

## Un model matemàtic per a l'anàlisi d'imatges\*

VICENT CASELLES, BARTOMEU COLL, JEAN-MICHEL MOREL

### 1 Introducció

Dins el camp de la visió per ordinador, el tractament o l'anàlisi de les imatges és una de les tasques fonamentals. Moltes són les aplicacions on es té una base d'imatges respecte de les quals es necessita fer una anàlisi per a una posterior interpretació o per poder extreure una informació que després podrem processar. En algunes aplicacions, com és el cas de les aplicacions industrials i mèdiques, s'obtenen seqüències d'imatges tridimensionals (3D). Pensem en el cas de la tomografia axial computada (TAC), la qual ens dona informació en 3D de les estructures internes del cos humà. Poder entendre aquestes imatges requereix la identificació i modelització de les superfícies dels objectes que són presents en les estructures en 3D. En el cas de les imatges per satèl·lit, un dels problemes prové del renou present en la imatge mateixa, degut potser a problemes de captació o transmissió. En aquest cas s'intenta fer un filtratge o preprocessament, intentant eliminar aquest renou i conservar la informació inherent a la pròpia imatge. El reconeixement automàtic de les formes és també un dels camps de la visió on l'anàlisi de les imatges és una part fonamental. Pensem en el reconeixement de la signatura o la lectura automàtica de caràcters. En el cas de la videovigilància potser es tracta d'identificar els trets principals en la fisonomia d'una persona o bé d'analitzar algun objecte que hi apareix. I no oblidem el camp de la robòtica, on l'anàlisi de les imatges en temps real que s'obtenen de l'entorn 3D en què es mou el robot, és d'una gran importància en la reconstrucció del seu entorn físic. En aquest cas, l'anàlisi o processament de les imatges seria un exemple del que anomenam *visió de baix nivell*, mentre que el problema de la robòtica aniria lligat a la *visió d'alt nivell*.

Aquests tan sols són alguns exemples d'aplicacions en el món de la visió per ordinador, en el qual l'anàlisi de les imatges té un paper fonamental. En aquesta anàlisi no és fàcil extreure la informació continguda en la imatge a partir dels nivells de grisos associats als píxels de la imatge. És per això que es fa necessari saber com

---

\* Conferència pronunciada a la Primera Trobada Matemàtica de la Societat Catalana de Matemàtiques, març 1998.

es formen les imatges de la vida real i quines són les operacions bàsiques que les conformen. En general, les imatges de la vida real o quotidiana que es projecten sobre la nostra retina, esdevenen un puzzle a causa de la descomposició arbitrària dels objectes que fan els uns sobre els altres a causa de les oclusions i ombres a què donen lloc. Aquest puzzle el resollem perllongant mentalment els objectes a partir de les ombres i les oclusions. Aquesta capacitat que tenim per *completar* els objectes s'anomena *completació amodal*, segons la teoria del *fenomenòleg* Kanizsa (vegeu [10]). Quan tenim una imatge natural, és a dir, una fotografia d'una escena exterior o interior presa amb una càmera estàndard, i la digitalitzam, és a dir, la *posam* dins la memòria de l'ordinador com si fos una matriu amb nombres, llavors ens plantejam si podem definir algorismes de tal manera que puguem dur a terme aquesta «completació amodal» (les juncions en T) i per tant, restituir una part de l'estructura física de l'escena fotografiada.

Estudiarem el *mapa topogràfic* de la imatge com l'estructura bàsica que ens dona la informació sobre els *objectes* de la mateixa imatge, això és, una descripció completa de la imatge per les fronteres dels conjunts de nivell i les juncions, o singularitats produïdes per aquestes fronteres ([3], [4]). Aquesta estructura serà invariant pels canvis de contrast de la imatge, propietat important imposada pel mateix model que estudiarem en la generació d'imatges. Aquest fet ens portarà a definir un tipus de singularitats essencials associades al mapa i de donar uns models matemàtics per poder simplificar la complexitat del mapa. Finalment, presentarem uns models de filtratge per a imatges segons els principis d'invariància que es demanen als diferents models, basats en l'anomenada *teoria de l'scale-space*, i que preservaran les singularitats bàsiques de la imatge.

## 2 La teoria de Kanizsa: generació de les imatges

### 2.1 Model simplificat per a la generació d'imatges naturals

Una qüestió que ens plantejam dins l'anàlisi d'imatges, prové del camí a partir del qual les imatges han estat físicament generades. En el món de l'acústica, la principal operació per a la qual els sons són generats és la vibració d'algun material rígid o elàstic. Si pensam en el cas del senyal i des d'un punt de vista matemàtic, els sons elementals són senyals oscil·latoris de manera que, quan diversos sons són emesos de manera simultània, el resultat final del senyal és simplement sumar els sons elementals. Com a conseqüència d'aquest fet, el món dels sons és modelitzat com a un espai vectorial amb descomposicions lineals.

En general, l'anàlisi dels sons intenta amb aquest tipus de descomposició trobar les diferents fonts amb el seu diferent temps i localització de les freqüències. Per una part, l'anàlisi de senyals descompon el senyal en sons elementals que millor poden ser localitzats en freqüència (anàlisi de Fourier) o en temps (anàlisi en ondetes o *wavelets*), o el dos alhora (Gabor, etc.), i per l'altra, ho fa de tal manera que aquesta descomposició sigui tan diferent com sigui possible (descomposició ortogonal). Dins l'anàlisi d'imatges també es tractaria de trobar aquests objectes bàsics (o sons, en el cas dels senyals) en els quals es pogués descompondre la imatge. Vegem-ho per al cas d'un model simple de generació d'imatges.

D'acord amb aquest model bàsic, la generació d'imatges naturals té lloc quan la llum solar o artificial reflectida pels objectes físics és obligada a passar pel forat d'una cambra *negra* (o caixa negra) i finalment s'imprimeix sobre la cara oposada

del forat de l'interior de la caixa negra. L'única hipòtesi que cal fer és que la llum es propaga en línia recta, fet cert per al problema que estem tractant. Llavors, si  $(x, y)$  és un punt genèric d'un pla fotosensitiu, anomenem  $u(x, y)$  el resultat d'aquesta impressió, el qual podem suposar que és un valor real que mesura la intensitat (o energia) de la imatge avaluada al punt  $(x, y)$ .

Segons la hipòtesi lambertiana, la llum natural o artificial que es reflecteix sobre la superfície dels objectes és la mateixa en totes les direccions. Aquesta hipòtesi posa uns problemes: diferents parts de la superfície dels objectes reflecteixen diferents intensitats de llum. Per altra part, els objectes poden ocultar a altres diverses fonts de llum. En aquest últim cas, apareixen les ombres sobre les seves superfícies.

## 2.2 Què és una imatge digital?

La diferència entre una fotografia o imatge natural i una imatge digital, és donada pel tipus de codificació. Per posar una fotografia dins la memòria d'un ordinador, es divideix la imatge en petits trossos quadrats, que anomenarem píxels, i a cada un d'aquests quadrats o píxels li associarem un número que representa la lluminància

Per a la codificació, cada número representa el nivell de gris associat al píxel  $(x, y)$ . Generalment, el negre està codificat pel zero, el 1 representa un color un poc menys negre, el 2 és encara un poc menys negre que el 1 i així. Dins la convenció que s'utilitza en el món de la informàtica, el 255 representa la codificació del color blanc.

Aquestes convencions no són absolutes. Pensem en una fotografia. Si no hi ha massa *llum*, la foto tendeix a saturar-se al negre mentre que si hi ha massa llum, la foto tendeix a saturar-se al blanc. I això és perquè la majoria dels captors de llum no són lineals i, encara pitjor, tenen un rang finit. Per això, quan la llum és massa forta (o feble) hi ha una saturació dels captors cap al blanc (o cap al negre).

Encara que el blanc i el negre absoluts no existeixen en la percepció, l'escala de les intensitats lluminoses intermèdies existeix, i normalment la codificam pels nombres enters creixents de 0 a 255. Llavors, amb aquesta divisió en píxels, una imatge digital és una matriu de números dins la memòria de l'ordinador.

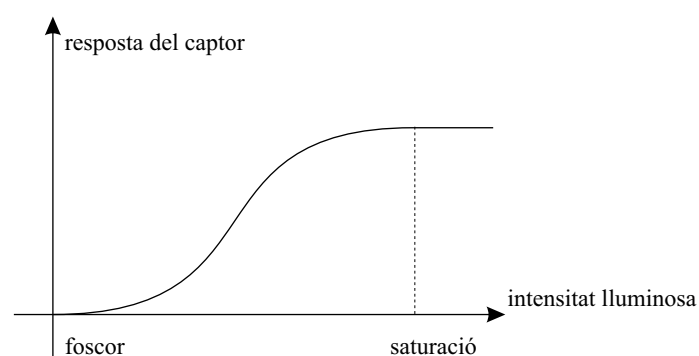


FIGURA 1: Resposta no lineal dels captors.

És evident que quan digitalitzam una fotografia, la imatge resultant digitalitzada es pareixerà més als nivells de grisos de la foto si la resolució és més alta. La resolució es pot entendre com la capacitat del sistema de la imatge per reproduir el contrast

dels objectes de mides diferents. Per *contrast*, entenem les diferències en intensitat dins un mateix objecte o bé entre un objecte i el fons que l'enrevolta. Notem que si una imatge digital ha estat digitalitzada amb una resolució suficient, ella mateixa té incorporada tota la complexitat i la riquesa de la fotografia.

### 2.3 Operacions bàsiques en la generació d'imatges: oclusions i transparències

Ben igual que en el cas de l'anàlisi de senyals o de sons, hem d'analitzar el fet mateix de la generació d'imatges naturals. A partir de l'anàlisi de les operacions bàsiques que donen lloc a les imatges naturals, podrem analitzar l'estructura dels objectes que conformen les imatges (vegeu [3]).

**2.3.1 Oclusió** Seguint la teoria de Kanizsa, generalment veiem les parts dels objectes davant de nosaltres que no ens queden ocultes per altres objectes. Ho podem formalitzar dient que els objectes han estat posats a l'escena, un per un o un darrere l'altre.

Donat un objecte  $\tilde{A}$  davant la càmera, anomenem  $A$  la regió de la imatge plana sobre la qual s'ha projectat l'objecte per la càmera. Anomenem  $u_A$  el nivell de gris de la imatge de l'objecte  $\tilde{A}$  així generat, el qual està definit en la regió plana  $A$ . Suposem que l'objecte  $\tilde{A}$  s'afegeix a una escena real  $\tilde{R}$  tal que la seva imatge era  $v$ . Observam una nova imatge, la qual depèn de quina part de l'objecte  $\tilde{A}$  està davant els objectes de  $\tilde{R}$  i quina darrera. Suposem que  $\tilde{A}$  oclou objectes de  $\tilde{R}$  i no està ocult per cap objecte de  $\tilde{R}$ . Tenim llavors una nova imatge  $u_{\tilde{R}\cup\tilde{A}}$  definida per:

$$\begin{aligned} u_{\tilde{R}\cup\tilde{A}} &= u_A \text{ en } A \\ u_{\tilde{R}\cup\tilde{A}} &= v \text{ en } \mathbb{R}^2 \setminus A . \end{aligned} \quad (1)$$

Notem que en aquest model bàsic no hem tengut en compte els efectes de l'ombra.

**2.3.2 Transparència** Suposem que una de les fonts de llum és un punt en l'espai euclidi i que un objecte  $\tilde{A}$  s'interposa entre una escena  $\tilde{R}$ , la imatge de la qual és  $v$ , i aquesta font de llum.

Anomenarem  $\tilde{S}$  l'ombra de  $\tilde{A}$  i  $S$  la regió que ocupa  $\tilde{S}$  en la imatge plana. Llavors la imatge resultant  $u$  es defineix com a:

$$\begin{aligned} u_{\tilde{R},\tilde{S},g} &= v \text{ en } \mathbb{R}^2 \setminus S \\ u_{\tilde{R},\tilde{S},g} &= g(v) \text{ en } S . \end{aligned} \quad (2)$$

on  $g$  denota una funció de contrast degut a l'ombra, la qual se suposa que és uniforme en  $\tilde{S}$ .

Clarament,  $g(s) \leq s$ , ja que la lluminositat decreix dins una ombra. L'única hipòtesi que podem fer sobre  $g$  és: punts d'igual nivell de gris  $s$  abans d'haver-hi l'ombra (mateix color) després de l'ombra tenen un nou nivell de gris  $g(s)$ , però el mateix pels diferents punts.

Una variant de l'efecte de l'ombra, el qual ha estat discutit per Kanizsa, és el fenomen de la transparència. En aquest cas, un objecte homogeni transparent  $\tilde{S}$  s'interposa entre part de l'escena i l'observador. Ja que  $\tilde{S}$  intercepta part de la llum enviada per l'escena, obtenim una relació equivalent per al cas de l'ombra i per

tant, des del punt de vista del processament d'imatges, transparència i ombra són equivalents. Si la transparència (o ombra) passa de manera uniforme sobre tota l'escena; llavors, la darrera relació serà

$$u_g = g(v), \quad (3)$$

la qual cosa ens diu que l'escala de nivell de grisos de la imatge s'ha alterat per una funció de canvi de contrast no decreixent.

Notem que en el món real tenim moltes fonts diverses de llums, ombres, autoil·luminacions dels objectes, llums difoses, etc. i per tant, el procés és molt més complex que el del model bàsic presentat.

#### 2.4 Requeriments per als operadors d'anàlisi d'imatges

Segons les operacions bàsiques definides abans en la generació d'imatges, oclusió i transparència, qualsevol processament de la imatge ha d'evitar destruir l'estructura de la imatge resultant d'aquestes operacions (1, 2). Llavors el nostre model simple de generació d'imatges ens imposa que les operacions d'anàlisi d'imatges han de ser invariants respecte a qualsevol canvi de contrast, un requeriment proposat per Matheron [14]. Direm que una operació  $T$  sobre una imatge és *invariant per contrast* si

$$T(g(u)) = g(T(u))$$

per a qualsevol funció de canvi de contrast  $g$  no decreixent. Exemples clàssics d'aquests tipus d'operadors són l'erosió, la dilatació, l'obertura i el tancament.

Ja s'ha comentat abans que la major part dels captors de llum són no lineals, i per això quan la llum és massa forta o feble, es produeix una saturació dels captors. Llavors els canvis de contrast no són causats només pels sensors sinó també pels canvis d'intensitat de la llum i els mateixos objectes presents en l'escena, el que dona lloc als canvis de contrast local.

L'evidència de la invariància de contrast en algunes branques de la visió per ordinador associades al reconeixement de formes és directa. Julesz, en la seva teoria, proposa els extrems de curvatura (*racons* o *terminacions* en la seva terminologia) així com les seves orientacions, com a claus per a la detecció de textures. També la teoria d'Attneave sobre el reconeixement humà de les formes, proposa els punts de curvatura i d'inflexió com a punts que contenen aquesta informació bàsica. En el cas de l'orientació, aquesta és donada per un vector tangent a la isofota i no s'altera per un canvi de contrast, igualment que la curvatura, la qual és calculada com la curvatura de les línies de nivell.

### 3 L'estructura bàsica per a l'anàlisi d'imatges. El mapa topogràfic

Anomenam *objectes bàsics* una classe d'objectes matemàtics, més simples que tota la imatge, els quals permeten descompondre la imatge i a partir dels quals la podem reconstruir. Dos exemples clàssics dins la descomposició d'una imatge són:

- *Descomposició additiva en ondes simples*: Els objectes bàsics de l'anàlisi de Fourier són el cosinus i el sinus; de l'anàlisi d'ondetes, ho són les ondetes o el paquet d'ondetes; els objectes bàsics de l'anàlisi de Gabor són les gaussianes modulades per sinus i cosinus. En tots aquests casos, la descomposició és

additiva i no s'adapta a l'estructura de les imatges. De fet, les operacions que ens porten a la construcció d'imatges reals són fortament no lineals i la més simple d'elles, el canvi de contrast, no preserva la descomposició additiva. Aquesta objecció no és certa en el cas de la compressió d'imatges, ja que en aquest cas les estructures de les escales fines de la imatge són les dominants i aquesta estructura és lineal.

- *Representació de la imatge per una segmentació*, és a dir, la descomposició de la imatge en regions homogènies separades per fronteres o contorns. La noció de *contorn* va lligada a la noció de *discontinuitat* i aquests contorns han estat considerats, i encara ho són en molts de casos, com als objectes bàsics de la imatge. Un resum dels models relacionats amb l'anàlisi variacional per a la definició de la noció de contorn, el podem trobar en [15], on s'argumenta que tots els mètodes de detecció de contorns són variacionals. La idea de contorn basat en la discontinuitat de la imatge no imposa cap restricció sobre el salt, però en la pràctica no ho podem calcular si no fixam la intensitat del contrast en els contorns, normalment un valor constant per a tota la imatge. Aquest criteri no és invariant respecte als canvis de contrast. Més encara, els clàssics detectors de contorns bàsicament consisteixen en la convolució de la imatge  $u$  amb un nucli  $k$ , seguit per un detector de contorns de tipus diferencial. Llavors si  $g$  és no lineal,  $k * (g(u)) \neq g(k * u)$ .

La morfologia matemàtica ofereix una alternativa: descompondre una imatge  $u$  en els seus conjunts de nivell, això és, en els conjunts  $X_\lambda u = \{\mathbf{x} \in \mathbb{R}^2 : u(\mathbf{x}) \geq \lambda\}$ ,  $\lambda \in \mathbb{R}$ , i els conjunts  $X_\lambda$  s'anomenen *conjunts de nivell* de  $u$ . Una imatge pot ser reconstruïda a partir dels conjunts de nivell per la fórmula:

$$u(\mathbf{x}) = \sup\{\lambda, u(\mathbf{x}) \geq \lambda\} = \sup\{\lambda, \mathbf{x} \in X_\lambda u\}.$$

La descomposició és, per tant, no lineal i redueix la imatge a una família de conjunts al pla  $\{X_\lambda\}$ . Òbviament, si transformam la imatge  $u$  en  $g(u)$  on  $g$  és una funció creixent, entesa com a una funció de contrast, llavors els conjunts de nivell de  $g(u)$  són els mateixos que  $u$ . Anem, doncs, a definir el mapa topogràfic associat a una imatge i a enunciar el resultat sobre la seva invariància per canvis de contrast locals (vegeu [4]).

Sigui  $\Omega$  un domini en  $\mathbb{R}^2$ . Sigui  $u: \Omega \rightarrow \mathbb{R}$  una imatge, és a dir, una funció mesurable i fitada.

1 DEFINICIÓ *Donada una imatge  $u$ , anomenarem conjunts de nivell superiors de  $u$  els conjunts de la forma  $[u \geq \lambda]$ , on  $\lambda \in \mathbb{R}$ .*

2 DEFINICIÓ *El mapa topogràfic superior d'una imatge és el conjunt format pels components connexos dels conjunts de nivell de  $u$ ,  $[u \geq \lambda]$ ,  $\lambda \in \mathbb{R}$ .*

Notem que, per la fórmula anterior, el mapa topogràfic superior associat amb  $u$  determina de manera única la funció  $u$ . També podríem treballar amb els conjunts de nivell inferiors de  $u$ ,  $[u \leq \lambda]$ .

Anomenam *línies de nivell* de  $u$  les fronteres dels conjunts de nivell superiors de  $u$ . Si suposam que podem determinar els conjunts de nivell de  $u$  a partir de les seves fronteres, les línies de nivell, llavors podem redefinir el mapa topogràfic de  $u$  com la

família de les línies de nivell de  $u$ . Aquest és el cas si la nostra imatge té la propietat que per a cada conjunt de nivell  $[u \geq \lambda]$ ,  $\lambda \in \mathbb{R}$ , la frontera  $\partial[u \geq \lambda]$  és donada com la unió numerable o finita de corbes de Jordan tancades. Llavors, les línies de nivell orientades defineixen els conjunts de nivell i, així, la funció  $u$ . Aquest fet restringeix el nostre model funcional per imatges contínues però no representa cap restricció al món de les imatges digitals. De fet, en el camp discret podem associar a cada conjunt de nivell un únic conjunt finit de corbes orientades de Jordan, les quals defineixen la seva frontera i, recíprocament, el conjunt de nivell es defineix de manera única a partir de les seves corbes de Jordan. En aquest cas, les anomenarem *corbes de la imatge*. En el que segueix, suposarem que aquestes corbes de nivell existeixen, bé sigui perquè la imatge és discreta o bé perquè pertany a un espai adequat de funcions.

3 DEFINICIÓ *Si  $u$  pertany a un espai de funcions tal que cada component connex d'un conjunt de nivell és fitat per un conjunt numerable o finit de corbes de Jordan orientades, llavors anomenarem mapa topogràfic de  $u$  la família d'aquestes corbes de Jordan.*

Noteu que, quan visualitzam el mapa topogràfic, tan sols visualitzarem les corbes de Jordan sense especificar els seus nivells  $\lambda$  o la seva orientació. Evidentment, per al cas de la reconstrucció de la imatge de  $u$ , necessitam aquesta informació.

Si suposam que els conjunts de nivell són subconjunts tancats de Cacciopoli de  $\mathbb{R}^2$ , això és, conjunts tancats els quals tenen llargària finita, llavors la seva frontera essencial és una unió numerable o finita de corbes tancades de Jordan i, possiblement, un conjunt de mesura de Hausdorff  $H^1$  nul·la. En aquest cas, podem descriure els components connexos dels conjunts de nivell per la seva frontera. Aquest és un cas interessant perquè cobreix el cas de les funcions de variació fitada, o simplement de classe  $BV$ , les quals han estat emprades en certs models de les imatges, com per exemple en el cas de la detecció de contorns, realçament de la imatge, etc. De fet, si  $u \in BV(\Omega)$ , llavors gairebé tots els conjunts de nivell  $[u \geq \lambda]$  són conjunts de Cacciopoli i el mapa topogràfic de  $u$  pot ser descrit en termes de les línies de nivell de  $u$ , per la fórmula de la reconstrucció.

### 3.1 Propietats d'invariància del mapa topogràfic superior

Provarem que el mapa topogràfic és invariant per contrast. Sigui  $\Omega$  un domini en el pla (per exemple, un rectangle). Donada una imatge  $u: \Omega \rightarrow \mathbb{R}$ ,  $\lambda \in \mathbb{R}$  i  $x \in [u \geq \lambda]$ , denotarem per  $cc([u \geq \lambda], x)$  el component connex de  $[u \geq \lambda]$  en  $\Omega$  que conté a  $x$ .

4 DEFINICIÓ *Direm que una aplicació multivaluada  $h: \Omega \times \mathbb{R} \rightarrow \mathcal{P}(\mathbb{R})$  és una multifunció monòtona si*

**MM1)**  $h(x, \lambda)$  és un interval de  $\mathbb{R}$  per a qualsevol  $x \in \Omega$  i  $\lambda \in \mathbb{R}$ .

*Sigui  $h^-(\cdot, \lambda) = \inf\{\mu: \mu \in h(x, \lambda)\}$ ,  $h^+(\cdot, \lambda) = \sup\{\mu: \mu \in h(x, \lambda)\}$ .*

**MM2)** *Si  $\lambda > \mu$ , llavors o bé  $h(x, \lambda) = h(x, \mu)$  o bé  $h^+(\cdot, \mu) \leq h^-(\cdot, \lambda)$ .*

**MM3)**  $\cup_{\lambda \in \mathbb{R}} h(x, \lambda)$  és un interval de  $\mathbb{R}$ .

5 DEFINICIÓ *Sigui  $u: \Omega \rightarrow [a, b]$  una imatge i sigui  $\{X_\lambda: \lambda \in [a, b]\}$  la família dels seus conjunts de nivell. Direm que una aplicació multivaluada  $h: \Omega \times \mathbb{R} \rightarrow \mathcal{P}(\mathbb{R})$  és un canvi de contrast local per a  $u$  si*

- H1)**  $h$  és una multifunció monòtona tal que per a cada  $\lambda \in \mathbb{R}$ ,  $h^-(\cdot, \lambda)$ ,  $h^+(\cdot, \lambda)$  són mesurables en  $\Omega$ .
- H2)** Si  $u(x) < \lambda$ , llavors  $h^+(x, u(x)) = h^-(x, \lambda) < h^+(x, \lambda)$ ,  $x \in \Omega$ ,  $\lambda \in \mathbb{R}$ .
- H3)**  $h^+(x, \lambda) = h^+(y, \lambda)$  per a tot  $x, y$  pertanyent al mateix component connex de  $[u \geq \lambda]$ ,  $\lambda \in \mathbb{R}$ .
- H4)** Sigui  $v(x) = h^+(x, u(x))$ . Si  $y \in cc([v \geq \mu], x)$  on  $\mu \in h(x, \lambda)$ ,  $x \in \Omega$ ,  $\lambda \in \mathbb{R}$ , llavors  $h(x, \lambda) = h(y, \lambda)$ .

6 DEFINICIÓ ([6]) Sigui  $u: \Omega \rightarrow [a, b]$  una imatge. Direm que  $v$  és una representació local de  $u$  si existeix algun canvi de contrast local  $h$  tal que  $v(x) = h^+(x, u(x))$ ,  $x \in \Omega$ .

La següent proposició prova que els canvis de contrast locals preserven el mapa topogràfic superior i per lo tant, el mapa topogràfic.

7 PROPOSICIÓ Sigui  $u: \Omega \rightarrow [a, b]$  i sigui  $v(x) = h^+(x, u(x))$ ,  $x \in \Omega$ , una representació local de  $u$ . Llavors:

- i)  $v(x) = \sup\{h^+(x, \lambda) : x \in X_\lambda u\}$ ,  $x \in \Omega$ . A més,  $x \in X_\lambda u$  si i només si  $x \in X_{h(x, \lambda)} v$ ,  $x \in \Omega$ ,  $\lambda \in \mathbb{R}$ .
- ii)  $v$  és una funció mesurable.
- iii) Sigui  $\Gamma$  (resp.  $\Gamma'$ ) un component connex de  $[v \geq \mu]$  (resp.  $[u \geq \lambda]$ ) el qual conté  $x$  i  $\mu = h^+(x, \lambda)$ . Llavors  $\Gamma = \Gamma'$ .
- iv) Per a cada component connex  $X$  de  $[u \geq \lambda]$  existeix  $\mu$  i un component connex  $Y$  de  $Y_\mu$  tal que  $X = Y$  i inversament.

El proper resultat ens diu que si dues imatges tenen el mateix mapa topogràfic superior, llavors estan relacionades per un canvi de contrast local. En aquest cas, les imatges poden ser considerades com a classes d'equivalència de funcions, mòdul canvis de contrast local.

8 TEOREMA Siguin  $u, v: \mathbb{R}^n \rightarrow \mathbb{R}$  dues funcions mesurables i fitades (imatges), i siguin  $X_\lambda$ , respectivament  $Y_\lambda$ , la família dels seus conjunts de nivell. Donat un conjunt de nivell  $X_\lambda$  (o  $Y_\lambda$ ) i un punt  $x \in \mathbb{R}^n$ , suposem que per a cada component connex  $X$  de  $X_\lambda$  existeix  $\mu$  i un component connex  $Y$  de  $Y_\mu$  tal que  $X = Y$  i això mateix amb  $X_\lambda$  i  $Y_\mu$  intercanviats. Llavors existeix un canvi de contrast local  $g(x, \lambda)$  tal que  $v(x) = g^+(x, u(x))$ .

Una interpretació intuïtiva del mapa topogràfic pot ser donada per les següents equivalències:

- Components connexos dels conjunts de nivell = Unió booleana d'objectes físics.
- Línies de nivell = Concatenació de trossos de fronteres d'objectes físics.
- Juncions alineades = Frontera en oclusió.



**3.2 Les singularitats de Kanizsa: les junctons en T**

Notem que en el procés de formació de la imatge, ja sigui contínua o digital, hi pot haver un canvi de contrast desconegut i no recuperable. Amb l'estructura donada pel mapa topogràfic, la imatge s'ha reduït a uns objectes que són invariants respecte als canvis de contrast, els components connexos dels conjunts de nivell de la imatge. A més, en el cas de les imatges digitals i per a segons quins espais de funcions en el cas analític, el mapa topogràfic d'una imatge pot ser descrit per la família de les seves corbes de nivell.

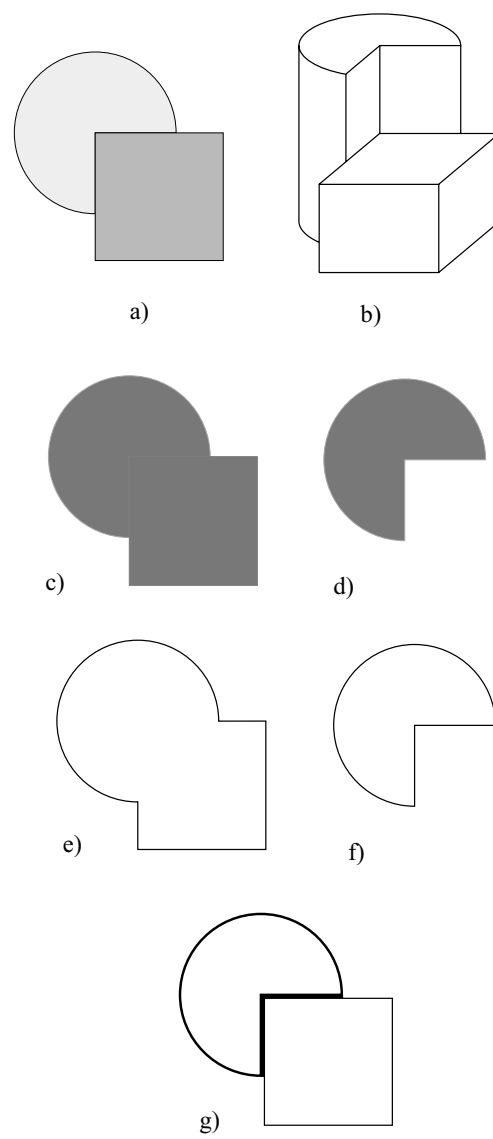


FIGURA 2: Anàlisi en relleu de la informació de la imatge.

La figura anterior és una mostra d'una imatge generada per una oclusió. En aquest cas, un disc gris és parcialment tapat per un quadrat més obscur. La figura b) ens dona una perspectiva d'aquest relleu. La imatge a) representa el quadrat sobre el disc i és una projecció de la figura b). Les figures c) i d) són dos conjunts de nivell a diferents nivells del relleu donat per b). Les figures e) i f) són les fronteres dels conjunts de nivell o les corbes de nivell de la imatge, així com ho hem definit abans. Finalment la figura g) ens dona la superposició de les dues línies de nivell e) i f). En aquest cas apareixen dues lletres T, o juncions, i l'estructura que es dona és la següent: Hi ha una part comú a les dues línies de nivell diferents i dues parts que no són comuns. Llavors a partir d'aquí podem deduir que una junció en T es produeix, quan es té un primer contorn, un segon que arriba fins el primer i es tallen, de tal manera que el primer es comporta com la barra horitzontal de la T i el segon, com la barra vertical. Basat en la teoria de Kanizsa, podem definir les parts significatives de les imatges com el resultat d'una segmentació de les línies de nivell per les juncions en T.

Llavors a partir d'aquí podem enunciar l'algorisme:

- Calcular els conjunts de nivell.
- Calcular les línies de nivell, o fronteres dels conjunts de nivell.
- Comparar dues a dues aquestes línies de nivell per trobar els punts on bifurquen i donen lloc als punts de junció en T.

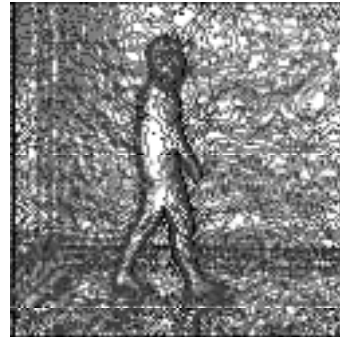
### 3.3 Càlcul de les línies de nivell i juncions sobre les imatges digitals

En una imatge digital, els conjunts de nivell són calculats per un simple llindar  $\lambda$  i en la tecnologia actual,  $\lambda = 0, 1, 2, \dots, 255$ , així que podem associar a la imatge digital 255 nivells de grisos diferents. Les corbes de Jordan frontera dels conjunts de nivell, les línies de nivell, són fàcilment calculables per un algorisme de càlcul de contorns, el qual ens dona cadenes de segments verticals o horitzontals limitant els píxels.

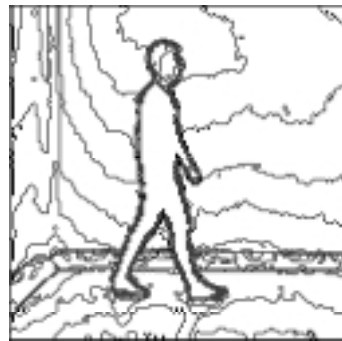
A nivell discret, definirem les juncions en T, com a qualsevol punt de la imatge plana a on dues línies de nivell (amb nivells diferents  $\lambda$ ) es troben. És a dir, considerarem que es troben en un punt si les línies de nivell divergeixen de manera significativa l'una de l'altra. Llavors per poder distingir aquesta situació de la situació en la qual dues línies de nivell paral·leles es troben en diferents punts a causa dels efectes de la quantificació, decidim tenir en compte una junció en T si i l'àrea de l'objecte que se superposa A, l'àrea de l'objecte ocult B i l'àrea del fons, són suficientment grans, és a dir, més grans que un cert llindar que imposarem. En qualsevol cas, el llindar d'àrea només depèn dels paràmetres de quantificació de la imatge digital i tendeix idealment a zero quan la qualitat de la imatge tendeix a infinit. A més, aquest llindar d'àrea ens delimita la talla dels objectes que considerem en aquesta anàlisi. Llavors, a partir d'aquestes premisses, l'algorisme per a la detecció de juncions, és el següent:



Imatge 3.1



Imatge 3.2



Imatge 3.3



Imatge 3.4



Imatge 3.5

FIGURA 3: Experiment del mapa topogràfic. La imatge original *El caminant* (autor: Paco Perales) és la imatge 3.1. La imatge 3.2 mostra el mapa topogràfic de la imatge 3.1 per les línies de nivell múltiples de 2. A la imatge 3.3, el mateix però pels múltiples de 10. La imatge 3.4 és la imatge original i la imatge 3.5 mostra el mapa topogràfic de la imatge 3.4, pels múltiples de 30.

### 3.3.1 Algorisme de detecció de juncions en T

- Fixem un llindar d'àrea  $n$  (a la pràctica,  $n = 40$  píxels sembla suficient per eliminar juncions associades a àrees molt petites) i un llindar  $b$  (a la pràctica,  $b = 2$  és suficient per eliminar els efectes de la quantificació dels nivells de grisos).
- A qualsevol punt  $x$  on dues línies de nivell es troben: definim  $\lambda_0 < \mu_0$  el valor mínim i el valor màxim de  $u$  dins l'entorn de 4-píxels al voltant del punt  $x$ .
- Denotam per  $L_\lambda$  el component connex de  $x$  en el conjunt  $\{y, u(y) \leq \lambda\}$  i per  $M_\mu$  el component connex de  $x$  en el conjunt  $\{y, u(y) \geq \mu\}$ . Calculam el més petit  $\lambda \geq \lambda_0$  tal que l'àrea de  $L_\lambda$  és més gran que  $n$ . Anomenam aquest valor  $\lambda_1$ . Trobam el més gran  $\mu, \lambda_1 \leq \mu \leq \mu_0$ , tal que l'àrea de  $M_\mu$  és més gran que  $n$ . Anomenam aquest valor  $\mu_1$ . Si  $\mu_1 - \lambda_1 \geq 2b$ , i si el conjunt  $\{y, \mu_1 - b \geq u(y) \geq \lambda_1 + b\}$  té un component connex el qual conté el punt  $x$  amb àrea més gran que  $n$ , aleshores ens quedam  $x$  com un punt de junció.



Imatge 4.1



Imatge 4.2



Imatge 4.3



Imatge 4.4

FIGURA 4: Experiment pel càlcul de les juncions a partir de l'algorisme anterior per diferents valors dels paràmetres  $n$  i  $b$ . En aquest experiment, imatge 4.1 és la imatge original. La imatge 4.2 ens mostra el càlcul de les juncions amb un llindar d'àrea  $n = 100$  i el paràmetre de quantificació  $b = 2$ . A la imatge 4.3,  $n = 100$  i  $b = 4$ . Finalment, a la imatge 4.4,  $n = 40$  i  $b = 2$ .

Una vegada definit el mapa topogràfic i enunciat l'algorisme per al càlcul de les juncions en  $T$ , és evident que tenim molta informació, i llavors el problema és com podem destriar la informació més bàsica del mapa respecte a la secundària. Aquest fet ens porta a donar una possible resposta per als objectes bàsics associats a l'anàlisi d'imatges:

*Els objectes bàsics de la imatge són les juncions en  $T$  i les parts de les línies de nivell que junten aquestes singularitats.*

**3.3.2 Visualització del mapa topogràfic** És clar que l'estructura del mapa topogràfic conté tota la informació. Però per tal de visualitzar aquesta estructura, podem definir diferents estratègies que ens permeten simplificar el mapa topogràfic sense perdre la informació bàsica de les principals línies de nivell i juncions.



Imatge 5.1



Imatge 5.2



Imatge 5.3



Imatge 5.4

FIGURA 5: La imatge 5.1 és la imatge 3.4 en la qual s'han eliminat tots els components connexos extrems d'àrea menor que 80 píxels. La imatge 5.2 mostra les línies de nivell de la imatge 5.1 que són múltiples de 20 (quantització). La imatge 5.3 és la imatge original i la imatge 5.4 és la imatge 5.3 després d'eliminar tots els components connexos extrems d'àrea menor que 100 píxels.

- Una primera eina per al filtratge de la imatge és l'aplicació de l'algorisme de Vincent-Serra [18]. Aquest algorisme invariant pel canvi de contrast elimina tots els components connexos extrems de la imatge que tinguin àrea menor que un cert nombre de píxels, normalment entre 10 o 30 píxels depenent de la talla de la imatge.
- Un altre camí en la simplificació del mapa topogràfic és donat pel procés de quantització. En aquest cas, s'intenta extreure la informació redundant que hi pugui haver al mapa topogràfic, particularment al cas de contorns on es pot tenir una acumulació de línies de nivell. Per això es presenten les línies de nivell múltiples d'un cert factor  $i$ , en conseqüència, eliminam totes les altres. Aquesta informació eliminada pot ser important, però el que es pretén és oferir una representació parcial però coherent de l'estructura de la imatge.

## 4 Filtratge de la imatge preservant les singularitats de Kanizsa

### 4.1 El model de l'*scale-space*

En aquesta primera part del treball s'ha presentat el mapa topogràfic com l'estructura bàsica de la imatge invariant per canvis de contrast locals. Donada una imatge, el problema que es planteja quan es proposa trobar les estructures, els objectes, les línies, etc., que componen la imatge mateixa, és un problema de difícil solució. Què entenem nosaltres mateixos per *objecte* o per *línia* per intentar-ho definir amb un algorisme i detectar-ho?



Imatge 6.1



Imatge 6.2

FIGURA 6: Experiment en l'aplicació de l'equació de la calor. En aquest experiment, la imatge imatge 6.1 és la imatge original. La imatge 6.2 és el resultat d'aplicar l'equació de la calor a la imatge 6.1 fins al temps  $t = 20$ .

La teoria clàssica està basada en la detecció dels contorns utilitzant la informació que ens donen els gradients alts associats a la imatge. És a dir, donada una imatge, podem calcular els gradients a cada punt i llavors considerar que estem en un punt de contorn si el valor absolut del gradient és gran. Però si ho fem directament sobre la imatge, podem trobar punts de gradient alt, però que estiguin associats al renou. Llavors una solució per extreure aquesta informació de la imatge és usar un filtre lineal. El filtre lineal més conegut és la convolució amb la funció gaussiana, i va ser

Koenderink ([12]) el que va adonar-se que la convolució amb la gaussiana era equivalent a resoldre l'equació de la calor. Després de passar aquest filtre lineal, es pot aplicar l'operador gradient per trobar els punts on el gradient és alt (veure [13]). És evident, que depenent de quina variància té la gaussiana, el resultat pot ser molt diferent, és a dir, la imatge estarà més o menys filtrada i per tant els detalls de la imatge poden quedar difuminats segons l'escala. Aquesta és la idea de l'anomenada *anàlisi multiescala*, que discutirem després. Si pensam en l'equació de la calor, agafam com a condició inicial la imatge i si cerquem la solució convolucionant amb la gaussiana, el resultat serà molt diferent depenent del temps allà on ens aturem. Asimptòticament, quan el temps tendeix a infinit, la solució tendeix a una constant.

En general, per calcular estructures en una imatge  $u(x, y)$ ,  $(x, y) \in \mathbb{R}^2$ , la teoria de l'*scale-space* o anàlisi multiescala proposa reemplaçar la imatge original per una versió filtrada  $u(t, x, y)$  (vegeu [19]). Com Witkin remarca, els contorns així trobats depenen de l'escala de la gaussiana. El paràmetre d'escala, que denotam per  $t$ , mesura el grau de filtratge que nosaltres volem donar a la imatge. Notem que per a  $t = 0$  ens referim a la imatge original i per a  $u(t, x, y)$  a la imatge filtrada al temps  $t$ . Ha estat provat en un treball d'Alvarez-Guichard-Lions-Morel [1], que el procés de filtratge local és representat per solucions d'equacions diferencials de tipus parabòlic, amb condició inicial  $u(0, x, y) = u(x, y)$ , la imatge que volem analitzar. Per distingir els diferents models que hi ha dins la teoria de l'*scale-space*, es tenen en compte les propietats d'invariància de l'operador associat a l'equació en derivades parcials: afí, morfològic o de contrast, etc. Aquest fet està en correspondència amb les propietats que li demanam a un bon model, que aplicam al filtratge de les imatges: que sigui invariant per una transformació afí de la imatge, que sigui invariant pels canvis de contrast aplicats a la imatge, etc. A partir de les diferents equacions proposades en la literatura, donem tres exemples, a causa de la seva invariància respecte a algunes propietats descrites anteriorment.

- L'equació de la calor [19]:

$$\frac{\partial u}{\partial t} = \Delta u \quad (4)$$

- L'equació de la curvatura mitja (equació d'Osher-Sethian [16], vegeu també [11])

$$\frac{\partial u}{\partial t} = |Du| \operatorname{curv}(u) \quad (5)$$

- El model AMSS ([17], [1])

$$\frac{\partial u}{\partial t} = |Du| (\operatorname{curv}(u))^{\frac{1}{2}} \quad (6)$$

on  $\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$  és la Laplaciana de  $u(t, x, y)$  respecte a les variables  $(x, y)$ ,  $Du = (\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y})$  denota el gradient espacial de  $u$ ,  $|Du|$  és la norma euclidiana de  $Du$  i  $\operatorname{curv}(u) = \operatorname{div}(\frac{Du}{|Du|})$ , la curvatura de les línies de nivell de  $u$ ,  $\operatorname{div} = \frac{\partial}{\partial x} + \frac{\partial}{\partial y}$  és l'operador divergència. Està provat en [1] que l'equació de la calor és l'única equació lineal isotròpica dels models de *scale-space*, mentre que el model donat per l'AMSS és l'únic model invariant per contrast i invariant afí dins els models donats per l'*scale-space* (vegeu també [17]). En el context de

L'*scale-space, invariant per contrast* significa que l'operador de filtratge  $T_t: u \rightarrow u(t) = T_t u$  commuta amb qualsevol canvi de contrast  $u \rightarrow g(u)$ , on  $g$  és qualsevol funció real no decreixent. L'equació d'Osher-Sethian és com un cas intermedi, en què les dues són quasilineals i invariants per contrast.

El problema que ens plantejam és el següent: hem de filtrar primer de tot la imatge per uns d'aquests models per després intentar cercar les estructures bàsiques de la imatge com per exemple les junccions? En primer lloc, qualsevol d'aquests models no té en compte la simetria de les junccions en T. A més, amb el filtratge de la imatge perdem informació que després no podem recuperar, les junccions en T es queden desplaçades o bé perdem la seva informació, ja que les junccions són un tipus de singularitat inherents al mateix procés de formació de la imatge. Per una altra part, sabem que l'algorisme que hem donat per al càlcul de junccions en T és invariant per contrast, que és un dels principis bàsics dins l'anàlisi d'imatges.

Llavors, el que proposam és detectar les singularitats de la imatge, prèviament a qualsevol procés de *scale-space* que li vulguem aplicar. Com ja hem dit abans, el problema que se'ns planteja quan miram el mapa topogràfic és la quantitat d'informació que hi ha i les oscil·lacions que es donen a les línies de nivell que ajunten les singularitats. Ja s'han donat uns models per a una millor visualització: eliminació de línies de nivell petites i la quantització. En el cas del filtratge multiescala o *scale-space*, donada una línia de nivell que ajunta dues singularitats, podríem utilitzar l'*scale-space* aplicat a les corbes per suavitzar aquests trossos de línies de nivell (vegeu [5]), però fixant els extrems.

#### 4.2 Models de filtratge per a imatges i corbes

El més invariant *scale-space* proposat en [1], és l'*scale-space* afimorfològic (AMSS), això és, una equació del tipus anterior la qual és invariant respecte als canvis de contrast i a les transformacions afins de la imatge plana,

$$\begin{aligned} \frac{\partial u}{\partial t} &= |Du|(\text{curv}(u))^{\frac{1}{3}} \\ u(0) &= u_0, \end{aligned} \quad (7)$$

on

$$\text{curv}(u)(x, y) = \frac{u_{xx}(u_y)^2 - 2u_{xy}u_x u_y + u_{yy}(u_x)^2}{(u_x^2 + u_y^2)^{\frac{3}{2}}}(x, y) \quad (8)$$

és la curvatura de la línia de nivell que passa per  $(x, y)$ . La interpretació del model AMSS és donada en [1] i correspon a la següent equació d'evolució per corbes proposada de manera independent en [17] i que anomenam ASS, o *scale-space* afíper corbes

$$\begin{aligned} \frac{\partial C}{\partial t}(t, s) &= (\text{curv}(C(t, s)))^{\frac{1}{3}} \vec{n}(t, s) \\ C(0, s) &= C_0(s), \end{aligned} \quad (9)$$

on  $C(0, s)$  és una corba inicial dins el pla i parametritzada per  $s$ ,  $\vec{n}(t, s)$  és el vector normal unitari a  $C(t, s)$  i  $\text{curv}(C(t, s))$  la curvatura de la corba a  $C(t, s)$ . La relació entre (AMSS) i (ASS) és formalment la següent: (AMSS) mou totes les línies



de  $u(t, x, y)$  com si cada una d'elles es mogué pel model (ASS). A nivell matemàtic, però, existència i unicitat no es donen al mateix grau. (AMSS) té una única solució en el sentit de les solucions de viscositat quan la condició inicial és contínua. (ASS) té una única solució regular quan la corba inicial és convexa [17]. *La major dificultat en la classificació proposada en [1] és la hipòtesi que la imatge inicial  $u_0(x)$  és contínua respecte a  $x$ .* Sense aquesta hipòtesi, el model (AMSS) i els relacionats models morfològics perden consistència matemàtica. Aquestes dificultats poden ésser obviades si empram el mètode d'Osher-Sethian [16]: Si volem aplicar el model (AMSS) a una imatge binària i per tant, discontinua  $\mathbf{1}_X$ , donada per la funció discontinua de  $X$  o funció característica, substituïm la funció que ens dóna la distància en signe  $\mathbf{1}_X$  de la imatge inicial  $u_0$  a  $X$ , la qual es Lipschitz i té  $X$  com a conjunt de nivell zero:  $X = \{x, u_0(x) \leq 0\}$ .

Apliquem a  $u_0$  el model (AMSS) i definim l'evolució de  $X$  com el conjunt de nivell zero de  $u(t)$ . Evans i Spruck [7] han provat que aquesta evolució no depèn de la funció de distància triada i coincideix amb (ASS) quan l'evolució de la frontera de  $X$  està ben definida i és regular. En altres paraules, poden emprar (AMSS) o bé per moviment de funcions contínues o bé per moure un conjunt simple, o una corba de Jordan simple, només movent la seva funció distància. Si pensam en funcions discontinues, podem deduir que l'evolució del (AMSS) serà ben definida només si ho podem reduir al moviment conjunt de les corbes de nivell simples. La pèrdua de continuïtat té però una conseqüència: la simetria es romp quan dos fronts es troben com és el cas de les juncions en T.

### 4.3 Un model de Kanizsa per a l'evolució d'imatges

En primer lloc, una correcta descomposició de la imatge implica una prèvia segmentació de les línies de nivell. Després d'aquesta descomposició donada pel mapa topogràfic, només el filtratge donat per l'equació del tipus (AMSS) o (ASS) pot ser aplicat. Això implica que podem moure les línies de nivell (ja sigui per suavitzar el mapa topogràfic o bé per aplicar l'anàlisi donat pe l'*scale-space*), però amb condicions a la frontera internes: les línies de nivell, durant el seu moviment donat per l'equació d'evolució, són obligades a passar pels punts extrems, les juncions en T.

En el cas de les imatges digitals, el conjunt de juncions és finit. Anomenam  $Z = \{z_1, z_2, \dots, z_n\}$  aquest conjunt. D'acord amb la discussió prèvia, hem d'adaptar el model d'evolució donat per l'equació (AMSS) a les imatges amb juncions en T. Llavors

**ALGORISME:** *Model de filtratge que preserva les juncions, invariant afí i invariant pels canvis de contrast.*

- Càlcul de les juncions  $z_1, z_2, \dots, z_n$  (per l'algorisme de detecció de juncions, el qual és invariant afí).
- Moviment de cada un dels trossos de les corbes de nivell  $C$  que junten les juncions pel model (ASS). Si la corba té punts extrems, es queden fixes i l'evolució donada pel (ASS) no permet a la corba creuar altres juncions. La corba final després de la seva evolució ve donada per  $C(t)$ .
- A partir de les corbes  $C(t)$ , es pot reconstruir una imatge filtrada, la qual té les corbes  $C(t)$  com a línies de nivell i les  $z_i$  com a juncions. Aquesta reconstrucció és possible gràcies a que el model (ASS) preserva la inclusió de corbes (vegeu [2]).

**4.3.1 Discussió analítica de l'algorisme** L'algorisme anterior és ben definit si es prova que el model (ASS) és matemàticament correcte; aquest punt però, no està completament solucionat, com ja s'ha dit abans. Llavors considerem la variant del model (AMSS) donat per l'equació en derivades parcials que implementa l'algorisme. Considerem la funció  $\mathbf{1}_{Z,\varepsilon}$  la qual és regular (és a dir,  $C^\infty$ ) i satisfà l'equació

$$\mathbf{1}_{Z,\varepsilon}(x, y) = 1 \quad (10)$$

si la distància de  $(x, y)$  a  $Z$  és més gran que  $\varepsilon$  i

$$\mathbf{1}_{Z,\varepsilon}(x, y) = 0 \quad (11)$$

si la distància de  $(x, y)$  a  $Z$  és més petita que  $\frac{\varepsilon}{2}$  i  $0 \leq \mathbf{1}_{Z,\varepsilon} \leq 1$  en qualsevol altre cas. Llavors el model considerat és

$$(AMSS_\varepsilon) \quad \frac{\partial u}{\partial t} = \mathbf{1}_{Z,\varepsilon} |Du| (\text{curv}(u))^{\frac{1}{3}}. \quad (12)$$

Aleshores, com està provat en [8], [9],  $(AMSS_\varepsilon)$  té una única solució de viscositat.

La interpretació de  $(AMSS_\varepsilon)$  és essencialment basada en la idea de preservar les singularitats donades per  $Z$  i moure les línies de nivell de  $u$  amb punts extrems sobre el conjunt de juncions  $Z$  per la variant del model (ASS) donat a l'algorisme anterior. Notem que la velocitat d'aquestes línies de nivell és alterada en un entorn de radi  $\varepsilon$  del conjunt de juncions  $Z$ .

**4.3.2 Anàlisi numèrica de les principals equacions** Com s'ha discutit abans, el model donat per  $(AMSS_\varepsilon)$  és el model correcte per ésser numèricament implementat. Un esquema numèric per a l'equació  $(AMSS)$  ha estat proposat per L. Álvarez i F. Guichard [2], basat en un mètode de diferències finites.

Seguint aquest model, posam l'equació  $(AMSS)$  en la forma

$$u_t = (|\nabla u|^3 \text{curv}(u))^{1/3} = (u_y^2 u_{xx} - 2u_x u_y u_{xy} + u_x^2 u_{yy})^{1/3}$$

Al nostre cas, per fixar l'evolució dels punts de junció, fem el següent canvi: Si  $(x, y)$  és un vertex de la malla que pertany a  $Z$ , simplement fixam a l'esquema donat els quatre píxels que envolten  $(x, y)$ . Aquest petit canvi a l'esquema numèric té dramàtiques conseqüències en l'evolució de la imatge dins l'*scale-space*, com veurem en els experiments numèrics.

4.3.3 Experiments



Imatge 7.1



Imatge 7.2



Imatge 7.3



Imatge 7.4



Imatge 7.5



Imatge 7.6

FIGURA 7: Algorisme de detecció de juncions i filtratge d'imatges que preserva les singularitats de Kanizsa. La imatge original és la imatge 7.1 *El caminant*. Imatge 7.2 dona el resultat de l'aplicació de l'algorisme de detecció de juncions aplicat a *El caminant* amb àrea  $n = 40$  i llindar  $b = 2$ . Les fronteres de les imatges generen gairebé una línia de juncions, com es veu entre la silueta de l'home i el fons de la imatge. Imatge 7.3: Anàlisi pel model (AMSS) de la imatge 7.1 a l'escala  $t = 12$ . Imatge 7.4: L'algorisme aplicat a la imatge 7.1, en la mateixa escala. Les juncions fixes són les de la imatge 7.2. L'aplicació de l'algorisme simplifica la forma de les línies de nivell i permet una millor visualització en l'organització del mapa topogràfic entre les juncions i les línies que junten juncions. Imatge 7.5: línies de nivell múltiples de 5 de la imatge original *El caminant*. Imatge 7.6 mostra les línies de la imatge 7.4, múltiples de 5.

## 5 Conclusions

A partir d'un model simple de generació d'imatges, hem analitzat l'estructura del mapa topogràfic com a estructura bàsica de la imatge invariant als canvis de contrast local. En aquest sentit, el mapa topogràfic conté tota la informació de la imatge, amb un algorisme de reconstrucció tan sols guardant la informació dels nivells de gris i l'orientació de cada línia de nivell. Podem fer simplificar aquesta informació eliminant totes les corbes de nivell petites respecte a un cert llindar o paràmetre, o bé per un procés de quantització. Per una altra part, l'estructura de les línies de nivell posa en evidència l'estructura dels objectes en oclusió, i dona lloc a les junctures en T. Aquestes singularitats resulten ser fonamentals en la generació d'imatges per oclusions o transparències d'objectes 3D.

S'ha presentat la teoria de l'*scale-space* com a un model per calcular les diferents estructures de tamanys diversos que es troben en una imatge. Però aquests models de filtratge no tenen en compte les singularitats de la imatge, com és el cas de les junctures en T. Una possible solució passaria per detectar aquestes singularitats de la imatge, com a pas previ a qualsevol procés de *scale-space* que li vulguem aplicar.

## Referències

- [1] ÁLVAREZ, L., GUICHARD, F., LIONS, P. L., MOREL, J. M. «Axioms and fundamental equations of image processing», *Arch. Rat. Mech. and Anal.*, 16, IX (1993), 200-257.
- [2] ÁLVAREZ, L., MOREL, J. M. «Formalization and Computational Aspects of Image Analysis», *Acta Numerica*, 1994, 1-59, Cambridge University Press.
- [3] CASELLES, V., COLL, B., MOREL, J. M. *A Kanizsa programme*. Preprint CEREMADE, Univ. Paris-Dauphiné, 1995.
- [4] CASELLES, V., COLL, B., MOREL, J. M. «Topographic Maps», Acceptat al *Int. J. of Computer Vision*, 1998.
- [5] CASELLES, V., COLL, B., MOREL, J. M. *Partial Differential Equations in Image Smoothing*. Seminaire de l'École Polytechnique, Palaiseau, Paris, 1996.
- [6] CASELLES, V., LISANI, J. L., MOREL, J. M., SAPIRO, G. *Shape Preserving Local Histogram Modification*. Technical Report HPL-97-58, April, 1997. També a *Proceedings ICIP-97*, IEEE Press, 1997.
- [7] EVANS, L. C., SPRUCK, J. «Motion of level sets by mean curvature», *J. Diff. Geom.*, 33 (1991), 635-681.
- [8] CRANDALL, M. G., ISHII, H., LIONS, P. L. «User's guide to viscosity solutions of second order partial differential equations», *Bull. Amer. Math. Soc.*, 27 (1992), 1-67.
- [9] GIGA, Y., GOTO, S., ISHII, H., SATO, M. M. «Comparison principle and convexity preserving properties for singular degenerate parabolic equations on unbounded domains», *Indiana Univ. Math. J.*, 40 (1991), 443-470.
- [10] KANIZSA, G. *Organization in Vision*. Nova York, Praeger, 1979.
- [11] KIMIA, B. B., TANNENBAUM, A., ZUCKER, S. W. «On the evolution of curves via a function of curvature, 1: the classical case», *J. of Math. Anal. and Appl.*, 163, 2, (1992).
- [12] KOENDERINK, J. J. «The structure of images», *Biol. Cybern.*, 50 (1994) 363-370.

- [13] MARR, D. *Vision*. Freeman and Co., 1981.
- [14] MATHERON, G. *Random Sets and Integral Geometry*. John Wiley, Nova York, 1975.
- [15] MOREL, J. M., SOLIMINI, S. *Variational methods in image processing*. Birkhäuser, 1994.
- [16] OSHER, S., SETHIAN, J. «Fronts propagating with curvature dependent speed: algorithms based on the Hamilton-Jacobi formulation», *J. Comp. Physics*, 79 (1988), 12-49.
- [17] SAPIRO, G., TANNENBAUM, A. «On affine plane curve evolution», *J. of Funct. Anal.*, 119, 1 (1994), 79-120.
- [18] VINCENT, L. «Grayscale area openings and closings, their efficient implementation and applications», *First Workshop on Mathematical Morphology and its Applications to Signal Processing*. J. Serra and Ph. Salembier, editors. Barcelona, 1993, 22-27.
- [19] WITKIN, A. P. «Scale-space filtering», *Proc. of IJCAI*, Karlsruhe 1983, 1019-1021.

VICENT CASELLES, BARTOMEU COLL  
DEPARTAMENT DE MATEMÀTIQUES I INFORMÀTICA,  
UNIVERSITAT DE LES ILLES BALEARS,  
07071 PALMA DE MALLORCA, BALEARS  
dmivca0@ps.uib.es, dmitcv0@ps.uib.es

JEAN-MICHEL MOREL  
CEREMADE,  
UNIVERSITÉ PARIS-DAUPHINE,  
PLACE DU MARECHAL LATTRE DE TASSIGNY,  
75775 PARIS CEDEX 16, FRANCE  
morel@paris9.dauphine.fr