat discriminativa del sistema. Aquesta combinació serà fins i tot possible amb mètodes desenvolupats per a l'anàlisi d'imatges inicialment acromàtiques.

Tornant al mecanisme d'aplicació del mètode de canals seleccionats, el procés de reconeixement que li segueix és anàleg al de filtrat i correlació multicanal RGB, però ara, en els nous canals. Un cop realitzada la descomposició de la informació del color d'una escena i d'un objecte concret a reconèixer, cal obtenir un filtre adaptat al patró per a cada una de les components si s'empra un correlador  $4f$  o derivats d'ell com el convergent. En el cas d'utilitzar un correlador de transformades conjuntes, la mateixa component de l'objecte a reconèixer i de l'escena s'inclouen en el pla d'entrada del correlador, i l'espectre conjunt obtingut en cada canal es pot modificar en un pas de càlcul digital intermedi (veure capítol 2). A continuació, sigui quin sigui el correlador emprat, es realitza, seqüencialment en cada canal, la correlació entre l'escena i l'objecte que es vol detectar. L'anàlisi conjunta dels valors d'intensitat dels pics de correlació obtinguts per cada canal, mitjançant criteris de decisió que inclouen l'aplicació de llindars i operacions lògiques, permetrà dir quins objectes de l'escena analitzada es reconeixen com el patró.

L'organització d'aquest capítol s'indica a continuació. La descripció del mètode de selecció adaptada d' $n$  canals quasimonocromàtics, es detalla en l'apartat 3.2, on es fa principal atenció a com s'han d'escollir els canals monocromàtics més adients donat un problema de reconeixement [168]. Un cas pràctic on la informació del color dels objectes és fins i tot més rellevant que la pròpia forma dels mateixos, se'ns planteja al portar a terme una tasca d'identificació de fruites, ja siguin de diferent varietats, o en diferent estat de maduració. En aquest cas, on els colors presents a l'escena pertanyen a una mateixa gamma de tonalitat, la capacitat de discriminació del sistema ha de ser superior en l'anàlisi de la informació del color que en el de la forma. En la secció 3.3 analitzarem la resposta dels filtres de potències fraccionàries (FPF) davant de l'anàlisi d'objectes de color uniforme i iguals en forma, cas en el qual ens trobem de manera aproximada si volem analitzar fruites en diferents estats de maduració o de diferents varietats. A continuació, en la secció 3.4, presentarem els resultats que s'obtenen en tots els casos plantejats i que permetran avaluar la viabilitat de la proposta efectuada. La secció 3.4 està dividida en diferents apartats per tal de poder descriure amb detall les escenes test creades artificialment per diferents tonalitats i l'escena natural de fruites emprada en aquesta anàlisi (apartat 3.4.1), l'obtenció experimental de la descomposició cromàtica en les configuracions RGB (apartat 3.4.2a) i dels  $n$  canals quasimonocromàtics (apartat 3.4.2b). Finalment, dins de la secció 3.4, aportarem pròpiament els resultats de simulació de les correlacions (apartat 3.4.3). En la secció 3.5, incloem la realització experimental de la correlació òptica, basant-nos en la descomposició cromàtica proposada i comparant-la novament amb el sistema de canals RGB. En aquesta secció 3.5 presentem separadament el muntatge òptic emprat (apartat 3.5.1) i els resultats de reconeixement experimentals per les escenes test i de fruites (apartat 3.5.2). Conclourem el capítol amb la secció 3.6, on s'aplica la proposta de descomposició adaptada d' $n$  canals quasimonocromàtics per al reconeixement d'objectes policromàtics (objectes amb variació interna de color). A més, s'estableix comparació amb la proposta de descomposició basada en els models de visió humana del color [114,185].

### **3.2 Correlació multicanal adaptada. Selecció d' $n$ canals quasimonocromàtics**

Una via per a la millora de la capacitat de discriminació cromàtica del sistema de reconeixement seria potenciant les diferències que es presentin en les reflectàncies espectrals dels colors dels objectes. Això es pot aconseguir utilitzant un nou conjunt de canals quasimonocromàtics que optimitzin el mostreig de l'espectre visible [168]. Ara bé, amb la finalitat de no augmentar innecessàriament el nombre de canals de correlació, cosa que complicaria la resolució dels problemes de reconeixement, és convenient disposar d'un coneixement previ dels colors de l'escena que permeti restringir el nombre de canals de la nostra proposta. Per exemple, si els colors dels objectes són de tonalitats molt diferents (vermell, verd, groc, blau), una descomposició RGB convencional serà probablement suficient. Però si els objectes presenten colors en una gamma de tonalitat comuna, llavors la descomposició RGB pot ser insuficient i la nostra proposta de canals seleccionats per aquesta gamma pot millorar la discriminació. Així doncs, aquesta solució es planteja especialment en aquells casos de reconeixement de formes en color on es requereixi fer més precisa la discriminació cromàtica dins una gamma de colors afins.

Per explicar la selecció de canals, suposem que els objectes de l'escena a analitzar són de color uniforme dins d'una gamma i que podem disposar de les seves corbes de reflectància espectral. Els canals més convenients per discriminar els objectes pel seu color, a més de per la seva forma, s'establiran a partir de l'anàlisi prèvia de les esmentades reflectàncies. Per unes tonalitats concretes trobarem probablement uns intervals de longituds d'ona on les diferències proporcionals de reflectància seran més accentuades. Així doncs, el que en un principi podria implicar un augment de les components cromàtiques d'una escena o objecte si es fes un mostreig regular en tota l'amplada de l'espectre visible (380 - 780 nm), es pot restringir en unes regions més estretes de l'espectre, estudiant per separat els colors que presentin problemes de discriminació, que molt possiblement coincidiran amb colors pertanyents a una mateixa gamma.

Escollirem com a nous canals per a la descomposició cromàtica, les longituds d'ona concretes de l'espectre on les diferències proporcionals de les reflectàncies dels objectes de color uniforme siguin més accentuades. D'aquesta manera, assignarem a cada objecte de color un nombre de components que coincidirà amb el nombre de canals espectrals escollits, i l'amplitud que tindrem en cada component serà proporcional al valor de la reflectància de l'objecte en la longitud d'ona corresponent. Els canals seran seleccionats de manera automàtica a partir de les corbes de reflectància espectral, comparades de dos en dos i, per tant, podran ser diferents per cada gamma.

Per obtenir les components de l'escena en els nous canals seleccionats podem pensar en la utilització d'uns filtres passa-banda que transmetin en regions de l'espectre en les quals vulguem obtenir informació, de la mateixa manera que per a obtenir la informació en les components RGB d'una càmera en color de 3 CCD, es fan servir filtres de banda passant davant dels detectors (secció 2.5.1 del capítol 2). La condició ara, serà que l'amplada de banda del filtre sigui suficientment reduïda (quasimonocromàtica) per assegurar la definició del canal seleccionat i l'efectivitat del mètode.

En particular, la utilització de filtres interferencials de transmitància en banda estreta (amplada de banda  $\approx 10$  nm) centrats en les diferents longituds d'ona de l'espectre que s'han seleccionat, ens poden permetre obtenir, en primera aproximació, els valors de reflectància dels objectes de l'escena per aquestes longituds d'ona. No obstant, els filtres interferencials emprats en la realització pràctica, seran aquells disponibles comercialment que tinguin la màxima transmitància en la longitud d'ona més propera a les  $\lambda$  escollides segons les diferències proporcionals entre les corbes de reflectància. Els filtres interferencials es situaran successivament davant de l'objectiu d'una càmera CCD en blanc i negre en el moment de la captació de l'escena test, amb la finalitat d'aconseguir les successives components de l'escena en els canals definits pels filtres de color interferencials.

En el mètode de la selecció adaptada de canals, com acabem de presentar, és aconsellable tenir un coneixement previ de les reflectàncies espectrals dels colors dels objectes presents a l'escena per tal de fer una bona selecció dels canals. En tot cas, si les corbes de reflectància espectral no estan disponibles, la selecció dels canals es pot fer a partir d'un test format per un conjunt de mostres de colors similars als colors que hi ha presents a l'escena. De fet, en aquest capítol presentarem resultats de la descomposició en canals quasimonocromàtics, tant per escenes artificials creades per avalar la viabilitat de la proposta presentada, com per una escena que inclou objectes naturals, analitzada a partir de la informació prèvia estreta d'una escena d'entrenament (gamma taronja) creada artificialment, on hi ha presents colors molt semblants als dels objectes reals (cítrics).

El sistema de descomposició cromàtica que acabem de descriure, es pot entendre per tant, com un pre-processament de la informació anterior a la correlació. Els passos que li segueixen són el filtratge adaptat, si es realitza la correlació amb un correlador de Vander Lugt, la correlació multicanal en els canals seleccionats i l'aplicació de normes de decisió. Els dos darrers passos són comuns tant si s'utilitza una arquitectura de correlador de Vander Lugt com si la correlació es porta a terme mitjançant l'anomenat correlador de transformades conjuntes, ambdós processadors descrits en la secció 2.2 del capítol anterior.

Per últim, notem que si la banda de transmitància dels filtres de color (interferencials) és suficientment estreta, la corba d'emissió espectral de la font de llum utilitzada com a il·luminant de l'escena en la descomposició multicanal no influeix de manera rellevant. A menys, clar està, que l'emissió de la font sigui nul·la (o estigui fora del rang dinàmic de la càmera) precisament en les longituds d'ona de transmitància del filtre. En efecte, els valors de l'emissió de la font en les longituds d'ona seleccionades només afecten com a constants globals a les distribucions d'intensitat captades per la càmera en cada canal. El fet que aquestes constants globals puguin variar d'un canal a un altre no és significatiu, ja que el procés de filtratge, correlació i aplicació d'un llindar, es realitza en cada canal per

separat i no s'efectua la comparació de valors de correlació relatius entre uns canals i els altres.

### **3.3 Resposta dels filtres de potències fraccionàries (FPF) per objectes similars en forma i de color uniforme**

Com ja s'ha vist en el capítol 2, els filtres de potències fraccionàries (FPF, de l'anglès *fractional power filters*) van ser introduïts per Vijaya Kumar i Hassebrook [6] a fi d'englobar un conjunt de filtres adaptats aplicats en el reconeixement de formes per correlació òptica i, generalitzar la seva definició a altres possibilitats de filtratge noves. Els FPF es poden implementar òpticament tant en un correlador de Vander Lugt [6], com en un de transformades conjuntes [12], de manera que és possible extreure'n conclusions independentment de l'arquitectura finalment utilitzada. Per a la descripció matemàtica d'aquest apartat, però, emprarem la notació corresponent al filtratge adaptat, pas només inclòs en el correlador en cas d'utilitzar l'arquitectura de Vander Lugt. Ho fem així, donat que la definició donada per Vijaya Kumar i Hassebrook en [6] està pensada per la generació d'un filtre adaptat.

El nostre objectiu és analitzar la codificació de la informació de color en aquest conjunt de filtres adaptats en un cas particular senzill, de manera que ens sigui possible escollir el tipus de filtratge adaptat que creiem més convenient pel problema de reconeixement plantejat. El cas a analitzar és el d'escenes que continguin objectes similars en forma i de color uniforme. De manera molt aproximada, aquest serà el cas, com veurem, d'analitzar fruites de diferents varietats.

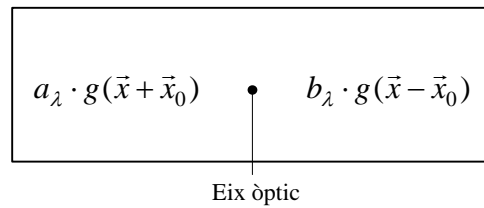
Suposem que partim de la transparència d'una escena,  $s(\vec{x})$ , en la qual hi ha representats dos objectes d'igual forma i tamany que presenten un color uniforme però distint, amb una transmitància espectral diferent. Això és equivalent a dir que aquests dos objectes es diferencien només pel seu color.

Quan aquests objectes són il·luminats per una llum monocromàtica de longitud d'ona  $\lambda$ , els dos objectes es podran representar, per tant, per una mateixa funció  $g(\vec{x})$ ,



multiplicada per un factor proporcional a la transmitància de cada un,  $a_\lambda$  i  $b_\lambda$  respectivament, per a la longitud d'ona  $\lambda$  il·luminant (Figura 3.1). L'escena, per aquesta longitud d'ona  $\lambda$ , s'escriu com la funció

$$s_\lambda(\vec{x}) = a_\lambda \cdot g(\vec{x} + \vec{x}_0) + b_\lambda \cdot g(\vec{x} - \vec{x}_0). \quad (3.1)$$



**Figura 3.1.** Transparència de l'escena d'hipòtesi, composta per dos objectes que només es diferencien pel seu color (uniforme). Cada objecte s'ha representat per la seva transmitància a la longitud d'ona de l'il·luminant  $\lambda$ .

Suposem que volem reconèixer el primer objecte, a l'esquerra de l'escena. En aquest cas haurem de prendre com a senyal la funció  $a_\lambda \cdot g(\vec{x})$  per generar el filtre adaptat.

Tenint en compte la definició dels FPF (Eq. 2.8), i que per la linealitat de la transformada de Fourier es satisfà que  $TF\{a_\lambda \cdot g(\vec{x})\} = a_\lambda \cdot TF\{g(\vec{x})\}$ , en el cas que ens ocupa, el filtre adaptat adquirirà la forma

$$H_{FPF}(\vec{\nu}, \lambda) = \begin{cases} a_\lambda^p \cdot |G(\vec{\nu})|^p \cdot e^{-i\Phi(\vec{\nu})} & \text{si } |G(\vec{\nu})| \neq 0 \\ 0 & \text{si } |G(\vec{\nu})| = 0, \end{cases} \quad (3.2)$$

on  $G(\vec{\nu})$  és la transformada de Fourier de la funció  $g(\vec{x})$  que descriu l'objecte que vol ser detectat, i  $p$  és un nombre real que defineix el tipus de filtre emprat.

Recordem que el filtre clàssic o CMF correspon al valor de  $p = 1$ , el filtre només de fase o POF a  $p = 0$ , i el filtre invers o IF pren el valor de  $p = -1$ .

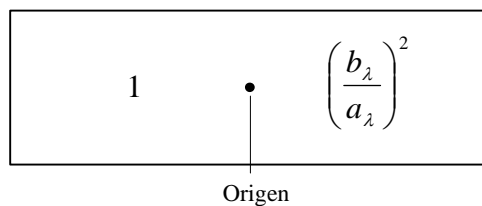
En el pla de freqüències del correlador òptic es realitza el producte de la TF de l'escena pel filtre

$$\begin{aligned}
 S_\lambda(\vec{\nu}) \cdot H_{FPF}(\vec{\nu}, \lambda) &= \\
 &= \left\{ a_\lambda \cdot |G(\vec{\nu})| \cdot e^{i\Phi(\vec{\nu})} \cdot e^{i2\pi\vec{\nu} \cdot \vec{x}_0} + b_\lambda \cdot |G(\vec{\nu})| \cdot e^{i\Phi(\vec{\nu})} \cdot e^{-i2\pi\vec{\nu} \cdot \vec{x}_0} \right\} \cdot H_{FPF}(\vec{\nu}, \lambda),
 \end{aligned} \tag{3.3}$$

i, després d'una segona transformació de Fourier, s'obté la distribució d'intensitat en el pla de correlació. En el pla de sortida del correlador s'obté una distribució d'amplitud que conté dos pics lluminosos en les posicions ocupades per cada un dels objectes, respectivament. Si representem per  $A_g(\vec{x})$  la funció d'autocorrelació de  $g(\vec{x})$ , es dedueix que el màxim d'amplitud dels pics és proporcional a  $a_\lambda^{p+1} \cdot A_g(\vec{x} + \vec{x}_0)$  per l'objecte de referència (autocorrelació) i  $a_\lambda^p \cdot b_\lambda \cdot A_g(\vec{x} - \vec{x}_0)$  pel segon objecte (correlació creuada). La correlació  $c_\lambda(\vec{x})$  en el pla de sortida és, excepte constants globals, igual a

$$c_\lambda(\vec{x}) = a_\lambda^{p+1} \cdot A_g(\vec{x} + \vec{x}_0) + a_\lambda^p \cdot b_\lambda \cdot A_g(\vec{x} - \vec{x}_0). \tag{3.4}$$

Donat que l'adquisició de dades en la zona de correlació del pla de sortida es realitza habitualment per un detector de potència (per exemple, una càmera CCD), els màxims finals de correlació normalitzats al valor màxim de l'autocorrelació de l'objecte de referència seran 1 i  $(b_\lambda/a_\lambda)^2$ , pels dos objectes de l'escena, respectivament (Figura 3.2). És sobre aquests valors normalitzats als quals s'aplica l'operació llindar.



**Figura 3.2.** Zona de correlació del pla de sortida. Màxims d'intensitat enregistrats per la càmera CCD un cop normalitzats al valor d'intensitat de l'autocorrelació de l'objecte de referència.

Veiem doncs, que els valors màxims de la correlació normalitzada són independents de  $p$ . Donat que els líndars s'apliquen sobre aquests valors màxims, això implica que qualsevol filtre de potències fraccionàries permet arribar a les mateixes conclusions en el cas teòric del reconeixement d'objectes d'igual forma i tamany, i de colors uniformes diferents. Aquesta conclusió concorda amb el resultats obtinguts per Millán i Yzuel en [188] per objectes de color uniforme sobre fons negre.

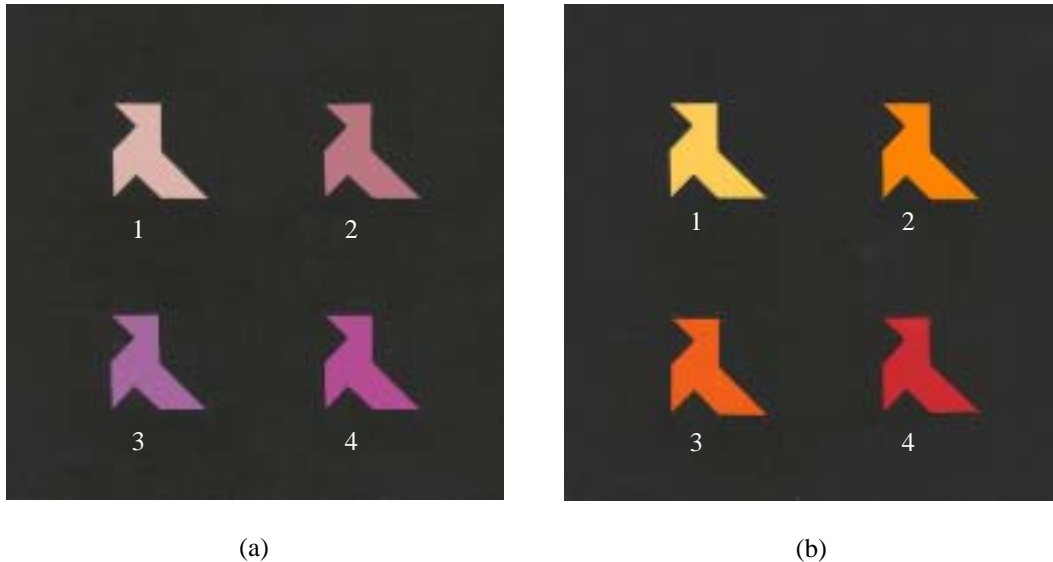
Així doncs, en aquest cas hipotètic, l'elecció del filtre vindria donada per altres paràmetres o circumstàncies com poden ser els relacionats amb la qualitat del pic de correlació en la realització experimental, més que pel valor màxim que aconseguim.

### 3.4 Aplicació del mètode de selecció adaptada de canals

#### 3.4.1 Selecció de canals adaptada a les característiques de les escenes test

Amb la finalitat de validar la proposta de la descomposició adaptada en  $n$  canals quasimonocromàtics, hem generat un conjunt d'escenes test, cada una de les quals conté quatre objectes iguals en forma i tamany, però de color uniforme diferent dins una mateixa gamma. El conjunt de les gammes estudiades abraça la major part de l'espectre visible [189], tot i que en aquest capítol, presentarem els resultats obtinguts per les gammes magenta (Figura 3.3a) i taronja (Figura 3.3b), com a més representatives dels casos estudiats donat que són les que presenten més dificultats en el reconeixement. D'altra banda, un objectiu no menys important d'aquestes

escenes artificials és la seva utilització, en un primer pas, per poder determinar les millors condicions de treball en l'anàlisi de possibles escenes reals. Un exemple d'això també serà introduït en aquest capítol.



**Figura 3.3.** (a) Escena de la gamma magenta, i (b) de la gamma taronja.

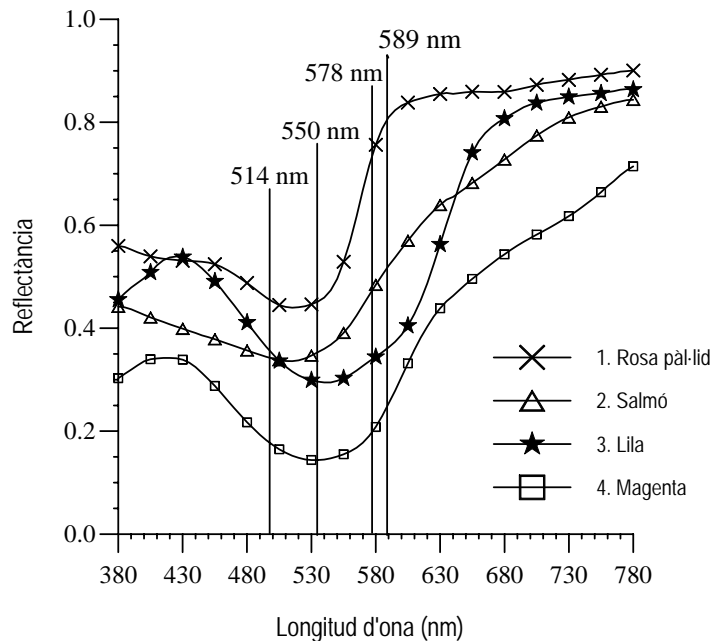
Els objectes que formen les escenes s'han construït amb cartolines de colors. De cada una de les cartolines, se n'ha obtingut la seva reflectància espectral mitjançant l'ús d'un espectrofotòmetre UV/VIS de Perkin Elmer, model Lambda 3B series PECSS. Aquest aparell mesura la quantitat de llum difosa a diferents longituds d'ona per una mostra opaca mitjançant la comparació d'un blanc patró. La mesura de la quantitat de llum difosa per una superfície rugosa es realitza mitjançant una esfera integradora. Hem analitzat les cartolines problema en el rang de 380 – 780 nm amb un pas de 5 nm en la longitud d'ona i una velocitat d'escombrat de 60nm/min. Per presentar aquests resultats s'han agrupat les corbes espectrals obtingudes per gammes de colors (Figures 3.4 i 3.5), de la mateixa manera que s'han agrupat els objectes en les escenes test.

En les corbes de reflectància espectral corresponents a les gammes magenta (Figura 3.4) i taronja (Figura 3.5) es senyalen també les longituds d'ona centrals dels filtres interferencials més adients per a la diferenciació dels colors. Aquestes longituds d'ona s'han obtingut seguint el procediment, ja descrit en la secció 3.2, que detallem a continuació.

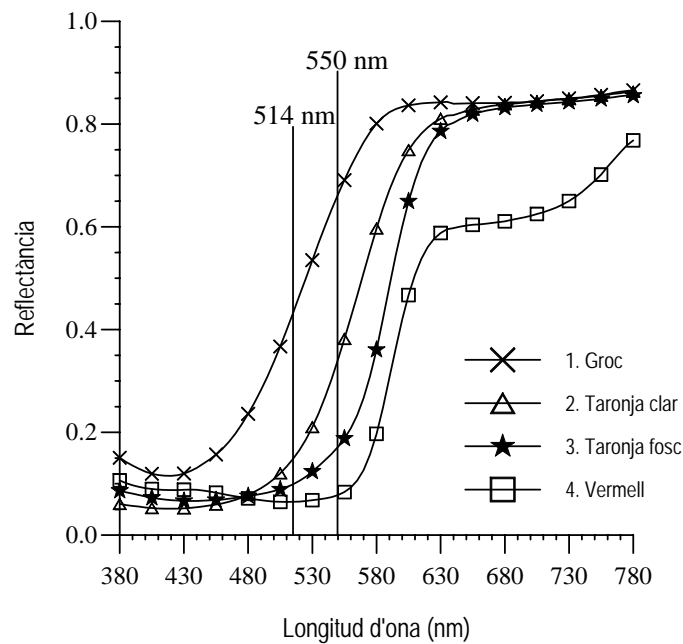
Per a la selecció de canals, s'han calculat les diferències proporcionals per a cada longitud d'ona entre totes les possibles parelles de corbes pertanyents a una mateixa gamma. I de tots aquests valors, s'ha considerat la longitud d'ona corresponent a la diferència proporcional major. En altres paraules, donat dos colors d'una gamma identificats com 1 i 2, hem buscat la longitud d'ona  $\lambda_0$  que compleixi

$$\frac{R_{1\lambda_0} - R_{2\lambda_0}}{\frac{1}{2}(R_{1\lambda_0} + R_{2\lambda_0})} = \max \left\{ \frac{|R_{1\lambda} - R_{2\lambda}|}{\frac{1}{2}(R_{1\lambda} + R_{2\lambda})} \right\} \quad \lambda \in \{380, 385, \dots, 780 \text{ nm}\}, \quad (3.5)$$

essent  $R_{i\lambda}$  el valor de reflectància pel color  $i$  en la longitud d'ona  $\lambda$ .



**Figura 3.4.** Corbes de reflectància espectral de cada un dels objectes de la gamma magenta (Figura 3.3a). S'indica el conjunt de longituds d'ona dels filtres interferencials escollits ( $\lambda_{fi}$ ).



**Figura 3.5.** Corbes de reflectància espectral de cada un dels objectes de la gamma taronja (Figura 3.3b). S'indica el conjunt de longituds d'ona dels filtres interferencials escollits ( $\lambda_{fi}$ ).

D'aquesta manera, i repetint el mateix procediment per cada parella de corbes de reflectància d'una mateixa gamma, obtenim un conjunt de longituds d'ona on hi ha les diferències proporcionals més grans entre els colors de la mateixa gamma de tonalitat. El nombre màxim de longituds d'ona que trobarem per a una gamma serà, per tant, de  $(N-1)!$  on  $N$  és el nombre de colors inclosos en la gamma. Per exemple, en la Taula 3.1 s'indica el conjunt de longituds d'ona  $\lambda_0$  representatives pels colors rosa pàl·lid (R) de l'objecte 1, salmó (S) de l'objecte 2, lila (L) de l'objecte 3 i magenta (M) de l'objecte 4 de la gamma magenta calculades a partir de l'Eq. (3.5). Per a totes aquestes longituds  $\lambda_0$ , també s'inclou la informació relativa a la diferència proporcional màxima per cada parella de corbes de reflectància, així com el valor mig de reflectància per aquesta longitud d'ona. La informació corresponent a aquest valor mig de reflectància ens és útil per avaluar si la zona de l'espectre escollida és prou energètica per enregistrar la corresponent component cromàtica de l'escena. Per últim, en la mateixa Taula 3.1 proporcionem les longituds d'ona  $\lambda_{fi}$  corresponents als filtres interferencials disponibles en la pràctica més propers a les  $\lambda_0$ . Informació anàloga es recull en la Taula 3.2 per als colors inclosos en la gamma taronja.

**Taula 3.1.** Canals seleccionats per a la gamma magenta (Figura 3.3a). Longituds d'ona  $\lambda_{fi}$  dels filtres interferencials escollides per a la gamma magenta tenint en compte les  $\lambda_0$  on s'obtenen les màximes diferències proporcionals entre les corbes de reflectància dels colors inclosos en aquesta gamma. Els colors pertanyents a aquesta gamma són el rosa pàl·lid (R), el salmó (S), el lila (L) i el magenta (M).

Colors	Dif. Proporcional (%)	Valor mig	$\lambda_0$ (nm)	$\lambda_{fi}$ (nm)
	$\left\{ \frac{ R_{1\lambda_0} - R_{2\lambda_0} }{\frac{1}{2}(R_{1\lambda_0} + R_{2\lambda_0})} \right\}$	$\frac{R_{1\lambda_0} + R_{2\lambda_0}}{2}$		
<b>R – S</b>	44	0.65	585	589
<b>R – L</b>	76	0.59	590	589
<b>R – M</b>	116	0.42	570	578
<b>S – L</b>	36	0.46	595	589
<b>S – M</b>	86	0.28	560	550
<b>L – M</b>	70	0.22	525	514

**Taula 3.2.** Canals seleccionats per a la gamma taronja (Figura 3.3b). Longituds d'ona  $\lambda_{fi}$  dels filtres interferencials escollides per a la gamma taronja tenint en compte les  $\lambda_0$  on s'obtenen les màximes diferències proporcionals entre les corbes de reflectància dels colors inclosos en aquesta gamma. Els colors pertanyents a aquesta gamma són el groc (G), el taronja clar (TC), el taronja fosc (TF) i el vermell (V).

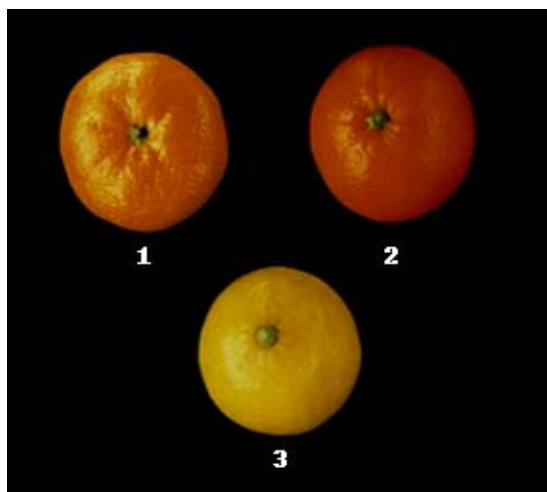
Colors	Dif. Proporcional (%)	Valor mig	$\lambda_0$ (nm)	$\lambda_{fi}$ (nm)
	$\left\{ \frac{ R_{1\lambda_0} - R_{2\lambda_0} }{\frac{1}{2}(R_{1\lambda_0} + R_{2\lambda_0})} \right\}$	$\frac{R_{1\lambda_0} + R_{2\lambda_0}}{2}$		
<b>G – TC</b>	103	0.20	495	514
<b>G – TF</b>	120	0.29	520	514
<b>G – V</b>	158	0.35	545	550
<b>TC – TF</b>	64	0.32	560	550
<b>TC – V</b>	128	0.26	560	550
<b>TF – V</b>	81	0.14	555	550

Observem com en ambdós casos, el nombre de longituds d'ona escollides de filtres interferencials diferents es redueix fortament, degut a que les diferències entre els colors d'una mateixa gamma es solen concentrar en unes regions concretes de l'espectre visible. A més, la necessitat d'aproximar les longituds d'ona  $\lambda_0$  per les  $\lambda_{fi}$  dels filtres interferencials comercials amb els quals, a la pràctica, materialitzem els canals seleccionats, porta a reduir encara més, el nombre de canals finalment emprats. Per aquest motiu, només considerarem quatre longituds d'ona en la gamma magenta ( $\lambda_1 = 514$  nm,  $\lambda_2 = 550$  nm,  $\lambda_3 = 578$  nm i  $\lambda_4 = 589$  nm), i dos en el cas de la gamma taronja ( $\lambda_1 = 514$  nm i  $\lambda_2 = 550$  nm).

En particular, la gamma taronja s'estudia també amb més detall perquè presentarem una aplicació més directa en la identificació d'objectes naturals (per exemple, fruites). La utilització de cartolines dins la gamma taronja ens possibilita fer un estudi amb colors semblants als que poden aparèixer en diferents cítrics (mandarines, taronges, llimones) o en diferents varietats d'un mateix tipus de cítric, de manera que, partint dels resultats obtinguts amb l'escena test, puguem estimar els filtres interferencials adequats per operar dins aquesta gamma quan en el problema de reconeixement es presenta una escena real com la de la Figura 3.6 [190]. En aquesta imatge apareixen una mandarina de la varietat *Satsuma Owari* (objecte 1) de color taronja clar, una taronja *Thomson* (objecte 2) de color taronja fosc, i una llimona de la varietat *Eureka* (objecte 3) de color groc. En el procés de reconeixement es vol detectar el cítric de color taronja clar, la mandarina de la varietat *Satsuma Owari*, i diferenciar-lo dels altres cítrics. Destaquem que l'escena escollida de cítrics constitueix un cas difícil de reconeixement i discriminació, ja que els objectes presents en ella tenen formes molt similars i es diferencien principalment pel color.

En tots els casos de reconeixement es comparen els resultats obtinguts mitjançant la descomposició multicanal proposada, amb els resultats que es deriven amb l'ús de la descomposició RGB mitjançant càmera 3CCD. D'aquesta manera es pot tenir una referència per avaluar la millora en la capacitat discriminativa del sistema proposat respecte el sistema RGB convencional.





**Figura 3.6.** Escena d'objectes naturals (cítrics). Imatges cedides per l'I.V.I.A. [190].

En l'annex I d'aquesta memòria, s'adjunten les imatges test corresponents a altres gammes de tonalitat (marró, blava i verda) estudiades en [189], conjuntament amb les corbes de reflectància espectral dels colors inclosos en cada una de les escenes test, indicant les longituds d'ona dels filtres interferencials escollides ( $\lambda_{fi}$ ). S'inclou també en l'annex I, una taula de valors similar a la Taula 3.1 per cada una de les gammes, on es dona la informació de la diferència proporcional màxima per cada parella de corbes d'una mateixa gamma, les  $\lambda_0$  on s'obtenen aquestes diferències màximes, el valor mig de reflectància per cada  $\lambda_0$  i, finalment, les  $\lambda_{fi}$  escollides per a realitzar la descomposició cromàtica en els canals quasimonocromàtics.

### 3.4.2 Descomposició multicanal experimental

A continuació, ja amb les escenes dels ocells de paper (Figura 3.3), els filtres interferencials reals, i les càmeres tant en blanc i negre com en color, s'han realitzat experimentalment les descomposicions multicanals, i posteriorment s'han portat a terme les correlacions, de manera simulada, a partir de les components de les escenes captades experimentalment. En aquest pas de descomposició cromàtica experimental estem introduint elements que modifiquen de diferent manera la informació de l'escena en quant a imperfeccions en els contorns degudes a la realització manual de les escenes, o a la resposta del sistema de captació.

Diverses fons de variació en la forma dels objectes, idealment iguals, són petits girs relatius entre els quatre objectes d'una escena, o bé lleugeres rotacions de l'escena respecte la horitzontal dels sensors CCD de la càmera de captació. Ambdós efectes, junt amb la discretització de l'escena, provoquen que, per exemple, els contorns rectes dels objectes deixin de ser-ho en la matriu de la imatge captada.

Altres problemes que allunyen les escenes captades de la idealitat poden provenir de la il·luminació. Si aquesta no és uniforme s'introdueixen variacions en els nivells de gris dels píxels dels objectes que es tradueixen en canvis de forma.

Tots aquests aspectes s'han intentat eliminar o minimitzar en la realització experimental de la descomposició cromàtica multicanal, però no es poden evitar totalment.

### 3.4.2a Sistema de descomposició RGB

Les components RGB s'han obtingut a partir de la captació de les escenes mitjançant una càmera en color de 3 CCD Sony, model DXC-930 P, que proporciona, per a cada imatge, tres matrius de píxels corresponents al senyal captat per tres sensors (que contenen detectors CCD), les respostes dels quals estan centrades respectivament en la zona dels vermells, dels verds i dels blaus (Figura 2.7 de la secció 2.5.1). La il·luminació en aquest cas s'ha realitzat amb llum fluorescent de 5000 K de temperatura de color (Figura 3.7).

La targeta d'adquisició MATROX, model MVP-AT, que digitalitza el senyal rebut pels sensors CCD, tant en el cas de la càmera en color com de la càmera en b/n, és controlada via software i permet fixar uns valors de *gain* ( $G$ ) i *offset* ( $O$ ) que cal escollir per captar en les millors condicions. Els dos paràmetres poden variar en el rang entre 0 i 255. En la digitalització de les components RGB amb càmera en color, hem pres de referència el treball realitzat per Corbalán [191] i s'ha escollit  $G = 0$  i  $O = 0$ , que correspon a la situació on menys es modifica la relació entre les tres components RGB, a la vegada que s'aconsegueixen uns millors contrastos entre els objectes i el fons, ja que el negre té les tres components nul·les.

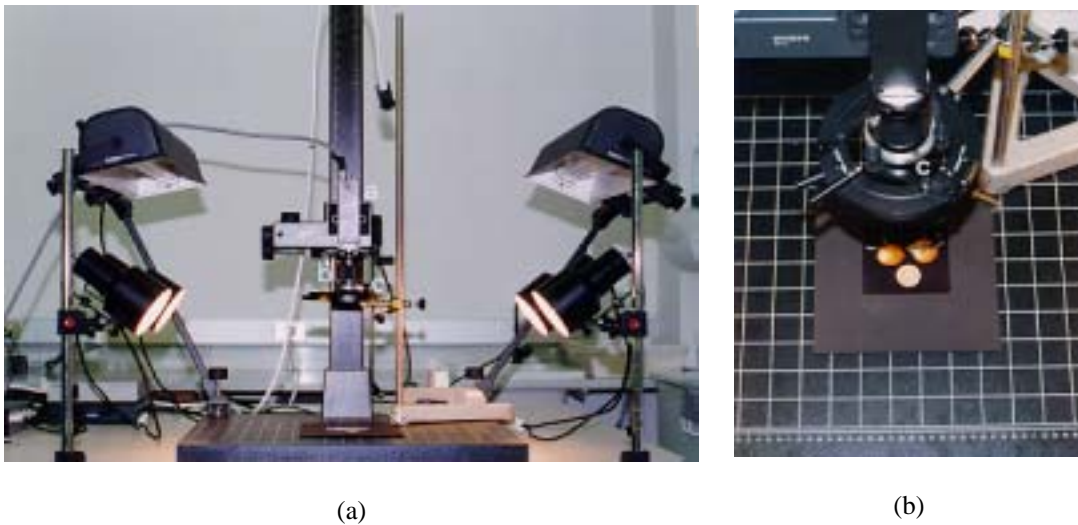


**Figura 3.7.** Sistema de captació per a l'obtenció de la descomposició multicanal RGB. Càmera en color de 3 CCD Sony DXC-930 P. Il·luminació amb llum fluorescent de 5000 K de temperatura de color.

### 3.4.2b Sistema de descomposició adaptada en $n$ canals quasimonocromàtics

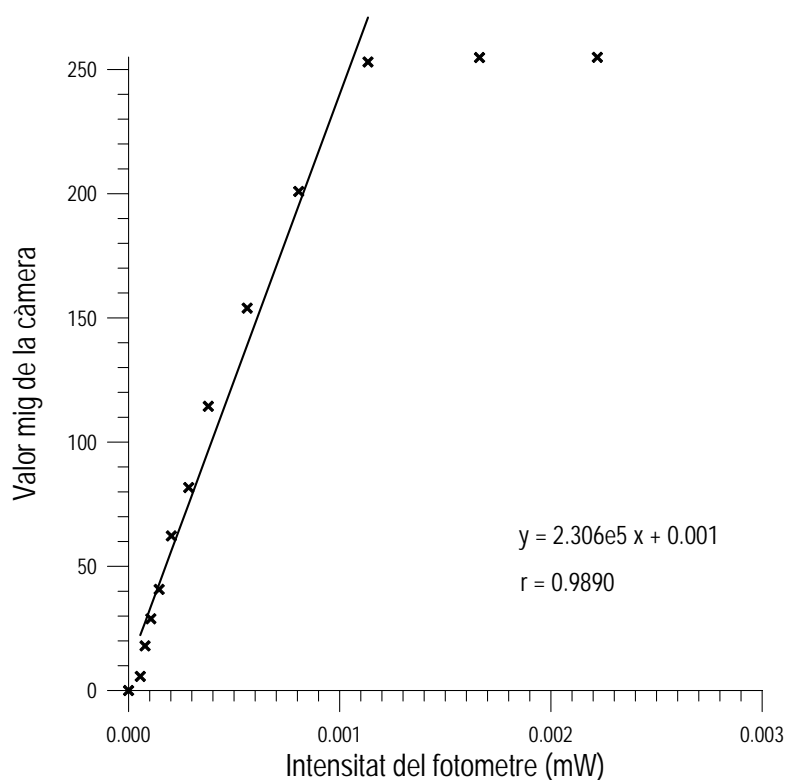
Per a la descomposició de l'escena en les components dels canals quasimonocromàtics seleccionats s'ha utilitzat una càmera en b/n CCD Pulnix, model TM-765 GL. Davant de l'objectiu de la càmera s'ha situat el filtre interferencial de transmitància màxima en la longitud d'ona escollida (Figura 3.8). S'ha utilitzat llum de tungstè de 500 W i de temperatura de color al voltant de 2700 K per a la il·luminació de les escenes. Els filtres interferencials només deixen passar el 40% de l'energia incident. Donat que la banda passant dels filtres és molt estreta (10 nm), i que els canals operen independentment, l'espectre de la il·luminació utilitzada proporciona un factor multiplicador global per canal, del qual podem prescindir tenint en compte que afecta per igual a tots els objectes de l'escena que són captats simultàniament.

Per caracteritzar la resposta de la càmera CCD en b/n s'ha realitzat una calibració [189]. A partir d'aquesta informació estem en condicions d'escollir els paràmetres de *gain* i *offset* (del mateix software d'adquisició que en el cas de la càmera en color), que siguin més adients per a la captació.



**Figura 3.8.** (a) Sistema de captació per a l'obtenció de la descomposició multicanal en els canals quasimonocromàtics seleccionats. Il·luminació amb llum de tungstè. (b) Perspectiva del mateix muntatge on es pot observar el filtre interferencial. En ambdues figures: *a*, Càmera en b/n CCD Pulnix TM-765 GL; *b*, objectiu i *c*, filtre interferencial de banda estreta.

La resposta més lineal de la càmera s'obté per  $G = 0$  i  $O = 70$  (Figura 3.9). Malgrat això, al captar les escenes en aquestes condicions hem observat que els contrastos amb el fons no són els més adequats ja que donen un senyal no nul del negre. Per aquest motiu, és preferible enregistrar les escenes amb uns valors de *gain* i *offset* nuls que ens assegurin un contrast ben diferenciat entre el fons i els objectes, tot i que la càmera no actua de manera absolutament lineal en un petit marge en els valors baixos d'intensitat. Quan fem la mateixa càmera en l'enregistrament dels plans de sortida del correlador, on és necessari prendre mesures dels màxims d'intensitat, i calcular-ne els valors relatius entre diferents pics, utilitzarem els valors de *gain* i *offset* esmentats ( $G = 0$  i  $O = 70$ ) que donen la resposta més lineal, i a més aplicarem una taula de correspondència o LUT (de l'anglès *look up table*) que ens permet corregir les divergències respecte a la linealitat en la resposta de la càmera.



**Figura 3.9.** Resposta de la càmera CCD Pulnix en b/n per als valors de *gain* = 0 i *offset* = 70. S'indica l'equació de la recta i el coeficient de regressió, que ens permet obtenir el LUT.

### 3.4.3 Correlacions simulades

#### 3.4.3a Gamma magenta

Primerament, considerem l'escena de la gamma magenta (Figura 3.3a) on l'objecte 2 és pres com a referència. Realitzem les correlacions simulades en dos sistemes multicanals: el sistema convencional RGB aplicat a les components de l'escena obtingudes experimentalment per la càmera en color de 3 CCD, i seguidament, el sistema de canals quasimonocromàtics proposat.

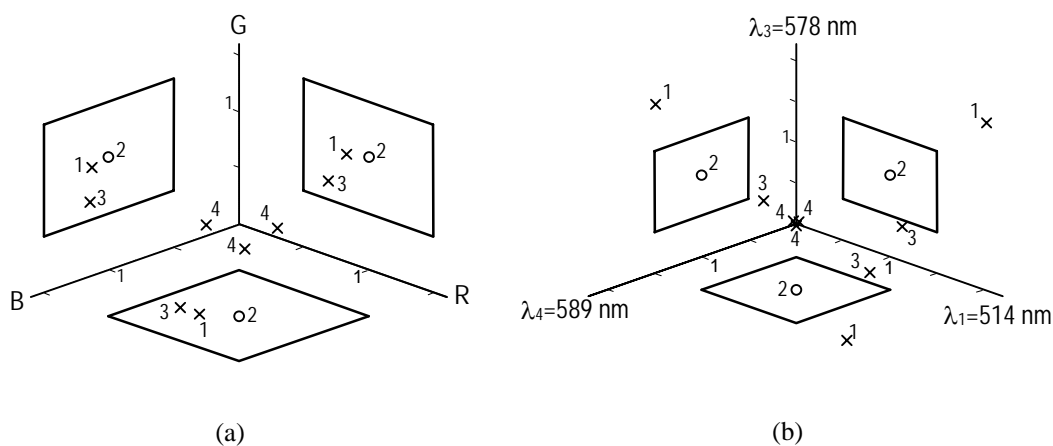
D'acord amb les conclusions derivades en la secció 3.3 per a escenes com les de la Figura 3.3, tots els filtres de potències fraccionàries proporcionarien, almenys teòricament, la mateixa distribució de màxims de correlació normalitzats. En el cas

de l'escena que analitzem ara (Figura 3.3a), utilitzarem el tipus de filtratge adaptat corresponent al filtre només de fase (POF), adaptat a l'objecte 2 a reconèixer. Escollim aquest tipus de filtre d'entre els filtres de potències fraccionàries, per les seves propietats favorables per a la posterior realització experimental. El POF presenta avantatges en quant a una major eficiència lluminosa respecte altres filtres inclosos en els FPF, com ara el CMF o el IF, s'obtenen valors alts del paràmetre PCE que indiquen una gran quantitat d'energia en el pic en comparació de l'energia total del pla de correlació, s'obté també una bona agudesesa del pic que afavoreix la detecció i localització de l'objecte buscat, i finalment, també s'aconsegueix una bona raó senyal-soroll [6].

La Taula 3.3 mostra els resultats numèrics d'intensitats de correlació quan s'aplica un doble llindar del 50% del valor de la intensitat de l'autocorrelació, en ambdós sistemes de canals, l'RGB i els canals quasimonocromàtics ( $\lambda_1=514$  nm,  $\lambda_2=550$  nm,  $\lambda_3=578$  nm i  $\lambda_4=589$  nm) seleccionats en la secció 3.4.1 (Taula 3.1). Donat que les components de l'escena a analitzar són enregistrades experimentalment, pels motius explicats amb anterioritat (secció 3.4.2), les formes dels objectes poden no coincidir exactament, de manera que els valors d'intensitat de correlació normalitzats poden diferir dels estimats teòricament mitjançant el coeficient  $(b_\lambda/a_\lambda)^2$  (secció 3.3). En cada un dels canals, els valors dels pics inclosos en la regió de reconeixement definida pel doble llindar ( $u = 0.5$ ), apareixen subratllats en la Taula 3.3. L'operació lògica AND s'aplica llavors, en els pics subratllats de cada objecte, considerant tots els canals, per obtenir el resultat de reconeixement. En l'espai d'intensitats de correlació, en la representació de la Figura 3.10, s'ha limitat a tres el nombre de canals quasimonocromàtics seleccionats ( $\lambda_1=514$  nm,  $\lambda_3=578$  nm i  $\lambda_4=589$  nm) per facilitar-ne la comparació amb el sistema RGB. Només s'han representat les projeccions a cada un dels hiperplans de manera que la visualització sigui així més fàcil.

**Taula 3.3.** Intensitats de correlació normalitzades per la gamma magenta en el sistema RGB i en els canals seleccionats  $\lambda_1=514$  nm,  $\lambda_2=550$  nm,  $\lambda_3=578$  nm i  $\lambda_4=589$  nm. Escena captada (Figura 3.3a) i simulació de la correlació amb filtre POF. L'objecte 2 es pren com a referència.

RGB	obj 1	obj 2	obj 3	obj 4	Canals seleccionats	obj 1	obj 2	obj 3	obj 4
<b>R</b>	<u>0.83</u>	<u>1.00</u>	<u>0.69</u>	0.30	$\lambda_1 = 514$	2.03	<u>1.00</u>	<u>1.13</u>	0.04
<b>G</b>	<u>0.96</u>	<u>1.00</u>	<u>0.66</u>	0.09	$\lambda_2 = 550$	<u>1.25</u>	<u>1.00</u>	<u>0.53</u>	0.01
<b>B</b>	<u>1.13</u>	<u>1.00</u>	<u>1.14</u>	0.25	$\lambda_3 = 578$	2.07	<u>1.00</u>	0.42	0.04
					$\lambda_4 = 589$	<u>1.49</u>	<u>1.00</u>	0.34	0.03
<b>Reconeix.</b>	SI	SI	SI	NO	<b>Reconeix.</b>	NO	SI	NO	NO



**Figura 3.10.** Resultats de correlació simulada per l'escena test de la gamma magenta (Figura 3.3a): (a) en el sistema RGB i (b) en el sistema adaptat de canals quasimonocromàtics. S'ha adaptat un filtre POF a l'objecte 2, que és la referència, en cada un dels canals. En la figura s'han emprat els símbols: o, per l'autocorrelació; x, per la correlació creuada, i —, per senyalar els límits de la superfície de decisió.

En el sistema RGB apareixen dues falses alarmes pels objectes 1 i 3, ja que els seus punts representatius estan inclosos en la regió de reconeixement (estan dins els tres rectangles de projecció en la Figura 3.10a), mentre que aquestes falses alarmes s'eliminen en el sistema de descomposició multicanal a partir de la selecció dels

canals quasimonocromàtics. No tant sols s'han eliminat les falses alarmes, sinó que també s'aconsegueix reduir el nombre de canals necessari per processar la informació. Per exemple, els canals  $\lambda_1=514$  nm i  $\lambda_4=589$  nm conjuntament, o només el canal  $\lambda_3=578$  nm són suficients per discriminar l'objecte 2 de la resta.

### 3.4.3b Gamma taronja

Presentem en segon lloc els resultats de reconeixement, per correlació simulada a partir de les components cromàtiques obtingudes de manera experimental, per l'escena test pertanyent a la gamma taronja (Figura 3.3b). Recordem que aquesta escena test és especialment interessant pel fet que d'ella extraurem la informació necessària per portar a terme un procés de reconeixement de l'escena de fruites naturals (Figura 3.6). Presentem els resultats per l'escena test quan prenem l'objecte 1 com a referència. S'ha emprat el filtre POF per emmagatzemar la informació necessària de l'objecte patró. En la Taula 3.4 i en Figura 3.11 presentem els resultats numèrics i gràfics en l'espai d'intensitats normalitzades, tant quan s'aplica la descomposició cromàtica en el sistema RGB, com en el sistema proposat de canals quasimonocromàtics ( $\lambda_1=514$  nm i  $\lambda_2=550$  nm) seleccionats en la secció 3.4.1 (Taula 3.2).

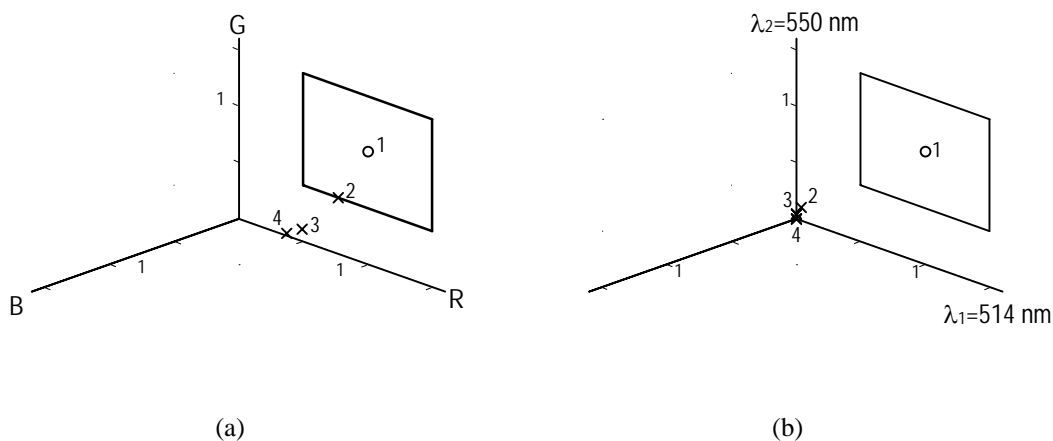
Observem que el canal B del sistema convencional no s'utilitza en el procés de reconeixement donat que les reflectàncies dels quatre objectes de l'escena en aquesta banda de l'espectre són molt baixes i fluctuants, i el soroll de la imatge emmascara el senyal provenint de cada objecte. Per tant, només són útils dos canals, l'R i el G. En aquests dos canals, doncs, veiem que el sistema RGB no és suficient per discriminar l'objecte buscat davant de la resta d'objectes presents a l'escena, donat que, concretament, apareix una falsa alarma per l'objecte 2. En el sistema de canals proposat, en canvi, s'obté una molt bona discriminació cromàtica entre els quatre ocells de paper de tonalitats semblants, i per tant la seva utilització és preferible també en aquest cas davant del sistema RGB. Observem també, que amb el sistema de canals quasimonocromàtics és possible una reducció dels canals de processament



de la informació. Només amb un dels canals, i sobretot considerant el canal  $\lambda_1=514$  nm, és suficient per identificar i discriminar la referència de manera molt relaxada.

**Taula 3.4.** Intensitats de correlació normalitzades per la gamma taronja en el sistema RGB i en els canals seleccionats  $\lambda_1=514$  nm i  $\lambda_2=550$  nm. Escena captada (Figura 3.3b) i simulació de la correlació amb filtre POF. L'objecte 1 es pren com a referència.

RGB	obj 1	obj 2	obj 3	obj 4	Canals seleccionats	obj 1	obj 2	obj 3	obj 4
<b>R</b>	<u>1.00</u>	<u>0.77</u>	0.49	0.37	$\lambda_1 = 514$	<u>1.00</u>	0.04	0.00	0.00
<b>G</b>	<u>1.00</u>	<u>0.50</u>	0.10	0.02	$\lambda_2 = 550$	<u>1.00</u>	0.12	0.04	0.00
<b>B</b>	-	-	-	-					
<b>Reconeix.</b>	SI	SI	NO	NO	<b>Reconeix.</b>	SI	NO	NO	NO



**Figura 3.11.** Resultats de correlació simulada per l'escena test de la gamma taronja (Figura 3.3b): (a) en el sistema RGB i (b) en el sistema adaptat de canals quasimonocromàtics. S'ha adaptat un filtre POF a l'objecte 1, que és la referència, en cada un dels canals. En la figura s'han emprat els símbols: o, per l'autocorrelació; x, per la correlació creuada, i —, per senyalar els límits de la superfície de decisió.

En aquesta memòria s'han escollit els resultats corresponents al reconeixement de l'objecte 2 de la gamma magenta, i el de l'objecte 1 de la gamma taronja, com a exemples més significatius, donat que permeten una major enriquiment de la discussió i els comentaris.

La proposta del sistema adaptat de selecció de canals s'ha aplicat en el reconeixement i discriminació de cada un dels objectes inclosos en la mateixa gamma magenta i taronja (Figura 3.3), davant de la resta d'ocells de paper amb color uniforme dins la mateixa tonalitat [189]. En l'anàlisi d'aquests resultats, s'observa una millora de la capacitat de discriminació cromàtica del sistema de reconeixement de selecció adaptada de canals davant la descomposició multicanal RGB. En la major part de casos, el sistema de canals quasimonocromàtics elimina les falses alarmes que esdevenen amb l'aplicació del sistema convencional RGB i, a més, afegeix uns resultats més relaxats que permeten realitzar el procés de reconeixement amb més comoditat. En general, també es veu la possibilitat de reduir el nombre de canals seleccionats, necessaris per portar a terme el procés de identificació d'un objecte i de discriminació de la resta satisfactòriament, en comparació amb el sistema RGB.

El sistema de canals quasimonocromàtics proposat, també s'ha emprat a altres gammes de colors (marró, verda i blava), seleccionant el conjunt de filtres interferencials apropiats en cada cas (veure annex I). S'ha aplicat el procediment multicanal pel sistema adaptat de canals seleccionats i pel sistema RGB, per a portar a terme el reconeixement de cada un dels objectes inclosos en cada una de les gammes estudiades. Per a totes les gammes, els resultats de correlació simulats prediuen un augment en la capacitat discriminativa cromàtica del sistema i també, pràcticament sempre, s'aconsegueix una reducció en el nombre de canals on es processa la informació de color de l'escena a analitzar, en comparació al sistema convencional dels tres canals RGB. A més, el reconeixement es pot portar a terme en unes condicions més relaxades donat que la distància relativa entre els pics de correlació creuada i d'autocorrelació en un mateix canal és major en el cas d'utilitzar el sistema de descomposició adaptada en canals quasimonocromàtics [189].

### 3.4.3c Escena d'objectes naturals (cítrics)

Per últim presentem els resultats per l'escena d'objectes reals (cítrics) de la Figura 3.6. Considerem els canals  $\lambda_1=514$  nm i  $\lambda_2=550$  nm seleccionats a partir de les corbes de reflectància espectral dels colors inclosos en l'escena test de la gamma taronja. En el cas dels objectes artificials, acabem de posar de manifest que aquests canals donen una capacitat de discriminació suficient en els resultats de simulació de la correlació. En l'escena dels cítrics, l'objecte 1 és el patró buscat. En aquest cas també es pot utilitzar el filtre POF com a mètode de filtratge adaptat a l'objecte a reconèixer, com en els casos anteriors. De totes maneres, per verificar que la nostra proposta pot millorar la discriminació cromàtica del sistema, utilitzarem els filtres CMF, per la seva major tolerància, comparat amb els filtres POF, a petites diferències en els contorns dels objectes naturals. D'aquesta manera assegurem que la discriminació aconseguida pel sistema estigui basada principalment en les diferències de color dels objectes, i no tant en les diferències en forma que no són especialment interessants en aquest exemple.

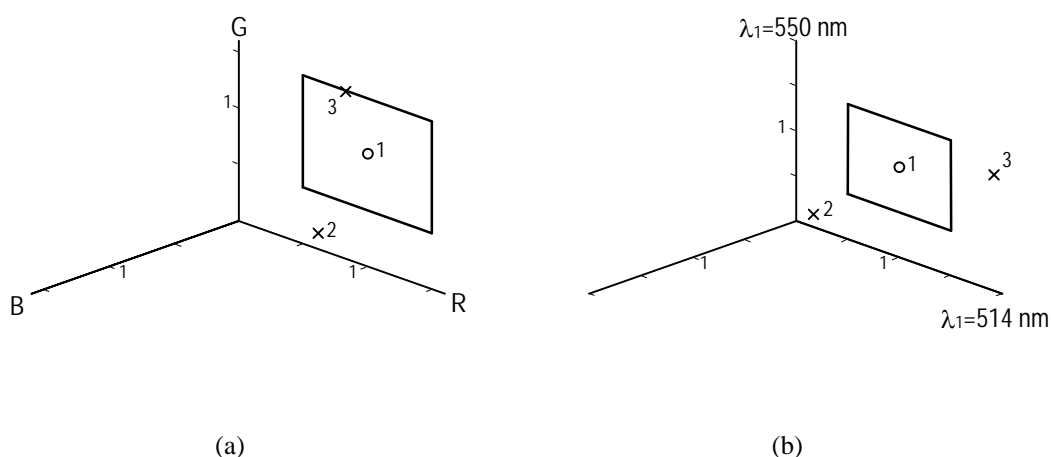
La Taula 3.5 mostra els resultats numèrics en l'espai d'intensitats normalitzades obtinguts en el cas d'emprar el filtre CMF, considerant la descomposició RGB i la de canals adaptats quasimonocromàtics. Els resultats subratllats indiquen els màxims de correlació inclosos dins la regió delimitada pel doble llindar ( $u = 0.5$ ). Aplicant l'operació lògica AND entre ells, obtenim els objectes que es reconeixen en el procés global. La representació gràfica dels resultats de correlació es mostren en la Figura 3.12. En ambdues representacions, observem per exemple, que en el sistema RGB, on novament el canal cromàtic B no aporta informació significativa, es reconeix el patró buscat, i a més apareix una falsa alarma en el cas de la llimona de color groc (objecte 3). En el sistema de canals proposat, s'elimina aquesta falsa alarma i amb el canal  $\lambda_1=514$  nm el reconeixement es porta a terme de manera relaxada.

Cal comentar en aquest punt, que la utilització d'objectes en l'escena test corresponents a la gamma taronja (Figura 3.3b) ens ha ajudat a trobar una regió de l'espectre on les tonalitats a discriminar tenien valors de reflectància més diferents, i

per tant, eren més fàcilment diferenciables. Però, els resultats així obtinguts per l'escena test, són només orientatius en el moment de passar a imatges amb objectes reals donat que mai apareixeran exactament les mateixes tonalitats en els dos casos. En els objectes reals fotografiats apareixen reflexes degut a la il·luminació utilitzada, i tenen petits detalls que fan que el color no sigui uniforme en tota la superfície. Per exemple, els reflexes de la primera mandarina (la referència) fan que en conjunt tingui una tonalitat taronja més clara, i per aquest motiu s'arriba a confondre amb la llimona en el canal de  $\lambda_2=550$  nm.

**Taula 3.5.** Intensitats de correlació normalitzades per l'escena de cítrics en el sistema RGB i en els canals seleccionats per la gamma taronja. Escena captada (Figura 3.6) i simulació de la correlació amb filtre CMF. L'objecte 1 es pren com a referència.

RGB	obj 1	obj 2	obj 3	Canals seleccionats	Obj 1	obj 2	obj 3
<b>R</b>	<u>1.00</u>	<u>0.62</u>	<u>0.83</u>	$\lambda_1 = 514$	<u>1.00</u>	0.17	1.92
<b>G</b>	<u>1.00</u>	0.14	<u>1.49</u>	$\lambda_2 = 550$	<u>1.00</u>	0.14	<u>1.29</u>
<b>B</b>	-	-	-				
<b>Reconeix.</b>	SI	NO	SI	<b>Reconeix.</b>	SI	NO	NO



**Figura 3.12.** Resultats de correlació simulada per l'escena d'objectes naturals (cítrics) (Figura 3.6): (a) en el sistema RGB i (b) en el sistema adaptat de canals quasimonocromàtics. S'ha adaptat un filtre CMF a l'objecte 1, que és la referència, en cada un dels canals. En la figura s'han emprat els símbols: o, per l'autocorrelació; x, per la correlació creuada, i —, per senyalar els límits de la superfície de decisió.

### 3.5 Resultats experimentals de reconeixement per correlació òptica

En aquest apartat descriurem la realització, seguint un procés completament experimental, del reconeixement d'objectes en color per correlació òptica multicanal utilitzant, per una banda, la descomposició en canals quasimonocromàtics seleccionats que s'ha proposat en aquest capítol i, per l'altra banda, la descomposició RGB. També s'intentarà corroborar els resultats de la correlació obtinguts per simulació numèrica en l'apartat anterior.

Haurem de tenir en compte, però, dos fets que intervenen en la realització experimental. En primer lloc, la generació del filtre adaptat i, en segon lloc, la posterior reproducció de les escenes i dels filtres adaptats en suports que puguin ser utilitzats en el correlador.

Els filtres adaptats són, en general, funcions complexes amb la informació tant de l'amplitud com de la fase de la transformada de Fourier de l'objecte a reconèixer. Existeixen diferents formes de codificar aquesta funció digitalment, en hologrames generats per ordinador, de manera que la informació completa de la funció s'emmagatzema dins de cada cel·la amb una representació binària, ja sigui mitjançant rectangles desplaçats dins una cel·la (en els hologrames de fase per desviació [192, 193]), o bé binaritzant unes franges d'interferència (en els mètodes interferomètrics [194, 195]). La codificació utilitzada en aquest treball és la de Lee [194, 195], que representa les franges d'interferència produïdes per holografia clàssica entre la transformada de Fourier del senyal a detectar i una ona de referència inclinada.

Per a la reproducció de les escenes i els filtres adaptats generats per ordinador sobre un suport material que permeti que siguin utilitzats en un banc òptic, existeixen diverses tècniques: impressora làser combinada amb fotoreducció posterior, impressora linotrònic, litografia, etc. Basant-nos en treballs que analitzen les propietats dels diferents dispositius [196], tenint en compte factors com la resolució, el temps d'impressió, el cost econòmic i els resultats en correlacions, hem escollit la impressora Linotrònic 630. Aquesta impressora s'utilitza habitualment en arts gràfiques i té una alta resolució (3251 punts per polzada) que permet realitzar els

hologrames directament sobre fotolit amb el tamany adequat per a la seva utilització en el banc òptic, evitant la pèrdua de resolució en processos de fotoreducció i, sense introduir defectes de revelats posteriors.

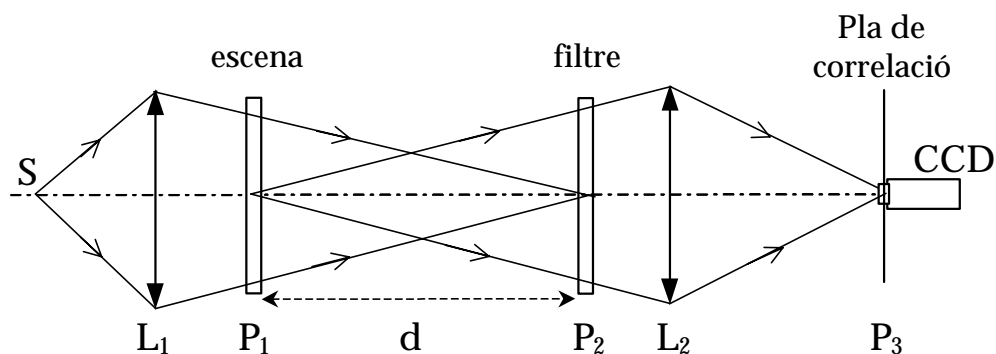
Les escenes i els filtres reproduïts estan formats per cel·les de 10x10 punts, amb un tamany aproximat per tota l'escena o filtre d'1 cm<sup>2</sup>. Amb aquestes dimensions es poden utilitzar directament sobre un banc òptic. Ambdues funcions s'imprimeixen sobre fotolit, un material transparent que es situa enmig de dues plaques de vidre per donar-li rigidesa. Per compensar les diferències de camí òptic introduïdes per les irregularitats superficials dels diferents medis, s'utilitza oli d'immersió de microscopia. Aquesta tècnica dona bons resultats de correlacions i té altres avantatges com són un reduït temps d'impressió, una fàcil accessibilitat a aquests tipus d'impressores i un baix cost econòmic.

### 3.5.1 Correlador òptic convergent

El muntatge més adient per realitzar la correlació òptica en el cas d'haver d'ajustar l'escala de la transformada de Fourier òptica de l'escena amb l'escala de la TF de l'objecte a detectar enregistrada en el filtre adaptat, és l'anomenat correlador convergent [125] (Figura 3.13) descrit en l'apartat 2.2.2 del capítol 2.

En la Figura 3.14 es presenta una visió general del correlador convergent experimental, i les Figures 3.15 i 3.16 mostren detalls del mateix correlador, tot indicant els elements que en formen part per tal de poder identificar-lo amb l'esquema gràfic anterior de la Figura 3.13.

Un test de franges fosques i clares (xarxa de difracció) i la seva TF digital enregistrada, ens facilita l'ajust dels elements del correlador ja que fem coincidir la transformada de Fourier òptica del test de franges situat en el pla P<sub>1</sub> del correlador, amb la seva transformada digital que prèviament hem reproduït sobre fotolit i hem situat en P<sub>2</sub>.



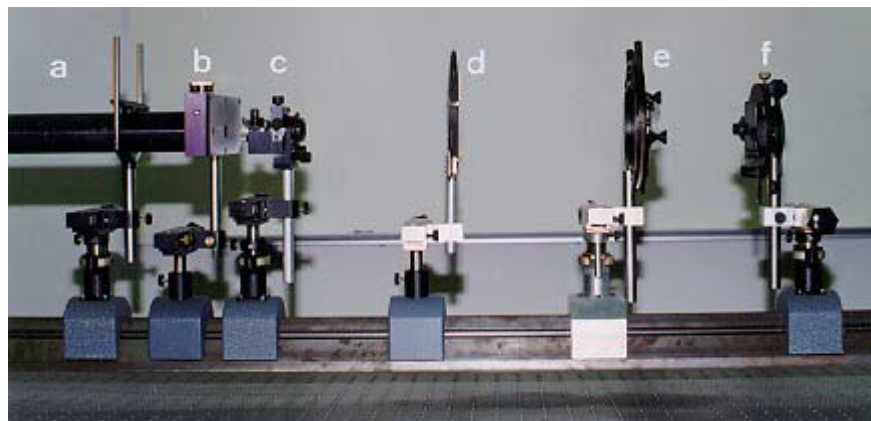
**Figura 3.13.** Esquema del correlador convergent.



**Figura 3.14.** Dispositiu experimental del correlador convergent.

Les correlacions experimentals s'han obtingut situant, en els plans corresponents, una escena i un filtre adaptat a l'objecte a reconèixer pertanyents a un mateix canal. El procés s'ha repetit per tots els canals necessaris. Totes les correlacions experimentals s'han fet a partir de les escenes captades amb les càmeres CCD (càmera en color pel sistema RGB i càmera en b/n amb filtre interferencial pel sistema de canals quasimonocromàtics seleccionats) de la manera descrita en els apartats 3.4.2a i 3.4.2b. El pla de correlació es capta amb la càmera Pulnix en b/n, fent coincidir el detector CCD amb el pla de sortida del correlador òptic. Sobre cada una de les distribucions d'intensitat de correlació enregistrades, caldrà fer les mesures dels valors màxims d'intensitat que apareixen, i ens interessarà sobretot conèixer els valors relatius entre ells. Donat que volem conèixer els valors relatius

entre els màxims mesurats, és important assegurar que la càmera CCD treballa en la seva zona de resposta lineal. Per la calibració feta anteriorment (apartat 3.4.2b), sabem que ens acostem a aquestes condicions si prenem el  $gain = 0$  i l' $offset = 70$  (Figura 3.9). A més, coneguda la corba de resposta, podem corregir els valors donats per la càmera tenint en compte una taula de correspondència o LUT.



**Figura 3.15.** Detall del dispositiu experimental (I). S'observa el sistema d'il·luminació amb feix làser (a) expandit i filtrat espacialment pel *pinhole* (c), l'atenuador (b), un diafragma (d), la lent  $L_1$  (e) i el pla d'entrada  $P_1$  del correlador on es situa l'escena (f).



**Figura 3.16.** Detall del dispositiu experimental (II): pla de Fourier  $P_2$  (g), la lent  $L_2$  (h) i el pla de correlació (i) del correlador convergent experimental. En (g) es situa el filtre adaptat a l'objecte a reconèixer. El pla del sensor CCD de la càmera en b/n coincideix amb el pla de correlació.



La font de llum monocromàtica que il·lumina el correlador és un làser d'He-Ne de 8 mW. La Figura 3.15 permet observar l'atenuador situat just davant del làser i abans d'expandir el seu feix amb el filtre espacial (*pinhole*). L'atenuador utilitzat és neutre i està format per una doble falca que ens permet modificar la intensitat del feix de manera que controlem la quantitat de llum que incideix sobre el detector CCD de la càmera situat en el pla de correlació. Així doncs, podem tenir cura de què els màxims de correlació obtinguts no superin el valor de saturació del detector i, en cada cas, podem controlar que les correlacions mesurades estiguin dins de la zona de resposta lineal de la càmera.

### 3.5.2 Resultats de reconeixement experimentals

#### 3.5.2a Gamma magenta

Primerament, mostrem els resultats de reconeixement per l'escena test de la gamma magenta (Figura 3.3a). Recordem que l'objecte a reconèixer és l'objecte 2 i es fa servir un filtre POF. La Figura 3.17 mostra el pla de sortida del correlador en el canal  $\lambda_1=514$  nm captat amb la càmera CCD b/n, on s'ha enregistrat tant la reconstrucció de l'escena (centre), com els termes fora d'eix, que corresponen a la correlació (dreta) i la convolució de les mateixes funcions (esquerra). La baixa energia de l'objecte 4 en el canal  $\lambda_1$  fa que no s'hagi pogut enregistrar adequadament en la Figura 3.17.



**Figura 3.17.** Imatge enregistrada amb la càmera CCD b/n del pla de sortida del correlador quan es realitza el reconeixement de l'objecte 2 de l'escena magenta utilitzant un filtre POF en el canal  $\lambda_1 = 514$  nm. En el centre s'observa la reconstrucció de l'escena, i els termes fora d'eix consisteixen en la correlació de l'escena per l'objecte de referència (dreta), i la convolució entre les mateixes funcions (esquerra).

A continuació, es presenten els resultats de la gamma magenta, quan es vol reconèixer l'objecte 2, conjuntament pels canals  $\lambda_1 = 514$  nm,  $\lambda_3 = 578$  nm i  $\lambda_4 = 589$  nm, i pels canals RGB. Tots aquests resultats es representen, en una taula numèrica (Taula 3.6), en l'espai de les intensitats de correlació (Figura 3.18), i en gràfics tridimensionals on s'observa la distribució de màxims d'intensitat obtinguda en el pla de correlació captat amb la CCD. En aquesta última representació tridimensional, la Figura 3.19 mostra els plans de correlació obtinguts amb la descomposició RGB, i la Figura 3.20 en els canals quasimonocromàtics proposats. En ambdós casos, al costat de cada pic s'indica el valor relatiu en referència a l'autocorrelació de l'objecte 2, i entre parèntesis es mostra el valor esperat segons la simulació de la correlació presentada en la Taula 3.3. Tant en la taula numèrica (Taula 3.6) com en la representació tridimensional dels plans de correlació (Figures 3.19 i 3.20), es subratllen els valors dels màxims de correlació normalitzats que es situen dins la banda de reconeixement definida pel doble llinzar ( $u = 0.5$ ). El procés de reconeixement es completa, aplicant l'operació lògica AND als resultats subratllats de cada objecte en tots els canals. Els resultats del sistema de selecció de canals quasimonocromàtics es poden comparar en totes les representacions amb els obtinguts en el sistema RGB.

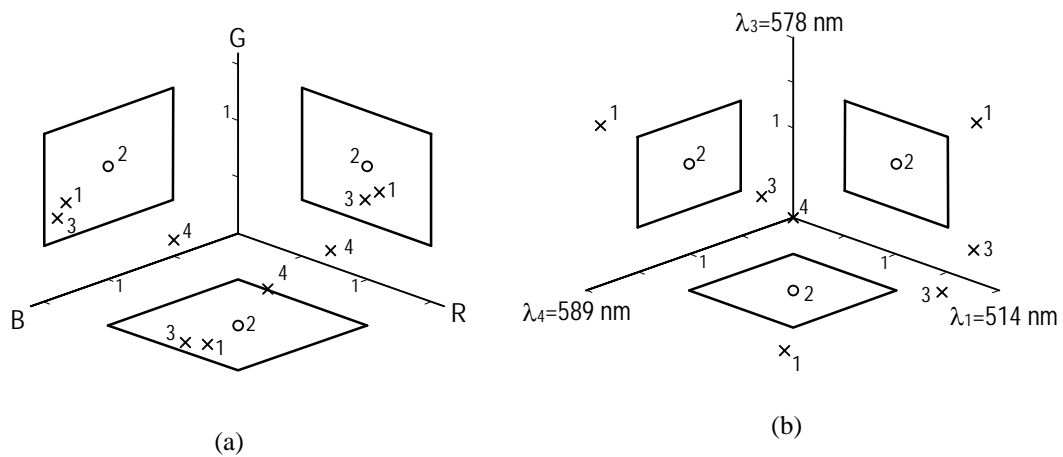
El canal  $\lambda_2 = 550$  nm, inicialment seleccionat per a la gamma magenta, no s'ha tingut en compte en aquesta realització experimental donat que els resultats de simulació han mostrat que la capacitat de discriminació del sistema en aquest canal és significativament més baixa, en comparació a la resta de canals ( $\lambda_1 = 514$  nm,  $\lambda_3 = 578$  nm i  $\lambda_4 = 589$  nm).

En qualsevol de les representacions es pot remarcar el fet que en el sistema convencional RGB, apareixen dues falses alarmes pel què fa als objectes 1 i 3, que no poden ser discriminats de l'objecte 2, pres com a referència a detectar. En canvi, el sistema adaptat de canals seleccionats porta a terme el procés de reconeixement sense dificultat ja que s'aconsegueix discriminar l'objecte patró de la resta d'ocells de paper inclosos en la mateixa escena. Observem, a més, que en el sistema de canals quasimonocromàtics, amb un únic canal d'entre qualsevol dels tres escollits,  $\lambda_1 = 514$

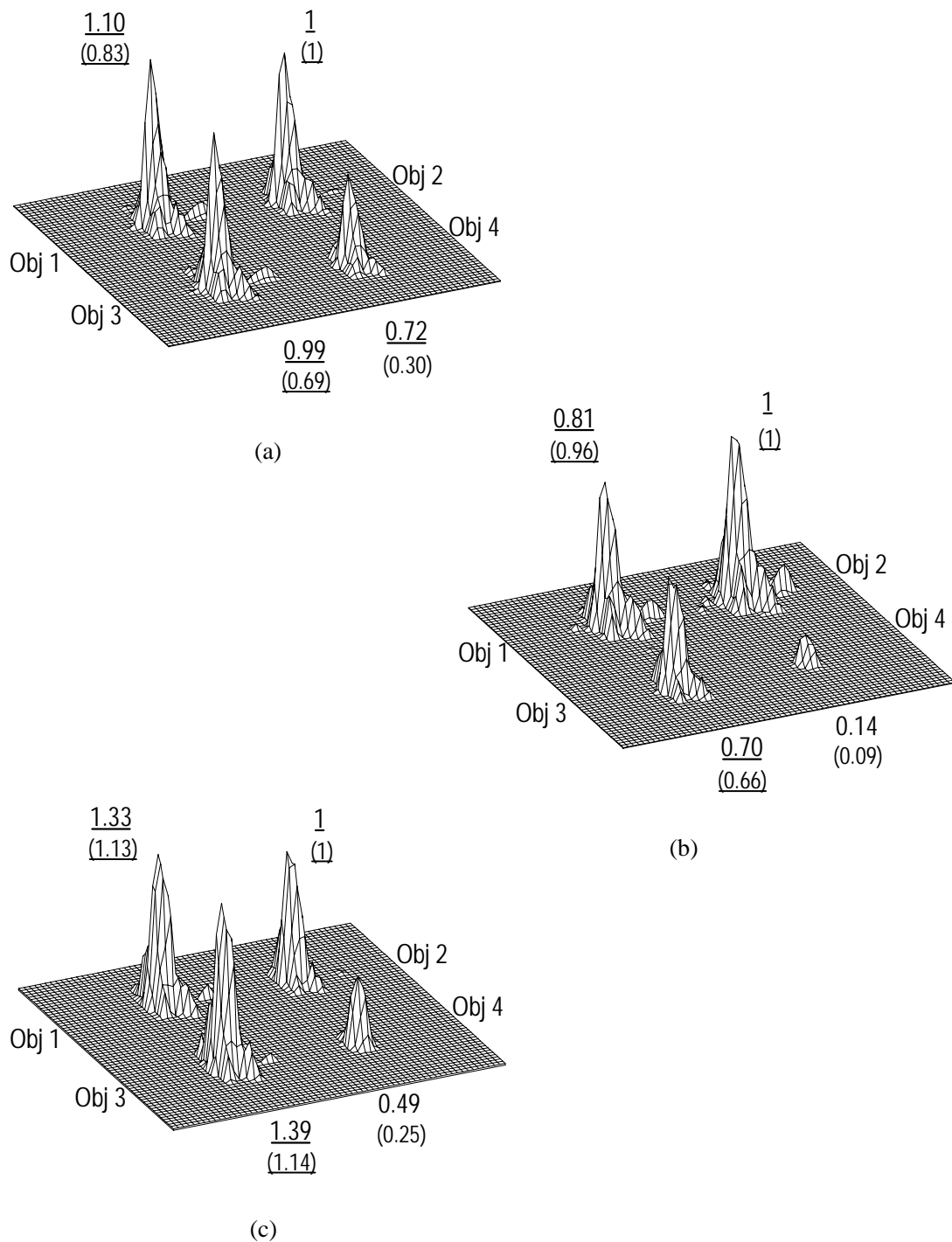
nm,  $\lambda_2=578$  nm i  $\lambda_3=589$  nm, és possible obtenir bons resultats en el procés de reconeixement.

**Taula 3.6.** Intensitats experimentals de correlació normalitzades per la gamma magenta en el sistema RGB i en els canals seleccionats. Escena captada (Figura 3.3a) i correlació experimental amb filtre POF. L'objecte 2 es pren com a referència.

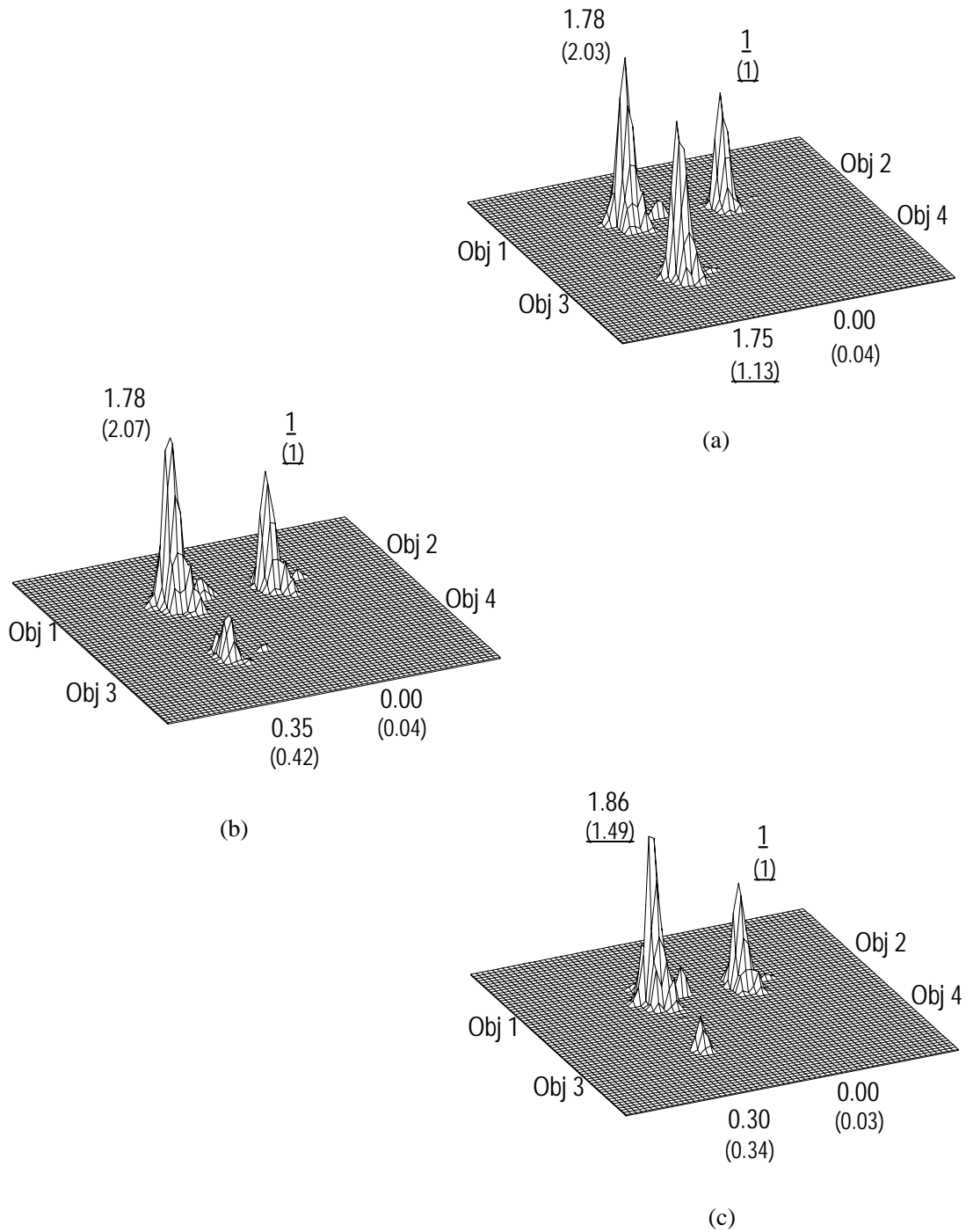
RGB	obj 1	obj 2	obj 3	obj 4	Canals seleccionats	obj 1	obj 2	obj 3	obj 4
<b>R</b>	<u>1.10</u>	<u>1.00</u>	<u>0.99</u>	<u>0.72</u>	$\lambda_1 = 514$	1.78	<u>1.00</u>	1.75	0.00
<b>G</b>	<u>0.81</u>	<u>1.00</u>	<u>0.70</u>	0.14	$\lambda_3 = 578$	1.78	<u>1.00</u>	0.35	0.00
<b>B</b>	<u>1.33</u>	<u>1.00</u>	<u>1.39</u>	0.49	$\lambda_4 = 589$	1.86	<u>1.00</u>	0.30	0.00
<b>Reconeix.</b>	SI	SI	SI	NO	<b>Reconeix.</b>	NO	SI	NO	NO



**Figura 3.18.** Resultats de correlació experimental per l'escena test de la gamma magenta (Figura 3.3a): (a) en el sistema RGB i (b) en el sistema adaptat de canals quasimonocromàtics. S'ha adaptat un filtre POF a l'objecte 2, que és la referència, en cada un dels canals. En la figura s'han emprat els símbols: o, per l'autocorrelació; x, per la correlació creuada, i —, per senyalar els límits de la superfície de decisió.



**Figura 3.19.** Representació en 3D del pla de correlació en el cas de reconèixer l'objecte 2 de la gamma magenta (Figura 3.3a) utilitzant filtres POF en els canals (a) R, (b) G i (c) B. Per a cada pic s'indiquen els valors relatius de la correlació creuada respecte l'autocorrelació obtinguts experimentalment, i els simulats (Taula 3.3) entre parèntesis.



**Figura 3.20.** Representació en 3D del pla de correlació en el cas de reconèixer l'objecte 2 de la gamma magenta (Figura 3.3a) utilitzant filtres POF en els canals (a)  $\lambda_1 = 514 \text{ nm}$ , (b)  $\lambda_3 = 578 \text{ nm}$  i (c)  $\lambda_1 = 589 \text{ nm}$ . Per a cada pic s'indiquen els valors relatius de la correlació creuada respecte l'autocorrelació obtinguts experimentalment, i els simulats (Taula 3.3) entre parèntesis.

La Taula 3.6 i la Figura 3.18, són representacions anàlogues a la Taula 3.3 i Figura 3.10, mostrades en l'apartat 3.4.3a pertanyents als resultats per les correlacions simulades. Comparant aquestes representacions, veiem que es confirma la millora de la capacitat discriminativa obtinguda mitjançant la proposta de canals seleccionats davant del sistema RGB convencional. Cal senyalar també, la bona concordància dels resultats experimentals amb els predits per les proves de simulació.

S'han realitzat també les proves i s'han obtingut els resultats experimentals de reconeixement en el cas en què l'objecte a detectar sigui qualsevol dels objectes presents a l'escena (és a dir, l'objecte 1, 3 i 4) [189]. En tots els casos els resultats de reconeixement pel sistema de selecció de canals adaptats demostren una millora en la capacitat de discriminació del sistema, l'eliminació de les falses alarmes que apareixen amb el sistema RGB, l'assoliment d'un procés de reconeixement més relaxat i, en general, una possible reducció del nombre de canals per portar a terme el procés de reconeixement. Els resultats obtinguts experimentalment al voler reconèixer qualsevol dels objectes pertanyents a la gamma magenta, s'adiuen amb els predits per simulació numèrica de la correlació.

### **3.5.2b Gamma taronja**

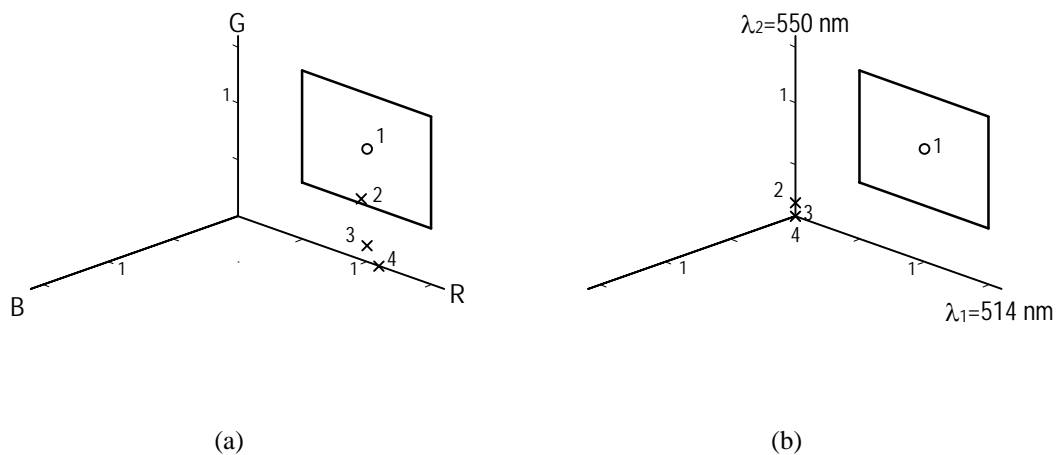
En aquesta gamma ens centrarem en el reconeixement de l'objecte 1 en l'escena de la Figura 3.3b. La Taula 3.7 conté els resultats experimentals de correlació quan s'utilitza un filtre POF. La Figura 3.21 conté la representació gràfica d'aquests mateixos resultats, que també es mostren en els gràfics tridimensionals dels plans de correlació experimentals de les Figures 3.22 i 3.23.

El sistema RGB confon els dos primers objectes i per tant la seva capacitat de discriminació en aquest cas no és acceptable. El canal R té pràcticament la mateixa quantitat d'energia procedent de cada un dels ocells de paper, i en el canal G no es diferencien suficientment els valors de reflectància dels objectes 1 i 2.

D'altra banda, en el sistema de descomposició en canals seleccionats, les diferències relatives entre els màxims de correlació són tant grans en el canal  $\lambda_1 = 514$  nm, que quan es rep senyal dels objectes 2, 3 i 4 de l'escena, el pic de correlació del primer objecte apareix saturat. Aquest fet també és apreciable en el canal  $\lambda_2 = 550$  nm, si bé en menys mesura. Cal dir, que en qualsevol dels dos canals proposats, el reconeixement del patró es porta a terme sense cap dificultat.

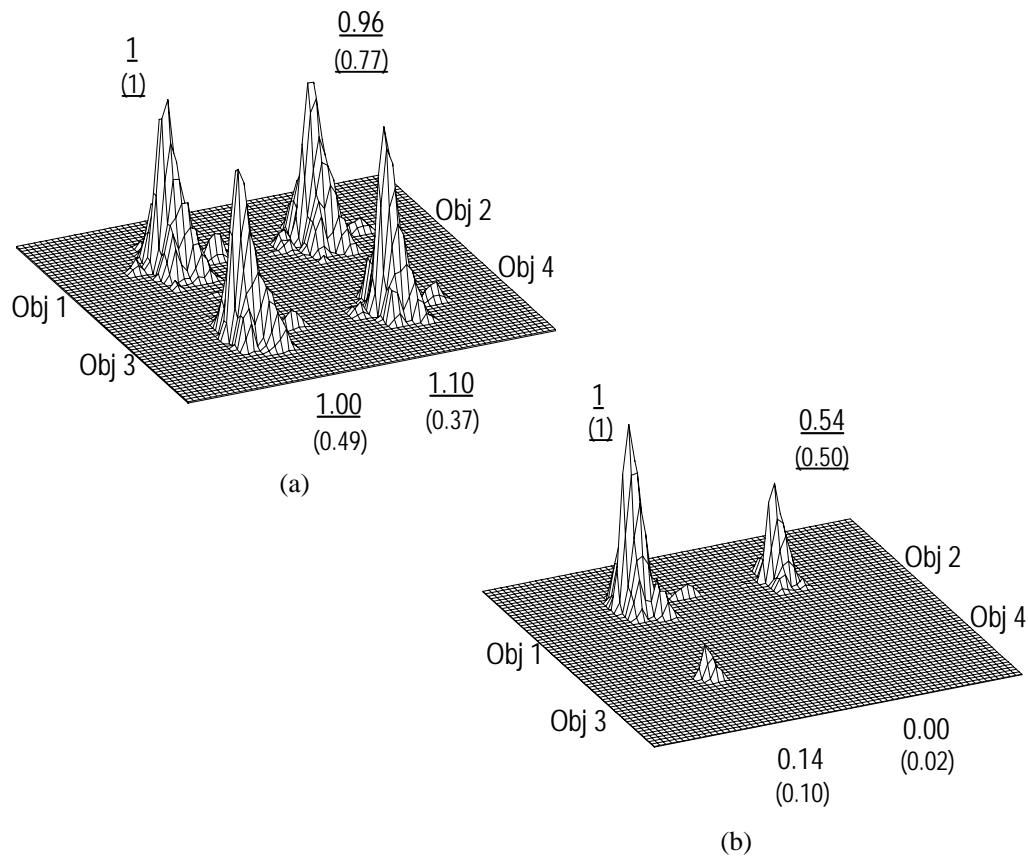
**Taula 3.7.** Intensitats experimentals de correlació normalitzades per la gamma taronja en el sistema RGB i en els canals seleccionats  $\lambda_1=514$  nm i  $\lambda_2=550$  nm. Escena captada (Figura 3.3b) i correlació experimental amb filtre POF. L'objecte a reconèixer és l'objecte 1.

RGB	obj 1	obj 2	obj 3	obj 4	Canals seleccionats	obj 1	obj 2	obj 3	obj 4
<b>R</b>	<u>1</u>	<u>0.96</u>	<u>1.00</u>	<u>1.10</u>	$\lambda_1 = 514$	<u>1.00</u>	0.00	0.00	0.00
<b>G</b>	<u>1</u>	<u>0.54</u>	0.14	0.00	$\lambda_2 = 550$	<u>1.00</u>	0.12	0.00	0.00
<b>B</b>	-	-	-	-					
<b>Reconeix.</b>	SI	SI	NO	NO	<b>Reconeix.</b>	SI	NO	NO	NO



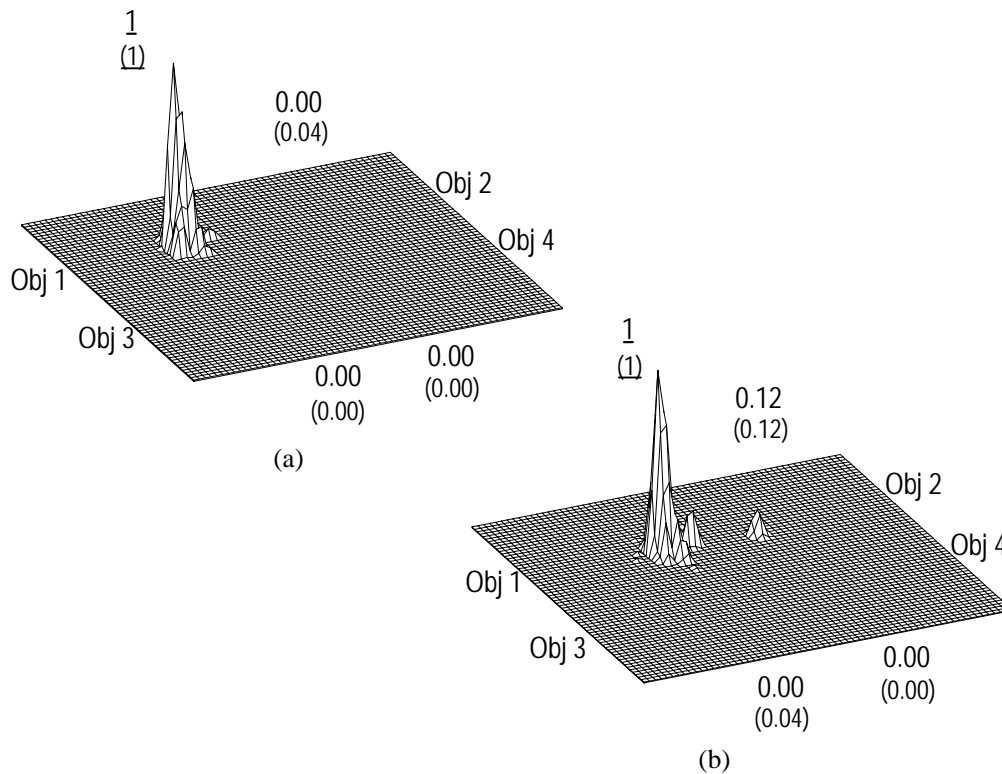
**Figura 3.21.** Resultats de correlació experimentals per l'escena test de la gamma taronja (Figura 3.3b): (a) en el sistema RGB i (b) en el sistema adaptat de canals quasimonocromàtics. S'ha adaptat un filtre POF a l'objecte 1, que és la referència, en cada un dels canals. En la figura s'han emprat els símbols: o, per l'autocorrelació; x, per la correlació creuada, i —, per senyalar els límits de la superfície de decisió.

Existeix en general, un bon acord entre els resultats presentats en l'apartat de simulació (Taula 3.4) i els donats en aquesta darrera taula (Taula 3.7), arribant en tots els casos d'utilització del sistema de descomposició en canals quasimonocromàtics seleccionats, a l'assoliment d'una millora de la capacitat discriminativa del sistema de reconeixement. La comparació entre resultats també és possible fer-la a partir de la representació gràfica en l'espai d'intensitats de correlació, entre la Figura 3.11 pels resultats de simulació i la Figura 3.21 pel que fa als resultats experimentals.



**Figura 3.22.** Representació en 3D del pla de correlació en el cas de reconèixer l'objecte 1 de la gamma taronja (Figura 3.3b) utilitzant filtres POF en els canals (a) R i (b) G. Per a cada pic s'indiquen els valors relatius de la correlació creuada respecte l'autocorrelació obtinguts experimentalment, i els simulats (Taula 3.4) entre parèntesis.





**Figura 3.23.** Representació en 3D del pla de correlació en el cas de reconèixer l'objecte 1 de la gamma taronja (Figura 3.3b) utilitzant filtres POF en els canals (a)  $\lambda_1 = 514$  nm i (b)  $\lambda_2 = 550$  nm. Per a cada pic s'indiquen els valors relatius de la correlació creuada respecte l'autocorrelació obtinguts experimentalment, i els simulats (Taula 3.4) entre parèntesis.

### 3.5.2c Escena d'objectes naturals (cítrics)

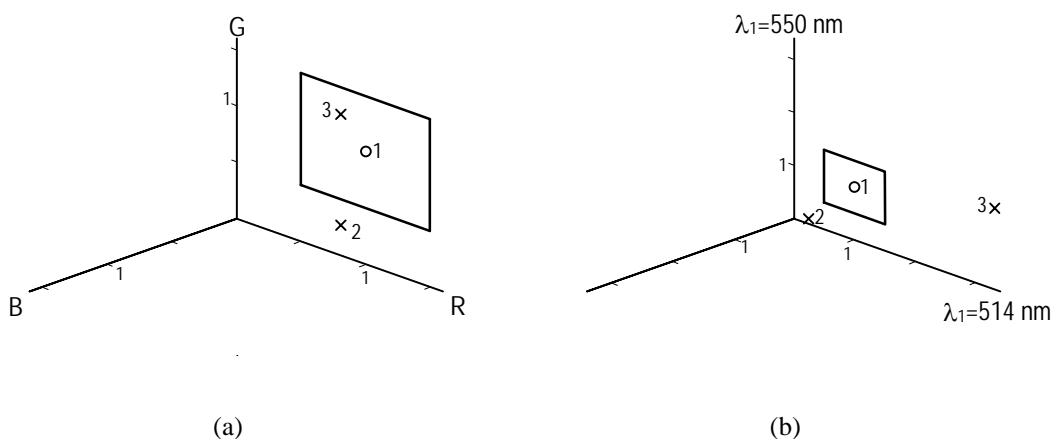
Recordem que en aquest cas, hem optat per la realització de la correlació amb el filtre CMF, per tal de poder assegurar que la discriminació es fa a partir de les diferències de color dels objectes, i no pas per petites variacions en la forma o detalls de les fruites naturals, a les quals és més sensible el filtre POF.

Per aquesta escena (Figura 3.6) ens proposem reconèixer l'objecte 1 (mandarina de la varietat *Satsuma Owari*) i discriminar-lo de la resta de cítrics de tonalitats diferents. Emprem les components escollides per l'escena test de la gamma taronja, és a dir, les corresponents als canals  $\lambda_1 = 514$  nm i  $\lambda_2 = 550$  nm. La Taula 3.8 presenta els resultats numèrics de la intensitat de correlació experimental en el sistema de canals RGB i en el sistema de canals quasimonocromàtics. La Figura 3.24 mostra aquests

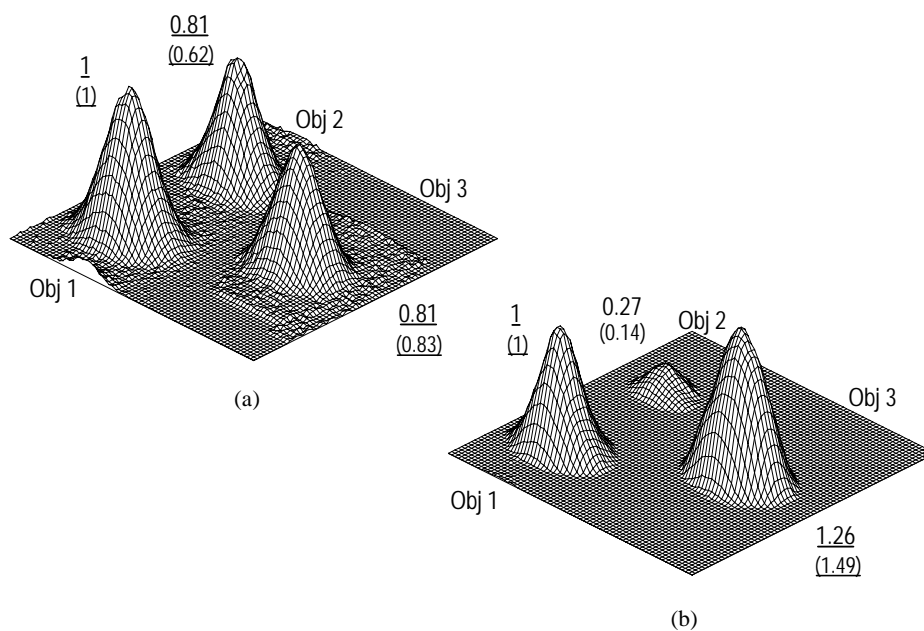
resultats gràficament i les representacions 3.25 i 3.26 són gràfics en tres dimensions dels pics de correlació obtinguts.

**Taula 3.8.** Intensitats experimentals de correlació normalitzades per l'escena de cítrics en el sistema RGB i en els canals seleccionats. Escena captada (Figura 3.6) i correlació experimental amb filtre CMF. L'objecte a reconèixer és l'objecte 1.

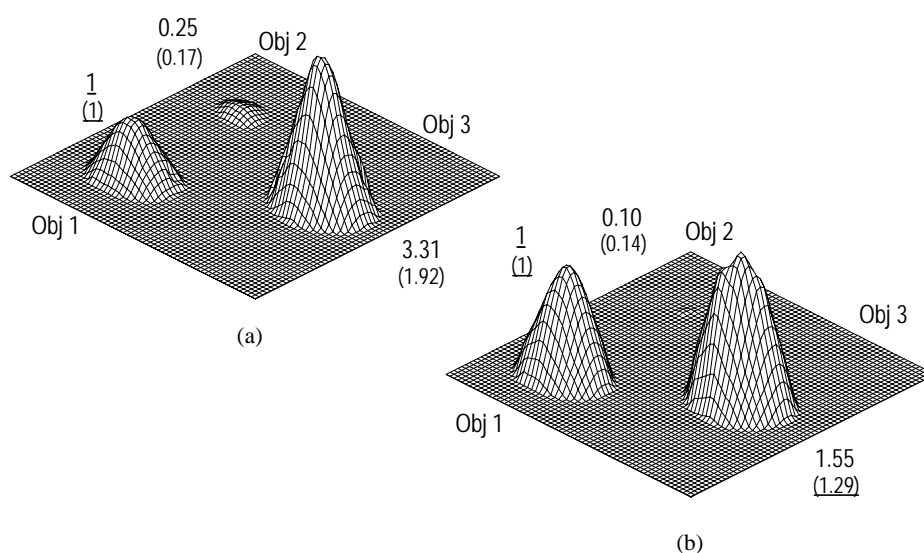
RGB	obj 1	obj 2	obj 3	Canals seleccionats	obj 1	obj 2	obj 3
<b>R</b>	<u>1.00</u>	<u>0.81</u>	<u>0.81</u>	$\lambda_1 = 514$	<u>1.00</u>	0.25	3.31
<b>G</b>	<u>1.00</u>	0.27	<u>1.26</u>	$\lambda_2 = 550$	<u>1.00</u>	0.10	1.55
<b>B</b>	-	-	-				
<b>Reconeix.</b>	SI	NO	SI	<b>Reconeix.</b>	SI	NO	NO



**Figura 3.24.** Resultats de correlació experimentals per l'escena d'objectes naturals (cítrics) (Figura 3.6): (a) en el sistema RGB i (b) en el sistema adaptat de canals quasimonocromàtics. S'ha adaptat un filtre CMF a l'objecte 1, que és la referència, en cada un dels canals. En la figura s'han emprat els símbols: o, per l'autocorrelació; x, per la correlació creuada, i —, per senyalar els límits de la superfície de decisió.



**Figura 3.25.** Representació en 3D del pla de correlació en el cas de reconèixer l'objecte 1 de l'escena de cítrics (Figura 3.6) utilitzant filtres CMF en els canals (a) R i (b) G. Per a cada pic s'indiquen els valors relatius de la correlació creuada respecte l'autocorrelació obtinguts experimentalment, i els simulats (Taula 3.5) entre parèntesis.



**Figura 3.26.** Representació en 3D del pla de correlació en el cas de reconèixer l'objecte 1 de l'escena de cítrics (Figura 3.6) utilitzant filtres CMF en els canals (a)  $\lambda_1 = 514$  nm i (b)  $\lambda_2 = 550$  nm. Per a cada pic s'indiquen els valors relatius de la correlació creuada respecte l'autocorrelació obtinguts experimentalment, i els simulats (Taula 3.5) entre parèntesis.

En les representacions mostrades es pot veure com el sistema RGB es incapaç de diferenciar la mandarina *Satsuma Owari* (objecte de referència) de la llimona *Eureka* (objecte 3), ja que tant pel canal R com pel G, els pics de correlació estan inclosos en la banda delimitada pel doble llinard. Si comparem els resultats experimentals obtinguts, amb els prèviament mostrats en la Taula 3.5, que corresponen a les proves de simulació, observem que la falsa alarma que es dona per l'objecte 3, és encara més inevitable en els resultats experimentals, ja que existeixen factors com la introducció del soroll, la perfecta alineació dels components, etc. que fan la realització experimental més difícil del previst. En referència als resultats obtinguts mitjançant el sistema adaptat de canals seleccionats, veiem que el reconeixement de l'objecte patró es porta a terme de manera relaxada, ja que les diferències amb els pics de correlació creuada dels altres objectes de l'escena, són suficientment grans per a ésser discriminat. Només amb un dels dos canals seleccionats, ja sigui  $\lambda_1 = 514$  nm o  $\lambda_2 = 550$  nm, es pot realitzar sense dificultat el procés de reconeixement plantejat. En tots els casos es posa de manifest el bon acord entre els resultats experimentals i els de simulació numèrica (Taula 3.5 i Figura 3.12).

### 3.6 Reconeixement d'objectes policromàtics

En el desenvolupament d'aquest capítol, hem mostrat principal interès en el mètode de descomposició adaptada de la informació del color en  $n$  canals quasimonocromàtics. Aquest mètode s'ha aplicat al reconeixement d'objectes amb la característica de tenir una distribució uniforme del color sobre el conjunt de la seva superfície. Si bé l'anàlisi realitzada ha estat molt útil per a l'obtenció de bons resultats de reconeixement al voler identificar, en condicions difícils, objectes naturals, ens proposem ara, un pas més endavant en l'estudi de l'abast d'aquesta proposta. Considerem ara objectes amb una distribució de colors variada en la seva superfície, és a dir, el que anomenem objectes policromàtics. Aquesta situació també és possible trobar-la en el cas de la identificació de fruites, com a tall d'exemple d'altres aplicacions afins. Existeixen fruites amb variació de color en la seva

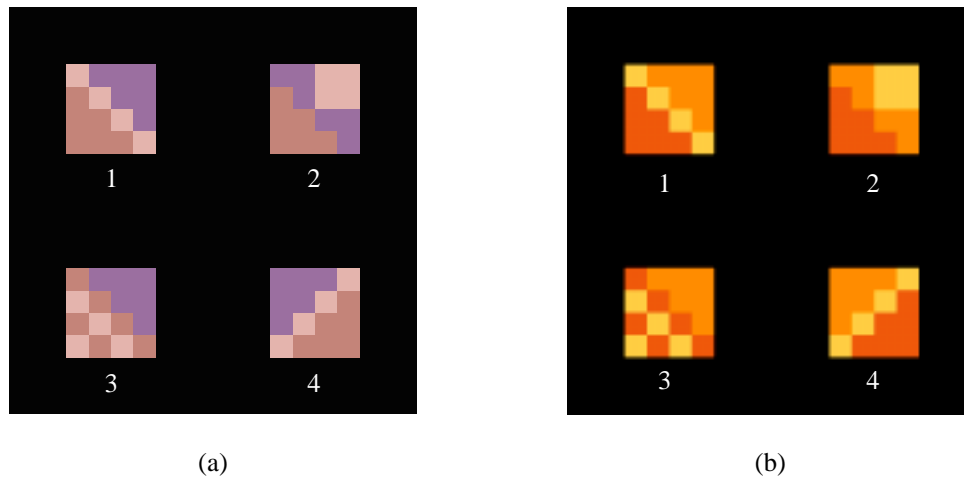
superfície. Per exemple, en els cítrics trobem les varietats *Sanguina*, o bé podem tractar amb altres fruites com algun tipus de préssecs, nectarines, etc. Per no fer tant particular l'exemple de reconeixement, també podem considerar els casos d'escenes on hi ha variacions en la il·luminació de la superfície acolorida, objectes originals i descolorits, etc.

En els casos esmentats, es manté com a fet comú que els colors que apareixen en les escenes són semblants i pertanyen, en general, a una mateixa tonalitat. Per tant, els requeriments de capacitat de discriminació del sistema de reconeixement són molt alts en quant a la discriminació del color i, el sistema convencional RGB no aconsegueix sempre resoldre aquests problemes de reconeixement.

Considerem les escenes test de la Figura 3.27, on els objectes policromàtics contenen diferents zones de colors dins la gamma magenta (rosa pàl·lid, salmó i lila) (Figura 3.27a) i dins la gamma taronja (groc, taronja clar i taronja fosc) (Figura 3.27b). Les corbes de reflectància espectral d'aquests colors es mostraven en les Figures 3.4 i 3.5, respectivament.

En aquest estudi, a partir dels colors captats amb les càmeres en les experiències prèvies, s'han generat per ordinador les matrius numèriques corresponents a les escenes de la Figura 3.27 tant pel sistema RGB, com pel de canals seleccionats.

La referència a reconèixer en ambdues escenes test, és l'objecte 1 i, a ell, s'ha adaptat un filtre POF. Els resultats de reconeixement, després de simular el procés de correlació, es recullen en la Taula 3.9 per a la gamma magenta i en la Taula 3.10 per la taronja. En les taules també s'afegeixen els resultats de reconeixement en el cas de considerar la descomposició convencional RGB. El llindar de reconeixement s'ha establert, com fins ara, al 50% del màxim d'autocorrelació. Els pics de correlació inclosos dins la regió de reconeixement delimitada pel doble llindar, apareixen subratllats en les taules de resultats. L'operació lògica AND ens permet obtenir els resultats finals de reconeixement.



**Figura 3.27.** (a) Escena d'objecte policromàtics de la gamma magenta, i (b) de la gamma taronja.

En el sistema convencional RGB, apareixen falses alarmes per tots els objectes de la gamma magenta, i pels objectes 2 i 4 de la gamma taronja. Per tant, la discriminació del sistema amb la descomposició multicanal RGB no és suficient per resoldre la tasca de reconeixement.

En les dues escenes test, el sistema de canals quasimonocromàtics permet eliminar les falses alarmes que apareixen en el mètode de descomposició RGB. Tant en l'escena de colors dins la gamma magenta, com en la de colors dins la gamma taronja, s'aconsegueix identificar de manera folgada el patró corresponent i discriminar-lo de la resta d'objectes. Al mateix temps, podem observar que també és possible una reducció en el nombre de canals a processar. Per exemple, per l'escena de la gamma magenta, els canals  $\lambda_3 = 578$  nm o  $\lambda_4 = 589$  nm individualment, resolten el problema de reconeixement plantejat de manera relaxada. En el cas de l'escena de la gamma taronja, també un sol canal,  $\lambda_1 = 514$  nm, és suficient per reconèixer l'objecte buscat i discriminar-lo de la resta.

Deduïm per tant, que el mètode de selecció de canals quasimonocromàtics potencia les diferències en reflectància dels colors dels objectes presents en l'escena, fent augmentar la capacitat discriminativa del sistema de reconeixement. En el cas de les escenes policromàtiques que ens ocupa, el fet d'augmentar les diferències en

reflectància no s'ha d'entendre només com a diferències en l'energia o nivell de gris, sinó que també poden arribar a ésser variacions de forma dels objectes en cada canal.

**Taula 3.9.** Intensitats simulades de correlació normalitzades per l'escena d'objectes policromàtics de la gamma magenta (Figura 3.27a) en el sistema RGB i en els canals seleccionats. L'objecte a reconèixer és objecte 1, a qui s'ha adaptat un filtre POF.

RGB	obj 1	obj 2	obj 3	obj 4	Canals seleccionats	obj 1	obj 2	obj 3	obj 4
<b>R</b>	<u>1.00</u>	<u>0.65</u>	<u>0.87</u>	<u>0.72</u>	$\lambda_1 = 514$	<u>1.00</u>	0.30	<u>0.88</u>	0.15
<b>G</b>	<u>1.00</u>	<u>0.50</u>	<u>0.64</u>	<u>0.56</u>	$\lambda_2 = 550$	<u>1.00</u>	0.14	<u>0.52</u>	0.10
<b>B</b>	<u>1.00</u>	<u>0.93</u>	<u>0.59</u>	<u>0.75</u>	$\lambda_3 = 578$	<u>1.00</u>	0.13	0.37	0.09
					$\lambda_4 = 589$	<u>1.00</u>	0.14	0.30	0.08
<b>Reconeix.</b>	SI	SI	SI	SI	<b>Reconeix.</b>	SI	NO	NO	NO

**Taula 3.10.** Intensitats simulades de correlació normalitzades per l'escena d'objectes policromàtics de la gamma taronja (Figura 3.27b) en el sistema RGB i en els canals seleccionats. L'objecte a reconèixer és l'objecte 1, a qui s'ha adaptat un filtre POF.

RGB	obj 1	obj 2	obj 3	obj 4	Canals seleccionats	obj 1	obj 2	obj 3	obj 4
<b>R</b>	<u>1.00</u>	<u>0.99</u>	<u>0.99</u>	<u>1.00</u>	$\lambda_1 = 514$	<u>1.00</u>	0.12	0.43	0.11
<b>G</b>	<u>1.00</u>	<u>0.83</u>	0.28	<u>0.51</u>	$\lambda_2 = 550$	<u>1.00</u>	<u>0.85</u>	0.26	0.36
<b>B</b>	-	-	-	-					
<b>Reconeix.</b>	SI	SI	NO	SI	<b>Reconeix.</b>	SI	NO	NO	NO

L'aplicació del mètode de canals seleccionats en el reconeixement i discriminació d'objectes policromàtics que acabem d'estudiar, té interès per ella mateixa perquè permet conèixer l'abast de la proposta en la millora de la capacitat de discriminació cromàtica del sistema de reconeixement.

A més, fem esment aquí, de l'estudi comparat que hem realitzat en aquest cas amb una altra proposta de correlació multicanal [185,186]. Concretament, amb la proposta realitzada per Millán *et al.* a [114,118] d'aplicar els models de visió humana del color, basats en la teoria dels colors oponents, per passar de les components RGB captades per la càmera, a les components ATD (A=clar-fosc, T=vermell-verd i D=blau-groc) mitjançant una transformació lineal amb coeficients positius i negatius. El mètode dels canals ATD s'ha aplicat també en [114,118] al reconeixement d'objectes policromàtics composts per colors pertanyents a una mateixa gamma de to aconseguint, en general, una millora en la capacitat discriminativa (respecte al sistema RGB) amb un nombre menor de canals.

En els casos de reconeixement analitzats en [185] per ambdues propostes (la selecció adaptada de canals quasimonocromàtics i el sistema de canals ATD), els resultats ens permeten dir que tant el sistema de selecció de canals monocromàtics, com el mètode basat en els models de visió humana arriben a conclusions similars. Per una banda s'augmenta la capacitat de discriminació del sistema de reconeixement i al mateix temps es redueix el temps de processament mitjançant una reducció del nombre de canals considerats.

Donat un problema de reconeixement, escollir una de les configuracions multicanals esmentades dependrà de diferents factors. Un d'ells és el requeriment d'informació prèvia, necessari només pel mètode de selecció de canals quasimonocromàtics. Els coneixements a priori de les corbes de reflectància dels colors que intervenen en el problema de reconeixement sovint no són fàcils d'obtenir, sobretot en el cas d'analitzar escenes amb objectes reals. En aquest capítol, però, hem posat de manifest que l'ús d'escenes test d'entrenament poden ajudar a resoldre satisfactòriament el problema.



Un altre factor a tenir en compte per escollir un o altre mètode, serà l'accessibilitat dels dispositius d'adquisició d'imatges i la reproducció de la informació cromàtica en els medis de reproducció més habituals. En el mètode de selecció de canals s'utilitza una càmera CCD en blanc i negre, i la descomposició cromàtica en els canals seleccionats es realitza en el mateix moment de l'adquisició, que es fa de manera seqüencial mitjançant els filtres interferencials de banda estreta situats davant de l'objectiu. En el mètode dels models de visió, es requereix una càmera CCD en color que proporciona les components RGB de la imatge enregistrada. En aquest cas, és necessari un pre-processat de la informació per aconseguir les noves components cromàtiques en els canals ATD mitjançant un pas de càlcul digital.

En quant a la reproducció de les components sobre un suport que es pugui emprar en el correlador òptic, el mètode de descomposició adaptada en  $n$  canals quasimonocromàtics, presenta l'avantatge d'obtenir unes noves components sempre amb valors numèrics positius. Aquesta característica facilita la reproducció de les components en diferents suports, com poden ser transparències o moduladors espacials de llum, sense necessitat de codificar la informació a representar. En els canals ATD, en canvi, s'obtenen unes components amb valors, en general, positius i negatius, per la qual cosa és necessari un procés de codificació d'aquesta informació per a ser introduïda en el correlador. En aquest sentit, la proposta dels canals quasimonocromàtics implica també una disminució en el temps de computació del sistema de reconeixement.

L'anàlisi d'aquestes i altres configuracions multicanals es detalla amb exemples en [186]. Es pot dir que, donat un problema de reconeixement, el tipus de descomposició multicanal haurà d'ésser escollida tenint en compte diversos factors, com ara, la facilitat d'obtenir informació prèvia, la disponibilitat d'aparells d'adquisició, la conveniència d'un procés seqüencial d'enregistrament o d'un pre-processat digital de la informació, la disposició dels aparells de reproducció, etc.

## Capítol 4

### Correlació no-lineal dual (DNC)

Més enllà d'una millora en la capacitat de discriminació com l'estudiada en el capítol 3 pel cas del reconeixement d'objectes en color, ens plantegem a partir d'ara la utilització d'un mètode general de reconeixement, que permeti parametritzar una àmplia varietat de mètodes de filtratge. Amb aquest mètode general es pretén disposar d'un marc que abasti no només mètodes de filtratge ja coneguts, sinó també, que possibiliti graus intermedis entre ells i doni cabuda a altres mètodes totalment nous. Es tracta de propiciar un control variable de les diverses propietats que millor caracteritzen el reconeixement per correlació òptica (detectabilitat dels màxims de correlació, localització de l'objecte, capacitat discriminativa, eficiència lluminosa, resistència al soroll, etc.) de manera que, davant les exigències i toleràncies en aquestes propietats plantejades per un problema concret, es determini el mètode de filtrat més adequat dins un únic marc global. Un mètode de reconeixement que contingui una varietat de filtres adaptats permet, a més d'assignar el filtre més indicat a un problema concret, la possibilitat de canviar de filtre, si convé, amb

relativa facilitat, és a dir, realitzant petits canvis que no afectin al disseny fonamental del mètode o a l'estructura bàsica del sistema que l'aplica.

Les operacions de globalització apareixen amb anterioritat a la literatura. Així, com exemple destacat, Vijaya Kumar i Hassebrook [6] van introduir en el 1990 els filtres de potències fraccionàries (FPF). Aquests filtres tenen la mateixa transmitància en fase (que coincideix amb el complex-conjugat de la fase de la transformada de Fourier de l'objecte a detectar) i difereixen en el mòdul de l'amplitud. Una característica comú d'aquests filtres adaptats és que proporcionen una correlació que és lineal amb l'energia de l'escena analitzada. A més dels filtres CMF, POF, i IF, ja coneguts prèviament en la literatura, els FPFs també incloïen en la seva definició noves operacions de filtratge adaptat. Posteriorment, Ersoy i Zeng [29] van proposar una generalització similar pel cas del processament no-lineal en el filtre adaptat, amb la característica que tant l'escena a analitzar, com l'objecte a reconèixer se'ls apliqués la mateixa operació no-lineal abans de realitzar el corresponent producte en el pla de Fourier (filtratge adaptat no-lineal simètric). Recentment, i fent un pas més endavant, Kotynski i Chalasinska-Macukow han proposat el mètode de la correlació no-lineal dual (DNC, de l'anglès *Dual Nonlinear Correlation*) [42] com una operació que, a més d'englobar mètodes de reconeixement per correlació òptica lineals i no-lineals, permet realitzar un processament no-lineal de l'escena a analitzar diferent del de l'objecte de referència.

En els propers capítols, ens centrarem en l'anàlisi i desenvolupament de les possibilitats de la DNC en el camp del reconeixement òptic d'objectes. Començarem, en aquest capítol, amb una descripció dels fonaments teòrics del mètode (seccions 4.1 i 4.2). Introduïrem també una anàlisi de la possibilitat de restricció de la regió de suport de les distribucions involucrades en el càlcul de la DNC com una eina útil en el reconeixement d'objectes texturats amb textures regulars senzilles [187] (secció 4.3). Finalment, en l'apartat 4.4, descriurem el dispositiu optoelectrònic capaç de realitzar la DNC experimental.

## 4.1 Definició de la DNC

Siguin  $s(\vec{x})$  i  $r(\vec{x})$  les funcions que representen, respectivament, l'escena d'entrada i l'objecte a reconèixer (referència). Siguin  $S(\vec{\nu})$  i  $R(\vec{\nu})$  llurs respectives transformades de Fourier. Representarem mitjançant  $N_P[z]$  una operació no-lineal sobre l'argument  $z$  que es defineix a partir de l'expressió

$$N_P[z] = \begin{cases} |z|^{P-1} \cdot z & \text{si } |z| > 0 \\ 0 & \text{si } |z| = 0. \end{cases} \quad (4.1)$$

El paràmetre real  $P$  indica l'ordre de la no-linealitat. Aquesta operació no-lineal s'aplicarà a funcions bidimensionals complexes com són els espectres de Fourier de l'escena i de l'objecte a reconèixer. Quan el paràmetre  $P = 1$ , l'Eq. (4.1) es converteix en una identitat on, per tant, no es modifica l'argument  $z$ . Si  $P = 0$ , l'operació no-lineal equival a una extracció de la fase de  $z$ . Suposem que el símbol  $IFT$  denota la transformada de Fourier inversa i que utilitzem una funció de pes  $T(\vec{\nu})$  arbitrària, real i positiva per controlar la regió de suport o domini a l'espai de Fourier sobre el que s'aplica la no-linealitat. La correlació no-lineal dual (DNC) es defineix a partir de l'expressió

$$DNC_{L,M}^{T(\vec{\nu})} \{s(\vec{x}), r(\vec{x})\} = IFT \left\{ N_L[S(\vec{\nu})] \cdot N_M[R^*(\vec{\nu})] \cdot T(\vec{\nu}) \right\}, \quad (4.2)$$

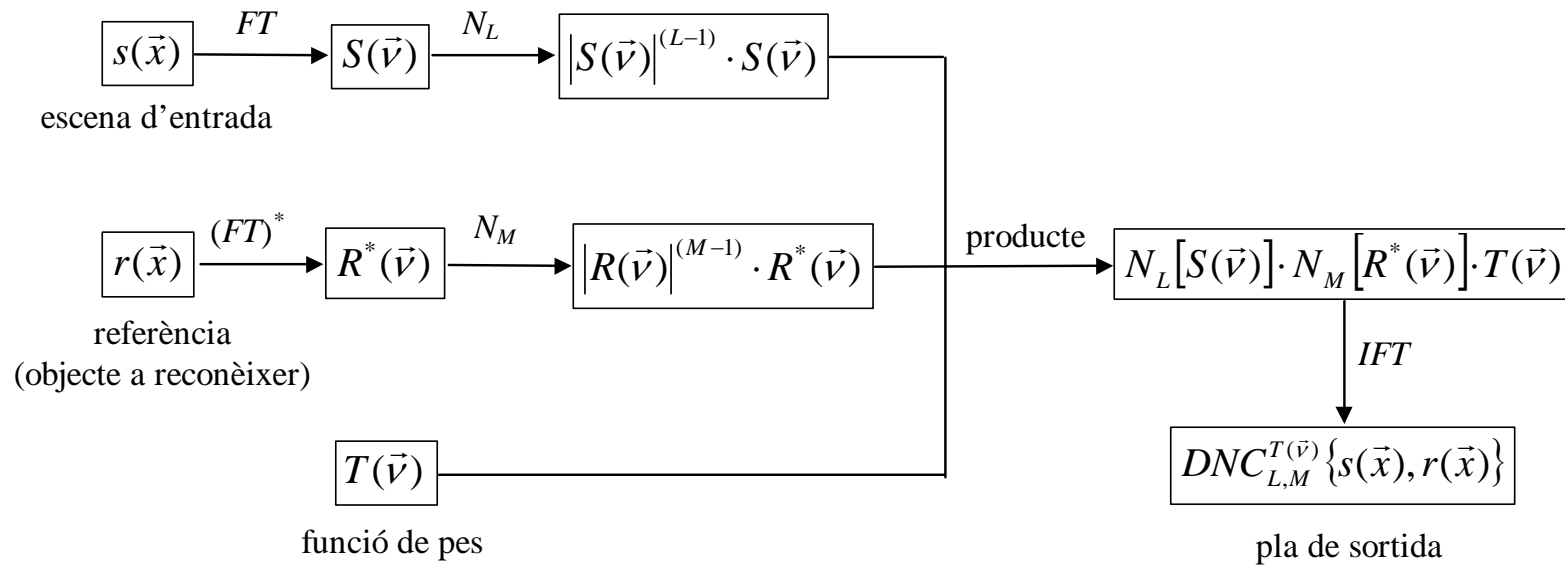
on el símbol (\*) representa el complex-conjugat,  $L$  és l'ordre de la no-linealitat que s'aplica a l'escena i  $M$  l'ordre de la no-linealitat que s'aplica a l'objecte a reconèixer [42]. El procés que condueix a la DNC, definida per l'Eq. (4.2), es representa esquemàticament en la Figura 4.1.

Si tenim en compte la definició de l'operador no-lineal (Eq. 4.1), l'expressió (4.2) es pot escriure com

$$DNC_{L,M}^{T(\vec{\nu})} \{s(\vec{x}), r(\vec{x})\} = IFT \left\{ |S(\vec{\nu})|^{L-1} \cdot S(\vec{\nu}) \cdot |R(\vec{\nu})|^{M-1} \cdot R^*(\vec{\nu}) \cdot T(\vec{\nu}) \right\}. \quad (4.3)$$

El tipus de no-linealitat aplicada a l'escena d'entrada i a l'objecte de referència es determina exclusivament a través dels paràmetres  $L$  i  $M$  i la funció de pes  $T(\vec{\nu})$ . Són valors particulars d'aquests paràmetres  $L$  i  $M$ , que recuperen alguns mètodes lineals i no-lineals de filtratge que s'han introduït prèviament en la literatura. Recollim alguns exemples de filtres adaptats ben coneguts en la Taula 4.1. Al confeccionar-la, s'ha considerat  $T(\vec{\nu}) = 1$  en tots els casos. Els exemples corresponen al filtre clàssic CMF [1], al filtre només de fase POF [4], al filtre invers IF [5], als filtres de potències fraccionàries FPF [6], als filtres localment no-lineals (de l'anglès *Locally nonlinear matched filters*, LNMF) [13], a la correlació pura de fase PPC (de l'anglès *Pure Phase Correlation* [30-33]), a la correlació de transformades conjuntes no-lineal NJTC (de l'anglès *Nonlinear Joint Transform Correlation* [158]) i al filtre òptim OF (de l'anglès *Optimal Filter* [197]). També s'inclou en la Taula 4.1 l'expressió particular que en cada cas adopta la DNC. Els paràmetres  $L$  i  $M$  varen ser inicialment introduïts en [42] com funcions de les freqüències espacials,  $L(\vec{\nu})$  i  $M(\vec{\nu})$ , i així s'indica per a la DNC en la Taula 4.1. No obstant, al llarg d'aquesta memòria,  $L$  i  $M$  es restringiran a valors reals.

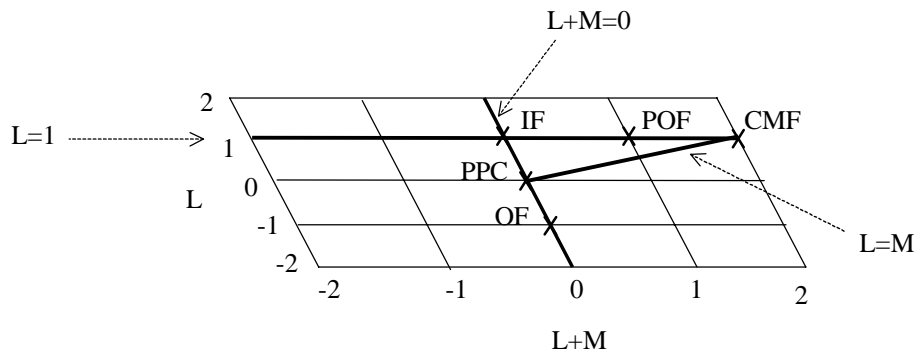
Les operacions particulars de filtratge incloses en la DNC també es poden representar gràficament com es mostra en la Figura 4.2. Sobre la recta  $L = 1$  trobaríem els filtres de potències fraccionàries (FPF) i, en particular, el filtre clàssic (CMF), el filtre només de fase (POF) i el filtre invers (IF). Tots ells, per ser lineals amb l'escena d'entrada, es poden obtenir òpticament mitjançant un correlador 4f o derivats d'ell, realitzant un procés de filtratge adaptat sobre l'objecte de referència corresponent a la operació que es vol dur a terme en cada cas. La resta d'operacions especificades en aquest gràfic contenen també un processament no-lineal de l'escena per la qual cosa requereixen realitzar un processament previ de la informació de l'escena abans de ser introduïda en el correlador, o bé requereixen d'altres tipus de correladors, com podria ser el correlador de transformades conjuntes, que permet fer el processat no-lineal i simultani de la informació de l'escena d'entrada i de l'objecte a reconèixer. La recta descrita per l'equació  $L = M$ , correspon al JTC no-lineal proposat per Javidi [158].



**Figura 4.1.** Representació esquemàtica del procés que condueix a la DNC.

**Taula 4.1.** Exemples de filtres adaptats expressats segons la formulació de la DNC donada per l'Eq. (4.3). En tots els casos s'ha considerat  $T(\vec{\nu}) = 1$ . No s'ha indicat la dependència espacial o freqüencial de les variables, a fi de simplificar les expressions.

Tipus de filtratge	$L$	$M$	DNC
Filtre adaptat clàssic (CMF)	1	1	$DNC_{1,1}\{s, r\} = IFT\{S \cdot R^*\}$
Filtre només de fase (POF)	1	0	$DNC_{1,0}\{s, r\} = IFT\left\{S \cdot \frac{R^*}{ R }\right\}$
Filtre invers (IF)	1	-1	$DNC_{1,-1}\{s, r\} = IFT\left\{S \cdot \frac{R^*}{ R ^2}\right\}$
Filtres de potències fraccionàries (FPF)	1	$p$ (nombre real)	$DNC_{1,p}\{s, r\} = IFT\{S \cdot  R ^{p-1} \cdot R^*\}$
Filtres adaptats localment no-lineals (LNMF)	1	$-(M(\vec{\nu}) - 1)$ (funció real i positiva)	$DNC_{1,-(M-1)}\{s, r\} = IFT\left\{S \cdot \frac{R^*}{ R ^M}\right\}$
Correlació només de fase (PPC)	0	0	$DNC_{0,0}\{s, r\} = IFT\left\{\frac{S}{ S } \cdot \frac{R^*}{ R }\right\}$
Correlador de transformades conjuntes no-lineal (NJTC)	$k \in [0,1]$ (nombre real)		$DNC_{k,k}\{s, r\} = IFT\{ S ^{k-1} S \cdot  R ^{k-1} \cdot R^*\}$
Filtre òptim (OF)	-1	1	$DNC_{-1,1}\{s, r\} = IFT\left\{\frac{S}{ S ^2} \cdot R^*\right\}$
Correlació no-lineal dual (DNC)	$L(\vec{\nu})$	$M(\vec{\nu})$	$DNC_{L,M}\{s, r\} = IFT\{ S ^{L-1} \cdot S \cdot  R ^{M-1} \cdot R^*\}$



**Figura 4.2.** Representació esquemàtica de diferents operacions de filtratge adaptat en funció dels paràmetres  $L$  i  $M$  definits en la DNC.

A nivell il·lustratiu, mostrem un exemple d'aplicació del procés de correlació no-lineal dual. Considerem la imatge de la Figura 4.3 com a escena a analitzar,  $s(\vec{x})$ , que conté dos caràcters, una lletra E i una H, ambdós uniformes i equienergètics. La nostra tasca de reconeixement consistirà en identificar el caràcter E, o objecte de referència  $r(\vec{x})$ , i discriminar-lo de la lletra H present a la mateixa escena.

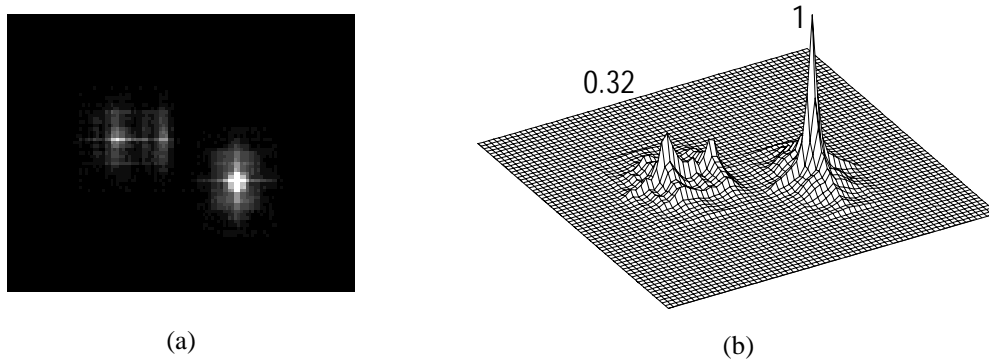


**Figura 4.3.** Escena a analitzar.

A partir de les transformades de Fourier d'ambdues funcions  $s(\vec{x})$  i  $r(\vec{x})$ , i considerant el cas més senzill on  $T(\vec{v})=1$ , realitzem el càlcul de les operacions indicades en l'Eq. (4.3) per a l'obtenció de la DNC. Realitzem el procés descrit per a dos parells de valors  $(L, M)$  diferents. En la Figura 4.4 es mostra el pla de correlació corresponent a la no-linealitat, equivalent al filtratge adaptat clàssic o CMF, descrita pels paràmetres  $L = M = 1$ . La Figura 4.4a ofereix una representació 2D en nivells de gris de la intensitat de l'esmentat pla. En la Figura 4.4b es mostra el mateix pla de

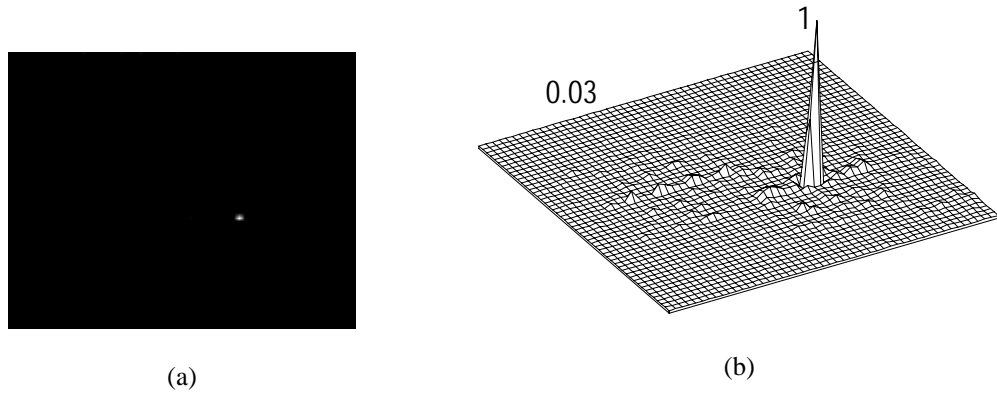


correlació en una representació tridimensional. En ambdues figures observem els característics pics de correlació localitzats en la mateixa posició on es troben els caràcters E i H dins l'escena a analitzar. Veiem com en el cas de la lletra E, el pic de correlació obtingut és de major intensitat i molt més estret, donat que la coincidència entre l'objecte analitzat i la referència és màxima (autocorrelació). En la posició ocupada per la lletra H, en canvi, la correlació és molt menys intensa, i està formada per dos lòbuls donades les característiques de forma de la lletra H. Aquest màxim d'intensitat rep el nom de correlació creuada entre dues funcions diferents.



**Figura 4.4.** Pla de correlació obtingut quan es vol reconèixer la lletra E dins l'escena de la Figura 4.3 aplicant la DNC pels paràmetres  $L = M = 1$  (cas del CMF). (a) Representació bidimensional, i (b) representació tridimensional, on s'indica el valor màxim d'intensitat de cada pic normalitzat a l'autocorrelació.

En la Figura 4.5 es mostra novament el pla de correlació, en representacions 2D i 3D, en el cas de considerar una no-linealitat definida pels paràmetres  $L = M = -1/2$ . En aquest cas, es pot observar que el sistema és capaç de reconèixer la lletra E i discriminar-la de la lletra H en condicions més relaxades que en el cas anterior ( $L = M = 1$  de la Figura 4.4). Observem que el pic d'autocorrelació és més estret i desapareixen el lòbuls laterals. A més, el pic de correlació creuada amb la lletra H és tan baix que es confon amb el nivell de soroll de fons.



**Figura 4.5.** Pla de correlació quan es vol reconèixer la lletra E dins l'escena de la Figura 4.3 aplicant la DNC pels paràmetres  $L = M = -1/2$ . (a) Representació 2D, i (b) representació 3D, on s'indica el valor màxim d'intensitat de cada pic normalitzat a l'autocorrelació.

## 4.2 Propietats de la DNC

### 4.2.1 Invariància a translacions

Aquesta característica permet que, donada una escena que inclogui l'objecte de referència en qualsevol posició de la imatge, l'operació de correlació ens doni el mateix pic de reconeixement, sempre centrat allà on estigui situat l'objecte buscat. En el cas de la DNC, la invariància a translacions es pot expressar de la següent manera,

$$DNC_{L,M}^{T(\vec{v})}\{s(\vec{x} - \vec{x}_0), r(\vec{x})\} = DNC_{L,M}^{T(\vec{v})}\{s(\vec{x}), r(\vec{x})\} * \delta(\vec{x} - \vec{x}_0), \quad (4.4)$$

on  $\vec{x}_0$  és la posició on s'ha desplaçat la funció  $s(\vec{x})$ , i  $*$  és el símbol que indica la convolució.

Aquesta propietat (4.4) es pot demostrar fàcilment, tenint en compte que la transformada de Fourier d'una funció  $s(\vec{x} - \vec{x}_0)$  compleix [128]

$$TF\{s(\bar{x} - \bar{x}_0)\} = S(\bar{\nu}) \cdot e^{-i2\pi\bar{x}_0\bar{\nu}}. \quad (4.5)$$

Tenint en compte (4.5), podem seguir el següent desenvolupament,

$$\begin{aligned} DNC_{L,M}^{T(\bar{\nu})}\{s(\bar{x} - \bar{x}_0), r(\bar{x})\} &= \\ &= IFT\{N_L[S(\bar{\nu}) \cdot \exp[-i2\pi\bar{x}_0\bar{\nu}]] \cdot N_M[R^*(\bar{\nu})] \cdot T(\bar{\nu})\} = \\ &= IFT\{|S(\bar{\nu})|^{L-1} \cdot S(\bar{\nu}) \cdot \exp[-i2\pi\bar{x}_0\bar{\nu}] \cdot |R(\bar{\nu})|^{M-1} \cdot R^*(\bar{\nu}) \cdot T(\bar{\nu})\} = \\ &= DNC_{L,M}^{T(\bar{\nu})}\{s(\bar{x}), r(\bar{x})\} * \delta(\bar{x} - \bar{x}_0). \end{aligned} \quad (4.6)$$

#### 4.2.2 Variació de la DNC amb la intensitat del feix de llum que il·lumina l'escena.

Considerem que l'escena a analitzar s'il·lumina amb un feix uniforme de manera que varia la seva intensitat en un factor  $I_{in}$ . En aquest cas, la intensitat de la correlació no-lineal dual es veu afectada de la següent manera,

$$\begin{aligned} |DNC_{L,M}^{T(\bar{\nu})}\{I_{in}^{1/2} \cdot s(\bar{x}), r(\bar{x})\}|^2 &= |IFT\{N_L[I_{in}^{1/2} \cdot S(\bar{\nu})] \cdot N_M[R^*(\bar{\nu})] \cdot T(\bar{\nu})\}|^2 = \\ &= \left| IFT\left\{ |I_{in}^{1/2} \cdot S(\bar{\nu})|^{L-1} \cdot I_{in}^{1/2} \cdot S(\bar{\nu}) \cdot |R(\bar{\nu})|^{M-1} \cdot R^*(\bar{\nu}) \cdot T(\bar{\nu}) \right\} \right|^2 = \\ &= \left| IFT\left\{ I_{in}^{L/2} \cdot |S(\bar{\nu})|^{L-1} \cdot S(\bar{\nu}) \cdot |R(\bar{\nu})|^{M-1} \cdot R^*(\bar{\nu}) \cdot T(\bar{\nu}) \right\} \right|^2 = \\ &= I_{in}^L \cdot |DNC_{L,M}^{T(\bar{\nu})}\{s(\bar{x}), r(\bar{x})\}|^2. \end{aligned} \quad (4.7)$$

Veiem que la intensitat de la DNC en el pla de sortida es veu afectada per un factor que és el valor  $I_{in}$  de variació en la intensitat del feix que il·lumina l'entrada, elevat al paràmetre  $L$  que regeix la no-linealitat aplicada a l'escena a analitzar. Per tant, sempre que considerem processos lineals amb l'escena d'entrada, és a dir quan  $L = 1$ , la variació en la intensitat de la DNC serà directament proporcional a  $I_{in}$ , com era d'esperar. Si  $L \neq 1$ , la relació deixa de ser proporcional i, en particular, si  $L = 0$ , la intensitat de la DNC en el pla de sortida és independent de les variacions uniformes d'intensitat del pla d'entrada.

### 4.2.3 Pic d'autocorrelació no-lineal dual

Per conèixer com és l'autocorrelació que produeix la DNC, analitzem escenes mono-objectes lliures de soroll, i considerem concretament una escena on només s'inclou l'objecte a reconèixer. En aquest cas, l'equació (4.3) es redueix a la següent expressió,

$$\begin{aligned} DNC_{L,M}^{T(\vec{v})} \{r(\vec{x}), r(\vec{x})\} &= IFT \left\{ |R(\vec{v})|^{L-1} \cdot R(\vec{v}) \cdot |R(\vec{v})|^{M-1} \cdot R^*(\vec{v}) \cdot T(\vec{v}) \right\} = \\ &= IFT \left\{ |R(\vec{v})|^{L+M} \cdot T(\vec{v}) \right\}, \end{aligned} \quad (4.8)$$

de manera que podem observar que l'autocorrelació no-lineal dual depèn de la suma dels paràmetres  $L$  i  $M$ . D'aquesta manera, les operacions que pertanyen a la recta descrita per  $L+M = 0$ , com poden ser IF, PPC o OF, tenen la particularitat que el pic d'autocorrelació és del tipus funció delta de Dirac, si a més es compleix que  $T(\vec{v}) = constant$ . El pic d'autocorrelació, en canvi, serà cada vegada més ample a l'augmentar el factor  $|L+M|$  (p. ex. el filtre POF amb  $L+M = 1$ , o el filtre CMF amb  $L+M = 2$ ).

#### 4.2.4 Escenes multiobjecte

Analitzarem en aquest apartat el comportament de la DNC davant d'escenes test que contenen múltiples objectes [42]. Considerarem el cas en què l'escena multiobjecte està constituïda per varis objectes localitzats en posicions diverses de l'escena i, tots ells amb la mateixa forma que l'objecte emprat com a referència. La funció que descriu aquest tipus d'escena es pot escriure com

$$s(\vec{x}) = \sum_k c_k \cdot r(\vec{x} - \vec{x}_k), \quad (4.9)$$

on considerem que tenim  $k$  objectes localitzats en les posicions descrites pels vectors  $\vec{x}_k$ , i amb diferent intensitat relativa entre ells donada pel valor  $|c_k|^2$ . L'expressió anterior també pot ésser escrita tenint en compte una distribució de funcions deltes de Dirac que ens proporcionen la localització dels diferents objectes,

$$s(\vec{x}) = r(\vec{x}) * \sum_k c_k \cdot \delta(\vec{x} - \vec{x}_k). \quad (4.10)$$

Si apliquem la DNC en aquesta escena, i denotant  $d(\vec{x}) = \sum_k c_k \cdot \delta(\vec{x} - \vec{x}_k)$ , obtenim l'expressió

$$\begin{aligned} DNC_{L,M}^{T(\vec{v})} \{s(\vec{x}), r(\vec{x})\} &= DNC_{L,M}^{T(\vec{v})} \{r(\vec{x}) * d(\vec{x}), r(\vec{x})\} = \\ &= DNC_{L,M}^{T(\vec{v})} \{r(\vec{x}), r(\vec{x})\} * IFT\{N_L[D(\vec{v})]\}, \end{aligned} \quad (4.11)$$

on  $D(\vec{v})$  és la transformada de Fourier de la funció  $d(\vec{x})$ . L'Eq. (4.11) es pot escriure de manera equivalent, com a la convolució de tres termes,

$$DNC_{L,M}^{T(\vec{v})}\{s(\vec{x}), r(\vec{x})\} = d(\vec{x}) * DNC_{L,M}^{T(\vec{v})}\{r(\vec{x}), r(\vec{x})\} * IFT\{|D(\vec{v})|^{L-1}\}. \quad (4.12)$$

El primer dels termes,  $d(\vec{x}) = \sum_k c_k \cdot \delta(\vec{x} - \vec{x}_k)$ , ens proporciona la localització dels pics d'autocorrelació en la posició on es troben situats els objectes en l'escena analitzada. Cada objecte que, recordem, és de forma idèntica a la referència, obtindrà un pic de correlació situat en el lloc de la funció delta de Dirac i d'alçada proporcional al seu coeficient  $c_k$ . El segon terme de l'expressió (4.12),  $DNC_{L,M}^{T(\vec{v})}\{r(\vec{x}), r(\vec{x})\}$ , correspon a l'autocorrelació no-lineal dual (Eq. 4.8). Aquest terme és el que dona la forma als pics de correlació. Concretament, en el cas en què  $T(\vec{v}) = constant$ , la forma del pic està determinada pel valor de  $L + M$ , de manera que només en el cas de considerar processaments on  $L + M = 0$  es recupera la forma de la funció delta en el pic d'autocorrelació.

Finalment, el tercer terme de l'expressió (4.12),  $IFT\{|D(\vec{v})|^{L-1}\}$ , provoca els efectes d'intermodulació característics en l'anàlisi d'escenes multiobjectes [36,38,198]. Aquest terme sol incloure una distribució de pics estrets que es poden aproximar per funcions deltes. La convolució amb aquest terme, és doncs, la font dels efectes d'intermodulació que provoca l'aparició de falses alarmes. La intermodulació depèn només del paràmetre  $L$ , i de la distribució dels objectes dins l'escena a analitzar, però no de la seva forma. Si la distribució dels objectes és periòdica, els resultats de correlació es veuen afectats pels efectes d'intermodulació. Tant en el cas de sistemes lineals (amb  $L = 1$ ), com d'escenes amb un sol objecte (amb  $|D(\vec{v})| = const$ ), aquest tercer terme no apareix en l'expressió (4.12), i els efectes d'intermodulació no intervenen en el procés de reconeixement. En el cas de considerar una distribució periòdica d'objectes dins l'escena a analitzar, s'han proposat mètodes que realitzen operacions de llindar adaptatives que milloren els resultats de reconeixement [16,40].

Cal remarcar que aquest estudi de l'aplicació de la DNC en escenes multiobjectes és vàlid en el cas d'imatges lliures de soroll. L'anàlisi en cas contrari, amb escenes

d'entrada que continguin soroll, s'ha de realitzar amb més detall, i s'obtidrien expressions més complicades.

Un altre problema a remarcar en l'anàlisi d'escenes multiobjecte, si s'utilitza l'arquitectura del correlador de transformades conjuntes, és el requeriment d'una amplada de banda espacial suficient per a compartir un mateix pla d'entrada entre l'escena multiobjecte a analitzar i l'objecte de referència. Aquesta amplada, en general, ha de ser suficient per evitar el solapament entre els termes de diferent ordre de difracció que s'obtenen a la sortida del JTC [199]. També s'ha de tenir en compte el fet que la intensitat del pic d'autocorrelació disminueix degut als diferents pics de correlacions creuades que apareixen entre els mateixos objectes de l'escena. Per tant, podem dir que la intensitat del pic d'autocorrelació depèn del nombre d'objectes continguts en l'escena a analitzar [46].

### **4.3 Modificació de la regió de suport mitjançant la funció $T(\vec{\nu})$ i els seus efectes en la discriminació d'objectes texturats**

Des de la primera descripció del model de correlació no-lineal dual [42], s'inclou dins l'equació matemàtica que descriu la DNC (Eq. 4.2) una funció de pes arbitrària, real i positiva,  $T(\vec{\nu})$ . La funció  $T(\vec{\nu})$ , està definida en el domini freqüencial i permet modificar el filtre de diferent manera segons interressi. Per exemple, la modificació introduïda per  $T(\vec{\nu})$  pot estar relacionada amb la informació relativa a l'obertura efectiva del sistema òptic, o es pot relacionar amb el nivell de soroll introduït durant el procés.

En particular, també és possible emprar la funció  $T(\vec{\nu})$  amb l'objectiu de modificar la regió de suport de les distribucions involucrades en el procés no-lineal. De fet, la introducció de nivells de transmitància nuls en el pla de freqüències ha estat prèviament proposada per diversos autor [200-201]. D'aquesta manera, és possible afegir un control addicional sobre la tolerància del filtre al soroll, o bé sobre l'agudesesa del pic.

Centrant-nos en el control de la regió de suport de les distribucions, podem pensar en aplicacions de la DNC cap a noves tasques de reconeixement on la informació rellevant que fa possible la identificació d'un objecte de referència està localitzada en una banda determinada de l'espectre de freqüències. Mitjançant una funció  $T(\vec{\nu})$  binària, podem aconseguir ressaltar la informació de la banda de freqüències adequada i eliminar al mateix temps altres freqüències amb contingut poc interessant pel problema plantejat, de manera que, a més de la identificació de l'objecte buscat, sigui possible també la seva discriminació davant d'altres objectes semblants.

En aquest sentit, doncs, volem posar de manifest la utilitat de la funció  $T(\vec{\nu})$ , i mostrarem alguns casos on la seva aplicació per a seleccionar unes determinades freqüències de l'espai de Fourier permet portar a terme un procés de reconeixement amb resultats satisfactoris. En concret, ens centrarem en el cas de treballar amb objectes amb nivell de gris uniforme i objectes amb una textura regular senzilla. És ben conegut que la informació d'una textura periòdica es concentra en una distribució regular de pics ben definits a l'espai de Fourier. Basat en aquest fet, una gran varietat de màscares supressores s'han dissenyat tradicionalment i s'han emprat eficientment per manipular el contingut espectral d'una imatge. Utilitzarem diferents funcions  $T(\vec{\nu})$  reals, positives i binàries que treballaran com a màscares supressores superposades a l'espectre de les distribucions involucrades en el càlcul de la DNC,  $S(\vec{\nu})$  i  $R(\vec{\nu})$ .

En primer lloc, farem una anàlisi matemàtica, que posi de manifest la utilitat de la funció de pes en l'exemple que acabem de fer esment. Suposem que ens interessa reconèixer un objecte independentment d'una textura periòdica que pugui tenir superposada, com per exemple, un patró de franges.

Representem l'objecte original, sense cap textura superposada, per la funció  $r(\vec{x})$ . L'objecte  $r(\vec{x})$  serà l'objecte de referència. Suposem que està limitat en freqüència. Això implica que el seu espectre,  $R(\vec{\nu})$ , és no-nul només sobre una regió finita de l'espai de Fourier. Suposem que el menor rectangle que inclou  $R(\vec{\nu})$  té per



dimensions  $B_u \times B_v$ . Sigui l'escena a analitzar  $s(\bar{x})$ , un únic objecte, similar a la referència en forma i energia total, però amb una textura periòdica superposada. Per simplificar l'anàlisi, suposem que la textura és un patró de franges verticals de freqüència  $u_0$ . En aquest cas, la funció que representa l'escena es descriu matemàticament per

$$\begin{aligned} s(\bar{x}) &= r(\bar{x}) [\cos(2\pi u_0 x) + 1] = \\ &= r(\bar{x}) \left\{ \frac{1}{2} [e^{i2\pi u_0 x} + e^{-i2\pi u_0 x}] + 1 \right\}, \end{aligned} \quad (4.13)$$

on  $\bar{x} = (x, y)$  són les coordenades cartesianes en el domini espacial. La transformada de Fourier de l'escena, que denotem per  $S(\bar{v})$ , té per expressió

$$S(\bar{v}) = R(\bar{v}) * \frac{1}{2} [\delta(u + u_0) + \delta(u - u_0) + 2\delta(u)], \quad (4.14)$$

on  $\bar{v} = (u, v)$  són les coordenades cartesianes en el domini de freqüències. En l'Eq. (4.14) s'obté que en l'espectre  $S(\bar{v})$  apareix l'espectre de l'objecte a reconèixer  $R(\bar{v})$ , repetit sobre els pics d'intensitat que genera la periodicitat de la textura, de freqüència fonamental  $u_0$ . Suposem també, que la freqüència de les franges  $u_0$  és major que l'amplada de banda  $B_u$  de l'objecte de referència  $r(\bar{x})$ , és a dir,

$$u_0 > B_u. \quad (4.15)$$

En conseqüència, l'espectre  $R(\bar{v})$  apareix repetit en  $S(\bar{v})$  però sense solapaments entre rèpliques adjacents.

Si apliquem el mètode de reconeixement basat en el model de la correlació no-lineal dual s'obté que l'expressió de la DNC entre les funcions  $s(\bar{x})$  i  $r(\bar{x})$  val

$$\begin{aligned}
DNC_{L,M}^{T(\bar{v})}\{s(\bar{x}), r(\bar{x})\} &= DNC_{L,M}^{T(\bar{v})}\left\{r(\bar{x})\left[\frac{1}{2}\left(e^{i2\pi u_0 x} + e^{-i2\pi u_0 x}\right) + 1\right], r(\bar{x})\right\} = \\
&= IFT\left\{N_L\left[R(\bar{v}) * \frac{1}{2}[\delta(u+u_0) + \delta(u-u_0) + 2\delta(u)]\right] \cdot N_M[R(\bar{v})] \cdot T(\bar{v})\right\} = \\
&= IFT\left\{N_L\left[\frac{1}{2}[R(\bar{v}+u_0) + R(\bar{v}-u_0)] + R(\bar{v})\right] \cdot N_M[R(\bar{v})] \cdot T(\bar{v})\right\} = \quad (4.16) \\
&= IFT\left\{\left[\frac{1}{2}[R(\bar{v}+u_0) + R(\bar{v}-u_0)] + R(\bar{v})\right]^{L-1} \cdot \right. \\
&\quad \left. \cdot \left[\frac{1}{2}[R(\bar{v}+u_0) + R(\bar{v}-u_0)] + R(\bar{v})\right] \cdot |R(\bar{v})|^{M-1} \cdot R^*(\bar{v}) \cdot T(\bar{v})\right\}.
\end{aligned}$$

Com pretenem reconèixer l'objecte  $r(\bar{x})$  independentment del patró de franges que se li superposa a l'escena  $s(\bar{x})$ , escollirem la funció de pes  $T(\bar{v})$  que restringeixi el domini de Fourier d'ambdues funcions a la zona que tenen en comú. D'aquesta manera, es potenciarà la informació que ambdues funcions comparteixen ( $r(\bar{x})$ ) i s'eliminarà la informació que les diferencia (la corresponent al patró de franges). Per exemple, podem prendre com a  $T(\bar{v})$  una funció rectangle de la forma

$$T_0(\bar{v}) = \text{rect}\left(\frac{u}{B_u}\right) = \begin{cases} 1 & |u| < B_u/2 \\ 1/2 & |u| = B_u/2, \\ 0 & \text{en la resta} \end{cases} \quad (4.17)$$

que genera una banda vertical de transmitància en el domini de Fourier d'amplada  $B_u$  igual a l'amplada de banda en l'eix  $u$  de l'objecte de referència  $r(\bar{x})$ . Substituint (4.17) en (4.16), s'obté

$$\begin{aligned}
DNC_{L,M}^{T_0(\bar{\nu})} \{s(\bar{x}), r(\bar{x})\} &= IFT \left\{ |R(\bar{\nu})|^{L-1} \cdot R(\bar{\nu}) \cdot |R(\bar{\nu})|^{M-1} \cdot R^*(\bar{\nu}) \right\} = \\
&= DNC_{L,M}^{T(\bar{\nu})=1} \{r(\bar{x}), r(\bar{x})\}.
\end{aligned} \tag{4.18}$$

L'Eq. (4.18) mostra que, en el cas de definir la funció  $T(\bar{\nu})$  d'acord amb (4.17), s'obtindria per l'escena  $s(\bar{x})$ , el mateix valor de la DNC que pel propi objecte de referència  $r(\bar{x})$ . Aquest valor coincideix amb l'autocorrelació no-lineal dual de l'objecte de referència  $r(\bar{x})$  realitzada sense restricció en la seva amplada de banda espectral ( $B_u \times B_v$ ). Tal resultat permet afirmar que l'escena  $s(\bar{x})$  seria reconeguda pel sistema com si fos l'objecte de referència  $r(\bar{x})$ , aconseguint així l'objectiu que es pretenia.

En un cas més general, no obstant, pot ser que la freqüència del patró de textura no superi l'amplada de banda espectral de l'objecte de referència. En l'exemple que acabem d'analitzar, implicaria que no es complís estrictament la desigualtat (4.15). En aquest cas, les rèpliques d' $R(\bar{\nu})$  adjacents a l'espectre d' $S(\bar{\nu})$ , si que es solaparien parcialment i, per tant, la igualtat (4.18) no es compliria. Aquest fet, no obstant, no té perquè invalidar el procediment. Definint  $T(\bar{\nu})$  com la banda

$$T(\bar{\nu}) = \text{rect} \left( \frac{u}{u_0} \right), \tag{4.19}$$

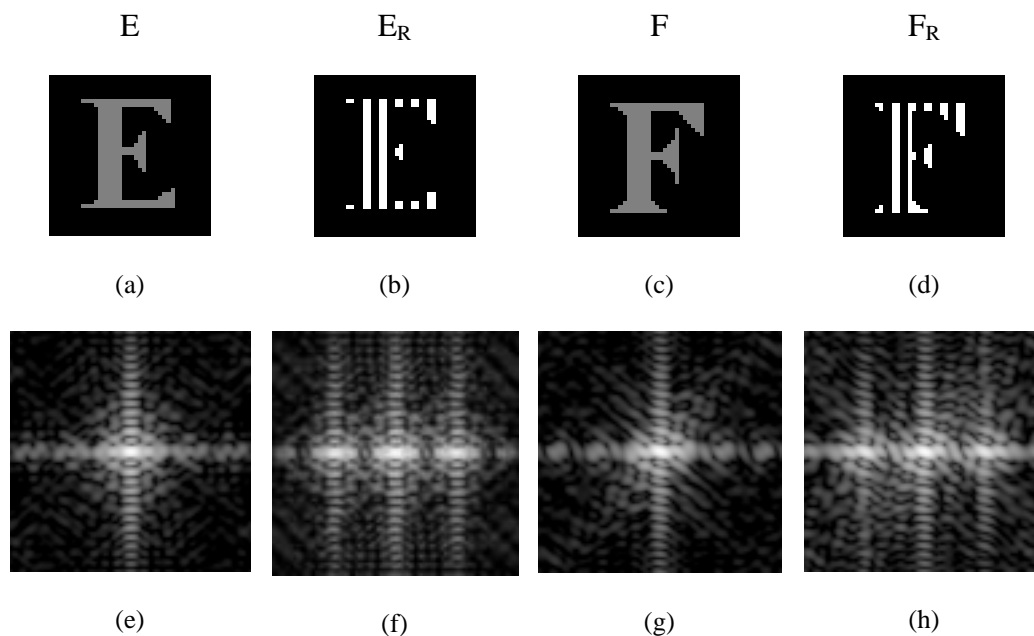
es podria seleccionar de l'espectre  $R(\bar{\nu})$  les freqüències baixes més energètiques, i per tant de major pes. Tot i que la igualtat (4.18) no s'arribi a complir, existeix una tolerància en quant a les variacions de la correlació (banda de reconeixement definida pel doble llindar), que pot, malgrat tot, permetre el reconeixement de l'escena  $s(\bar{x})$  com si fos l'objecte  $r(\bar{x})$ . Clarament, aquesta situació serà tant més

favorable quant més alta sigui la freqüència fonamental de la textura,  $u_0$ , acostant-se al valor de  $B_u$ .

Sobre aquesta base, analitzem alguns casos il·lustratius de la utilitat de la funció  $T(\bar{v})$  en el reconeixement d'objectes texturats.

### 4.3.1 Cas de reconeixement I

En el primer cas, volem reconèixer les lletres E independentment de la seva textura, i discriminar-les de la resta de lletres incloses en el conjunt test mostrat en la Figura 4.6 (a)-(d). Es tracta d'un problema clàssic en reconeixement, on la forma E es vol discriminar davant d'una forma molt semblant com és la F. En aquest cas, a més, pretenem que el sistema sigui sensible a la variació de forma de les lletres E i F, però sense ser sensible al patró de la textura que se'ls superposa en la Figura 4.6. Per tal de facilitar la comparació dels resultats, totes les lletres de la Figura 4.6 tenen la mateixa energia. La Figura 4.6 també conté una representació logarítmica de la distribució d'intensitat de la transformada de Fourier de cada lletra del test (Figura 4.6 (e)-(h)).



**Figura 4.6.** (a)–(d) Representació de les escenes test consistents en caràcters equienergètics, uniformes i altres texturats. (e)–(h) Representació logarítmica de les corresponents transformades de Fourier en intensitat.

Intentem resoldre el problema plantejat aplicant primerament alguns dels mètodes de filtratge adaptat convencionals, com són el CMF, el POF i el PPC. Els resultats obtinguts per simulació numèrica quan la lletra E amb nivell de gris uniforme (Figura 4.6a) es pren com a referència es mostren en la Taula 4.2, mentre que en la Taula 4.3 estan inclosos els resultats de simulació quan el procés de reconeixement es realitza prenent com a patró la lletra E amb la textura de ratlles verticals (Figura 4.6b). En ambdues taules, els màxims d'intensitat dels pics de correlació creuada ( $cc$ ) estan normalitzats al màxim de l'autocorrelació ( $ac$ ), és a dir, es presenten els valors  $cc/ac$  en les Taules 4.2 i 4.3. Com en els capítols anteriors, es determina la capacitat de discriminació del sistema de reconeixement, aplicant l'expressió

$$DC = \left| 1 - \frac{cc}{ac} \right| \leq u, \quad (4.20)$$

on  $u$  és el valor llindar que limita la banda de valors de correlació pels quals el sistema reconeix un objecte i l'identifica amb la referència. Fora d'aquesta banda l'objecte és discriminat. De l'Eq. (4.20), calculem el valor de  $DC$  pels resultats de correlació normalitzats de les Taules 4.2 i 4.3, i considerem un valor llindar comú de  $u = 0.5$ . Tots aquells objectes amb llur resultat de correlació que compleixi  $DC \leq u$  (subratllats en les Taules 4.2 i 4.3) són reconeguts i identificats amb la referència. Per contra, els que no ho compleixen (sense subratllar en les Taules 4.2 i 4.3) no són reconeguts i el sistema els discrimina de la referència.

**Taula 4.2.** Resultats de correlació per a les escenes analitzades de la Figura 4.6 (a)-(d), quan la lletra E uniforme és l'objecte de referència (Figura 4.6a). Els valors dels màxims de correlació creuada normalitzats que apareixen subratllats, estan en la banda de reconeixement i, per tant, indiquen que les lletres corresponents es reconeixen com a similars al patró.

<b>Correlació normalitzada (<math>cc/ac</math>)</b>				
<b>Mètode de filtratge</b>	<b>E (ref.)</b>	<b>E<sub>R</sub></b>	<b>F</b>	<b>F<sub>R</sub></b>
<b>CMF</b>	<u>1.00</u>	<u>1.01</u>	<u>0.80</u>	<u>0.86</u>
<b>POF</b>	<u>1.00</u>	<u>1.00</u>	<u>0.50</u>	<u>0.55</u>
<b>PPC</b>	<u>1.00</u>	0.38	0.24	0.13

**Taula 4.3.** Resultats de correlació per a les escenes analitzades de la Figura 4.6 (a)-(d), quan la lletra E<sub>R</sub> ratllada és l'objecte de referència (Figura 4.6b). Els valors dels màxims de correlació creuada normalitzats que apareixen subratllats, estan en la banda de reconeixement i, per tant, indiquen que les lletres corresponents es reconeixen com a similars al patró.

<b>Correlació normalitzada (<math>cc/ac</math>)</b>				
<b>Mètode de filtratge</b>	<b>E</b>	<b>E<sub>R</sub> (ref.)</b>	<b>F</b>	<b>F<sub>R</sub></b>
<b>CMF</b>	0.32	<u>1.00</u>	0.28	<u>0.84</u>
<b>POF</b>	0.26	<u>1.00</u>	0.18	<u>0.67</u>
<b>PPC</b>	0.22	<u>1.00</u>	0.12	0.30

Dels resultats continguts en les Taules 4.2 i 4.3 observem que, amb els mètodes de filtratge convencionals CMF i POF, no es pot realitzar el reconeixement de la lletra E independentment de la seva textura.

Quan el caràcter E uniforme es pren com a referència (Taula 4.2), els màxims de correlació normalitzats dins la banda de reconeixement apareixen per les lletres E uniforme i ratllada buscades, però també per les dues lletres F, uniforme i ratllada, del conjunt d'escenes test. Apareixen, per tant, falses alarmes en les lletres F. Tot i així, el filtratge POF és el que obté en la Taula 4.2 uns resultats que més s'acosten, sense aconseguir-ho, al tipus de reconeixement que es desitja. En el cas del PPC, el mètode de filtratge és més sensible a les diferències entre els objectes i les falses alarmes desapareixen, però tampoc es possible de reconèixer simultàniament les dues lletres E amb diferent textura (Figures 4.6a i 4.6b).

Quan es pren la lletra E ratllada com a patró (Taula 4.3), els mètodes de filtratge convencionals tampoc aconseguen resoldre el problema plantejat. Els filtres CMF i POF no reconeixen la lletra E uniforme i, de nou, presenten falses alarmes pel cas de la lletra F ratllada. En el cas del PPC, en canvi, no apareixen falses alarmes, però tampoc aconseguen identificar la lletra E uniforme.

Partint d'aquests resultats, si volguéssim assolir el nostre objectiu de reconèixer la lletra E independent de la textura amb que ha estat omplerta, hauríem d'aplicar reiteradament el mètode de filtratge corresponent al PPC, prenent com a referència el caràcter a reconèixer amb cada una de les textures amb què es presenta a l'escena. Intentarem, però, agilitzar aquest procés ja que el que ens interessa és obtenir bons resultats amb tant sols un únic procés de reconeixement.

Si examinem l'espectre de Fourier de la lletra E ratllada del test, que es mostra en representació logarítmica de la Figura 4.6f, podem observar que es tracta de la transformada de Fourier del caràcter E uniforme (Figura 4.6a) que es repeteix d'acord amb la convolució per tres deltes de Dirac (la central corresponent al terme dc i les dues laterals corresponents a la freqüència fonamental del patró de línies verticals) (Eq. 4.14). Per tant, proposem aplicar un mètode de filtratge molt sensible

als petits detalls que mantingui la informació del caràcter en si mateix, però limitat al lòbul central del domini freqüencial per evitar la informació corresponent a la textura periòdica.

La resolució del problema de reconeixement plantejat la portem a terme aplicant el model de la DNC, amb els valors dels paràmetres  $L = M = 0$  (equivalent al PPC) i la funció  $T_1(\vec{v})$  igual a una funció rectangle al llarg de l'eix horitzontal,  $T_1(\vec{v}) = \text{rect}(u)$ , amb  $u$  la coordenada en l'eix horitzontal, que genera una banda vertical d'alta transmitància (veure Figura 4.7a). Per observar la limitació de la regió de suport a l'espai de freqüències introduïda per  $T_1(\vec{v})$ , representem la distribució en intensitat de la funció corresponent al PPC modificat (Figura 4.7b),

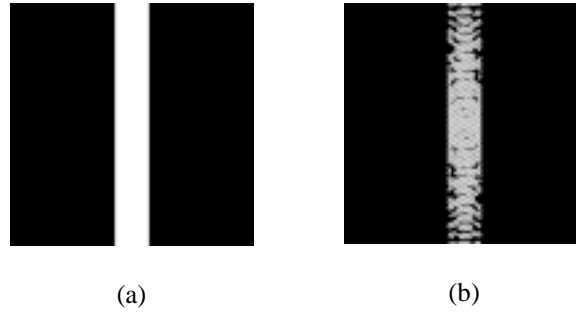
$$FT\{DNC_{0,0}^{T_1(\vec{v})}\{s(\vec{x}), r(\vec{x})\}\} = |S(\vec{v})|^{-1} \cdot S(\vec{v}) \cdot |R(\vec{v})|^{-1} \cdot R^*(\vec{v}) \cdot T_1(\vec{v}). \quad (4.21)$$

Per a la representació logarítmica de la Figura 4.7b s'ha pres en l'Eq. (4.21) la lletra E uniforme (Figura 4.6a) simultàniament com a escena  $s(\vec{x})$  i com a referència  $r(\vec{x})$ . Els resultats de simulació numèrica de la DNC es poden obtenir realitzant el càlcul indicat en l'Eq. (4.3), on intervenen les transformades de Fourier de l'escena a analitzar i de l'objecte de referència. Calculant la transformada de Fourier inversa de l'expressió (4.21), obtenim els resultats de correlació seguint el procés descrit. Aquests resultats s'inclouen en la Taula 4.4.

Observem que, amb el mètode de la DNC, introduint una funció binària senzilla  $T_1(\vec{v})$ , és possible reconèixer el caràcter E independentment de la seva textura, i al mateix temps discriminar-lo d'altres caràcters molt semblants en forma, com la lletra F, tot i tenir el mateix tipus de textura uniforme i periòdica de ratlles verticals que les lletres E buscades. A més, cal destacar, que el procés de reconeixement es realitza de manera satisfactòria tant si considerem la lletra E uniforme com a patró de referència, o bé si utilitzem el mateix caràcter E però recobert amb la textura de



franges. En ambdós casos, el reconeixement es porta a terme amb una única aplicació de la DNC.



**Figura 4.7.** (a) Funció  $T_1(\bar{\nu})$ . (b) Representació logarítmica de  $\left|FT\left\{DNC_{0,0}^{T_1(\bar{\nu})}\{s(\bar{x}), r(\bar{x})\}\right\}\right|^2$  quan la lletra E uniforme (Figura 4.6a) es pren simultàniament com a escena i referència durant el procés de la DNC.

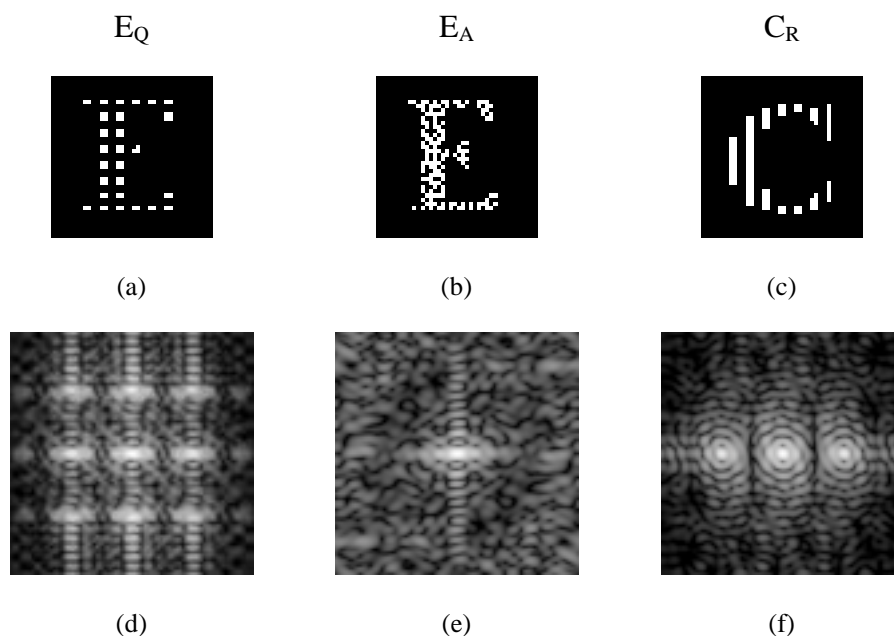
**Taula 4.4.** Resultats de reconeixement per les escenes analitzades de la Figura 4.6 (a)-(d), quan s'aplica el model de la DNC amb la funció  $T_1(\bar{\nu})$  representada en la Figura 4.7a.

Referència	Procediment de la DNC	Correlació normalitzada ( $cc/ac$ )			
		E	E <sub>R</sub>	F	F <sub>R</sub>
E	PPC ( $L = 0, M = 0$ ) i $T_1(\bar{\nu})$	<u>1.00</u>	<u>0.80</u>	0.38	0.30
E <sub>R</sub>	PPC ( $L = 0, M = 0$ ) i $T_1(\bar{\nu})$	<u>0.92</u>	<u>1.00</u>	0.42	0.37

### 4.3.2 Cas de reconeixement II

Considerem ara, un segon cas. Volem reconèixer tots els caràcters inclosos dins un conjunt d'escenes test, que continguin una determinada textura (el patró de franges

verticals). En aquest cas, no tenim en compte la forma del caràcter i només ens fixem en el seu recobriment interior. Per poder analitzar amb major profunditat el problema plantejat, afegim tres lletres amb diferents textures i amb la mateixa energia (Figura 4.8 (a)-(c)) que les del conjunt d'escenes equienergètiques emprat en l'exemple anterior (Figura 4.6(a)-(d)). Les Figures 4.8a i 4.8b mostren la mateixa forma de lletra E però en el primer cas conté una textura quadriculada, i en la Figura 4.8b una distribució de punts aleatòria. La darrera lletra considerada, la Figura 4.8c, conté la mateixa textura de franges verticals que la lletra E de la Figura 4.6b, però amb una forma completament diferent. Les Figures 4.8 (d)-(f) contenen en representació logarítmica, la intensitat de la transformada de Fourier de les lletres de les Figures 4.8 (a)-(c).

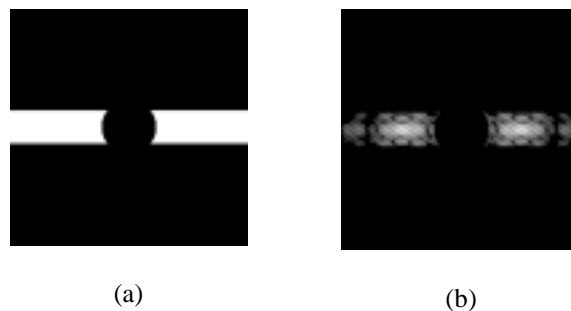


**Figura 4.8.** (a)-(c) Representació binària de les escenes test consistents en caràcters texturats. (d)-(f) Representació logarítmica de llur corresponent transformada de Fourier en intensitat.

Si la tasca de reconeixement consisteix, tal com hem esmentat, a trobar les lletres que contenen una determinada textura, ens convé conservar la informació relativa a la periodicitat de la textura. Sabem que les franges verticals que configuren la textura,

presenten pics d'intensitat corresponents a la freqüència fonamental en la direcció horitzontal de l'espectre i en els successius ordres de difracció múltiples d'aquesta freqüència (veure Figures 4.6f, 4.6h i 4.8f). Si ens proposem identificar la textura independentment del caràcter que la contingui, serà la informació d'aquests pics la que convé que es transmeti. Per contra, serà convenient eliminar la informació de l'ordre zero de difracció, on es contenen les freqüències baixes que defineixen a grans trets la forma de l'objecte, i eliminar també bona part de la informació de freqüències situada en els hemisferis superior i inferior del pla de Fourier.

Per això, en aquest cas proposem utilitzar la funció  $T_2(\vec{v})$  (Figura 4.9a), en el model de la DNC. La funció  $T_2(\vec{v})$  és de la forma  $T_2(\vec{v}) = \text{rect}(v) \cdot [1 - \text{circ}(|\vec{v}|)]$ , amb  $v$  la coordenada en l'eix vertical de l'espai de freqüències. En la Figura 4.9b es mostra la distribució en intensitat de la funció  $FT\{DNC_{1,1}^{T_2(\vec{v})}\{s(\vec{x}), r(\vec{x})\}\}$  corresponent a la lletra E amb textura de franges verticals (Figura 4.6b), considerada simultàniament com a escena i com a referència, prenent valors  $L = M = 1$  (equivalent al CMF), i emprant la funció  $T_2(\vec{v})$ .



**Figura 4.9.** (a) Funció  $T_2(\vec{v})$ . (b) Representació logarítmica de  $|FT\{DNC_{1,1}^{T_2(\vec{v})}\{s(\vec{x}), r(\vec{x})\}\}|^2$  quan la lletra  $E_R$  amb textura de franges verticals (Figura 4.6b) es pren com a escena a analitzar i com a referència durant el procés de la DNC.

La Taula 4.5 conté els resultats de correlació per les diferents lletres contingudes en el conjunt d'escenes test, i considerant diferents paràmetres  $L$  i  $M$  en la DNC conjuntament amb la màscara  $T_2(\vec{v})$ . El millor resultat correspon a la DNC

equivalent al filtre CMF, el qual representa el filtre adaptat menys sensible als petits detalls, entre els mètodes de filtratge estudiats. Aconsegueix el reconeixement de totes les lletres amb la textura de franges verticals, que són l' $E_R$ , l' $F_R$  i la  $C_R$ . També es reconeix el caràcter  $E_Q$ , que constitueix un cas difícil donat que la textura de quadres inclou el patró de franges verticals buscat.

Si analitzem en la Taula 4.5 els resultats de la DNC equivalents al filtratge amb POF observem que el caràcter  $C_R$  texturat amb franges és rebutjat en el procés de reconeixement. Aquest fet, es pot explicar tenint en compte la menor tolerància a variacions de forma del POF en comparació amb el CMF, de manera que només lletres amb la mateixa textura que la buscada i d'igual forma que la referència o d'altres molt semblants, com pot ser la lletra F, són reconegudes pel sistema.

**Taula 4.5.** Resultat de reconeixement per les escenes test de les Figures 4.6 (a)-(d) i 4.8 (a)-(c), quan s'aplica el model de la DNC amb la funció  $T_2(\vec{v})$  representada en la Figura 4.9a. El caràcter  $E_R$  és l'objecte de referència.

Procediment de la DNC	Correlació normalitzada ( $cc/ac$ )						
	E	$E_R$ (ref.)	F	$F_R$	$E_Q$	$E_A$	$C_R$
CMF ( $L=1, M=1$ ) i $T_2(\vec{v})$	0.04	<u>1.00</u>	0.03	<u>0.84</u>	<u>0.94</u>	0.03	<u>0.58</u>
POF ( $L=1, M=0$ ) i $T_2(\vec{v})$	0.06	<u>1.00</u>	0.03	<u>0.83</u>	<u>0.93</u>	0.05	0.33
PPC ( $L=0, M=0$ ) i $T_2(\vec{v})$	0.20	<u>1.00</u>	0.05	0.31	<u>0.90</u>	0.10	0.10

Finalment, el PPC proporciona un reconeixement molt més acurat en quant a l'anàlisi conjunta de la textura i de la forma. Els resultats per aquesta no-linealitat concorden perfectament amb la ben coneguda poca tolerància a petits canvis de

forma del PPC. Així doncs, només la referència i la mateixa  $E_Q$  quadriculada passen el procés de reconeixement i la resta de caràcters són rebutjats.

En conclusió, si volem trobar en una escena una determinada textura sense donar importància a la forma del caràcter, cal recórrer al filtre CMF per la seva tolerància a canvis de forma. En canvi, si el reconeixement ha de ser molt més estricte en el sentit que ens plantegem trobar igual textura i forma simultàniament, cal utilitzar no-linealitats com el PPC. Entre ambdós extrems, el filtre POF proporciona un reconeixement amb una capacitat de discriminació intermitja.

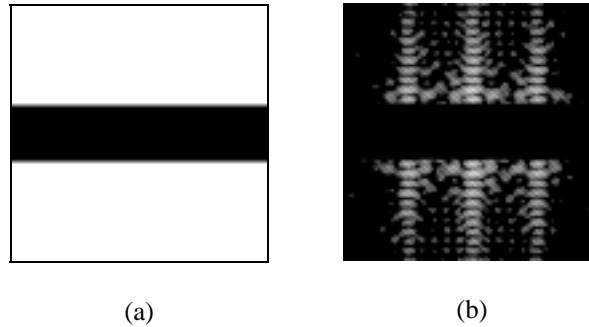
Veiem també en aquest exemple, que l'elecció de la funció  $T(\vec{v})$  ens permet una anàlisi correcta de l'escena segons el problema que hem plantejat, amb només una única aplicació de la DNC.

### 4.3.3 Cas de reconeixement III

Com a últim cas per mostrar les utilitats de la funció anomenada  $T(\vec{v})$ , analitzem quina regió de suport a l'espai de freqüències és l'adequada per discriminar la lletra E amb franges verticals ( $E_R$  en la Figura 4.6b) de la mateixa E però amb quadrícula ( $E_Q$  en la Figura 4.8a).

A l'espai de Fourier, la diferència més rellevant entre ambdues lletres, es posa de manifest fora dels eixos centrals, tant en la direcció horitzontal com vertical (Figures 4.6f i 4.8d). Per tant, és convenient eliminar almenys la franja central horitzontal on estant contingudes la informació de les baixes freqüències de la forma del caràcter, així com la periodicitat de les franges verticals. La Figura 4.10a mostra la funció  $T_3(\vec{v})$  escollida, que és del tipus  $T_3(\vec{v}) = 1 - \text{rect}(v)$ , amb  $v$  la coordenada en l'eix vertical. La Figura 4.10b representa la intensitat de la distribució  $FT\{DNC_{1,1}^{T_3(\vec{v})}\{s(\vec{x}), r(\vec{x})\}\}$  corresponent a la lletra  $E_R$  amb textura de línies verticals

(Figura 4.6b) agafada com a escena i com a referència pel cas de la no-linealitat equivalent al CMF ( $L = M = 1$ ).



**Figura 4.10.** (a) Funció  $T_3(\vec{v})$ . (b) Representació logarítmica de  $\left|FT\{DNC_{1,1}^{T_3(\vec{v})}\{s(\vec{x}), r(\vec{x})\}\}\right|^2$  quan la lletra  $E_R$  amb textura de franges verticals (Figura 4.6b) es pren com a escena a analitzar i com a referència durant el procés de la DNC.

La Taula 4.6 facilita els valors numèrics de la correlació normalitzada  $cc/ac$  en aquest cas, és a dir, per a la no-linealitat equivalent al CMF quan s'aplica el mètode de la DNC, conjuntament amb la funció  $T_3(\vec{v})$ . Els resultats demostren que la discriminació en un cas difícil com el de les dues lletres E amb els mateixos contorns i amb textures diferents, però amb la periodicitat coincident en una de les direccions, també és possible mitjançant aquest mètode.

En resum, al llarg dels tres casos presentats, hem comprovat que amb la DNC, escollint adequadament els paràmetres d' $L$  i  $M$ , així com la forma de la funció de pes  $T(\vec{v})$ , es pot obtenir un reconeixement d'objectes amb un control variable de la seva capacitat discriminativa respecte a possibles patrons de textura regular continguts en els objectes. La simplicitat en el disseny de  $T(\vec{v})$  va associada a la simplicitat del patró de textura en el domini de Fourier.

**Taula 4.6.** Resultat de reconeixement per les escenes test de les Figures 4.6b i 4.8a, quan s'aplica el model de la DNC amb la funció  $T_3(\bar{v})$  representada en la Figura 4.10a. El caràcter  $E_R$  és l'objecte de referència.

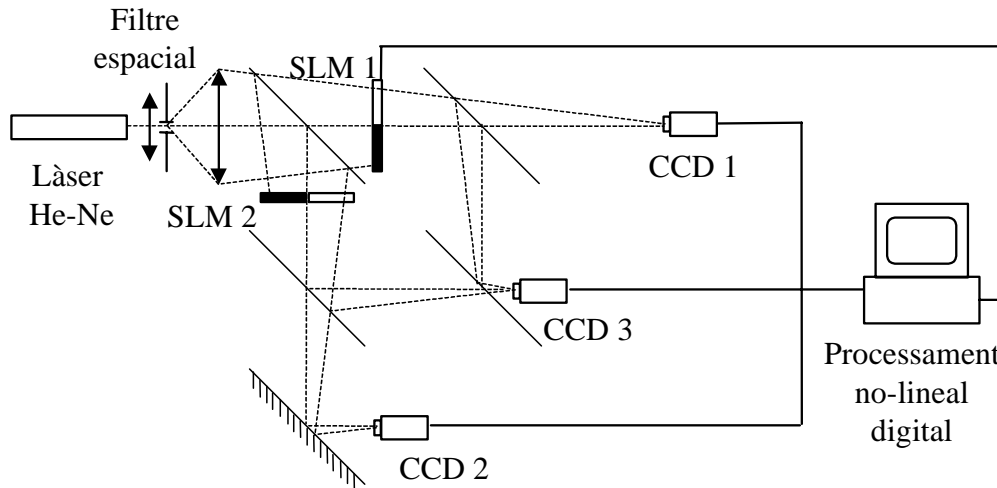
<b>Correlació normalitzada (<math>cc/ac</math>)</b>		
<b>Procediment de la DNC</b>	<b><math>E_R</math> (referència)</b>	<b><math>E_Q</math></b>
<b>CMF (<math>L=1, M=1</math>) i <math>T_3(\bar{v})</math></b>	<u>1.00</u>	1.77

#### 4.4 Realització òptico-digital de la DNC

Les arquitectures de correlador del tipus  $4f$ , o variacions d'aquesta, com pot ser l'arquitectura de correlador convergent, no permeten processar la informació continguda a l'escena un cop ha estat introduïda en el pla d'entrada del correlador. És a dir, aquestes arquitectures es restringeixen als casos on les operacions portades a terme en el procés de reconeixement són lineals amb la intensitat de l'escena d'entrada. En tot cas, amb una arquitectura de correlador  $4f$  o similar, només és possible aplicar operacions no-lineals a l'escena a analitzar si es realitza prèviament un pre-processat de la informació abans d'introduir-la en el pla d'entrada del correlador. Amb el correlador  $4f$ , per tant, l'únic processament es fa a través del filtre adaptat a l'objecte a reconèixer.

En aquest punt, el JTC presenta una clara avantatge respecte les arquitectures esmentades, donat que tant l'escena com la referència es situen, inicialment sense modificar, en el pla d'entrada del correlador i, és en un pas intermig on es porta a terme un processat digital no-lineal, del mateix tipus o, independentment, sobre la informació procedent de l'escena i de l'objecte. En aquest sentit, l'arquitectura escollida per la implementació opto-electrònica de la correlació no-lineal dual, és una configuració basada en el correlador de transformades conjuntes o JTC (Figura 4.11) [42]. Pel fet que la DNC engloba en la seva definició tant operacions lineals

com no-lineals, en aquesta configuració és possible el processament lineal i no-lineal ja sigui a l'escena analitzada, com a l'objecte de referència, o ambdós a la vegada.



**Figura 4.11.** Dispositiu òptic JTC per a la realització de la DNC amb tres càmeres CCD per a la captació simultània de tres transformades de Fourier,  $I_S(\vec{\nu})$ ,  $I_R(\vec{\nu})$  i  $I(\vec{\nu})$ .

El procés complet de reconeixement amb la DNC es porta a terme en dos passos consecutius. En el primer, l'escena i la referència es situen en els moduladors SLM 1 i SLM 2 (Figura 4.11) respectivament, separats en direccions contràries de l'eix òptic per una quantitat  $|\vec{b}|$  (Figura 4.12). Cada modulador correspon a una de les branques en les quals s'ha dividit l'arquitectura bàsica del JTC. Al final de cada una d'elles, hi ha situada una càmera CCD que, suposant que estan ajustades en llur zona de resposta lineal, enregistren les distribucions d'intensitat obtingudes en el pla de freqüències: la CCD 1 capta la transformada de Fourier de l'escena a analitzar,

$$I_S(\vec{\nu}) = |S(\vec{\nu})|^2, \quad (4.22)$$

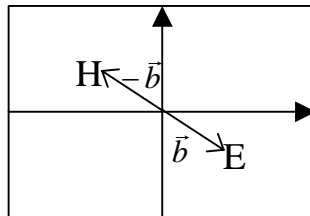
i la càmera CCD 2 enregistra la transformada de Fourier de l'objecte de referència,



$$I_R(\vec{\nu}) = |R(\vec{\nu})|^2. \quad (4.23)$$

A l'esquema de la Figura 4.11 es pot observar com les dues branques descrites pels moduladors espacials, es retroben novament, mitjançant làmines divisòries del feix, en el pla on es situa una tercera càmera. La CCD 3 enregistra la distribució d'intensitat conjunta a l'espai de freqüències de l'escena a analitzar i de l'objecte de referència. Com a resultat de la interferència dels dos feixos provenint de cada un dels moduladors, pot escriure's

$$\begin{aligned} I(\vec{\nu}) &= |FT\{s(\vec{x} + \vec{b}) + r(\vec{x} - \vec{b})\}|^2 = \\ &= |S(\vec{\nu})|^2 + |R(\vec{\nu})|^2 + S(\vec{\nu}) \cdot R^*(\vec{\nu}) \cdot \exp(i4\pi\vec{\nu} \cdot \vec{b}) + \\ &+ S^*(\vec{\nu}) \cdot R(\vec{\nu}) \cdot \exp(-i4\pi\vec{\nu} \cdot \vec{b}). \end{aligned} \quad (4.24)$$



**Figura 4.12.** Disposició de l'escena a analitzar (lletra H a l'esquerra) i de l'objecte de referència (lletra E a la dreta) en el modulador espacial de llum.

El fet d'enregistrar en paral·lel les tres distribucions d'intensitat, a part de permetre realitzar l'operació de reconeixement de manera més ràpida, evita problemes d'estabilitat de la intensitat del làser ja que les imatges es capten simultàniament.

Un cop captades les diferents funcions, representades per les Eqs. (4.22), (4.23) i (4.24), en el pla de freqüències, es realitza el procés no-lineal a través del qual

obtenim una nova distribució que anomenem la distribució d'intensitat conjunta modificada, MJPS (de l'anglès *Modified Joint Power Spectrum*),

$$P(\vec{v}) = [I(\vec{v}) - I_S(\vec{v}) - I_R(\vec{v})] \cdot I_S(\vec{v})^{(L-1)/2} \cdot I_R(\vec{v})^{(M-1)/2} + C, \quad (4.25)$$

on s'inclou una constant C per poder assegurar que la funció  $P(\vec{v})$  sigui positiva i es pugui, per tant, reproduir en el modulador espacial de llum en el segon pas del procés de la DNC. La transformada de Fourier inversa de la funció descrita en l'Eq. (4.25), tenint en compte les expressions (4.22), (4.23) i (4.24), es pot escriure com

$$\begin{aligned} IFT\{P(\vec{v})\} &= \\ &= IFT\left\{ [S \cdot R^* \cdot \exp(i4\pi\vec{v} \cdot \vec{b}) + R \cdot S^* \cdot \exp(-i4\pi\vec{v} \cdot \vec{b})] \cdot \right. \\ &\quad \left. \cdot |S|^{(L-1)} \cdot |R|^{(M-1)} + C \right\} = \\ &= DNC_{L,M}\{s, r\} * \delta(\vec{x} + 2\vec{b}) + \\ &\quad + DNC_{M,L}\{r, s\} * \delta(\vec{x} - 2\vec{b}) + \\ &\quad + C \cdot \delta(x, y). \end{aligned} \quad (4.26)$$

Els dos primers sumands de l'Eq. (4.26), ambdós fora d'eix, corresponen als senyals de sortida de l'operació DNC. En concret, aquests dos termes són simètrics i estan separats una distància  $2\vec{b}$  en direccions oposades respecte la funció delta central, donat que en el JTC sempre hi cap la possibilitat d'interpretar el pla d'entrada del correlador intercanviant els papers de l'escena a analitzar i de l'objecte de referència. Així doncs, de manera simètrica a l'ordre zero del pla de sortida, apareixen sempre dos termes idèntics corresponents a la correlació de l'escena amb la referència i viceversa. El primer d'ells, ens facilitarà el resultat de la correlació no-lineal dual. El

darrer terme de l'Eq. (4.26) consisteix en una funció delta de Dirac situada en eix que correspon al terme d'ordre zero del pla de sortida del correlador. El fet de ser una funció delta permet assegurar que no existiran problemes degut al solapament d'aquesta funció amb els termes fora d'eix de la DNC. En aquest cas, veiem que la captació de la informació de les tres distribucions per separat (Eqs. (4.22), (4.23) i (4.24)), permet un processat digital intermig que porta a la minimització de l'amplada del terme central del pla de sortida del processador. D'aquesta manera, s'obté a més, un millor senyal de correlació en la posició del primer ordre de difracció.

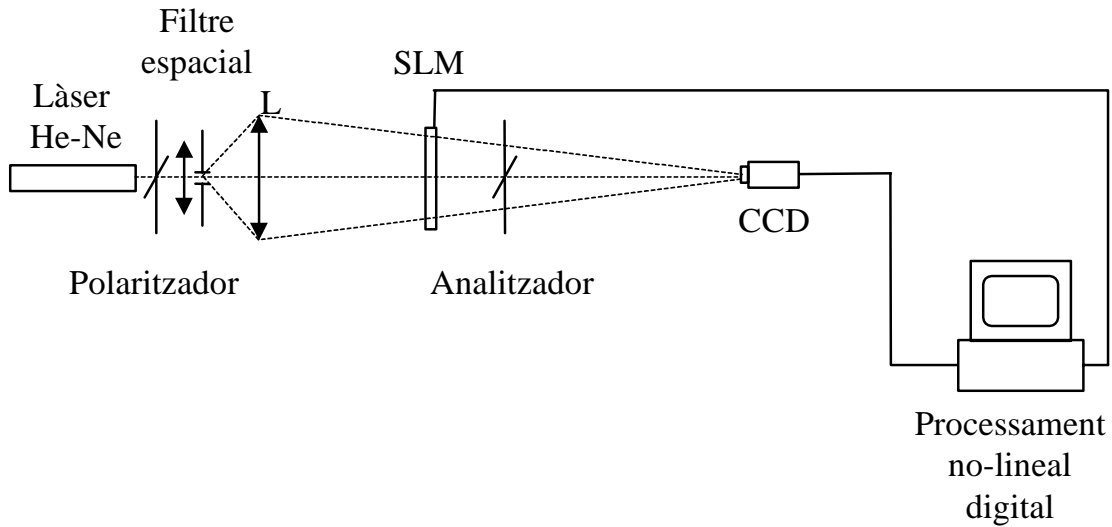
Cal fer esment, també, que donat que òpticament no es pot obtenir la transformada de Fourier inversa, el procediment a seguir en el segon pas experimental de la DNC és reproduir la funció  $P(\vec{v})$  en el modulador indicat com SLM 1 en la Figura 4.11, i enregistrar la seva transformada de Fourier directa mitjançant la càmera CCD 1. La relació d'aquesta amb la TF inversa ve donada per una inversió de la funció, que es pot neutralitzar per un canvi de sistema de coordenades ( $\vec{x} \rightarrow -\vec{x}$ ).

Remarquem doncs, que amb aquest dispositiu experimental (Figura 4.11), podem obtenir la correlació no-lineal dual i, per tant, qualsevol altre tipus de correlació definida a partir d'una operació de filtratge inclosa en la DNC. Per aconseguir-ho caldrà variar simplement els paràmetres  $L$  i  $M$  del processament no-lineal intermig. El fet que el dispositiu experimental es basi en l'arquitectura del correlador de transformades conjuntes, afegeix al sistema l'avantatge de poder obtenir diferents processaments de reconeixement en un mateix correlador, sense la necessitat d'utilitzar, ni posicionar, filtres adaptats per cada situació. Una dificultat, però, d'aquest dispositiu està relacionada amb la operació de substracció de funcions  $I(\vec{v})$ ,  $I_S(\vec{v})$ ,  $I_R(\vec{v})$ , quan aquestes es reproduïen en dispositius diferents amb la possibilitat que la resposta d'aquests no coincideixi exactament i la il·luminació de cada un d'ells, també pot tenir variacions. Per tant, cal controlar de manera precisa aquestes magnituds i tenir en compte la possible necessitat de compensació de les diferències entre les funcions.

#### 4.4.1 Dispositiu i processament simplificat

El dispositiu experimental (Figura 4.11) i les operacions involucrades en el processament digital (Eq. 4.25) que hem descrit en l'apartat anterior, poden ser simplifiades notablement sense afectar de manera significativa els resultats de correlació. Per una banda es pot reduir el nombre d'elements constituents del muntatge òptic, reduint-ne per tant el seu cost i facilitant-ne la seva alineació, i per altra banda és possible una reducció en el temps de càlcul durant el processat no-lineal intermig. Ambdues simplifiacions han estat presentades a [202].

La Figura 4.13 mostra l'esquema del correlador de transformades conjuntes simplificat per a realitzar la DNC, on només s'ha utilitzat un sol modulador espacial de llum i una única càmera CCD. El primer pas de la DNC és el que es veu més modificat, en el sentit que ara l'escena i la referència són reproduïts conjuntament en el mateix modulador. Escena i referència es presenten separats de l'eix òptic en direccions contràries una mateixa distància  $|\vec{b}|$  en l'SLM (Figura 4.12). A la CCD s'enregistra la distribució d'intensitat de la transformada de Fourier conjunta ( $I(\vec{\nu})$ ) com a resultat de les interferències produïdes pels raigs que travessen ambdues imatges. Posteriorment, i de manera seqüencial, es mostra només l'escena a analitzar per captar la corresponent distribució d'intensitat de la seva transformada de Fourier ( $I_S(\vec{\nu})$ ) i, per últim, es mostra en l'SLM l'objecte de referència, captant amb la càmera la tercera i última distribució espectral ( $I_R(\vec{\nu})$ ). Utilitzant aquesta arquitectura simplifiada de correlador no-lineal dual, cal doncs assegurar el mateix nivell d'il·luminació uniforme en el pla d'entrada del correlador al llarg de tot el procés de captació.



**Figura 4.13.** Dispositiu simplificat JTC per a l'obtenció experimental de la DNC.

Per altra banda, una segona modificació pot ser introduïda en l'algorisme del procés no-lineal per reduir el temps de càlcul. En lloc de l'expressió (4.25) com a punt de partida per a l'obtenció de la DNC, utilitzarem una altra expressió,

$$P'(\vec{\nu}) = I(\vec{\nu}) \cdot I_S(\vec{\nu})^{(L-1)/2} \cdot I_R(\vec{\nu})^{(M-1)/2}. \quad (4.27)$$

Aquesta expressió (4.27) també permet obtenir la DNC com veurem a continuació, i a més, assegura valors no-negatius per a la distribució d'intensitat de l'espectre conjunt modificat no-linealment,  $P'(\vec{\nu})$ , més fàcilment reproduïbles en el modulador espacial de llum. Calculant la transformada de Fourier inversa de l'expressió anterior (4.27), s'obté

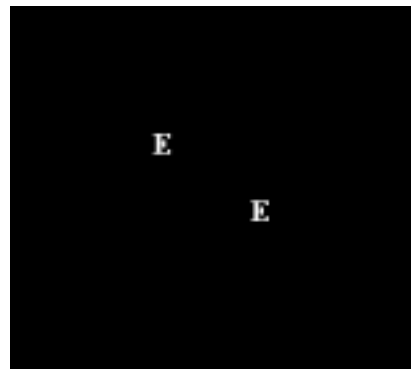
$$\begin{aligned}
IFT\{P'(\vec{v})\} &= \\
&= IFT\left\{\left[|S|^2 + |R|^2 + S \cdot R^* \cdot \exp(i4\pi\vec{v} \cdot \vec{b}) + R \cdot S^* \cdot \exp(-i4\pi\vec{v} \cdot \vec{b})\right] \cdot |S|^{(L-1)} \cdot |R|^{(M-1)}\right\} = \\
&= IFT\left\{|S|^{(L+1)} \cdot |R|^{(M-1)} + |S|^{(L-1)} \cdot |R|^{(M+1)}\right\}_+ \tag{4.28} \\
&+ DNC_{L,M}\{s(x, y), r(x, y)\} * \delta(\bar{x} + 2\bar{b}) + \\
&+ DNC_{M,L}\{r(x, y), s(x, y)\} * \delta(\bar{x} - 2\bar{b}).
\end{aligned}$$

Observem doncs, que en l'expressió (4.28) s'obtenen els dos termes de la DNC, com en l'expressió (4.26) anterior, més un terme central. Aquest terme central és diferent i, en general, més ample que el que s'obtenia en l'expressió (4.26). Els termes de la correlació no-lineal dual no es veuen afectats pels canvis introduïts en l'algorisme, però sí que és necessari controlar l'amplada del terme en eix (primer sumand de l'Eq. 4.28) per tal d'evitar possibles solapaments amb els termes fora d'eix, amb la conseqüent alteració dels resultats de correlació.

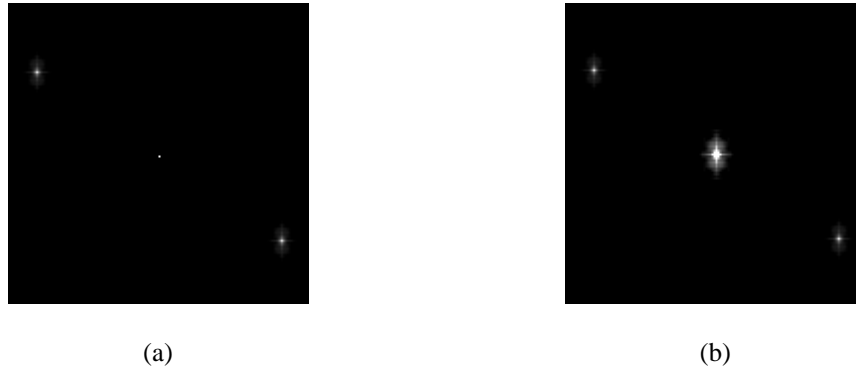
Per a facilitar la comprensió de les implicacions que comporta l'ús alternatiu de les funcions  $P(\vec{v})$  o  $P'(\vec{v})$  en el resultat que s'obté en el pla de sortida, considerem un exemple on volem reconèixer la lletra E en una escena on només hi ha present aquest mateix caràcter. La Figura 4.14 mostra el pla d'entrada del correlador no-lineal dual, on l'escena a analitzar es situa en la zona superior esquerra del pla, i l'objecte a reconèixer es troba en la zona inferior dreta de la imatge. Mitjançant simulació numèrica, reproduïm els resultats obtinguts a partir dels dos algorismes presentats per a realitzar la DNC. En el càlcul hem considerat una DNC amb  $L = 1$  i  $M = 1$  equivalent al cas del CMF. La Figura 4.15a correspon al pla de sortida del JTC en el cas de calcular l'espectre de transformades conjuntes modificat no-linealment,  $P(\vec{v})$ , a través de l'expressió (4.25). Dins d'aquesta expressió es realitza la substracció de les distribucions d'intensitat de l'escena,  $I_S(\vec{v})$ , i de l'objecte de referència,  $I_R(\vec{v})$ , de la distribució d'intensitat conjunta  $I(\vec{v})$ . Observem com el terme central està molt

ben delimitat i és estret tal i com correspon a una funció delta de Dirac (Eq. 4.26). Si el processat digital no-lineal es realitza a partir de l'Eq. (4.27), sense fer la substracció de les funcions, s'obté la Figura 4.15b en el pla de sortida del correlador. En aquest cas, l'ordre zero s'eixampla ja que correspon a la transformada de Fourier inversa d'una funció de més amplada de banda en freqüències (primer sumand de l'Eq. 4.28). Pel fet d'ocupar una certa amplada, és necessari assegurar que aquest terme no s'està solapant als resultats de la DNC en el pla de correlació obtingut (termes que s'observen a l'esquerra i a la dreta de les Figures 4.15a i 4.15b, simètricament respecte el terme central).

Aquest segon algorisme simplificat presenta l'avantatge addicional d'ocupar un període de temps més reduït en el càlcul digital. Afegit al fet que no és necessari realitzar la substracció entre funcions, que la distribució MJPS ( $P'(\vec{v})$ ) obtinguda després del procés no-lineal no conté valors negatius i, valorant altres avantatges que comentarem en el capítol d'anàlisi de les condicions experimentals, farà que sigui utilitzat pels càlculs posteriors en els diferents estudis i aplicacions presentats en aquest treball.



**Figura 4.14.** Exemple de pla d'entrada del processador on l'escena a analitzar només conté el caràcter de referència. L'escena a analitzar es situa a la meitat esquerra, i la referència a la meitat dreta.



**Figura 4.15.** Pla de sortida del processador quan la Figura 4.14 es reproduïda a l'SLM en el primer pas de la DNC, i la no-linealitat aplicada correspon al cas del CMF. En (a) s'ha calculat l'espectre conjunt modificat no-linealment  $P(\vec{\nu})$  de l'Eq. 4.25, i en (b) s'ha utilitzat l'expressió (4.27) corresponent a  $P'(\vec{\nu})$ .





## **Capítol 5**

# **Anàlisi numèrica de les condicions experimentals per a la realització de la DNC**

### **5.1 Introducció**

En el capítol anterior, hem proposat un sistema opto-electrònic simplificat (Figura 4.13, en la secció 4.4.1) per a la realització experimental de la DNC. En aquest dispositiu s'han tingut en compte principalment dos elements optoelectrònics: la càmera CCD, que s'utilitza tant per a l'enregistrament de les distribucions d'intensitat espectrals necessàries en el processament no-lineal de la DNC, com també per l'enregistrament de la correlació final d'on s'extreuen els resultats de reconeixement. El segon element optoelectrònic present en el dispositiu, és la

pantalla de cristall líquid actuant com a modulador espacial de llum. La seva utilitat, com ja s'ha esmentat en el capítol anterior, es centra d'una banda, en el primer pas de la DNC, amb la reproducció de l'objecte a reconèixer i de l'escena a analitzar en el pla d'entrada del correlador i, per altra banda, en el segon pas de la DNC, en la reproducció de l'espectre de potències conjunt modificat (MJPS, de l'anglès *Modified Joint Power Spectrum*). Aquests dos elements optoelectrònics, tenen un paper clau en el correlador, i de les seves característiques de treball podrien dependre els resultats que s'obtenen en el procés de reconeixement complet. Per aquest motiu, és molt important analitzar les possibilitats que ofereixen aquests elements, així com les limitacions que poden introduir en el sistema. En altres paraules, l'estudi de les condicions experimentals introduïdes pels elements optoelectrònics inclosos en el sistema òptic proposat per a la realització de la DNC, ens pot permetre conèixer les limitacions reals de la proposta teòrica realitzada en el capítol 4.

En la literatura, trobem treballs on es mostren les possibles limitacions del sistema òptic donades per dispositius, com càmeres CCD i moduladors espacials de llum. En aquest sentit, Purwosumarto i Yu [138], estudien les limitacions de dos tipus d'arquitectures: el correlador de Vander Lugt o  $4f$  (VLC), i el correlador de transformades conjuntes (JTC), en quant a la degradació dels resultats de correlació provocats per la desalineació transversal i longitudinal d'alguns elements del sistema. Per aquesta anàlisi, Purwosumarto i Yu tenen en consideració les característiques dels sistemes de captació i dels moduladors espacials de llum, en diferents marques i models accessibles actualment en el mercat. Bàsicament, introdueixen paràmetres relacionats amb la resolució espacial d'aquests sistemes, com pot ser el nombre de píxels i llur tamany. Les conclusions a les quals s'arriba en aquest estudi permeten dir que el JTC és més robust als desalineaments de l'SLM emprat en el pla de freqüències que el VLC, donat que, en quant a desalineació transversal només el correlador de Vander Lugt veu afectats els resultats de correlació de manera important, mentre que el JTC no depèn en cap manera d'aquesta condició. En canvi, si es consideren desalineaments longitudinals de l'SLM, és a dir al llarg de l'eix òptic, ambdós correladors veuen degradats els resultats de correlació, i en major

mesura en el cas del JTC. De totes maneres, la tolerància en la posició del modulador en el JTC, està al voltant d'1 mm, distància factible d'ésser controlada a la pràctica.

Treballs anteriors de Javidi i col·laboradors es centren en alguns condicionaments experimentals que poden afectar els resultats del correlador de transformades conjuntes binari (BJTC, de l'anglès *Binary Joint Transform Correlator*) [14,158].

En el BJTC s'enregistra l'espectre d'intensitat conjunt (JPS) en el pla de Fourier del correlador mitjançant una càmera CCD. La mitjana de la distribució enregistrada s'estableix com a valor llindar per a la binarització del JPS i, posteriorment, amb una segona transformada de Fourier d'aquesta funció binaritzada, s'obté el pla de sortida del correlador amb els característics pics d'intensitat que detecten i localitzen l'objecte de referència en el cas de ser present a l'escena que s'analitza. Durant l'enregistrament del JPS, en el pas previ a la binarització, existeix una limitació en els nivells de quantització en els quals s'enregistra la informació, donada per l'efecte conjunt del rang dinàmic de la càmera CCD i de la precisió en bits de la targeta digitalitzadora o interfase entre la càmera i el modulador.

Javidi i col·laboradors han estudiat l'efecte de la quantització de la informació en el procés d'enregistrament de les distribucions d'intensitat [203,204], i de la compressió no-lineal de la informació deguda a la saturació dels sensors de la càmera CCD [204]. En [203] s'avaluen característiques com la intensitat del pic d'autocorrelació, la seva relació amb la intensitat dels lòbuls principals que apareixen al voltant del pic i l'amplada del mateix pic d'autocorrelació. Aquestes característiques s'utilitzen per determinar la qualitat dels resultats de reconeixement i comparar-los amb els resultats del correlador de transformades conjuntes lineal o clàssic (apartat 2.2.3 del capítol 2) [2,3]. El BJTC es veu afectat per la quantització en el pla de Fourier de manera que requereix un nombre mínim de nivells en l'enregistrament del JPS per a poder obtenir els pics de correlació en el pla de sortida del dispositiu. Si el nombre de nivells de quantització és superior al mínim esmentat, els resultats de correlació s'acosten en gran mesura als obtinguts per un BJTC teòric que no inclogui la quantització en l'enregistrament de l'espectre d'intensitat conjunt. Els resultats del BJTC es comparen amb els obtinguts amb el JTC lineal, i es posa de manifest com

aquest darrer, veu disminuir la qualitat dels pics de correlació a mesura que es redueix el nombre de nivells de quantització del sistema d'adquisició.

Javidí i Wang [204] realitzen una descripció matemàtica de l'efecte de la quantització de l'espectre d'intensitat conjunt en el JTC binari, i s'utilitza el paràmetre de la relació senyal-soroll (SNR) com un indicatiu de la qualitat dels resultats de correlació. En el mateix treball, es porta a terme un estudi de l'efecte d'una compressió no-lineal de la informació enregistrada, deguda a la saturació dels sensors de la càmera CCD durant el procés de captació. Es demostra que, en presència de quantització en el pla de Fourier, la saturació permet disminuir el nombre mínim de nivells de quantització requerit per a l'obtenció dels pics de correlació.

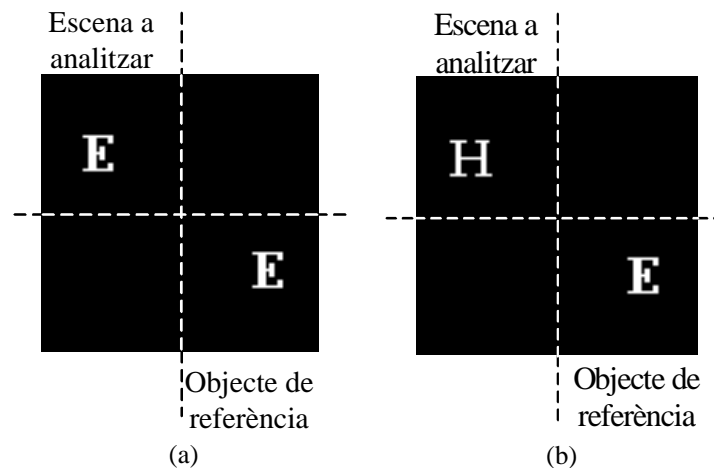
En aquest capítol, i seguint aquesta línia, el nostre treball es centrarà en l'anàlisi d'alguns aspectes experimentals que poden condicionar els resultats de la DNC que s'obtinguin en un JTC. Concretament, considerarem primer la possibilitat d'admetre un nivell de saturació de les distribucions espectrals enregistrades per la càmera CCD. Aquesta saturació afectarà, per tant, al MJPS calculat segons l'expressió simplificada (Eq. 4.27) i, en conseqüència, també afectarà als resultats de la DNC. Estudiarem les variacions de la DNC en funció del nivell de saturació i establirem el nivell més adient (Secció 5.2).

A continuació, afegirem l'efecte de la quantització de la càmera CCD segons els nivells de gris accessibles del sistema de captació. Aquesta quantització afecta en gran mesura a l'enregistrament de les distribucions espectrals que serveixen per calcular el MJPS. Establirem, en la secció 5.3, les condicions de quantització més idònies tenint en compte les característiques tècniques del dispositiu optoelectrònic del qual disposarem en la realització experimental.

En la secció 5.4, mostrarem la possibilitat de modificar la regió de suport de les distribucions espectrals i, per tant, del MJPS, a través d'un paràmetre ( $\varepsilon$ ) involucrat en el càlcul no-lineal de l'Eq. (4.27). Escollirem, a partir dels resultats obtinguts, el valor d' $\varepsilon$  més convenient en la implementació de la DNC.

Per últim, considerarem l'efecte de la quantització introduïda per l'SLM (secció 5.5). Aquest dispositiu ha de mostrar, en el primer pas de la DNC, l'escena a analitzar i l'objecte de referència. Primerament, tractarem amb escenes binàries, on la quantitat de nivells de quantització no influirà en la realització del primer pas de la DNC. No obstant, durant el segon pas de la DNC, caldrà tenir en compte que es vol reproduir un espectre amb un rang d'intensitats molt ampli. Això, exigiria disposar, dins la zona de resposta lineal de l'SLM, d'un nombre suficient de nivells de gris del dispositiu per tal de tenir una reproducció fiable del MJPS en el correlador. Més endavant, també es treballarà amb escenes i objectes amb nivells de gris.

Iniciarem l'anàlisi digital partint d'escenes senzilles, segmentades, binàries i lliures de soroll. Ens centrarem en les escenes a analitzar que es mostren en la Figura 5.1, però l'estudi és vàlid, en general, per a qualsevol forma d'objecte. Les escenes estan constituïdes per caràcters de diferent forma i nivell de gris uniforme. Hem escollit les lletres E i H com exemple representatiu d'una situació general. En tots els càlculs numèrics, s'ha utilitzat sempre l'algorisme descrit per l'Eq. (4.27), corresponent al càlcul simplificat de l'espectre conjunt modificat,  $P'(\vec{v})$ .



**Figura 5.1.** Plans d'entrada del correlador no-lineal dual: (a) per l'estudi de l'autocorrelació de la lletra E; (b) per l'estudi de la correlació creuada de la lletra E amb la lletra H.

En molts casos ens interessarà estudiar les característiques del pic d'autocorrelació i la seva evolució amb els altres paràmetres que s'aniran variant. L'autocorrelació l'obtindrem quan la Figura 5.1a es mostri en l'SLM durant el primer pas del processament. Un paràmetre que ens permet fer-ne una valoració numèrica és el PCE (apartat 2.3.3 del capítol 2), que es pot considerar com un criteri estàndard entre els que més s'utilitzen per a l'avaluació del comportament dels diferents filtres de correlació òptics [6]. Recordem que el PCE es defineix per l'expressió (Eq. 2.13),

$$PCE = \frac{\text{energia del pic de correlació}}{\text{energia del pla de correlació}} = \frac{|c(\vec{0})|^2}{\iint |c(\vec{x})|^2 d\vec{x}}. \quad (5.1)$$

El PCE dóna informació de la distribució d'energia en el pla de correlació de sortida i es pot considerar com una mesura de l'agudesesa del pic de correlació. Si la referència està present a l'escena, aquesta serà detectada amb facilitat i localitzada amb precisió a través d'un pic d'autocorrelació estret que sobresurti clarament d'un fons energèticament baix. En aquesta circumstància, el PCE s'acostarà al valor unitat. Si pel contrari, el pic de correlació té una amplada considerable o bé existeix un fons energètic important, el valor obtingut pel paràmetre PCE serà proper a zero. En aquest cas el pic de correlació no permetrà la detecció de l'objecte dins a l'escena, ni la seva discriminació respecte d'altres objectes.

En el cas de la DNC, si tenim en compte el teorema de Parseval [127] per al càlcul de l'energia total del pla de correlació, el paràmetre PCE adquireix l'expressió [42]

$$PCE = \frac{\left| \iint N_L[S] \cdot N_M[R^*] \cdot T d\vec{v} \right|^2}{\iint |N_L[S] \cdot N_M[R^*] \cdot T|^2 d\vec{v}}. \quad (5.2)$$

## 5.2 Saturació de la càmera CCD

El primer punt on intervé la càmera CCD en la correlació no-lineal dual, és en l'enregistrament de les tres distribucions d'intensitat,  $I$ ,  $I_S$  i  $I_R$ . En general, aquestes funcions presenten un rang d'intensitats molt ample de manera que el seu enregistrament lineal des de les freqüències més baixes fins a les altes freqüències resulta difícil amb el rang dinàmic en nivells de gris assequible actualment pels dispositius CCD. Si es vol captar linealment les baixes freqüències és inevitable una pèrdua d'informació en les freqüències més altes, i viceversa. Aquest inconvenient ja s'ha intentat resoldre amb diferents propostes descrites a la literatura [161,39], i podem dir que és força comú intentar ressaltar la informació continguda en les altes freqüències, donat que, tot i ésser normalment poc energètica, conté informació qualitativament rellevant que es vol preservar. Així doncs, és habitual voler enregistrar una major quantitat d'informació en la zona d'altres freqüències tot i perdre la informació de la part central de l'espectre, i per tant de les baixes freqüències, acceptant un cert nivell de saturació d'aquestes.

En la nostra proposta, on és necessari enregistrar tres distribucions d'intensitat en el pla de Fourier ( $I_S$ ,  $I_R$  i  $I$ ), ens cal incloure un nivell de saturació de les funcions per poder conservar la informació de les altes freqüències de les tres distribucions. Donat llur alt rang de nivells d'intensitat, cal saturar les freqüències baixes molt energètiques, i enregistrar així linealment una zona mitja de l'espectre.

La saturació de la càmera CCD en el procés d'enregistrament d'imatges, es pot introduir intencionadament augmentant la intensitat dels punts de la imatge. Una manera d'aconseguir aquest propòsit en un sistema òptic és regulant la intensitat del feix uniforme de llum que il·lumina el sistema mitjançant un filtre neutre de transmissió variable. Variant la transmissió en intensitat del filtre s'aconsegueix resituar els valors de la distribució d'intensitat que es pretén captar dins el rang dinàmic de resposta del sensor de la càmera. A l'augmentar la intensitat transmesa pel filtre, s'aconsegueix augmentar l'energia de tots els píxels constituents de la imatge enregistrada, de manera que és possible captar un nivell no-nul en regions on

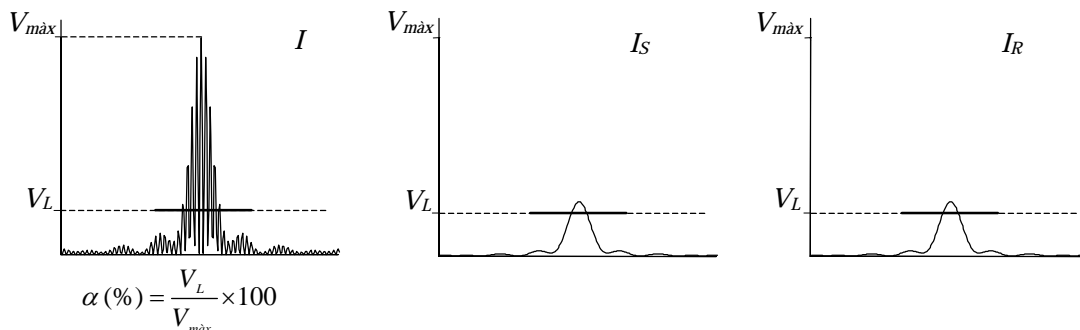


no hi havia resposta inicialment. A més, es pot provocar un enregistrament no-lineal amb saturació en altres zones de la imatge on s'assoleix i es supera el nivell màxim de resposta lineal de la CCD. Per tant, en aquestes zones saturades, es perd la informació de la intensitat relativa amb la resta de punts de la imatge.

En aquest estudi, volem establir el nivell de saturació, comú per les tres distribucions enregistrades, que sigui més convenient per a la millora dels resultats de correlació amb la DNC. El nivell de saturació escollit el determinem a partir del que anomenem llindar de saturació, que ve representat pel paràmetre  $\alpha$  ( $0\% \leq \alpha \leq 100\%$ ), i fa referència al valor  $V_L$  a partir del qual es trunquen les funcions  $I$ ,  $I_S$  i  $I_R$ , expressat en percentatge respecte el valor màxim  $V_{m\grave{a}x}$  del conjunt de les tres distribucions  $I$ ,  $I_S$  i  $I_R$  (Figura 5.2),

$$\alpha(\%) = \frac{V_L}{V_{m\grave{a}x}} \times 100. \quad (5.3)$$

En general, el valor màxim ( $V_{m\grave{a}x}$ ) pertany a la distribució d'intensitat de l'espectre conjunt  $I$ . Aquells punts de les distribucions que tinguin valors d'intensitat superiors a  $V_L$  estaran saturats i modificaran la seva intensitat pel valor màxim permès de nivell de gris (255 en el nostre cas). Els punts amb intensitat inferior al llindar de saturació es repartiran proporcionalment els valors de nivell de gris assequibles pel rang dinàmic de la càmera ( $2^8$  nivells).



**Figura 5.2.** Procés de saturació, caracteritzat pel paràmetre  $\alpha$ , de les distribucions  $I$ ,  $I_S$  i  $I_R$ .

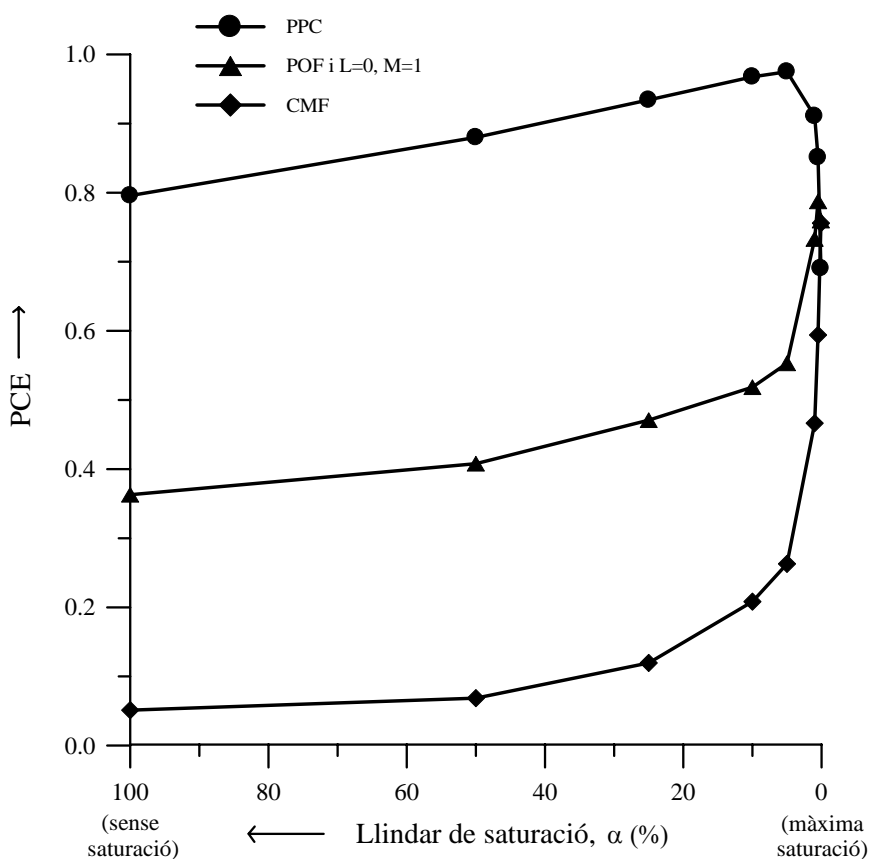
En la Figura 5.3 representem, en el cas d'autocorrelació de la lletra E (Figura 5.1a) la variació del PCE en funció de la variació del nivell de saturació introduït en les funcions  $I$ ,  $I_S$  i  $I_R$  al ser enregistrades per la càmera CCD. Estudiem l'efecte de la saturació en quatre tipus de processament: el filtre clàssic o CMF amb paràmetres  $L=1$  i  $M=1$ , el filtre només de fase o POF amb  $L=1$  i  $M=0$ , el processat simètric al POF, és a dir amb  $L=0$  i  $M=1$ , i per últim, la correlació només de fase o PPC amb  $L=0$  i  $M=0$ .

En la Figura 5.3, el llinard  $\alpha$ , representat en l'eix d'abscisses (creixent de dreta a esquerra), s'ha fet variar entre un valor màxim d' $\alpha=100\%$  quan les funcions no són truncades per l'efecte de la saturació, fins a un valor mínim del 0.1%, quan s'assoleix un efecte màxim de la saturació de les distribucions. En l'eix d'ordenades de la Figura 5.3 es mostra el valor del PCE assolit en cada cas. Cal remarcar que, degut al fet que l'escena a analitzar només conté l'objecte a reconèixer, els resultats del filtre POF, i del seu simètric amb  $L=0$  i  $M=1$ , coincideixen per tots els valors de saturació considerats.

En general, el fet de disminuir el llinard a partir del qual es saturen les funcions, fa augmentar el PCE i, per tant, millora la detecció i localització de la referència. Per valors d' $\alpha$  entre 100% i 10%, aquest comportament és comú pels quatre processaments no-lineals portats a terme mitjançant l'algorisme de la DNC, tot i que els valors absoluts del PCE difereixen per cada un dels processaments. Sempre s'obté un valor del PCE superior en el cas del PPC, que dóna pics de correlació més alts i estrets. En el cas del POF o el seu simètric, s'obtenen valors intermedis del PCE. Finalment, el filtre CMF és el que proporciona uns pics més amples i que menys sobresurten sobre el fons energètic del pla de correlació i, per tant, aconsegueix els valors més baixos del PCE.

En particular pel cas del PPC, a mesura que la saturació augmenta i, en conseqüència, disminueix el llinard de tall  $V_L$  de les funcions, s'assoleix un valor màxim del PCE quan el paràmetre  $\alpha$  està al voltant del 5%. A partir d'aquest nivell, si la saturació continua augmentant, el PCE empitjora i disminueix fins a un valor on

les diferents no-linealitats aplicades conflueixen. Això succeeix quan  $\alpha \approx 0.1\%$ . En aquest moment podem dir que el sistema perd la sensibilitat a les no-linealitats aplicades.

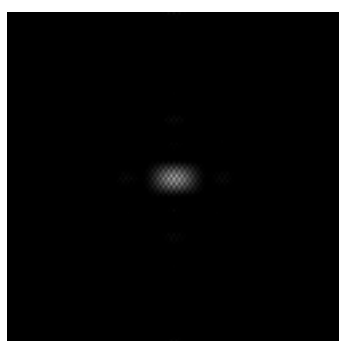


**Figura 5.3.** Variació del PCE en funció del llindar  $\alpha$  de saturació de les funcions enregistrades per la càmera CCD quan el pla d'entrada és el de la Figura 5.1a.

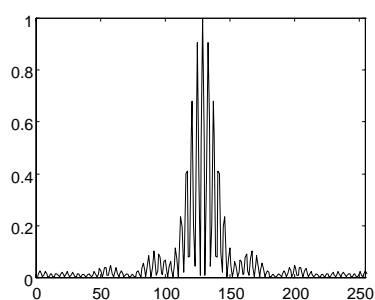
Els resultats descrits fins ara, es poden entendre més fàcilment observant l'efecte de la saturació directament en el pla de Fourier sobre la distribució d'intensitat conjunta,  $I$ . La Figura 5.4a mostra una representació 2D d'aquesta distribució, que correspon al cas on la Figura 5.1a es reproduïx en el pla d'entrada del correlador. La Figura 5.4b mostra una secció longitudinal de l'anterior coincidint amb el seu màxim central. La Figura 5.4c representa en escala logarítmica la mateixa informació que en la Figura 5.4a per facilitar la visualització de la funció  $I$ .

En general, una distribució com la representada té un alt rang de nivells d'intensitat. A l'intentar enregistrar-la linealment mitjançant una càmera CCD provoca una pèrdua d'informació en alguna zona de l'espai de freqüències. Si la zona central es capta correctament, els petits canvis que es donen entre punts propers en les freqüències més elevades, no es poden enregistrar degut a la insuficiència de nivells de quantització per a ésser discriminats.

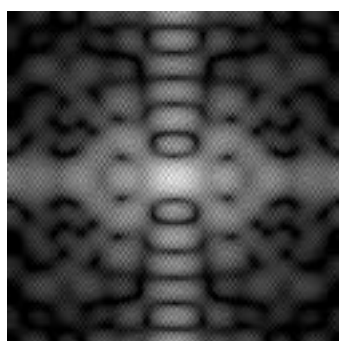
Com hem dit, una manera d'obtenir més informació corresponent a la regió de les altes freqüències, és permetre un cert nivell de saturació durant l'enregistrament de les funcions. D'aquesta manera aconseguim captar més informació procedent de les freqüències altes de l'espectre, però, per contra perdem informació de les zones saturades, normalment les baixes freqüències. Per mostrar-ne un exemple, les Figures 5.5, 5.6 i 5.7 mostren representacions equivalents a la Figura 5.4 considerant diferents llindars de saturació. La Figura 5.5 representa el cas de considerar un llindar de saturació  $\alpha = 50\%$ , la Figura 5.6 correspon al valor d' $\alpha = 10\%$  i, per últim, la Figura 5.7 considera  $\alpha = 1\%$ . El contrast entre la zona central i la corona més externa disminueix pel fet de truncar la funció. Podem observar com el fet de perdre informació en la regió de les baixes freqüències, es contraresta amb el reforçament aconseguït en les freqüències més altes. És més fàcil enregistrar més informació dels canvis de valors entre punts consecutius de les altes freqüències. Això explica la millora dels resultats del PCE a mesura que augmentem l'efecte de la saturació en la captació de les funcions. En el cas on la saturació és més extrema, per  $\alpha = 1\%$ , l'amplificació de les altes freqüències és molt gran, de manera que s'aconsegueix pràcticament la binarització de la distribució o *hard clipping* (Figura 5.7). En aquest moment, l'efecte és pràcticament el mateix per totes les no-linealitats, de manera que podem dir que la variació en el funcionament del sistema segons la no-linealitat aplicada s'esvaeix i s'obté un resultat comú a totes elles.



(a)

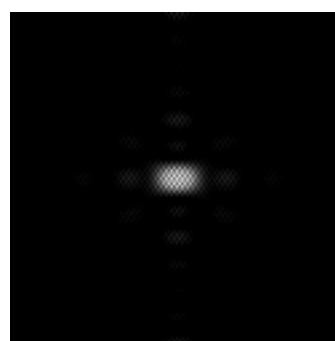


(b)

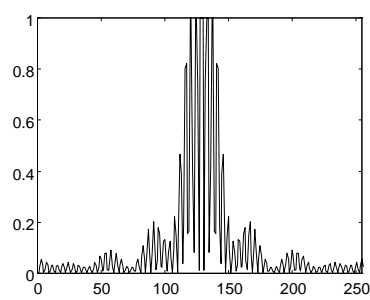


(c)

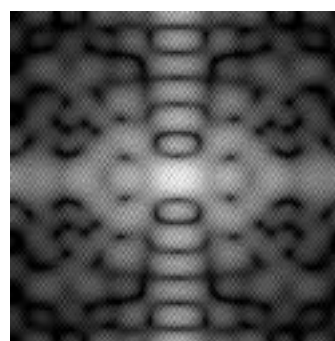
**Figura 5.4.** Representació de la distribució d'intensitat conjunta  $I$  en el pla de Fourier quan la Figura 5.1a es reproduïx en el pla d'entrada del processador. Es suposa un enregistrament lineal del sistema de captació amb  $\alpha = 100\%$ . (a) Representació lineal 2D de la distribució, (b) secció en 1D de la mateixa funció incloent el seu valor màxim central  $i$ , (c) representació logarítmica 2D de la distribució.



(a)

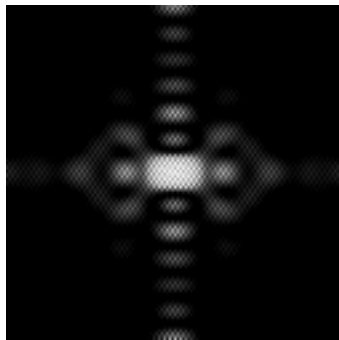


(b)

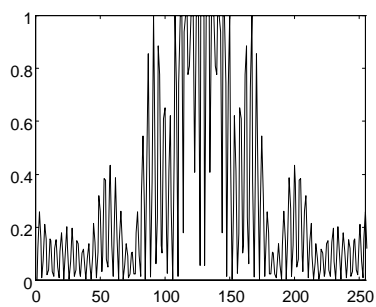


(c)

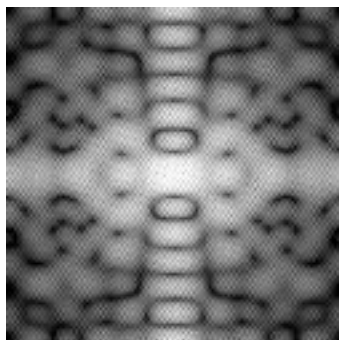
**Figura 5.5.** Ídem per  $\alpha = 50\%$ .



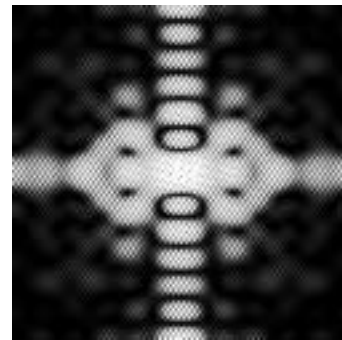
(a)



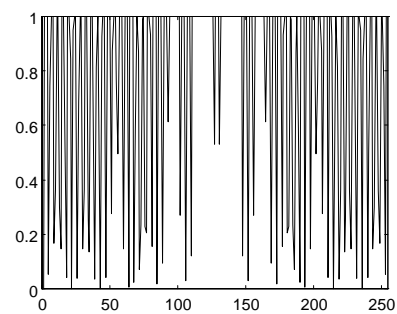
(b)



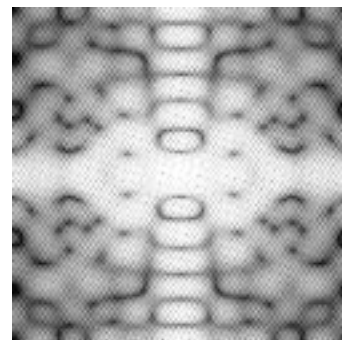
(c)



(a)



(b)



(c)

**Figura 5.6.** Representació de la distribució d'intensitat conjunta  $I$  en el pla de Fourier quan la Figura 5.1a es reproduïx en el pla d'entrada del processador. Es suposa un enregistrament lineal del sistema de captació amb  $\alpha = 10\%$ . (a) Representació lineal 2D de la distribució, (b) secció en 1D de la mateixa funció incloent el seu valor màxim central  $i$ , (c) representació logarítmica 2D de la distribució.

**Figura 5.7.** Ídem per  $\alpha = 1\%$ .

La deducció que es pot fer arrel d'aquest estudi és que ambdós elements, saturació de la càmera CCD per un costat, i la mateixa no-linealitat aplicada per l'altre, comporten un efecte paral·lel en el processament de la informació: modificar el pes de les altes freqüències en el contingut espectral general. El nostre objectiu d'obtenir un processador sensible a les no-linealitats, ens obliga a prendre un compromís. Cal introduir un cert nivell de saturació per millorar els resultats de correlació i concretament aconseguir un augment en el paràmetre PCE, però és necessari escollir-lo de manera acurada, ja que un nivell excessiu de saturació pot provocar una pèrdua de sensibilitat del dispositiu optoelectrònic respecte la variació dels paràmetres  $L$  i  $M$  que descriuen la no-linealitat aplicada. Si el sistema no fos sensible a les variacions d' $L$  i  $M$  de la DNC, les possibilitats del nostre dispositiu en la seva aplicació en diferents tasques de reconeixement quedarien molt limitades.

Com a motiu principal, doncs, intentarem conservar la dependència del sistema amb les diferents no-linealitats, millorant al mateix temps el paràmetre PCE com indicatiu d'una millora de la qualitat del pic de correlació i, en conseqüència, dels resultats de reconeixement. Per això, hem escollit el llindar de saturació de valor d' $\alpha = 10\%$ . A partir dels resultats obtinguts en el nostre exemple (Figura 5.3), veiem que aquest llindar permet assolir conjuntament els objectius proposats i, per tant, s'ha fixat el llindar de saturació en el 10% per les posteriors aplicacions.

### 5.3 Quantització del sistema de captació

El sistema d'adquisició quantitza la informació enregistrada de les distribucions  $I$ ,  $I_S$  i  $I_R$  en els diferents nivells de gris discrets que li són accessibles. Els dispositius CCD comercials acostumen a tenir valors entre 256 nivells de quantització, que correspon a les càmeres més esteses sobretot pel seu baix cost, i 1024 nivells de gris diferents (càmeres d'alt rang dinàmic emprades, per exemple, en observacions astronòmiques). Nosaltres disposarem d'un sistema d'adquisició (càmera i targeta digitalitzadora) amb 256 nivells de quantització, que prèviament calibrarem per poder assegurar que el sistema treballa en un rang aproximadament lineal. Sovint la

resposta més lineal de la càmera no abraça el rang màxim de nivells de gris, sinó que la zona de treball lineal inclou un nombre inferior de nivells de quantització.

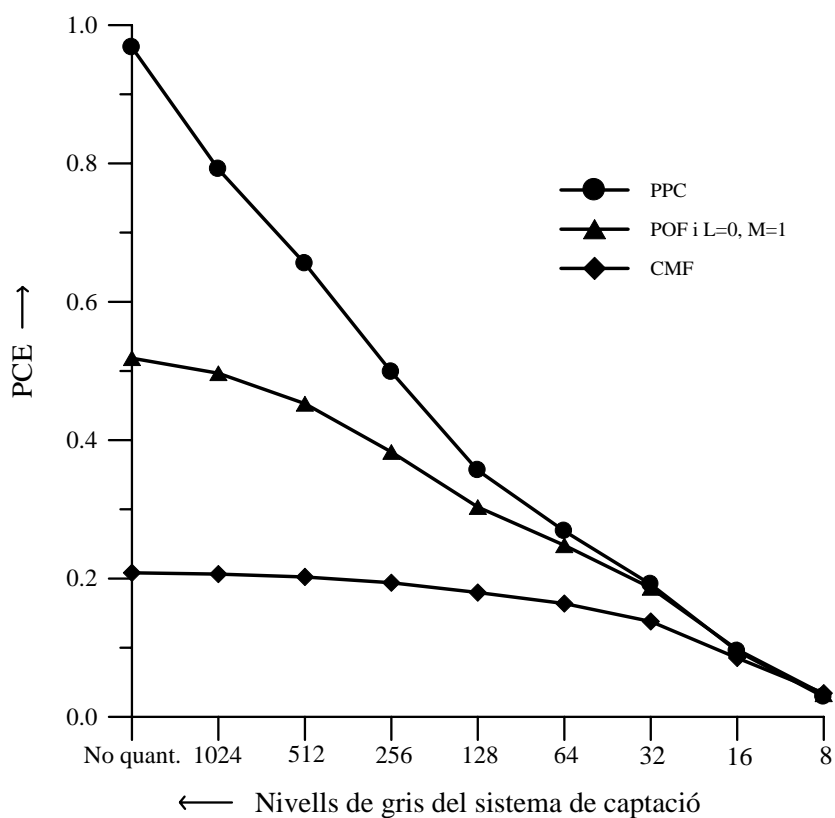
Novament, s'ha emprat la Figura 5.1a com a exemple representatiu del procés de reconeixement d'objectes binaris, i s'ha pres el paràmetre PCE com a referència per a l'anàlisi de la influència de la quantització de la càmera en els resultats de correlació. Hem mantingut fixat un nivell de saturació d' $\alpha = 10\%$  en l'enregistrament de les tres distribucions  $I$ ,  $I_S$  i  $I_R$ . S'ha partit d'un hipotètic i ideal enregistrament continu del senyal (sense quantització de la informació), i s'ha anat disminuint el nombre de nivells de gris fins a un nombre molt reduït, concretament 8.

La representació dels resultats obtinguts es mostra en la Figura 5.8, on novament s'ha graficat simultàniament els quatre tipus de processament diferents escollits (CMF, POF,  $L=0$  i  $M=1$  i PPC). Observem en aquesta representació, la forta dependència del PCE amb el nombre de nivells de quantització de la càmera. A mesura que decreixen els valors disponibles del sistema d'adquisició, la qualitat del pic de correlació empitjora ràpidament, tenint un efecte directe sobre els valors del PCE obtinguts. El comportament és similar en els quatre tipus de processament estudiats, observant un decreixement més brusc en el cas del PPC. Per sota de 16 nivells de gris s'obté gairebé el mateix valor del PCE per qualsevol paràmetre  $L$  i  $M$  escollit i, per tant, el sistema esdevé insensible al tipus de no-linealitat aplicada.

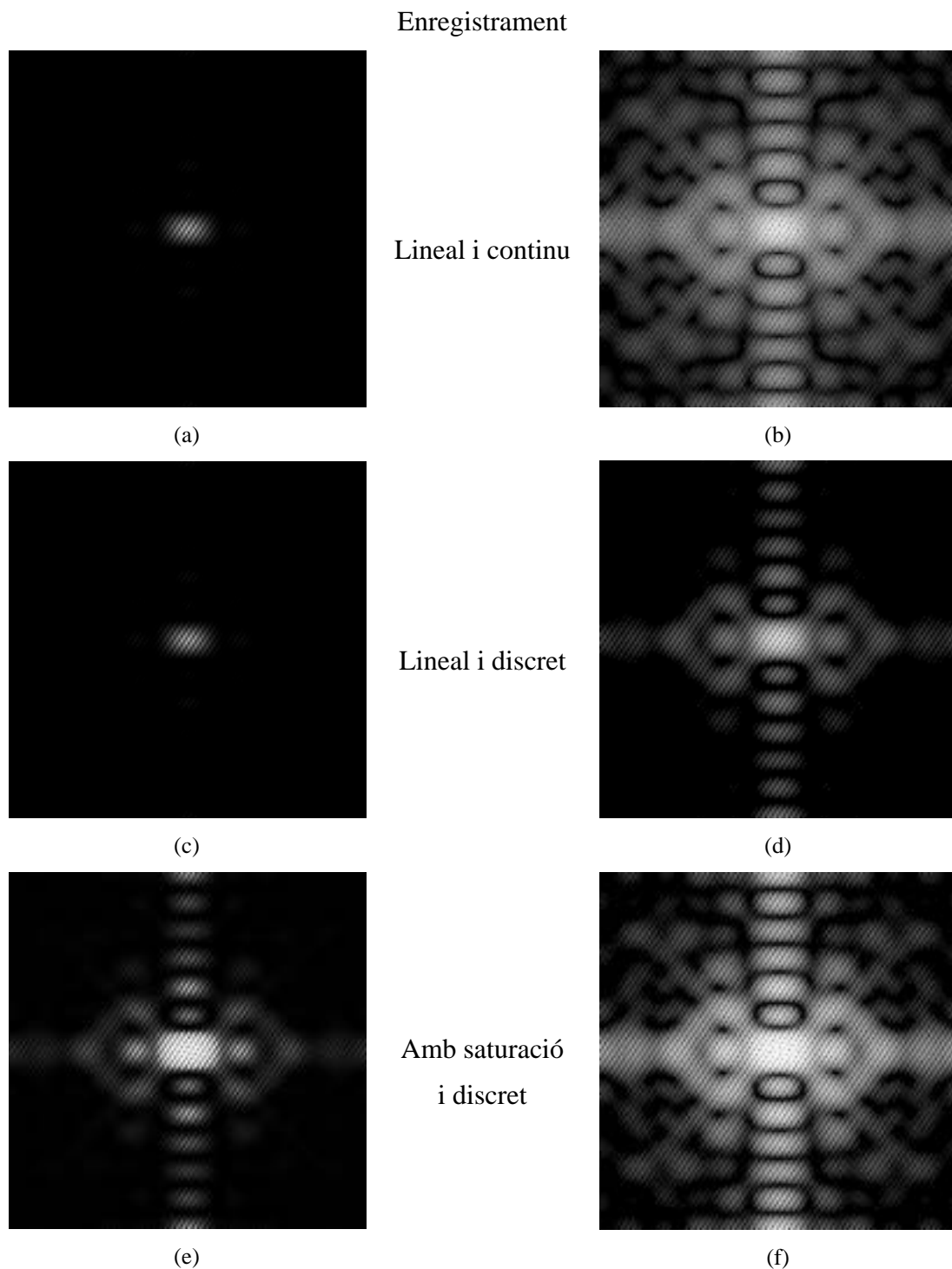
Podem explicar el comportament posat de manifest en la Figura 5.8, a partir de l'espectre de transformades conjuntes,  $I$ , corresponent al cas de l'autocorrelació de la lletra E (Figura 5.1a) quan l'observem en els tres casos següents: primer, enregistrament lineal (sense saturació) i continu (sense quantització); segon, lineal i discret amb 256 nivells de gris del sistema de captació, i tercer, saturat amb un llindar d' $\alpha = 10\%$  i al mateix temps discret amb la quantització de 256 nivells de gris. Aquests espectres es troben representats en la Figura 5.9, on també s'inclouen llurs representacions logarítmiques per facilitar-ne la visualització. Les Figures 5.9a i 5.9b són idèntiques a les anteriors Figures 5.4a i 5.4c, respectivament, però s'han mantingut en aquesta representació per facilitar la comparació visual amb els altres



casos. Si comparem la Figura 5.9d, on s'inclou la limitació experimental de la quantització en 256 nivells de gris del sistema de captació, amb la Figura 5.9b, que correspon a la informació de l'espectre conjunt suposant un enregistrament lineal i continu, s'observa una pèrdua important de la informació de les altes freqüències. En la Figura 5.9d, la regió de suport de l'espectre capturat es troba principalment centrat en les baixes freqüències. En canvi, al permetre saturació de l'espectre, en particular amb un llindar del 10% en la Figura 5.9f, es recupera en gran mesura la informació corresponent a les freqüències més altes.



**Figura 5.8.** Variació del PCE en funció de la quantització en nivells de gris introduïda pel sistema d'adquisició, quan el pla d'entrada és el de la Figura 5.1a. El llindar de saturació és  $\alpha = 10\%$ .



**Figura 5.9.** Distribució d'intensitat de l'espectre conjunt  $I$ , per a l'escena de la Figura 5.1a, en representació lineal en les figures de l'esquerra i logarítmica a la dreta. En (a) i (b) es suposa enregistrament lineal i continu de la càmera CCD, en (c) i (d) l'enregistrament és lineal però existeix quantització de la càmera CCD en 256 nivells de gris, i en (e) i (f) s'afegeix a la quantització en 256 nivells de gris, un efecte de saturació amb un llindar d' $\alpha = 10\%$ .

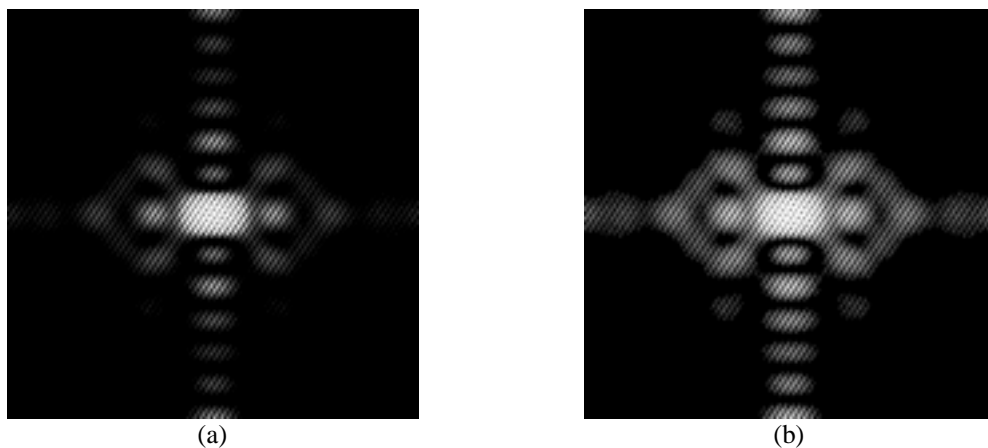
El comportament que acabem de descriure per la funció  $I$  (espectre de transformades conjuntes), és comú pels dos altres espectres que s'han d'enregistrar,  $I_S$  i  $I_R$ . D'aquest estudi podem extreure'n doncs, que tot i la inevitable pèrdua d'informació en les funcions deguda a la quantització del sistema d'adquisició, no esdevé inviable el mètode de la correlació no-lineal dual. Es pot contrarestar l'efecte negatiu de la quantització amb un procés de saturació de les baixes freqüències de l'espectre, de manera que sigui possible mantenir la informació més rellevant en les distribucions.

Dels resultats aportats en la secció 5.2, vam concloure que un llindar de saturació escollit de manera apropiada permet ressaltar un ampli rang de freqüències altes durant l'enregistrament de les distribucions  $I$ ,  $I_S$  i  $I_R$ . Concretament, vam escollir un llindar de saturació d' $\alpha = 10\%$ . Aquest fet es tradueix en una ampliació de la regió de suport de l'espectre conjunt modificat i, com a conseqüència, augmenta la qualitat del pic de correlació juntament amb l'eficiència lluminosa del dispositiu.

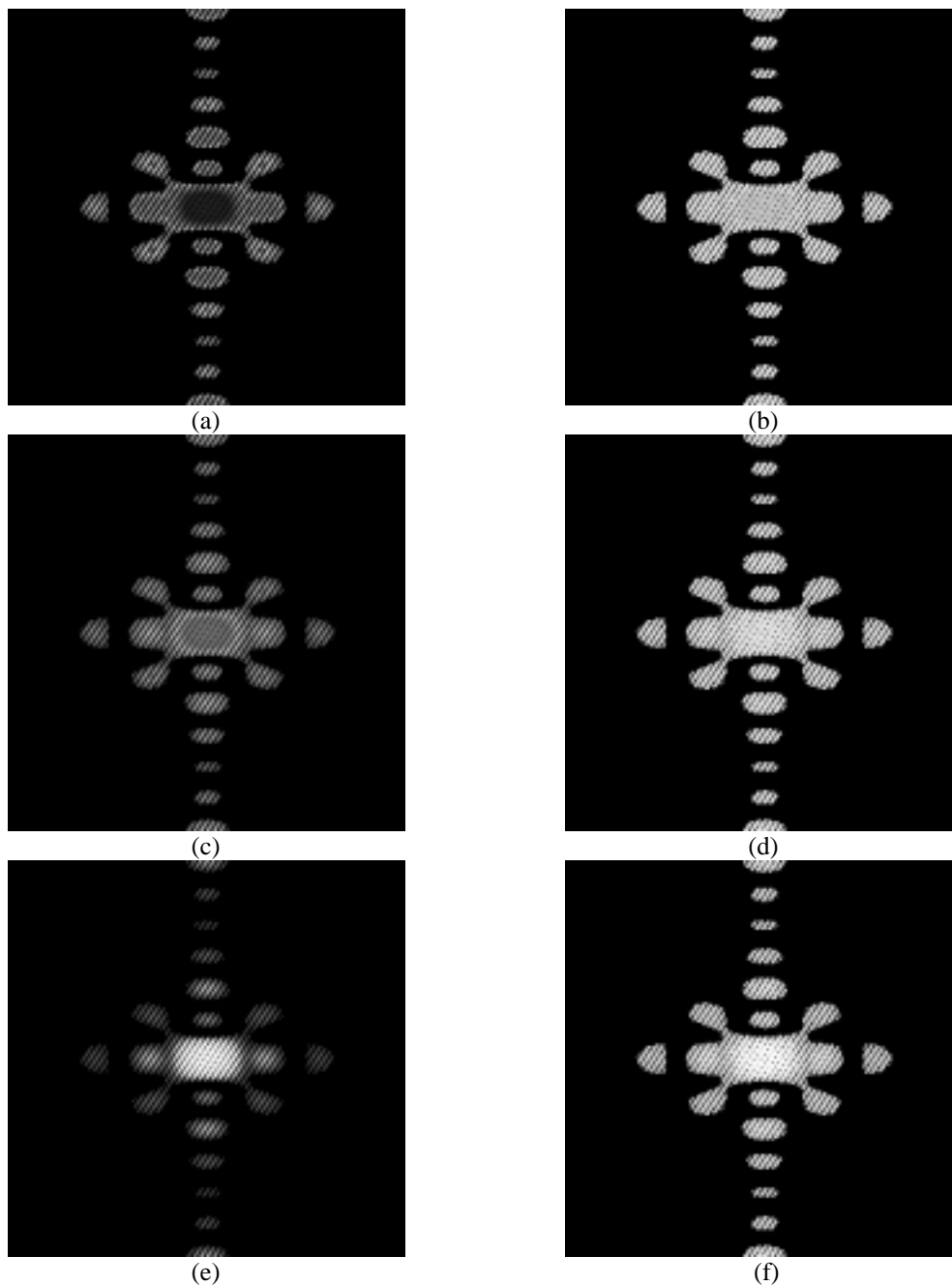
Tenint en compte el dispositiu amb el qual treballarem en la realització experimental de la DNC, que és un sistema de captació sobre el qual s'haurà aplicat una saturació de la càmera d' $\alpha = 10\%$  i que a més discretitza la imatge amb 256 nivells de quantització, és important fixar-nos que els resultats obtinguts i representats en la Figura 5.8 d'aquesta secció 5.3, són prou satisfactoris en els quatre parells  $(L, M)$  de la DNC estudiats. Tot i que la quantització produeix una disminució important del paràmetre PCE, que ens indica una disminució de la qualitat del pic de correlació respecte del cas ideal continu (no quantització), podem observar que les diferències entre els quatre tipus de processament aplicats són encara prou importants com per considerar que el sistema de la DNC és sensible a les variacions en els paràmetres  $L$  i  $M$ . En gran mesura, la conservació d'aquesta variació amb els valors d' $L$  i  $M$  és deguda a la saturació permesa en la zona central de l'espectre.

Aquesta conclusió és certa sempre i quan l'enregistrament de les funcions esmentades no es faci amb un nombre molt petit de nivells de gris (per exemple, 16 o 8). En aquest cas, la pèrdua d'informació és tan elevada que és responsable de la insensibilitat del sistema a les no-linealitats aplicades. Per exemple, si considerem el

cas de tenir només 16 nivells de gris en el sistema d'adquisició, l'espectre de transformades conjuntes  $I$  que es digitalitzaria correspondria al de la Figura 5.10a. Per facilitar-ne la completa visualització de tota la informació inclosa en aquest espectre, s'adjunta també una representació logarítmica de mateixa funció  $I$  (Figura 5.10b). Un cop enregistrades les tres distribucions d'intensitat  $I$ ,  $I_S$  i  $I_R$ , amb un nivell de saturació del 10% i 16 nivells de gris, aquestes s'utilitzarien per a calcular l'espectre de transformades conjuntes modificat no-linealment,  $P'(\vec{v})$  (Eq. 4.27). Comparant l'espectre conjunt modificat MJPS en els casos del PPC ( $L=0$  i  $M=0$ ), POF ( $L=1$  i  $M=0$ ), i CMF ( $L=1$  i  $M=1$ ) (Figura 5.11), observem que la regió de suport és molt reduïda pel fet de no disposar d'un nombre més elevat de nivells de quantització en la digitalització. La zona central saturada és equivalent pels tres casos estudiats, i ocupa gran part de la regió on la distribució és no-nul·la. Per tant, en aquests espectres modificats no-linealment, la informació continguda és molt minsa i és pràcticament equivalent per qualsevol no-linealitat que es vulgui aplicar. En aquestes figures és pot veure l'efecte que es denomina en anglès *hard clipping*, on l'efecte conjunt de la saturació i de la no-linealitat aplicada, aconsegueixen pràcticament binaritzar la informació continguda en l'espectre conjunt modificat no-linealment, MJPS (veure representacions logarítmiques de la Figura 5.11).



**Figura 5.10.** Distribució d'intensitat de l'espectre conjunt  $I$  corresponent al cas on l'escena a analitzar i la referència són la mateixa lletra E (Figura 5.1a). (a) Representació lineal, i (b) logarítmica. Durant la captació es suposa un nivell de saturació d' $\alpha = 10\%$ , i 16 nivells de gris del sistema d'adquisició.



**Figura 5.11.** Funcions MJPS en el cas on l'escena a analitzar i la referència són la mateixa lletra E (Figura 5.1a), i s'enregistren les distribucions espectral  $I$ ,  $I_S$  i  $I_R$  amb un nivell de saturació  $\alpha = 10\%$  i amb una quantització de 16 nivells de gris. Les tres figures de l'esquerra són representacions lineals i les de la dreta logarítmiques. (a) i (b) corresponen al cas del PPC, (c) i (d) al POF, i (e) i (f) al CMF.

Tant pel que fa al nivell de saturació com als nivells de quantització de les distribucions en el procés d'adquisició mitjançant la càmera CCD, podem comparar els valors que hem escollit per les posteriors aplicacions de la DNC, amb els valors

proposats en el cas del JTC binari (BJTC) [204]. Si considerem 256 nivells de gris disponibles en el nostre sistema d'adquisició, el BJTC requereix un nivell màxim de saturació del 27% [204]. Només si el coeficient de saturació es manté per sota d'aquest valor, el BJTC proporcionarà els pics de correlació. Tenint en compte que el BJTC és un cas particular del JTC no-lineal, i que aquest es pot incloure dins de la DNC (veure Taula 4.1 del capítol 4), podem pensar que el nivell de saturació del 10% que hem escollit pot ser un nivell apropiat per dur a terme la DNC.

Si ens fixem en el cas del PPC (Figura 5.11a), observem una zona central enfosquida, coincident aproximadament amb la regió d' $I$  on hi ha saturació durant el procés de captació (veure Figura 5.10a). En el càlcul del MJPS, en el producte  $I \cdot I_S^{(L-1)/2} \cdot I_M^{(M-1)/2}$  (Eq. 4.27) amb  $L = M = 0$ , que correspon al cas del PPC, s'inverteix el mòdul de  $\sqrt{I_S}$  i  $\sqrt{I_R}$ , fent que els valors majors d' $I_S$  i  $I_R$  (és a dir, les zones saturades en la regió central de l'espectre) generin els valors més petits del MJPS després de la no-linealitat.

#### 5.4 Nivell de zero en el càlcul de la no-linealitat (paràmetre $\varepsilon$ )

Un cop enregistrades les distribucions  $I$ ,  $I_S$  i  $I_R$  mitjançant el sistema d'adquisició, el següent pas de la DNC correspon al càlcul de l'espectre conjunt modificat,  $P'(\vec{v})$ , segons l'expressió (4.27) del capítol anterior. Recordem que aquesta expressió s'obté de la definició de la DNC (Eq. 4.2) on s'aplica a les funcions  $I_S$  i  $I_R$  un operador no-lineal definit per l'Eq. (4.1). Per evitar punts de singularitat en el càlcul no-lineal, cal establir la condició que el resultat de l'operador serà nul quan el mòdul del seu argument sigui zero. Aquesta condició implica que la funció  $P'(\vec{v})$  també s'anul·li en el cas en que els mòduls d' $I_S$  i/o  $I_R$  siguin zero. Per tant, podem dir que el càlcul involucrat en el procés no-lineal sobre les funcions  $I_S$  i  $I_R$ , per qualsevol valor dels paràmetres  $L$  i  $M$ , es realitza només en la regió dels espectres on s'han enregistrat valors no nuls de la intensitat que per definició coincideix amb les regions de suport de les distribucions en l'espai de Fourier. D'aquesta manera, es mantenen nuls els

píxels amb nivell de gris zero on inicialment no hi ha hagut resposta del sistema de captació.

El problema de les singularitats que hem esmentat, és comú en tots aquells mètodes de filtratge tipus invers [205]. Basant-nos en la proposta del pròpiament anomenat filtre invers (IF), inicialment conegut com a filtre d'amplitud compensada [5], podem pensar en una nova expressió de l'operador no-lineal tal com

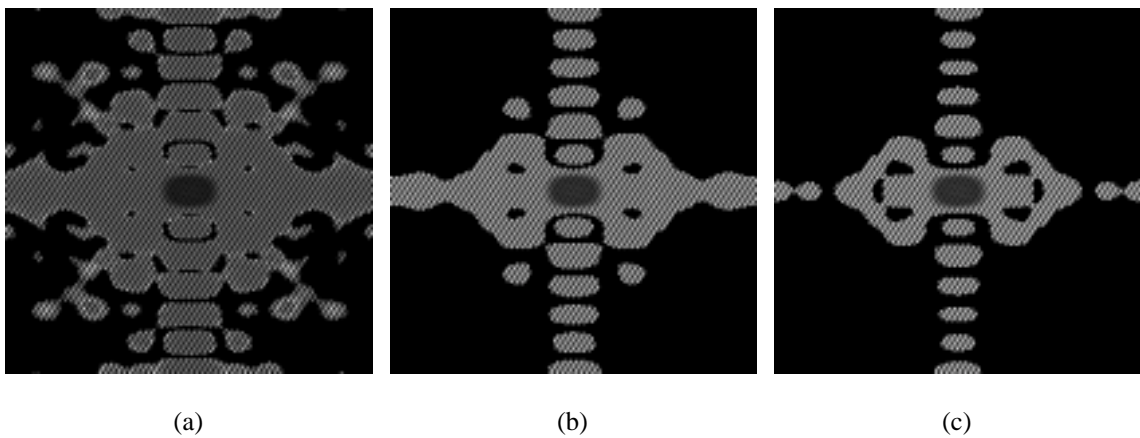
$$N_P[z] = \begin{cases} |z|^{P-1} \cdot z & \text{si } |z| > \varepsilon \\ 0 & \text{si } |z| \leq \varepsilon, \end{cases} \quad (5.4)$$

on  $\varepsilon$  és un paràmetre que representa el nivell de zero de la intensitat enregistrada per la càmera CCD en el pla de Fourier. Aquest paràmetre està relacionat amb el valor d'*offset* fixat pel sistema de captació d'imatges.

El paràmetre  $\varepsilon$  ens permet, a més, eliminar el soroll present en les distribucions enregistrades i també limitar llurs regions de suport [5, 152].

Per tal d'escollir un valor apropiat d' $\varepsilon$ , ens és indispensable conèixer la seva influència en el càlcul de la DNC. Per exemple, en la Figura 5.12 mostrem les funcions MJPS calculades a partir de la lletra E considerada com escena i com referència (pla d'entrada de la Figura 5.1a), quan  $I$ ,  $I_S$  i  $I_R$  es saturen amb un llinard d' $\alpha = 10\%$  i s'enregistren amb 256 nivells de gris del sistema de captació. La no-linealitat aplicada correspon als valors  $L = M = 0$  representatius del PPC. La diferència entre les tres funcions representades es troba en el valor del paràmetre  $\varepsilon$  escollit. En el primer cas (Figura 5.12a) es pren  $\varepsilon = 0$ , en (b)  $\varepsilon = 5$  i, finalment en (c)  $\varepsilon = 10$ . Cal tenir en compte que les funcions sobre les quals s'aplica la no-linealitat estan quantitzades en 256 nivells de gris, de manera que l' $\varepsilon$  escollit, en valor relatiu representa en (a) el 0%, en (b) el 2% i en (c) el 4% del valor màxim de les distribucions.

La variació de la distribució MJPS en funció del paràmetre  $\varepsilon$  es troba bàsicament en dos fets: per una banda, la regió de suport del MJPS es redueix lògicament a mesura que augmentem el valor d' $\varepsilon$  i, en segon lloc, el contrast del patró de franges es pot veure alterat amb  $\varepsilon$  (per exemple, en el cas del PPC el contrast de les franges augmenta amb  $\varepsilon$ ).



**Figura 5.12.** Representació logarítmica de les distribucions MJPS amb  $L = M = 0$  calculades a partir d' $I$ ,  $I_S$  i  $I_R$  enregistrades amb un llindar de saturació del 10%, 256 nivells de gris, i un valor del paràmetre  $\varepsilon$  del sistema d'adquisició de (a) 0, (b) 5 i (c) 10.

El proper pas d'aquesta anàlisi, serà el de determinar el rang adequat per l' $\varepsilon$  del processament no-lineal dual que ens permeti uns bons resultats de reconeixement. És a dir, estudiarem què té un pes més gran en el procés de correlació no-lineal dual: obtenir una àmplia regió de suport per disposar de més quantitat d'informació, o bé, obtenir unes bones franges d'interferència ben contrastades per facilitar la seva reproducció en el modulador durant el segon pas de la DNC (La influència de les característiques del modulador en la reproducció d'aquesta distribució s'estudiaran en el proper apartat 5.5).



Realitzem un estudi numèric senzill per poder determinar quin factor té més importància. Si considerem l'escena de la Figura 5.1a, en la qual escena i referència corresponen a la mateixa lletra E, podem determinar l'expressió de la funció  $P'(\vec{\nu})$  a partir de l'Eq. (4.27) del capítol anterior,

$$P'(\vec{\nu}) = 2|R(\vec{\nu})|^{L+M} [1 + \cos(4\pi\vec{\nu} \cdot \vec{b})], \quad |R(\vec{\nu})| > \varepsilon, \quad (5.5)$$

on el cosinus indica el patró de franges que ve determinat per la posició dels objectes,  $|R(\vec{\nu})|$  és el mòdul de la transformada de Fourier de l'objecte (lletra E) quan està centrat a l'origen de coordenades, i  $\vec{b}$  és la distància de l'escena i de la referència a l'eix òptic, en direccions oposades per cada lletra. L'*offset* representat pel paràmetre  $\varepsilon$  s'aplica a  $|R(\vec{\nu})|$  prèviament al càlcul de  $P'(\vec{\nu})$ . Per estudiar separatament com afecta la regió de suport i el contrast del patró de franges sobre el resultat final de la DNC introduïm un paràmetre  $K$  que moduli l'amplitud de les franges d'interferència de l'espectre conjunt, de manera que per un  $\varepsilon$  fixat i variant  $K$  entre  $[0,1]$ , variem el contrast de les franges sense afectar l'energia total en el pla de freqüències. Amb la introducció del paràmetre  $K$ , l'expressió de la funció  $P'(\vec{\nu})$  és la següent:

$$P'(\vec{\nu}) = 2|R(\vec{\nu})|^{L+M} [1 + K \cos(4\pi\vec{\nu} \cdot \vec{b})], \quad |R(\vec{\nu})| > \varepsilon. \quad (5.6)$$

Considerem la no-linealitat corresponent al PPC ( $L = M = 0$ ), i fem variar el paràmetre  $\varepsilon = \{0, 5, 10\}$  i el paràmetre  $K$  dins l'interval  $[0,1]$ . Quan  $K = 0$  s'obté una distribució uniforme sense franges, i amb  $K = 1$  s'obté el màxim contrast de les franges, de manera que per un valor d' $\varepsilon$  fixat i variant  $K$  es manté constant l'energia de la distribució. D'aquesta manera, es poden comparar els resultats obtinguts directament en el pla de sortida del correlador, analitzant el màxim d'autocorrelació (*ac*).

En la Taula 5.1 es mostren els resultats del màxim d'autocorrelació ( $ac$ ), i del paràmetre PCE per diferents valors de  $K$ , calculant la no-linealitat del PPC sobre les funcions captades amb un  $\varepsilon = 0$ . En la Taula 5.2 apareixen els mateixos resultats però havent considerat un  $\varepsilon = 10$  en el procés d'enregistrament. La comparació d'ambdues taules, posa clarament de manifest que els valors obtinguts pel PCE decauen ràpidament a l'augmentar l'*offset* del sistema de captació a través del paràmetre  $\varepsilon$ . També disminueix el valor màxim del pic d'autocorrelació, però no de manera tant acusada. Per tant, la regió de suport de la funció juga un paper molt important en el procés de reconeixement, possibilitant l'obtenció de bons resultats si la regió on s'inclou informació rellevant de la funció és el més gran possible.

Una segona observació important s'obté a l'analitzar els resultats obtinguts per a un mateix  $\varepsilon$ , quan es modifica el paràmetre  $K$  de control del contrast de les franges d'interferència. Tant en la Taula 5.1, com en la Taula 5.2, veiem que és necessari considerar valors molt baixos de  $K$ , per obtenir una degradació molt important dels resultats. Altrament dit, la informació continguda en les franges d'interferència es manté, tot i tenir un contrast molt baix. Això permet obtenir un bon pic d'autocorrelació en la posició esperada, excepte en el cas extrem on  $K = 0$  pel qual la funció  $P'(\bar{v})$  passaria a ser completament uniforme i el pic d'autocorrelació quedaria inclòs en l'ordre zero de difracció.

Concloem doncs, que per als posteriors resultats d'aquest treball, és millor utilitzar el paràmetre  $\varepsilon$  inclòs en la no-linealitat aplicada en la DNC amb el valor mínim permès, de manera que s'enregistren les distribucions en una regió de suport el màxim d'extensa, encara que es pugui perdre contrast en les franges d'interferència. La pèrdua de contrast d'aquestes franges, excepte en el cas d'ésser quasi total, es pot tolerar sense que afecti en excés al resultat de correlació.

Fixant el paràmetre  $\varepsilon$  en un valor baix, concretament prenem  $\varepsilon = 0$ , també aconseguim que les modificacions experimentals de la regió de suport de les funcions  $I$ ,  $I_S$  i  $I_R$  siguin degudes, sense retalls afegits, a l'efecte conjunt de la saturació i de la quantització del sistema de captació.

**Taula 5.1.** Resultats de correlació en quant al paràmetre PCE i al màxim d'autocorrelació ( $ac$ ) per a  $\varepsilon = 0$  i diferents valors de  $K$ .

$\varepsilon = 0$		
$K$	PCE	$ac$
1	0.50	$1.38 \cdot 10^{10}$
0.75	0.52	$8.48 \cdot 10^9$
0.5	0.53	$3.72 \cdot 10^9$
0.25	0.51	$1.02 \cdot 10^9$
0	0.01	- (no hi ha visible un màxim de correlació)

**Taula 5.2.** Resultats de correlació en quant al paràmetre PCE i al màxim d'autocorrelació ( $ac$ ) per a  $\varepsilon = 10$  i diferents valors de  $K$ .

$\varepsilon = 10$		
$K$	PCE	$ac$
1	0.17	$6.44 \cdot 10^9$
0.75	0.19	$4.94 \cdot 10^9$
0.5	0.20	$3.03 \cdot 10^9$
0.25	0.19	$6.96 \cdot 10^8$
0	0.01	- (no hi ha visible un màxim de correlació)

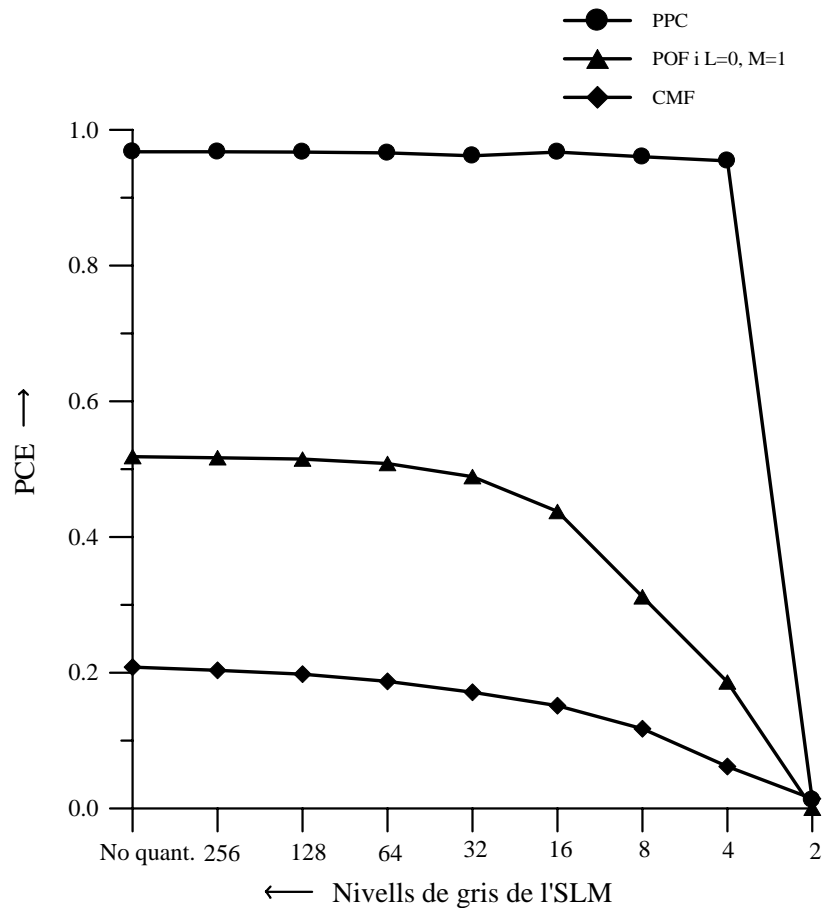
## 5.5 Quantització del modulador espacial de llum (SLM)

En aquesta secció s'estudia la influència de la quantització introduïda pel modulador. L'SLM intervé en el primer pas de la DNC al reproduir els plans d'entrada del correlador (l'escena a analitzar, l'objecte de referència i ambdós conjuntament). En un primer cas més senzill, les imatges que reproduïm en el modulador, consisteixen en diferents caràcters binaris i, per tant, no es veuran influenciats pel nombre de nivells de gris disponibles en el dispositiu optoelectrònic. La quantització introduïda pel modulador espacial de llum té un paper important en el cas de treballar amb escenes en nivells de gris si volem diferenciar els objectes segons el seu nivell energètic, i també en el cas de reproduir l'espectre conjunt modificat no-linealment (MJPS) amb un rang d'intensitats molt ampli. Com la reproducció del MJPS en nivells de gris és un requisit comú a totes les experiències, tant si el pla d'entrada conté objectes binaris com si no, és important assegurar un cert nombre de nivells de quantització reproduïbles amb fiabilitat en el modulador. També serà important la linealitat de la reproducció, ja que aquest condicionament experimental participarà igualment en el resultat final de la correlació.

En la Figura 5.13 es mostra la influència del nombre de nivells de gris, hipotèticament disponibles en un modulador, sobre el paràmetre PCE en el cas de prendre l'autocorrelació de la lletra E (Figura 5.1a). Tenint en compte les conclusions dels apartats anteriors, hem fixat un nivell de saturació d' $\alpha = 10\%$  i una quantització de 256 nivells de gris en el sistema d'adquisició. Es considera una variació decreixent en el nombre de nivells de quantització de l'SLM des d'un cas ideal de reproducció contínua del senyal MJPS, és a dir, sense quantització (equivalent a un nombre infinit de nivells de gris), passant per valors més habituals en les experiències reals de 64/32 nivells de gris, i es continua disminuint fins a una reproducció binària amb tant sols 2 nivells de quantització. Les variacions del PCE són molt suaus per a tots els tipus de processament representats (PPC, POF,  $L=0$  i  $M=1$  i CMF), de manera que es pot afirmar que el PCE és pràcticament constant excepte en el cas on l'SLM treballa amb molts pocs nivells de gris (per sota de 16). El processament menys afectat per aquest condicionament experimental, és el PPC.

Els altres tres presenten una variació més significativa del PCE, que decreix suaument a mesura que el nombre de nivells de gris és més reduït.

Per explicar aquest comportament, cal estudiar més detalladament com és la distribució del MJPS, calculada mitjançant l'expressió de  $P'(\vec{v})$  (Eq. 4.27), que es reproduïx en l'SLM durant el segon pas de la DNC. Aquesta distribució  $P'(\vec{v})$  consisteix bàsicament en un patró de franges d'interferències modulad per una funció envolupant. Comentarem tres aspectes: un és el patró de franges, un altre són els contorns que limiten l'envolupant i, el tercer, els pendents més o menys suaus amb els quals l'envolupant modula el màxim central i lòbuls laterals respecte del fons nul.



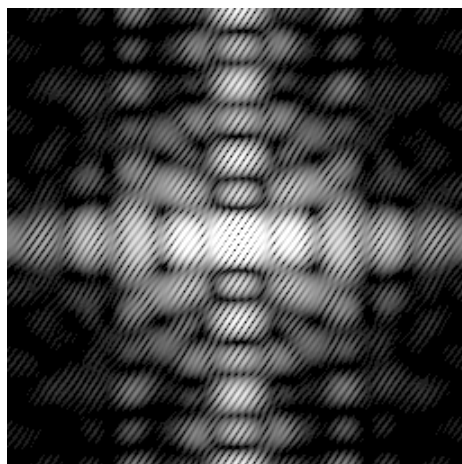
**Figura 5.13.** Variació del PCE en funció del nombre de nivells de gris de l'SLM, quan es considera el cas de l'autocorrelació de la lletra E (Figura 5.1a). El llindar de saturació és  $\alpha = 10\%$  i el sistema d'adquisició introdueix una quantització en 256 nivells de gris.

El patró de franges prové de la funció  $I$ . El nombre d'elements i llur posició relativa en el pla d'entrada determina principalment la orientació, la freqüència i la distribució de les franges d'interferència observades en la distribució  $I$  (Figura 5.4c). La Figura 5.4c, en representació logarítmica, correspon al cas d'introduir dues lletres E idèntiques (l'escena i la referència) en el pla d'entrada del processador (Figura 5.1a). Si s'enregistra la distribució d'intensitat  $I$  amb un llindar de saturació del 10% i amb 256 nivells de gris del sistema de captació, s'obté la Figura 5.9f. L'orientació i freqüència de les franges que apareixen al calcular el MJPS coincideixen amb les de  $I$ , independentment dels paràmetres  $L$  i  $M$  emprats (veure Figura 5.12a pel cas del PPC).

Els contorns de l'envolupant del màxim central i els dels lòbuls laterals que s'observen en la distribució  $I$  i, en conseqüència també en la funció MJPS, depenen de la forma dels elements continguts en l'escena a analitzar i en l'objecte de referència. Per exemple, podríem comparar la Figura 5.9f corresponent a la distribució  $I$  en representació logarítmica pel cas de tenir dues lletres E simultàniament en el pla d'entrada del correlador, amb la Figura 5.14 que s'obté amb les mateixes condicions experimentals (llindar de saturació d' $\alpha = 10\%$  i quantització amb 256 nivells de gris) quan situem en el pla d'entrada un caràcter H al costat de la referència E (Figura 5.1b). Les diferències en els lòbuls laterals també es poden posar de manifest si comparem els espectres modificats no-linealment (MJPS) pel cas del PPC ( $L = M = 0$ ) (veure Figures 5.12a i 5.15).

Per últim, el tercer aspecte a remarcar, pròpiament del MJPS, està relacionat amb els pendents de l'envolupant que modula els diferents lòbuls i les franges d'interferència. En la nostra anàlisi, aquest és l'aspecte més interessant, perquè les no-linealitats aplicades canvien el pendent de l'envolupant dins la regió de suport sense modificar de manera essencial la forma dels lòbuls. Per exemple, el PPC possibilita que la informació de les franges d'interferència sigui pràcticament uniforme en tota la regió de suport de l'espectre, de manera que l'envolupant de les franges és quasi binària (Figura 5.12a). Si considerem una altra no-linealitat, com ara la corresponent al POF, l'envolupant de modulació és més suau que en el cas anterior

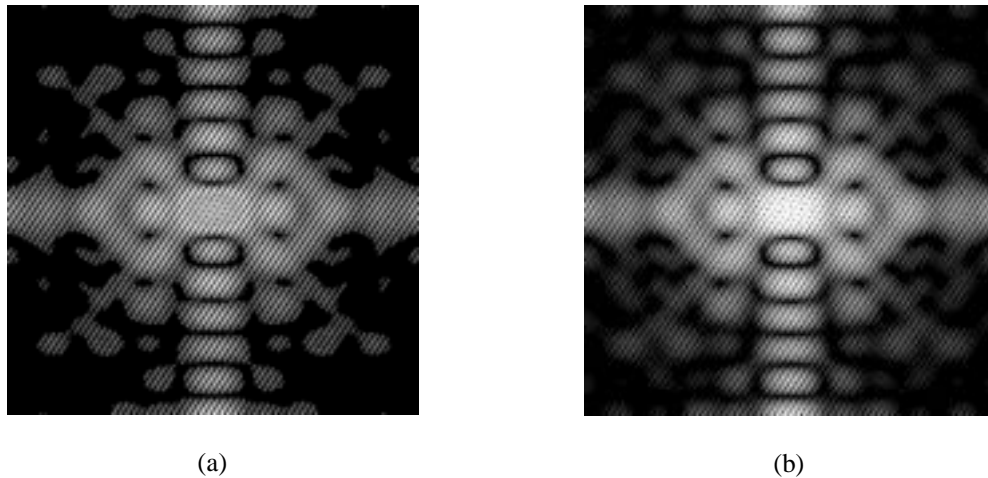
(Figura 5.16a), i encara és molt més suau si la no-linealitat aplicada és la del CMF (Figura 5.16b).



**Figura 5.14.** Representació logarítmica de la distribució  $I$  per a l'escena de la Figura 5.1b quan s'enregistra amb un nivell de saturació d' $\alpha = 10\%$  i 256 nivells de quantització.



**Figura 5.15.** Representació logarítmica de la distribució de l'espectre conjunt modificat no-linealment (MJPS) obtinguda pel cas del PPC ( $L = M = 0$ ), quan es reproduueix la Figura 5.1b en el pla d'entrada del correlador. Es considera un llindar de saturació del 10% i una quantització del sistema de captació de 256 nivells de gris.



**Figura 5.16.** Representació logarítmica de les distribucions de l'espectre conjunt modificat no-linealment (MJPS) en els casos (a)  $L=1$  i  $M=0$  (POF), i (b)  $L=1$  i  $M=1$  (CMF), quan es reproduïx la Figura 5.1a en el pla d'entrada del correlador. Es considera un llindar de saturació del 10%, i una quantització del sistema de captació de 256 nivells de gris.

Donat que les variacions de l'envolupant (lòbuls centrals i laterals) en tota la regió de suport de les distribucions MJPS són més suaus pel POF i pel CMF que pel PPC, fa falta un major nombre de nivells de quantització de l'SLM per a reproduir el POF i el CMF de manera fiable. En canvi, el PPC, on la distribució de franges està modulada per una envolupant quasi binària, permet una bona reproducció encara que es disposi d'un nombre petit de nivells de gris. Aquest fet explica que en la Figura 5.13 sigui el PPC la no-linealitat en la qual el paràmetre PCE es veu menys afectat en funció del nombre de nivells de gris reproduïbles en el modulador.

Aquests comentaris també ens porten a concloure que un sistema que sigui sensible a les no-linealitats introduïdes (a través de diferents valors d' $L$  i  $M$ ) requereix poder reproduir l'envolupant de les franges de la funció  $P'(\vec{v})$  amb suficient fiabilitat. Això requereix un nombre suficient de nivells de gris accessibles per l'SLM (per sobre de 16). Altrament, el sistema esdevé menys sensible i, per tant, menys versàtil. Prova d'això és que les tres corbes de la Figura 5.13 s'acosten per sota de 32 nivells de gris reproduïbles en l'SLM i s'ajunten per SLM binaris.



Per a les posteriors experiències d'aplicació de la DNC en el reconeixement per correlació òptica, hem escollit 32 nivells de gris per a la reproducció de la informació en el modulador, considerant aquest valor com a factible en la realització pràctica a partir del dispositiu que s'emprarà.

## **Capítol 6**

# **Sistema de reconeixement amb capacitat de discriminació variable**

### **6.1 Introducció**

En nombroses ocasions, els mètodes de filtratge convencionals doten als sistemes de reconeixement d'una capacitat discriminativa inadequada, bé sigui per ser massa baixa o bé per ser excessiva. En la seva evolució històrica, la tendència ha estat d'elevat la capacitat discriminativa dels sistemes de reconeixement, la qual cosa els permetia ser cada vegada més sensibles a petites diferències entre els objectes i aconseguir una discriminació més efectiva. Aquesta propietat, no obstant, també portava associats alguns efectes no volguts. Precisament per l'augment de sensibilitat, la resistència al soroll i a les petites distorsions disminuïa, de manera que el sistema de reconeixement apreciava diferències allà on, en realitat, eren els efectes deguts al soroll i altres distorsions.

Des d'un altre punt de vista, donat que la relació de similitud és arbitrària, pot interessar reconèixer com a iguals dos objectes que difereixin en detalls que desestimem. Un sistema de reconeixement amb capacitat de discriminació variable és capaç d'adaptar-se a un problema plantejat emprant el nivell de discriminació que requereixi. Això es podria aconseguir canviant el tipus de filtre en el correlador, però tot i amb això, només es podria escollir entre uns pocs nivells de discriminació amb els quals el sistema operés. La introducció de la tolerància també es podria fer variant el valor del llindar que s'aplica en el pla de correlació per decidir si l'objecte buscat és a l'escena i discriminar-lo d'altres objectes. Una mesura tan immediata com aquesta, però, generaria un augment de falses alarmes, a més de ser indiscriminada, no permetent l'establiment d'una tolerància selectiva a determinats aspectes dels objectes en anàlisi (contorns, intensitat, color, textura, soroll, etc.).

En aquest capítol mostrarem que un sistema que apliqui la DNC, segons el model definit en el capítol 4, pot assolir una capacitat de discriminació variable, que es controla a través dels valors (reals) assignats a dos paràmetres ( $L$  i  $M$  en la Eq. 4.27). Amb això es pretén adaptar el sistema de reconeixement a les exigències discriminatives del problema plantejat, podent variar d'un cas a un altre sense major dificultat. Aquesta nova propietat permetrà augmentar clarament les prestacions i la versatilitat d'un sistema de reconeixement.

Al llarg dels capítols que segueixen considerarem tres tipus de variacions en els objectes en front dels quals provarem la capacitat discriminativa del sistema de reconeixement. Aquestes variacions seran: variacions de forma en la silueta o contorn dels objectes, variacions d'intensitat (o nivell de gris) i variacions de color. En aquest capítol, ens centrarem en les dues primeres. L'anàlisi de la capacitat discriminativa davant variacions d'intensitat en l'objecte buscat té interès en sí mateixa i, a més, amplia el camp d'aplicació al reconeixement d'objectes en color. Això s'explica, com es veurà en el capítol 8, perquè el reconeixement d'objectes en color es sol realitzar mitjançant una correlació multicanal, on la informació del color es processa en canals separats (per exemple, els canals RGB proporcionats per una

càmera en color). En cada un dels canals, el problema es redueix a la discriminació de patrons en imatges acromàtiques amb diferents nivells de gris.

Per analitzar com varia la capacitat discriminativa del sistema en funció de les no-linealitats aplicades dins del model de la DNC, utilitzarem tests especialment elaborats pel tipus de diferències que presenten els objectes que els componen (diferències de forma, de nivell de gris i de color). El fet de tenir un únic sistema òptic i la possibilitat de modificar el tipus de no-linealitat aplicada simplement escollint adequadament els valors d' $L$  i  $M$ , ens portarà a buscar en quina regió dels valors d' $L$  i  $M$  és més convenient resoldre un problema concret de reconeixement per aconseguir una bona capacitat de discriminació, o bé, obtenir una certa tolerància a petits canvis.

Recordem que la capacitat de discriminació d'un sistema es defineix a partir de la condició per a què un caràcter sigui reconegut com l'objecte de referència durant el procés de reconeixement. Tal com hem vist en el capítol 2, aquesta condició es pot expressar com

$$\left| \frac{cc}{ac} - 1 \right| \leq u, \quad (6.1)$$

on  $cc$  indica el màxim de correlació creuada entre l'objecte a analitzar i l'objecte de referència, i  $ac$  indica l'autocorrelació de l'objecte de referència. El símbol  $u$  denota el llindar i prenem  $u = 0.5$ .

Si considerem la igualtat de l'Eq. (6.1),

$$\left| \frac{cc}{ac} - 1 \right| = u, \quad (6.2)$$

es defineixen els límits de reconeixement superior i inferior. De manera compacta en l'Eq. (6.1), s'expressa que el llindar té en compte possibles pics de correlacions

creuades tant superiors com inferiors al pic d'autocorrelació. És a dir, amb l'expressió (6.1) establim una banda de reconeixement que té en compte que, en algunes ocasions, els màxims de les correlacions creuades poden superar el màxim de l'autocorrelació de l'objecte de referència. Aquest cas el podem trobar, per exemple, quan apliquem el procés de reconeixement per a la identificació d'objectes en nivell de gris o color (veure apartat 2.4.3 del capítol 2).

Comencem l'apartat 6.2 amb l'estudi de les possibilitats discriminatives que presenta la DNC davant les variacions de forma dels objectes a analitzar. Continuarem l'estudi de l'aplicació de la DNC en l'apartat 6.3 pel cas de considerar objectes amb diferències en el seu nivell de gris. En ambdós apartats, el 6.2 i el 6.3, s'ha seguit la mateixa metodologia: en primer lloc, introduïm els tests. Seguidament, estudiem els resultats proporcionats pels mètodes de filtratge convencionals, concretament el CMF, el POF, l'IF i el PPC. I per últim, presentem els resultats de l'aplicació de la DNC.





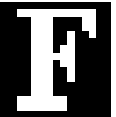
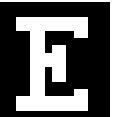
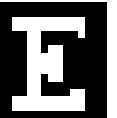

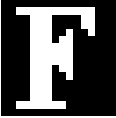
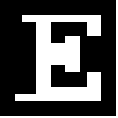
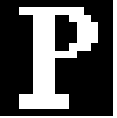



## **6.2 Variacions de forma entre els objectes a analitzar**

En aquesta secció realitzarem un estudi de l'aplicació de la DNC en el reconeixement d'objectes que poden presentar variacions en la forma. L'objectiu serà obtenir un sistema que pugui adaptar-se a diferents graus de discriminació respecte a aquesta característica. Partirem d'unes escenes test elaborades amb la finalitat d'analitzar el comportament de la DNC al variar els paràmetres  $L$  i  $M$  que governen les no-linealitats aplicades, quan només existeixen diferències entre els objectes degudes a la forma.

### **6.2.1 Descripció del test 1**

En aquest apartat considerem un test format per un conjunt d'objectes que configuraran un total de 14 escenes segmentades. Els objectes del test es mostren en

la Figura 6.1 i consisteixen en caràcters (lletres) binaris sobre fons negre. S'han dissenyat de manera que tots els objectes tinguin la mateixa àrea i, per tant, siguin equienergètics, amb la intenció de facilitar la comparació dels resultats de reconeixement que s'obtinguin a partir d'ells.

Etiqueta	$E_{ref}$	$E_1$	$E_2$	$F_1$	$F_2$	$E_3$	$E_4$
Objecte (lletra)							
Producte escalar $cs = \frac{\bar{s} \cdot \bar{r}}{ \bar{r} ^2}$	1.0	0.95	0.92	0.89	0.86	0.85	0.84
Etiqueta	$E_5$	$F_3$	$E_6$	$P_1$	$R_1$	$F_4$	$E_7$
Objecte (lletra)							
Producte escalar $cs = \frac{\bar{s} \cdot \bar{r}}{ \bar{r} ^2}$	0.84	0.84	0.77	0.76	0.74	0.73	0.70

**Figura 6.1.** Test 1: Escenes binàries i equienergètiques per a l'anàlisi de la DNC davant variacions en la forma dels objectes.

La primera lletra ( $E_{ref}$ ) es pren com a patró a identificar. La resta de caràcters estan ordenats segons un paràmetre de similitud respecte al patró de referència. Com paràmetre orientatiu del grau de semblança s'ha pres el coeficient de similitud  $cs$  que es defineix com el producte escalar entre l'escena a analitzar ( $\bar{s}$ ), i el caràcter de referència ( $\bar{r}$ ), ambdós considerats com a vectors, normalitzat al valor del mòdul al quadrat de la referència [206]. L'expressió matemàtica corresponent és

$$cs = \frac{\bar{s} \cdot \bar{r}}{|\bar{r}|^2}. \quad (6.3)$$

El paràmetre coeficient de similitud ( $cs$ ) s'ha establert en base al solapament entre una lletra i el caràcter de referència ( $E_{ref}$ ). Un valor de  $cs$  proper a 1 indica una gran semblança entre la lletra a analitzar i la referència, mentre que valors baixos de  $cs$  són indicatius de majors diferències entre els caràcters comparats.

En el propòsit de reconeixement per correlació òptica, el coeficient de similitud pot ser acceptat tenint en compte que la correlació clàssica establerta per Vander Lugt mesura el grau de semblança entre dos objectes en tant el seu màxim de solapament. El màxim del pic de correlació (en amplitud) obtingut en el cas del CMF coincideix amb el  $cs$  definit mitjançant l'Eq. (6.3). No obstant, admetem que el coeficient  $cs$  s'ha d'utilitzar de manera orientativa i amb certes precaucions. Per exemple, un petit desplaçament d'una lletra dins el quadre de fons negre a la Figura 6.1 faria variar considerablement el seu valor en el coeficient de similitud respecte a la referència. Aquest desplaçament, per contra, no alteraria el màxim del pic de correlació establert per Vander Lugt. Per aquesta raó, en el nostre cas, hem tingut en compte el màxim solapament que es podria obtenir entre una lletra i la referència de manera que el valor  $cs$  fos el màxim possible. En la Figura 6.1, s'ha explicitat aquest valor màxim del paràmetre  $cs$  obtingut per cada lletra test.

Observem en la Figura 6.1 que l'interval de valors de  $cs$  considerats oscil·la entre 0.95 i 0.70, de manera que s'inclouen lletres E molt semblants a la referència (per exemple,  $E_1$  i  $E_2$ ), lletres F amb gran similitud ( $F_1$  i  $F_2$ ), i fins a lletres més diferents com poden ser la R o la P. Aquesta elecció d'objectes amb diferent grau de similitud respecte a la referència ens permetrà provar fins a on arriba la capacitat discriminativa del sistema i estudiar la seva variació amb la DNC aplicada.

### 6.2.2 Mètodes de filtratge convencionals (CMF, POF, IF, PPC)

Considerem primerament, un sistema de correlació on s'utilitzen tècniques convencionals de filtratge adaptat per al reconeixement. Per exemple, són prou coneguts els casos del CMF, el POF, el IF i el PPC amb capacitat discriminativa diferent.

Els resultats corresponents a aquests quatre mètodes de filtratge a l'intentar reconèixer l'objecte de referència en els objectes del test 1 es mostren en la Taula 6.1. Els valors de correlació donats en la Taula 6.1 corresponen a la correlació creuada normalitzada, és a dir, al quocient  $cc/ac$ , on  $cc$  és el màxim del pic de correlació creuada entre l'escena a analitzar i la referència, i  $ac$  és el màxim del pic d'autocorrelació de la referència amb ella mateixa. Establim un llindar  $u=0.5$ , a partir del qual si la raó  $cc/ac$  el supera, o bé és inferior a 1.5, es considera que la lletra analitzada és similar a la referència i, com a tal, apareix al final del procés de reconeixement. En canvi, per qualsevol altra valor de la raó  $cc/ac$ , el valor del màxim de correlació creuada normalitzada queda fora de la banda de reconeixement establerta pel doble llindar i la lletra test és considerada suficientment diferent de la referència, essent rebutjada en el procés de reconeixement. Aquest criteri de decisió, pel fet d'aplicar-se sobre els resultats de correlació d'objectes equienergètics, és un cas particular inclòs en el criteri general definit per l'Eq. (6.1).

Els resultats de reconeixement, també els representarem de forma més compacta utilitzant una única escena contenint tots els caràcters test. D'aquesta manera, la Figura 6.2 mostra les lletres contingudes en el test 1 que superen satisfactòriament el procés de reconeixement, és a dir, que el valor màxim del corresponent pic de correlació està inclòs en la banda de reconeixement establerta pel doble llindar. La resta de caràcters són rebutjats, i com a conseqüència no apareixen en la figura. En la Figura 6.2 es mostren els resultats pels quatre mètodes de filtratge considerats. Cal comentar, que els plans representats en la Figura 6.2 no corresponen al que seria pròpiament el pla de sortida d'un sistema de reconeixement. Si així fos, en les posicions on s'ha detectat un pic de correlació, el valor màxim del qual està inclòs en



la banda de reconeixement definida pel doble llinar, hauria d'aparèixer una lletra idèntica a la referència, donat que el sistema les reconeixeria com a iguals. No obstant, per facilitar la comprensió dels resultats, hem optat per representar el caràcter analitzat, i no la referència, en les posicions on es detecten els objectes similars al patró.

**Taula 6.1.** Resultats de correlació per alguns mètodes de filtratge convencionals. Els valors numèrics que apareixen corresponen a la correlació creuada normalitzada ( $cc/ac$ ), i aquells que es troben dins la banda de reconeixement definida pel doble llinar  $u=0.5$ , estan subratllats per indicar que són les lletres reconegudes al final del procés de reconeixement. Els objectes considerats es presenten en la Figura 6.1.

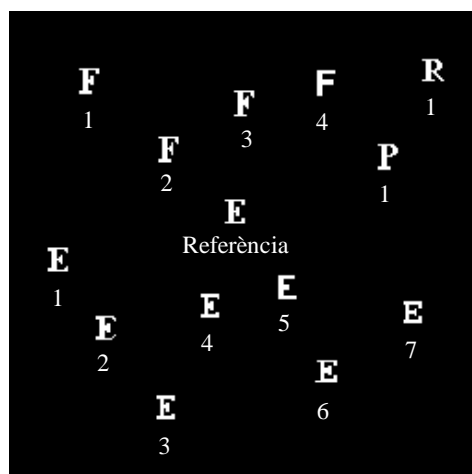
<b>Obj.</b>	<b>E<sub>ref</sub></b>	<b>E<sub>1</sub></b>	<b>E<sub>2</sub></b>	<b>F<sub>1</sub></b>	<b>F<sub>2</sub></b>	<b>E<sub>3</sub></b>	<b>E<sub>4</sub></b>	<b>E<sub>5</sub></b>	<b>F<sub>3</sub></b>	<b>E<sub>6</sub></b>	<b>P<sub>1</sub></b>	<b>R<sub>1</sub></b>	<b>F<sub>4</sub></b>	<b>E<sub>5</sub></b>
<b>cs</b>	<b>1.0</b>	<b>0.95</b>	<b>0.92</b>	<b>0.89</b>	<b>0.86</b>	<b>0.85</b>	<b>0.84</b>	<b>0.84</b>	<b>0.84</b>	<b>0.77</b>	<b>0.76</b>	<b>0.74</b>	<b>0.73</b>	<b>0.70</b>
<b>CMF</b>	<u>1</u>	<u>0.90</u>	<u>0.85</u>	<u>0.79</u>	<u>0.75</u>	<u>0.73</u>	<u>0.71</u>	<u>0.71</u>	<u>0.71</u>	<u>0.60</u>	<u>0.57</u>	<u>0.55</u>	<u>0.54</u>	<u>0.50</u>
<b>POF</b>	<u>1</u>	<u>0.74</u>	<u>0.62</u>	<u>0.61</u>	<u>0.53</u>	<u>0.51</u>	0.45	0.34	0.40	0.30	0.31	0.30	0.18	0.18
<b>IF</b>	<u>1</u>	<u>0.52</u>	0.40	0.49	0.36	0.36	0.29	0.09	0.19	0.12	0.20	0.18	0.04	0.04
<b>PPC</b>	<u>1</u>	0.39	0.31	0.41	0.29	0.29	0.18	0.09	0.17	0.08	0.15	0.12	0.03	0.03

Dels resultats obtinguts i presentats en la Taula 6.1 i en la Figura 6.2, veiem que pel filtre CMF, el més tolerant a canvis de forma, tots els caràcters considerats es confonen amb la referència quan el llinar aplicat és del 50% ( $u=0.5$ ). La raó entre  $cc/ac$  va decreixent de manera directament proporcional al  $cs^2$  degut a què el valor que es dona al coeficient de similitud, al igual que el filtre clàssic, mesura el màxim solapament entre dos caràcters. La relació amb el quadrat del  $cs$  és deguda a què la mesura dels pics de correlació la fem en intensitat. Donat que el filtre CMF és lineal amb la transmitància de l'escena i de la referència i que les lletres analitzades són equienergètiques, els resultats numèrics de correlació apareixen en la Taula 6.1

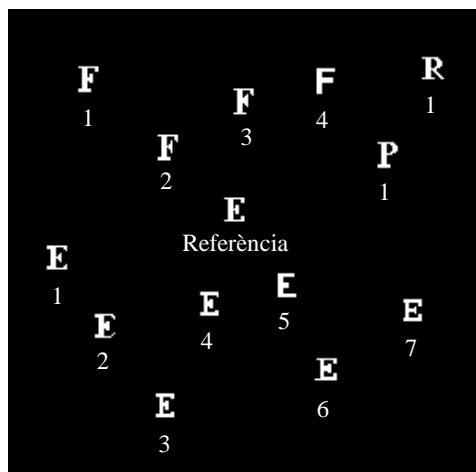
ordenats segons el valor del coeficient de similitud  $cs$ . Així doncs, si ens proposem variar la capacitat discriminativa del sistema quan apliquem el filtre CMF, podem pensar en modificar el valor llindar  $u$  de manera que obtinguem un sistema amb una capacitat discriminativa més acurada i, per tant, passarem a reconèixer un nombre menor de lletres que coincidiran amb les més similars a la referència segons el paràmetre  $cs$ , o bé, podem aconseguir un sistema més tolerant i permetem així, reconèixer un nombre més elevat de lletres amb diferències amb la referència cada vegada més apreciables.

Si considerem els resultats obtinguts pel filtre POF, veiem que la capacitat discriminativa d'aquest mètode és superior en comparació de l'anterior, i només es detecten les lletres amb coeficient de similitud igual o superior a 0.85. Veiem que si mantenim el llindar de  $u=0.5$ , tant es reconeixen les lletres E més similars al patró, com alguns caràcters F molt semblants a l' $E_{ref}$ . Donat que en el cas del POF no només es té en compte el solapament entre dos objectes, sinó que es dóna major importància a la coincidència dels contorns, els resultats numèrics de reconeixement no són directament proporcionals a la variació de  $cs^2$ , de manera que un canvi en el llindar  $u$  no implica sempre discriminar o reconèixer un major número de lletres en ordre creixent o decreixent al seu coeficient de similitud.

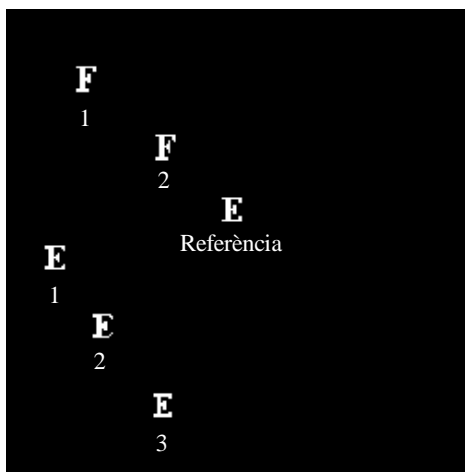
La discriminació del sistema és molt més bona si s'utilitza el filtre invers, IF, en el procés de correlació. Amb el llindar establert al 50% només es reconeix conjuntament amb el patró la lletra més similar en forma, la  $E_1$ . Observem que en aquest cas, si relaxéssim lleugerament el llindar per tenir més tolerància a petits canvis en la forma, passa a ésser reconeguda primer la lletra  $F_1$ , per davant de la  $E_2$ , tot i tenir aquesta darrera un coeficient de similitud major. Aquest fet, és conseqüència de la no-linealitat del filtratge adaptat utilitzat del filtre invers.



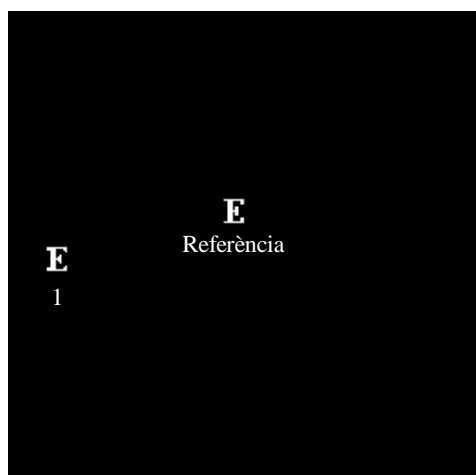
(a)



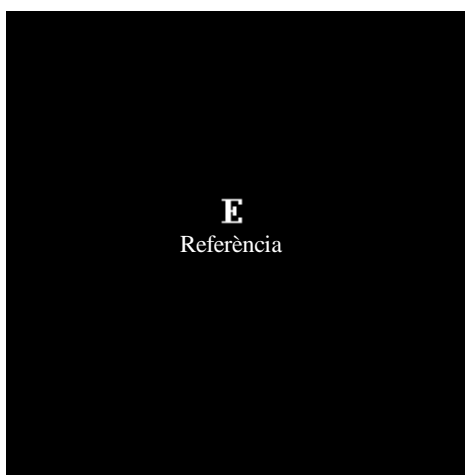
(b)



(c)



(d)



(e)

**Figura 6.2.** Resultats de reconeixement per diferents mètodes convencionals de filtratge a l'analitzar les lletres del test 1 (Figura 6.1). (a) Test 1. (b) Caràcters que es reconeixen a l'utilitzar el CMF, (c) ídem pel POF, (d) ídem per l'IF, i (e) ídem pel PPC.

Per últim, el PPC és el mètode, dins dels quatre convencionals considerats, amb una capacitat discriminativa més alta, ja que només es reconeix la lletra E de referència, i la resta de lletres són rebutjades amb valors de  $cc/ac$  relativament baixos. Un petit canvi en el llindar de reconeixement tampoc aconseguiria una variació suau de la capacitat discriminativa del sistema.

Amb aquest exemple, posem de manifest que, tot i disposar d'un conjunt de mètodes de filtratge adaptat que proporcionen diferents nivells de discriminació, el sistema de reconeixement no tindrà gaire flexibilitat en poder variar de manera contínua i adaptar aquesta capacitat de discriminació a un problema concret.

### 6.2.3 Aplicació de la DNC

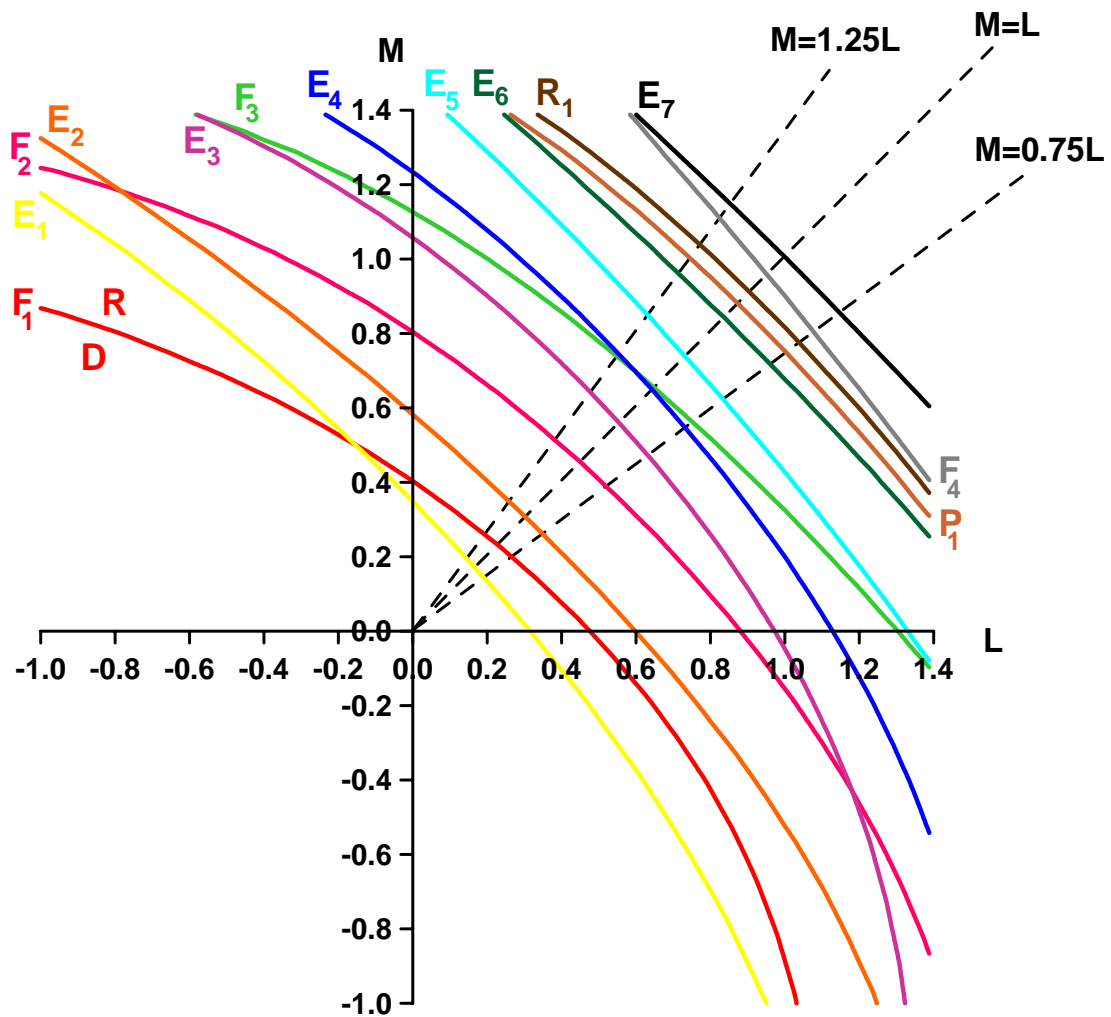
Analitzem ara els resultats obtinguts si volem resoldre el mateix problema aplicant la correlació no-lineal dual. Aquest mètode representa una solució més àmplia del problema que, com s'ha vist en el capítol anterior (capítol 4), també inclou els mètodes de filtratge convencionals tot just analitzats en la secció 6.2.2. Els paràmetres reals  $L$  i  $M$  en l'Eq. (4.27) són els que ens permeten modificar de manera contínua el tipus de processament no-lineal. Recordem també que  $L$  constitueix la no-linealitat aplicada a l'escena a analitzar, i  $M$  a l'objecte de referència. Ens restringirem en un interval de valors d'aquests paràmetres. En concret, hem aplicat la DNC en l'interval  $L, M \in [-1, 1.4]$ , considerant salts de 0.2. L'interval escollit inclou els mètodes convencionals pels valors: CMF ( $L = 1, M = 1$ ), POF ( $L = 1, M = 0$ ), IF ( $L = 1, M = -1$ ) i PPC ( $L = 0, M = 0$ ).

Primerament, hem considerat condicions experimentals ideals per a l'aplicació de la DNC, és a dir, les distribucions espectrals de freqüències espacials es registren, es processen i es reproduïxen sense considerar efectes de rang dinàmic limitat (saturació) ni de quantització dels dispositius experimentals.

Una representació compacta dels resultats de reconeixement obtinguts en l'interval escollit, es mostra en la Figura 6.3. Per portar a terme la discriminació o el reconeixement d'una lletra en front del patró, hem fixat el valor llindar d' $u = 0.5$ , que

defineix el que anomenem el límit de l'àrea de reconeixement. Donat que les lletres estudiades són equienergètiques, els valors corresponents a llur correlació creuada normalitzada mai superen la unitat i, per aquest motiu, les línies límits corresponents al llindar superior de les bandes de reconeixement ( $cc/ac = 1.5$  en l'Eq. 6.2) no apareixen en la Figura 6.3. Així doncs, per valors del màxim de la correlació creuada normalitzada superiors o iguals al llindar ( $cc/ac \geq 0.5$ ), obtenim reconeixement de la lletra analitzada. En el cas contrari, per valors inferiors al llindar ( $cc/ac < 0.5$ ), la lletra es discrimina de la referència. En l'espai bidimensional generat pels paràmetres  $L$  i  $M$ , hem representat per a cada lletra test, la línia que uneix els punts  $(L, M)$  pels quals s'obté el límit de l'àrea de reconeixement, és a dir, pels quals s'obté  $cc/ac = 0.5$ . Les línies s'han traçat per interpolació a partir dels valors obtinguts variant  $L$  i  $M$  en increments de 0.2. Per tant, es divideix l'espai  $(L, M)$ , entre la regió on el caràcter considerat s'identifica com el patró buscat, i la regió on el mateix caràcter seria rebutjat. En la Figura 6.3, per a una de les línies representades, concretament per la de la lletra  $F_1$ , s'afegeix la informació de quina zona correspon al reconeixement del caràcter (amb una lletra  $R$ ), i quina a la zona de discriminació ( $D$ ). Per a cada caràcter pertanyent al conjunt d'escenes test, s'ha representat la corresponent línia límit de reconeixement amb un color diferent. L'orientació de les zones de reconeixement i discriminació per totes elles és la mateixa que la indicada per la línia de la lletra  $F_1$ .

Observem que les línies  $cc/ac = 0.5$  representades en la Figura 6.3 apareixen aproximadament ordenades segons el valor del coeficient de similitud  $cs$  calculat per a cada un dels caràcters de la Figura 6.1. Donat un caràcter amb coeficient de similitud proper a 1 (gran semblança amb la referència), és necessari utilitzar valors positius baixos d' $L$  i  $M$  (entre 0.2 i 0), o bé utilitzar valors negatius d'aquest paràmetres, per poder-lo diferenciar del patró. En canvi, com a cas contrari, si es vol reconèixer com a l'objecte buscat, una lletra amb coeficient de similitud relativament baix, cal recórrer a valors alts d' $L$  i  $M$  ( $L, M \geq 1$ ) per aconseguir aquest objectiu.



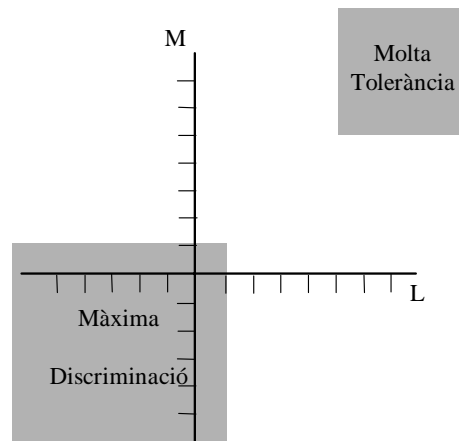
**Figura 6.3.** Resultats de reconeixement mitjançant DNC en condicions experimentals ideals. Representació de les línies llindars de reconeixement pel valor  $cc/ac = 0.5$ , per a cada una de les lletres que conformen el conjunt test 1 (Figura 6.1). En la línia corresponent a la lletra  $F_1$  s'indiquen la zona de reconeixement (R) i la zona de discriminació (D) d'aquesta lletra respecte a la referència ( $E_{ref}$ ). L'orientació d'aquestes zones és la mateixa per la resta de corbes.

Delimitem doncs, una regió en l'espai  $(L, M)$  on trobem màxima discriminació dels caràcters equienergètics amb forma diferent, i una altra regió on la tolerància a canvis de forma és més gran. Aquestes dues regions diferenciades les representem sobre el gràfic de la Figura 6.4.

La posició d'aquestes dues zones en el pla  $LM$ , la podem entendre tenint en compte el resultat d'aplicar les diferents no-linealitats sobre la informació continguda a

l'escena i a la referència. És a dir, si volem discriminar dues lletres molt similars en forma, caldrà potenciar les petites diferències entre els caràcters comparats. Això s'aconsegueix ressaltant les altes freqüències on està continguda la informació dels contorns i dels detalls dels objectes. Els valors dels paràmetres  $L$  i  $M$  que més afavoreixen les altes freqüències atenuant les baixes freqüències, corresponen a valors nuls o negatius. En el cas contrari, trobem la situació en què es volen tolerar canvis més importants en la forma dels objectes comparats. La regió d' $L$  i  $M$  amb valors positius iguals i superiors a la unitat es caracteritza per l'amplificació de les freqüències més baixes, generalment més energètiques per si mateixes, separant encara més la diferència energètica amb les altes freqüències. Observem també, a partir del gràfic Figura 6.3, que la comparació de dos caràcters equienergètics per aconseguir la seva confusió o discriminació en el procés de reconeixement, s'ha de portar a terme aplicant aproximadament el mateix tipus de no-linealitat tant a l'escena a analitzar com a l'objecte de referència. Això es veu també en la Figura 6.4 donat que totes dues àrees, l'àrea de molta tolerància i la de màxima discriminació, estan centrades sobre l'eix  $L = M$ . Pels motius que acabem d'explicar es pot entendre que cal comparar el mateix tipus d'informació dels objectes considerats per poder aplicar un criteri de decisió.

Fixem-nos en la regió  $(-0.2 < L < 1.4, -0.2 < M < 1.4)$ , que inclou el quadrant d' $L$  i  $M$  positius, en la Figura 6.3. Veiem que és possible recórrer de manera contínua aquesta regió i obtenir una capacitat discriminativa variant i progressiva des de la màxima discriminació fins a una tolerància molt gran. D'un extrem a l'oposat de l'interval, aconseguim passar de la discriminació total de la lletra de referència davant la resta de caràcters, fins a obtenir la identificació de tots ells amb la lletra de referència. Per exemple, si escollim la recta  $L = M$ , que implica la mateixa no-linealitat aplicada a l'escena i a la referència, podem obtenir un eix de variació molt ampli en la capacitat de discriminació del sistema. Altres rectes no simètriques, com les que compleixen l'equació  $M = kL$ , amb  $k \in [0.75, 1.25]$ , tenen propietats similars (Figura 6.3).



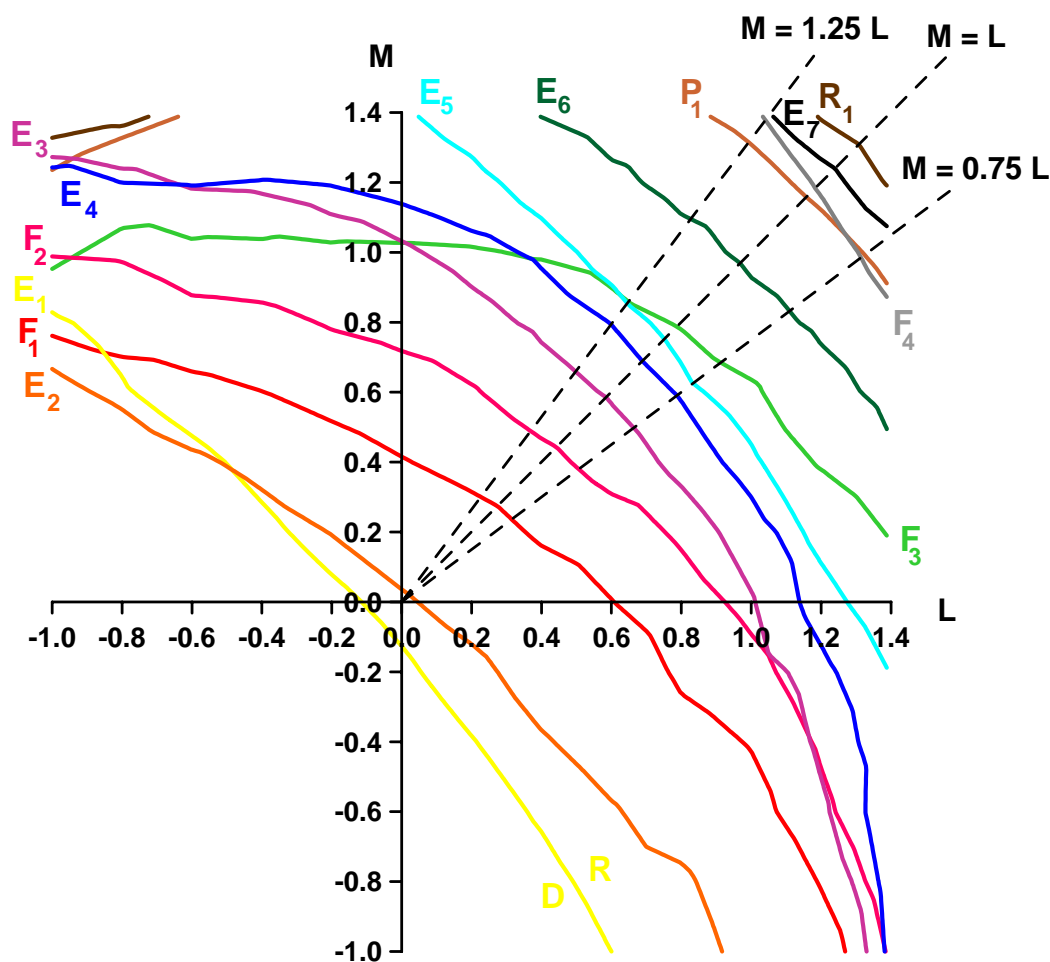
**Figura 6.4.** Delimitació de les zones de màxima discriminació, i màxima tolerància a canvis de forma en els caràcters analitzats (Test 1, Figura 6.1) mitjançant la DNC.

La representació de la Figura 6.3 es pot obtenir també pel cas on tinguem en compte els condicionaments experimentals, ja mencionats en el capítol 5. Això vol dir, permetre un nivell de saturació dels sensors CCD de la càmera durant l'enregistrament dels espectres ( $\alpha=10\%$ ), considerar el sistema de captació amb un nombre limitat de nivells de gris (256) amb els quals es quantitza la informació enregistrada, i finalment, suposant un nombre finit de nivells de gris (32) per a la reproducció en el modulador espacial de llum de l'espectre conjunt modificat no-linealment (MJPS) durant el segon pas de la DNC. Els resultats corresponents als de la Figura 6.3 però modificats al tenir en compte les condicions experimentals, es mostren en la Figura 6.5.

Observem que el comportament és molt similar en ambdues figures (Figura 6.3 i Figura 6.5), sobretot per les lletres amb coeficient de similitud intermig, que pràcticament romanen invariants. En els casos més extrems, els de lletres molt semblants a la referència, i els de les lletres més diferents, són on trobem més separació entre les línies del llinard obtingudes en condicions ideals i les obtingudes tenint en compte les condicions experimentals. En general, les línies de  $cc/ac = 0.5$  tendeixen a separar-se entre elles quan introduïm les restriccions experimentals, aconseguint que facin falta valors més baixos d' $L$  i  $M$  ( $L$  i  $M < 0$ ) per a discriminar



les lletres més semblants, i valors més alts dels paràmetres ( $L$  i  $M \geq 1.4$ ) per reconèixer les més diferents. Aquest comportament pot ésser degut a la pèrdua d'informació en les distribucions espectrals capturades, tant pel procés de quantització que elimina les freqüències més altes, com el procés de saturació que uniformitza les freqüències més baixes. D'aquesta manera, amb una limitació més gran de la informació a processar, cal anar a no-linealitats més extremes per accentuar les diferències de caràcters molt semblants i aconseguir discriminar-los, o bé trobar les similituds entre caràcters més diferents i poder així identificar-los com a la referència buscada.



**Figura 6.5.** Resultats de reconeixement mitjançant DNC al considerar una saturació d'  $\alpha=10\%$ , una quantització del sistema de captació de 256 nivells de gris, i 32 nivells de gris de l'SLM. Representació de les línies llindar de reconeixement pel valor  $cc/ac = 0.5$ , per a cada una de les lletres que conformen el conjunt test de la Figura 6.1. En la línia corresponent a la lletra  $E_1$  s'indiquen la zona de reconeixement (R) i la zona de discriminació (D) d'aquesta lletra respecte a la referència ( $E_{ref}$ ). L'orientació d'aquestes zones és la mateixa per la resta de corbes.

Tot i les discrepàncies que acabem d'esmentar, els resultats ens indiquen que tenint en compte les condicions experimentals, també és possible trobar rectes sobre les quals, la capacitat de discriminació del sistema variï gradualment des d'una discriminació absoluta de tots els caràcters davant de la referència, fins a la identificació de tots ells com a similars al patró.

A tall d'exemple, indicarem sobre la recta marcada en la Figura 6.5, d'equació  $M = 0.75 L$ , els resultats de reconeixement i discriminació que obtindríem en diferents punts (Taula 6.2). En primer lloc, destacar que amb la recta escollida, podem variar de manera contínua i gradual la capacitat de discriminació del sistema de reconeixement. D'aquesta manera, segons la tasca de reconeixement que ens plantegem, en un únic correlador podem escollir el valor dels paràmetres  $L$  i  $M$  característics de la no-linealitat aplicada, i controlar d'aquesta manera el grau de tolerància del sistema. Aquesta és un clar avantatge davant dels mètodes de filtratge convencionals, com els que hem considerat en l'apartat 6.2.2, que restringeixen severament les possibilitats d'adaptació de la capacitat discriminativa del sistema de reconeixement.

També és necessari esmentar que, tot i haver considerat un conjunt d'eixos definits per  $M = k L$  amb  $k \in [0.75, 1.25]$  que permeten una variació progressiva des de la discriminació total de la referència fins a una tolerància gran a les variacions de forma, aquesta elecció no és única i podem optar per eixos amb altres característiques. Per exemple, amb eixos horitzontals amb  $M = \text{constant}$ , o verticals amb  $L = \text{constant}$ , la variació de la capacitat discriminativa del sistema pot ser més lenta. Fins i tot, podríem modificar les no-linealitats aplicades sense variar en gran mesura el grau de discriminació del sistema de reconeixement. Aquest seria el cas si prenguéssim eixos del tipus  $M = -k L + c$ , de manera que escollint apropiadament els valors de les constants  $k$  i  $c$ , la capacitat de discriminació del sistema pot ser quasi constant.

**Taula 6.2.** Resultats de reconeixement de la DNC sobre la recta  $M = 0.75 L$  tenint en compte condicions experimentals.



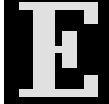




<b>Paràmetres (<math>L, M</math>)</b>	<b>Lletres que es reconeixen</b>	<b>Paràmetres (<math>L, M</math>)</b>	<b>Lletres que es reconeixen</b>
<b>(-0.2, -0.15)</b>	$E_{\text{ref}}$	<b>(0.8, 0.6)</b>	$E_{\text{ref}}, E_1, E_2, F_1, F_2, E_3, E_4$
<b>(0, 0)</b>	$E_{\text{ref}}, E_1$	<b>(1, 0.75)</b>	$E_{\text{ref}}, E_1, E_2, F_1, F_2, E_3, E_4, E_5, F_3$
<b>(0.2, 0.15)</b>	$E_{\text{ref}}, E_1, E_2$	<b>(1.2, 0.9)</b>	$E_{\text{ref}}, E_1, E_2, F_1, F_2, E_3, E_4, E_5, F_3, E_6$
<b>(0.4, 0.3)</b>	$E_{\text{ref}}, E_1, E_2, F_1$	<b>(1.4, 1.05)</b>	$E_{\text{ref}}, E_1, E_2, F_1, F_2, E_3, E_4, E_5, F_3, E_6, P_1, F_4, E_7$
<b>(0.6, 0.45)</b>	$E_{\text{ref}}, E_1, E_2, F_1, F_2$	<b>(1.6, 1.2)</b>	$E_{\text{ref}}, E_1, E_2, F_1, F_2, E_3, E_4, E_5, F_3, E_6, P_1, F_4, E_7, R_1$

### 6.3 Variacions d'intensitat (nivell de gris) entre els objectes a analitzar

#### 6.3.1 Descripció del test 2

Considerem en aquest apartat el test 2 format per el conjunt d'escenes mostrades en la Figura 6.6. Totes les lletres representades en el test 2 tenen idèntica forma a la lletra de referència, que a la vegada, coincideix amb la referència del test 1 (Figura 6.1). Les diferències entre les lletres del test 2 radiquen en el nivell de gris amb el qual s'ha recobert cada un dels caràcters. En la mateixa figura, i com a subíndex de les lletres considerades, es dona el valor numèric del nivell de gris corresponent, que varia entre el 255 de la referència i el 50 de la lletra menys energètica.

També en aquesta ocasió, és possible establir una mesura de la semblança entre els caràcters i el patró a partir del producte escalar normalitzat (Eq. 6.3). Aquest paràmetre no serà representatiu ara del solapament entre les dues lletres comparades, donat que sempre serà màxim, però sí que ens proporciona una mesura de la diferència energètica entre els dos caràcters considerats. De fet, si comparem el coeficient de similitud obtingut per cada lletra (Figura 6.6), podem observar que coincideix numèricament amb el quocient entre els nivells del gris del caràcter test i la referència, respectivament.

Etiqueta	$E_{255 \text{ (ref)}}$	$E_{240}$	$E_{225}$	$E_{200}$	$E_{150}$	$E_{100}$	$E_{50}$
Objecte (lletra)							
Producte escalar $cs = \frac{\bar{s} \cdot \bar{r}}{ \bar{r} ^2}$	1.0	0.94	0.88	0.78	0.59	0.39	0.20

**Figura 6.6.** Test 2: lletres amb variacions en el nivell de gris.  $E_{255 \text{ (ref)}}$  coincideix amb  $E_{\text{ref}}$  del test 1 (Figura 6.1) i constitueix el patró de referència.

Amb els objectes del test 2, seguim el mateix procediment que en la secció anterior. Analitzarem, primerament, el comportament dels mètodes de filtratge convencionals (CMF, POF, IF, PPC) amb aquest conjunt d'escenes test. Observarem les seves possibilitats en quant a la capacitat discriminativa del sistema, i plantejarem i comprovarem la utilitat de la DNC en aquest problema de reconeixement.

### 6.3.2 Mètodes de filtratge convencionals (CMF, POF, IF, PPC)

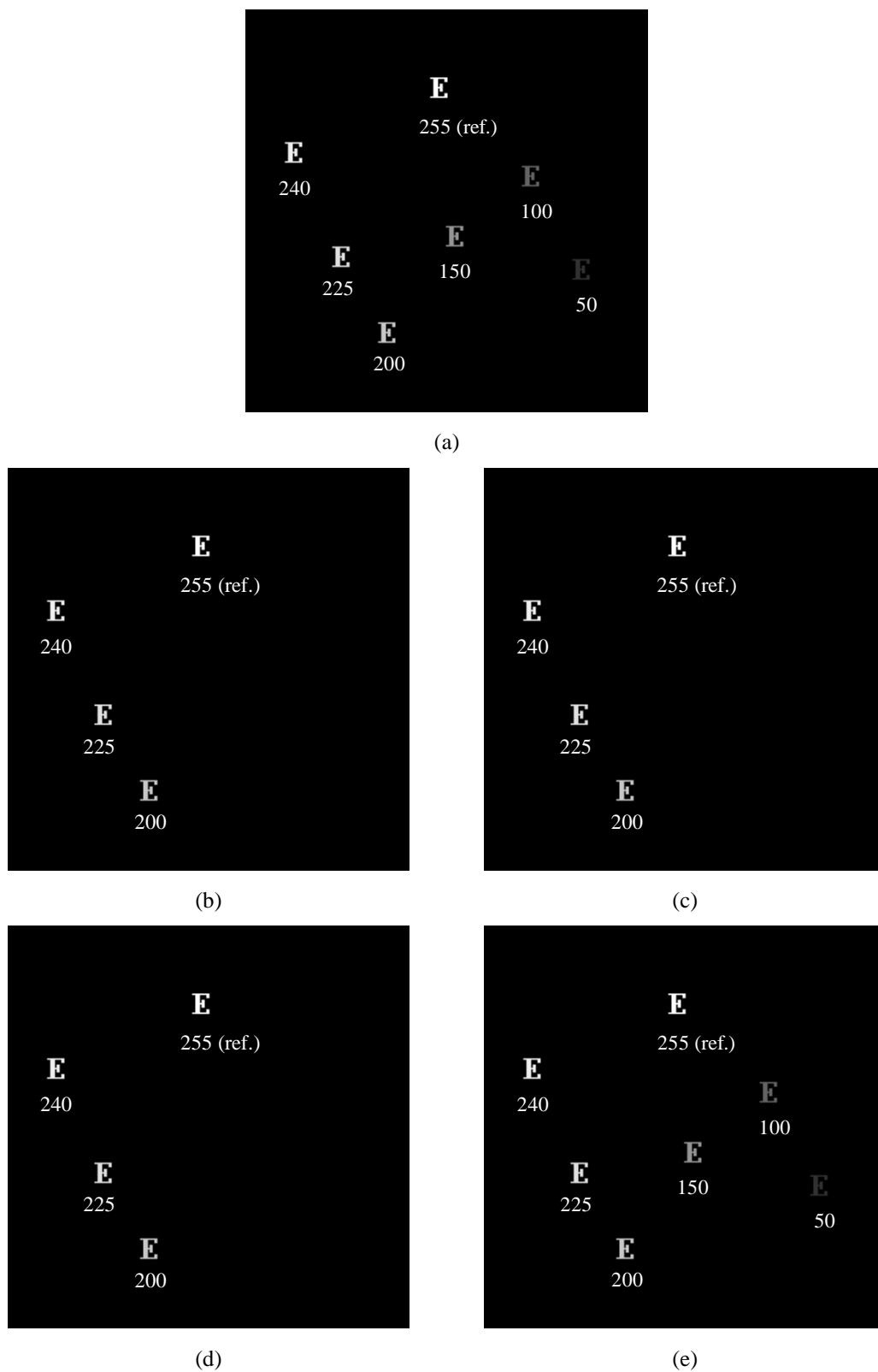
La Taula 6.3 reuneix els resultats numèrics obtinguts en el procés de reconeixement utilitzant els mètodes de filtratge convencionals CMF, POF, IF, PPC. Els valors de la taula corresponen al màxim de la correlació creuada normalitzada ( $cc/ac$ ) i,

apareixen subratllats aquells que es troben dins la banda de reconeixement definida pel llindar establert d' $u=0.5$ . Els mateixos resultats es mostren de manera gràfica en la Figura 6.7 per facilitar-ne la seva visualització.

**Taula 6.3.** Resultats de correlació per alguns mètodes de filtratge convencionals. Els valors numèrics que apareixen corresponen al màxim de correlació creuada normalitzada ( $cc/ac$ ), i aquells que es troben dins la banda de reconeixement definida pel doble llindar  $u=0.5$ , estan subratllats per indicar que són les lletres identificades al final del procés de reconeixement. Els objectes considerats es presenten en la Figura 6.6.

<b>Objecte</b>	<b><math>E_{255}</math> (ref)</b>	<b><math>E_{240}</math></b>	<b><math>E_{225}</math></b>	<b><math>E_{200}</math></b>	<b><math>E_{150}</math></b>	<b><math>E_{100}</math></b>	<b><math>E_{50}</math></b>
<b><i>cs</i></b>	<b>1.0</b>	<b>0.94</b>	<b>0.88</b>	<b>0.78</b>	<b>0.59</b>	<b>0.39</b>	<b>0.20</b>
<b>CMF</b>	<u>1</u>	<u>0.89</u>	<u>0.78</u>	<u>0.62</u>	0.35	0.15	0.04
<b>POF</b>	<u>1</u>	<u>0.89</u>	<u>0.78</u>	<u>0.62</u>	0.35	0.15	0.04
<b>IF</b>	<u>1</u>	<u>0.89</u>	<u>0.78</u>	<u>0.62</u>	0.35	0.15	0.04
<b>PPC</b>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>

Dels resultats obtinguts, és de destacar el fet que els filtres CMF, POF i IF donen idèntics resultats de reconeixement. Recordem que tots ells pertanyen al grup més genèric anomenat filtres de potències fraccionàries (FPF) [6], amb la característica comú que el màxim de correlació és lineal amb l'energia de l'escena d'entrada. Per això, hem demostrat en el capítol 3 (apartat 3.3), que si s'apliquen els filtres FPF pel reconeixement d'objectes d'igual forma amb nivell de gris uniforme s'obté un valor relatiu entre els màxims de correlació creuada i d'autocorrelació coincident amb el quocient  $(b/a)^2$ , sent  $b$  el nivell de gris de l'objecte a analitzar i  $a$  el nivell de gris de la referència [168]. Aquest resultat és el mateix per a qualsevol tipus de filtre inclòs en la definició dels FPF, i tenint en compte el valor del coeficient de similitud, veiem que el resultat normalitzat de reconeixement coincideix en el nostre cas amb  $cs^2$ .



**Figura 6.7.** Resultats de reconeixement per diferents mètodes convencionals de filtratge a l'analitzar les lletres del test 2 (Figura 6.6). (a) Test 2. (b) Caràcters que es reconeixen a l'utilitzar el CMF, (c) ídem pel POF, (d) ídem per l'IF, i (e) ídem pel PPC.

De la Taula 6.3 i la Figura 6.7 es dedueix que la capacitat discriminativa del sistema de reconeixement que utilitza els mètodes de filtratge CMF, POF, IF i PPC admet molt poca variació quan els objectes presenten diferències en el seu nivell d'energia. De fet, les possibilitats de variar la capacitat discriminativa del sistema són molt menors davant d'aquest tipus de diferències que no pas de les diferències en forma estudiades en l'apartat 6.2.2.

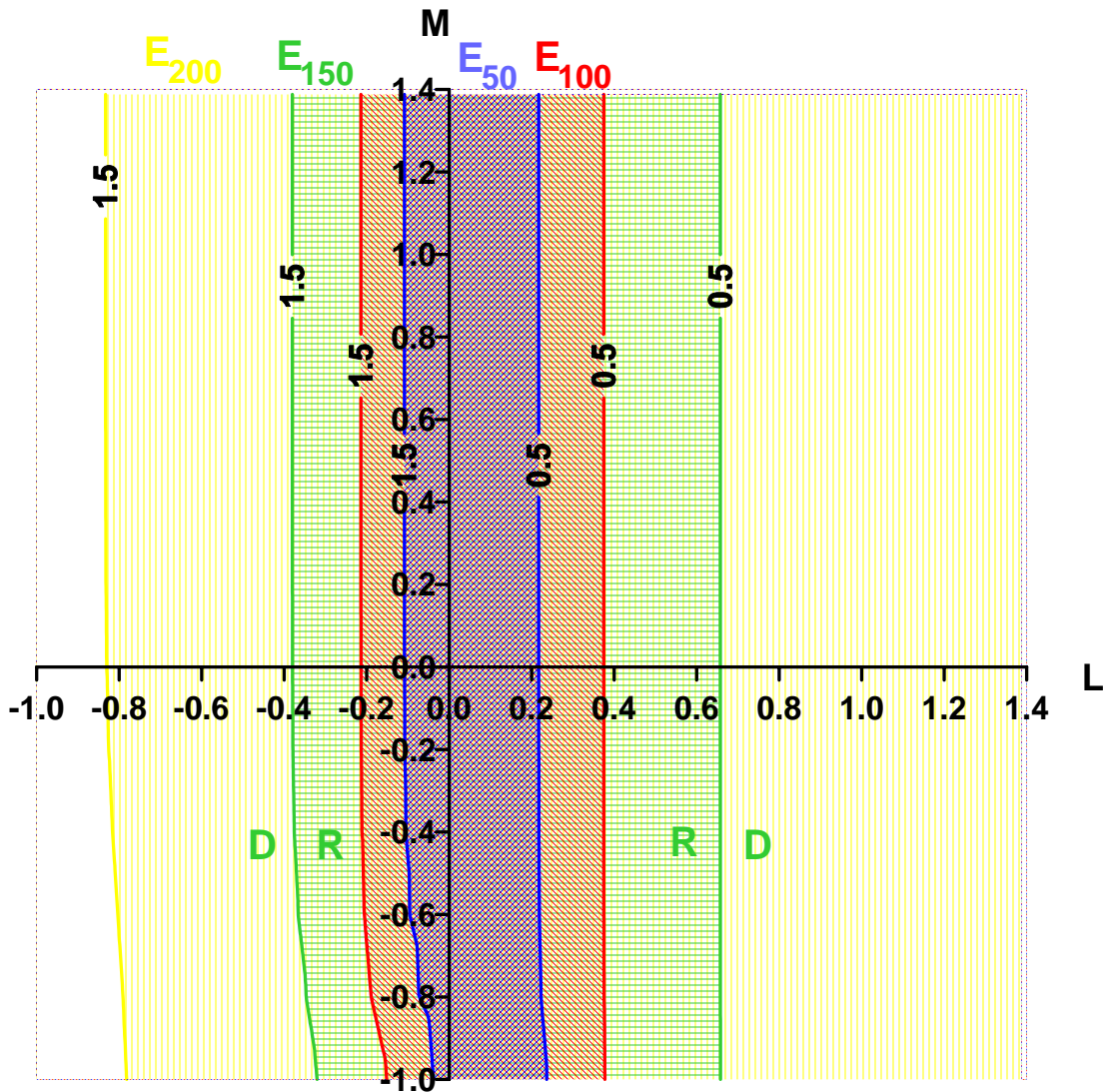
Pels resultats obtinguts en els filtres CMF, POF i IF, observem que la capacitat de discriminació del sistema és exactament la mateixa. L'única possibilitat de variació implicaria la modificació del llindar de reconeixement, fixat inicialment al 50% en qualsevol dels filtres considerats.

En la correlació només de fase (PPC), al processar només la informació de fase de l'escena i de la referència en el procés de reconeixement, s'elimina la informació d'amplitud relacionada amb l'energia dels caràcters. En l'exemple tractat, l'energia és l'única informació distintiva entre les lletres i, per tant, no és sorprenent que els resultats de reconeixement obtinguts per aquest mètode siguin idèntics a la unitat per a qualsevol lletra analitzada. Deduïm doncs, que aquest tipus de filtratge no permet discriminar en un problema de reconeixement on la única informació que diferencia els caràcters entre ells, és el seu nivell de gris.

### 6.3.3 Aplicació de la DNC

Analitzem ara els resultats que s'obtenen amb la DNC en el cas del reconeixement d'objectes d'igual forma però amb nivell de gris uniforme diferent (test 2, Figura 6.6). Partim de la mateixa zona o domini de paràmetres  $L$  i  $M$  que hem utilitzat anteriorment, és a dir, l'interval  $L, M \in [-1, 1.4]$ . Considerem novament salts de 0.2 dins l'interval. En primer lloc, analitzarem un sistema en condicions experimentals ideals, lliure dels efectes de saturació i quantització de les distribucions. Els resultats obtinguts es presenten de manera compacte sobre l'espai  $(L, M)$  (Figura 6.8). En aquesta figura, s'ha assenyalat per a cada lletra les línies límits de reconeixement generades per l'aplicació del doble llindar, prenent  $u = 0.5$  en l'Eq. (6.1). Cada lletra

està representada per un color diferent, i en una d'elles ( $E_{150}$ ) s'ha explicitat quina zona correspon al reconeixement de la lletra (R, banda central), i quina a la discriminació (D, en els laterals). L'orientació d'aquestes dues zones respecte la línia lliendar, coincideix per a totes les lletres analitzades.



**Figura 6.8.** Resultats de reconeixement obtinguts mitjançant DNC a partir del càlcul de la funció  $P'(\vec{v})$  (Eq. 4.27) en condicions experimentals ideals. Representació de les línies lliendar de reconeixement corresponents a  $\left|1 - \frac{cc}{ac}\right| = u$ , amb  $u = 0.5$ , per les lletres que conformen el test 2 de la

Figura 6.6. En la banda de reconeixement corresponent a la lletra  $E_{150}$  s'indica la zona de reconeixement (R) i la zona de discriminació (D) d'aquesta lletra respecte a la referència ( $E_{ref}$ ). L'orientació d'aquestes zones és la mateixa per la resta de corbes.



Recordem que a l'analitzar escenes amb variació en el nivell de gris i aplicar diferent no-linealitat a l'escena i a l'objecte de referència, el resultat del màxim d'autocorrelació de la referència ( $E_{255}$ ) no té perquè ser en tots els casos superior a tots els màxims de les correlacions creuades. La condició de reconeixement establerta en l'Eq. (6.1) amb el valor del llindar fixat a  $u = 0.5$ , limita doblement la zona de reconeixement, de manera que els objectes, el màxim de correlació creuada dels quals satisfaci  $0.5 \leq \frac{cc}{ac} \leq 1.5$ , seran identificats com a la referència. Veiem doncs, en la Figura 6.8 com apareixen per la majoria de lletres dos línies llindar, cada una de les quals porta l'etiqueta corresponent al valor  $cc/ac$  que representa. La zona de reconeixement d'una lletra està constituïda pels valors d' $L$  i  $M$  situats en la zona compresa entre ambdues línies límit.

En la Figura 6.8 s'obté que, fixat un valor d' $L$  en l'operador que actua sobre l'escena, el resultat de reconeixement és pràcticament independent de la no-linealitat aplicada a l'objecte de referència i definida per  $M$ . Per justificar analíticament aquesta afirmació, considerem una escena amb un únic objecte d'igual forma que l'objecte de referència. Ambdós objectes tenen en tota la seva superfície un nivell de gris uniforme, però d'energia diferent per l'escena a analitzar i per l'objecte de referència. Aquesta situació la podem descriure analíticament per les expressions

$$\begin{aligned} r(\vec{x}) &= a \cdot g(\vec{x}), \\ s(\vec{x}) &= b \cdot g(\vec{x}), \end{aligned} \tag{6.4}$$

que descriuen l'objecte de referència i l'escena a analitzar, respectivament, i  $a$  i  $b$  són nombres reals que indiquen la transmitància, uniforme en tota la seva superfície, de cada un d'ells. La funció  $g(\vec{x})$  representa la forma de l'objecte comú contingut a l'escena i també a la referència.

Si sobre aquestes funcions, apliquem la DNC amb la no-linealitat descrita pel paràmetre  $L$  sobre l'escena, i per  $M$  sobre la referència, obtenim

$$\begin{aligned} DNC_{L,M}^{T(\vec{v})=1}\{s(\vec{x}), r(\vec{x})\} &= IFT\left\{\left(b|G(\vec{v})|\right)^{L-1} \cdot b G(\vec{v}) \cdot \left(a|G(\vec{v})|\right)^{M-1} \cdot a G^*(\vec{v})\right\} = \\ &= IFT\left\{b^L \cdot a^M \cdot |G(\vec{v})|^{L+M}\right\}, \end{aligned} \quad (6.5)$$

on  $G(\vec{v}) = TF\{g(\vec{x})\}$ , i s'ha pres el cas més senzill on  $T(\vec{v}) = 1$  en la definició de la DNC (Eq. 4.3).

El resultat de la correlació creuada entre l'escena a analitzar i l'objecte de referència (Eq. 6.5), el podem normalitzar respecte el valor de l'autocorrelació donat per

$$DNC_{L,M}^{T(\vec{v})=1}\{r(\vec{x}), r(\vec{x})\} = IFT\left\{a^{L+M} \cdot |G(\vec{v})|^{L+M}\right\}. \quad (6.6)$$

Si denotem per  $A_g = IFT\left\{|G(\vec{v})|^{L+M}\right\}$  i, tenint en compte la linealitat de la transformada de Fourier, trobem que el valor del màxim d'intensitat de la correlació creuada normalitzada entre les funcions  $s(\vec{x})$  i  $r(\vec{x})$  compleix la relació

$$\frac{cc}{ac} = \left(\frac{b}{a}\right)^{2L}, \quad (6.7)$$

que ens indica que, en el cas considerat d'objectes d'igual forma i color uniforme, la DNC proporciona uns pics de les correlacions creuades normalitzades depenents únicament del paràmetre  $L$  característic de la no-linealitat aplicada a l'escena.

El resultat donat per l'Eq. (6.7) correspon a un cas més general que l'estudiat pels filtres de potències fraccionàries en l'apartat 3.3 d'aquesta memòria. Recordem que

els filtres de potències fraccionàries es poden implementar en un correlador de transformades conjuntes [12] i, en el cas de la DNC equival a prendre els valors dels paràmetres  $L = 1$  i  $M = p$  (veure Taula 4.1 del capítol 4), segons la nomenclatura seguida per Vijaya Kumar i Hassebrook [6]. Si així ho fem, el resultat deduït en el capítol 3 (Figura 3.2) esdevé un cas particular de l'Eq. (6.7).

Observem també, que coneguts  $a$  i  $b$ , és a dir les transmitàncies o nivells de gris de la referència i de l'escena, respectivament, podem trobar quin valor del paràmetre  $L$  ens dóna el grau de discriminació del sistema més adient.

Les dues línies que limiten la banda de reconeixement sempre es situen a banda i banda de l'eix  $L = 0$  (Figura 6.8). En els valors de la DNC obtinguts per l'eix  $L = 0$ , no influeix la informació energètica dels objectes d'igual forma que la referència, ja que d'acord amb l'Eq. (6.7) s'obté la raó constant  $cc/ac = 1$ . Quan la referència és més intensa que l'escena que s'analitza ( $a > b$ ), observem també que la línia corresponent al llindar inferior (és a dir,  $cc/ac = 0.5$ ) es situa en els valors positius d' $L$ . En canvi, la línia llindar superior ( $cc/ac = 1.5$ ) es troba en els valors negatius d' $L$ . A més, la banda de reconeixement no és simètrica respecte el valor d' $L = 0$ . Totes aquestes conclusions que s'extreuen de l'Eq. (6.7), s'adiuen amb les característiques estretes de l'observació de la Figura 6.8.

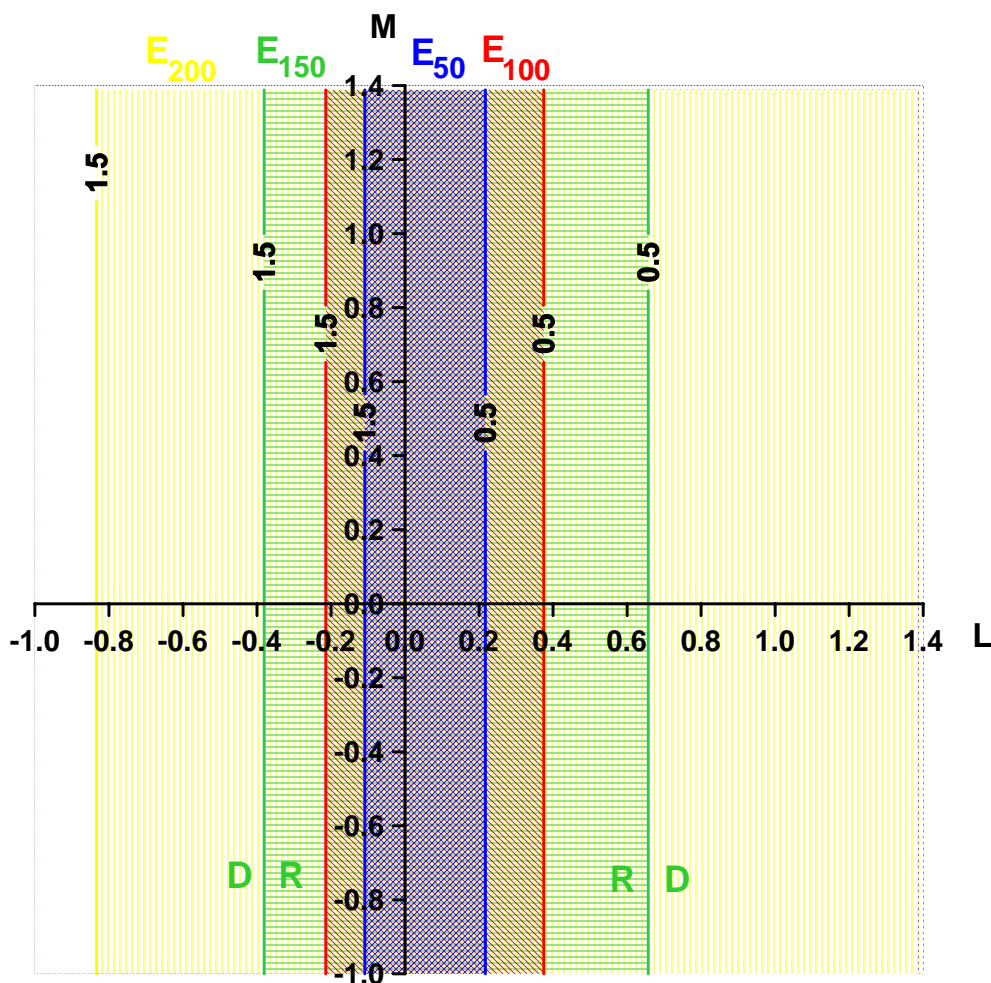
En el cas de la lletra  $E_{200}$ , en l'amplada de la zona de l'espai  $(L, M)$  analitzada, només s'inclou el llindar superior, mentre que el llindar  $cc/ac = 0.5$  queda per fora de l'interval  $[-1, 1.4]$  estudiat. Resolent l'Eq. (6.7) podem comprovar que aquesta línia llindar inferior tallaria l'eix d'abscisses en el punt  $L = 1.43$ . Observem també en el gràfic de la Figura 6.8, que quan més similar és el caràcter analitzat respecte a l'objecte buscat no només en forma sinó també en intensitat, l'amplada de la regió de reconeixement és més gran, i viceversa, bandes de reconeixement estretes i properes al valor  $L = 0$  són representatives de caràcters d'igual forma que la referència però amb nivell energètic més diferent. Analitzant els resultats obtinguts en l'interval de valors de  $L$  i  $M$  estudiat, veiem primerament que hi ha dos lletres que és impossible discriminar de la referència dins l'interval estudiat. Aquestes són l' $E_{240}$  i l' $E_{225}$ , que

consisteixen en els caràcters amb diferències energètiques inferiors al 12% respecte a la referència. Per discriminar aquestes lletres, i d'acord amb l'Eq. (6.7), cal prendre valors d'  $L < -1.6$  o  $L > 2.77$ , pel cas de la lletra E<sub>225</sub>, i  $L < -3.34$  o  $L > 5.2$ , per la lletra E<sub>240</sub>.

Si examinem amb detall la Figura 6.8, però, observem que les línies lliard es desvien lleugerament del comportament predit per l'Eq. (6.7) en la regió de l'espai on  $M$  pren valors negatius més baixos ( $-1 < M < -0.4$ ) i  $L$  correspon a qualsevol valor negatiu. S'han realitzat proves per comprovar l'origen d'aquestes desviacions. En el cas de prendre valors molt negatius d'ambdós paràmetres  $L$  i  $M$ , amb la consegüent potenciació de les freqüències més altes, s'ha comprovat com en el pla de sortida del correlador apareixen unes franges d'interferència com a efecte de les no-linealitats aplicades que provoquen el solapament del terme central del pla de sortida, sense interès en les aplicacions de reconeixement, i el terme dels ordres  $\pm 1$  on es localitza la correlació. Per assegurar que la separació respecte el comportament ideal de les corbes de la Figura 6.8 és degut a aquest solapament, hem obtingut de nou les mateixes corbes pel cas on l'espectre conjunt modificat es calculava segons l'expressió (4.25), de la funció denominada  $P(\vec{\nu})$ . Recordem que mitjançant aquest procediment, aconseguíem un ordre central del tipus funció delta de Dirac i, per tant, s'evitaria el solapament amb altres ordres. Les corbes així obtingudes es representen en la Figura 6.9. Ara, el comportament deduït a partir de l'anàlisi de l'expressió (6.7), amb les línies límit de les zones de reconeixement perfectament paral·leles a l'eix  $L = 0$ , es correspon exactament amb l'observat en aquesta representació.

Si introduïm les restriccions experimentals establertes en el capítol anterior, podem obtenir un gràfic equivalent a la Figura 6.8, i més proper a les condicions experimentals amb les que implementarem la DNC en un dispositiu optoelectrònic. Considerem novament un 10% de saturació en els espectres enregistrats, 256 nivells de gris del sistema de captació i 32 nivells de gris per reproduir l'espectre conjunt modificat en el modulador. Amb aquestes condicions, els resultats de reconeixement es modifiquen sensiblement, i així queda reflectit en la Figura 6.10. Per una banda, es perd la invariància dels resultats de correlació respecte al paràmetre  $M$ . A més, les

bandes de reconeixement que s'estableixen amb el llindar superior i inferior, ja no es troben al voltant de l'eix  $L = 0$ , sinó que depenen conjuntament d' $L$  i  $M$ , i no sempre estan incloses una dins l'anterior a mesura que el nivell de gris s'allunya respecte al de la referència (o el coeficient de similitud disminueix).



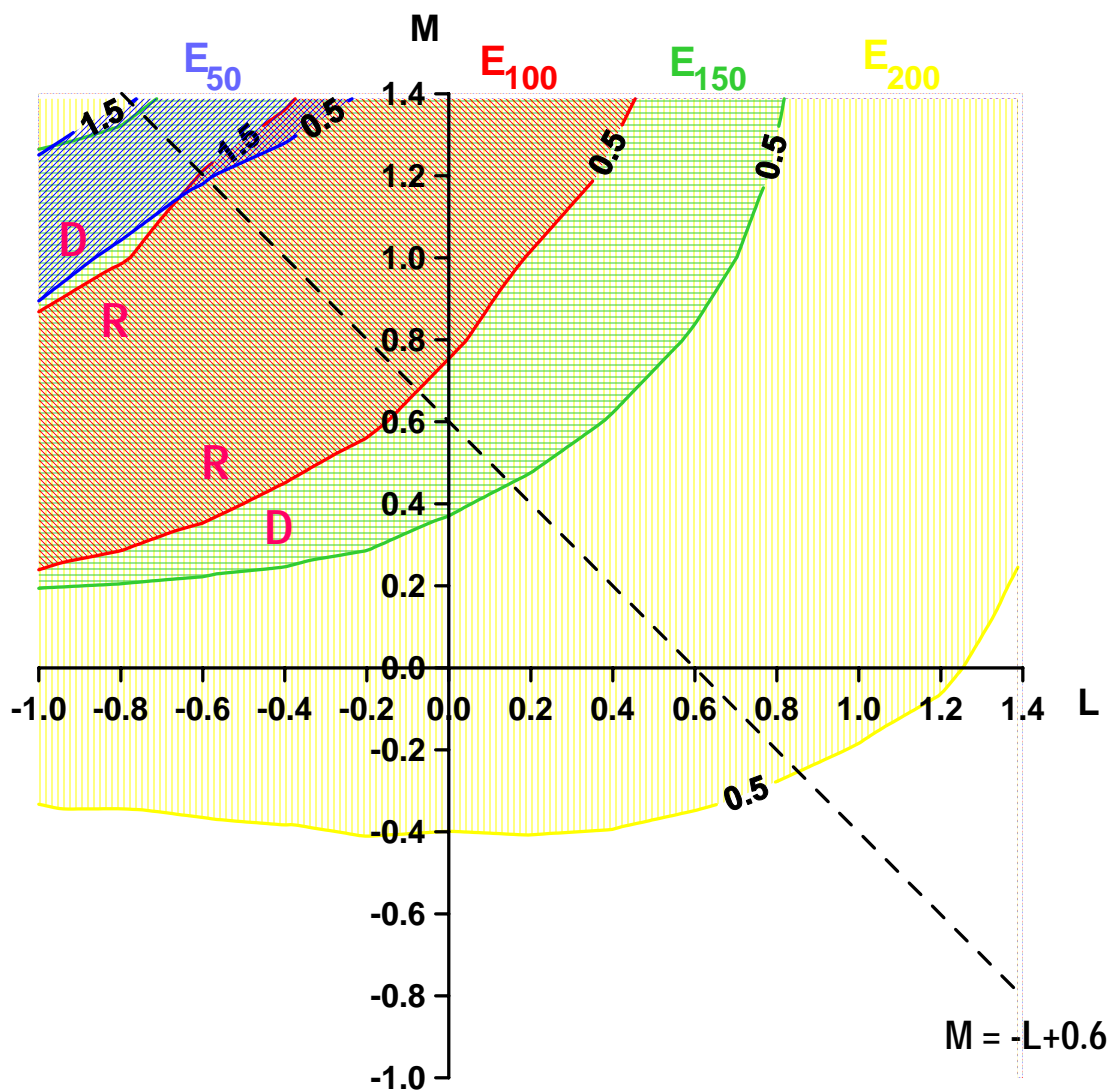
**Figura 6.9.** Resultats de reconeixement obtinguts mitjançant DNC a partir del càlcul de la funció  $P(\vec{v})$  (Eq. 4.25) en condicions experimentals ideals. Representació de les línies llindar de reconeixement corresponents a  $\left|1 - \frac{cc}{ac}\right| = u$ , amb  $u = 0.5$ , per les lletres que conformen el test 2 de la

Figura 6.6. En la banda de reconeixement corresponent a la lletra  $E_{150}$  s'indica la zona de reconeixement (R) i la zona de discriminació (D) d'aquesta lletra respecte a la referència ( $E_{ref}$ ). L'orientació d'aquestes zones és la mateixa per la resta de corbes.

Les diferències entre els gràfics de la Figura 6.8 i de la Figura 6.10 són principalment degudes a l'efecte de la quantització del sistema d'adquisició. Per comprovar-ho, hem realitzat proves aïllant l'efecte de la quantització de la càmera i de la saturació del sensor de la CCD quan enregistrem les distribucions d'intensitat. Només quan s'introdueix la quantització, i encara que es considerin una gran quantitat de nivells de gris disponibles pel sistema d'adquisició, els resultats es veuen sensiblement afectats, acostant-se al comportament que es posa de manifest en la Figura 6.10. Hem comprovat que l'efecte de la quantització del sistema sobre les distribucions captades, es tradueix en un retall de les freqüències més altes, de manera que s'introdueix una limitació de la regió de suport de les funcions. Si comparem dues lletres amb energies molt diferents (per exemple, l' $E_{50}$  quan es vol confondre o discriminar de l' $E_{255}$ ), el retall de freqüències en la distribució menys intensa és molt més important, donat que mantenim la informació de l'energia relativa de les dues funcions. Aquest és el motiu pel qual varien tan sensiblement les zones de reconeixement i discriminació, només per haver introduït condicionaments experimentals.

De la Figura 6.8 podríem extreure que la zona de l'espai  $LM$  on la DNC disposa d'una major tolerància als canvis en el nivell de gris dels objectes és teòricament per valor d' $L = 0$ . Concretament, el sistema seria invariant a canvis en el nivell energètic dels objectes. Però hem vist que la introducció dels condicionaments experimentals no permeten assegurar-ho a la pràctica.

Tenint en compte les diferències més importants que s'introdueixen amb les condicions experimentals, podem determinar en la gràfica de la Figura 6.10 una recta sobre la qual la capacitat discriminativa del sistema per objectes d'igual forma i nivell de gris diferent, va variant progressivament. De totes maneres, el fet que ara les bandes de reconeixement no estiguin incloses unes dins les altres, ens impedeix a la pràctica poder arribar a uns valors concrets d' $L$  i  $M$  que proporcionin una tolerància total amb la possibilitat de reconèixer totes les lletres i identificar-les amb la referència.

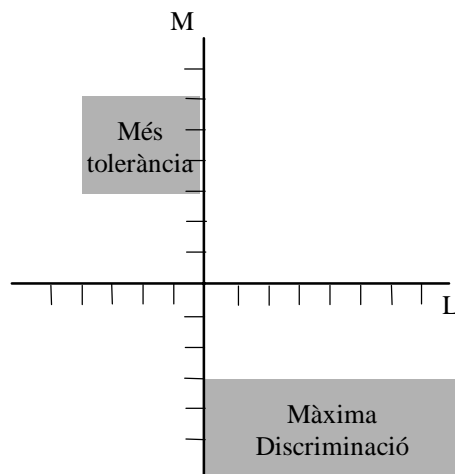


**Figura 6.10.** Resultats de reconeixement obtinguts mitjançant DNC a partir del càlcul de la funció  $P'(\bar{v})$  (Eq. 4.27) al considerar les condicions experimentals donades per una saturació d'  $\alpha=10\%$ , una quantització del sistema de captació de 256 nivells de gris, i 32 nivells de gris de l'SLM. Representació de les línies llindar de reconeixement corresponents a  $\left|1 - \frac{cc}{ac}\right| = u$ , amb  $u = 0.5$ , per les lletres que conformen el test 2 de la Figura 6.6. En la banda de reconeixement corresponent a la lletra  $E_{100}$  s'indica la zona de reconeixement (R) i la zona de discriminació (D) d'aquesta lletra respecte a la referència ( $E_{ref}$ ). L'orientació d'aquestes zones és la mateixa per la resta de corbes.

Observem en la Figura 6.10, que podem recórrer l'espai generat per  $(L, M)$  a través, per exemple, de la recta d'equació  $M = -L + 0.6$  i variar la capacitat de discriminació del sistema. En la zona on  $L$  és gran ( $L \geq 1$ ) i  $M$  té valors més negatius ( $M \leq -0.6$ ), la

discriminació de la referència és prou bona i només es reconeixien les lletres  $E_{255}$  (referència),  $E_{240}$  i  $E_{225}$ . En canvi, en la zona oposada de la mateixa recta, quan tenim  $L$  negatives, i  $M$  grans i positives ( $0.8 \leq M \leq 1$ ), la tolerància és molt gran i la majoria dels caràcters més energètics (exceptuant l' $E_{50}$ ) poden ser identificats a la referència. En zones intermèdies el grau de discriminació del sistema va variant de forma contínua.

Dins l'interval d' $L$  i  $M$  estudiat, les regions de treball on trobem, respectivament, màxima discriminació i major tolerància a petits canvis en el nivell de gris dels caràcters, estan representades en la Figura 6.11.



**Figura 6.11.** Delimitació de les zones de màxima discriminació, i màxima tolerància a canvis de nivell de gris en els caràcters del test 2 (Figura 6.6).

## 6.4 Conclusió

Com a conclusió d'aquests primers apartats on hem estudiat l'aplicació de la DNC a objectes equienergètics de forma diferent i, posteriorment, a objectes d'igual forma i amb nivell de gris uniforme i diferent, observem algunes característiques comunes. En ambdós casos, aconseguim un sistema de reconeixement amb una capacitat de discriminació variable de manera contínua amb els valors d' $L$  i  $M$ . Observem també,



que la zona de millor discriminació i la de màxima tolerància s'associen amb el valor de la no-linealitat aplicada a la referència ( $M$ ). Per  $M$  negatius s'assoleix la màxima discriminació en les dues situacions, mentre que són valors alts d' $M$  els que permeten tenir un sistema amb més tolerància. El paràmetre  $L$  per la seva banda, sembla estar associat amb el tipus d'informació que és més important en cada cas, bé sigui la forma o el nivell de gris. En el cas de variacions de forma, és necessari considerar  $L$  negatius per poder discriminar amb més precisió la referència de qualsevol altra forma present com a objecte a analitzar. En canvi, si volem obtenir certa tolerància a les variacions de forma dels objectes, cal recórrer a valors alts i positius d' $L$  (Figura 6.4). Pel reconeixement d'objectes on l'única informació distintiva és el nivell de gris uniforme que els recobreix, el comportament és just al contrari. Valors d' $L$  alts i positius, conjuntament amb  $M$  negatius, com ja hem fet esment, donen una màxima discriminació del sistema a variacions energètiques d'un objecte. Al contrari, valors negatius d' $L$  i positius d' $M$ , permeten tenir més tolerància a aquestes variacions (Figura 6.11). En forma de resum, englobem aquestes conclusions en la Taula 6.4.

**Taula 6.4.** Resum esquemàtic de les conclusions obtingudes amb l'aplicació de la DNC al reconeixement de caràcters amb variacions de forma i de nivell de gris.

	<b>M</b> alts i positius	<b>M</b> baixos en valor absolut i negatius
<b>L</b> alts i positius	Tolerància a canvis en la forma	Màxima discriminació pel nivell de gris
<b>L</b> baixos en valor absolut i negatius	Tolerància a canvis en el nivell de gris	Màxima discriminació per la forma

# Capítol 7

## Realització experimental de la DNC

### 7.1 Introducció

En aquest capítol mostrarem la realització experimental del processador DNC amb capacitat de discriminació variable davant de canvis en la forma i en la intensitat dels objectes. Per això, descriurem en el següent apartat (secció 7.2), el dispositiu optoelectrònic emprat, amb els elements que el componen i els passos consecutius que cal realitzar per a obtenir la DNC en el pla de sortida del correlador. Donat que utilitzarem una càmera CCD en blanc i negre per enregistrar l'espectre de potència conjunt (JPS) i el pla de correlació, és necessari portar a terme una calibració del sistema de captació, incloent la càmera CCD i la targeta digitalitzadora (apartat 7.3). En el següent punt d'aquest capítol, en l'apartat 7.4, descriurem el funcionament de les pantalles de cristall líquid com a moduladors espacials de llum, i també detallarem la calibració d'aquestes pel seu funcionament més idoni en el nostre dispositiu experimental.

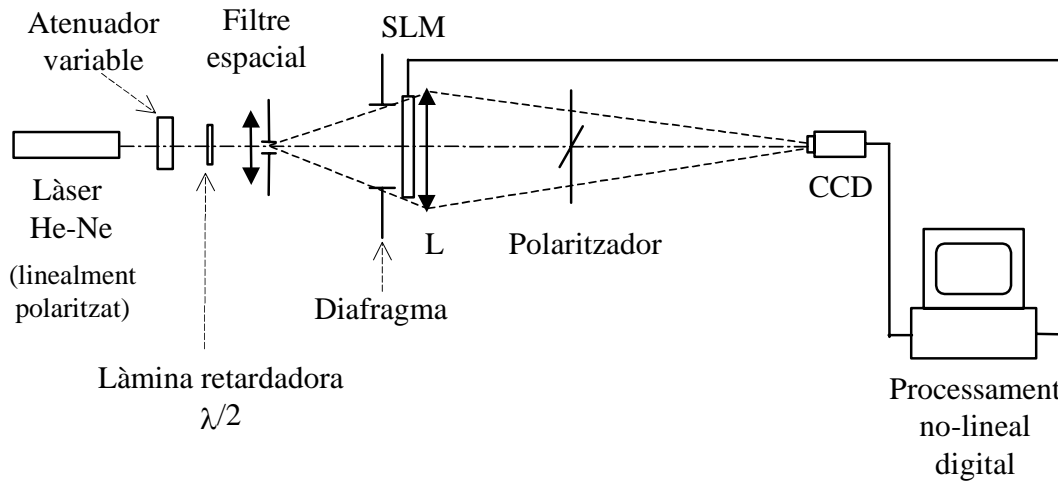
En l'apartat 7.5 descriurem els passos indispensables per posar en funcionament el conjunt de tot el dispositiu òptico-digital i, finalment, l'apartat 7.6 mostrarà els resultats experimentals de la DNC, separadament, per quan considerem variacions de forma en objectes equienergètics, i pel cas on tots els objectes tenen exactament la mateixa forma i les úniques variacions són d'intensitat o nivell de gris. Amb aquests resultats validarem la proposta realitzada en el capítol anterior. Així mateix, tractarem d'explicar les discrepàncies que puguin haver entre els resultats de simulació i els obtinguts experimentalment.

## 7.2 Descripció del dispositiu experimental

La realització experimental de la DNC, s'ha portat a terme tenint en compte el dispositiu optoelectrònic simplificat que es mostra en la Figura 7.1. L'esquema del muntatge coincideix essencialment amb el representat de manera més general en la Figura 4.13, però en la Figura 7.1 s'han detallat tots els elements emprats i s'ha situat el modulador espacial de llum davant de la lent, sense que això afecti essencialment als resultats obtinguts. El dispositiu DNC emprat es basa en l'arquitectura del correlador de transformades conjuntes (JTC), i permet obtenir els resultats finals de la correlació no-lineal dual en dos passos consecutius. En el primer pas, el modulador espacial de llum reproduïx el pla d'entrada del correlador mostrant, de manera successiva, l'escena monobjecte a analitzar, l'objecte de referència i ambdós conjuntament, i s'enregistren amb la càmera CCD les tres distribucions d'intensitat  $I_S$ ,  $I_R$  i  $I$  corresponents a les respectives transformades de Fourier. En el segon pas, el modulador reproduïx l'espectre conjunt modificat no-linealment (MJPS), del qual se n'obté la seva transformada de Fourier en el pla de sortida del processador. En aquesta informació, que també és enregistrada per la càmera CCD, estan inclosos un terme central sense interès pel nostre treball i, a més, la DNC. L'anàlisi de la distribució de la DNC a través de l'aplicació de llindars de reconeixement i, si s'escau, operacions lògiques establertes pel criteri de decisió, permet dir si l'objecte inclòs a l'escena problema correspon o no al patró buscat.

L'ús del dispositiu simplificat per a portar a terme la DNC experimentalment, implica un control de la intensitat de llum que il·lumina els successius plans d'entrada del correlador per a mantenir-la constant, però presenta l'avantatge d'abaratir el cost del sistema òptic i facilitar l'alineació dels seus components. A més, recordem que en la proposta realitzada en el capítol 4, també s'inclou una simplificació del procés de càlcul no-lineal sobre l'espectre de transformades conjuntes (JPS), reduint-ne d'aquesta manera, el seu temps computacional. Aquesta variació en el càlcul, com ja s'ha vist, no varia el resultat de la correlació sempre i quan l'ordre central del pla de sortida sigui prou estret per no solapar i, per tant, afectar als pics de correlació. Per altra banda, cal recordar també, que el processament no-lineal simplificat evita la subtracció de les funcions  $I_S$  i  $I_R$  de la funció  $I$ , tasca inclosa en la formulació inicial del MJPS, i difícil de realitzar experimentalment de manera acurada [207,208].

En la Figura 7.1, observem la font de llum monocromàtica emprada, un làser d'He-Ne polaritzat linealment de 7 mW, expandit i filtrat espacialment mitjançant una unitat que integra un objectiu de microscopi (60x) i un filtre espacial ( $5 \mu\text{m}$ ). La intensitat del feix que il·lumina el sistema es controla amb un atenuador neutre de densitat variable. Una obertura rectangular variable situada just davant del modulador, limita la regió de la pantalla del cristall líquid (LCTV) que és il·luminada. En el modulador es reproduïxen, controlades per ordinador, les imatges a analitzar en el primer pas de la DNC, i el MJPS en el segon i darrer pas. Com més endavant detallarem, és important fixar la orientació del pla de polarització de la llum incident sobre la pantalla LCTV per obtenir el règim adequat de treball del modulador. Això s'aconsegueix amb la làmina retardadora  $\lambda/2$  situada en un suport giratori graduat abans de l'expansor del feix. En el nostre cas, hem treballat amb el que s'anomena règim principalment d'amplitud, on s'aconsegueix una modulació majoritàriament en l'amplitud del feix mitjançant l'acció conjunta de la pantalla de cristall líquid i d'un polaritzador situat just darrera d'aquesta.



**Figura 7.1.** Dispositiu optoelectrònic basat en el JTC utilitzat en la realització experimental de la DNC.

La transformada de Fourier (TF) de la imatge reproduïda en la pantalla de cristall líquid, es realitza òpticament mitjançant la lent  $L$  de distància focal  $f' = 300$  mm. La TF s'obté en el pla conjugat de la font puntual lluminosa proporcionada pel *pinhole* del filtre espacial i, és en aquest pla on es situa la càmera CCD que enregistra les distribucions d'intensitat amb una targeta digitalitzadora controlada per ordinador. La configuració del dispositiu presentat, amb il·luminació divergent, permet ajustar l'escala de les TF en intensitat captades, variant la distància del modulador a la lent. Quan la il·luminació és divergent, i l'objecte (SLM) es situa per davant de la lent, la TF de l'objecte presenta una escala que depèn de la distància  $\vec{d}$  entre l'objecte i la lent [132],

$$G(\vec{x}) = C e^{i \frac{k}{2d} |\vec{x}|^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\vec{\xi}) e^{i 2\pi \left( \vec{\xi} \cdot \frac{\vec{x}}{\lambda d} \right)} d\vec{\xi}, \quad (7.1)$$

on  $g(\vec{\xi})$  descriu la distribució d'amplitud de l'objecte reproduït en el modulador, i  $C$  agrupa tots els termes independents d'  $\vec{x}$  i  $\vec{\xi}$ . El fet d'apropar al màxim el modulador

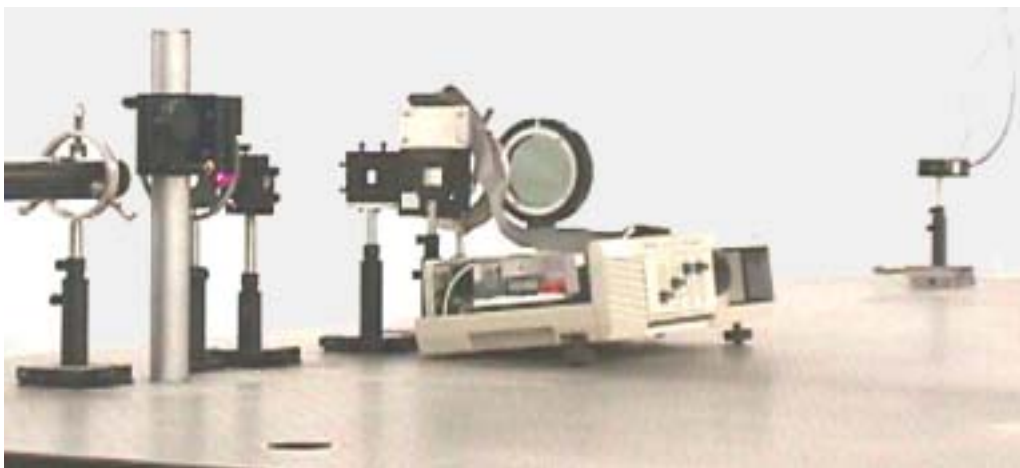
a la lent, permet, per una banda, obtenir la TF amb el tamany més gran possible fixats els plans conjugats *pinhole*-càmera CCD i, per altra banda, es minimitza el retall en freqüències produït per la pupil·la de la lent. El control de l'escala de la TF, en aquest muntatge, només es fa per aprofitar la resolució de mostreig espacial del sensor de la càmera.

Finalment, i només en el segon pas de la DNC, utilitzem un segon polaritzador just davant el sensor CCD, que no s'ha representat en la Figura 7.1, amb la funció d'un atenuador variable addicional que ens permet enregistrar els pics de correlació obtinguts amb la DNC dins el rang lineal del sistema de captació. Donat que per diferents mètodes de correlació no-lineal dual, les característiques del pic de correlació, en particular la seva intensitat màxima, varien, el polaritzador situat en aquest punt, permet modificar en cada enregistrament la quantitat de llum que arriba a la càmera, de manera que s'evita la saturació del sensor en la posició del pic de correlació. Només en el cas en què és necessària la comparació dels resultats entre diferents escenes monobjecte a analitzar, la posició del polaritzador es manté fixada per poder obtenir els valors relatius dels diferents màxims de correlació creuada.

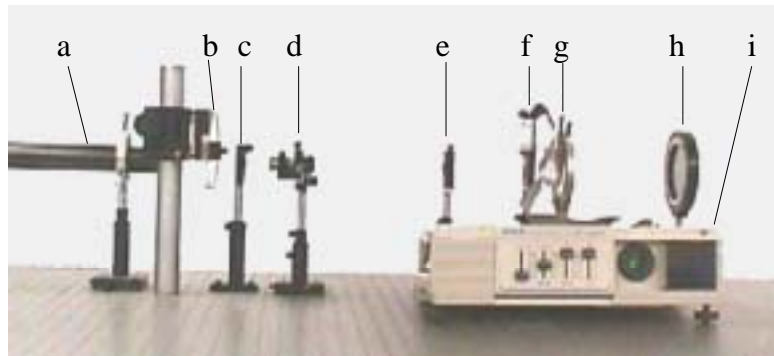
Una altra possibilitat d'enregistrament del terme corresponent a la correlació dins el rang lineal de la càmera, i evitant la saturació del sensor en la posició de l'ordre central de la sortida del correlador, és mitjançant l'anomenat bloqueig suau del terme dc [209]. Aquest mètode es caracteritza per l'ús d'un material fotoluminiscent i feixos làser de diferent  $\lambda$ . En la referència [209], Jutamulia i Gregory proposen una tècnica de subtracció de senyals enlloc de l'ús d'un diafragma que bloquegi el terme dc en l'ordre zero, donat que és molt difícil conèixer amb exactitud l'amplada d'aquest terme. Basat en aquesta mateixa idea, però emprant materials més assequibles i un sol feix làser, Ledesma *et al.* proposen l'ús d'una placa hologràfica impressionada de manera que contingui el terme central que constitueix la regió més intensa. L'ús d'aquesta placa just davant del sensor CCD en les posteriors captacions de resultats experimentals, atenua l'ordre central, evitant així la saturació del sensor, i enregistra en millors condicions el terme de correlació [27].

En el nostre cas, no hem bloquejat el terme dc central en l'enregistrament del pla de sortida del correlador, tot i que l'ús d'una de les tècniques que acabem d'exposar, especialment l'exposada en [27] per la seva accessibilitat, seria una possibilitat de millora del dispositiu experimental emprat.

En les Figures 7.2 i 7.3 es mostra, respectivament, una vista general del correlador no-lineal dual, i un detall del mateix on no s'ha inclòs el sistema de captació. Veiem que en el muntatge utilitzat, hi ha presents elements purament òptics com el filtre espacial, la làmina  $\lambda/2$ , la lent, etc., i d'altres optoelectrònics, com el modulador espacial de llum i la càmera CCD, que contenen una interfase de transformació del senyal òptic en elèctric i viceversa i, requereixen en general d'uns circuits electrònics i d'un *software* pel seu funcionament i control. L'electrònica i el *software* disposen sovint d'uns paràmetres de control que possibiliten modificar el règim de treball dels aparells. Per aquest motiu, és indispensable una calibració prèvia de l'SLM i de la càmera CCD. A partir d'ella es pot conèixer el seu funcionament i determinar les millors condicions de treball segons l'aplicació en la qual es volen utilitzar.



**Figura 7.2.** Vista general del correlador no-lineal dual experimental.



**Figura 7.3.** Detall del correlador no-lineal dual experimental. En ell s'observa: *a*, làser He-Ne; *b*, atenuador variable; *c*, làmina  $\lambda/2$ ; *d*, filtre espacial; *e*, obertura rectangular variable; *f*, pantalla de cristall líquid connectada al videoprojector; *g*, lent; *h*, polaritzador; *i*, videoprojector amb els controls de l'SLM.

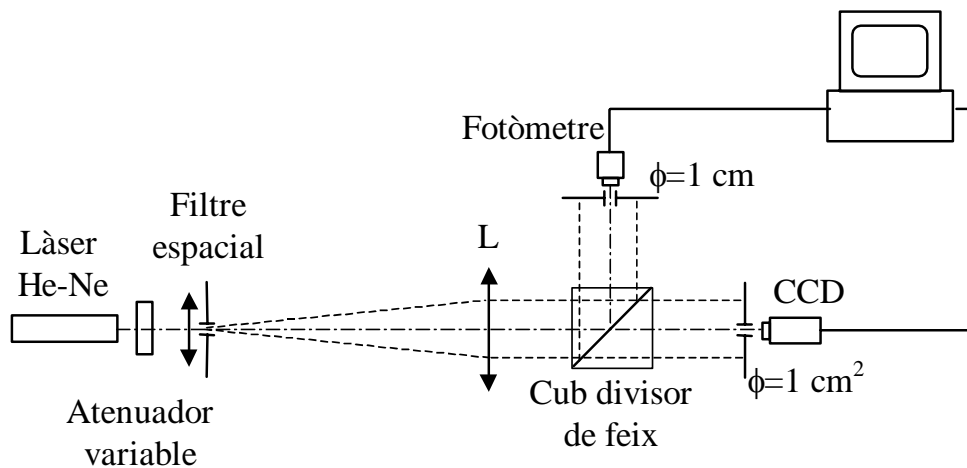
### 7.3 Calibració del sistema de captació

El sistema de captació del muntatge està constituït pel detector que és una càmera CCD b/n marca PULNIX model TM-765 amb resposta lineal ( $\gamma=1$ ), i una targeta digitalitzadora MATROX model COMET. La càmera CCD s'utilitza en el muntatge del correlador no-lineal dual sense l'objectiu, de manera que el pla de correlació es fa coincidir amb la matriu de sensors CCD de la càmera. Aquesta matriu té una àrea activa d'aproximadament  $8.8 \times 6.6 \text{ mm}^2$ , en la qual es distribueixen  $756 \times 581$  píxels. El tamany de cada sensor elemental és de  $11.0 \times 11.0 \text{ }\mu\text{m}^2$ . Prèviament a la utilització en el muntatge, s'ha comprovat que el sistema de captació no mostreja d'igual manera en les direccions vertical (V) i horitzontal (H), sinó que les imatges captades presenten una distorsió  $V/H = 1.13$ . Aquesta distorsió no influirà en el procés de reconeixement donat que afecta per igual a l'escena i a la referència, i a totes les distribucions d'intensitat captades en el primer pas de la DNC. Les dades rebudes pels sensors CCD es transmeten a la targeta digitalitzadora de manera entrellaçada, la qual cosa suposa un augment en el temps de processament, però al mateix temps pot significar una pèrdua d'informació, ja que no es capten simultàniament la resposta de tots els píxels. El fet de treballar amb imatges

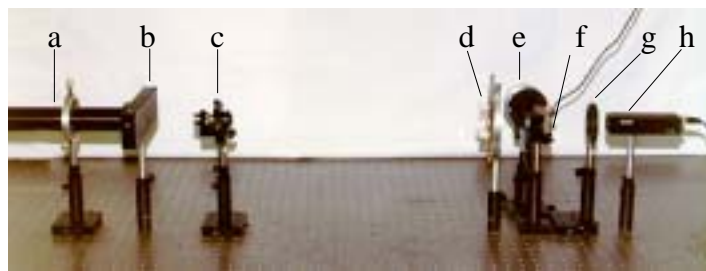


estàtiques minimitza aquesta possible pèrdua d'informació. El *software* de captació permet regular dos paràmetres que són els denominats *brillo* i *contrast*, que actuen com un voltatge afegit, i com un guany multiplicat al senyal, respectivament. Les corbes de resposta de la càmera obtingudes per a diferents valors del *brillo* i del *contrast*, ens permetran escollir la regió de resposta més lineal i amb més rang dinàmic del sistema de captació.

El dispositiu en el qual s'han portat a terme les mesures de calibració, es detalla de manera esquemàtica en la Figura 7.4, i es mostra una fotografia del mateix en la Figura 7.5. Un feix làser d'He-Ne de 17 mW proporciona la il·luminació del sistema. S'expandeix utilitzant un objectiu de microscopi invertit (60x), i es filtra espacialment amb un *pinhole* (5  $\mu\text{m}$ ) situat sobre el disc d'Airy que s'obté darrere de l'objectiu. Una lent de  $f = 500$  mm proporciona un feix col·limat de diàmetre suficientment gran per obtenir una superfície d'il·luminació sobre els sensors aproximadament uniforme. Un cub separador o *beamsplitter* divideix el feix en dos braços. La llum d'un braç incideix sobre el sensor CCD i, la llum de l'altre braç, sobre un fotòmetre. Davant ambdós elements, situem un diafragma d'obertura circular de diàmetre  $\phi = 1$  cm, que ens selecciona una àrea d'il·luminació aproximadament igual sobre els sensors. El fotòmetre (Newport, model 818-SL), amb la seva targeta convertidora analògica-digital, permet mesurar les variacions en la intensitat relativa que rep en cada moment la càmera. L'atenuador variable situat darrere el làser possibilita modificar la intensitat que passa cap als sensors adaptant-la al seu rang dinàmic. Mitjançant el *software* de captació es realitza la lectura del valor mig de la distribució enregistrada per la càmera. Després de repetir el mateix procediment per a diferents valors de *brillo* i *contrast*, determinem la resposta més lineal de la càmera CCD, tot assegurant l'aprofitament del seu rang dinàmic. La corba seleccionada correspon als paràmetres *brillo* = 60 i *contrast* = 0 (Figura 7.6).

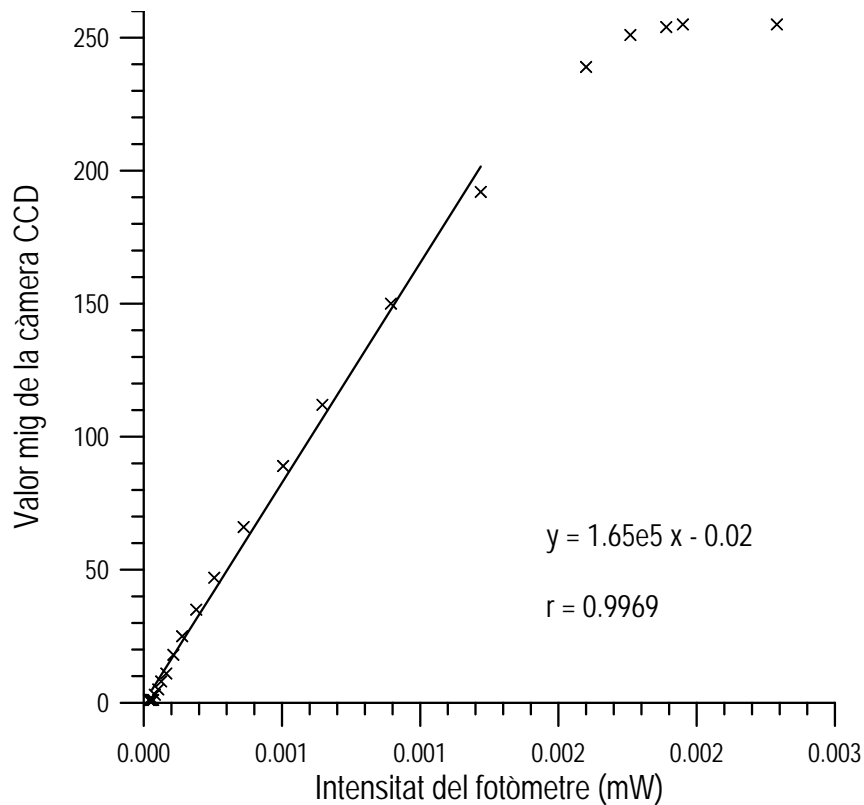


**Figura 7.4.** Esquema del muntatge per a la calibració del sistema de captació.



**Figura 7.5.** Dispositiu emprat en la calibració de la càmera CCD. S'observa: a, làser; b, atenuador variable; c, filtre espacial; d, lent; e, fotòmetre i diafragma; f, cub divisor de feix; g, diafragma, i h, càmera CCD.

En aquestes condicions observem que el negre dona un nivell nul, de manera que assegurem un bon contrast i, a més, la linealitat de la resposta de la càmera s'estén del zero fins aproximadament un valor de 200. Aquesta característica es tindrà especialment en compte al captar els plans de correlació de la DNC. És important captar els pics d'autocorrelació i de correlació creuada dins una regió de resposta lineal, donat que els seus màxims són els que compararem numèricament per determinar si un objecte és detectat o rebutjat. Per tant, en el segon pas de la realització de la DNC procurarem no sobrepassar el valor de 200 en els màxims de correlació.



**Figura 7.6.** Resposta aproximadament lineal del sistema de captació pels valor de *brillo* = 60 i *contrast* = 0.

Cal esmentar també, el nivell de soroll detectat en l'enregistrament d'imatges amb el sistema d'adquisició. En els components optoelectrònics és habitual la presència de petites oscil·lacions en els valors numèrics enregistrats degut a petites inestabilitats en els senyals elèctrics, corrents paràsites o a altres causes. La càmera CCD emprada en les proves experimentals, conjuntament amb la targeta digitalitzadora i el *software* d'adquisició presenta un tipus de soroll de freqüència periòdica, totalment inevitable en qualsevol captació. Per aïllar l'efecte del soroll i mostrar-lo aquí, hem enregistrat una imatge corresponent al nivell nul d'energia, obtingut al captar amb tot l'entorn a les fosques i tapant el sensor de la càmera. La imatge digital enregistrada mostra l'aspecte representat en la Figura 7.7, que correspon a una zona 12 x 12 píxels de la imatge total de 512 x 512 píxels. Les franges horitzontals es presenten de forma

idèntica sobre tota la imatge. Hem pogut observar que aquest patró de franges és un soroll que s'afegeix a qualsevol imatge captada mitjançant el sistema de captació.

1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0

**Figura 7.7.** Nivell de soroll periòdic captat en tots els enregistraments realitzats amb la càmera Pulnix TM-765, controlada amb la targeta Matrox Comet.

## 7.4 Caracterització del modulador espacial de llum

El modulador espacial de llum emprat s'ha extret del videoprojector comercial EPSON model VP-100PS. Aquest aparell conté tres pantalles de cristall líquid (LCTV) a les quals s'envia separatament les components cromàtiques RGB de la imatge en color que es vol reproduir. L'electrònica de l'aparell regula les característiques de cada component de la imatge, així com la relació entre elles. El senyal elèctric enviat a les pantalles LCTV pot ser modificat externament a través de tres potenciòmetres situats en la part frontal del videoprojector. Aquests potenciòmetres reben el nom de *brillo*, que actua com un voltatge constant afegit al senyal enviat, el *contrast*, que actua com un guany en el voltatge aplicat a cada pantalla, i per últim el *color*, que modifica la relació de voltatges aplicats sobre cada una de les tres pantalles de l'aparell.

Les tres pantalles LCTV incloses en aquest model de videoprojector es poden extreure de l'aparell, i les connexions poden ser adaptades per al seu ús en un sistema òptic. Els dispositius LCTV en general han estat àmpliament emprats en

correladors òptics coherents ja que permeten reproduir i modificar en temps real les escenes analitzades o la informació corresponent a l'objecte de referència buscat. Existeixen diferents treballs que en detallen les seves característiques tècniques [177], i s'han proposat diferents mètodes per a la seva caracterització tant en règim de modulació d'amplitud com de fase, a fi i efecte d'utilitzar-los adequadament en els correladors òptics [175-184].

#### **7.4.1 Característiques generals de les LCTV del videoprojector Epson model VP-100PS. Paràmetres i comportament espacial**

Els panells del videoprojector Epson model VP-100PS són dispositius de cristall líquid del tipus nemàtic, amb una estructura molecular de forma helicoidal (en anglès *twisted-nematic*). En aquesta configuració, l'eix director del cristall es situa paral·lel a la superfície del substrat pla i transparent on hi ha dipositats els elèctrodes, i descriu un angle de  $90^\circ$  des de l'entrada fins a la sortida del cristall. Donada l'elasticitat del cristall líquid, l'aplicació d'un camp elèctric comporta l'alineació de les molècules en la direcció del camp aplicat, amb més força quan més intens és el camp elèctric. Un cristall líquid nemàtic pot ser considerat com un rotatori variable de la polarització de la llum incident. És a dir, si en el cristall incideix una ona polaritzada lineal, inicialment en la direcció del director del LCTV, el pla de polarització d'aquesta ona girarà un cert angle en funció del camp elèctric aplicat sobre les molècules del cristall. En el cas en què no hi hagi cap camp aplicat en la direcció de propagació de la llum, l'ona incident descrita, girarà el seu pla de polarització  $90^\circ$ , el mateix angle que descriuen les molècules dins el cristall. En canvi, si s'aplica un camp molt fort en la direcció de propagació de la llum, les molècules del cristall s'alineen en la direcció del camp, i el pla de polarització de la llum es manté constant al travessar la pantalla. Un camp d'intensitat intermèdia, implicarà la rotació del pla de polarització de l'ona incident en un angle determinat per l'orientació de les molècules i el gruix del cristall. Un canvi en l'orientació de les molècules del cristall comporta, doncs, un canvi en les propietats òptiques de les cel·les [174].

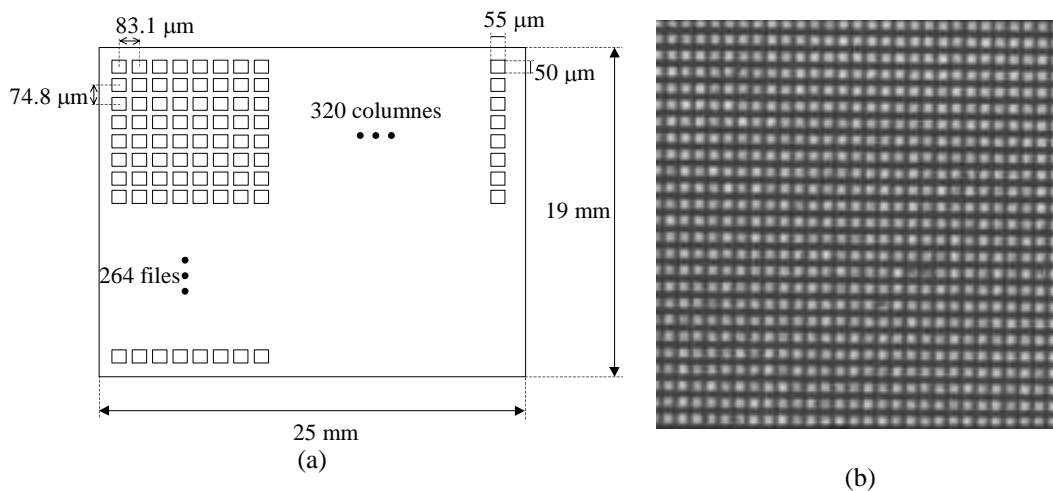
Si a la sortida de la LCTV situem un polaritzador, obtindrem una modulació en amplitud del feix, ja que la llum transmesa pel polaritzador serà dependent de la orientació del pla de polarització que tingui a la sortida de la LCTV. És en aquest règim de treball, com actuen les pantalles dins el videoprojector.

En general, donada una orientació de la llum polaritzada lineal incident, i uns valors dels potenciòmetres *brillo*, *contrast* i *color*, s'obté un règim de treball on coexisteixen els modes de fase i d'amplitud acoblats. No obstant, modificant convenientment les condicions de polarització de la llum d'entrada i els potenciòmetres dels videoprojectors, és possible obtenir un mode d'operació quasi de fase (amb la polarització de la llum incident paral·lela al vector director de l'SLM), o bé quasi d'amplitud (amb la polarització perpendicular al vector director).

Les pantalles del videoprojector Epson VP-100PS tenen un tamany aproximat de 25 x 19 mm. Aquest, està dividit en cel·les, també anomenats píxels, de manera que conté 320 columnes i 264 files. Cada píxel, de forma rectangular, mesura 55 x 50  $\mu\text{m}$ , i els centres de dos píxels consecutius disten entre ells 83.1  $\mu\text{m}$  en horitzontal i 74.8  $\mu\text{m}$  en vertical. Això suposa tant sols una àrea del 44% que és activa dins la pantalla LCTV [179]. En la Figura 7.8a es representa aquesta distribució de píxels dins la pantalla, indicant les distàncies i tamanys esmentats. També s'inclou una imatge real augmentada de la pantalla de cristall líquid captada amb la càmera CCD, quan en ella s'ha representat una imatge uniforme de nivell de gris 255 (Figura 7.8b). S'observa clarament que no tota l'àrea de la pantalla constitueix la part activa de la mateixa. A més, es pot apreciar en algunes cel·les una àrea activa que no és totalment uniforme.

Trobem també referència de la mesura de l'àrea activa d'altres LCTV del tipus *twisted-nematic* en [175]. En [210] es porta a terme una anàlisi teòrica de l'efecte de les zones mortes o no-actives del modulador, suposant-les amb transmitància nul·la, i considerant al mateix temps una transmissió uniforme sobre tota la zona activa dels píxels. En aquest estudi, s'aporten resultats numèrics en quant a l'efecte d'aquestes zones mortes en el pic de correlació, i en altres paràmetres d'avaluació del

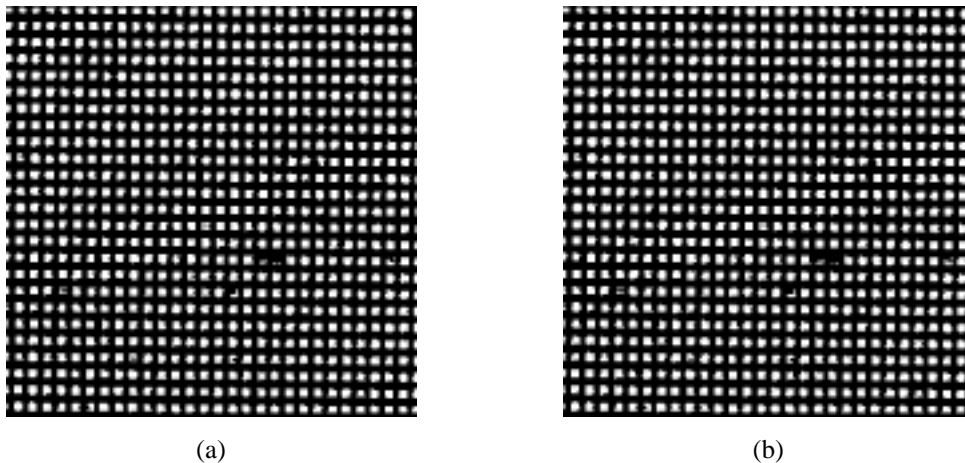
comportament dels filtres CMF, POF i BPOF, com ara l'eficiència i la relació senyal-soroll (SNR). Les conclusions extretes en aquest treball són vàlides estrictament pel correlador de Vander Lugt, però de manera orientativa podem tenir en compte el comportament del sistema en el nostre cas. En general, pels tres filtres adaptats considerats, s'observa una clara disminució del pic de correlació a mesura que augmenta la zona morta del modulador, ja sigui en el pla d'entrada o en el del filtre a l'espai de Fourier del VLC. El mateix comportament s'observa per l'energia del pla d'entrada que arriba al pla de correlació. Per escenes lliures de soroll, s'observa una variació del paràmetre SNR inferior al 10% quan es modifica el tamany de les zones mortes dels moduladors.



**Figura 7.8.** (a) Esquema d'una pantalla LCTV d'Epson amb els tamany i distàncies característics del model VP-100PS. (b) Detall de la pantalla LCTV quan en ella es reproduceix una imatge uniforme.

Les imatges reproduïdes en el modulador són enviades a través d'una targeta digitalitzadora MATROX MVP-AT. Aquestes imatges tenen un tamany de 512 x 512 píxels i cada un d'ells pren valors entre 0 i 255 nivells de gris, que és l'equivalent a una representació de 8 bits per cada píxel. La reproducció d'una imatge de 512 x 512 píxels de la targeta digitalitzadora en el modulador es realitza a

través de dues presentacions successives. La primera presentació correspon a les files parells, i a continuació es reproduïxen les files senars, de manera que cada una de les exposicions contribueix aproximadament per igual a la intensitat de la imatge total. Això ho comprovem representant una imatge de files alternes amb un nivell de gris fix però enviant primer només la informació de les línies parells, i posteriorment de les línies senars. En ambdues situacions projectem la imatge reproduïda pel modulador i observem una imatge amb nivell de gris uniforme per tots els píxels del modulador. Aquest valor correspon, aproximadament, a la meitat del valor màxim de la imatge original.



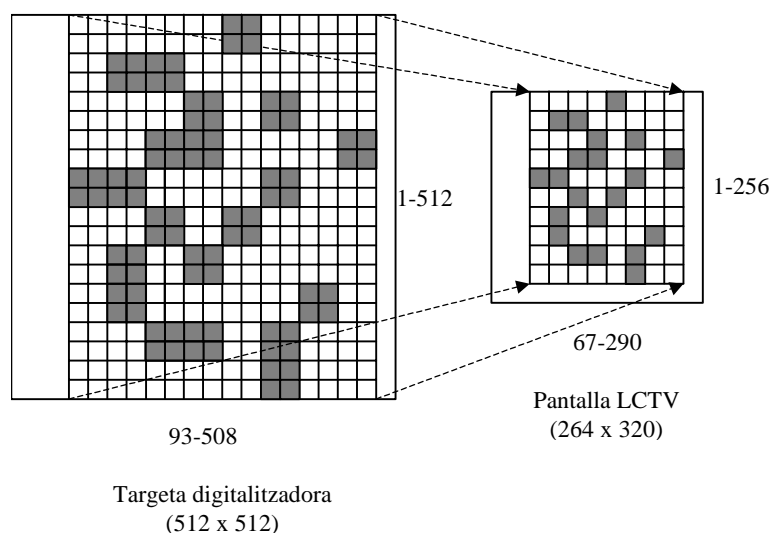
**Figura 7.9.** Part de la imatge reproduïda pel modulador, en representació binària, (a) amb la informació de les línies parells, enviada per la targeta. (b) Una imatge similar s'obté si s'envien les línies senars.

Aquesta manera de reproduir la imatge en el modulador comporta que no totes les línies del la LCTV són utilitzades. Les 512 línies de la imatge digital es converteixen en 256 línies en el modulador, concretament són les 256 primeres files de píxels. De la fila 257 a la 264 de l'SLM, no s'enregistra cap informació.

Si el mateix estudi es porta a terme per les columnes de la imatge, es determina que el comportament general coincideix amb el de les files, és a dir dues columnes de la imatge digital envien la informació a una mateixa columna del modulador. També



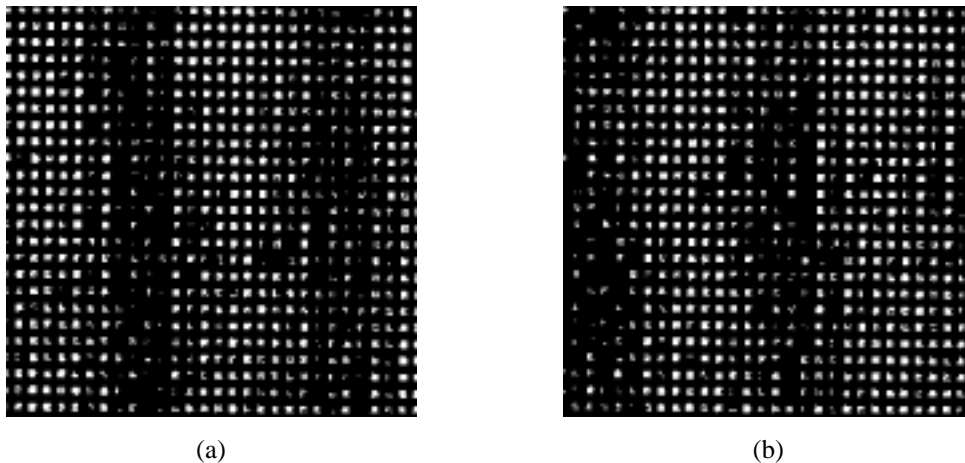
trobem un nombre de columnes en l'SLM sense utilitzar que no són activades en cap cas. Quan la reproducció de la imatge es fa per columnes, també observem que un nombre limitat de columnes, concretament les primeres i les últimes que són activades del modulador, tenen un comportament poc fiable degut a variacions temporals en la seva transmitància. Així doncs, podem dir que les columnes que s'aprofiten en la imatge digital corresponen a les columnes 93 a la 508, que activen (aproximadament de dues en dues), les columnes 67 a la 290 del modulador. Aquest comportament es descriu de manera gràfica en la Figura 7.10.



**Figura 7.10.** Reproducció en la pantalla LCTV d'una imatge digital 512 x 512 píxels emmagatzemada en la targeta de l'ordinador.

El comportament descrit coincideix de manera aproximada amb l'observat per les mateixes LCTV del videoprojector Epson model VP-100PS utilitzades en [211]. En el treball referenciat en [211] es posa de manifest la relació entre files i columnes de la targeta digitalitzadora i de la pantalla de cristall líquid, tot i que el valor numèric de les files i les columnes aprofitades en l'escriptura d'una imatge en l'SLM no coincideix. Aquest fet es pot deure a la intervenció del *software* i la targeta de digitalització, diferents en ambdós casos.

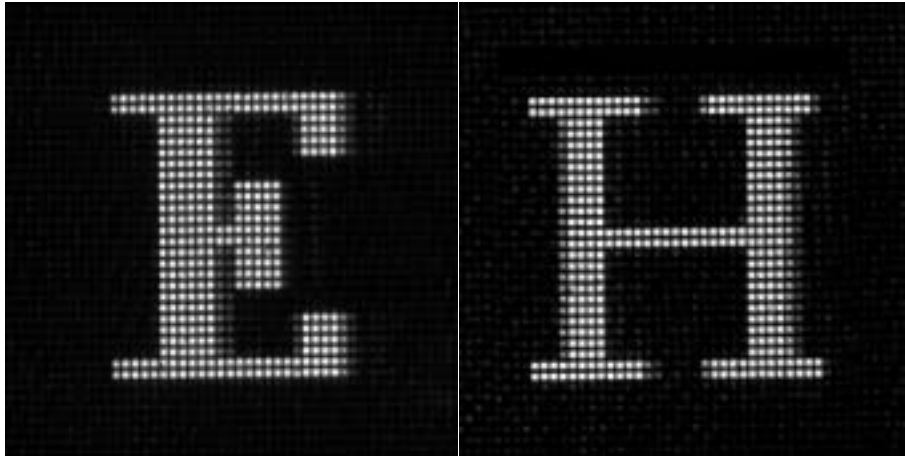
Una diferència important del comportament en la reproducció per columnes de les imatges en l'SLM, en comparació a la representació per files, radica en el fet que la contribució de dues columnes de la imatge digital no és exactament al 50% de la transmitància dels píxels. Ho comprovem també, enviant només les columnes parells, i posteriorment les columnes senars, d'un imatge de nivell de gris uniforme. En ambdós casos la imatge presentada consisteix en columnes de diferent intensitat entre elles. Observem també, com les columnes parells amb una transmitància més alta, corresponen a regions on les columnes senars contribueixen en menys mesura a la transmitància de l'LCTV (Figura 7.11).



**Figura 7.11.** Part de la imatge reproduïda pel modulador, en representació binària, (a) amb la informació de les columnes parells enviada per la targeta (b) Imatge corresponent amb la informació de les columnes senars.

En aquest estudi, també s'ha observat el fet que la informació inclosa en una columna de la imatge digital, activa, en alguns casos i sense periodicitat aparent, dues columnes consecutives de la LCTV. Aquest fet, que depèn de l'electrònica de l'aparell, fa de moment impossible un control columna per columna en l'escriptura de les imatges en el modulador, i és el responsable de no poder tenir un control acurat sobre els contorns verticals dels objectes reproduïts en l'LCTV. La Figura 7.12 vol posar de manifest aquest inconvenient. En ella es mostren dues lletres que

han estat enviades separatament al modulador. En ambdós casos es pot observar com apareixen algunes columnes tènues en el exterior dels contorns del caràcter.



**Figura 7.12.** Caràcters representats en la pantalla LCTV on s'observen les deficiències en el control de les columnes de la imatge.

Binaritzant les mateixes imatges de la Figura 7.12 a partir d'un llindar en el valor de gris 195, podem observar que la informació captada no és uniforme en tots els píxels activats de l'SLM (Figura 7.13). A més, s'observa que alguns píxels activats en les columnes sorolloses arriben a assolir una transmitància elevada, superior al llindar (veure la columna dreta de la lletra H).

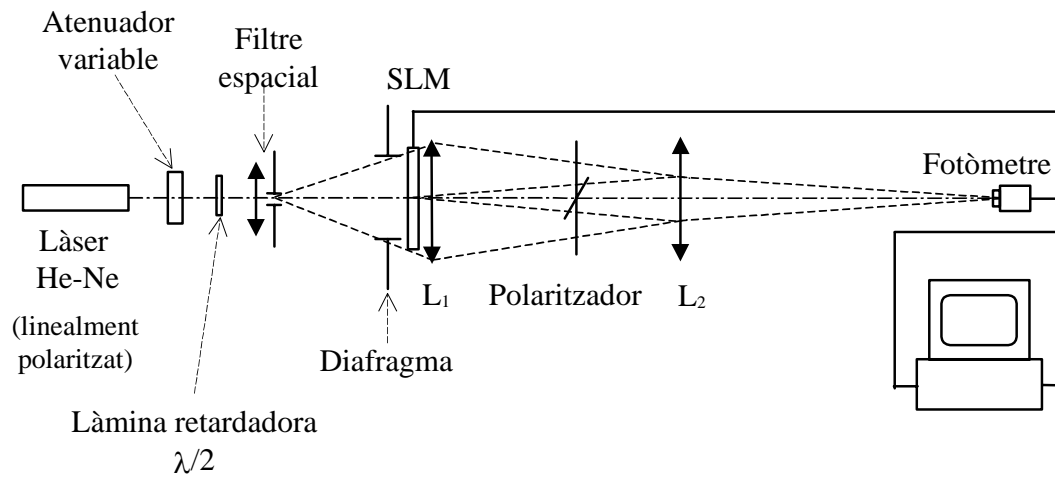


**Figura 7.13.** Caràcters de la Figura 7.12 binaritzats amb el llindar de 195 en el nivell de gris.

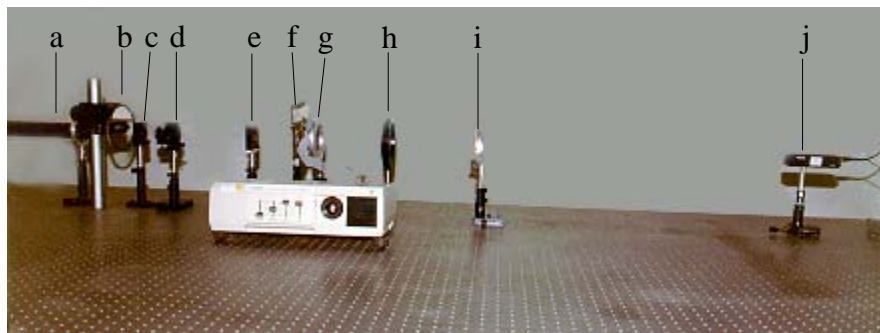
### 7.4.2 Determinació del règim de modulació d'amplitud de la pantalla LCTV

En la nostra realització experimental, hem utilitzat la pantalla corresponent al canal vermell del videoprojector i, en tot moment, s'ha enviat el senyal elèctric pel mateix canal cromàtic de la targeta digitalitzadora connectada al canal R de l'electrònica de l'aparell. Hem treballat en totes les experiències realitzades en el règim denominat de modulació principalment d'amplitud. Per estar en les condicions idònies de treball, cal incidir amb un feix de llum polaritzat lineal perpendicular a la direcció de les molècules del cristall líquid en la superfície d'incidència. La direcció del l'eix director de la LCTV però, no vé especificada pel fabricant. Per aquest motiu, el primer pas de la caracterització de la pantalla de cristall líquid és conèixer la direcció de les molècules dins el cristall. Per fer-ho utilitzarem les tècniques òptiques descrites en la referència [183].

El muntatge emprat, es mostra de forma esquemàtica en la Figura 7.14. I en la Figura 7.15 es mostra una fotografia general del dispositiu òptic. El làser emprat emet un feix polaritzat linealment amb potència de 7 mW. El procediment per trobar la direcció de l'eix de l'SLM consisteix en variar el pla de polarització de la llum incident mitjançant una làmina  $\lambda/2$ . Només en el cas en que el pla de polarització de la llum incident coincideixi amb la direcció del director del cristall, o bé sigui perpendicular a ell, la llum a la sortida de la LCTV seguirà sent polaritzada lineal. En qualsevol altre cas obtindrem llum polaritzada el·líptica. Per tant, si situem un polaritzador a la sortida de l'SLM, podrem girar el seu eix d'extinció fins a detectar un mínim d'energia mitjançant un fotòmetre. L'orientació del polaritzador en el moment de detectar aquest mínim, ens indica la posició de l'eix director del modulador i la seva perpendicular. Aquesta operació es va portar a terme inicialment amb el modulador apagat.



**Figura 7.14.** Esquema del muntatge per a la calibració de la pantalla LCTV.



**Figura 7.15.** Vista general del dispositiu òptic emprat en la calibració del modulador espacial de llum. Destaquem: *a*, làser He-Ne; *b*, atenuador variable; *c*, làmina  $\lambda/2$ ; *d*, filtre espacial; *e*, obertura rectangular; *f*, pantalla LCTV; *g*, lent; *h*, polaritzador; *i*, lent i *j*, càmera CCD.

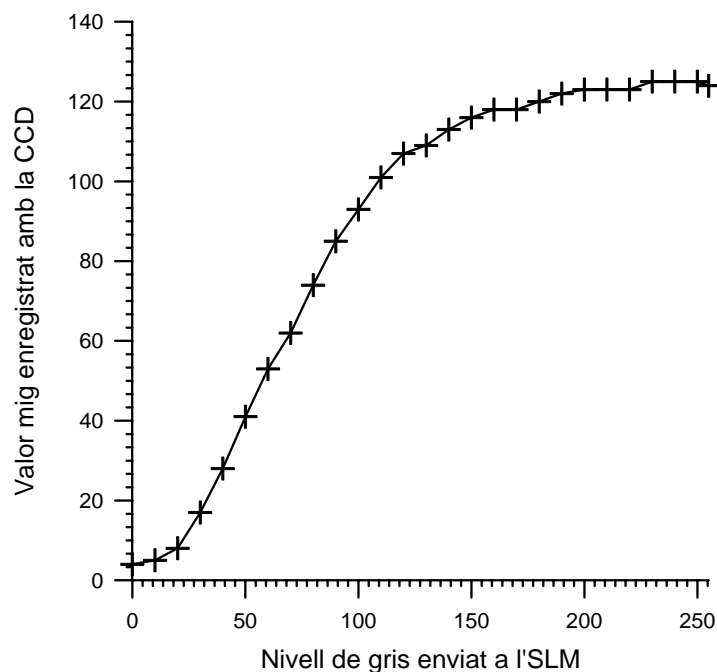
Comentem a continuació una observació experimental que ens ha portat a corregir lleugerament les posicions de la làmina  $\lambda/2$  i del polaritzador per a fixar el mode de funcionament dels SLM. A l'engegar el modulador i reproduir en ell una distribució uniforme de nivell de gris zero, s'enregistrava un augment en l'energia transmesa respecte la situació en la que l'SLM estava apagat. El fet que la pantalla del modulador tingui una composició mixta, amb zones actives i zones no actives i, que, a la vegada, les zones actives puguin variar lleugerament en el seu comportament quan el modulador està apagat i quan està operatiu i s'envia un valor digital zero,

permet explicar la variació registrada en l'energia transmesa. Amb aquest ajust del modulador està clar que ni tant sols enviant un nivell de gris zero, s'assoleix un bon mínim energètic, la qual cosa afecta de manera important al contrast de les imatges. Aquest efecte ha estat també descrit per Soutar *et al.* [179]. Considerant que és important obtenir un bon nivell de zero per poder disposar d'un millor contrast en les imatges, reajustem la làmina  $\lambda/2$ , de manera que amb la pantalla connectada i reproduint una imatge uniforme de nivell de gris zero, obtenim el mínim de llum transmesa pel sistema. La variació en graus de la posició de la làmina ha estat de  $4^\circ$ .

El següent pas consisteix en determinar les corbes de transmissió de la pantalla LCTV per a diferents valors dels potenciòmetres *brillo*, *contrast* i *color*. Per realitzar aquesta segona part de la caracterització, substituïm el fotòmetre per la càmera CCD prèviament calibrada. Amb la càmera captem la imatge del modulador, quan en ell es representen imatges de nivell de gris uniforme que anirem variant, i calculem el valor mig de la imatge enregistrada.

Tenint en compte, el tipus d'imatges que s'han de reproduir en el modulador, en alguns casos binàries, i en d'altres imatges en nivells de gris, ens cal trobar una corba de treball del modulador que ens proporcioni un bon contrast de les imatges, i una reproducció el més lineal possible dels diferents nivells de gris. Per a obtenir el millor contrast i una bona reproducció del nivell de zero, també és important controlar la intensitat del feix que il·lumina la pantalla de l'SLM. Amb l'atenuador variable situat just a la sortida del làser (Figura 7.14) regulem el nivell d'intensitat de manera que la transmissió del modulador al reproduir un nivell de gris uniforme igual a zero, sigui captat per la càmera i digitalitzat amb valor mínim, i augmenti linealment amb les variacions de nivell de gris. La corba obtinguda experimentalment que compleix aquests requisits es mostra en la Figura 7.16 i correspon als valors de *brillo* = 4, *contrast* = 9.5 i *color* = 10 per a una posició fixada de l'atenuador. En aquesta posició de l'atenuador, la il·luminació de l'SLM és aproximadament uniforme de  $0.237 \text{ mW/cm}^2$ .

Com a resultat d'aquesta calibració prèvia, s'han fixat les posicions més convenientes de tots els elements, així com la posició dels potenciòmetres del videoprojector (*Brillo=4*, *Contrast =9.5* i *Color =10*). Comparant amb treballs anteriors d'altres autors relatius a la caracterització del mateix tipus de pantalles Epson LCTV en règim majoritàriament d'amplitud [183,184], els nostres valors concorden amb els seus.



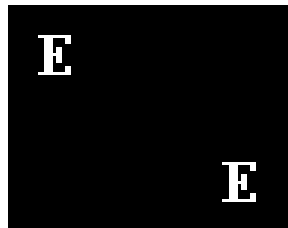
**Figura 7.16.** Corba de calibració de la pantalla LCTV en règim quasi d'amplitud.

## 7.5 Posada en funcionament del correlador no-lineal dual

El següent pas després de la caracterització dels elements optoelectrònics del sistema, és l'alineació i posada a punt final del dispositiu per a la consecució de bons resultats de correlació aplicant la DNC.

Per a realitzar el primer pas de la DNC, cal adequar la posició i el tamany dels objectes representats en el pla d'entrada del correlador, per a poder enregistrar correctament la informació continguda en les distribucions d'intensitat  $I$ ,  $I_S$  i  $I_R$ .

Situem doncs, per assolir aquest propòsit, el pla d'entrada consistent en la lletra  $E_{ref}$  com a patró i a la vegada com escena segmentada a analitzar (Figura 7.17). Per poder enregistrar la distribució espectral conjunta (JPS) i enregistrar-la amb la càmera, és necessari reproduir el pla d'entrada de manera que els objectes tinguin un àrea important en relació a l'àrea del fons. Recordem que aquesta àrea no té un transmitància totalment nul·la amb la qual cosa la seva contribució en el JPS pot ésser important en comparació amb la dels caràcters. Per solucionar aquest inconvenient es pot augmentar, per una banda, l'àrea ocupada pels caràcters, o bé, es pot emprar una obertura rectangular que limiti la llum transmesa en les regions laterals del LCTV no-actives. Amb aquestes condicions si que és possible d'enregistrar la distribució espectral conjunta  $I$ , i les distribucions espectrals de l'escena i de la referència per separat,  $I_S$  i  $I_R$ .



**Figura 7.17.** Pla d'entrada del correlador no-lineal dual.

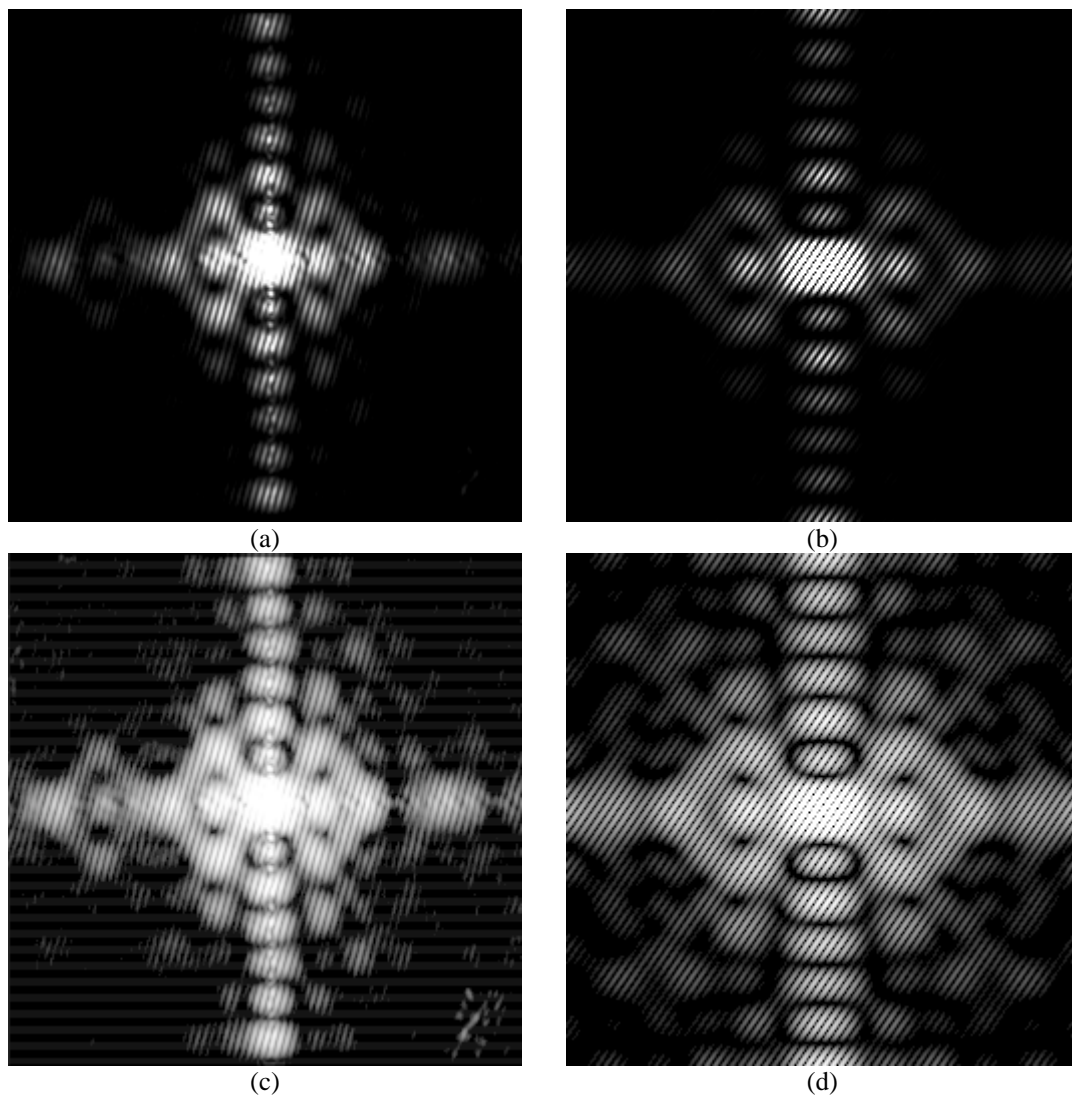
Fixat el tamany de les lletres, també cal fixar la distància entre elles. Recordem que quan més properes estiguin, obtindrem franges d'interferència més amples i en conseqüència enregistrades per la CCD amb un major número de píxels, i posteriorment reproduïdes en el modulador amb més fiabilitat i qualitat. El problema



però està en el solapament dels ordres de difracció en el pla de sortida del correlador, si els caràcters del pla d'entrada es situen massa propers. Analitzant doncs, ambdós factors, amplada de franges i solapament dels ordres de difracció, determinem i fixem la millor posició de compromís de l'escena a analitzar i l'objecte de referència en el pla d'entrada de la DNC.

Mostrem en la Figura 7.18a la distribució d'interferències de l'espectre conjunt  $I$  obtingut experimentalment quan figuren dues lletres  $E_{ref}$  en el pla d'entrada de l'SLM (Figura 7.17). Per facilitar-ne la comparació, reproduïm en la Figura 7.18b el corresponent espectre  $I$  que s'ha obtingut per simulació numèrica considerant una saturació d' $\alpha=10\%$ , i un enregistrament amb 256 nivells de gris del sistema de captació. Les representacions logarítmiques d'ambdues figures, que es mostren en les Figures 7.18c i 7.18d respectivament, permeten apreciar la distribució d'intensitat amb més detall.

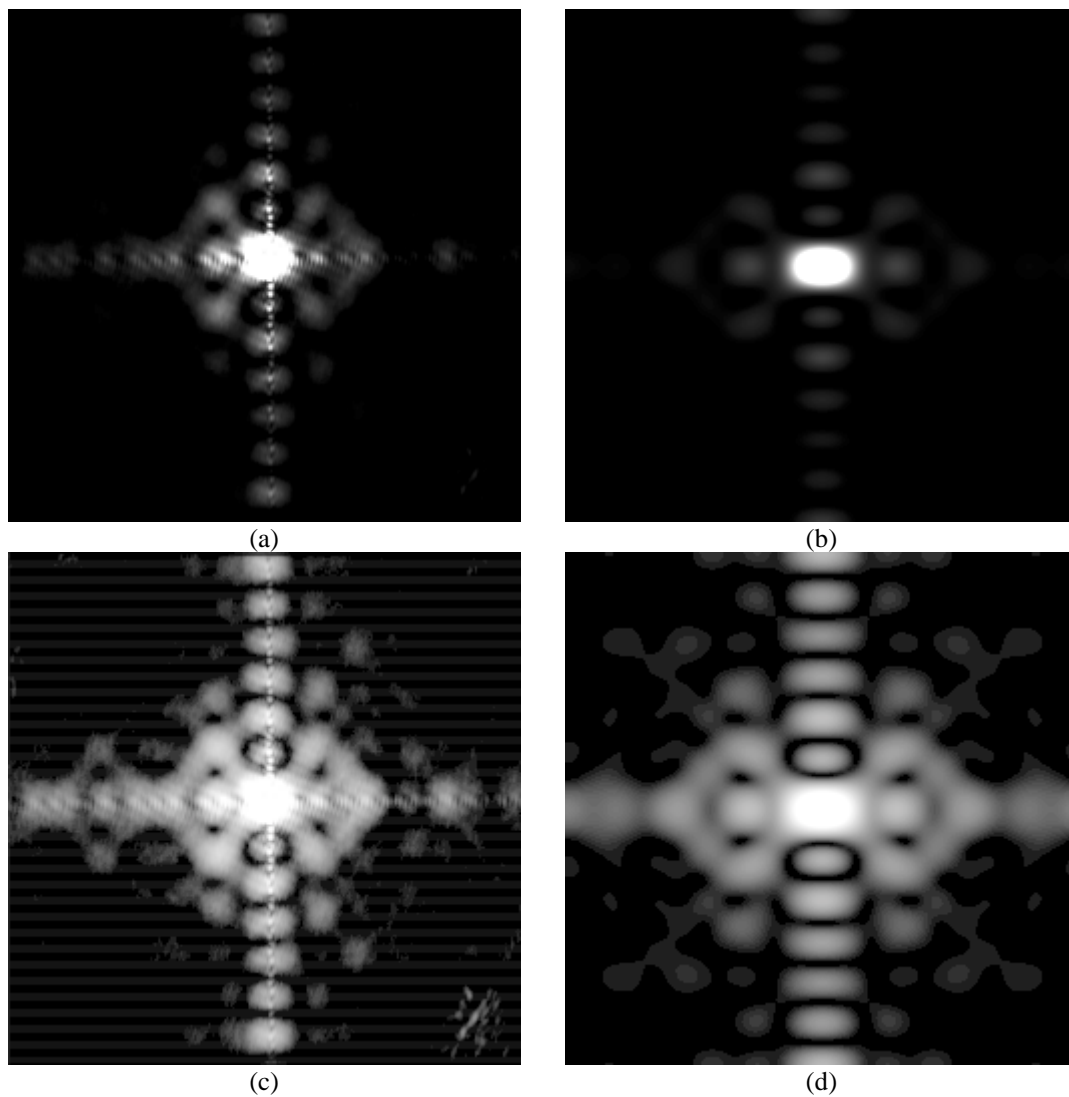
Observem en la Figura 7.18 la gran semblança entre els dos espectres tant en quant a l'àrea que conforma la regió de suport, com a l'envolupant de l'espectre. En les imatges òptiques (Figures 7.18a i 7.18c) s'aprecia un encongiment en la direcció horitzontal que és degut a l'asimetria present tant en la pantalla de cristall líquid ( $V/H=0.82$ ), que es reproduïx en sentit invers en l'espectre de transformades conjunt ( $V/H=1.22$ ), com de la càmera CCD ( $V/H=1.13$ ) mitjançant la qual s'enregistren les imatges. L'efecte conjunt dels dos encongiments horitzontals, provoca una relació asimètrica en les imatges representades en les Figures 7.18a i 7.18c, aproximadament de  $V/H=1.38$ . En aquestes mateixes distribucions obtingudes experimentalment, podem observar un patró de difracció al llarg de l'eix vertical i horitzontal de les figures, corresponent a l'obertura rectangular que limita la zona il·luminada sobre la LCTV. Aquesta figura de difracció, en canvi, no està present en les distribucions obtingudes per simulació numèrica donat que no s'ha considerat l'efecte del diafragma.



**Figura 7.18.** Distributions d'intensitat de l'espectre conjunt  $I$  quan es reproduïx el pla d'entrada (SLM) mostrat en la Figura 7.17. (a) obtingut experimentalment i (b) obtingut per simulació numèrica ( $\alpha=10\%$  i 256 nivells de gris de la ccd); (c) i (d) contenen les representacions logarítmiques d'(a) i (b), respectivament.

També veiem que la zona saturada central en ambdós espectres ocupa una àrea molt similar. Donat que la distribució experimental i simulada no s'allunyen massa en quant a la regió saturada d'ambdues funcions, decidim mantenir la posició de l'atenuador prèviament fixada en la calibració del modulador.

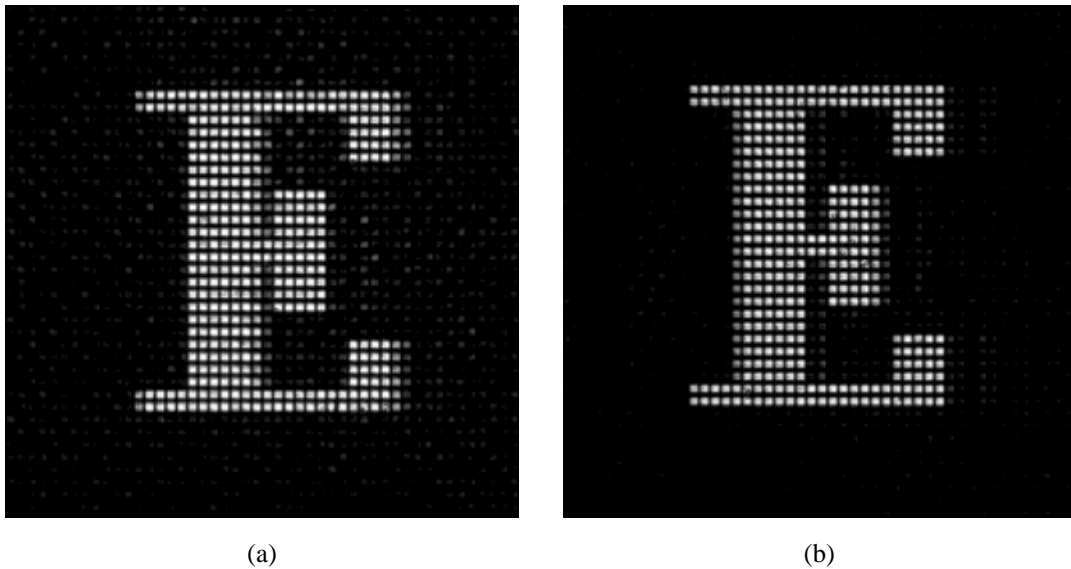
També és necessari enregistrar la distribució espectral de l'escena i de la referència, separatament. La Figura 7.19a mostra la funció  $I_S$  obtinguda en el procés experimental i la Figura 7.19b consisteix en la mateixa distribució calculada per simulació.



**Figura 7.19.** Distribucions d'intensitat a)  $I_S$  ( $I_R$ ) experimentals i b) simulades; (c) i (d) són representacions logarítmiques d'(a) i (b), respectivament.

Donat que l'escena i el patró són idèntics, la  $I_S$  simulada és una funció idèntica a la funció  $I_R$  (Figura 7.19). Experimentalment però, s'observen algunes lleugeres

variacions entre  $I_S$  i  $I_R$  degudes principalment a la poca definició dels contorns verticals dels caràcters pels problemes de l'SLM en la reproducció de les columnes (Figura 7.20).








**Figura 7.20.** Reproducció del caràcter  $E_{ref}$  en la posició a) de l'escena a analitzar i b) de l'objecte de referència.

## 7.6 Resultats experimentals de la correlació no-lineal dual

Amb el muntatge descrit gràficament en la Figura 7.1, realitzem de manera experimental la DNC. Analitzarem primerament el comportament experimental del sistema per diferents valors d' $L$  i  $M$ , així com la variació en la capacitat discriminativa davant de petits canvis de forma en els objectes analitzats quan aquests són tots ells equienergètics. Posteriorment, realitzarem la DNC experimentalment, per poder obtenir diferents graus de capacitat de discriminació quan els caràcters a analitzar són d'igual forma, però difereixen en el nivell de gris uniforme que els recobreix.

### 7.6.1 Variació de la capacitat de discriminació del correlador no-lineal dual davant variacions de forma

El conjunt d'escenes segmentades utilitzades com a test en aquesta realització experimental de la DNC, està format per 4 caràcters dels ja emprats en l'apartat 6.2 del capítol anterior. Representem en la Figura 7.21 les lletres seleccionades a partir de les del test 1 de la Figura 6.1 del capítol 6.

Etiqueta	$E_{ref}$	$E_2$	$F_2$	$E_5$	$P_1$
Objecte (lletra)					
Producte escalar $cs = \frac{\vec{s} \cdot \vec{r}}{ \vec{r} ^2}$	1	0.92	0.86	0.84	0.76

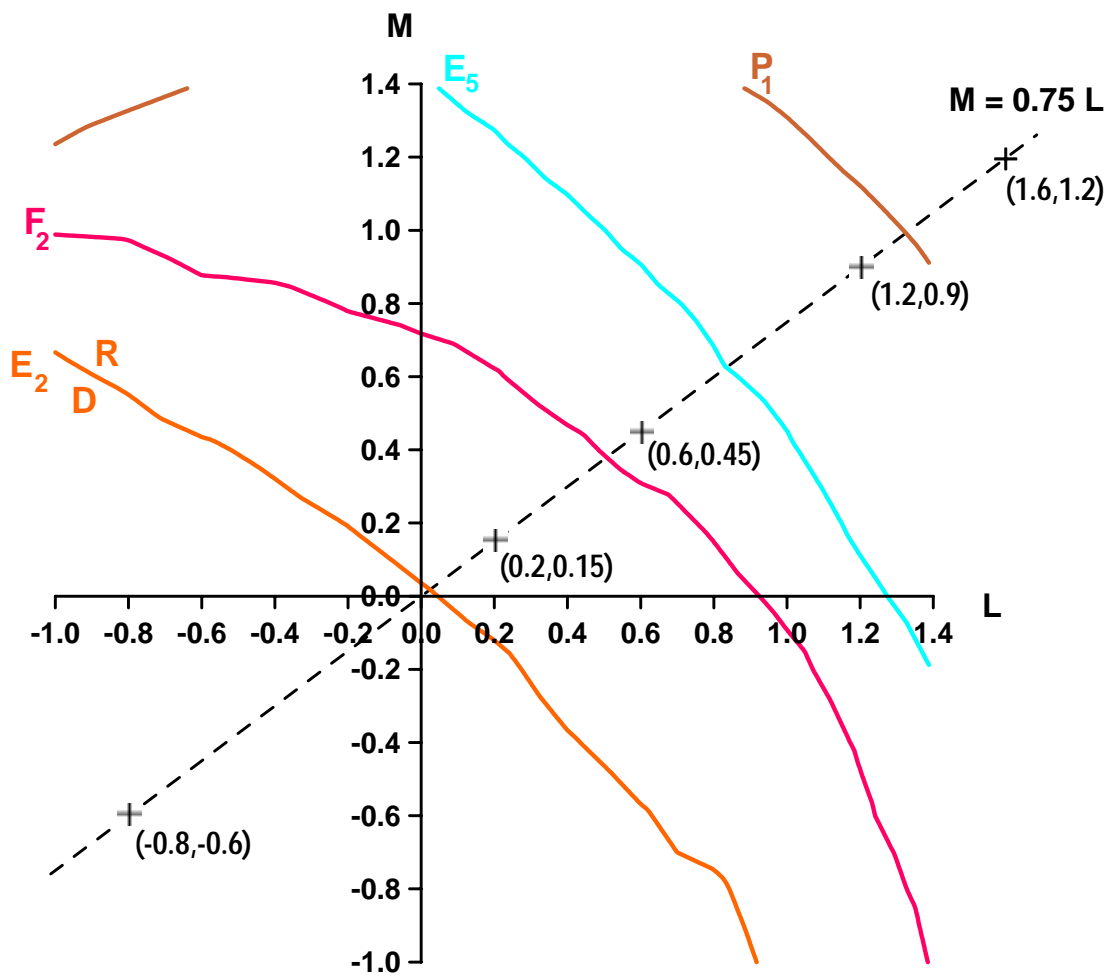
**Figura 7.21.** Conjunt d'escenes test per a la realització experimental de la DNC amb capacitat de discriminació variable davant de canvis de forma.

En la Figura 7.22 es representen, per a cada lletra del test de la Figura 7.21, la corba del llinar de reconeixement que es va obtenir per simulació numèrica aplicant la DNC en condicions experimentals (Figura 6.5). Cada corba uneix els parells ( $L$ ,  $M$ ) pels quals el valor de la correlació creuada normalitzada (al valor d'autocorrelació de la referència) iguala el valor llinar de reconeixement, és a dir, aquells parell ( $L$ ,  $M$ ) pels quals s'obté  $cc/ac = 0.5$ . Recordem que les escenes són equienergètiques i, en general, el pic de correlació creuada és inferior al d'autocorrelació en el marge de no-linealitats considerades. Per aquest motiu, les corbes llinars superiors de la banda de reconeixement ( $cc/ac=1.5$ ) no apareixen en aquesta representació gràfica. Cada corba  $cc/ac = 0.5$ , per tant, separa la zona de valors d' $L$  i  $M$  pels quals s'aconsegueix el reconeixement de la lletra com la lletra de referència (marcada amb


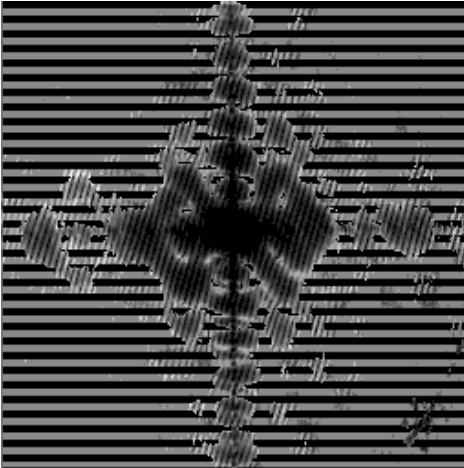
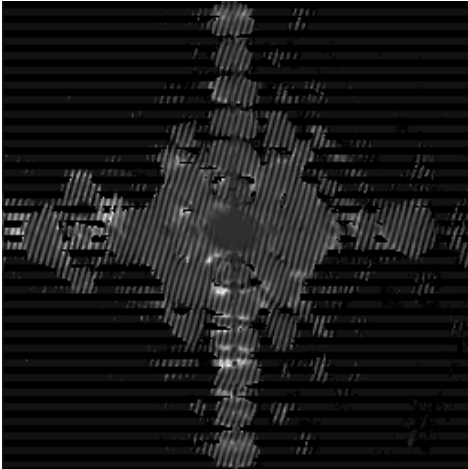
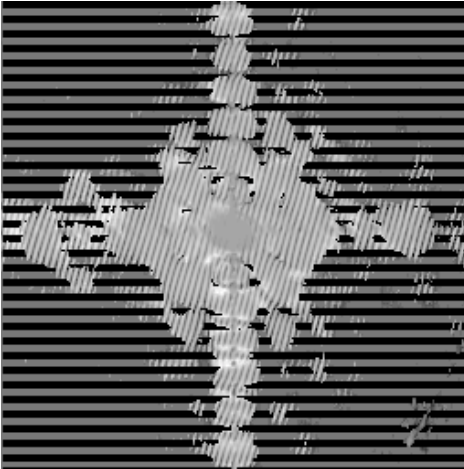
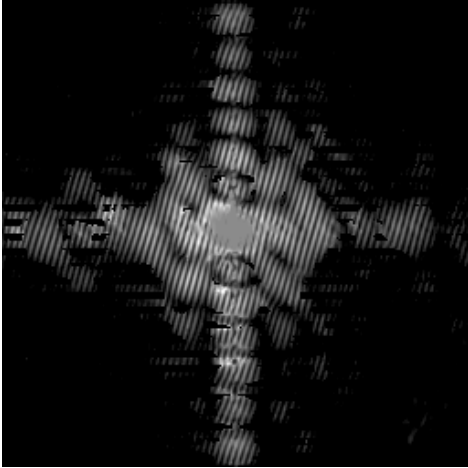
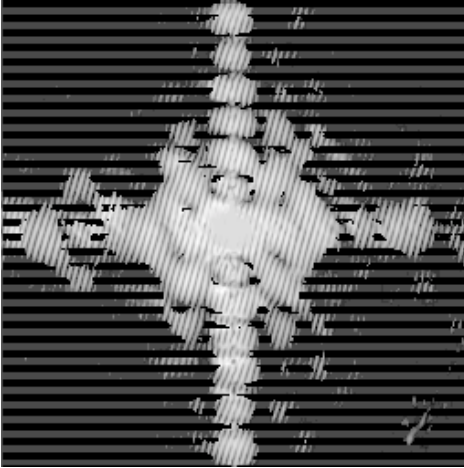
R en la Figura 7.22), de la zona de valors d' $L$  i  $M$  pels quals la lletra es discrimina de la lletra de referència (marcada amb D en la Figura 7.22). També s'ha assenyalat la línia  $M = 0.75 L$  sobre la qual prendrem, en les proves experimentals, els valors dels paràmetres  $L$  i  $M$  de les no-linealitats a aplicar sobre l'escena a analitzar i l'objecte de referència, respectivament. Sobre la línia, s'han marcat els punts  $(L, M)$  que utilitzarem experimentalment en la DNC per aconseguir diferent capacitat de discriminació. En un extrem, pel valor  $(-0.8, -0.6)$ , tots els caràcters haurien de ser discriminats de la referència i, en l'altre extrem, per  $(1.6, 1.2)$ , tots els caràcters haurien de ser reconeguts com a iguals al patró. Entremig d'aquests valors extrems, s'han escollit altres parells  $(L, M)$  que corresponen a una discriminació teòricament intermèdia i que haurien de permetre un reconeixement progressiu dels caràcters.

En el primer pas de la DNC, enregistrem de manera successiva les distribucions d'intensitat de l'espectre de cada una de les lletres test, quan són reproduïdes conjuntament amb la referència en el pla d'entrada del correlador ( $I$ ) i quan es mostren individualment ( $I_S$ ). Enregistrem també la funció d'intensitat de l'espectre de referència ( $I_R$ ), que és una distribució comú per a totes les lletres test.

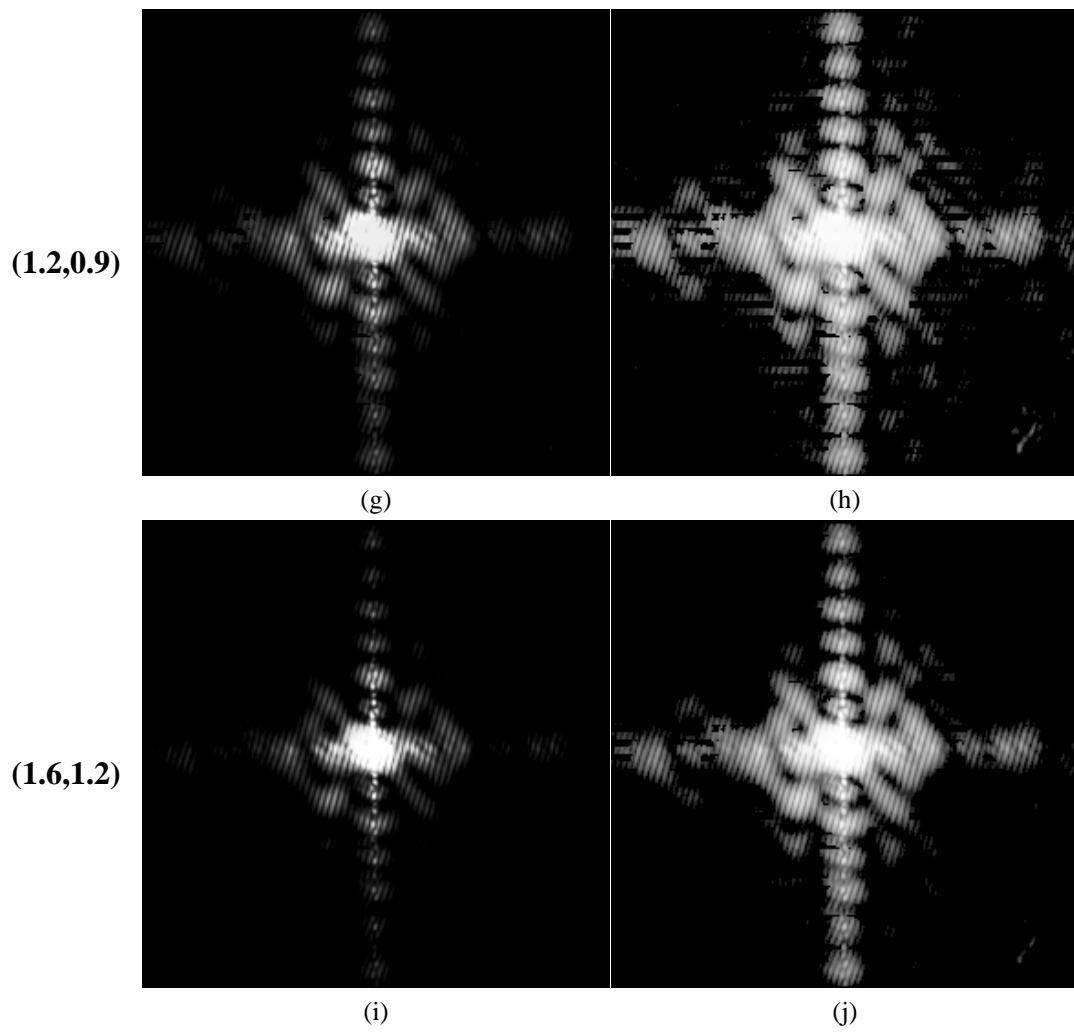
Amb la informació enregistrada, realitzem el càlcul de l'espectre conjunt no-linealment modificat, pels parells  $(L, M)$  seleccionats, i per a cada una de les lletres test de la Figura 7.21. La Figura 7.23 mostra, a tall d'exemple, els MJPS corresponents a diferents  $(L, M)$  representats de manera lineal i logarítmica, per a una de les lletres del test (pel caràcter  $F_2$ ). Per facilitar-ne la comparació, es representa en la Figura 7.24 els corresponents MJPS obtinguts per simulació numèrica. En la simulació s'ha emprat l'algorisme corresponent a la funció  $P'(\vec{v})$  (Eq. 4.27), i s'han considerat els condicionaments experimentals raonats en el capítol 5 d'aquesta memòria (llindar de saturació d' $\alpha=10\%$  i 256 nivells de quantització del sistema de captació).



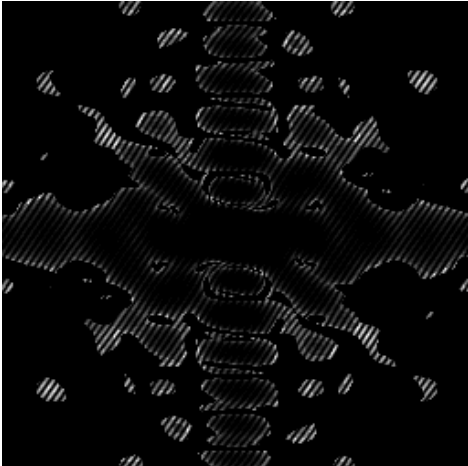
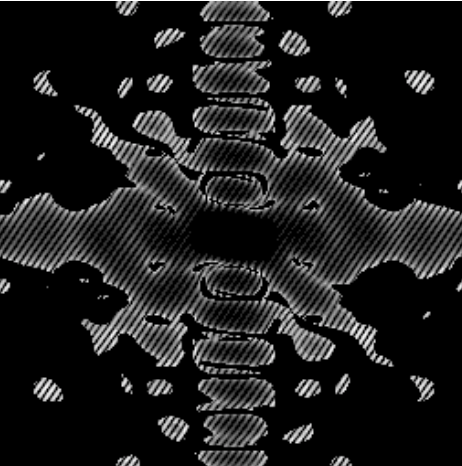
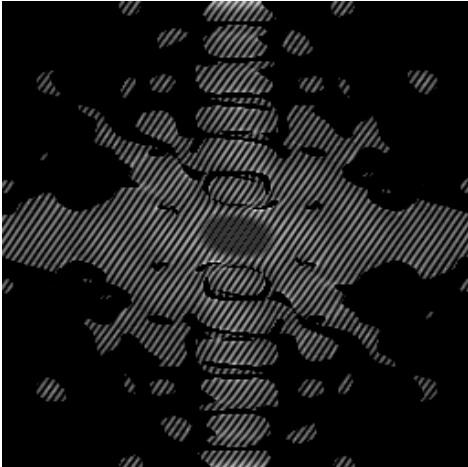
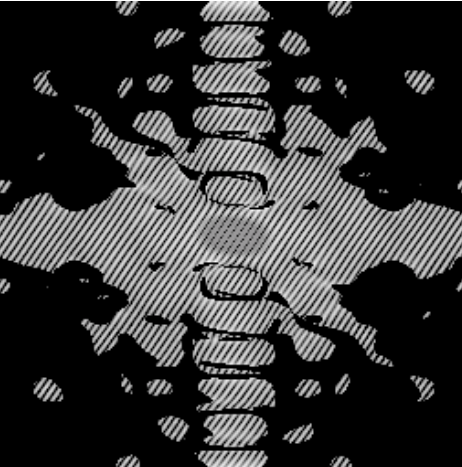
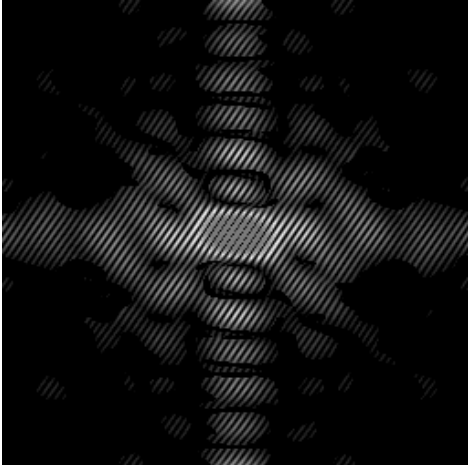
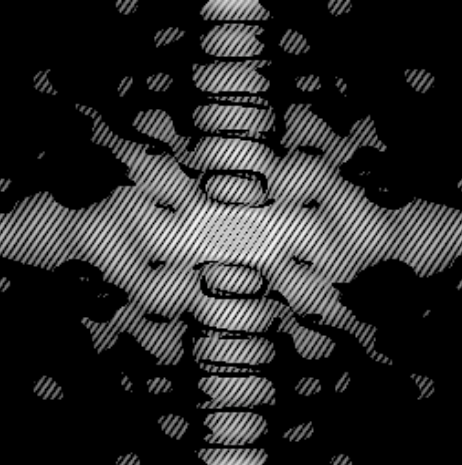
**Figura 7.22.** Pla  $LM$  amb les corbes  $cc/ac = 0.5$  de reconeixement i/o discriminació pels caràcters test seleccionats (Figura 7.21). S'ha assenyalat la recta  $M = 0.75 L$  sobre la qual s'han seleccionat els valors dels paràmetres  $L$  i  $M$  per a la realització experimental de la DNC.

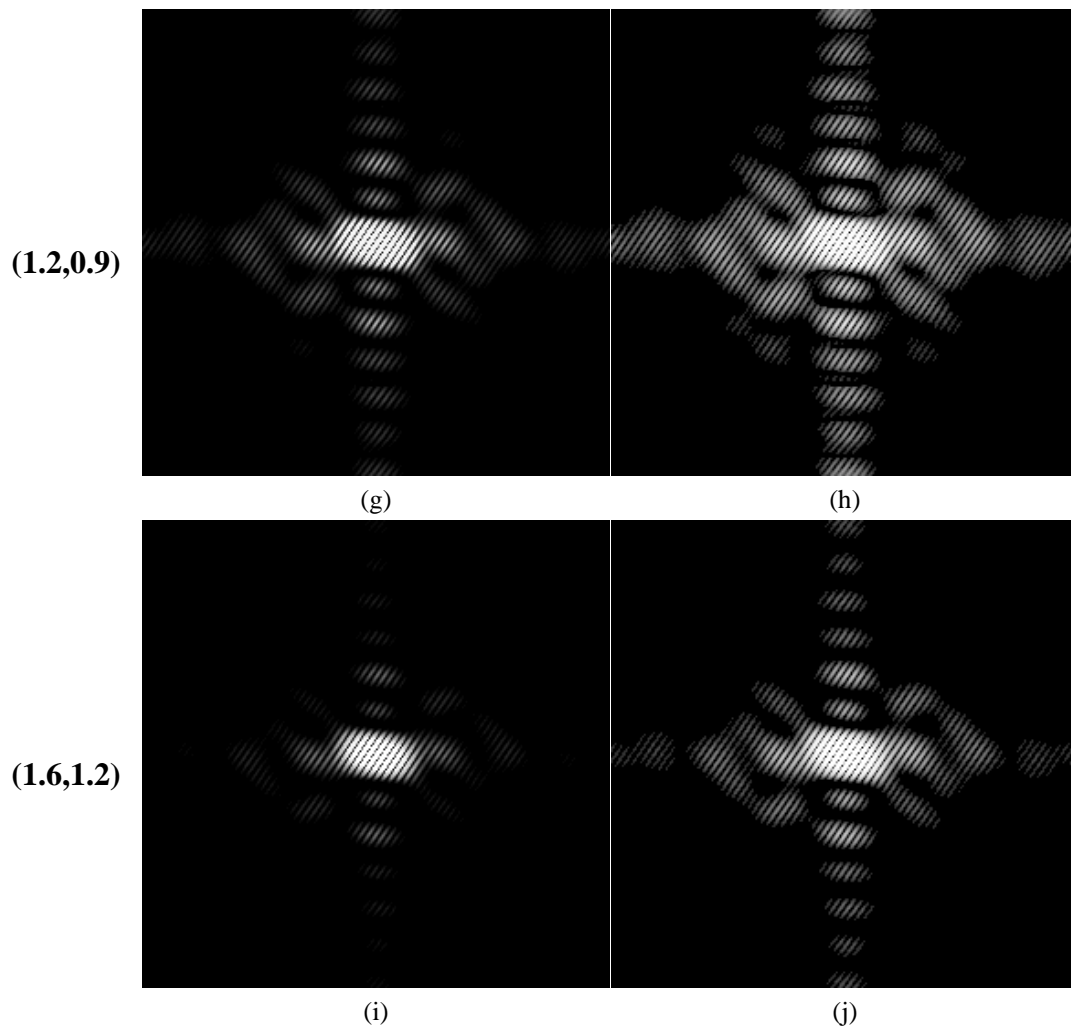
$(L, M)$	Representació lineal	Representació logarítmica
$(-0.8, -0.6)$	 (a)	 (b)
$(0.2, 0.15)$	 (c)	 (d)
$(0.6, 0.45)$	 (e)	 (f)





**Figura 7.23.** MJPS obtinguts experimentalment per a la lletra  $F_2$  (escena) i la lletra  $E_{ref}$  (referència), aplicant diferents no-linealitats. A la columna de l'esquerra la representació de les distribucions d'intensitat és lineal, i a la dreta, és logarítmica.

(L, M)	Representació lineal	Representació logarítmica
(-0.8,-0.6)	 (a)	 (b)
(0.2,0.15)	 (c)	 (d)
(0.6,0.45)	 (e)	 (f)



**Figura 7.24.** MJPS obtinguts per simulació considerant condicions experimentals (saturació  $\alpha=10\%$  i 256 nivells de gris del sistema de captació) per a la lletra  $F_2$  (escena) i la lletra  $E_{ref}$  (referència), aplicant diferents no-linealitats. A la columna de l'esquerra la representació de les distribucions d'intensitat és lineal, i a la dreta, logarítmica.

En línies generals, podem dir que els espectres conjunts modificats (MJPS) predits per simulació s'acosten en gran mesura als obtinguts experimentalment, donat que la regió de suport en ambdós casos és molt similar per a cada no-linealitat aplicada, i el contrast de les franges varia seguint el mateix comportament. És a dir, a l'aplicar diferents no-linealitats, partint de valors negatius d' $L$  i  $M$  i augmentant-los fins a assolir valors positius i majors que la unitat, observem, d'una banda, com la regió de suport va variant pel fet que en el primer cas s'accentua la contribució de les

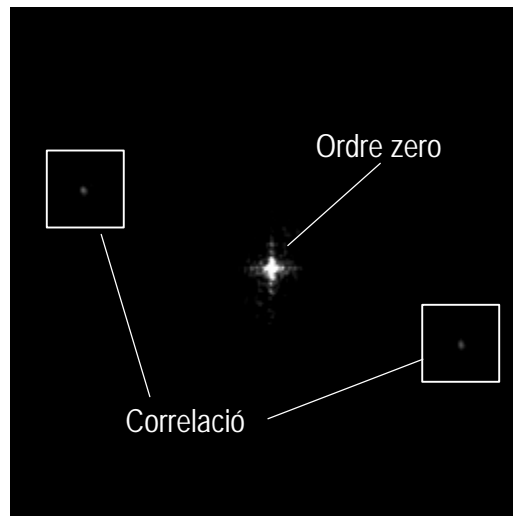
freqüències més altes, i quan  $L, M > 1$ , són les freqüències més baixes les que adquireixen un paper més rellevant. Per altra banda, el contrast de les franges s'incrementa a l'augmentar els paràmetres  $L$  i  $M$  cap a valors positius i superiors a 1.

Les diferències principals entre les distribucions MJPS experimentals, i les obtingudes per simulació, radiquen en l'encongiment de les funcions experimentals en la direcció horitzontal per les raons abans esmentades, la presència del patró de difracció de l'obertura rectangular només present en les distribucions òptiques, i un reforçament important del patró periòdic de franges en el cas dels valors més baixos i negatius d' $L$  i  $M$ , degut al soroll del sistema de captació sempre present en qualsevol enregistrament.

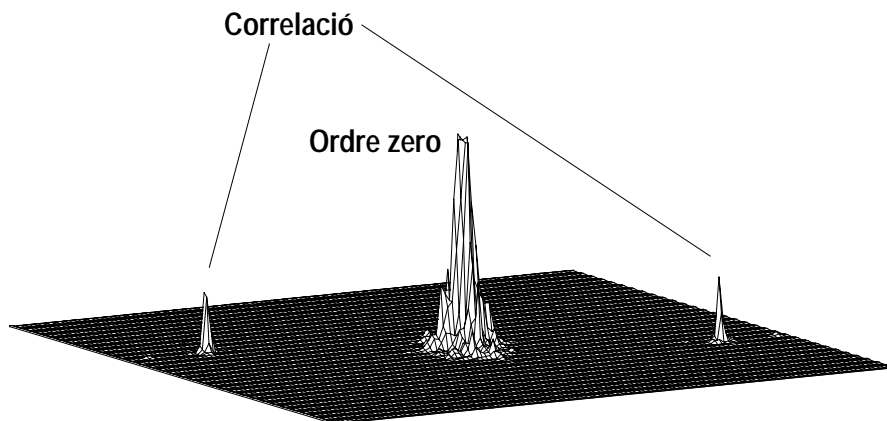
Les distribucions MJPS per a cada parell  $(L, M)$  escollit, es reproduïxen en el modulador en el segon pas de la DNC. Una segona transformada de Fourier ens permet obtenir el pla de sortida del correlador no-lineal dual per a cada lletra i parell  $(L, M)$  amb els corresponents pics de correlació.

Com a exemple, si la lletra  $E_{\text{ref}}$  es pren com a escena test i també com a referència, el pla de sortida del correlador no-lineal dual en el cas de reproduir en el modulador el MJPS calculat pels paràmetres  $L=0.6$  i  $M=0.45$ , és el que es presenta en la Figura 7.25. En ell s'observa bàsicament tres termes. L'ordre central o zero de difracció, on s'obté la suma de diferents termes relatius a la convolució de l'escena amb ella mateixa, i de la referència amb sí mateixa, i que no aporta informació rellevant en el procés de correlació. Els dos termes que apareixen a banda i banda d'aquest terme central corresponen als ordres  $\pm 1$  de difracció, i és en ells on observem clarament el pic de correlació entre l'escena i la referència en una banda, i viceversa, entre la referència i l'escena sobre la diagonal en el costat contrari.

En la Figura 7.26 representem la funció 3D d'intensitats de correlació. S'ha truncat l'ordre central per fer possible l'enregistrament dels pics de correlació.

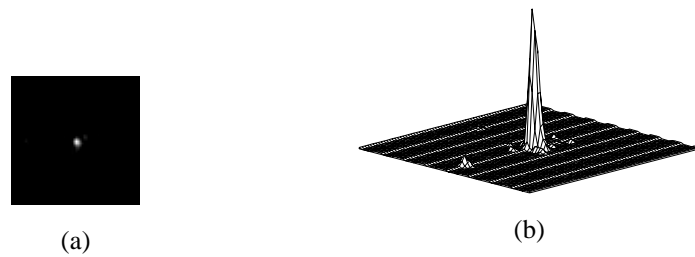


**Figura 7.25.** Pla de correlació obtingut quan el MJPS calculat prenent  $E_{\text{ref}}$  com a escena i com a referència, amb  $L = 0.6$  i  $M = 0.45$ , es reproduïx en el SLM en el segon pas de la DNC.



**Figura 7.26.** Pla de correlació de la Figura 7.25 en representació 3D. S'ha truncat l'ordre central per poder visualitzar els pics de correlació.

En la Figura 7.27a mostrem pròpiament el pla de correlació en una imatge 2D, i en la Figura 7.27b el representem com una superfície tridimensional. Ambdues representacions s'han extret de les Figures 7.25 i 7.26, respectivament.

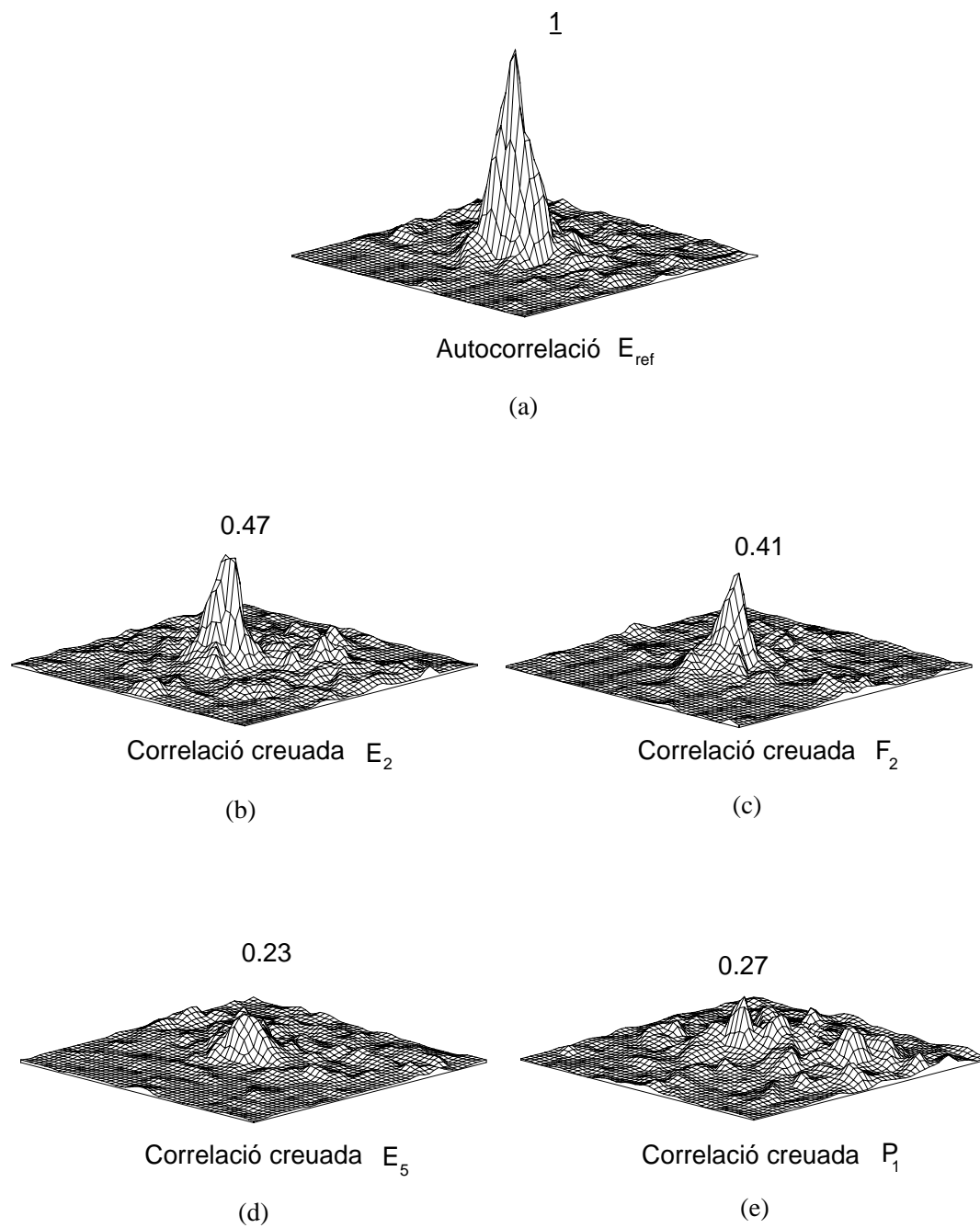


**Figura 7.27.** Intensitat de correlació estreta de les Figures 7.25 i 7.26, en a) una representació 2D i en b) com a una superfície tridimensional.

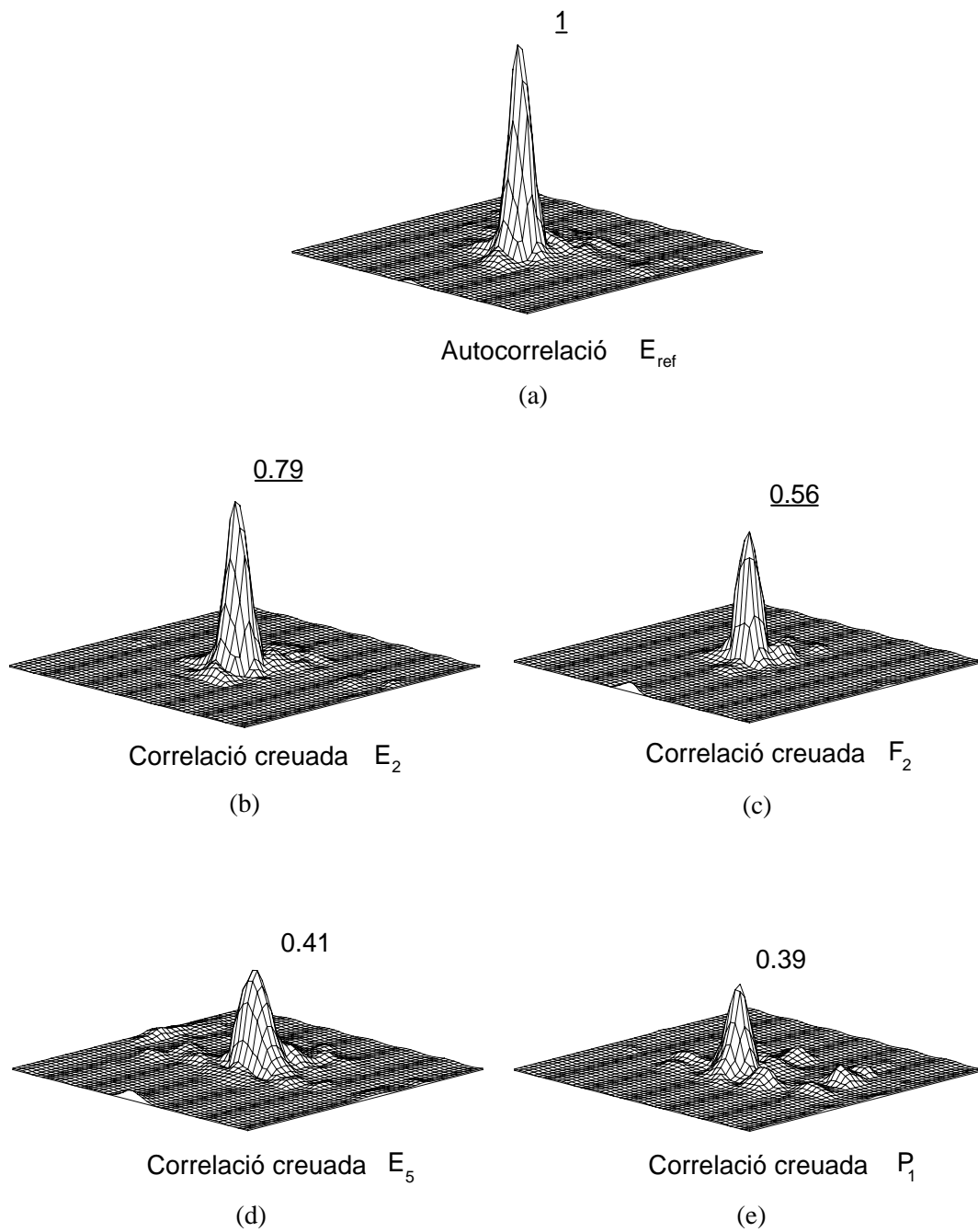
En la Figura 7.28 es mostren els pics de correlació (autocorrelació (a) i correlació creuada (b) – (e)) normalitzats a l'autocorrelació de la referència (a), obtinguts per a cada una de les lletres analitzades, quan la no-linealitat aplicada correspon als paràmetres  $(-0.8, -0.6)$ . Al costat dels pics es proporciona el seu valor màxim, subratllat en el cas en què aquest valor estigui inclòs dins la banda de reconeixement de l'objecte corresponent.

Les Figures 7.29 i 7.30 mostren una representació equivalent en els casos de les no-linealitats  $(L, M) = (0.6, 0.45)$  i  $(L, M) = (1.6, 1.2)$ .

Els resultats numèrics obtinguts en la realització experimental de la DNC, es mostren en la Taula 7.1. Per a cada parell  $(L, M)$  es donen els valors de  $cc/ac$  corresponents als màxims de les correlacions creuades per cada lletra normalitzats a l'autocorrelació, també obtinguda experimentalment. Entre parèntesi figuren els resultats d' $cc/ac$  obtinguts per simulació numèrica. Els valors numèrics que superen el llindar de reconeixement fixat al 50% de l'autocorrelació ( $u = 0.5$ ) han estat subratllats en la Taula 7.1. Els valors experimentals que compleixen la condició de superar el llindar indiquen que el caràcter analitzat en l'escena s'identifica amb el patró de referència. Per tant, el reconeixement és afirmatiu, figurant un “Si” en la casella corresponent. La resta de valors, corresponen a màxims dels pics de correlació creuada que difereixen més d'un 50% del màxim del pic d'autocorrelació, i per tant, són rebutjats en el procés de reconeixement. En aquests casos figura un “No” en la casella corresponent.

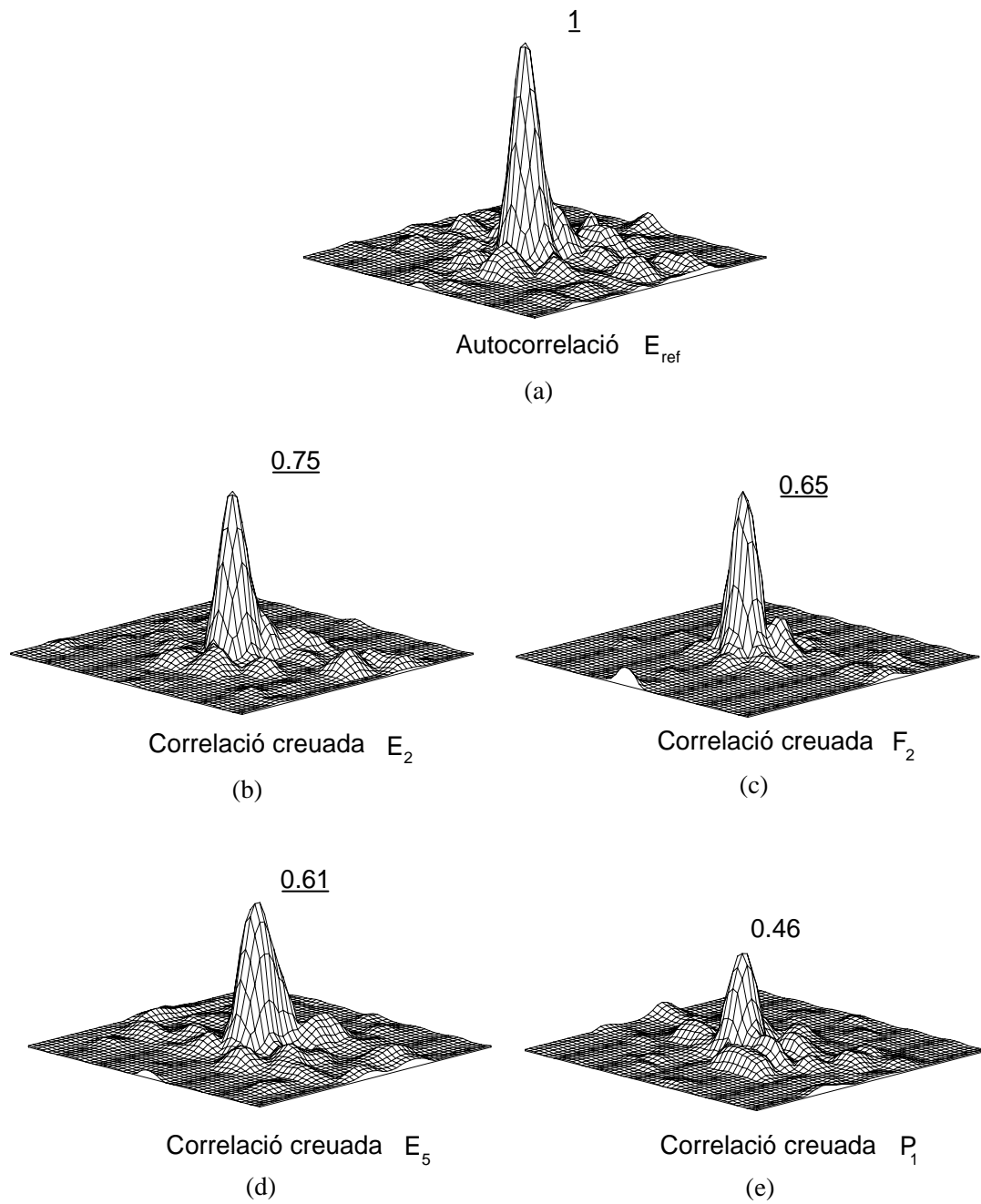


**Figura 7.28.** Pics d'autocorrelació (a) i de correlació creuada (b)–(e) normalitzats, obtinguts experimentalment pels paràmetres  $L = -0.8$  i  $M = -0.6$ , per a les cinc lletres que constitueixen el conjunt d'escenes test (Figura 7.21).



**Figura 7.29.** Pics d'autocorrelació i de correlació creuada normalitzats, obtinguts experimentalment pels paràmetres  $L = 0.6$  i  $M = 0.45$ , per a les cinc lletres que constitueixen el conjunt d'escenes test (Figura 7.21).





**Figura 7.30.** Pics d'autocorrelació i de correlació creuada normalitzats, obtinguts experimentalment pels paràmetres  $L = 1.6$  i  $M = 1.2$ , per a les cinc lletres que constitueixen el conjunt d'escenes test (Figura 7.21).

**Taula 7.1.** Resultats de reconeixement de la DNC experimental per a diferents valors ( $L$ ,  $M$ ) a l'analitzar les escenes test de la Figura 7.21. Els valors numèrics corresponen al paràmetre  $cc/ac$ . Entre parèntesi figuren els resultats de  $cc/ac$  obtinguts per simulació. Els valors que figuren subratllats, superen el llindar de correlació  $u = 0.5$ .

<b>Caràcter</b>	<b>E<sub>ref</sub></b>	<b>E<sub>2</sub></b>	<b>F<sub>2</sub></b>	<b>E<sub>5</sub></b>	<b>P<sub>1</sub></b>
<b>(-0.8, -0.6)</b> <b>Reconeixement</b>	<u>1</u> (1) Si	0.47 (0.24) No	0.41 (0.17) No	0.23 (0.04) No	0.27 (0.05) No
<b>(0.2, 0.15)</b> <b>Reconeixement</b>	<u>1</u> (1) Si	<u>0.67</u> (0.56) Si	0.49 (0.40) No	0.34 (0.21) No	0.30 (0.19) No
<b>(0.6, 0.45)</b> <b>Reconeixement</b>	<u>1</u> (1) Si	<u>0.79</u> (0.67) Si	<u>0.56</u> (0.52) Si	0.41 (0.37) No	0.39 (0.27) No
<b>(1.2, 0.9)</b> <b>Reconeixement</b>	<u>1</u> (1) Si	<u>0.82</u> (0.77) Si	<u>0.62</u> (0.67) Si	<u>0.54</u> (0.70) Si	0.45 (0.46) No
<b>(1.6, 1.2)</b> <b>Reconeixement</b>	<u>1</u> (1) Si	<u>0.75</u> (0.75) Si	<u>0.65</u> (0.70) Si	<u>0.61</u> (0.87) Si	0.46 (0.59) No

En la Taula 7.1 podem observar que s'ha aconseguit la variació en el grau de tolerància a canvis en la forma dels objectes analitzats del sistema DNC. Els resultats experimentals obtinguts s'adiuen en gran mesura als predits per simulació, si bé no exactament amb el resultat numèric obtingut, però si en el comportament del sistema. Per valors negatius dels paràmetres  $L$  i  $M$  s'aconsegueix una màxima discriminació de la lletra E que es vol detectar, davant d'altres lletres E similars i d'altres caràcters diferents. A mesura que augmentem els valors dels paràmetres sobre la recta indicada, el grau de tolerància del sistema és major i varia progressivament, de manera que passen a ser reconeguts com a similars a la referència els caràcters amb coeficient de similitud menor, que abans eren diferenciats del patró. Podem dir, per tant, que en aquestes condicions disposem d'un sistema de reconeixement amb capacitat de discriminació intermèdia. Els valors més elevats considerats,

$(L, M) = (1.6, 1.2)$ , configuren el sistema DNC amb un grau de tolerància alt, on es reconeix el major nombre de caràcters.

Veiem doncs, a partir de la valoració qualitativa dels resultats obtinguts, que la realització experimental de la DNC concorda en bona mesura amb els resultats previstos en la simulació numèrica del procés de reconeixement.

Si comparem els valors numèrics dels màxims de correlació normalitzats al pic d'autocorrelació que s'obtenen experimentalment i per simulació, observem petites discrepàncies i, fins i tot, en el cas en què el màxim de correlació sigui proper al valor llindar establert, provoca que aquestes diferències numèriques signifiquin un canvi en el reconeixement final d'un caràcter (veure lletra  $P_1$  en la no-linealitat  $L = 1.6$ , i  $M = 1.2$ ). Les discrepàncies que acabem de comentar poden ser explicades pel retall de freqüències realitzat pel dispositiu experimental com a conseqüència de les dimensions finites dels components òptics utilitzats. Recordem que en l'apartat 6.2 del capítol anterior, vam comprovar com la introducció de les limitacions experimentals introduïen una major separació entre les bandes de reconeixement dels diferents caràcters. Aquest fet es podia explicar a partir de la pèrdua d'informació en les freqüències menys energètiques de l'espectre (coincidents amb les freqüències més altes) com a conseqüència de la quantització introduïda pel sistema de captació. En el cas que considerem ara, el retall de freqüències es produeix en més mesura per l'efecte de la pupil·la del sistema òptic. En aquest sentit, seria possible millorar els resultats amb el disseny d'un correlador més compacte.

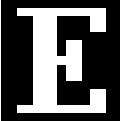




Com a comentari menor, observem en les Figures 7.28-7.30, que és en el cas de la no-linealitat intermèdia (és a dir, els valors d' $L$  i  $M$  positius i inferiors a la unitat), on s'obtenen uns pics de correlació de millor qualitat, en quant a l'amplada i el nivell de soroll. Aquest comportament es pot comparar amb els pics obtinguts en les altres dues no-linealitats considerades,  $(L, M) = (-0.8, -0.6)$  i  $(L, M) = (1.6, 1.2)$ , on sobretot destaca l'augment del soroll en el pla de correlació.

Concloem doncs, que amb el dispositiu experimental descrit en el punt 7.2 d'aquest mateix capítol, hem pogut obtenir un correlador no-lineal dual, sensible a variacions en les no-linealitats aplicades a l'escena i a la referència, i que permet un control en la capacitat de discriminació del sistema a canvis de forma en els objectes.

### 7.6.2 Variació de la capacitat de discriminació del correlador no-lineal dual davant variacions d'intensitat (nivell de gris)

Volem portar a terme un procés de correlació òptico-digital en el cas on analitzem les escenes contenint objectes d'igual forma amb nivell de gris uniforme diferent, el que és equivalent a canvis en la transmitància dels mateixos un cop reproduïts en l'SLM. L'objectiu és també el d'aconseguir un sistema de reconeixement amb capacitat de discriminació variable a canvis en el nivell de gris.

El conjunt d'escenes segmentades test utilitzades en aquesta realització experimental de la DNC, està formada per cinc caràcters dels ja emprats en el test 2 (Figura 6.6) del capítol anterior. Representem en la Figura 7.31 les lletres seleccionades.

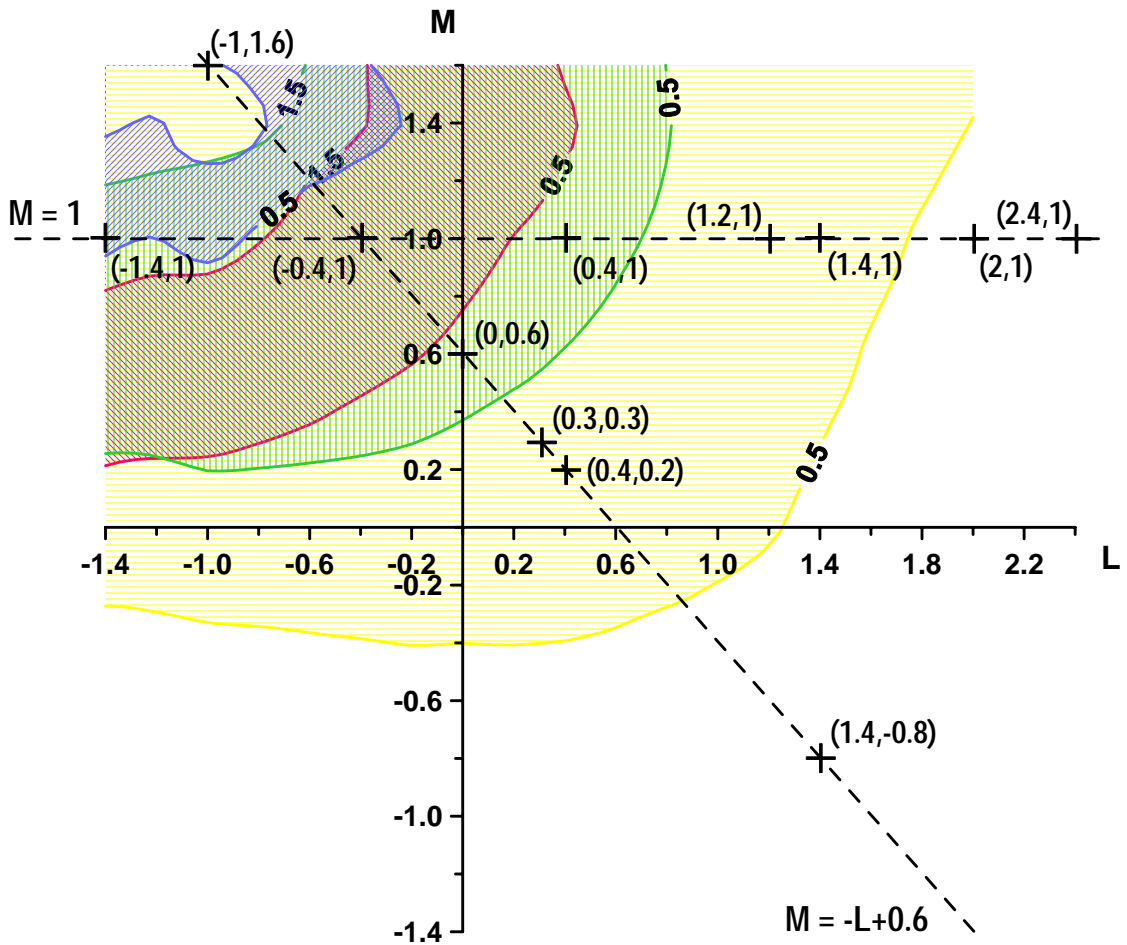
Etiqueta	$E_{ref}$	$E_{200}$	$E_{150}$	$E_{100}$	$E_{50}$
Objecte (lletra)					
Nivell de gris	255	200	150	100	50
Producte escalar $cs = \frac{\bar{s} \cdot \bar{r}}{ \bar{r} ^2}$	1	0.78	0.59	0.39	0.20

**Figura 7.31.** Conjunt d'escenes test per a la realització experimental de la DNC amb variació de la capacitat discriminativa davant de canvis en el nivell de gris uniforme dels objectes.

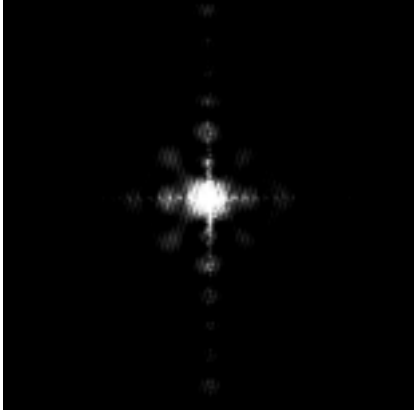
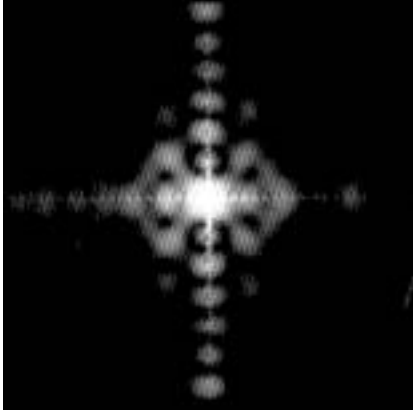
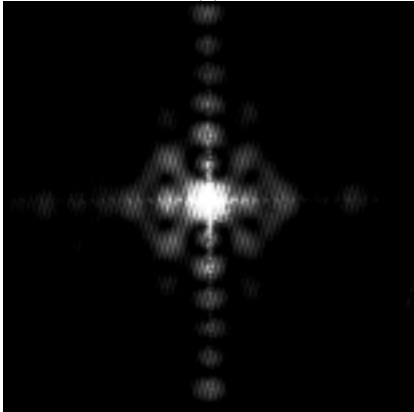
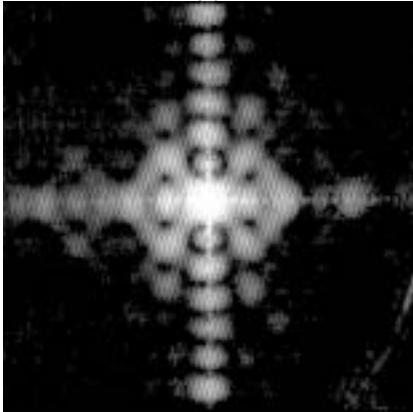
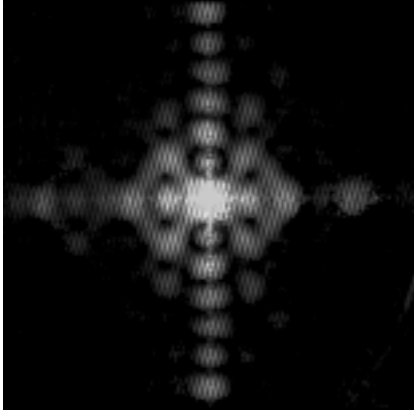
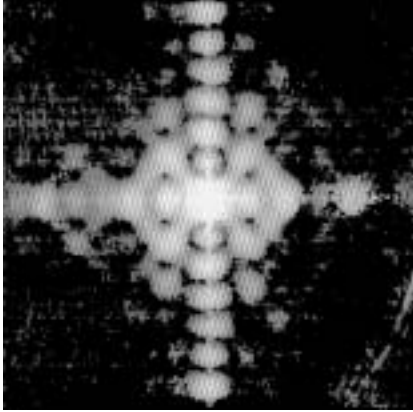
Representem en la Figura 7.32, molt similar a l'anterior Figura 6.10, el gràfic en el pla  $LM$  de les corbes lliures de les bandes de reconeixement de cada una de les lletres incloses en el test, obtingudes per simulació numèrica en la secció 6.3. Recordem que, en aquesta representació es marquen, per a cada lletra, el límit superior de la banda de reconeixement amb una etiqueta "1.5" i el límit inferior amb "0.5". Donat un caràcter, les dues corbes esmentades delimiten la banda de reconeixement dins la qual el caràcter és detectat com a similar al patró buscat, i fora de la banda, el caràcter es considera prou diferent al patró com per ser rebutjat en el procés d'identificació. En la mateixa representació (Figura 7.32), assenyallem també els punts  $(L, M)$  en els quals portarem a terme la realització experimental de la DNC. En aquesta ocasió, prenem diversos punts pertanyents a dos eixos diferents del pla  $LM$ . Concretament, considerem les rectes  $M = -L + 0.6$  i  $M = 1$ , per estudiar en elles la variació progressiva de la capacitat discriminativa del sistema enfront de variacions d'intensitat dels objectes. S'han escollit aquests dos eixos d'acord amb els resultats obtinguts per simulació. En ambdós eixos les proves de simulació prediuen una variació del grau de discriminació des d'un reconeixement exclusiu del patró pres com a referència, fins al reconeixement de la majoria de caràcters del test, indicant així una situació on la tolerància del sistema a canvis en el nivell de gris és prou elevada. Les diferències que les proves de simulació prediuen entre els dos eixos les podem trobar en què el canvi en la capacitat discriminativa és més lent per la recta  $M = 1$ , i que també en aquest eix es compleix que les no-linealitats aplicades mitjançant la DNC es realitzen només sobre la informació de l'escena, mentre es manté la informació de la referència com en el cas lineal del JTC.

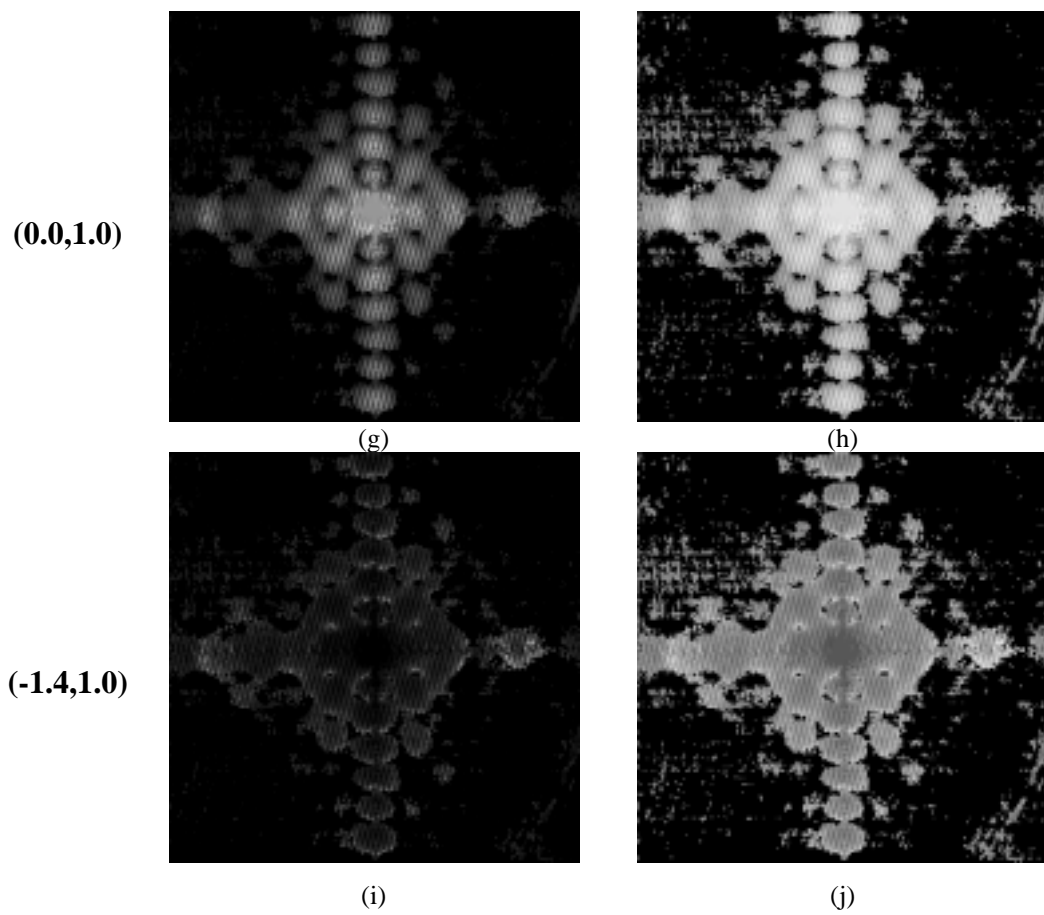
La Figura 7.33 mostra les distribucions MJPS obtingudes experimentalment per a diferents parells  $(L, M)$  sobre la recta  $M = 1$ , considerant el caràcter  $E_{150}$  com a escena i  $E_{ref}$  com a referència. En aquesta representació gràfica s'inclouen tant representacions lineals dels MJPS, com les logarítmiques. En la Figura 7.34 es mostren les funcions MJPS que s'obtenen per simulació numèrica de la DNC pels mateixos valors d' $L$  i  $M$  que en la Figura 7.33, i inclouent els condicionaments experimentals. Les distribucions presentades en la Figura 7.33 i les corresponents a

les altres lletres incloses en el test, són representades en el modulador durant el segon pas de la DNC. A partir d'aquesta informació, realitzem una segona TF òptica que ens proporciona el pla de sortida del processador, on s'inclouen els pics de correlació.



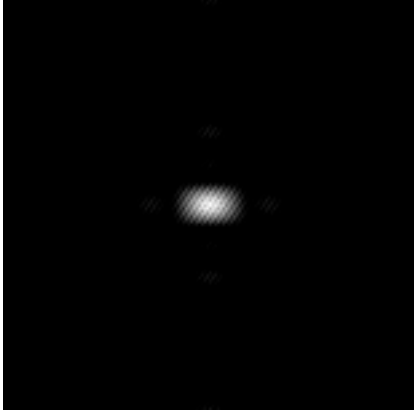
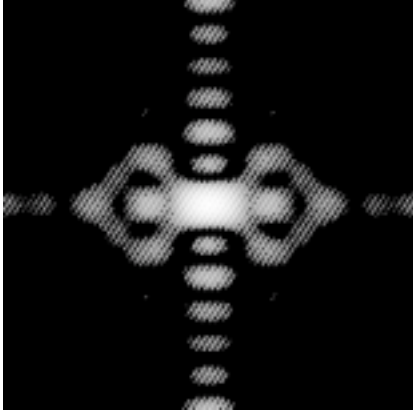
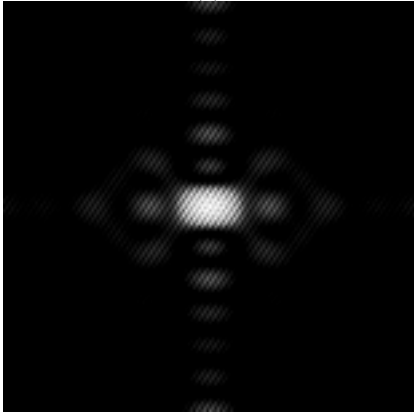
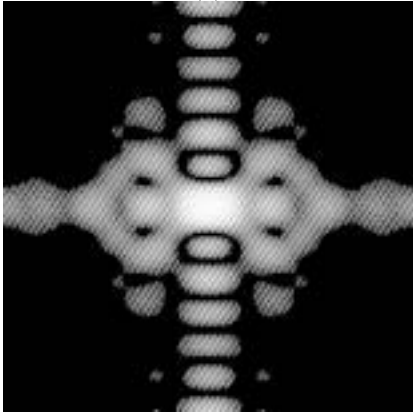
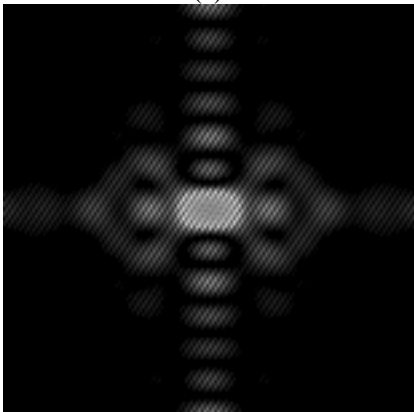
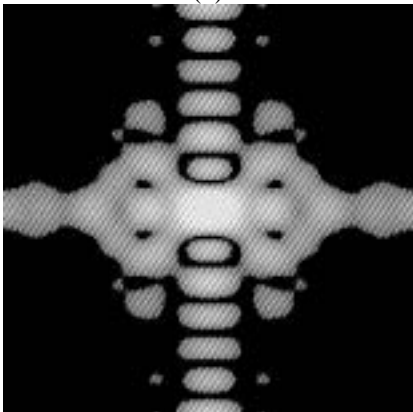
**Figura 7.32.** Pla  $LM$  amb les corbes de reconeixement i/o discriminació pels caràcters test seleccionats (Figura 7.31). També és representen les línies  $M = -L + 0.6$  i  $M = 1$  sobre les quals s'han escollit els punts  $(L, M)$  per a la realització experimental de la DNC.

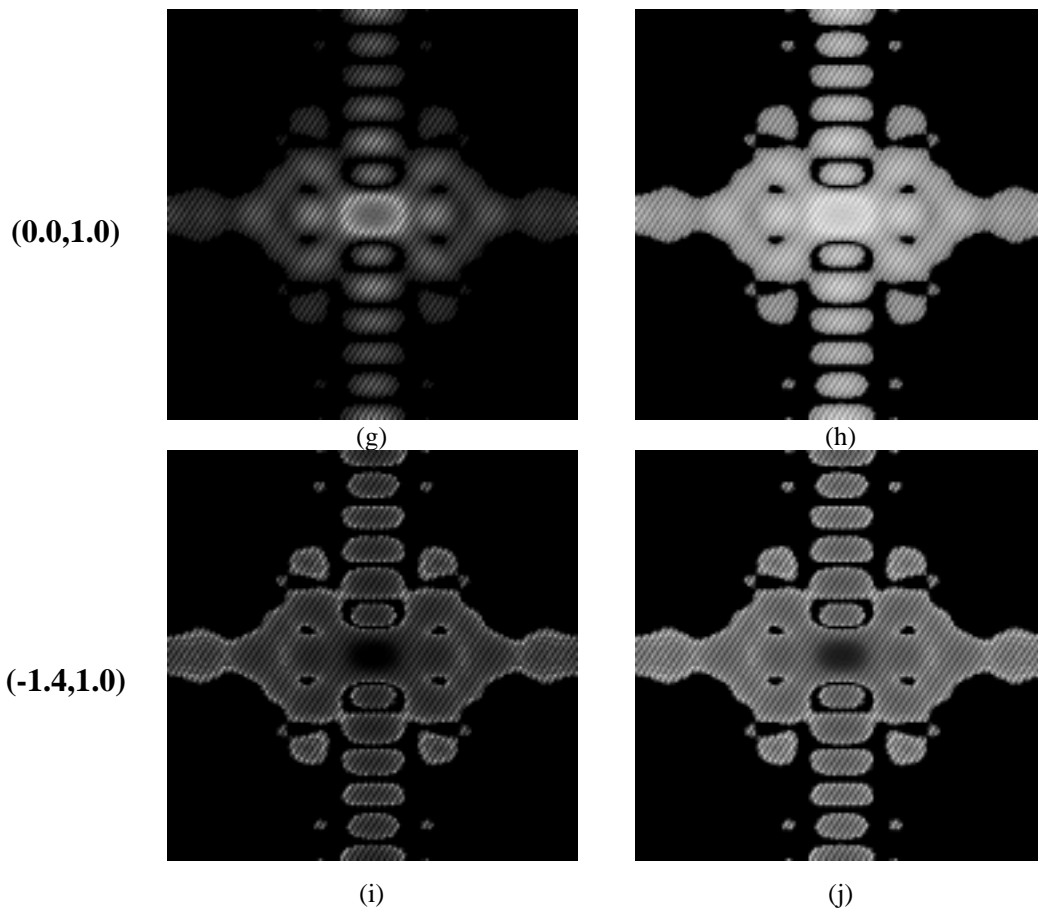
<b>(L, M)</b>	<b>Representació lineal</b>	<b>Representació logarítmica</b>
<b>(2.4,1.0)</b>		
<b>(1.2,1.0)</b>		
<b>(0.4,1.0)</b>		



**Figura 7.33.** MJPS obtinguts experimentalment per a la lletra  $E_{150}$  com a escena i  $E_{ref}$  com a referència i diferents no-linealitats. A la columna de l'esquerra la representació de les funcions és lineal i, a la dreta, logarítmica.



(L, M)	Representació lineal	Representació logarítmica
(2.4,1.0)	 (a)	 (b)
(1.2,1.0)	 (c)	 (d)
(0.4,1.0)	 (e)	 (f)



**Figura 7.34.** MJPS obtinguts per simulació per a la lletra  $E_{150}$  com a escena i  $E_{ref}$  com a referència i diferents no-linealitats. A la columna de l'esquerra la representació de les funcions és lineal i, a la dreta, logarítmica.

De les Figures 7.33 i 7.34, observem una bona concordància entre les distribucions MJPS obtingudes en el procés de DNC simulat, i les obtingudes experimentalment. Destacar novament, l'encongment en la direcció horitzontal sofert per les distribucions enregistrades experimentalment, donat la asimetria dels píxels del modulador i de la càmera, tant pel seu tamany com per la distància entre píxels contigus. En les distribucions obtingudes òpticament apareix, també, el patró de difracció de l'obertura rectangular que limita la pantalla del modulador i en alguns casos la influència del soroll periòdic (bandes horitzontals) introduït per la càmera.

Veiem, que mantenint constant el paràmetre  $M$  que actua sobre la informació provenint de l'objecte a detectar, el fet de modificar el paràmetre  $L$  corresponent a la

no-linealitat aplicada a l'escena, prenent valors alts i positius, i disminuint el seu valor cap a valors positius baixos, i fins i tot, molt negatius, dins la regió estudiada, implica en el sentit descrit, un augment inicial de la regió de suport on s'aplica la no-linealitat. En la representació logarítmica, però, veiem que aquesta assoleix una extensió màxima. El progrés del paràmetre  $L$ , provoca un augment en el pes de les altes freqüències, i en el valor més petit considerat,  $L = -1.4$ , es produeix una disminució del contrast de les franges apreciable.

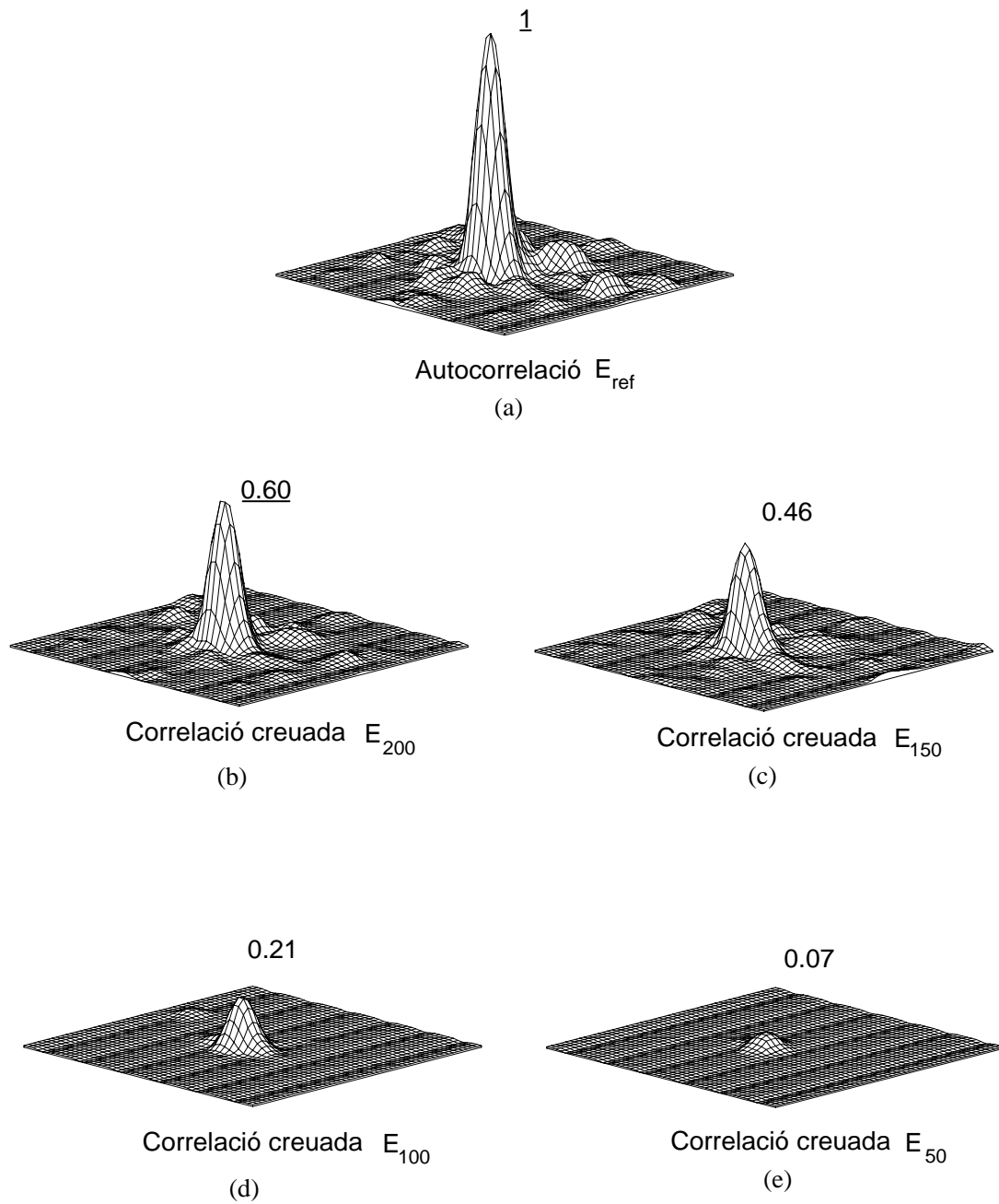
Alguns resultats òptics de correlació pels valors  $(L, M) = \{(2.4, 1.0), (1.2, 1.0), (-1.4, 1.0)\}$  es mostren gràficament en les Figures 7.35, 7.36 i 7.37, respectivament. Els resultats experimentals per a totes les lletres del test, i les diferents no-linealitats considerades, es detallen en les Taules 7.2 i 7.3. La primera d'elles (Taula 7.2), correspon a les no-linealitats contingudes en la recta  $M = -L + 0.6$ , i la Taula 7.3 mostra els resultats dels parells  $(L, M)$  de la recta  $M = 1$ . En totes les representacions s'ha indicat el valor màxim dels pics de correlació creuada normalitzats al valor de l'autocorrelació. S'assenyalen amb el subratllat, aquells pics que estan inclosos dins la banda de reconeixement. En les Taules 7.2 i 7.3, s'inclouen els valors numèrics obtinguts en l'anàlisi per simulació, i s'indica amb un "Si" quines lletres superen el procés de reconeixement en cada una de les no-linealitats aplicades. Els caràcters que són rebutjats, i per tant discriminats del patró, estan indicats mitjançant un "No".

En la Taula 7.2, els resultats de correlació experimentals per a la línia  $M = -L + 0.6$  mostren un capacitat de discriminació variable. El comportament general coincideix amb el predit en el capítol anterior. És a dir, en el quadrant inferior dret de la regió de l'espai  $LM$  considerat, s'aconsegueix una discriminació màxima, i en el costat oposat de la recta, aconseguim el grau de tolerància més gran a canvis de nivell de gris. Veiem que s'obté una discriminació variable però en un rang menor a l'esperat per simulació. En general, els valors dels pics de correlació normalitzats s'acosten més a la unitat en el cas d'obtenció de la correlació òptica, que en el procés simulat. D'aquesta manera, en la zona de màxima discriminació, el sistema no arriba a diferenciar el caràcter  $E_{200}$  de la referència amb un nivell de gris de 255, tot i obtenir un valor del màxim de correlació creuada proper al llindar de la banda de

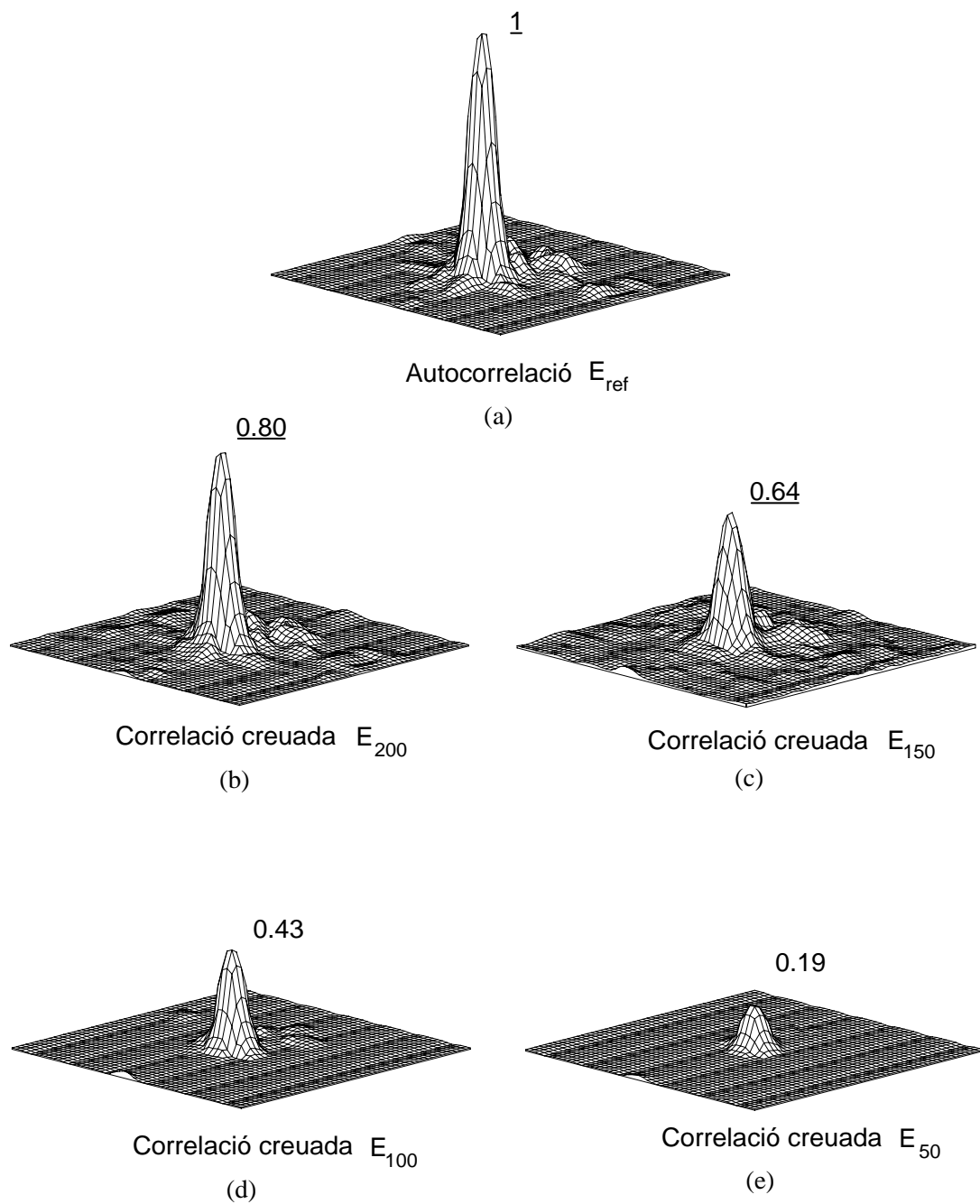
reconeixement. De  $(L, M) = (1.4, -0.8)$  a  $(L, M) = (0.4, 0.2)$  la tolerància a canvis d'intensitat augmenta, passant a reconèixer les lletres  $E_{150}$  i  $E_{100}$  de manera simultània. Finalment, tots els caràcters del test s'identifiquen al patró per les no-linealitats  $(L, M) = (-0.4, 1.0)$  i  $(L, M) = (-1.4, 1.6)$ . En aquest darrer cas, però, el sistema hauria de discriminar la majoria de caràcters del test, excepte l' $E_{200}$  que és el més energètic fora de la referència.

En la Taula 7.3 hem recollit els resultats experimentals quan només es modifica la no-linealitat aplicada a l'escena ( $L$  variable) mentre que la informació de l'objecte a detectar, en canvi, roman sense modificar com en el cas del JTC lineal (recta  $M = 1$ ). Ho hem fet així, per a què les diferències obtingudes en els valors de la correlació siguin deguts exclusivament a l'efecte dels diferents operadors aplicats sobre l'escena en cada lletra del test (diferents valors d' $L$ ). De totes maneres, els resultats obtinguts en aquest segon cas, no mostren ni millores ni empitjoraments significatius respecte als primers que s'han presentat al llarg de l'eix  $M = -L + 0.6$ .

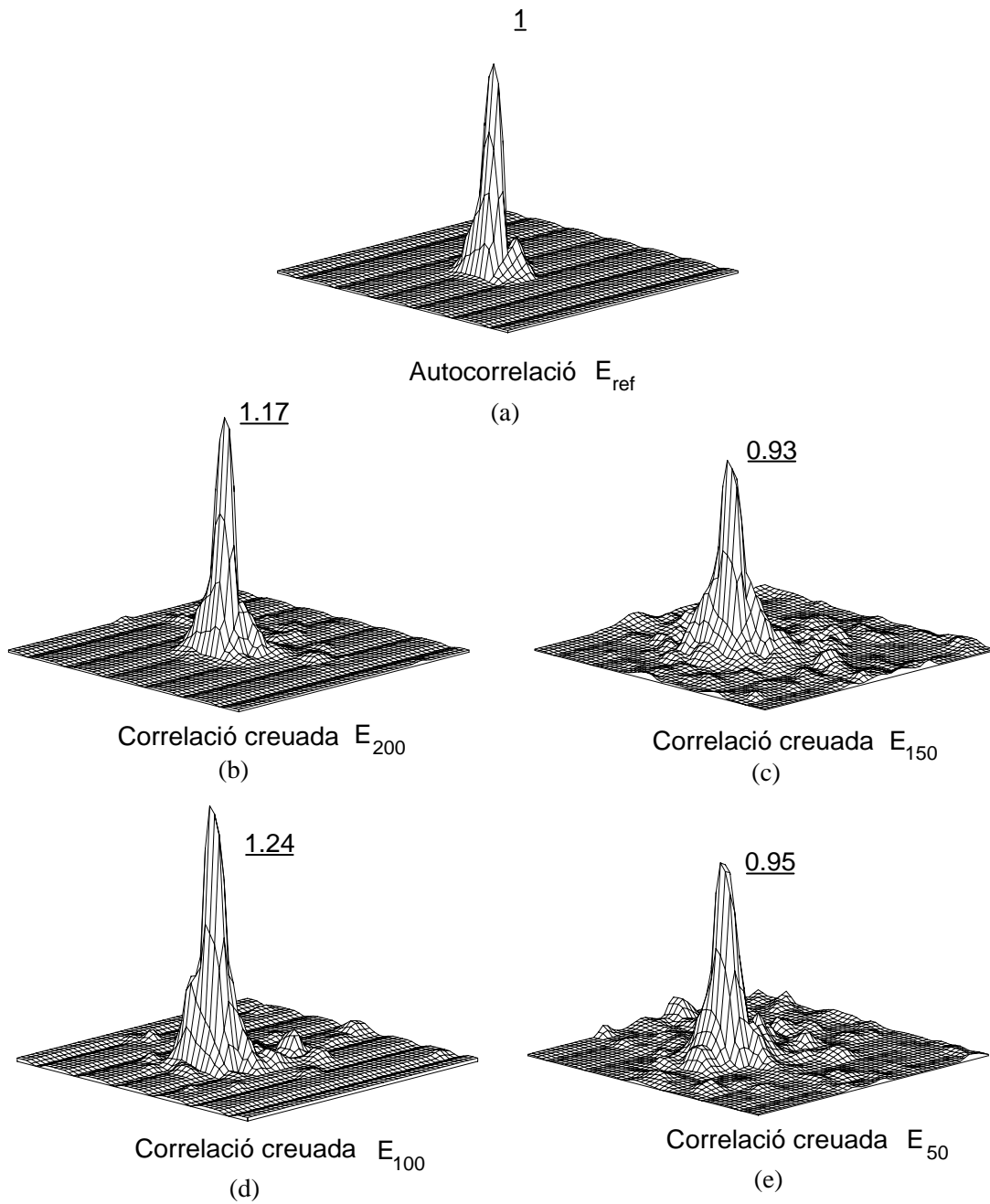
Tractem d'explicar la desavenència parcial dels resultats experimentals amb els obtinguts prèviament per simulació a partir de dues consideracions. Un efecte que no s'ha tingut en compte en la simulació del procés de reconeixement per correlació òptica, és la diferència de fase introduïda pel modulador al reproduir les escenes i distribucions d'intensitat amb un rang de nivells de gris. Tot i escollir els paràmetres més propers al règim de treball només en amplitud del modulador, sabem que sempre hi ha associada una modulació de fase, que no podem evitar. En el primer exemple, on s'han analitzat el comportament del sistema davant de canvis de forma en escenes monobjectes equienergètiques, la modulació de fase, no intervé de manera significativa en els resultats, però si analitzem les variacions en nivell de gris, és quan més es potencia aquest efecte de la fase acoblada. Un altre efecte que no s'ha tingut en compte en la simulació és el de l'estructura quadriculada amb àmplies zones inactives (al voltant del 44% [210]) de l'SLM. Podria ser que la combinació d'ambdós efectes no permeti una millor concordància dels resultats experimentals amb els de simulació.



**Figura 7.35.** Pics d'autocorrelació i de correlació creuada normalitzats, obtinguts experimentalment pels paràmetres  $L = 2.4$  i  $M = 1.0$ , per a les cinc lletres que constitueixen el conjunt d'escenes test (Figura 7.31).



**Figura 7.36.** Pics d'autocorrelació i de correlació creuada normalitzats, obtinguts experimentalment pels paràmetres  $L = 1.2$  i  $M = 1.0$ , per a les cinc lletres que constitueixen el conjunt d'escenes test (Figura 7.31).



**Figura 7.37.** Pics d'autocorrelació i de correlació creuada normalitzats, obtinguts experimentalment pels paràmetres  $L = -1.4$  i  $M = 1.0$ , per a les cinc lletres que constitueixen el conjunt d'escenes test (Figura 7.31).

**Taula 7.2.** Resultats de reconeixement de la DNC experimental per a diferents valors ( $L$ ,  $M$ ) sobre la recta  $M = -L + 0.6$ , a l'analitzar les escenes test de la Figura 7.31. Els valors numèrics corresponen al paràmetre  $cc/ac$ . Entre parèntesi figuren els resultats de  $cc/ac$  obtinguts per simulació. Els valors que figuren subratllats, superen el llindar de correlació  $u = 0.5$ .

Caràcter	$E_{ref}$	$E_{200}$	$E_{150}$	$E_{100}$	$E_{50}$
<b>(1.4, -0.8)</b> <b>Reconeixement</b>	<u>1</u> (1) Si	<u>0.53</u> (0.38) Si	0.32 (0.12) No	0.19 (0.02) No	0.05 (0.0) No
<b>(0.4, 0.2)</b> <b>Reconeixement</b>	<u>1</u> (1) Si	<u>0.78</u> (0.64) Si	<u>0.64</u> (0.38) Si	<u>0.55</u> (0.18) Si	0.29 (0.01) No
<b>(0.3, 0.3)</b> <b>Reconeixement</b>	<u>1</u> (1) Si	<u>0.89</u> (0.67) Si	<u>0.67</u> (0.43) Si	<u>0.60</u> (0.23) Si	0.31 (0.02) No
<b>(0.0, 0.6)</b> <b>Reconeixement</b>	<u>1</u> (1) Si	<u>0.90</u> (0.77) Si	<u>0.77</u> (0.60) Si	<u>0.77</u> (0.42) Si	0.45 (0.05) No
<b>(-0.4, 1.0)</b> <b>Reconeixement</b>	<u>1</u> (1) Si	<u>1.01</u> (0.96) Si	<u>0.97</u> (0.97) Si	<u>1.03</u> (1.04) <u>Si</u>	<u>0.75</u> (0.22) Si
<b>(-1.0, 1.6)</b> <b>Reconeixement</b>	<u>1</u> (1) Si	<u>1.26</u> (1.34) Si	<u>1.19</u> (2.0) Si	<u>1.40</u> (3.77) Si	<u>1.42</u> (10.0) Si



**Taula 7.3.** Resultats de reconeixement de la DNC experimental per a diferents valors  $(L, M)$  sobre la recta  $M = 1$ , a l'analitzar les escenes test de la Figura 7.31. Els valors numèrics corresponen al paràmetre  $cc/ac$ . Entre parèntesi figuren els resultats de  $cc/ac$  obtinguts per simulació. Els valors que figuren subratllats, superen el llindar de correlació  $u = 0.5$ .

Caràcter	$E_{ref}$	$E_{200}$	$E_{150}$	$E_{100}$	$E_{50}$
<b>(2.4, 1.0)</b> Reconeixement	<u>1</u> (1) Si	<u>0.60</u> (0.39) Si	0.46 (0.10) No	0.21 (0.01) No	0.07 (0.0) No
<b>(2.0, 1.0)</b> Reconeixement	<u>1</u> (1) Si	<u>0.69</u> (0.42) Si	<u>0.52</u> (0.13) Si	0.28 (0.02) No	0.09 (0.0) No
<b>(1.4, 1.0)</b> Reconeixement	<u>1</u> (1) Si	<u>0.78</u> (0.54) Si	<u>0.60</u> (0.25) Si	0.34 (0.07) No	0.17 (0.01) No
<b>(1.2, 1.0)</b> Reconeixement	<u>1</u> (1) Si	<u>0.80</u> (0.59) Si	<u>0.64</u> (0.31) Si	0.43 (0.12) No	0.19 (0.01) No
<b>(0.4, 1.0)</b> Reconeixement	<u>1</u> (1) Si	<u>0.93</u> (0.80) Si	<u>0.77</u> (0.60) Si	<u>0.71</u> (0.39) Si	0.40 (0.05) No
<b>(0.0, 1.0)</b> Reconeixement	<u>1</u> (1) Si	<u>0.96</u> (0.88) Si	<u>0.84</u> (0.80) Si	<u>0.86</u> (0.64) Si	<u>0.58</u> (0.11) Si
<b>(-0.4, 1.0)</b> Reconeixement	<u>1</u> (1) Si	<u>1.01</u> (0.96) Si	<u>0.97</u> (0.97) Si	<u>1.03</u> (1.04) <u>Si</u>	<u>0.75</u> (0.22) Si
<b>(-1.4, 1.0)</b> Reconeixement	<u>1</u> (1) Si	<u>1.16</u> (1.05) Si	<u>0.93</u> (1.35) Si	<u>1.24</u> (2.52) Si	<u>0.95</u> (1.24) Si

## **Capítol 8**

### **Aplicació de la DNC al reconeixement d'objectes en color**

Els resultats obtinguts en els capítols anteriors ens permetran ara aplicar el model de la DNC al reconeixement d'objectes en color. Recordem-los breument. En el capítol 6, hem aplicat la DNC, per una banda, al reconeixement d'objectes equienergètics amb variacions de forma, i per una altra banda, al reconeixement d'objectes d'igual forma però amb nivell de gris uniforme i diferent entre ells. D'aquest estudi, hem observat que és possible delimitar un zona en l'espai definit pels paràmetres  $L$  i  $M$  de les no-linealitats aplicades a l'escena i a la referència respectivament, en la qual la capacitat de discriminació del sistema davant les diferències comentades, varia de manera gradual des d'una discriminació absoluta de la referència de qualsevol altre caràcter lleugerament diferent, fins a tenir un nivell de tolerància a canvis força

elevat. En el capítol 7, per altra banda, hem demostrat la viabilitat de la proposta en el cas d'aplicar la DNC a escenes on els objectes presenten canvis de forma, aconseguint un dispositiu optoelectrònic sensible a les diferents no-linealitats aplicades i, com a conseqüència, amb un grau de discriminació variable de manera controlada. En quant a les variacions de nivell de gris, el dispositiu emprat en la realització experimental, ens ha permès també modificar la capacitat de discriminació del sistema. Les dificultats associades a aquestes darreres proves per igualar els resultats predits per simulació, semblen estar relacionades a les limitacions del modulador espacial de llum utilitzat, més que no pas, al mètode de la DNC. A partir dels resultats d'aquests dos darrers capítols, ens plantegem fer un pas més en el reconeixement d'objectes en color, si més no, a nivell de simulació, sabent que la seva realització experimental depèn en part de les prestacions que puguin donar els moduladors.

En aquesta ocasió, ens plantegem estudiar l'aplicació de la DNC en problemes de reconeixement on apareixen tant canvis en la forma com en la distribució espectral dels objectes. De fet, no només parlarem de nivells de gris dels l'objectes, sinó que considerarem un test d'objectes en color, amb els quals portarem a terme un procés de correlació multicanal, que implica, l'anàlisi en cada un dels canals cromàtics, d'un problema de reconeixement conjuntament per la forma i pel nivell de gris dels objectes. El processament global de les dades extretes de cada canal cromàtic, aplicant llindars de reconeixement i operacions lògiques, donarà lloc als resultats de reconeixement d'objectes en color.

En aquest capítol inicialment aplicarem una descomposició RGB a la informació cromàtica de l'escena a analitzar i de l'objecte a detectar, i en aquests tres canals realitzarem el procediment de correlació DNC multicanal (apartat 8.1). Centrarem, però, la nostra atenció en objectes acolorits amb tonalitats properes, pertanyents a una mateixa gamma, tals com els colors tinguts en compte en les escenes test del capítol 3. Ens plantegem, per tant, un problema de reconeixement en el qual la

capacitat de discriminació cromàtica del sistema, en alguns casos, ha de ser especialment elevada.

Posteriorment, en l'apartat 8.2, proposarem la combinació de la DNC amb la descomposició adaptada en  $n$  canals quasimonocromàtics, com a una millora del sistema de reconeixement, principalment adequada quan es requereix una bona discriminació cromàtica.

## 8.1 Variacions de forma i color entre els objectes a analitzar

En aquest estudi, volem determinar, si és possible, alguna regió de l'espai  $LM$  que permeti modificar de manera controlada la capacitat de discriminació del sistema a canvis en la forma i en el color dels objectes, bé sigui de forma independent per cada una de les característiques, bé sigui conjuntament. Els casos podran entremesclar-se i així ens plantejarem, per exemple, com obtenir un reconeixement d'objectes amb un nivell de discriminació alt davant de variacions de forma mantenint una certa tolerància a variacions de color i, viceversa, com obtenir tolerància davant de petites variacions de forma tenint una discriminació elevada davant de variacions de color [212,213].

Aquests plantejaments amb una exigència mixta en quant al tipus i grau de capacitat discriminativa apareixen en operacions pràctiques de reconeixement. Per exemple, molts objectes de color pateixen degradacions (descoloriment, enfosquiment) o corresponen a tintades diferents (amb una inadequada reproducció de color) que fa que apareguin de colors lleugerament diferents dins d'una mateixa gamma. Pot interessar llavors, que el sistema de reconeixement sigui insensible a aquestes possibles variacions de color (dins un rang) que considerarem poc rellevants. No obstant, si els objectes formen part d'un codi mixt de formes acolorides (bandes,

banderes, logotips, caràcters alfanumèrics, etc.) pot interessar, al mateix temps, una alta discriminació de la forma que permeti la lectura automàtica.

També és possible trobar una aplicació en el cas contrari, és a dir, on sigui necessària una bona discriminació al color, però calgui admetre certa tolerància a la forma. Aquest seria el cas, per exemple, de voler identificar peces deliberadament acolorides. Els colors amb els quals es pinten les peces poden ser de la mateixa gamma i molt uniformes, però l'àrea sobre la qual es pinta, o la forma de la peça pot tenir petites variacions.










En el cas de voler detectar degradacions en els objectes, per exemple degudes al pas del temps o a la manipulació i l'ús dels mateixos, caldrà disposar d'un sistema de reconeixement amb una elevada capacitat de discriminació a la forma, si apareixen petites deformacions o desgast en les peces analitzades, i al mateix temps, pot interessar mantenir el grau de discriminació elevat respecte al color, si els colors dels objectes s'han descolorit.

Per últim, cal comentar també alguns exemples on seria necessària una tolerància a petites variacions tant en la forma, com en el color dels objectes. En aquesta situació, podem considerar aplicacions a classificació de fruites on la forma entre les diferents unitats analitzades serà semblant però no igual, i l'estat de maduració de la fruita pot fer variar la seva tonalitat. També és un exemple englobat dins aquest cas, la lectura de codis de resistències elèctriques, on tant l'àrea pintada com el seu color pot variar de manera apreciable.

En els casos que acabem de plantejar, els mètodes més convencionals de reconeixement mitjançant filtratge adaptat, tenen una capacitat de discriminació massa rígida per poder adaptar-se a la necessitat d'admetre un cert valor de tolerància a petits canvis respecte l'objecte de referència. Per resoldre aquest inconvenient, també introduïm la DNC en el reconeixement de formes en color.

### 8.1.1 Descripció del test 3

Fem una selecció dels caràcters emprats en els apartats anteriors per tal de presentar un exemple prou representatiu de les possibilitats de l'aplicació de la DNC en el reconeixement d'objectes en color. La Figura 8.1 conté el test 3 amb 9 elements segmentats a estudi. En aquesta figura, s'inclou informació de la forma dels caràcters, així com del color uniforme amb el qual s'ha acolorit cada un d'ells. El nom atorgat a cada caràcter coincideix amb l'indicat en el test 1 (Figura 6.1) de l'anàlisi de la DNC davant de variacions de forma (apartat 6.2). Mantenim en la Figura 8.1 el coeficient de similitud, que ja es va presentar en la Figura 6.1, que es calcula mitjançant el producte escalar de l'escena i de la referència, quan ambdues eren equienergètiques. Aquest valor ens dóna una idea si més no de la semblança de la forma del caràcter analitzat amb el caràcter patró. A més d'aquesta informació, també s'indica el color del caràcter (R, rosa pàl·lid; L, lila; S, salmó; i TF, taronja fosc). Com a referència o objecte a reconèixer considerem la lletra  $E_{ref}$ , la forma de la qual coincideix amb la de la referència dels test anteriors (Figures 6.1 i 6.6) i que és de color rosa pàl·lid.

Etiqueta	$E_{ref}$	$E_1$	$E_{2A}$	$E_{2B}$	$F_1$	$F_2$	$E_3$	$E_5$	$R_1$
Caràcter									
Producte escalar $cs = \frac{s \cdot F}{ F ^2}$	1	0.95	0.92	0.92	0.89	0.86	0.85	0.84	0.74
Color	R	L	S	TF	L	R	R	R	R

**Figura 8.1.** Test 3: escena test formada per diferents caràcters en color. S'inclou informació del producte escalar normalitzat considerant formes equienergètiques, i la denominació dels colors emprats (R: rosa pàl·lid; L: lila; S: salmó i TF: taronja fosc).

Els colors dels objectes varien mantenint-se dins una gamma que, en aquest cas, és la magenta. Però, també s'ha insertat un objecte acolorit de taronja fosc, tonalitat pertanyent a la gamma taronja. Les corbes de reflectància pels colors rosa pàl·lid, salmó i lila, que aquí s'utilitzen i estan inclosos a la gamma magenta, es mostren en la Figura 3.4 del capítol 3. Pel color taronja fosc de la gamma taronja, la corba de reflectància s'ha mostrat prèviament en la Figura 3.5, també del capítol 3. En la Taula 8.1 es detallen els valors numèrics de les corresponents components cromàtiques proporcionades per la càmera en coordenades R, G, B. S'adjunten com a valors orientatius de la similitud entre els colors les diferències proporcionals dels valors mitjos obtinguts en cada canal RGB calculades a partir de les expressions,

$$\Delta R(\%) = \frac{R_s - R_r}{(R_s + R_r)/2} \times 100,$$

$$\Delta G(\%) = \frac{G_s - G_r}{(G_s + G_r)/2} \times 100, \quad (8.1)$$

$$\Delta B(\%) = \frac{B_s - B_r}{(B_s + B_r)/2} \times 100,$$

on els subíndexs  $s$  i  $r$  indiquen que la component cromàtica correspon a l'escena i a la referència, respectivament.

Podem establir dos grans grups dins el conjunt d'escenes test que hem presentat en la Figura 8.1. El primer grup on el color dels objectes és diferent al de la referència i que, a més, les lletres incloses tenen com a característica comú una gran semblança en forma amb l'objecte buscat. Aquest és el cas de les lletres  $E_1$ ,  $E_{2A}$ ,  $E_{2B}$ , i  $F_1$ , totes elles, en general, amb un coeficient de similitud per la forma més gran que 0.89. Per aquests caràcters, les variacions proporcionals en el nivell de gris de les components cromàtiques, respecte de la referència es situen entre l'1 % i el 191 % en el cas més extrem del caràcter  $E_{2B}$  de color taronja fosc, i entre un 8 % i un 47 % pels colors pertanyents a la mateixa gamma (Taula 8.1).

El segon grup de caràcters està format per les lletres  $F_2$ ,  $E_3$ ,  $E_5$  i  $R_1$ , que tenen el mateix color que la referència (rosa pàl·lid) i que, per tant, la seva semblança amb la referència segons aquesta característica és màxima. En canvi, la semblança en forma és en tots els casos inferior al grup de caràcters enumerats anteriorment (entre 0.86 i 0.74).

**Taula 8.1.** Valors de les components cromàtiques RGB corresponents a les reflectàncies mitges que proporciona una càmera en color, dels colors emprats en el test 3 (Figura 8.1). També s'indiquen les diferències proporcionals per als valors RGB dels colors dels objectes analitzats ( $R_s$ ,  $G_s$ ,  $B_s$ ) respecte als valors RGB del color de la referència ( $R_r$ ,  $G_r$ ,  $B_r$ ).

		Color			
		Rosa pàl·lid (R)	Lila (L)	Salmó (S)	Taronja fosc (TF)
<b>Components cromàtiques</b>	<b>R</b>	228	156	196	225
	<b>G</b>	181	112	132	83
	<b>B</b>	174	161	122	4
<b>Diferències proporcionals</b>	$\Delta R$ (%)	0	38	15	1
	$\Delta G$ (%)	0	47	31	74
	$\Delta B$ (%)	0	8	35	191

El nostre propòsit en aquest estudi de reconeixement, consisteix en determinar les condicions necessàries per aconseguir un sistema amb capacitat discriminativa variable i controlable segons el problema plantejat. Considerem quatre casos. Volem ser capaços de detectar les lletres en forma similar a la referència amb tolerància a les variacions que es puguin introduir en el color dins una mateixa gamma de to, com també d'obtenir el cas contrari, identificar com si fos la referència aquelles lletres que tinguin exactament el mateix color, però admetre una certa tolerància a petites variacions de forma. Variant doncs, el grau de tolerància al color i a la forma, podem



considerar, també, dues tasques de reconeixement més, aquella en què sigui necessari un alt grau de discriminació en ambdues característiques forma i color, i al contrari, portar a terme un procés de reconeixement amb tolerància a petits canvis en la forma i en el color dels objectes respecte a la referència.

No obstant, en els quatre casos plantejats, també demanem que el sistema de reconeixement segueixi tenint una capacitat de discriminació davant d'objectes molt diferents en forma i/o en color. Per aquest motiu, dins el conjunt d'escenes test, hem introduït alguns caràcters de control ( $E_{2B}$ ,  $F_1$ ,  $F_2$  i  $R_1$ ) que ens permetran avaluar el grau de tolerància a la forma i al color del sistema en cada cas.

### 8.1.2 Mètodes de filtratge convencionals (CMF, POF, IF, PPC)

A l'igual que s'ha procedit en els capítols anteriors, considerem primerament un sistema de reconeixement amb el mètode de filtratge adaptat a l'objecte a reconèixer i, concretament, prenem els tipus de filtratges més convencionals: CMF, POF, IF i PPC. Es realitza la simulació numèrica de la correlació en cada canal cromàtic per separat en condicions ideals de realització, és a dir, sense incloure restriccions experimentals. En la Taula 8.2 es donen els resultats obtinguts corresponents a la raó  $cc/ac$ , on  $cc$  és el màxim del pic de correlació creuada, i  $ac$  és el màxim del pic d'autocorrelació. En cada canal es considera un llindar del 50% per a decidir si l'objecte considerat s'identifica amb la referència o, pel contrari, és rebutjat. Aquest criteri s'aplica a cada un dels canals per separat, i per decidir si passa el procés de reconeixement, es realitza la operació lògica AND als resultats obtinguts en cada canal. D'aquesta manera, si un objecte obté els màxims de correlació normalitzats dins l'interval  $0.5 \leq cc/ac \leq 1.5$  (valors que apareixen subratllats en la Taula 8.2), en els tres canals cromàtics simultàniament, es considera similar al patró i passa el procés de reconeixement. En cas contrari, l'objecte és rebutjat.

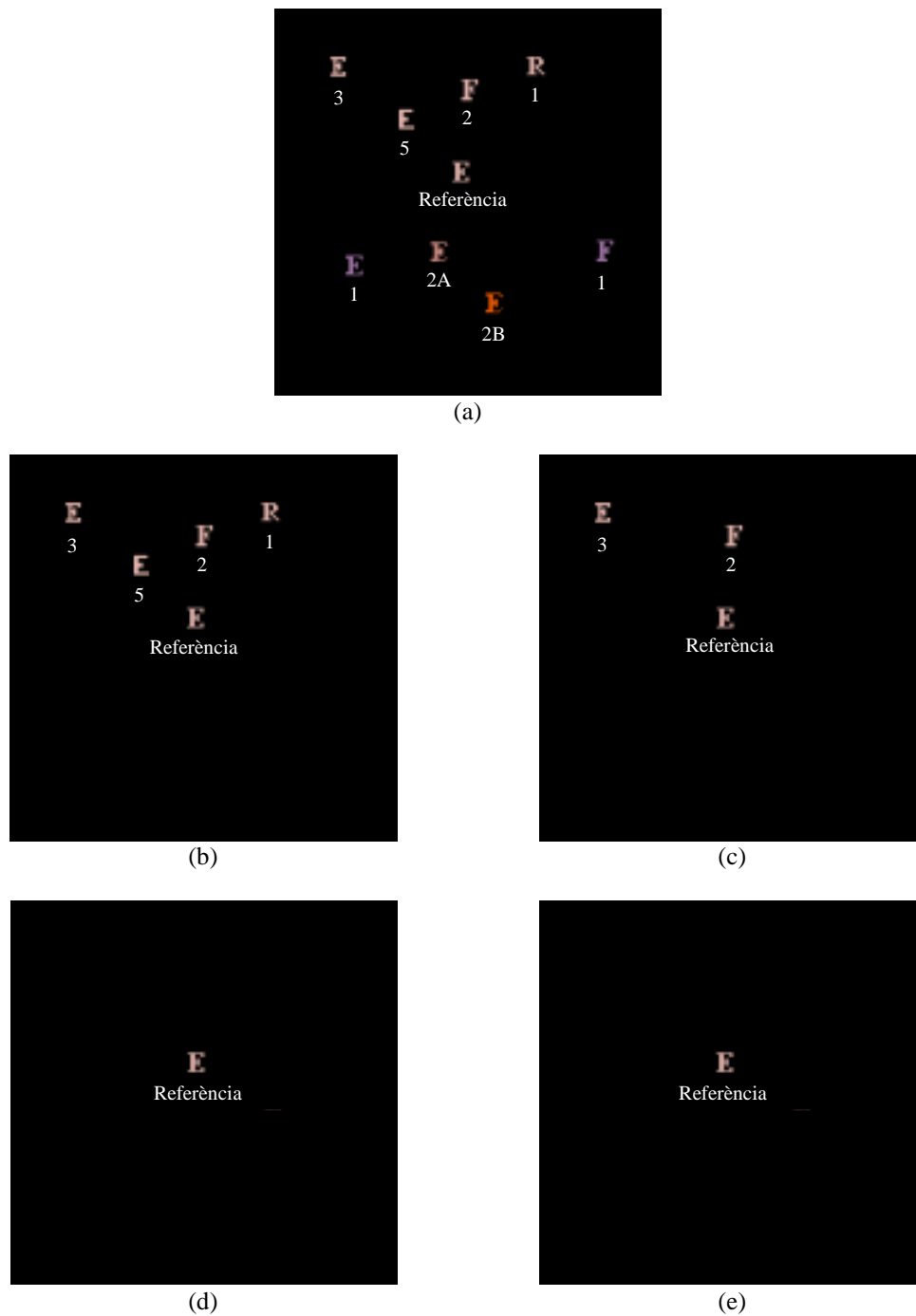
Una manera més immediata de visualització dels resultats obtinguts, és representant sobre la mateixa escena els caràcters que arriben a superar tot el procés de reconeixement. La Figura 8.2 els mostra en els quatre tipus de filtratge analitzats.

Observem que amb els quatre filtres emprats tenim diferents capacitats discriminatives del sistema, però rígides i que no admeten la introducció d'un major o menor grau de tolerància en relació a un aspecte concret dels objectes, com seria la forma o el color. Així, passem des de la capacitat de discriminació més alta quan només s'obté el reconeixement del patró en el cas del PPC i de l'IF, fins a tenir toleràncies a variacions de forma en el cas del POF, i encara nivells de tolerància a la forma més elevats en el cas del CMF. De totes maneres, en aquests darrers casos, el nivell de tolerància és exclusivament a variacions de forma i no de color.

Donat que el llindar s'estableix de manera totalment arbitrària, podríem pensar en modificar-lo per obtenir altres nivells de tolerància en el sistema. Per exemple, en els casos del CMF i del POF, és possible considerar un llindar més restrictiu augmentant-ne el seu valor. En el cas de l'IF i del PPC seria necessari relaxar el llindar, prenent valors més baixos, per augmentar la tolerància del sistema. Per escollir els valors adients del llindar  $u$  en cada cas, retornaríem als resultats inclosos en la Taula 8.2. Després de la pròpia lletra de referència, observem que la lletra  $F_2$  obté el valor de  $cc/ac$  més alt, o dels més alts, en els quatre mètodes de filtratge adaptat. Aquest fet és equivalent a establir que serà un dels objectes que es reconeixerà sempre si augmentem el llindar en el CMF o el POF, o bé si el disminuïm en l'IF o en PPC.

En conclusió, veiem que el procés de reconeixement per correlació òptica mitjançant filtratge adaptat, considerant mètodes convencionals, no aconsegueix resoldre el problema de reconeixement amb discriminació variable a la forma i/o al color que s'ha plantejat. Fixat el llindar en el 50%, els resultats són fixats pel mètode de filtratge utilitzat. En el cas del CMF i del POF, obtenim una certa tolerància a les variacions de forma i una elevada discriminació del color, que no podem canviar variant lleugerament el llindar. Pels casos de l'IF i del PPC, on ja s'obté una elevada





**Figura 8.2.** Reconeixement en condicions ideals dels caràcters del test 3 emprant mètodes de filtratge convencionals. (a) Caràcters del test 3. (b) Caràcters que es reconeixen a l'utilitzar el CMF, (c) ídem pel POF, (d) ídem per l'IF, (e) ídem pel PPC.

### 8.1.3 Aplicació de la DNC

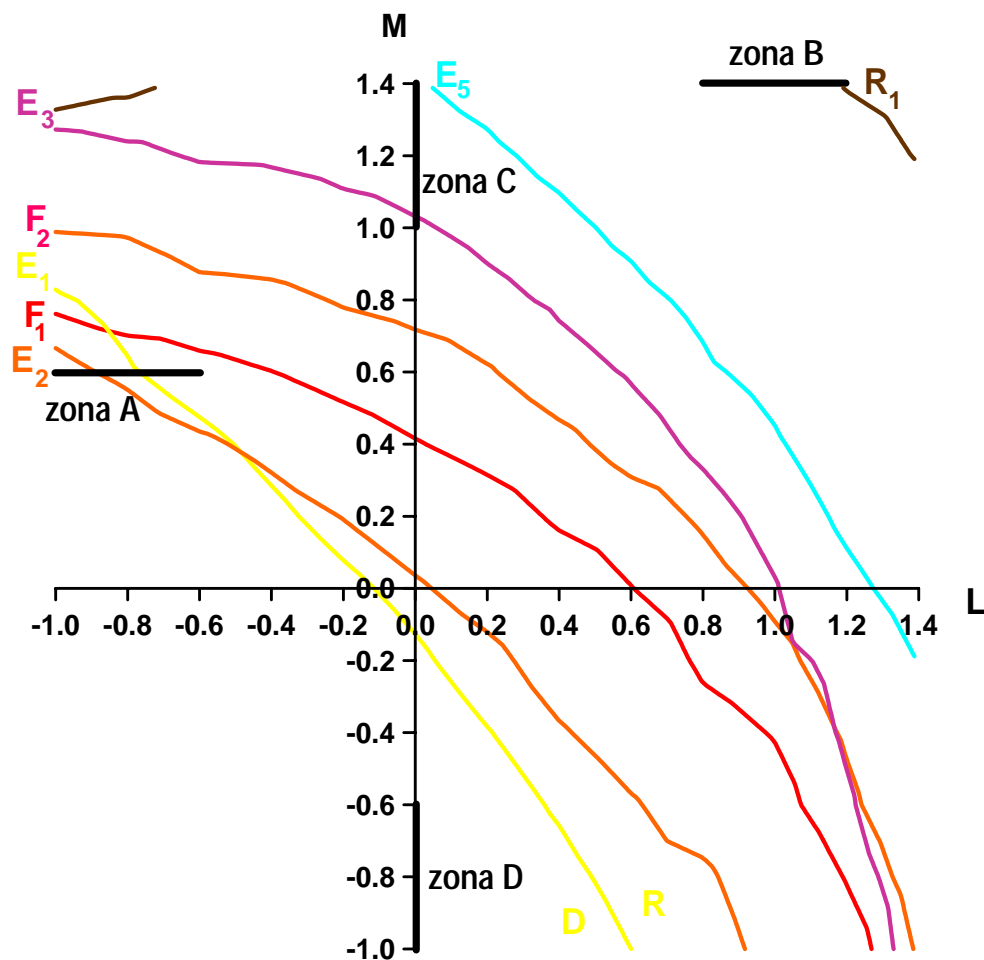
Estudiem les possibilitats de la DNC per dotar al sistema de reconeixement d'una capacitat de discriminació amb tolerància selectiva respecte a la forma i/o al color.

Representem part de la Figura 6.5 en la Figura 8.3, és a dir, les corbes llindar de reconeixement considerant només les formes dels caràcters inclosos en el conjunt d'escenes del test 3. També reproduïm la Figura 6.10 en la Figura 8.4 on s'estudia la DNC en front a variacions de nivell de gris quan les lletres són idèntiques en forma. Recordem que aquestes representacions corresponen a l'aplicació de la DNC tenint en compte condicions experimentals com són el nivell del 10% de saturació del sistema de captació, la quantització en 256 nivells de gris de la càmera, i la quantització en 32 nivells del modulador (veure capítol 5).

Cal tenir en compte que la informació d'ambdues gràfiques és només orientativa pels caràcters dels quals disposem en aquest exemple. Recordem que les lletres del test 3 que s'utilitzen en aquest apartat tenen només la forma coincident amb algunes lletres del test 1, sobre les quals es va construir la Figura 6.5. A més, els nivells de gris de les components de les lletres del test 3, no coincideixen en general amb els de la referència, com succeïa en el test 1. Per altra banda, els nivells de gris de les components RGB dels caràcters del test 3 no coincideixen exactament amb els nivells de gris analitzats en els caràcters que componen el test 2 i a partir dels quals es construeix la Figura 8.4.

De totes maneres, la informació que ens proporcionen els dos gràfics (Figures 8.3 i 8.4) ens serà de molta utilitat ja que permet preveure de manera aproximada el comportament del sistema en el cas estudiat. En les Figures 8.3 i 8.4, s'han marcat quatre zones corresponents a uns valors concrets dels paràmetres  $L$  i  $M$ . La primera regió, denominada com a zona A, està delimitada pels valors  $-1 \leq L \leq -0.6$  i  $M = 0.6$ . La segona zona del pla correspon als valors  $0.8 \leq L \leq 1.2$  i  $M = 1.4$  i s'ha anomenat zona B. Hem denominat com a zona C, la regió de l'espai  $LM$  al voltant dels valors

$L = 0$  i  $1.0 \leq M \leq 1.4$ . Per últim, la zona D correspon a la regió descrita per  $L = 0$  i  $-0.6 \leq M \leq -1.0$ .

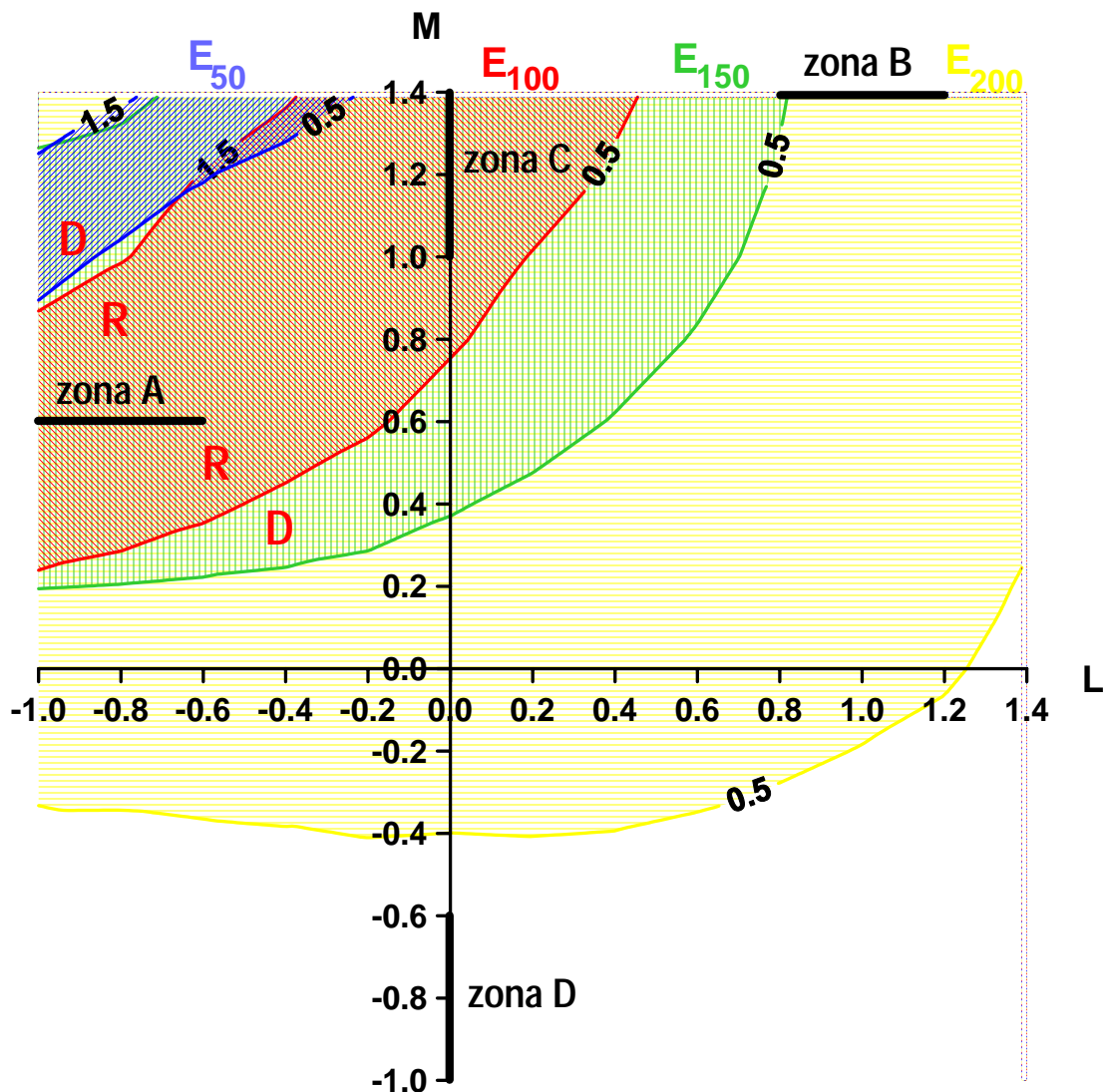


**Figura 8.3.** Zones A, B, C i D per a l'anàlisi de la DNC davant de variacions de forma i color. Es presenten també les corbes llindar de reconeixement ( $u = 0.5$ ) extretes de la Figura 6.5 pels caràcters del test 1 (Figura 6.1) que tenen la mateixa forma que els que constitueixen el test 3 (Figura 8.1).

Analitzem aquestes quatre regions A, B, C i D per conèixer el comportament del correlador DNC davant de canvis en la forma dels objectes, i canvis en el seu nivell de gris.

En la zona denominada com a A (Figura 8.3), estem en una regió de valors d' $L$  i  $M$  on la discriminació per la forma del caràcter és prou elevada ja que tant sols interceptem les corbes de reconeixement amb un llindar  $u=0.5$  pels caràcters  $E_1$  i  $E_2$ , els més similars en forma al patró. En canvi, estudiant la Figura 8.4, ens adonem que la zona A és una regió de valors d' $L$  i  $M$  als quals correspon una certa tolerància a canvis en el nivell de gris. Per altra banda, la zona B pertany a una regió de l'espai  $LM$  on la tolerància a canvis de forma és important, donat que està situada en les àrees de reconeixement de la major part dels caràcters de test 3 (Figura 8.3), i al mateix temps s'obté un sistema amb menor tolerància en la discriminació del nivell de gris (Figura 8.4). En el cas de la zona C, el sistema obté tolerància a ambdues característiques, donat que correspon a una zona intermèdia entre les diferents corbes llindar pertanyents a caràcters amb diferents formes (Figura 8.3), i a la vegada, també és un zona intermèdia entre les corbes llindar de reconeixement pels caràcters que presenten només variacions en la seva intensitat (Figura 8.4). Finalment, en la zona D, trobem el comportament oposat, donat que aquesta àrea està inclosa en un comportament molt discriminatiu tant a canvis en la forma (Figura 8.3), com en el nivell de gris (Figura 8.4).

En concordància amb aquesta anàlisi, si apliquem la DNC dins les regions esmentades, hem d'obtenir un sistema amb diferents comportaments en quant a la seva capacitat de discriminació davant canvis en la forma i en el nivell de gris dels objectes. Si a més, generalitzem l'anàlisi enfront de les variacions de nivell de gris, al conjunt de canals RGB, això ens conduirà a una tolerància o discriminació del sistema a variacions en el color dels objectes.



**Figura 8.4.** Zones A, B, C i D per a l'anàlisi de la DNC davant de variacions de forma i color. Es presenten també les corbes lliardar de reconeixement ( $u = 0.5$ ) extretes de la Figura 6.10 per a objectes d'igual forma i diferent nivell de gris constituents del test 2 (Figura 6.6).

### 8.1.3.a Discriminació a canvis de forma i tolerància a canvis en el color dels objectes analitzats (zona A)

La Taula 8.3 mostra els resultats de la DNC pels valors d' $L$  i  $M$  dins la zona A ( $-1 \leq L \leq -0.6$  i  $M = 0.6$ , veure Figures 8.3 i 8.4).



**Taula 8.3.** Resultats de reconeixement aplicant la DNC als objectes en color del test 3 (Figura 8.1) per valors d' $L$  i  $M$  dins la zona A ( $-1 \leq L \leq -0.6$  i  $M = 0.6$ ). Sistema molt discriminant a canvis en la forma i tolerant a petits canvis de color.

<b>Etiqueta</b>	<b>E<sub>ref</sub></b>	<b>E<sub>1</sub></b>	<b>E<sub>2A</sub></b>	<b>E<sub>2B</sub></b>	<b>F<sub>1</sub></b>	<b>F<sub>2</sub></b>	<b>E<sub>3</sub></b>	<b>E<sub>5</sub></b>	<b>R<sub>1</sub></b>
<b>Color</b>	<b>R</b>	<b>L</b>	<b>S</b>	<b>TF</b>	<b>L</b>	<b>R</b>	<b>R</b>	<b>R</b>	<b>R</b>
<b>Pr. Escalar</b> $cs = \frac{s \cdot r}{ r ^2}$	<b>1</b>	<b>0.95</b>	<b>0.92</b>	<b>0.92</b>	<b>0.89</b>	<b>0.86</b>	<b>0.85</b>	<b>0.84</b>	<b>0.74</b>
<b>L = -0.6</b> <b>M = 0.6</b>	<b>R</b>	<u>1</u>	<u>0.73</u>	<u>0.65</u>	<u>0.56</u>	<u>0.53</u>	0.40	0.30	0.17
	<b>G</b>	<u>1</u>	<u>0.74</u>	<u>0.70</u>	<u>0.81</u>	<u>0.53</u>	0.40	0.30	0.17
	<b>B</b>	<u>1</u>	<u>0.57</u>	<u>0.72</u>	0.0	<u>0.50</u>	0.40	0.30	0.17
<b>Reconeix.</b>	SI	SI	SI	NO	SI	NO	NO	NO	NO
<b>L = -0.8</b> <b>M = 0.6</b>	<b>R</b>	<u>1</u>	<u>0.74</u>	<u>0.63</u>	<u>0.53</u>	<u>0.57</u>	0.38	0.28	0.15
	<b>G</b>	<u>1</u>	<u>0.79</u>	<u>0.73</u>	<u>0.93</u>	<u>0.58</u>	0.38	0.28	0.15
	<b>B</b>	<u>1</u>	<u>0.54</u>	<u>0.76</u>	0.0	0.48	0.38	0.28	0.15
<b>Reconeix.</b>	SI	SI	SI	NO	NO	NO	NO	NO	NO
<b>L = -1</b> <b>M = 0.6</b>	<b>R</b>	<u>1</u>	<u>0.75</u>	<u>0.62</u>	<u>0.50</u>	<u>0.58</u>	0.36	0.25	0.12
	<b>G</b>	<u>1</u>	<u>0.82</u>	<u>0.75</u>	<u>1.08</u>	<u>0.60</u>	0.36	0.25	0.12
	<b>B</b>	<u>1</u>	<u>0.50</u>	<u>0.78</u>	0.0	0.46	0.36	0.25	0.12
<b>Reconeix.</b>	SI	SI	SI	NO	NO	NO	NO	NO	NO

Analitzant aquests resultats, veiem que en els casos d' $L = -0.8$  o  $-1$  i  $M = 0.6$ , aconseguim detectar conjuntament amb la referència, les lletres  $E_1$  i  $E_{2A}$ , que corresponen als caràcters més similars en forma a la lletra E de referència, i que estan acolorits uniformement amb tons lila (L) i salmó (S), respectivament, diferents al de la referència (rosa pàl·lid, R), però tots ells dins la mateixa gamma magenta. Veiem també que la lletra  $E_{2B}$ , de forma idèntica a la  $E_{2A}$ , acolorida amb taronja fosc (TF) no s'identifica amb la referència i això ens indica que tot i tenir un sistema tolerant a

canvis de color, no ha perdut la seva capacitat de discriminació de colors més dispers. Així mateix, la discriminació per la forma també és molt bona donat que la lletra  $F_1$  amb un coeficient de similitud molt alt, aconsegueix, encara que amb poc marge, ésser discriminada de la referència. La resta de lletres ( $F_2$ ,  $E_3$ ,  $E_5$  i  $R_1$ ), molt més diferents en forma, encara que tenen el mateix color que el patró (Rosa pàl·lid, R), no passen el procés de reconeixement i són rebutjades.

Veiem que els resultats comentats considerant els valors  $L = -0.8$  o  $-1$  i  $M = 0.6$ , es modifiquen lleugerament en l'extrem de l'interval  $L = -0.6$ ,  $M = 0.6$ , en el sentit que, la lletra  $F_1$  passa a ésser reconeguda com a similar a la referència donat que tot just iguala el llindar establert en  $u = 0.5$ . Això és així, donat que s'assoleixen, en aquest cas, valors de cc/ac molt propers al llindar de reconeixement establert. Obtenim, per tant, en l'interval  $-1 \leq L \leq -0.8$  i  $M = 0.6$ , un sistema discriminant a canvis de forma en els objectes, i tolerant al mateix temps a petits canvis en el color dins un mateix rang de tonalitats.

La Figura 8.5 mostra sobre la mateixa escena els caràcters que superen tot el procés de reconeixement en el cas d' $L = -0.8$  o  $-1$  i  $M = 0.6$ .



**Figura 8.5.** (a) Caràcters del test 3. (b) Reconeixement simulat tenint en compte condicions experimentals dels caràcters del test 3 emprant les no-linealitats  $L = -0.8$  (o  $-1$ ) i  $M = 0.6$  incloses en la zona A de l'espai  $LM$  (Figures 8.3 i 8.4).

### 8.1.3.b Tolerància a canvis en la forma i discriminació a canvis en el color dels objectes analitzats (zona B)

Considerem la segona zona escollida del pla ( $L, M$ ) indicada en les Figures 8.3 i 8.4 com a zona B. Fem variar els paràmetres  $L$  i  $M$  en passos de 0.2 dins l'interval descrit per  $0.8 \leq L \leq 1.2$  i  $M = 1.4$ . Mostrem en la Taula 8.4, els resultats de la DNC pels valors indicats d' $L$  i  $M$  inclosos dins la zona B.

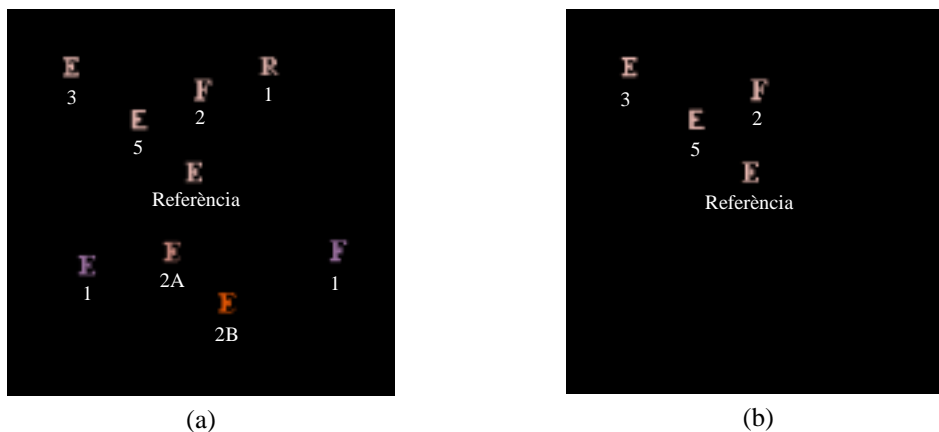
**Taula 8.4.** Resultats de reconeixement aplicant la DNC als objectes en color del test 3 (Figura 8.1) per valors d' $L$  i  $M$  dins la zona B ( $0.8 \leq L \leq 1.2$  i  $M = 1.4$ ). Sistema tolerant a petits canvis en la forma i molt discriminant a canvis de color.

<b>Etiqueta</b>	<b>E<sub>ref</sub></b>	<b>E<sub>1</sub></b>	<b>E<sub>2A</sub></b>	<b>E<sub>2B</sub></b>	<b>F<sub>1</sub></b>	<b>F<sub>2</sub></b>	<b>E<sub>3</sub></b>	<b>E<sub>5</sub></b>	<b>R<sub>1</sub></b>	
<b>Color</b>	<b>R</b>	<b>L</b>	<b>S</b>	<b>TF</b>	<b>L</b>	<b>R</b>	<b>R</b>	<b>R</b>	<b>R</b>	
<b>Pr. Escalar</b> $cs = \frac{\bar{s} \cdot \bar{r}}{ r ^2}$	<b>1</b>	<b>0.95</b>	<b>0.92</b>	<b>0.92</b>	<b>0.89</b>	<b>0.86</b>	<b>0.85</b>	<b>0.84</b>	<b>0.74</b>	
<b>L = 0.8</b> <b>M = 1.4</b>	<b>R</b>	<u>1</u>	<u>0.57</u>	<u>0.68</u>	<u>0.80</u>	0.46	<u>0.70</u>	<u>0.69</u>	<u>0.70</u>	0.45
	<b>G</b>	<u>1</u>	<u>0.50</u>	<u>0.56</u>	0.18	0.41	<u>0.70</u>	<u>0.69</u>	<u>0.70</u>	0.45
	<b>B</b>	<u>1</u>	<u>0.85</u>	<u>0.53</u>	0.0	<u>0.66</u>	<u>0.70</u>	<u>0.69</u>	<u>0.70</u>	0.45
<b>Reconeix.</b>	SI	SI	SI	NO	NO	SI	SI	SI	NO	
<b>L = 1</b> <b>M = 1.4</b>	<b>R</b>	<u>1</u>	0.49	<u>0.63</u>	<u>0.77</u>	0.40	<u>0.70</u>	<u>0.71</u>	<u>0.75</u>	0.47
	<b>G</b>	<u>1</u>	0.42	0.49	0.13	0.34	<u>0.70</u>	<u>0.71</u>	<u>0.75</u>	0.47
	<b>B</b>	<u>1</u>	<u>0.82</u>	0.46	0.0	<u>0.64</u>	<u>0.70</u>	<u>0.71</u>	<u>0.75</u>	0.47
<b>Reconeix.</b>	SI	NO	NO	NO	NO	SI	SI	SI	NO	
<b>L = 1.2</b> <b>M = 1.4</b>	<b>R</b>	<u>1</u>	0.44	<u>0.60</u>	<u>0.77</u>	<u>0.56</u>	<u>0.71</u>	<u>0.74</u>	<u>0.80</u>	<u>0.50</u>
	<b>G</b>	<u>1</u>	0.35	0.43	0.09	0.41	<u>0.71</u>	<u>0.74</u>	<u>0.80</u>	<u>0.50</u>
	<b>B</b>	<u>1</u>	<u>0.81</u>	0.40	0.0	0.37	<u>0.71</u>	<u>0.74</u>	<u>0.80</u>	<u>0.50</u>
<b>Reconeix.</b>	SI	NO	NO	NO	NO	SI	SI	SI	SI	

En aquesta ocasió observem que pel cas  $L = 1$  i  $M = 1.4$ , es detecten a més de la referència, els caràcters  $F_2$ ,  $E_3$  i  $E_5$ , que corresponen a les lletres més semblants segons el coeficient de similitud en forma (producte escalar en la Figura 8.1), de les que tenen el mateix color que la referència (R). Els caràcters  $E_1$  i  $E_{2A}$ , tot i tenir un coeficient de similitud en forma més proper a la unitat, són discriminants donat que estan acolorits amb una tonalitat diferent (lila i salmó, respectivament). La resta de caràcters, molt més diferents en color ( $E_{2B}$ ), en forma ( $R_1$ ), o en color i en forma conjuntament ( $F_1$ ), són també rebutjats en el procés de reconeixement.

Observem que per la resta de parells ( $L, M$ ) considerats, apareixen falses alarmes en quant a les lletres  $E_1$  i  $E_{2A}$ , de color diferent al de la referència, per la no-linealitat (0.8, 1.4). La lletra  $R_1$ , de forma considerablement diferent a la referència, es reconeix en el cas de (1.2, 1.4). Per aquest motiu, el sistema de reconeixement tolerant a petits canvis en la forma i discriminant a canvis en el color dels objectes, s'obté en un marge de valors d' $L$  i  $M$  limitat. En qualsevol cas, en el proper apartat realitzarem una proposta que ens permetrà augmentar aquest interval.

Representem en la Figura 8.6 els caràcters que superen en procés de reconeixement a l'aplicar les no-linealitats definides per  $L = 1$  i  $M = 1.4$  sobre l'escena a analitzar i a la referència, respectivament. En aquesta figura es mostren les lletres detectades com a similars al patró sobre la mateixa escena.



**Figura 8.6.** (a) Caràcters del test 3. (b) Reconeixement simulat tenint en compte condicions experimentals dels caràcters del test 3 emprant la no-linealitat  $L = 1$  i  $M = 1.4$  inclosa en la zona B de l'espai  $LM$  (Figures 8.3 i 8.4).

### 8.1.3.c Tolerància a canvis en la forma i en el color dels objectes analitzats (zona C)

La següent regió estudiada en aquesta anàlisi correspon a la que hem anomenat zona C en les Figures 8.3 i 8.4. Els paràmetres  $L$  i  $M$  varien dins l'interval descrit per  $L = 0$  i  $1 \leq M \leq 1.4$ . En la Taula 8.5, es detallen els resultats de la DNC aplicats als objectes en color del test 3 (Figura 8.1), quan s'apliquen les no-linealitats incloses en la zona C.

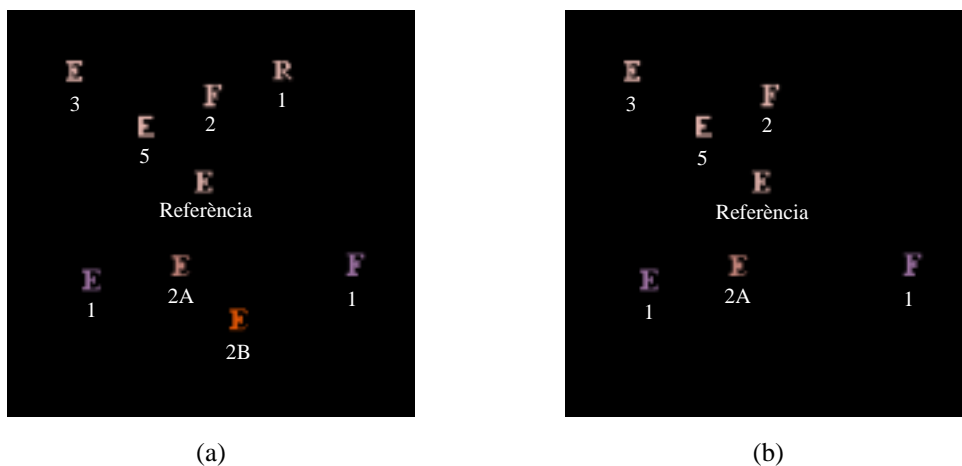
**Taula 8.5.** Resultats de reconeixement aplicant la DNC als objectes en color del test 3 (Figura 8.1) per valors d' $L$  i  $M$  dins la zona C ( $L = 0$  i  $1 \leq M \leq 1.4$ ). Sistema tolerant a petits canvis de forma i color.

Etiqueta	$E_{ref}$	$E_1$	$E_{2A}$	$E_{2B}$	$F_1$	$F_2$	$E_3$	$E_5$	$R_1$
<b>Color</b>	<b>R</b>	<b>L</b>	<b>S</b>	<b>TF</b>	<b>L</b>	<b>R</b>	<b>R</b>	<b>R</b>	<b>R</b>
<b>Pr. Escalar</b> $cs = \frac{\bar{x} \cdot \bar{r}}{ \bar{r} ^2}$	<b>1</b>	<b>0.95</b>	<b>0.92</b>	<b>0.92</b>	<b>0.89</b>	<b>0.86</b>	<b>0.85</b>	<b>0.84</b>	<b>0.74</b>
<b><math>L = 0</math></b> <b><math>M = 1.2</math></b>	<b>R</b>	<u>1</u>	<u>0.85</u>	<u>0.80</u>	<u>0.79</u>	<u>0.65</u>	<u>0.62</u>	<u>0.54</u>	0.44 0.36
	<b>G</b>	<u>1</u>	<u>0.84</u>	<u>0.80</u>	<u>0.73</u>	<u>0.66</u>	<u>0.62</u>	<u>0.54</u>	0.44 0.36
	<b>B</b>	<u>1</u>	<u>0.81</u>	<u>0.80</u>	0.0	<u>0.66</u>	<u>0.62</u>	<u>0.54</u>	0.44 0.36
<b>Reconeix.</b>	SI	SI	SI	NO	SI	SI	SI	NO	NO
<b><math>L = 0</math></b> <b><math>M = 1.4</math></b>	<b>R</b>	<u>1</u>	<u>0.91</u>	<u>0.85</u>	<u>0.83</u>	<u>0.73</u>	<u>0.67</u>	<u>0.59</u>	<u>0.50</u> 0.41
	<b>G</b>	<u>1</u>	<u>0.92</u>	<u>0.87</u>	<u>0.86</u>	<u>0.76</u>	<u>0.67</u>	<u>0.59</u>	<u>0.50</u> 0.41
	<b>B</b>	<u>1</u>	<u>0.87</u>	<u>0.86</u>	0.001	<u>0.73</u>	<u>0.67</u>	<u>0.59</u>	<u>0.50</u> 0.41
<b>Reconeix.</b>	SI	SI	SI	NO	SI	SI	SI	SI	NO
<b><math>L = 0</math></b> <b><math>M = 1.6</math></b>	<b>R</b>	<u>1</u>	<u>1.01</u>	<u>0.92</u>	<u>0.86</u>	<u>0.86</u>	<u>0.71</u>	<u>0.67</u>	<u>0.55</u> 0.48
	<b>G</b>	<u>1</u>	<u>1.03</u>	<u>0.96</u>	<u>1.00</u>	<u>0.88</u>	<u>0.71</u>	<u>0.67</u>	<u>0.55</u> 0.48
	<b>B</b>	<u>1</u>	<u>0.90</u>	<u>0.96</u>	0.002	<u>0.78</u>	<u>0.71</u>	<u>0.67</u>	<u>0.55</u> 0.48
<b>Reconeix.</b>	SI	SI	SI	NO	SI	SI	SI	SI	NO

A partir de l'anàlisi dels resultats numèrics de  $cc/ac$  per als diferents caràcters inclosos en el test 3, veiem que per  $L = 0$  i  $M = 1.4$  o  $1.6$ , aconseguim un sistema de reconeixement tolerant a petits canvis en la forma i al mateix temps, tolerant a petits canvis en el color. D'aquesta manera, lletres del mateix color que la referència, són detectades tot i tenir una forma més diferent ( $F_2$ ,  $E_3$  i  $E_5$ ), i al contrari, lletres molt similars en forma i, per tant, amb coeficient de similitud proper a la unitat, són reconegudes com al patró tot i tenir variacions en el color més apreciables ( $E_1$ ,  $E_{2A}$  i  $F_1$ ). Tot i així, el sistema manté un cert grau de discriminació que permet rebutjar lletres amb colors més diferents, pertanyents a una altra gamma de tonalitat ( $E_{2B}$ ), i lletres amb canvis importants en la forma ( $R_1$ ).

La diferència en el cas d' $L = 0$  i  $M = 1.2$ , ve donada per la discriminació de la lletra  $E_5$  que obté un valor de  $cc/ac$  lleugerament inferior al llindar establert.

La Figura 8.7 mostra els caràcters que superen el procés de reconeixement sobre la mateixa escena del test 3. Aquests resultats són equivalents per a les no-linealitats  $L = 0, M = 1.4$  i  $L = 0, M = 1.6$ .

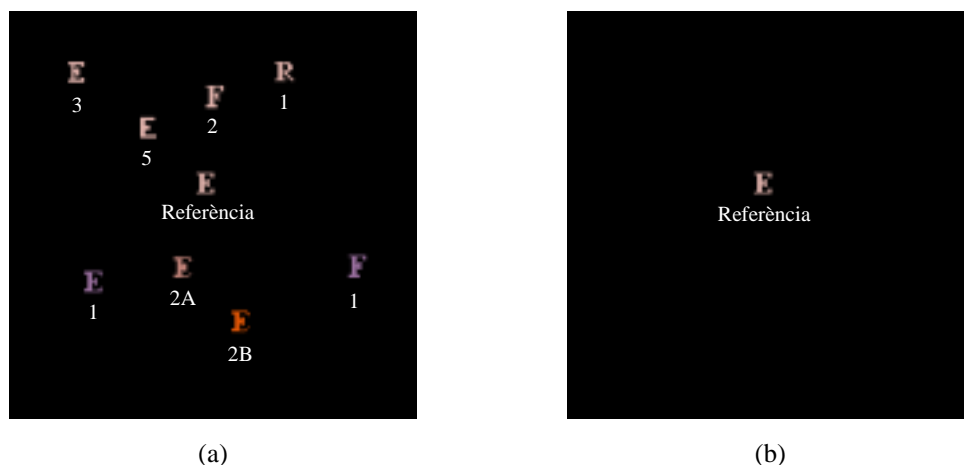


**Figura 8.7.** (a) Caràcters del test 3. (b) Reconeixement simulat tenint en compte condicions experimentals dels caràcters del test 3 emprant les no-linealitats  $L = 1, M = 1.4$  (o  $M = 1.6$ ) incloses en la zona C de l'espai  $LM$  (Figures 8.3 i 8.4).



En aquest cas, qualsevol parell  $(L, M)$  considerat dins la regió D, proporciona una màxima discriminació davant canvis en la forma i en el color dels objectes, donat que només supera el procés de reconeixement l'objecte escollit com a patró. Observem també, que en l'interval considerat, s'inclou com a cas particular el filtre adaptat IF, que, seguint el procediment de la DNC i tenint en compte els condicionaments experimentals, proporciona una màxima discriminació al sistema de reconeixement, tal com es preveia a l'analitzar aquest mètode de filtratge en condicions ideals (Taula 8.2).

Representem en aquest cas també, els caràcters inclosos en el test 3 que superen el procés de reconeixement. Aquesta representació la realitzem sobre la mateixa escena a estudi (Figura 8.8).

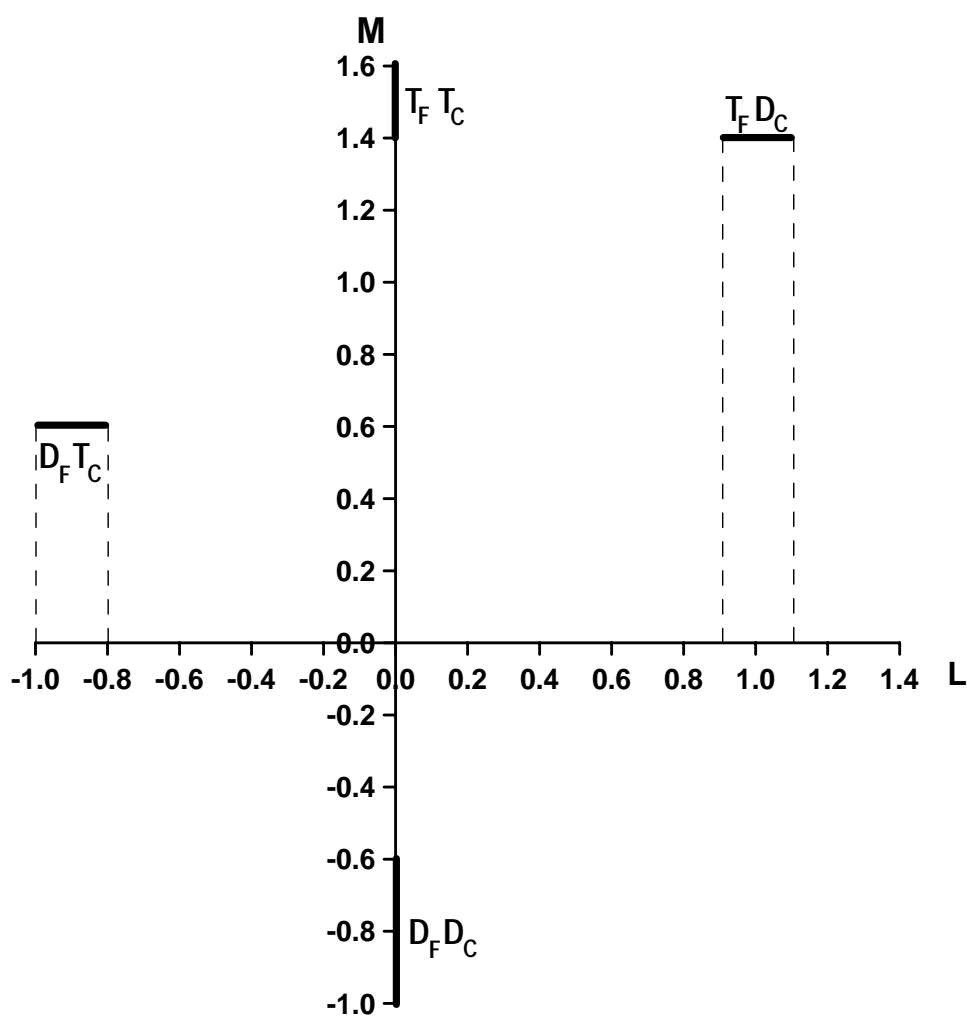


**Figura 8.8.** (a) Caràcters del test 3. (b) Reconeixement simulat tenint en compte condicions experimentals dels caràcters del test 3 emprant les no-linealitats incloses en l'interval  $L = 0, -0.6 \leq M \leq -1$ , de la zona D de l'espai  $LM$  (Figures 8.3 i 8.4).

A manera de resum, podem representar sobre l'espai  $LM$  definit per les no-linealitats aplicades independentment sobre la informació referent a l'escena a analitzar ( $L$ ) i sobre la informació de l'objecte a detectar ( $M$ ), les regions que proporcionen un comportament diferent del sistema de reconeixement en quant a la seva tolerància a canvis en la forma i en el color dels objectes. Hem vist, amb els exemples aportats en



aquest apartat, que podem delimitar regions on existeix una alta discriminació a canvis de forma i tolerància a canvis en el color ( $D_F T_C$ ); la viceversa, tolerància a canvis en la forma i discriminació a canvis en el color ( $T_F D_C$ ); tolerància conjuntament a canvis de forma i de color ( $T_F T_C$ ), i per últim, una gran discriminació a ambdós tipus de canvis (forma i color) ( $D_F D_C$ ). Marquem les regions esmentades en la Figura 8.9.



**Figura 8.9.** Delimitació en l'espai LM de les regions:  $D_F T_C$ , d'alta discriminació a canvis en la forma i tolerància a canvis en el color;  $T_F D_C$ , de tolerància a canvis en la forma i discriminació a canvis en el color;  $T_F T_C$ , de tolerància a canvis en la forma i en el color;  $D_F D_C$ , d'alta discriminació a canvis en la forma i en el color.

## **8.2 Descomposició adaptada en canals quasimonocromàtics aplicada al reconeixement d'objectes en color mitjançant la DNC**

Amb l'objectiu de millorar els resultats de reconeixement presentats fins ara en aquest capítol, proposem combinar el mètode de correlació multicanal amb el de la descomposició adaptada en canals quasimonocromàtics del capítol 3.

Hem posat de manifest en el capítol 3, les avantatges de la descomposició cromàtica adaptada en  $n$  canals quasimonocromàtics, sobretot en els casos on els objectes presents a l'escena tenen colors molt propers entre ells, és a dir, pertanyents a una mateixa gamma de tonalitat. Donat que en el test 3, la majoria de caràcters estan acolorits amb els colors agrupats en la gamma magenta, l'aplicació de la descomposició en canals quasimonocromàtics prèviament a l'aplicació de la DNC, serà especialment adequada quan els requeriments en discriminació pel color siguin alts. Dit d'una altra manera, utilitzarem l'acció combinada de la descomposició cromàtica en els canals seleccionats i la DNC, en els casos on volem obtenir una alta discriminació davant de canvis en els colors dels objectes. Aquests casos corresponen a les regions denominades  $T_F D_C$  i  $D_F D_C$ , respectivament. Donat que les zones de treball del sistema de reconeixement on s'obtenen capacitats de discriminació variable, s'obtenen en regions limitades en quant a les variacions dels paràmetres de les no-linealitats, amb aquesta proposta pretenem aconseguir unes regions de treball més àmplies, que ens permetin un procés de reconeixement més relaxat.

Tenint en compte, que els colors de la gamma magenta presents en el test 3 són el rosa pàl·lid (R), el salmó (S) i el lila (L), podem determinar a partir de la Taula 3.1, quines longituds d'ona són les més adequades per portar a terme la descomposició cromàtica. Concretament, les parelles de colors R-S i R-L obtenen les màximes diferències proporcionals en reflectància en unes longituds d'ona molt properes, 585 nm i 590 nm, respectivament. Considerant els filtres interferencials disponibles, considerarem en ambdós casos la longitud d'ona de 589 nm per a portar a terme la descomposició adaptada.

### 8.2.1 Tolerància a canvis de forma i discriminació a canvis en el color dels objectes analitzats

En la Taula 8.7 mostrem els resultats obtinguts per simulació en l'aplicació de la DNC a la component  $\lambda=589$  nm de la descomposició cromàtica adaptada. Es donen els valors del màxim de correlació creuada normalitzats al màxim d'autocorrelació per a cada una de les escenes segmentades representades en el test 3.

**Taula 8.7.** Resultats de reconeixement aplicant la DNC dins l'interval  $0.8 \leq L \leq 1.2$  i  $M = 1.4$ , a la component  $\lambda=589$  nm de la descomposició adaptada realitzada a la informació dels objectes en color del test 3 (Figura 8.1). Sistema tolerant a canvis en la forma i molt discriminant a canvis en el color.

<b>Etiqueta</b>	<b>E<sub>ref</sub></b>	<b>E<sub>1</sub></b>	<b>E<sub>2A</sub></b>	<b>E<sub>2B</sub></b>	<b>F<sub>1</sub></b>	<b>F<sub>2</sub></b>	<b>E<sub>3</sub></b>	<b>E<sub>5</sub></b>	<b>R<sub>1</sub></b>	<b>H<sub>1</sub></b>
<b>Color</b>	<b>R</b>	<b>L</b>	<b>S</b>	<b>TF</b>	<b>L</b>	<b>R</b>	<b>R</b>	<b>R</b>	<b>R</b>	<b>R</b>
<b>Pr. Escalar</b> $cs = \frac{s \cdot F}{ F ^2}$	<b>1</b>	<b>0.95</b>	<b>0.92</b>	<b>0.92</b>	<b>0.89</b>	<b>0.86</b>	<b>0.85</b>	<b>0.84</b>	<b>0.74</b>	<b>0.57</b>
<hr/>										
<b><math>\lambda=589</math> nm</b>										
<b>L = 0.8</b>	<u>1</u>	0.31	0.47	0.43	0.24	<u>0.70</u>	<u>0.69</u>	<u>0.70</u>	0.45	0.29
<b>M = 1.4</b>										
<b>Reconeix.</b>	SI	NO	NO	NO	NO	SI	SI	SI	NO	NO
<hr/>										
<b><math>\lambda=589</math> nm</b>										
<b>L = 1</b>	<u>1</u>	0.22	0.40	0.35	0.18	<u>0.70</u>	<u>0.71</u>	<u>0.75</u>	0.47	0.32
<b>M = 1.4</b>										
<b>Reconeix.</b>	SI	NO	NO	NO	NO	SI	SI	SI	NO	NO
<hr/>										
<b><math>\lambda=589</math> nm</b>										
<b>L = 1.2</b>	<u>1</u>	0.19	0.33	0.28	0.13	<u>0.71</u>	<u>0.74</u>	<u>0.80</u>	<u>0.50</u>	0.35
<b>M = 1.4</b>										
<b>Reconeix.</b>	SI	NO	NO	NO	NO	SI	SI	SI	SI	NO

La Taula 8.7 ens permet posar de manifest que en l'interval  $0.8 \leq L \leq 1$  i  $M = 1.4$ , s'obtenen resultats de reconeixement equivalents, on es detecten com a similars al patró els caràcters  $F_2$ ,  $E_3$  i  $E_5$ . La resta de caràcters del test 3 són rebutjats, ja sigui pel color diferent ( $E_1$ ,  $E_{2A}$ ,  $E_{2B}$ ), per les diferències acusades en la forma ( $R_1$ ), o bé pels dos motius conjuntament ( $F_1$ ). Veiem que, amb la descomposició en canals quasimonocromàtics, hem aconseguit en l'exemple mostrat, augmentar lleugerament l'interval de valors  $L$  i  $M$  en els quals és possible l'obtenció d'un sistema amb tolerància a canvis en la forma dels objectes, i al mateix temps aconseguir una bona discriminació a canvis en el color. A més, els resultats de la raó  $cc/ac$  per les lletres amb colors diferents al de la referència, s'allunyen més de la unitat en comparació dels resultats obtinguts amb el sistema RGB (Taula 8.4), indicant unes majors diferències entre els pics de correlació creuada i d'autocorrelació, la qual cosa facilita un procés de reconeixement més relaxat. Per últim, cal comentar també, que pels objectes de la gamma magenta que apareixen en el test 3, només ha estat necessari l'ús d'un sol canal quasimonocromàtic, accelerant i simplificant molt el procés global de reconeixement.

### 8.2.2 Discriminació elevada respecte a la forma i al color dels objectes analitzats

En el segon cas on es requereix una alta discriminació correspon a l'aplicació del mètode DNC per a l'obtenció d'un sistema amb bona discriminació davant canvis tant de forma com de color en els objectes analitzats. Apliquem la descomposició en canals quasimonocromàtics seleccionats dins l'interval  $L = 0$  i  $-0.6 \leq M \leq -1$ . La Taula 8.8 detalla els resultats de reconeixement en aquest cas, i proporciona els valors de la raó  $cc/ac$ , que apareixen subratllats en el cas en què es trobin dins la banda de reconeixement delimitada pel doble llinar ( $u = 0.5$ ).

Emprant el canal quasimonocromàtic  $\lambda = 589$  nm seleccionat, obtenim una màxima discriminació en el sistema de reconeixement, rebutjant tots els caràcters presents a l'escena test, excepte la referència. Aquest resultat de reconeixement es va obtenir també amb els canals RGB (Taula 8.6). L'avantatge d'emprar la descomposició adaptada en canals quasimonocromàtics, radica en l'obtenció d'uns marges més



# Capítol 9

## Conclusions

En el panorama actual dels processadors d'imatges basats en la correlació òptica, es pot dir que s'ha avançat molt en torn a aconseguir un augment de la capacitat discriminativa del sistema de reconeixement que permeti diferenciar el patró buscat de qualsevol altre objecte que presenti algun tipus de diferència. En una altra línia de desenvolupament, trobem els sistemes que emmagatzemen informació del patró en diferents situacions, o atorguen al sistema la capacitat de reconeixement invariant a distorsions, com ara rotacions, canvis d'escala, d'il·luminació, etc. En aquesta situació de diversitat de mètodes i sistemes, un canvi d'objectiu o de tasca de reconeixement, implica, inevitablement, un canvi important del procediment a seguir i, tal vegada, del dispositiu òptic a emprar.

En aquest treball, hem proposat un sistema de reconeixement basat en la correlació òptica que té, com a característica més destacada, la capacitat de discriminació variable en funció dels requeriments del problema plantejat. Sense necessitat de modificar el dispositiu optoelectrònic i variant de manera senzilla els valors d'uns paràmetres inclosos en el processat digital intermedi, s'aconsegueixen reproduir

diferents mètodes de correlació, lineals i no-lineals convencionals, d'altres intermedis, i també permet, dins el mateix model, nous processaments de la informació. D'aquesta manera, s'obté un sistema molt versàtil, en què és possible controlar graus de tolerància diferents en funció de vàries característiques dels objectes.

En la línia d'adaptar la capacitat de discriminació a les necessitats del problema de reconeixement plantejat, incloem el cas on la discriminació dels objectes requereix diferenciar colors similars, pertanyents a una mateixa gamma de tonalitat. En aquest sentit, hem proposat un mètode de descomposició multicanal de la informació cromàtica de l'escena a analitzar i de l'objecte a reconèixer, que cal aplicar prèviament a la correlació òptica. La descomposició multicanal es fa mitjançant una selecció de canals quasimonocromàtics que s'adapta la gamma de tonalitat dels colors implicats.

Per l'anàlisi d'escenes acromàtiques on es pretengui un reconeixement d'objectes amb una tolerància variable a canvis de forma, nivell de gris i textures, hem proposat la utilització del model de la correlació no-lineal dual (DNC). Aquest model presenta l'avantatge de l'adaptabilitat, ja que, introduint petits canvis que no afecten el seu mecanisme de funcionament, s'aconsegueix reconduir la solució de reconeixement amb un grau de tolerància variable. S'ha vist que aquest model de la DNC pot també estendre la seva aplicació al reconeixement d'objectes en color.

Ambdues propostes, la descomposició cromàtica multicanal en canals quasimonocromàtics seleccionats i la DNC, s'han estudiat separatament. No obstant això, també s'ha proposat la seva aplicació conjunta en els casos on els requeriments de discriminació cromàtica són elevats i, a l'hora, es permet un cert grau de tolerància respecte a un altre aspecte de l'objecte, com per exemple, la seva forma.

Del treball de recerca desenvolupat, s'extreuen les conclusions que a continuació es detallen.

1. Primerament, s'ha establert un criteri de decisió basat en un doble llindar que té en compte el fet que l'objecte buscat no té perquè ser el més energètic entre els que componen l'escena. Aquest doble llindar restringeix el reconeixement als objectes que, a més de tenir una forma similar a la de l'objecte buscat, tenen també una energia similar a la seva. Aquest criteri és particularment convenient en les aplicacions desenvolupades pel reconeixement multicanal d'objectes en color. Tot i així, és possible la seva generalització a altres casos de reconeixement, tal com s'ha aplicat en aquesta memòria.
2. El mètode de descomposició multicanal adaptada ha permès l'augment selectiu de la capacitat de discriminació cromàtica del sistema de reconeixement. El mètode consisteix en l'elecció d' $n$  canals quasimonocromàtics on les longituds d'ona de màxima transmissió dels quals, es centren en les regions de l'espectre on els colors d'una gamma determinada de tonalitat presenten les majors diferències proporcionals. Al ser una solució adaptada, gammes de colors diferents requereixen conjunts de canals quasimonocromàtics seleccionats també diferents. La selecció dels canals és més precisa quan es disposa d'un coneixement previ dels colors dels objectes, i és aproximada quan s'obté d'una altra escena d'entrenament de colors similars.

S'ha realitzat experimentalment la descomposició multicanal en els canals quasimonocromàtics seleccionats. La component en cada canal s'ha obtingut mitjançant una càmera CCD en blanc i negre, davant de l'objectiu de la qual s'ha posicionat un filtre interferencial de banda estreta. El filtre té la transmissió màxima en una longitud d'ona propera a la calculada mitjançant les diferències proporcionals de les corbes de reflectància dels colors analitzats.

L'ús d'un correlador convergent ha facilitat la realització experimental de la correlació multicanal d'escenes multiobjecte formades per diferents objectes de colors dins una mateixa gamma de tonalitat. S'ha demostrat que els filtres adaptats del conjunt FPF obtenen resultats finals de reconeixement teòricament equivalents quan els objectes són iguals en forma i estan acolorits amb tonalitats uniformes. S'ha escollit el filtre adaptat, doncs, en funció de les característiques en la seva



realització experimental. En totes les proves realitzades, s'ha posat de manifest una millora significativa dels resultats de reconeixement obtinguts pel mètode dels canals quasimonocromàtics seleccionats davant de la descomposició RGB. En tots els casos, s'han eliminat les falses alarmes produïdes pel mètode RGB, s'han augmentat les diferències entre els màxims d'autocorrelació i de correlació creuada de manera que el procés de reconeixement s'ha portat a terme de forma més relaxada, i en molts casos, s'ha aconseguit reduir el nombre de canals necessari respecte els tres emprats pel sistema RGB.

S'ha estudiat l'aplicació del mètode de descomposició multicanal adaptada quan les escenes a analitzar són tests especialment dissenyats formats per objectes de color uniforme, quan les escenes contenen objectes naturals i, fins i tot, quan contenen objectes policromàtics. En tots els casos, els resultats obtinguts amb els canals quasimonocromàtics escollits han millorat respecte als resultats obtinguts amb la descomposició RGB.

3. Per a l'obtenció del sistema de reconeixement òptic amb capacitat de discriminació variable, hem emprat la correlació no-lineal dual (DNC) com a mètode de processament base del nostre sistema. Les operacions lineals i no-lineals aplicades independentment sobre l'escena a analitzar i l'objecte de referència, poden ser modificades de manera contínua variant uns paràmetres reals,  $L$  i  $M$ , inclosos en un processat digital intermedi. L'arquitectura sobre la que es realitza la DNC, és el correlador de transformades conjuntes (JTC), que presenta l'avantatge de no requerir filtres adaptats a l'objecte a reconèixer.
4. Les simplificacions proposades del dispositiu optoelectrònic que realitza la DNC i de l'algorisme de càlcul no-lineal aplicat a l'espectre de transformades conjuntes (JPS), respecte del seu disseny i formulació inicials, no han afectat significativament els resultats obtinguts.
  - 4.1 Per una banda, el dispositiu emprat en aquest treball, redueix el nombre de components òptics i optoelectrònics, aconseguint un abaratiment del sistema òptic, i facilitant la seva realització experimental i alineació dels seus

---

components òptics. Aquesta variant del dispositiu inicial, requereix un control més acurat de les oscil·lacions en la il·luminació del sistema, control que és factible en la realització experimental, però també redueix els errors produïts quan s'utilitzen varis moduladors i càmeres CCD per reproduir i enregistrar senyals que cal comparar i operar conjuntament durant les operacions no-lineals portades a terme.

- 4.2 Per altra banda, s'ha proposat una modificació de l'algorisme del càlcul digital. Aquesta modificació ha consistit en una simplificació del procés que redueix el temps de càlcul i evita les dificultats associades a la subtracció, òptica o digital, de les distribucions d'intensitat truncades per l'efecte de la saturació del sensor CCD del sistema de captació. S'ha comprovat que l'eixamplament de l'ordre central del pla de sortida del correlador provocat per aquest canvi en l'algorisme, no afecta significativament als resultats de correlació situats en els ordres de difracció contigus, sempre i quan existeixi una distància suficient entre l'escena i l'objecte a detectar en el pla d'entrada del correlador.
5. S'han estudiat els efectes d'algunes limitacions experimentals del sistema DNC simplificat, sobre els resultats de correlació i sobre el paràmetre PCE (*Peak-to-correlation energy*) que avalua la qualitat del pic de correlació. Els condicionaments experimentals que s'han tingut en compte són deguts fonamentalment a les limitacions tècniques dels dispositius optoelectrònics, com la saturació del sistema de captació en l'enregistrament de distribucions d'intensitat amb rang de valors molt ampli, l'efecte de la quantització en un nombre limitat de nivells de gris introduït pel sistema d'adquisició, el nivell de zero escollit en l'enregistrament de les distribucions d'intensitat abans esmentades, i l'efecte de quantització en un cert nombre de nivells de gris que introdueix el modulador en la reproducció del JPS. Dels condicionaments esmentats, destaca com a més important l'efecte de la quantització del sistema d'adquisició. La reducció de la informació en les altes freqüències de l'espectre com a conseqüència de la quantització en pocs nivells, pot ser prou important

perquè elimini la sensibilitat del sistema DNC a variacions en les no-linealitats aplicades. S'ha comprovat que la introducció d'un nivell de saturació del sistema de captació pot contrarestar en part l'efecte negatiu de la quantització. En aquest cas, s'ha determinat el nivell de saturació més adient per a la posterior realització experimental. L'efecte degut a les modificacions introduïdes en el nivell de zero de les distribucions enregistrades, modifica la regió de suport i el contrast de les franges del JPS. S'ha establert el nivell de zero de captació en el valor mínim possible per aconseguir una màxima informació en les freqüències més altes, tot i perdre contrast en les franges d'interferència del JPS modificat, ja que es demostra que en aquestes condicions s'obtenen uns millors resultats de correlació. Finalment, en quant a la quantització introduïda pel modulador espacial de llum, s'observa que afecta més a les no-linealitats que generen distribucions amb un rang de nivells de gris més ampli pel JPS modificat. Tanmateix, només en el cas de treballar amb un nombre molt reduït de nivells de quantització en el modulador, el PCE es veu afectat de manera important. Considerant els dispositius optoelectrònics disponibles per a la realització experimental de la DNC, s'han establert els valors dels paràmetres experimentals que millor descriuen el comportament del sistema. S'ha comprovat la viabilitat de la proposta presentada, donat que el sistema DNC simplificat es manté sensible a les variacions en les no-linealitats aplicades.

6. S'ha estudiat el control en la capacitat de discriminació variable del sistema DNC davant de modificacions en la forma, en la intensitat, en la textura i en el color dels objectes a analitzar, aplicant el procediment de reconeixement a un conjunt de tests que a continuació descrivim.

6.1 S'han considerat caràcters alfanumèrics uniformes i texturats per a conèixer el comportament del sistema davant de la presència de diferents textures periòdiques senzilles. S'han escollit diferents funcions de pes que modifiquen la regió de suport de les distribucions sobre la qual s'aplica la no-linealitat. S'ha demostrat que amb funcions màscara elementals i diferents no-linealitats es superen àmpliament les possibilitats dels mètodes convencionals

---

caracteritzats per una capacitat de discriminació fixada. S'ha modificat independentment el grau de tolerància a la forma i a la variació del patró de textura dels objectes.

- 6.2 En una segona tasca de reconeixement, s'han considerat escenes monobjectes consistents en caràcters alfanumèrics que presenten diferències en la seva forma. El grau de similitud dels objectes analitzats, s'ha mesurat a partir del producte escalar normalitzat. Dins el model de la DNC, s'han delimitat les regions de reconeixement i de discriminació dels caràcters estudiats en un espai generat pels paràmetres  $L$  i  $M$  de les no-linealitats aplicades, establint un valor llindar per a la capacitat de discriminació. En la regió denominada de reconeixement, l'objecte corresponent és detectat com a similar al patró, i fora d'ella, l'objecte és rebutjat. S'han marcat diferents àrees a l'espai  $LM$  on és possible modificar de manera contínua el grau de discriminació del sistema davant dels canvis en la forma. El sistema DNC també permet escollir altres regions de l'espai  $LM$  on la capacitat de discriminació varia més lentament amb els paràmetres  $L$  i  $M$  i, fins i tot, regions on es manté pràcticament constant malgrat modificar el valor dels paràmetres del càlcul no-lineal.
- 6.3 S'ha aplicat el sistema de reconeixement proposat en un altre test consistent en caràcters d'igual forma i nivell de gris uniforme, però diferent entre ells. Aquest test, s'ha emprat per avaluar el comportament de la DNC davant de canvis en la intensitat dels objectes analitzats. S'ha comprovat que els mètodes de filtratge convencionals no aporten resultats satisfactoris. En el model de la DNC, però, s'ha determinat a l'espai  $LM$  les bandes de reconeixement per a cada un dels caràcters en el cas ideal, i també en el cas de considerar condicions experimentals. S'han observat modificacions importants al considerar els condicionaments experimentals esmentats, respecte els resultats obtinguts en el cas ideal. S'ha constatat la quantització del sistema de captació com la causa principal d'aquestes divergències, donat el retall de freqüències altes que aquest fet provoca. S'ha determinat una regió en l'espai  $LM$  on la capacitat discriminativa del sistema varia de manera

progressiva des d'una discriminació màxima fins a un grau de tolerància relativament elevat. Es poden determinar també, i s'han proposat, altres regions de l'espai  $LM$  on la discriminació del sistema té una variació més lenta, o pràcticament es manté constant.

7. S'ha portat a terme experimentalment el sistema de reconeixement DNC i s'han obtingut resultats que demostren la possibilitat d'obtenir realment una capacitat de discriminació variable del sistema davant de canvis de forma i d'intensitat dels objectes. S'han inclòs dispositius optoelectrònics que permeten la realització del procés de reconeixement en un temps només limitat pel temps de computació de la no-linealitat aplicada a la informació de l'escena a analitzar i a l'objecte a detectar. En les proves experimentals, s'han tingut en compte, separadament, les diferències en la forma i en la intensitat dels objectes, i s'han aplicat les no-linealitats corresponents a les regions delimitades per simulació on la capacitat de discriminació del sistema passa pels valors més alts, fins a un cert grau de tolerància. S'han buscat zones intermèdies en la capacitat de discriminació.
8. En quant al sistema DNC davant de variacions en la forma dels objectes, s'ha aconseguit reproduir experimentalment el comportament predit en simulació, i s'han determinat els paràmetres  $L$  i  $M$  que permeten obtenir una variació contínua de la discriminació del sistema. Els resultats experimentals s'adiuen amb els previstos per simulació, demostrant-se la viabilitat de la proposta en aquest cas.
9. S'ha aconseguit una capacitat de discriminació variable en el reconeixement experimental d'objectes que presenten diferències en el seu nivell de gris, equivalent a canvis en la intensitat dels mateixos. Els resultats experimentals no són totalment equivalents als predits per simulació, sinó que les possibilitats de la DNC apareixen més limitades. En aquest cas, considerem important la influència de les característiques tècniques del modulador espacial de llum en la reproducció del JPS modificat.

10. S'ha plantejat l'aplicació del processador DNC pel reconeixement d'objectes en color amb control sobre la capacitat discriminativa del sistema. S'han analitzat diferents tasques de reconeixement segons la tolerància a canvis de forma i/o a canvis en el color dels objectes. S'han aconseguit bons resultats de simulació en els següents quatre problemes plantejats: un sistema amb alta discriminació a la forma i tolerant a canvis de color; l'operació inversa, és a dir, un sistema tolerant a canvis en la forma, i discriminant en relació al color; un sistema amb alta discriminació a la forma i al color dels objectes, i viceversa, un sistema amb tolerància a ambdós aspectes.
  
11. En la consecució d'un sistema de reconeixement amb capacitat de discriminació controlable respecte a la forma dels objectes i, al mateix temps, amb una alta discriminació cromàtica, s'ha proposat l'aplicació conjunta de la DNC, sobre les components cromàtiques obtingudes a partir de la descomposició adaptada en  $n$  canals quasimonocromàtics seleccionats. Sobre un exemple on els caràcters alfanumèrics de prova contenien colors uniformes propers en tonalitat entre ells, s'ha demostrat que la descomposició adaptada prèvia a l'aplicació de la DNC permet obtenir uns resultats de reconeixement satisfactoris de manera més relaxada que la descomposició RGB i, a la vegada, s'aconsegueix augmentar la regió a l'espai  $LM$  on el sistema té una alta discriminació davant de canvis de color.



# **Annex I**

## **Escenes i corbes de reflectància espectral**

En aquest annex es presenten les escenes incloses en [189] que s'han utilitzat com a test en el procés de reconeixement d'objectes en color per correlació multicanal pel mètode de selecció adaptada de canals quasimonocromàtics. A part de les escenes per les gammes magenta i taronja que ja s'han presentat en el capítol 3 d'aquesta memòria, l'estudi del mètode proposat es va realitzar per tres gammes de tonalitat més: marró, blava i verda.

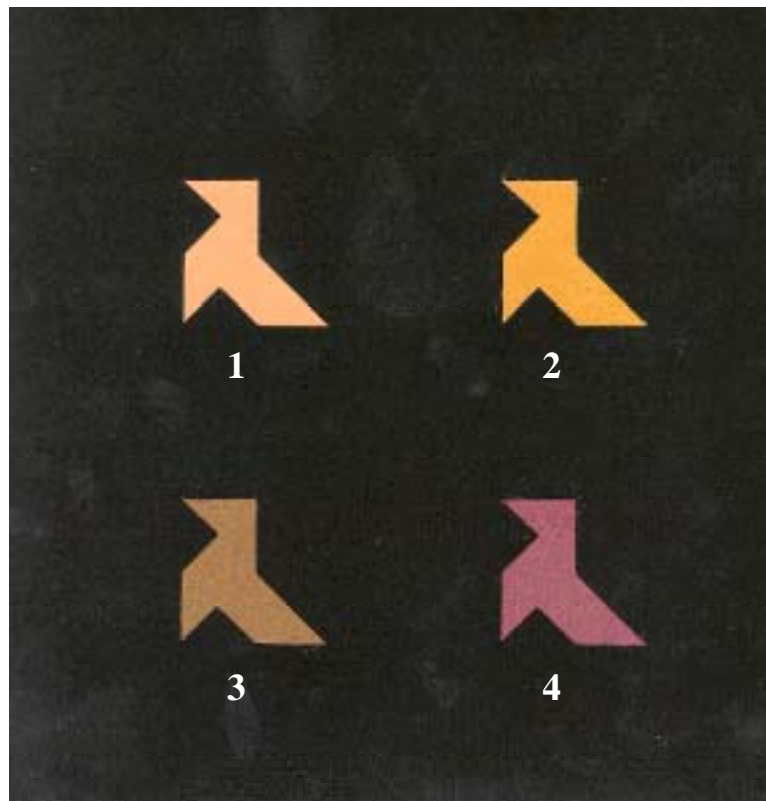
Junt a cada una de les escenes es mostren les corbes de reflectància espectral de les cartolines de colors amb què s'han generat els quatre objectes, obtingudes mitjançant un espectrofotòmetre UV/VIS de Perkin Elmer, model Lambda 3B series PECSS. A cada gràfic s'ha senyalat també les longituds d'ona dels filtres interferencials escollits per a obtenir la descomposició de la informació cromàtica de les escenes en els canals seleccionats.

S'adjunta, per cada una de les escenes test, una taula amb la informació relativa a la diferència proporcional més gran entre les cada parella de corbes d'una mateixa

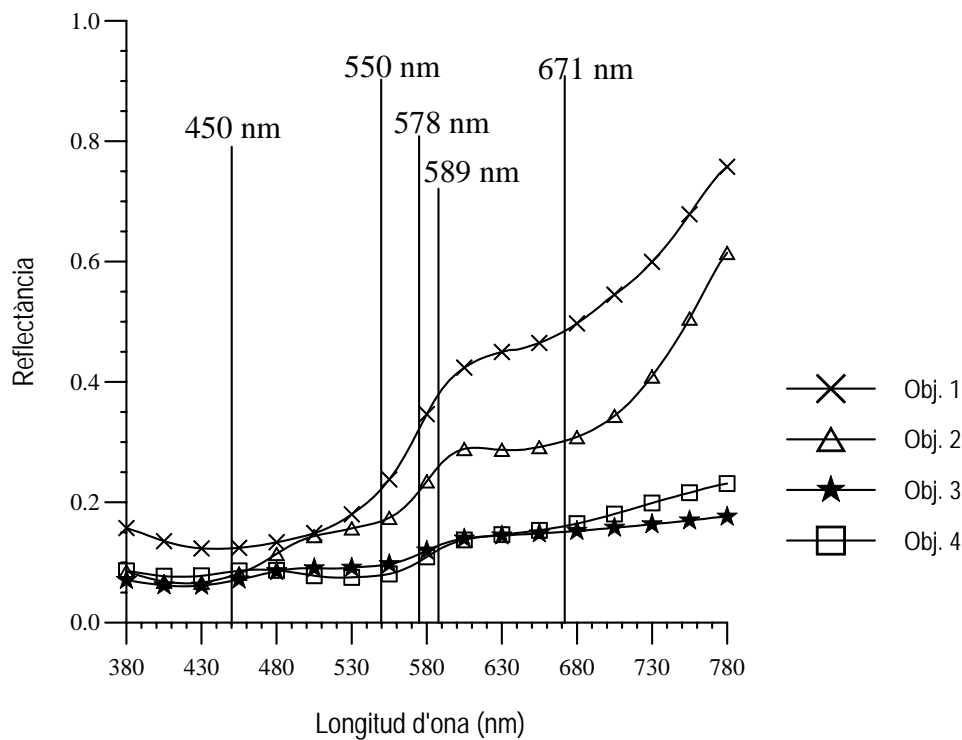


gamma de color, el valor de  $\lambda_0$  (Eq. 3.5) on s'obté la diferència proporcional màxima. També es proporciona el valor mig de les dues corbes en la longitud d'ona  $\lambda_0$  per tal d'indicar si la banda de l'espectre serà prou energètica. I per últim, es dona el valor  $\lambda_{fi}$  que coincideix amb la longitud d'ona del filtre interferencial escollit per realitzar experimentalment la descomposició cromàtica.

### I.1 Gamma marró



**Figura I.1.** Escena de la gamma marró.

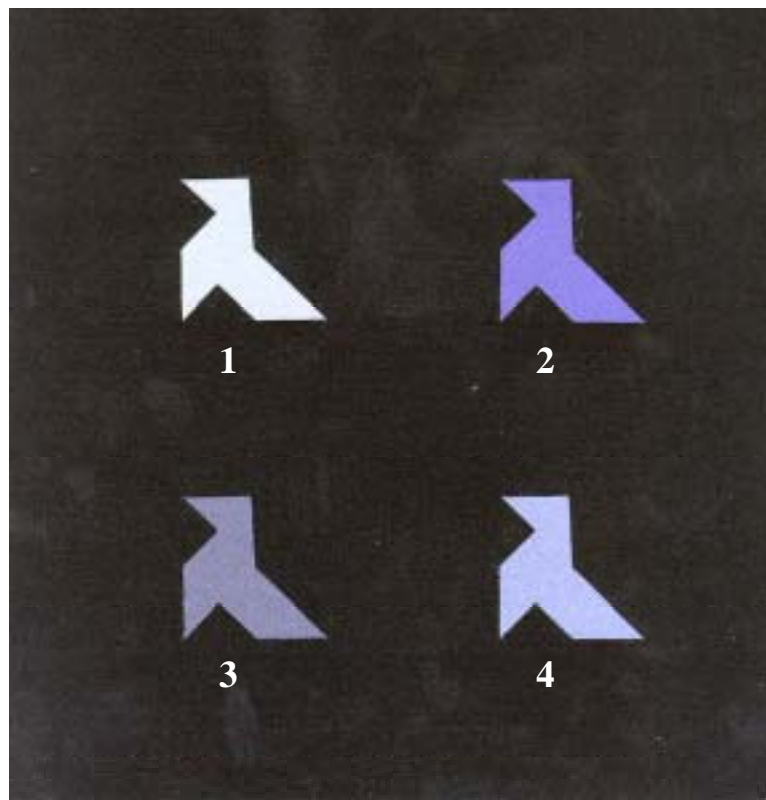


**Figura I.2.** Corbes de reflectància espectral de cada un dels objectes de la gamma marró (Figura I.1). S'indica el conjunt de longituds d'ona dels filtres interferencials escollits ( $\lambda_{fi}$ ).

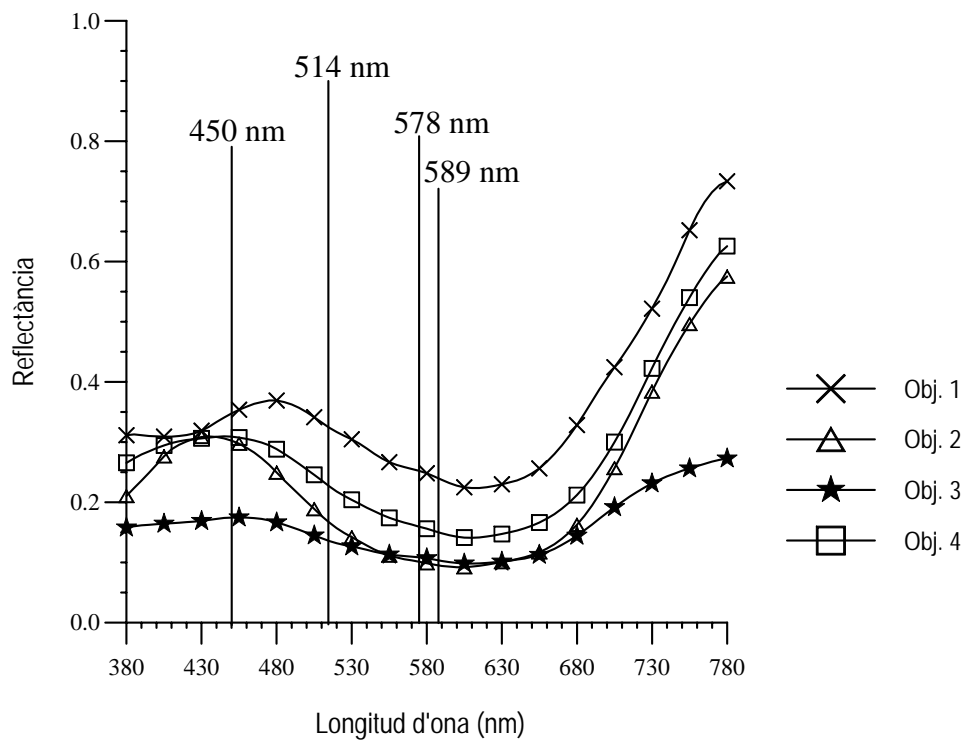
**Taula I.1.** Canals seleccionats per a la gamma marró (Figura I.1). Longituds d'ona  $\lambda_{fi}$  dels filtres interferencials escollides per a la gamma marró tenint en compte les  $\lambda_0$  on s'obtenen les màximes diferències proporcionals entre les corbes de reflectància dels colors inclosos en aquesta gamma.

Colors (n° obj)	Dif.Proporcional(%) $\left\{ \frac{ R_{1\lambda_0} - R_{2\lambda_0} }{\frac{1}{2}(R_{1\lambda_0} + R_{2\lambda_0})} \right\}$	Valor mig $\frac{R_{1\lambda_0} + R_{2\lambda_0}}{2}$	$\lambda_0$ (nm)	$\lambda_{fi}$ (nm)
1 – 2	66	0.10	405	450
1 – 3	105	0.32	670	671
1 – 4	104	0.21	575	578
2 – 3	70	0.21	600	589
2 – 4	73	0.13	560	550
3 – 4	23	0.07	430	450

## I.2 Gamma blava



**Figura I.3.** Escena de la gamma blava.

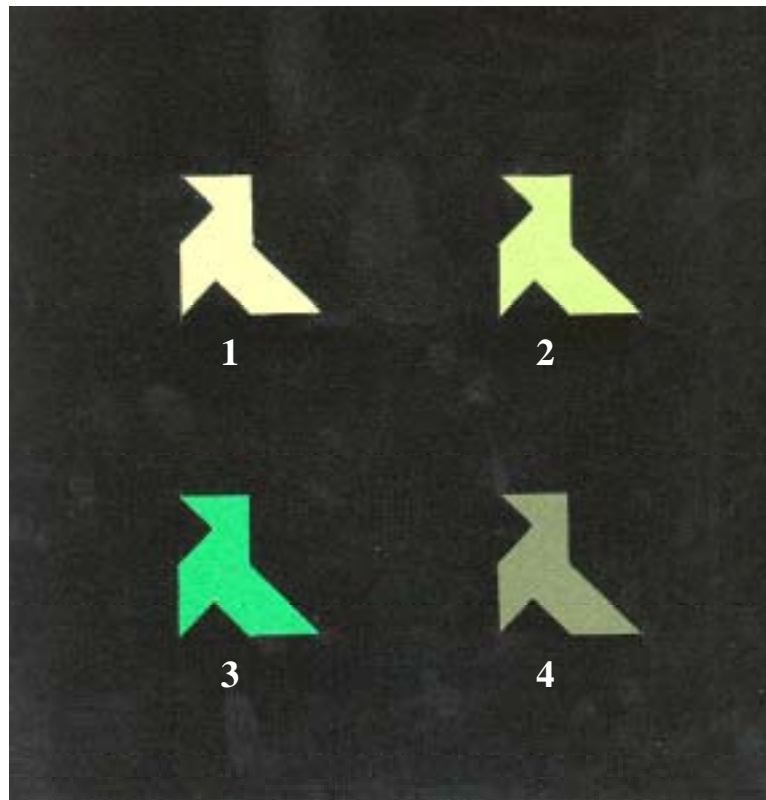


**Figura I.4.** Corbes de reflectància espectral de cada un dels objectes de la gamma blava ( Figura I.3). S'indica el conjunt de longituds d'ona dels filtres interferencials escollits ( $\lambda_{fi}$ ).

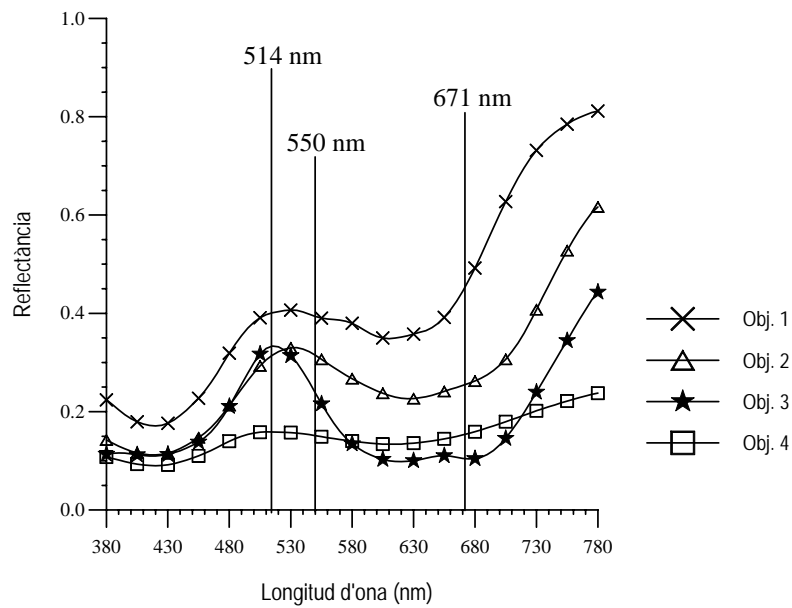
**Taula I.2.** Canals seleccionats per a la gamma blava ( Figura I.3). Longituds d'ona  $\lambda_{fi}$  dels filtres interferencials escollides per a la gamma blava tenint en compte les  $\lambda_0$  on s'obtenen les màximes diferències proporcionals entre les corbes de reflectància dels colors inclosos en aquesta gamma.

Colors (nº obj)	Dif.Proporcional(%) $\left\{ \frac{ R_{1\lambda_0} - R_{2\lambda_0} }{\frac{1}{2}(R_{1\lambda_0} + R_{2\lambda_0})} \right\}$	Valor mig $\frac{R_{1\lambda_0} + R_{2\lambda_0}}{2}$	$\lambda_0$ (nm)	$\lambda_{fi}$ (nm)
1 – 2	94	0.17	585	589
1 – 3	83	0.22	530	514
1 – 4	46	0.19	590	589
2 – 3	54	0.23	430	450
2 – 4	54	0.13	570	578
3 – 4	58	0.24	420	450

### I.3 Gamma verda



**Figura I.5.** Escena de la gamma verda.



**Figura I.6.** Corbes de reflectància espectral de cada un dels objectes de la gamma taronja ( Figura I.5). S'indica el conjunt de longituds d'ona dels filtres interferencials escollits ( $\lambda_{fi}$ ).

**Taula I.3.** Canals seleccionats per a la gamma verda ( Figura I.5). Longituds d'ona  $\lambda_{fi}$  dels filtres interferencials escollides per a la gamma verda tenint en compte les  $\lambda_0$  on s'obtenen les màximes diferències proporcionals entre les corbes de reflectància dels colors inclosos en aquesta gamma.

Colors (n° obj)	Dif.Proporcional(%) $\left\{ \frac{ R_{1\lambda_0} - R_{2\lambda_0} }{\frac{1}{2}(R_{1\lambda_0} + R_{2\lambda_0})} \right\}$	Valor mig $\frac{R_{1\lambda_0} + R_{2\lambda_0}}{2}$	$\lambda_0$ (nm)	$\lambda_{fi}$ (nm)
1 – 2	54	0.35	670	671
1 – 3	124	0.27	670	671
1 – 4	98	0.30	670	671
2 – 3	84	0.18	670	671
2 – 4	73	0.24	540	550
3 – 4	71	0.25	515	514



## **Annex II**

### **Publicacions relacionades amb el treball**

Relacionem a continuació les publicacions i contribucions a congressos sorgides a partir del treball de recerca realitzat i presentat en aquesta memòria.

#### **Publicacions en revistes arbitrades**

Pérez, E., Millán, M. S. (1997) “Adaptive channel selection for improving chromatic discrimination in colour pattern recognition”, *Opt. Commun.*, **134**, 273-280, ISSN 0030-4018.

Pérez, E., Chalasinska-Macukow, K., Styczynski, K., Kotynski, R., Millán, M. S. (1997) “Dual nonlinear correlator based on computer controlled joint transform processor: digital analysis and optical results”, *J. of Mod. Opt.*, **44** (8), 1535-1552, ISSN 0950-0340/97.



Millán, M. S., Pérez, E., Chalasinska-Macukow, K. (1998) "Pattern recognition with variable discrimination capability by dual non-linear optical correlation" Accepted at Optics Communications.

### Capítols de llibre

Campos, J., Yzuel, M. J., Yaroslavsky, L. P., Millán, M. S., Moreno, I., Corbalán, M., Pérez, E. (1997) "Channel transformations for improving color pattern recognition", a *Optoelectronic Information Processing*, B. Javidi i Ph. Réfrégier, eds. SPIE Press, Washington, pàgs. 105-135, ISBN 0-8194-2675-X.

### Publicacions en Proceedings

Millán, M. S., Corbalán, M., Pérez, E., Romero, J., Yzuel, M. J. (1996) "Improvement of the color discrimination capability in optical pattern recognition", *Second Iberoamerican Meeting in Optics*. Proc. SPIE, Vol. **2730**, 207-210, ISBN 0-8194-2111-1.

Chalasinska-Macukow, E. Pérez, K. Styczynski, R. Kotynski, M. S. Millán (1996) "Optoelectronic realization of dual nonlinear correlation" *XVII Congress of the International Commission for Optics (ICO)*, Proc. SPIE, **2778**, 455-456, ISBN 0-8194-2164-2.

Pérez, E., Millán, M. S., Chalasinska-Macukow, K. (1998) "Dual nonlinear correlation applied to textured and colour object recognition" *Electronic Imaging Processing, Printing and Publishing in Colour*, Proc. SPIE, **3409**, 444-455, ISBN 0277-786X/98.

Chalasinska-Macukow, K., Kotynski, R., Pérez, E., Millán, M. S. (1998) "Dual nonlinear correlation: optoelectronic realization and application to pattern recognition" Proc. SPIE of the International Conference on Optics and Optoelectronics (en premsa).

---

## Contribucions a congressos

### Internacionals

Millán, M. S., Corbalán, M., Pérez, E., Romero, J., Yzuel, M. J., “Improvement of the color discrimination capability in optical pattern recognition”, *Second Iberoamerican Meeting in Optics*, Guanajuato (Mèxic), setembre 1995. Comunicació oral.

Chalasinska-Macukow, K., Pérez, E., Styczynski, K., Kotynski, R., Millán, M. S., “Optoelectronic realization of dual nonlinear correlation”, *XVII Congress of the International Commission for Optics (ICO)*, Taejon (Corea), agost 1996. Comunicació oral.

Campos, J., Yzuel, M. J., Yaroslavski, L. P., Millán, M. S., Moreno, I., Corbalán, M., Pérez, E., “Channel transformations for improving color pattern recognition”, *II Euro-American Workshop on Optoelectronic Information Processing*, Sitges (Barcelona), juny 1997. Ponència convidada.

Pérez, E., Chalasinska-Macukow, K., Kotynski, R., Millán, M. S., “Optical pattern recognition by dual nonlinear correlation. Digital analysis of experimental conditions”, *II Euro-American Workshop on Optoelectronic Information Processing*, Sitges (Barcelona), juny 1997. Póster.

Pérez, E., Millán, M. S., Chalasinska-Macukow, K., “DNC applied to colour object recognition with tolerance to light shape and colour variations”, *Electronic Imaging Processing, Printing and Publishing in Colour*, Zürich (Suïssa), maig 1998. Póster.

Chalasinska-Macukow, K., Kotynski, R., Pérez, E., Millán, M. S. (1998) “Dual nonlinear correlation: optoelectronic realization and application to pattern recognition” *International Conference on Optics & Optoelectronics*. Dehradun (India), desembre 1998. Ponència convidada.

### Nacionals

M. S. Millán, J. Pladellorens, J. Escofet, M. Corbalán, E. Pérez, N. Vila (1994) “Procesado de imagen aplicado al reconocimiento en color y a la inspección textil en la Universidad Politécnica de Cataluña”, Actes de la IV Reunión Nacional de Óptica celebrada a Granada, pp. 3-4, ISBN 84-338-1931-3. Comunicació oral.

M. S. Millán, J. Pladellorens, J. Escofet, M. Corbalán, E. Pérez, N. Vila (1994) “Adquisición y manipulación de la información óptica en sistemas de reconocimiento de formas en color” Actes de les III Jornadas de Visión Artificial celebrades a Màlaga, pp. 241-249. Comunicació oral.

E. Pérez, M. S. Millán (1995) “Narrow-band multichannel optical correlation to improve colour object discrimination in pattern recognition” Actes del VI Simposium Nacional de Reconocimiento de Formas y Análisis de Imágenes celebrat a Còrdova, pp. 731-732. Pòster.

E. Pérez, K. Chalasinska-Macukow, R. Kotynski, M. S. Millán (1997) “Reconocimiento óptico de objetos mediante correlación no lineal dual. Análisis digital de condiciones experimentales”, Actes de la V Reunión Nacional de Óptica celebrada a València, pp. 137-138. Pòster.

### **Treball de recerca de 3<sup>er</sup> cicle**

Pérez, E. (1995) “Selecció de canals quasimonocromàtics per a la millora de la capacitat de discriminació cromàtica en el reconeixement de formes per correlació òptica” Treball de recerca de 3<sup>er</sup> cicle, Universitat Politècnica de Catalunya.

## Bibliografia

- [1] Vander Lugt, A. B. (1964) "Signal detection by complex filtering", IEEE, Trans. Inf. Theory, IT-10, 139-145.
- [2] Weaver, C. S., Goodman, J. W. (1966) "A technique for optically convolving two functions", Appl. Opt., 5 (7), 1248-1249.
- [3] Rau, J. E. (1966) "Detection of differences in real distributions", J. Opt. Soc. Am., 56 (11), 1490-1494.
- [4] Horner, J. L., Gianino, P. D. (1984) "Phase-only matched filtering", Appl. Opt., 23 (6), 812-816.
- [5] Mu, G. G., Wang, X. M., Wang, Z. Q. (1988) "Amplitude-compensated matched filtering", Appl. Opt., 27 (16), 3461-3463.
- [6] Vijaya Kumar, B. V. K., Hasebrook, L. (1990) "Performance measures for correlation filters", Appl. Opt., 29 (20), 2997-3006.
- [7] Horner, J. L., Leger, J. (1985) "Pattern recognition with binary phase-only filters", Appl. Opt., 24 (5), 609-611.
- [8] Flannery, D. L., Loomis, J. S. i Milkovich, M. E. (1988) "Transform-ratio ternary phase-amplitude filter formulation for improved correlation discrimination", Appl. Opt., 27 (19), 4079-4083.

- [9] Mahalanobis, A., Vijaya Kumar, B. V. K., Casasent, D. (1987) "Minimum average correlation energy filters", *Appl. Opt.*, **26** (17), 3633-3640.
- [10] Campos, J., Turon, F., Yaroslavsky, L. P., Yzuel, M. J. (1991) "Some filters for reliable recognition and localization of objects by optical correlators: a comparison", *Int. J. Opt. Comp.* **2**, 341-365.
- [11] Feng, D., Zhao, H., Xia, S. (1991) "Amplitude-modulated JTC for improving correlation discrimination", *Opt. Commun.*, **86**, 260-264.
- [12] Inbar, H., Marom, E. (1993) "Matched, phase-only, or inverse filtering with joint-transform correlators", *Opt. Lett.*, **18** (19), 1657-1659.
- [13] Gualdron, O., Arsenault, H. (1993) "Locally nonlinear matched filtering", *Opt. Lett.*, **18** (22), 1949-1951.
- [14] Javidi, B., Kuo, C. J. (1988) "Joint transform image correlation using a binary spatial light modulator at the Fourier plane", *Appl. Opt.*, **27** (4), 663-665.
- [15] Javidi, B., Wang, J. (1991) "Binary nonlinear joint transform correlation with median and subset median thresholding", *Appl. Opt.*, **30** (8), 967-976.
- [16] Javidi, B., Wang, J., Tang, Q. (1991) "Multiple-object binary joint transform correlation using multiple-level threshold crossing", *Appl. Opt.*, **30** (29), 4234-4244.
- [17] Wang, J., Javidi, B. (1994) "Multiobject detection using the binary joint transform correlator with different types of thresholding methods", *Opt. Eng.*, **33** (6), 1793-1805.
- [18] Rogers, S. K., Cline, J. D., Kabrisky, M., Mills, J. P. (1990) "New binarization techniques for joint transform correlation", *Appl. Opt.*, **29** (9), 1088-1093.
- [19] Vallmitjana, S., Juvells, I., Carnicer, A. (1992) "Evaluation of a suitable threshold for binarization of power spectrum in a noise-free joint transform correlator", *Opt. Commun.*, **90**, 221-226.
- [20] Inbar, H., Mendlovic, D., Marom, E. (1993) "Error-diffusion binarization for joint transform correlators", *Appl. Opt.*, **32** (5), 707-714.
- [21] Alam, M. S., Karim, M. A. (1993) "Fringe-adjusted joint transform correlation", *Appl. Opt.*, **32** (23), 4344-4350.

- 
- [22] Alam, M. S., Karim, M. A. (1993) "Joint-transform correlation under varying illumination", *Appl. Opt.*, **32** (23), 4351-4356.
- [23] Michel, H. E., Awwal, A. A. (1998) "Joint Fourier transform correlation with phase thresholding in the Fourier domain", *Opt. Eng.*, **37** (1), 33-37.
- [24] Vallmitjana, S., Carnicer, A., Martín-Badosa, E., Juvells, I. (1995) "Nonlinear filtering in object and Fourier space in a joint transform optical correlator: comparison and experimental realization", *Appl. Opt.*, **34** (20), 3942-3949.
- [25] Lu, G., Yu, F. T. S. (1996) "Performance of a phase-transformed input joint transform correlator", *Appl. Opt.*, **35** (2), 304-313.
- [26] Erbach, P. S., Gregory, D. A., Hammock, J. B. (1996) "Phase-only joint-transform correlator", *Appl. Opt.*, **35** (17), 3091-3096.
- [27] Ledesma, S., Iemmi, C., Campos, J., Yzuel, M. J. (1998) "Joint transform correlator architecture with a single LCTV operating in phase-mostly mode", *Opt. Commun.*, **151**, 101-109.
- [28] Styczynski, K., Campos, J., Yzuel, M. J., Chalasinska-Macukow, K. (1995) "New arrangement for limited intensity invariance pattern recognition with high diffraction efficiency", *Opt. Commun.*, **118**, 193-198.
- [29] Ersoy, O. K., Zeng, M. (1989) "Nonlinear matched filtering", *J. Opt. Soc. Am., A*, **6** (5), 636-648.
- [30] Rosen, J., Kotzer, T., Shamir, J. (1991) "Optical implementation of phase extraction pattern recognition", *Opt. Commun.*, **83** (1,2), 10-14.
- [31] Kotzer, T., Rosen, J., Shamir, J. (1992) "Phase extraction pattern recognition", *Appl. Opt.*, **31** (8), 1126-1137.
- [32] Chalasinska-Macukow, K., Baranska, E. (1990) "Discrimination of character using phase information only", *J. Optics*, **21**, 261-266.
- [33] Chalasinska-Macukow, K., Turon, F., Yzuel, M. J., Campos, J. (1990) "Performance of the pure phase-only correlation method for pattern recognition", *Proc. SPIE*, **1347**, 262-273.

- [34] Yu, F. T. S., Cheng, F., Nagata, T., Gregory, D. A. (1989) "Effects of fringe binarization of multiobject joint transform correlation", *Appl. Opt.*, **28** (15), 2988-2990.
- [35] Davis, J. A., Merrill, E. A., Cottrell, D. M., Bunch, R. M. (1990), "Effects of sampling and binarization in the output of the joint Fourier transform correlator", *Opt. Eng.*, **29** (9), 1094-1100.
- [36] Chalasinska-Macukow, K., Turon, F., Yzuel, M. J., Campos, J. (1991) "Intermodulation effects in pure phase-only correlation method", *Proc. SPIE*, **1564**, 285-293.
- [37] Tanone, A., Uang, C. M., Yu, F. T. S., Tam, E. C., Gregory, D. A. (1992) "Effects of thresholding in joint-transform correlation", *Appl. Opt.*, **31** (23), 4816-4822.
- [38] Turon, F., Ahouzi, E., Campos, J., Chalasinska-Macukow, K., Yzuel, M. J. (1994) "Nonlinearity effects in the pure phase correlation method in multiobject scenes", *Appl. Opt.*, **33** (11), 2188-2191.
- [39] Carnicer, A., Juvells, I., Vallmitjana, S. (1992) "Effects of thresholding level variation in fringe binarization of multiobject joint transform correlation", *Appl. Opt.*, **31** (8), 1012-1014.
- [40] Kotzer, T., Rosen, J., Shamir, J. (1993) "Multiple-object input in nonlinear correlation", *Appl. Opt.*, **32** (11), 1919-1932.
- [41] Carnicer, A., Martín-Badosa, E., Juvells, I., Vallmitjana, S. (1995) "Spatial envelop-free nonlinear joint transform correlation" *Opt. Commun.*, **114**, 336-343.
- [42] Kotynski, R., Chalasinska-Macukow, K. (1996) "Optical correlator with dual nonlinearity", *J. Mod. Opt.*, **43** (2), 295-310.
- [43] Lu, G., Zhang, Z., Wu, S., Yu, F. T. S. (1997) "Implementation of a non-zero-order joint transform correlator by use of phase-shifting techniques", *Appl. Opt.*, **36** (2), 470-483.
- [44] Li, C. T., Yin, S., Yu, F. T. S. (1998) "Non-zero order joint transform correlator", *Opt. Eng.*, **37** (1), 58-65.

- 
- [45] Yu, F. T. S., Li, C. T., Yin, S. (1998) "Comparison of detection efficiency for nonzero-order and conventional joint transform correlator", *Opt. Eng.*, **37** (1), 52-57.
- [46] Alam, M. S., Karim, M. A. (1994) "Multiple target detection using a modified fringe-adjusted joint transform correlator", *Opt. Eng.*, **33** (5), 1610-1617.
- [47] Tang, Q., Javidi, B. (1993) "Technique for reducing the redundant and self-correlation terms in joint transform correlators", *Appl. Opt.*, **32** (11), 1911-1918.
- [48] Grycewicz, T. J. (1995) "Fourier-plane windowing in the binary joint transform correlator for multiple target detection", *Appl. Opt.*, **34** (20), 3933-3941.
- [49] Réfrégier, Ph. (1990) "Filter design for optical pattern recognition: multicriteria optimization approach", *Opt. Lett.*, **15** (15), 854-856.
- [50] Réfrégier, Ph. (1991) "Optimal trade-off filters for noise robustness, sharpness of the correlation peak, and Horner efficiency", *Opt. Lett.*, **16** (11), 829-831.
- [51] Figue, J., Réfrégier, Ph. (1992) "Influence of the noise model on correlation filters: peak sharpness and noise robustness", *Opt. Lett.*, **17** (20), 1476-1478.
- [52] Réfrégier, Ph., Laude, V., Javidi, B. (1994) "Nonlinear joint Fourier transform correlation: an optimal solution for adaptive image discrimination and input noise robustness", *Opt. Lett.*, **19**, 405-407.
- [53] Réfrégier, Ph., Laude, V., Javidi, B. (1995) "Basic properties of nonlinear global filtering techniques and optimal discriminant solutions", *Appl. Opt.*, **34** (20), 3915-3923.
- [54] García, J. Szoplik, T., Ferreira, C. (1992) "Shift-and-scale invariant pattern recognition using an elliptic coordinate-transformed phase-only filter", *Appl. Opt.*, **31** (23), 4823-4828.
- [55] Cojoc, D., Molina, M. T., García, J., Ferreira, C. (1997) "Coordinate-transformed filter for shift-invariant and scale-invariant pattern recognition", *Appl. Opt.*, **36** (20), 4812-4815.



- [56] Arsenault, H. H., Hsu, Y. N. (1983) "Rotation-invariant discrimination between almost similar objects", *Appl. Opt.*, **22**, 130-132.
- [57] Arsenault, H. H., Hsu, Y. N., Chalasinska-Macukow, K. (1984) "Rotation-invariant pattern recognition", *Opt. Eng.*, **23**, 705-709.
- [58] Arsenault, H. H., Sheng, Y. (1986) "Properties of the circular harmonic expansion for rotation-invariant pattern recognition", *J. Opt. Soc. Am. A*, **25**, 3225-3229.
- [59] Arsenault, H. H., Ferreira, C., Levesque, M., Szoplik, T. (1986) "Simple filter with limited rotation invariance", *J. Opt. Soc. Am. A*, **25**, 3230-3234.
- [60] García-Martínez, P., García, J., Ferreira, C. (1995) "A new criterion for determining the expansion center of circular harmonic filters", *Opt. Commun.*, **117**, 339.
- [61] Arsenault, H. H., Belisle, C. (1985) "Contrast-invariant pattern recognition using circular harmonics", *Appl. Opt.*, **24**, 2072-2075.
- [62] Honiball, H. R., Roux, F. S. (1994) "Correlation peak degradation for out-of-plane rotation in an optical pattern recognition system", *Opt. Eng.*, **33** (9), 2849-2853.
- [63] Moya, A., Mendlovic, D., García, J., Ferreira, C. (1996) "Projection-invariant pattern recognition with a phase-only logarithmic-harmonic-derived filter", *Appl. Opt.*, **35** (20), 3862-3867.
- [64] Caulfield, H. J., Maloney, W. T. (1969) "Improved discrimination in optical character recognition", *Appl. Opt.*, **8** (11), 2354-2356.
- [65] Caulfield, H. J. (1980) "Linear combinations of filters for character recognition: a unified treatment", *Appl. Opt.*, **19**, 3877-3879.
- [66] Hester, C. F., Casasent, D. (1980) "Multivariant technique for multiclass pattern recognition", *Appl. Opt.*, **19** (11), 1758-1761.
- [67] Casasent, D. (1984) "Unified synthetic discriminant function computational formulation", *Appl. Opt.*, **23** (10), 1620-1627.
- [68] Vijaya Kumar, B. V. K. (1992) "Tutorial survey of composite filter designs for optical correlators", *Appl. Opt.*, **31** (23), 4773-4801.

- 
- [69] Javidi, B., Wang, W., Zhang, G. (1997) "Composite Fourier-plane nonlinear filter for distortion-invariant pattern recognition", *Opt. Eng.*, **36** (10), 2690-2696.
- [70] Hassebrook, L. G., Lhamon, M. E., Wang, M., Chatterjee, J. P. (1997) "Postprocessing of correlation for orientation estimation", *Opt. Eng.*, **36** (10), 2710-2718.
- [71] Vallmitjana, S., Juvells, I., Carnicer, A., Campos, J. (1991) "Pattern detection using a modified composite filter with nonlinear joint transform correlator", *Proc. SPIE*, **1564**, 266-274.
- [72] Tang, Q., Javidi, B. (1992) "Sensitivity of the nonlinear joint transform correlator: experimental investigations", *Appl. Opt.*, **31** (20), 4016-4024.
- [73] García, J., Szoplik, T., Ferreira, C. (1993) "Optoelectronic morphological image processor", *Opt. Lett.*, **18**, 1952-1954.
- [74] García-Martínez, P. Mas, D., García, J., Ferreira, C. (1998) "Nonlinear morphological correlation: optoelectronic implementation", *Appl. Opt.*, **37** (11), 2112-2118.
- [75] Khoury, J., Kane, J. S., Asimellis, G., Gronin-Golom, M., Woods, C. (1994) "All-optical nonlinear joint Fourier transform correlator" *Appl. Opt.*, **33** (35), 8216-8225.
- [76] Sánchez-de-la-Llave, J. D., Pommet, D. A., Fiddy, M. A. (1998) "Novel joint transform correlator architecture using bacteriorhodopsin optically addressable spatial light modulators", *Opt. Eng.*, **37** (1), 27-32.
- [77] Grycewicz, T. J. (1997) "Lensless joint transform correlator", *Opt. Eng.*, **36**(10), 2781-2786.
- [78] Comastri, S. A., Simon, J. M., Fucci, F. (1997) "Joint transform correlator: aberrations in the joint spectrum and their influence on the correlation peak", *Pure Appl. Opt.*, **6**, 173-189.
- [79] Le Grand, Y. (1972) "Optique physiologique", Vol. II *Lumière et couleurs*. Masson, Paris. Capítol 7.

- [80] Vargas, A., Campos, J., Navarro, R. (1995) "An application of the Gabor multiscale decomposition of an image to pattern recognition", Proc. SPIE, **2730**, 622-625.
- [81] Vargas, A., Campos, J., Navarro, R. (1996) "Non-linear process applied to multiple correlations in pattern recognition of defocused images", Proc. SPIE, **2778**, 519-520.
- [82] Vargas, A. (1998) "Correlador óptico multicanal. Aplicación simultánea de bancos de filtros para análisis de imagen y reconocimiento de formas". Tesis Doctoral, Universitat Autònoma de Barcelona.
- [83] Badiqué, E., Komiya, Y., Ohya, N., Tsujiuchi, J., Honda, T. (1987) "Color image correlation", Opt. Commun., **61**, 181-186.
- [84] Badiqué, E., Komiya, Y., Ohya, N., Honda, T., Tsujiuchi, J. (1988) "Use of color image correlation in retrieval of gastric surface topography by endoscopic stereopair matching", Appl. Opt., **27**, 941.
- [85] Badiqué, E., Ohya, N., Honda, T., Tsujiuchi, J. (1988) "Color image correlation for spatial/spectral recognition and increased selectivity", Opt. Commun., **68**, 91-96.
- [86] Badiqué, E., Ohya, N., Honda, T., Tsujiuchi, J. (1988) "A complex synthetic discriminant filter for the parallel recognition of color patterns", Opt. Commun., **67** (5), 335-340.
- [87] Pei, S. C., Hsu, M. J. (1993) "Two-dimensional invariant color pattern recognition using a complex log mapping", Opt. Eng., **32** (7), 1616-1622.
- [88] Yu, F. T. S. (1985) "White-light optical signal processing". John Wiley and sons, New York.
- [89] Yu, F. T. S. (1986) "Principle of optical processing with partially coherent light", a Progress in Optics, Vol. **23**. E. Wolf, editor, North-Holland Pub., Amsterdam.
- [90] Yu, F. T. S. (1984) "Color image recognition by spectral-spatial matched filtering", Opt. Eng., **23** (6), 690-694.
- [91] Yu, F. T. S., Hsu, F. K. (1985) "Complex signal detection with broad spectral band processing", Appl. Opt., **24**, 2135.

- 
- [92] Yu, F. T. S., Javidi, B. (1986) "Experiments on real-time polychromatic signal detection by matched spatial filtering", *Opt. Commun.*, **56** (6), 384-388.
- [93] Javidi, B., Yu, F. T. S. (1986) "Real-time image recognition by partially coherent light", *Appl. Opt.*, **25** (14), 2365-2371.
- [94] Warde, C., Caulfield, H. J., Yu, F. T. S., Ludman, J. E. (1984) "Real-time joint spectral-spatial matched filtering", *Opt. Commun.*, **49**, 241.
- [95] Ludman, J. E., Javidi, B., Yu, F. T. S., Caulfield, H. J., Warde, C. (1984) "Real-time color-pattern recognition", *Proc. SPIE*, **465**, 143-149.
- [96] Javidi, B., Kuo, C. J., Chen, Y. F., Ludman, J. E. (1988) "Color object identification by monochromatic binary correlation", *Appl. Opt.*, **27** (5), 949-953.
- [97] Yu, F. T. S., Jutamulia, S., Lin, T. W. (1987) "Real-time polychromatic signal detection using a color liquid crystal television" *Opt. Eng.*, **26**, 453-459.
- [98] Yu, F. T. S., Jutamulia, S., Yelamarty, R. V., Gregory, D. A. (1989) "Adaptive joint transform correlator for real-time color pattern recognition", *Opt. Laser Technol.*, **21**, 189-192.
- [99] Wang, Z. Q., Cartwright, C. M., Soutar, C., Gillespie, W. A. (1993) "Real-time color image correlation with a color liquid-crystal television and a Fresnel holographic filter", *Appl. Opt.*, **32** (5), 715-717.
- [100] Gu, Z. H., Lee, S. H., Fainman, Y. (1987) "Statistical recognition of color images", *Appl. Opt.*, **26** (15), 3145-3152.
- [101] Millán, M. S., Campos, J., Ferreira, C., Yzuel, M. J. (1989) "Matched filter and phase only filter performance in colour image recognition", *Opt. Commun.*, **73** (4), 277-284.
- [102] Ferreira, C., Millán, M. S., Yzuel, M. J., Campos, J. (1992) "Experimental results in color pattern recognition by multichannel matched filtering", *Opt. Eng.*, **31** (10), 2231-2238.
- [103] García, J., Ferreira, C., Campos, J., Bosch, S., Millán, M. S., Yzuel, M. J. (1992) "Multichannel rotation-invariant pattern recognition for polychromatic objects using circular harmonic filters", *Opt. Commun.*, **91**, 425-432.

- [104] Millán, M. S., Yzuel, M. J., Campos, J., Ferreira, C. (1992) "Different strategies in optical recognition of polychromatic images", *Appl. Opt.*, **31** (14), 2560-2567.
- [105] Millán, M. S., Vila, N., Yzuel, M. J. (1992) "Colour reversal films in multichannel optical correlators for polychromatic image recognition", *Pure Appl. Opt.*, **1**, 199-218.
- [106] Corbalán, M., Millán, M. S., Yzuel, M. J. (1996) "Color image acquisition by charge-coupled-device cameras in polychromatic pattern recognition", *Opt. Eng.*, **35** (3), 754-760.
- [107] Zhu, Z. (1994) "Color pattern recognition in a image system with chromatic distortion", *Opt. Eng.*, **33** (9), 3047-3051.
- [108] García, J., Campos, J., Ferreira, C. (1994) "Multichannel color pattern recognition using a minimum average correlation energy filter", *Appl. Opt.*, **33**, 221-224.
- [109] Lancis, J., Tam, E., Pons, A., Andrés, P. (1993) "Achromatic white-light joint transform correlator: theory and experiment", *Proc. SPIE*, **1983**, 419-420.
- [110] García, J., Campos, J., Ferreira, C. (1994) "Circular-harmonic minimum average correlation energy filter for color pattern recognition", *Appl. Opt.*, **33** (11), 2180-2187.
- [111] Deutsch, M., García, J., Mendlovic, D. (1996) "Multichannel single-output color pattern recognition by use of a joint-transform correlator", *Appl. Opt.*, **35** (35), 6976-6982.
- [112] Zalevsky, Z., Raveh, I., Shabtay, G., Mendlovic, D. (1997) "Single-output color pattern recognition using a fractional correlator", *Opt. Eng.*, **36** (8), 2127-2136.
- [113] Yamaba, K., Miyake, Y. (1993) "Color character recognition method based on human perception", *Opt. Eng.*, **32** (1), 33-40.
- [114] Millán, M. S., Corbalán, M., Romero, J., Yzuel, M. J. (1995) "Optical pattern recognition based on color vision models", *Opt. Lett.*, **20** (16), 1722-1724.

- 
- [115] Hurvich, L. M., Jameson, D. (1957) "An opponent-process theory of color vision", *Psychol. Rev.*, **64**, 384-402.
- [116] Guth, S. L., Massof, R. W., Benzschawel, T. (1980) "Vector model for normal and dichromatic color vision", *J. Opt. Soc. Am.*, **70**, 197-212.
- [117] Boynton, R. M. (1986) "A system of photometry and colorimetry based on cone excitations", *Color Res. & Appl.*, **11**, 244-252.
- [118] Corbalán, M. (1997) "Métodos de procesado de la información en color para la mejora del reconocimiento de objetos por correlación óptica". *Tesi Doctoral*, Universitat Politècnica de Catalunya.
- [119] Mendlovic, D., García-Martínez, P., García, J., Ferreira, C. (1995) "Color encoding for polychromatic single channel optical pattern recognition", *Appl. Opt.*, **34** (32), 7538-7544.
- [120] Mendlovic, D., Deutsch, M., Ferreira, C., García, J. (1996) "Single channel polychromatic pattern recognition using a joint transform correlator", *Appl. Opt.*, **35** (32), 6382-6389.
- [121] Moreno, I., Kober, V., Lashin, V., Campos, J., Yaroslavski, L. P., Yzuel, M. J. (1996) "Color pattern recognition with circular component whitening", *Opt. Lett.*, **21** (7), 498-500.
- [122] Kober, V., Lashin, V., Moreno, I., Campos, J., Yaroslavsky, L. P., Yzuel, M. J. (1997) "Color component transformation for optical pattern recognition", *J. Opt. Soc. Am. A*, **14** (10), 2656-2669.
- [123] Caulfield, H. J., Moreno, I., Campos, J., Yzuel, M. J. (1997) "Coherent recognition of colored patterns", *Opt. Commun.*, **133**, 77-81.
- [124] Moreno, I., Campos, J., Yzuel, M. J., Kober, V. (1998) "Implementation of bipolar real-valued input scenes in a real-time optical correlator: application to color pattern recognition", *Opt. Eng.*, **37** (1), 144-150.
- [125] Fielding, K. H., Horner, J. L., Makekau, C. K. (1990) "Modified two-focal-length optical correlator", *Appl. Opt.*, **29** (29), 4332-4335.
- [126] Yu, F. T. S., Song, Q. W., Cheng, Y. S. i Gregory, D. A. (1990) "Comparison of detection efficiencies for VanderLugt and joint transform correlators", *Appl. Opt.*, **29** (2), 225-232.

- [127] Goodman, J. W. (1968) "Introduction to Fourier optics", McGraw-Hill, Inc., New York.
- [128] Gaskill, J. D. (1987) "Linear systems, Fourier transforms, and optics", John Wiley & Sons, New York.
- [129] Casasent, D., editor (1978) "Optical data processing. Applications", Springer-Verlag, Berlín.
- [130] Lee, S. H., editor (1981) "Optical information processing. Fundamentals". Springer-Verlag, Berlín.
- [131] Goodman, J. W. (1968) "Introduction to Fourier optics", McGraw-Hill, Inc., New York. Capítol 5.
- [132] Casasent, D., editor (1978) "Optical data processing. Applications", Springer-Verlag, New York. Pàg. 23.
- [133] Javidi, B., Horner, J. L. (1989) "Single spatial light modulator joint transform correlator", *Appl. Opt.*, **28** (5), 1027-1032.
- [134] Javidi, B., Gregory, D. A., Horner, J. L. (1989) "Single modulator joint transform correlator architectures", *Appl. Opt.*, **28** (3), 411-413.
- [135] Yu, F. T. S., Gregory, D. A. (1996) "Optical pattern recognition: architectures and techniques", *Proc. IEEE*, **84** (5), 733-752.
- [136] Fielding, K. H., Horner, J. L. (1990), "1-f binary joint transform correlator", *Opt. Eng.*, **29** (9), 1081-1087.
- [137] Horner, J. L., Makekai, C. K. (1989) "Two-focal-length optical correlator", *Appl. Opt.*, **28** (24), 5199-5201.
- [138] Purwosumarto, P., Yu, F. T. S. (1997) "Robustness of joint transform correlator versus VanderLugt correlator", *Opt. Eng.*, **36** (10), 2775-2780.
- [139] Yu, F. T. S., Lu, X. J. (1984) "A real-time programmable joint transform correlator", *Opt. Commun.*, **52** (1), 10-16.
- [140] Yu, F. T. S., Jutamulia, S., Lin, T. W., Gregory, D. A. (1987) "Adaptive real-time pattern recognition using a liquid crystal TV based joint transform correlator", *Appl. Opt.*, **26**, 1370.

- 
- [141] Lu, X. J., Yu, F. T. S., Gregory, D. A. (1990) "Comparison of Vander Lugt and joint transforms correlators", *Appl. Phys. B*, **51**, 153-164.
- [142] Jutamulia, S. (1992) "Joint transform correlators and their applications", *Proc. SPIE*, **1812**, 233-243.
- [143] Gregory, D. A., Loudin, J. A. (1989) "Joint transform correlator limitations", *Proc. SPIE*, **1053**, 198-207.
- [144] Vander Lugt, A. B., Rotz, F. B. i Klooster, A. (1965) "Optical and electro-optical information processing". Eds. J. T. Tippett et al., MIT Press, Cambridge.
- [145] Vander Lugt, A. (1967) "The effects of small displacements of spatial filters", *Appl. Opt.*, **6**, 1221.
- [146] Casasent, D., Furman, A. (1977) "Sources of correlation degradation", *Appl. Opt.*, **16** (6), 1652-1661.
- [147] Vijaya Kumar, B. V. K., Pochapsky, E. (1986) "Signal to noise ratio considerations in modified matched spatial filters", *J. Opt. Soc. Am. A*, **3**, 777.
- [148] Gianino, P. D., Horner, J. L. (1984) "Additional properties of the phase-only filter", *Opt. Eng.*, **23**, 695.
- [149] Zhou, H., Zhao, F., Yu, F. T. S., Chao, T. (1993) "Improved interclass multiobject discrimination with phase-difference prewhitening technique", *Opt. Eng.*, **32** (11), 2720-2725.
- [150] Ahouzi, E., Campos, J., Yzuel, M. J. (1994) "Phase only filter with improved discrimination", *Opt. Lett.*, **19** (17), 1340-1342.
- [151] Campos, J., Bosch, S., Sallent, J. (1991) "Optical pattern recognition in defocused image using correlations filters", *Opt. Commun.*, **82** (3,4), 370-379.
- [152] Awwal, A. A. S., Karim, M. A., Jahan, S. R. (1990) "Improved correlation discrimination using an amplitude-modulated phase-only filter", *Appl. Opt.*, **29** (2), 233-236.
- [153] Cohn, R. (1993) "Adaptive real-time architectures for phase-only correlation", *Appl. Opt.*, **32** (5), 718-725.



- [154] Ahouzi, E., Campos, J., Chalasinska-Macukow, K., Yzuel, M. J. (1994) "Optoelectronic pure phase correlator", *Opt. Commun.*, **110**, 27-32.
- [155] Ahouzi, E., Iemmi, C., Ledesma, S., Lashin, V., Chalasinska-Macukow, K., Campos, J., Yzuel, M. J. (1997) "Pattern recognition with a phase-shifting interferometric correlator. Discrimination-capability enhancement", *Appl. Phys. B*, **64**, 331-338.
- [156] Chalasinska-Macukow, K., Gorecki, C. (1992) "Optoelectronic implementation of the quasi-phase correlator", *Opt. Commun.*, **93**, 11-18.
- [157] Martín-Badosa, E., Carnicer, A., Juvells, I., Vallmitjana, S. (1997) "Complex modulation characterization of liquid crystal devices by interferometric data correlation", *Meas. Sci. Technol.*, **8**, 764-772.
- [158] Javidi, B. (1989) "Nonlinear joint power spectrum based optical correlator", *Appl. Opt.*, **28** (12), 2358-2367.
- [159] Javidi, B., Tang, Q., Gregory, D. A., Hudson, T. D. (1991) "Experiments on nonlinear joint transform correlator using an optically addressed spatial light modulator in the Fourier plane", *Appl. Opt.*, **30** (14), 1772-1776.
- [160] Javidi, B. (1990) "Generalization of the linear matched filter concept to nonlinear matched filters", *Appl. Opt.*, **29** (8), 1215-1224.
- [161] Javidi, B., Wang, J., Ruiz, C., Ruiz, J. (1991) "Nonlinearity transformed baseband filters for optical pattern recognition", *Appl. Opt.*, **30** (14), 1776-1780.
- [162] Javidi, B., Zhang, G. (1992) "Experiments on nonlinearly transformed matched filters", *Opt. Eng.*, **31** (5), 934-938.
- [163] Javidi, B. (1990) "Comparison of the nonlinear JTC and the nonlinearly transformed matched filter based correlator for noisy input scenes", *Opt. Eng.*, **29** (9), 1013-1020.
- [164] Javidi, B., Wang, J., Tang, Q. (1994) "Nonlinear joint transform correlators", *Pattern Recognition*, **27** (4), 523-542.
- [165] Javidi, B. (1994) "Nonlinear joint transform correlators", in *Real-Time Optical Information Processing*, ed. B. Javidi i J. L. Horner. Academic Press, Inc., San Diego, California. Capítol 4.

- 
- [166] Javidi, B., Gianino, P. D., Rhodes, W. T., Hendrickson, B. M. (1995) "Nonlinear techniques in optical pattern recognition: introduction by the feature editors", *Appl. Opt.*, **34** (20), 3856-3857.
- [167] Horner, J. L. (1982) "Light utilization in optical correlators", *Appl. Opt.*, **21** (24), 4511-4514.
- [168] Pérez, E., Millán, M. S. (1997) "Adaptive channel selection for improving chromatic discrimination in colour pattern recognition", *Opt. Commun.*, **134**, 273-280.
- [169] Goodman, J. W. (1968) "Introduction to Fourier optics", McGraw-Hill, Inc., New York. Pàg. 181.
- [170] Boyle, W. S., Smith, G. E. (1970) "Charge coupled semiconductor devices", *Bell Systems Technical Journal*, **49**, 587-593.
- [171] Holst, G. C. (1996) "CCD arrays, cameras and displays", SPIE Optical Engineering Press, Washington.
- [172] Corbalán, M., Millán, M. S., Yzuel, M. J. (1998) "Analysis of the compensating mechanism of a colour CCD camera for changes of light", *Proc. SPIE*, **3409**, 456-466.
- [173] Efron, U. Editor (1995) "Spatial light modulator technology. Materials, devices, and applications" Marcel Dekker, Inc., New York.
- [174] Goodman, J. W. (1996) "Introduction to Fourier optics", McGraw-Hill, Inc., New York. Segona edició revisada.
- [175] Boreman, G. D., Raudenbush, E. R (1988) "Modulation depth characteristics of a liquid crystal television spatial light modulator", *Appl. Opt.*, **27** (14), 2940-2943.
- [176] Lu, K., Saleh, B. E. A. (1990) "Theory and design of the liquid crystal TV as an optical spatial phase modulator", *Opt. Eng.*, **29** (3), 240-246.
- [177] Kirsch, J. C., Gregory, D. A., Thie, M. W., Jones, B. K. (1992) "Modulation characteristics of the Epson liquid crystal television", *Opt. Eng.*, **31** (5), 963-970.

- [178] Laude, V., Mazé, S., Chavel, P., Réfrégier, Ph. (1992) "Amplitude and phase coding measurements of a liquid crystal television", *Opt. Commun.*, **103**, 33-38.
- [179] Soutar, C., Monroe Jr., S. E., Knopp, J. (1994) "Measurement of the complex transmittance of the Epson liquid crystal television", *Opt. Eng.*, **33** (4), 1061-1068.
- [180] Soutar, C., Lu, K. (1994) "Determination of the physical properties of an arbitrary twisted-nematic liquid crystal cell", *Opt. Eng.*, **33** (8), 2704-2712.
- [181] Dou, R., Giles, M. K. (1996) "Simple technique for measuring the phase property of a twisted nematic liquid crystal", *Opt. Eng.*, **35** (3), 808-812.
- [182] McClain, J. L., Erbach, P. S., Gregory, D. A., Yu, F. T. S. (1996) "Spatial light modulator phase depth determination from optical diffraction information", *Opt. Eng.*, **35** (4), 951-954.
- [183] Moreno, I. (1996) "Diseño de sistemas multicanales en tiempo real para el reconocimiento de objetos policromáticos". *Tesi Doctoral, Universitat Autònoma de Barcelona.*
- [184] Martín Badosa, E. (1998) "Correlador óptico para el reconocimiento de objetos basado en las propiedades de modulación de los dispositivos de cristal líquido". *Tesi Doctoral, Universitat de Barcelona.*
- [185] Millán, M. S., Corbalán, M., Pérez, E., Romero, J., Yzuel, M. J. (1995) "Improvement of the color discrimination capability in optical pattern recognition", *Proc. SPIE*, Vol. **2730**, 207-210, ISBN 0-8194-2111-1.
- [186] Campos, J., Yzuel, M. J., Yaroslavsky, L. P., Millán, M. S., Moreno, I., Corbalán, M., Pérez, E. (1997) "Channel transformations for improving color pattern recognition", a *Optoelectronic Information Processing*, B. Javidi i Ph. Réfrégier, eds., SPIE Press, Washington, pàgs. 105-135.
- [187] Pérez, E., Millán, M. S., Chalasinska-Macukow, K. (1998) "Dual nonlinear correlation applied to textured and colour object recognition", *Proc. SPIE*, **3409**, 444-455.

- 
- [188] Millán, M. S. i Yzuel, M. J. (1993) "High-pass matched filtering for optical character recognition with character-background contrast invariance", Proc. SPIE, **1983**, 421-422.
- [189] Pérez, E. (1995) "Selecció de canals quasimonocromàtics per a la millora de la discriminació cromàtica en el reconeixement de formes per correlació òptica multicanal". Treball de recerca, Universitat Politècnica de Catalunya.
- [190] Bono, R., Soler, J. i Fernández de Córdoba, L. (1985) "Variedades de agrios cultivadas en España". Conselleria de Agricultura, Pesca y Alimentación de la Generalidad Valenciana. Instituto Valenciano de Investigaciones Agrarias (I.V.I.A.).
- [191] Corbalán, M. (1993) "Estudio de la imagen captada mediante cámara CCD aplicado al reconocimiento óptico de objetos policromáticos". Treball de recerca. Universitat Autònoma de Barcelona.
- [192] Lohmann, A. W. i Paris, D. P. (1967) "Binary Fraunhofer holograms, generated by computer", Appl. Opt., **6**, 1739.
- [193] Burckhardt, C. B. (1979) "A simplification of Lee's method of generating holograms by computer", Appl. Opt., **9**, 1949.
- [194] Lee, W. H. (1974) "Binary synthetic holograms", Appl. Opt., **13** (7), 1677-1682.
- [195] Lee, W. H. (1979) "Binary computer-generated holograms", Appl. Opt., **18** (21), 3661-3669.
- [196] Moreno, I., Gorecki, C., Campos, J., Yzuel, M. J. (1995) "Comparison of computer-generated holograms produced by laser printers and lithography: application to pattern recognition", Opt. Eng., **34** (12), 3520-3525.
- [197] Yaroslavsky, L. P. (1992) "Is the phase-only filter and its modifications optimal in terms of the discrimination capability in pattern recognition?", Appl. Opt., **31** (11), 1677-1679.
- [198] Chalasinska-Macukow, K., Kotynski, R., Styczynski, K., Turon, F., Campos, J., Yzuel, M. J., Ahouzi, E., Gorecki, C. (1994) "Pure phase correlation. Application to optical pattern recognition", a *Optical pattern recognition*, B. Javidi i Ph. Réfrégier, eds., SPIE Press, Washington, pàgs. 275-292.

- [199] Javidi, B., Horner, J. L. (1994) "Nonlinear joint transform correlators" a *Real-time optical information processing*. Academic Press, Inc., Cap. 4, 115-183.
- [200] Downie, J. D. (1994) "Design of ternary correlation filters to reduce probability error", *Opt. Eng.*, **33**, 1777-1784.
- [201] Vijaya Kumar, B. V. K., Shi, W. i Hendrix, C. (1991) "Partial information correlation filters with maximally sharp correlation peaks", *Opt.Comp. and Proc.*, **1** (1), 29-46.
- [202] Pérez, E., Chalasinska-Macukow, K., Styczynski, K., Kotynski, R., Millán, M. S. (1997) "Dual nonlinear correlator based on computer controlled joint transform processor: digital analysis and optical results", *J. Mod. Opt.*, **44** (8), 1535-1552.
- [203] Javidi, B., Ruiz, J., Ruiz, C. (1991) "Performance of the binary nonlinear joint transform correlators in the presence of the Fourier plane quantization", *Opt. Commun.*, **80** (3,4), 275-284.
- [204] Javidi, B., Wang, J. (1991) "Quantization and truncation effects on binary joint transform correlation", *Opt. Commun.*, **84** (5,6), 374-383.
- [205] Yu, F. T. S. (1983) "Optical information processing", Wiley-Interscience, New York. Pàgs. 198-202.
- [206] Arsenault, H. H., Sheng, Y. (1991) "Pattern recognition, similarity, neural nets and optics" a *International trends in optics*, J. W. Goodman, ed., Academic Press, San Diego. Capítol 30.
- [207] Liu, H. K., Chao, T.H. (1986) "Optical image subtraction techniques", *Proc. SPIE*, **638**, 55-65.
- [208] Pati, G. S., Unnikrishnan, G., Singh, K. (1998) "Multichannel image addition and subtraction using joint-transform correlator architecture", *Opt. Commun.*, **150**, 33-37.
- [209] Jutamulia, S., Gregory, D. A. (1998) "Soft blocking of the dc term in Fourier optical systems", *Opt. Eng.*, **37** (1), 49-51.
- [210] Gianino, P. D., Woods, C. L. (1992) "Effects of spatial light modulator opaque dead zones on optical correlation", *Appl. Opt.*, **31** (20), 4025-4033.

- [211] Styczynski, K. (1995) "Liquid crystal spatial light modulator as a dynamic optical phase and amplitude information carrier", Publicació interna del Dept. de Física de la Universitat de Varsòvia.
- [212] Millán, M. S., Pérez, E., Chalasinska-Macukow, K. (1998) "Pattern recognition with variable discrimination capability by dual non-linear optical correlation", Accepted at Optics Communication.
- [213] Chalasinska-Macukow, K., Kotynski, R., Pérez, E., Millán, M. S. (1998) "Dual nonlinear correlation: optoelectronic realization and application to pattern recognition", Proc. SPIE of the International Conference on Optics and Optoelectronics (en premsa).

