

# PROCESSAMENT D'IMATGES. SEGMENTACIÓ

*ANTONI GASULL, FERRAN MARQUES*

Dept. de Teoria del Senyal i Comunicacions. E.T.S.E.T.B. - U.P.C.  
Apartat 30002, 08080 Barcelona

## INTRODUCCIÓ

L'origen de l'aplicació de les tècniques de processament d'imatges es pot datar al principi dels anys vint quan es realitza la primera digitalització d'una imatge per a ésser transmesa per un cable submarí recentment instal·lat entre Londres i Nova York. No és però fins a mitjan dels anys seixanta que experimenta un impuls important [1]. Això és degut fonamentalment a dos fets, un que afecta el processament del senyal en general i l'altre específic del cas de la imatge.

En primer lloc, és en aquesta època quan es desenvolupa de manera important el processat digital del senyal en general, la qual cosa es deu, d'una banda al desenvolupament del poderós algorisme de la FFT que permet de treballar digitalment en el domini transformat i d'altra banda a l'aparició de les primeres màquines informàtiques que poden donar suport a aquestes tècniques.

En segon lloc, la meitat dels anys seixanta és el moment d'expansió dels projectes d'exploració de l'espai, la qual cosa significa que es produeixen inversions molt fortes en aquest camp. Lògicament en aquesta exploració una part molt important rau en l'ús de les imatges, que han d'ésser adquirides, transmeses i tractades.

No és però fins al final dels anys setanta i fonamentalment als vuitanta quan el processament d'imatge adquireix una dimensió important en molts àmbits. Això és degut exclusivament a la tecnologia que amb els seus avenços permet que el tractament d'imatge ja no sigui una cosa exclusiva d'equips de recerca que disposen de grans ordinadors. Amb les novetats de la tecnologia l'adquisició, l'emmagatzemament i el tractament d'imatges es poden fer actualment amb equips de cost relativament baix, la qual cosa fa que els àmbits d'aplicació d'aquestes tècniques siguin innombrables i creixin dia rera dia.

El processament d'imatges es pot entendre com una particularització del processament del senyal, encara que normalment aquesta darrera expressió se sol aplicar, per raons històriques al cas de senyals temporals, és a dir unidimensionals. La característica fonamental de les imatges és que són senyals multidimensionals i és precisament aquest aspecte el que, com s'ha comentat, en fa retardar historicament la difusió científica, pel volum d'informació que suposen. Com a exemple comparatiu es pot dir que l'emmagatzemament digital directe (sense compressió) d'una sola imatge (típicament de 512 x 512 píxels de 8 bits) és equivalent a l'emmagatzemament digital directe de gairebé 33 segons de senyal de veu, i ja no en parlem si es treballa amb seqüències.

De tota manera, com també ja s'ha dit, el nombre de camps en què actualment s'aplica el tractament d'imatges és molt gran, la qual cosa no ha de sorprendre tenint en compte que les imatges són potser el mitjà més important que té l'èsser humà de percebre el món que l'envolta. Medicina, comunicacions, automatització industrial o robòtica són només alguns dels exemples més habituals.

Dins el tractament d'imatges es pot establir una subdivisió de tècniques que de manera general es poden classificar en tres àrees: codificació, millora-restauració i anàlisi. Les tècniques de codificació cerquen la reducció o compressió de la informació ja sigui per emmagatzemament o per transmissió [1].

Les tècniques de millora, com el seu nom indica, cerquen de manera genèrica l'augment de qualitat de les imatges. Com a exemples es poden citar l'eliminació o disminució del soroll o la restauració [1,2]. La restauració que en general es tracta com una àrea a part, de fet no és més que un cas particular de millora en què la degradació és produïda per un sistema modelable. La qualitat esmentada, però, no s'ha d'entendre només perquè a l'observador li resulti globalment més agradable, sinó que també hi ha casos en què es tracta de facilitar la cerca de la informació desitjada sense que això hagi de representar una millora visual global. En certes situacions les tècniques de millora poden distorsionar una imatge però en canvi això pot facilitar la detecció de les característiques volgudes.

El tercer grup de tècniques mereix un nou apartat ja que de fet és en aquesta àrea on s'aplica fonamentalment la segmentació. Són precisament els mètodes de segmentació el nucli principal d'aquest treball i per tant l'anàlisi d'imatges servirà de introducció a la darrera part que es destinarà íntegrament a les tècniques de segmentació d'imatges.

## ANÀLISI D'IMATGES

A diferència de les dues àrees comentades breument en la introducció, en què les tècniques s'apliquen a una imatge per a obtenir una altra imatge, l'anàlisi d'imatges parteix també d'imatges però per a obtenir paràmetres o

característiques, és a dir una representació simbòlica de la imatge tractada. Es per tant en aquí on conflueixen les tècniques de processament del senyal amb les de intel·ligència artificial en el apartat de visió [3]. En general es pot dir que en el nivell baix de l'anàlisi intervé la part de processament, mentre que en el nivell alt intervé la part de intel·ligència artificial. No son però només aquests camps els implicats ja que normalment les tècniques d'anàlisi requereixen coneixements específics de l'aplicació, la qual cosa les converteix en interdisciplinars.

Aquesta part d'anàlisi és la més propera al que realitzem amb el nostre sistema de visió. Es a dir, a partir d'una imatge s'extraiu d'ella la informació de la composició (a nivell d'objecte) i a partir d'un aprenentatge previ es realitza el reconeixement. Un exemple d'aquest procés que es pot intentar fer de manera automàtica ens el trobem en una aplicació biomèdica. La figura 1 mostra una imatge adquirida amb microscopi d'un conjunt de certs tipus de cèl·lules. Encara que depèn de

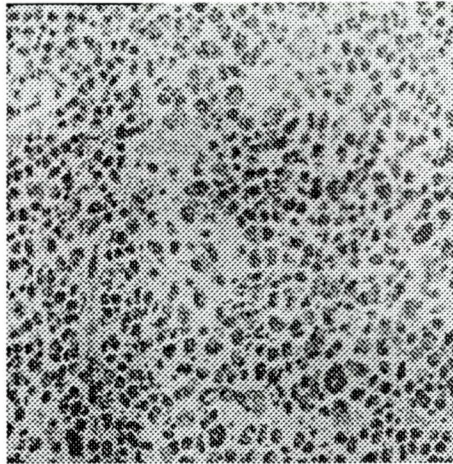


Fig. 1. Imatge biològica d'un conjunt de cèl·lules.

l'aplicació, en general la informació que s'ha d'extreure d'aquest tipus d'imatges és el recompte de la població de cèl·lules així com l'obtenció de certs paràmetres morfològics. Per a realitzar aquesta tasca de forma automàtica cal extreure primerament, i com a pas més important, les zones de la imatge que corresponen a les cèl·lules. Part d'aquesta tasca és precisament el que es coneix com a segmentació i que serà presentada en detall a continuació. Cal dir, encara que no sigui motiu del present treball, que per a completar la tasca d'anàlisi caldrà finalment obtenir els paràmetres pertinents segons la informació desitjada i també que habitualment es requereix prèviament algun tipus de preprocessament per a millorar la imatge, en el sentit de facilitar la tasca del mètode de segmentació que s'utilitzi [3].

## SEGMENTACIÓ

Tornant a l'exemple de la figura 1, la segmentació d'aquella imatge ens ha de permetre descompondre-la en diferents regions per a poder decidir posteriorment quines d'aquestes regions són les cèl·lules d'interès. Per tant una definició general del problema és que la segmentació d'una imatge és la descomposició d'aquesta en regions homogenies en una o diverses característiques. Aquesta definició es prou ambigua com per a cobrir plantejaments determinístics o plantejaments aleatoris (homogeneïtat implicarà estacionarietat), així com també per a fixar-nos en el color o el nivell de gris o la textura, etc.

Cal destacar el plantejament ampli que es dona aquí de la segmentació perquè habitualment es cau en la particularització, en el sentit de considerar la segmentació com la divisió en regions uniformes en el nivell de gris, quan en realitat ens podem trobar casos en què són les propietats estadístiques les que es mantenen dins cada regió, o bé que les regions es diferencien per textura. Com a exemple d'això, el resultat de la segmentació de la figura 2 a) i de la 2 b) ha d'ésser el mateix, figura 2 c).

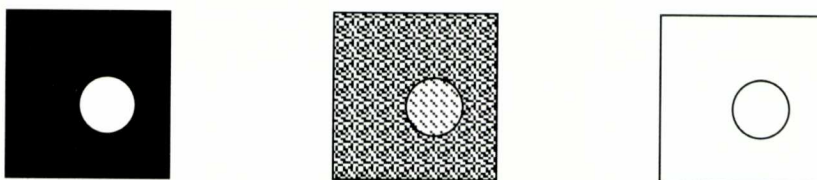


Fig. 2. Exemples de segmentació a) nivell de gris, b) textura.

Malgrat l'ambigüitat de la definició és possible donar una formulació a l'objectiu de la segmentació. Les regions en què es descompon la imatge han de verificar les següents condicions:

- A)  $I = \bigcup_j R_j$
- B)  $R_j \cap R_i = \emptyset \quad \forall j \neq i$
- C)  $R_j$  connectat  $\forall j$
- D)  $P(R_j) = \text{Cert} \quad \forall j$
- E)  $P(R_j \cup R_i) = \text{Fals} \quad \forall j \neq i$

on  $I$  representa tota la imatge,  $R_i$  és la  $i$ -èsima regió en què es descompon la imatge i  $P(.)$  significa la característica o característiques d'homogeneïtat escollides per a fer la segmentació.

La primera condició significa que la segmentació ha d'ésser completa, és a dir que cada píxel pertanyi a alguna regió i només a una (condició B). La tercera condició diu que tots els píxels de la regió han d'estar connectats, la qual cosa significa que una regió no pot quedar tallada per altres. Això també significa que cada regió queda limitada per un sol contorn. Hi ha l'excepció del cas que una regió en tingui inscrita una altra, però així i tot el límit de la regió interior serveix per a limitar per dins la regió exterior i per tant només es necessita el límit exterior d'aquesta per a descriure-la completament.

Les dues darreres condicions controlen la homogeneïtat de cada regió. Aquestes dues condicions estan dient que tots els píxels d'una regió han de complir una/es característica/es i que si s'agafa qualsevol altra regió (que pot ésser un sol píxel) veïna ja no es compleix. La indefinició rau precisament en l'elecció i comprovació de  $P(.)$ .

De fet l'ambigüïtat de la definició de segmentació és una conseqüència directa de l'ampli ventall d'aplicacions que té, ja que d'aquesta manera es poden unificar diferents tècniques que d'una manera o altra persegueixen objectius semblants.

Com ja s'ha dit la major part de les aplicacions on s'utilitza la segmentació són de l'àrea d'anàlisi d'imatges. Així per exemple la segmentació es necessita en citologia per a fer l'anàlisi automàtica de cèl·lules, i realitza la separació de les cèl·lules de la resta de la imatge. O bé en processos d'automatització industrial per a la inspecció automàtica de peces on cal extreure els objectes a analitzar de la resta de l'imatge. També s'utilitza la segmentació per al reconeixement automàtic de caràcters ja que primerament s'ha de delimitar cada caràcter a classificar. En robòtica s'utilitza la segmentació per a delimitar el conjunt d'objectes que componen l'escena. També és necessari fer segmentació en les innombrables aplicacions de la teledetecció, per tal de delimitar i classificar cadascun dels terrenys que formen la imatge obtinguda per satèl·lit.

L'anàlisi d'imatges no és però l'únic camp del tractament d'imatges on hom necessita la segmentació. També està adquirint cada vegada més importància en codificació. Així s'utilitza la segmentació per a calcular el vector de moviment en codificació de seqüències. També és la base dels mètodes de codificació anomenats de segona generació o mètodes de codificació basats en la forma, en què es codifica i es transmet cada regió de manera separada [4].

## TÈCNiques DE SEGMENTACIÓ

Un cop analitzat el problema de la segmentació, a continuació es fa una breu revisió d'alguns dels mètodes més habituals. De fet com a tècniques

pròpiament de segmentació només es poden considerar les tres darreres, que són les que es coneixen com a mètodes regionals; les dues primeres, encara que molt populars, són tècniques generals de processament d'imatges però no específiques de segmentació.

– *Segmentació per amplitud*

Aquest tipus de segmentació utilitza la característica més senzilla per a determinar l'homogeneïtat, com és el nivell de gris. Cada regió es forma amb el conjunt de píxels connectats que estan entre dos nivells de gris (llindars) predeterminats [1,2]. L'exemple més clàssic és la binarització.

La simplicitat del mètode és evident i això explica la seva popularitat. Per aquest motiu s'ha inclòs dins aquest treball. Però és l'exemple més clar de per què s'ha hagut de donar una definició tan àmplia de la segmentació. Si s'hagués considerat una definició clarament regional, com per exemple considerar la descomposició de la imatge en regions estacionàries, aquest mètode no seria un procés de segmentació. Això fa que en aplicacions de segmentacions complexes, com podria ésser la segmentació d'escenes, no tingui cap sentit l'ús d'aquest mètode. De fet és més un mètode de classificació que de segmentació. Es classifiquen els píxels que estan en certes franjes de nivell de gris i posteriorment es procedeix a la segmentació agrupant els píxels veïns que siguin de la mateixa classe.

El problema més important que s'ha de resoldre amb aquest mètode és l'elecció dels llindars òptims. Aquests valors es dedueixen a partir de l'histograma de la imatge, però només és senzill en alguns casos en què l'histograma és clarament multimodal. En general per la llei dels grans nombres l'histograma d'una imatge de mida gran tendeix a ésser unimodal i per tant complica molt l'elecció dels llindars. Aquest és un motiu més per no considerar aquest mètode si no és en situacions senzilles. Una alternativa habitual és la de calcular els llindars a partir d'un histograma local, per tal d'evitar el comentat efecte integrador dels grans nombres. Aquests llindars poden servir per a tota l'imatge o només per a cada zona d'on han estat extrets. En aquest darrer cas els llindars serien variables per al global de la imatge.

– *Tècniques basades en contorns*

Aquestes tècniques suposen que existeix una discontinuïtat entre regions i tracten de trobar-la. Per això cal realitzar quatre passos. El primer pas consisteix a aplicar detectors de contorns. Com que és una de les eines bàsiques del processament d'imatge, existeixen innumbrables algorismes per a realitzar la cerca dels contorns: els més típics són els que es basen en els operadors clàssics de gradient com el Sobel [2], Roberts [5], Prewitt [6], etc., encara que en l'actualitat s'estan usant operadors morfològics [7,8]. A la figura 3 es mostra el resultat del gradient morfològic sobre la imatge "Cameraman".



Figura 3. Resultat del gradient morfològic sobre el "Cameraman".

El següent pas consisteix en l'eliminació de falsos contorns. Aquests apareixen a causa de la sensibilitat dels detectors de gradient al soroll que es tradueix en l'aparició de petits contorns aïllats, com es pot apreciar en la part inferior dreta de la figura 3. Ja que corresponen a gradients baixos, aquests contorns es poden eliminar binaritzant la imatge, la qual cosa s'ha de fer igualment ja que els valors dels gradients es corresponen a nivells de gris.

Com és lògic els contorns finals, que seran els límits de les regions, han de tenir un gruix d'un píxel, per tant el tercer pas consisteix a aplicar un algorisme per a aprimar les línies obtingudes fins al moment [9]. Cal tenir present que en aquest pas no s'ha de trencar la connectivitat dels segments. El darrer pas consisteix a tancar els contorns de manera que cada regió quedi completament limitada i es completi la partició de la imatge [10].

Aquest mètode funciona bé amb imatges de grans contrastos, és a dir, quan clarament les regions estan separades per alts gradients, però aquesta situació no es dona sempre, i així, per exemple, hi pot haver variació estadística però no de gradient, o encara és més comú a l'inrevés, és a dir, que hi hagi variacions de gradient però de manera uniforme i per tant dins una mateixa regió. A més s'ha de destacar també la sensibilitat al soroll per la indefinició de gradient en el cas digital.

#### – *Formació de les regions per creixement*

Aquest sí que és un mètode exclusivament de segmentació. Es comença suposant un gran conjunt inicial de regions (habitualment cada píxel una

regió). La segmentació es realitza per la unió de regions veïnes segons el criteri d'homogeneïtat decidit a priori fins a arribar a una situació en què ja no és possible fer més unions [11].

El gran avantatge d'aquest mètode és que explota simultàniament les propietats de connectivitat i d'homogeneïtat. Per contra, presenta alguns problemes. El principal inconvenient del mètode és que depèn molt de les condicions inicials i de l'ordre en què es realitza el procés d'unió. A la figura 4 es

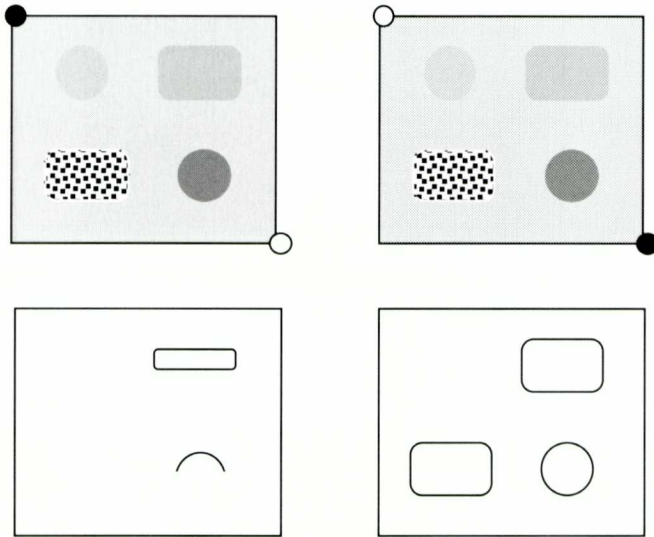


Fig. 4. A sota resultat de la segmentació segons que es comenci el procés d'unió des d'un punt o un altre.

poden veure els resultats de la segmentació d'una imatge sintètica. El criteri que s'ha utilitzat és que la distància (norma 1) entre un pixel i la mitjana de la regió sigui inferior a un determinat llindar. A la part esquerra el resultat obtingut fent un escombrat d'esquerra a dreta i de dalt a baix (inici en el punt negre i final en el blanc), a la part dreta el resultat de fer l'escombrat en sentit contrari.

Encara que potser és la tècnica de segmentació més popular, en general en certes aplicacions, com pot ésser el cas de la codificació, els resultats no són prou satisfactoris, sobretot per la tendència a la sobresegmentació (figura 5) i també pel cost computacional. A la figura 5 es mostra la imatge original a l'esquerra i la imatge segmentada a la dreta, on s'ha omplert cada regió amb la mitjana.





Fig. 5. Resultat de la segmentació per creixement de regions.

– *Formació de les regions per divisió i unió*

Aquest mètode naixia amb la idea d'evitar el problema principal present en el mètode anterior, realitzant un procés invers [12]. En lloc de començar amb moltes regions i anar unint, aquí es comença amb una de sola (tota la imatge) i es va dividint en regions més petites. Si una regió no compleix el criteri d'homogeneïtat predeterminat es divideix en quatre i així successivament.

Òbviament les tres primeres condicions donades en la definició de segmentació es compleixen directament. El procés de separació que s'acaba de

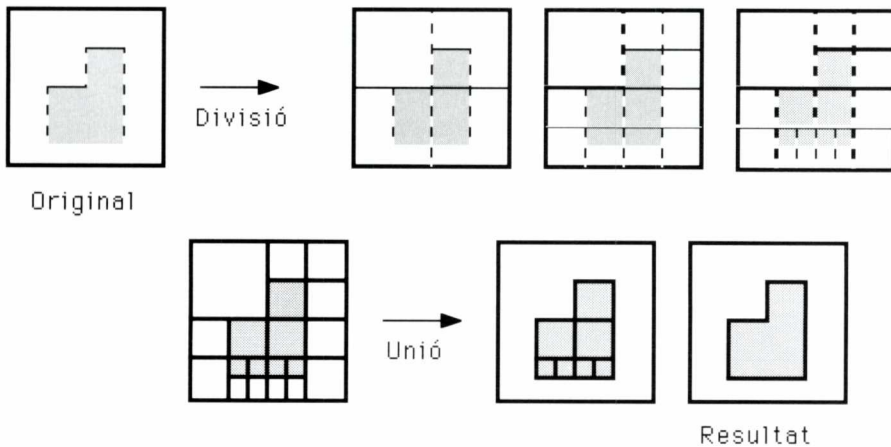


Fig. 6. Exemple del procés de formació de les regions per divisió i unió.

comentar obeeix a la condició quarta però no assegura el compliment de la cinquena. Això obliga a fer un nou pas com és el d'ajuntar regions veïnes que compleixin el criteri d'homogeneïtat. A la figura 6 es pot seguir tot el procés de separació i unió amb un senzill exemple [13].

Encara que aquest mètode és un pas més cap al mètode òptim, també presenta alguns problemes. El principal és l'efecte de blocs que dona la imatge segmentada a causa de la manera de realitzar les divisions. També cal mencionar la complexitat de l'estructura de dades per a la realització de l'algorisme.

#### – Segmentació per camp compost Gaus-Markov

El darrer mètode que es presenta obeeix a un model matemàtic més complex, però en canvi obté uns resultats molt més satisfactoris [14]. El plantejament inicial és el de considerar el problema de la segmentació com la partició òptima de la imatge en regions estacionàries. És a dir, que una imatge s'entén com una realització d'un procés estocàstic no estacionari de la qual s'ha de trobar la partició més probable. La partició es pot entendre com un camp superposat a la imatge d'elements unitaris o nuls segons que siguin o no del contorn respectivament. Això suposa definir un model jeràrquic o compost de la imatge format per un nivell observable (pixels) i per un segon nivell amagat dels contorns.

Un cop establert el model cal definir les estadístiques de cada nivell. La suposició més lògica, habitual i confirmada empíricament és que els valors dels pixels de cada regió segueixen una distribució gaussiana. D'altra banda, per raó de la dependència entre un punt del contorn i el seu veí, és d'esperar que el nivell de dalt segueixi un model de Markov de segon ordre, la qual cosa significa que els elements de la partició seguiran una estadística de Gibbs.

L'obtenció de la partició òptima s'obté aplicant el criteri MAP, és a dir, buscant la partició més probable donada una realització concreta. Aplicant Bayes això significa calcular el màxim del producte de la funció de probabilitat de la partició (funció de Gibbs) per la funció de probabilitat de la realització donada la partició.

Per a accelerar el procés de recerca de l'òptim es proposen solucions subòptimes però que tenen l'inconvenient d'una certa dependència amb la partició inicial. Per a eliminar aquesta dependència es pot fer una descomposició piramidal i anar projectant el resultat de la segmentació d'un nivell a l'altre de la piràmide [15]. Això no incrementa el cost computacional i en canvi aconsegueix uns resultats realment bons, com es pot observar a la figura 7.

Cal dir que precisament per la quantitat d'aplicacions de la segmentació que dia a dia estan apareixent, el nombre de tècniques també va dia a dia creixent i per tant és impossible presentar-les totes. També mereixen especial menció tècniques relativament recents basades en morfologia matemàtica [8]



Fig. 7. Exemple de segmentació mitjançant el model compost Gaus-Markov.

o d'altres basades en segmentació piramidal [16], però a part de no ésser tan habituals no arriben a resultats tan satisfactoris com la darrera tècnica proposada.

## CONCLUSIONS

Després d'una breu introducció al món del processament d'imatges, s'ha destacat la importància de l'anàlisi d'imatges. Dins aquesta àrea tot i no ésser tècniques exclusives d'ella, una de les parts més importants és la segmentació.

La primera cosa que s'ha de destacar de la segmentació és la inexistència d'una definició concreta. En aquest treball s'ha fet un plantejament prou ampli per a poder fer una unificació de tècniques. Com a conseqüència de l'ambigüitat en la definició de segmentació és molt difícil valorar la qualitat dels resultats dels diferents mètodes, que d'altra banda poden diferir bastant en els seus objectius. Fins i tot visualment la valoració és difícil, tot i que el sistema humà de visió és la millor eina per a fer la segmentació [17].

S'han presentat cinc tipus de tècniques diferents. Les dues primeres, segmentacions per amplitud i per contorns, pel fet de no estar orientades a l'objecte proporcionen resultats una mica pobres, i per contra són relativament senzilles de realització. Les tres darreres són tècniques estrictament de segmentació i tot i la seva complexitat proporcionen resultats millors, en especial la tècnica basada en el model compost de Gaus-Markov, que pot servir en el futur com una alternativa als dos mètodes típics de formació de les regions per creixement o per divisió i unió.

## REFERÈNCIES

- [1] R.C. GONZALEZ, P. WINTZ, *Digital Image Processing*, Addison-Welsey, Massachusetts, 1987.
- [2] W. K. PRATT, *Digital Image Processing*, Wiley-Interscience, Nova York, 1991.
- [3] D.H. BALLARD, C. M. BROWN, *Computer Vision*, Prentice Hall, New Jersey, 1982.
- [4] M. KOCHER, *Codage d'Images à Haute Compression Basé sur un Modèle Contour-Texture*, PhD Thesis, École Polytechnique Fédérale de Lausanne, Lausana, 1983.
- [5] L. G. ROBERTS, «Machine perception of three-dimensional solids,» in *Optical and Electro-Optical Information Processing*, J. T. Tippett et al., MIT Press, Cambridge, MA, pp. 159-197.
- [6] J. M. S. PREWITT, «Object Enhancement and Extraction,» in *Picture Processing and Psychopictorics*, B. S. Lipkin and A. Rosenfeld, academic Press, Nova York, 1970.
- [7] J. SERRA, *Image Analysis and Mathematical Morphology*, I, Academic Press, Londres, 1982.
- [8] BEUCHER, *Segmentation d'images et morphologie mathématique*, PhD Thesis, Paris School of Mines, Paris, juny, 1990.
- [9] R. J. SCHALKOFF, *Digital Image Processing and Computer Vision*, John Willey & Sons, Nova York, 1989.
- [10] D. VERNON, *Machine Vision. Automated Visual Inspection and Robot Vision*, Prentice Hall International (UK) Ltd, 1991.
- [11] J.L. MUERLE, D.C. ALLEN, «Experimental Evaluation of Techniques for Automatic Segmentation of Objects in a Complex Scene,» in *Pictorial Pattern Recognition*, G. C. Cheng et al., Eds., pp. 3-13, Thompson, Washington, 1968.

- [12] S.L. HOROWITZ, T. PAVLIDIS, «Picture Segmentation by a directed split-and-merge procedure,» in *Proc. 2nd Int. Joint Conf. on Pattern Recognition*, pp. 424-433, 1974.
- [13] K. S. FU, R. C. GONZALEZ AND C. S. G. LEE, *Robotics: Control, Sensing, Vision and Intelligence*, McGraw-Hill, Nova York, 1987.
- [14] F. MARQUES, A. GASULL, T. R. REED, M. KUNT, «Coding-oriented segmentation based on Gibbs-Markov random fields and human visual system knowledge», in *Proc. of ICASSP-91*, pp. 2749-2752, Toronto, Canadà, maig 1991.
- [15] F. MARQUES, J. CUNILLERA, A. GASULL, «Hierarchical segmentation using compound Gaus-Markov random fields» in *Proc. of ICASSP-92*, vol.III, pp. 53-56, San Francisco, USA, març 1992.
- [16] R. WILSON AND M. SPANN, *Image Segmentation and Uncertainty*, Research Studies Press Ltd., Letchmore, Hertfordshire, Anglaterra, 1988.
- [17] D. MARR, *Vision*, W. H. Freeman, Nova York, 1982.