

Anàlisi d'imatges

Índice

1. Contorns	1
2. Segmentació	1
2.1. Segmentació de Mumford-Shah	1
3. Descripció de la imatge per les seves línies de nivell: el mapa topogràfic	1
3.1. El principi d'invariància per canvi de contrast	1
3.1.1. Canvis de contrast i filtratge d'imatges	3
3.2. Conjunts de nivell i canvis de contrast	5
3.3. Línies de nivell i mapa topogràfic	8
3.4. Organització del mapa topogràfic en una estructura d'arbre	9
3.5. Mapa topogràfic i conjunts de nivell	9
3.6. Mapa topogràfic d'una imatge digital	11
3.7. Aplicacions del mapa topogràfic	13
3.7.1. Comparació morfològica d'imatges	13
3.7.2. Filtratge: filtre de gra [GuichardMonasse]	17
3.7.3. Posada en correspondència d'imatges (<i>registration</i>)	19
4. Referències	25

1. Contorns

2. Segmentació

2.1. Segmentació de Mumford-Shah

3. Descripció de la imatge per les seves línies de nivell: el mapa topogràfic

3.1. El principi d'invariància per canvi de contrast

Una propietat important de la visió humana, enunciada pels psicòlegs gestaltistes ja en 1923 és que la percepció dels objectes d'una escena és independent de la il·luminació o del color. La Figura 1 il·lustra aquesta propietat.

En la Figura 1 observam tres imatges que representen la mateixa escena. Una d'elles és una imatge en color, y les altres dues són imatges de nivell de gris amb il·luminacions diferents. A pesar d'aquestes diferències en el color i la il·luminació no tenim cap dificultat en reconèixer com a idèntics els objectes que

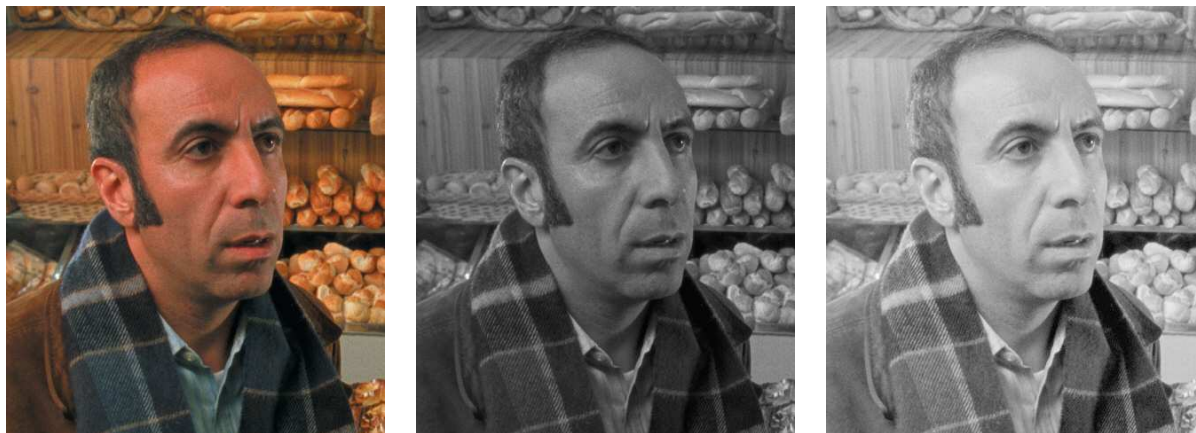


Figura 1: Invariància de la percepció als canvis de color o contrast: en les tres imatges som capaços de reconèixer els objectes que hi apareixen.

apareixen en les tres imatges. En aquest sentit deim que la informació de *forma* continguda en les imatges és invariant als canvis de color i contrast.

Si ens centram en el cas de les imatges de nivell de gris, ens podem demanar si els models per a l'anàlisi d'imatges presentats en les seccions anteriors (contorns i segmentació) prenen en compte aquesta invariància. Dit en altres paraules, les característiques bàsiques d'una imatge, obtingudes per segmentació o extracció de contorns, són les mateixes quan la imatge és sotmesa a un canvi de contrast? La resposta és "no" i les següents figures ho il·lustren.

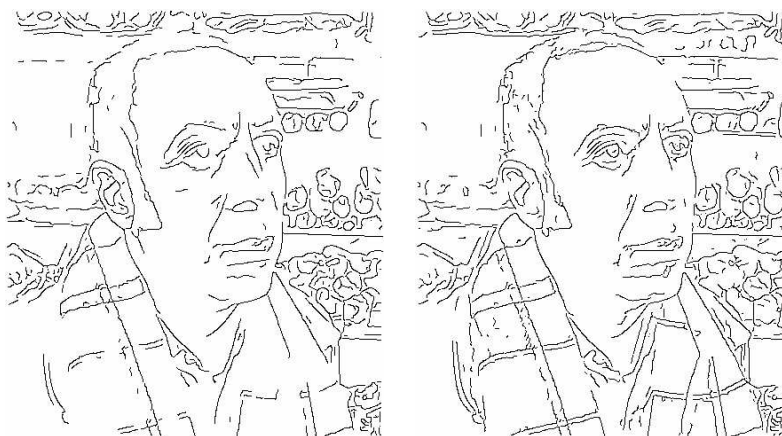


Figura 2: Els contorns obtinguts per a dues imatges relacionades per un canvi de contrast són diferents. En conseqüència, podem afirmar que la representació de les imatges basada en els contorns no és invariant als canvis de contrast.

Les figures anteriors mostren com els models presentats fins ara per a l'anàlisi d'imatges són incapaços de complir un dels requeriments bàsics de la percepció humana, el de la invariància per canvi de contrast. Per aquest motiu cerquem un nou model que sí el compleixi. Aquest model està basat en la representació de la imatge pels seus conjunts de nivell.



Figura 3: Les segmentacions obtingudes per a dues imatges relacionades per un canvi de contrast són diferents. En conseqüència, podem afirmar que la representació de les imatges basada en la segmentació no és invariant als canvis de contrast.

3.1.1. Canvis de contrast i filtratge d'imatges

Abans de passar a estudiar aquesta nova representació de les imatges dedicam unes línies a parlar dels canvis de contrast des del punt de vista de la seva modelització matemàtica. El model més simple consisteix en considerar un canvi de contrast com una funció monòtona creixent - estrictament o no - (veure la Figura 4, centre).

$$g : \mathbb{R} \longrightarrow \mathbb{R} \quad g \text{ monòtona creixent} \quad \Longleftrightarrow \quad g \text{ és un canvi de contrast}$$

Aquest model captura la característica essencial d'un canvi de contrast: els valors de nivell de gris són modificats, però no la seva relació d'ordre. En aquest sentit, si un objecte és més clar que un altre en la imatge original, el mateix passarà en la nova imatge (veure les imatges esquerra i dreta de la Figura 4). Com a molt, si la funció de canvi de contrast no és estrictament monòtona pot passar que dos nivells que abans eren diferents tinguin ara assignat el mateix valor.

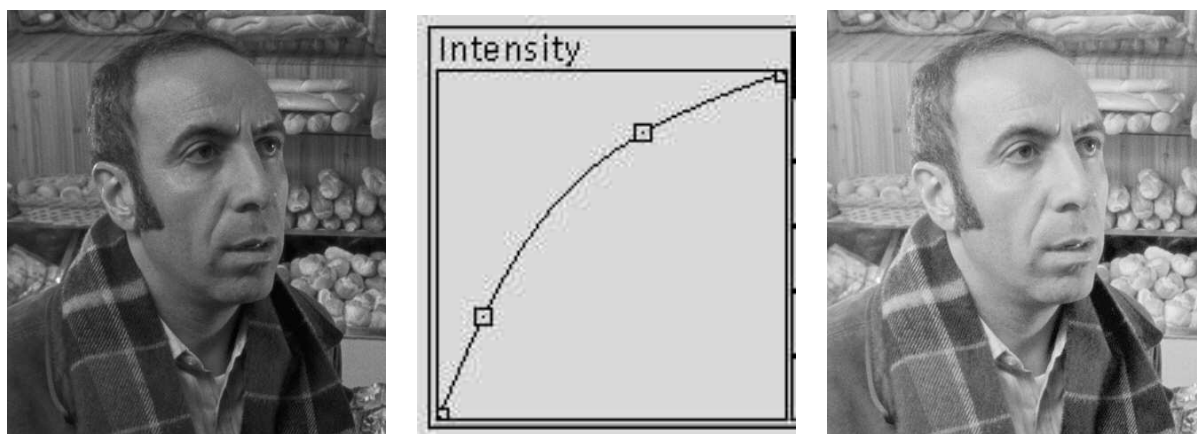


Figura 4: Imatge original, canvi de contrast i imatge resultant d'aplicar el canvi de contrast a la primera imatge.

Definició. Deim que un **operador** (filtre) T que associa a una imatge u una nova imatge Tu (dins un determinat espai d'imatges) és **invariant per canvi de contrast** (o **morfològic**) si compleix la propietat

$$T(g(u)) = g(Tu) \quad (1)$$

on g és un canvi de contrast.

És a dir, T és invariant per canvi de contrast si commuta amb totes les funcions monòtones creixents g .

Exercici. Recordem (Tema 5) que el filtratge gaussià es pot modelar mitjançant l'equació de la calor:

$$\frac{\partial u(x, y, t)}{\partial t} = \nabla^2 u(x, y, t)$$

on $u_0 = u(x, y, 0)$ és la imatge original, ∇^2 és l'operador laplacià i $u(x, y, t)$ és la imatge filtrada a l'escala t , la qual és solució de l'anterior equació en derivades parcials.

Demostrar que l'operador T_t que assigna a l'imatge original $u_0 = u(x, y, 0)$ la nova imatge $T_t u(x, y, 0) = u(x, y, t)$ no és invariant per canvi de contrast.

Solució. Si T_t és invariant per canvi de contrast s'ha de complir que

$$T_t g(u(x, y, 0)) = g(T_t u(x, y, 0)) = g(u(x, y, t))$$

és solució de l'equació de la calor. Aplicant la regla de la cadena i el fet que $u(x, y, t)$ és solució de l'equació de la calor obtenim (suposant que g és C^2):

$$\frac{\partial g(u(x, y, t))}{\partial t} = g'(u) \frac{\partial u(x, y, t)}{\partial t} = g'(u) \nabla^2 u(x, y, t) \quad (2)$$

$$\nabla^2 g(u(x, y, t)) = g'(u) \nabla^2 u(x, y, t) + g''(u) \|\nabla u\|^2 \quad (3)$$

Comparant aquestes dues equacions podem veure com, en general, $\frac{\partial g(u(x, y, t))}{\partial t} \neq \nabla^2 g(u(x, y, t))$, i per tant el filtratge gaussià no és invariant per canvis de contrast.

Observacions.

1. En general, les operacions damunt imatges definides per una convolució no són invariants per canvis de contrast. Això implica que les operacions amb imatges realitzades amb filtres lineals i invariants no són invariants per canvis de contrast, ja que aquestes operacions es modelen com la convolució de la imatge original per la resposta impulsional del filtre (veure el Tema 3).
2. Si a més de la invariància per canvi de contrast exigim als filtres altres propietats relacionades amb la percepció, com la invariància per translació, per rotació, zoom, etc. podem arribar a deduir l'expressió general de les operacions perceptualment acceptables damunt les imatges. S'ha demostrat que aquestes operacions es poden relacionar amb la família d'equacions diferencials estudiades en el Tema 5. Recordem que aquestes equacions tenen la forma

$$\frac{\partial u}{\partial t} = |\nabla u| \text{curv}(u)^\alpha$$

3.2. Conjunts de nivell i canvis de contrast

Donada una imatge $u(x, y) : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ definim el **conjunt de nivell (superior)** λ de u com:

$$\mathcal{X}_\lambda u = \mathcal{X}_\lambda^{\text{sup}} u = \{(x, y) \in \Omega : u(x, y) \geq \lambda\} \quad (4)$$

De manera anàloga podem definir el **conjunt de nivell inferior** λ com:

$$\mathcal{X}_\lambda^{\text{inf}} u = \{(x, y) \in \Omega : u(x, y) \leq \lambda\} \quad (5)$$

Observació. Un conjunt de nivell (superior o inferior) pot estar format per una o vàries components connexes¹ (veure la Figura 5).

La Figura 5 mostra un exemple de conjunt de nivell superior d'una imatge.

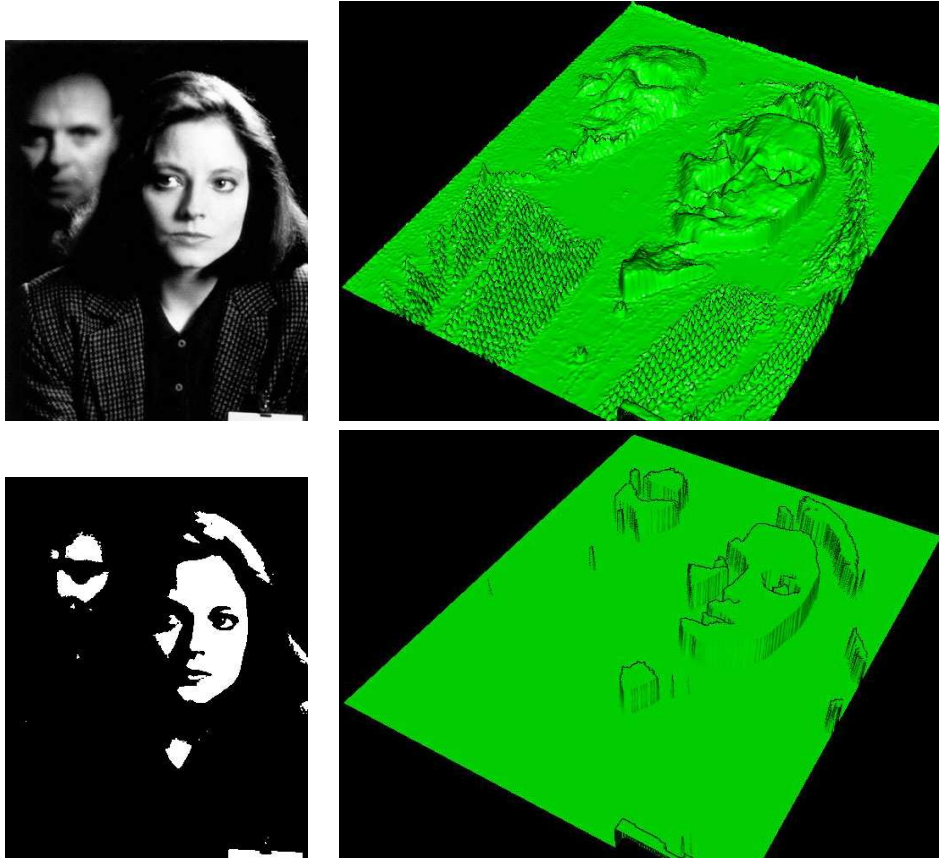


Figura 5: Representació tridimensional d'una imatge i d'un dels seus conjunts de nivell. En la part superior, imatge original i representació 3D. En la part inferior, un conjunt de nivell de la imatge i la seva representació tridimensional. Podem observar con el conjunt de nivell està compost per vàries components connexes.

¹Es diu que un espai topològic X és **connex** si no es pot escriure com la unió de dos conjunts disjunts no buits. Un conjunt C de X s'anomena **component connexa** si és connex i per a tot subconjunt connex C_1 de X tal que $C \subseteq C_1$, llavors $C_1 = C$. La noció de connexitat en el cas d'imatges digitals ha estat definida en el Tema 4.

Propietats. Donada una imatge $u(x, y) : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$, si $X_\lambda = \mathcal{X}_\lambda$ denota el conjunt de nivell superior λ , llavors

I) **Propietats d'inclusió.**

- a) $X_\lambda \subseteq X_\mu$, per a tot $\lambda > \mu$
- b) $X_\lambda = \cap_{\mu < \lambda} X_\mu$, per a tot $\lambda \in \mathbb{R}$

II) **Fòrmula de reconstrucció.**

$$u(x, y) = \sup\{\lambda \in \mathbb{R} : (x, y) \in X_\lambda\} \quad (6)$$

Demostració.

Propietat i)a). La propietat es dedueix directament de la definició de conjunt de nivell, ja que, com $\lambda > \mu$,

$$\mathcal{X}_\mu u = \{(x, y) \in \Omega : u(x, y) \geq \mu\} =$$

$$= \{(x, y) \in \Omega : u(x, y) \geq \lambda\} \cup \{(x, y) \in \Omega : \mu \geq u(x, y) < \lambda\} \supseteq \{(x, y) \in \Omega : u(x, y) \geq \lambda\} = \mathcal{X}_\lambda u$$

Propietat i)b). (Exercici).

Propietat ii). (Exercici).

Exercici. Conjunts de nivell d'una imatge digital. En el cas d'una imatge digital, el conjunt de valors que pren $u(x, y)$ està quantitzat i, normalment, $u(x, y) \in \{0, 1, 2, \dots, 255\}$. Es demana demostrar que en una imatge digital existeixen un màxim de 255 conjunts de nivell diferents.

Demostració. Basta demostrar que per a tot $0 < x < 1$, $i \in \{0, 1, \dots, 254\}$:

$$\begin{aligned} X_{i+x} &= \{(x, y) \in \Omega : u(x, y) \geq i + x\} = \\ &= \{(x, y) \in \Omega : i + x \leq u(x, y) < i + 1\} \cup \{(x, y) \in \Omega : u(x, y) \geq i + 1\} = \\ &= \emptyset \cup \{(x, y) \in \Omega : u(x, y) \geq i + 1\} = X_{i+1} \end{aligned}$$

La Figura 6 mostra (marcats en blanc) alguns exemples de conjunts de nivell per a una imatge. Observar la propietat d'inclusió d'aquests conjunts a mida que augmenta λ . Observar que cada conjunt de nivell està format per una o varies components connexes.

Teorema [GuichardMorel]. Conjunts de nivell i canvis de contrast. Donada una imatge u i una funció contínua per la dreta no decreixent g (canvi de contrast continu per la dreta), llavors

$$\begin{aligned} \forall \lambda \in \mathbb{R} \quad \exists \mu \in \mathbb{R} \quad \text{tal que} \quad \mathcal{X}_\lambda u &= \mathcal{X}_\mu g(u) \quad \text{i} \\ \forall \mu \in \mathbb{R} \quad \exists \lambda \in \mathbb{R} \quad \text{tal que} \quad \mathcal{X}_\lambda u &= \mathcal{X}_\mu g(u) \end{aligned}$$

És a dir, tot conjunt de nivell de u és un conjunt de nivell de $g(u)$ i viceversa. Per tant, u i $g(u)$ tenen, globalment, els mateixos conjunts de nivell.



Figura 6: Imatge original i conjunts de nivell superiors per a $\lambda = 100, 150$ i 200 . Els pixels que pertanyen al conjunt de nivell s'han marcat en color blanc. Observar la propietat d'inclusió entre els conjunts i l'existència de vàries components connexes en cada un d'ells.

Demostració. Definim $g^{(-1)}(\lambda)$ com

$$g^{(-1)}(\lambda) = \inf\{r : g(r) \geq \lambda\}$$

Com que g és contínua, també serà contínua per la dreta i tenim que per a tota successió de nombres r_n tals que $r_n > g^{(-1)}(\lambda)$

$$\lim_{r_n \rightarrow g^{(-1)}(\lambda)} g(r_n) = g(g^{(-1)}(\lambda))$$

Com que $r_n > g^{(-1)}(\lambda)$, per la definició de $g^{(-1)}(\lambda)$ tenim que $g(r_n) \geq \lambda$, llavors $\lim_{r_n \rightarrow g^{(-1)}(\lambda)} g(r_n) \geq \lambda$ i per tant

$$g(g^{(-1)}(\lambda)) \geq \lambda$$

És fàcil demostrar (*exercici*) que si aquesta condició es compleix i g és no decreixent, llavors

$$g(s) \geq \lambda \iff s \geq g^{(-1)}(\lambda)$$

En conseqüència, podem escriure la igualtat següent:

$$\mathcal{X}_\lambda(g(u)) = \{(x, y) : g(u(x, y)) \geq \lambda\} = \{(x, y) : u(x, y) \geq g^{(-1)}(\lambda)\} = \mathcal{X}_\mu u$$

on $\mu = g^{(-1)}(\lambda)$.

Per una altra banda, si g és estrictament creixent, llavors

$$\{(x, y) : u(x, y) \geq \mu\} = \{(x, y) : g(u(x, y)) \geq g(\mu)\}$$

i seguint un raonament similar a l'anterior tenim que

$$\mathcal{X}_\mu u = \{(x, y) : u(x, y) \geq \mu\} = \{(x, y) : g(u(x, y)) \geq g(\mu)\} = \{(x, y) : u(x, y) \geq g^{(-1)}(g(\mu))\} = \mathcal{X}_\lambda$$

on $\lambda = g^{(-1)}(g(\mu))$.

Propietats.

1. Donades dues imatges u i v que tenen globalment els mateixos conjunts de nivell, podem trobar un canvi de contrast g creixent tal que $v = g(u)$. A més, g es pot definir per la fórmula

$$g(\lambda) = \sup\{\mu, \mathcal{X}_\mu v \supset \mathcal{X}_\lambda u\}$$

2. Si g no és estrictament creixent ni contínua però és contínua per la drete², llavors podem al menys assegurar que tots els conjunts de nivell de $g(u)$ també ho són de u , però no a la inversa.

La Figura 7 mostra dues imatges relacionades per un canvi de contrast i un conjunt de nivell comú a totes dues, en el primer cas corresponent al nivell 235 i en l'altre al nivell 130.



Figura 7: D'esquerra a dreta, dues imatges relacionades per un canvi de contrast i un conjunt de nivell comú a ambdues. El conjunt de nivell correspon al nivell 235 per a la primera imatge i al nivell 130 per a la segona.

Conclusió. Hem demostrat que els conjunts de nivell ens donen una descripció de la imatge invariant per canvis de contrast. Significa això que els podem considerar com les estructures bàsiques de la imatge? En la secció següent veurem que és possible simplificar la descripció de la imatge donada pels conjunts de nivell si ens quedam únicament amb les fronteres d'aquests conjunts. Aquesta simplificació dóna lloc al **mapa topogràfic** de la imatge.

3.3. Línies de nivell i mapa topogràfic

Donada una imatge $u : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$, definida en un subconjunt tancat Ω de \mathbb{R}^2 , anomenam **línia de nivell** λ de u al conjunt de punts

$$u^{-1}(\lambda) = \{(x, y) \in \Omega : u(x, y) = \lambda\} \quad (7)$$

²Recordem que una funció g és contínua per la drete en un punt s si

$$\lim_{r \rightarrow s, r > s} g(r) = g(s)$$

Teorema. Si $u : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$, amb Ω tancat, és C^1 llavors, per a quasi tots els nivells λ , la línia de nivell

$$u^{-1}(\lambda) = \{(x, y) \in \Omega : u(x, y) = \lambda\}$$

està formada per una unió finita de corbes de Jordan³ tancades de classe C^1 . Els únics valors de λ per als quals no es compleix aquesta propietat són aquells per als quals existeix un $(x, y) \in u^{-1}(\lambda)$ tal que $|\nabla u|(x, y) = 0$. Aquests nivells reben el nom de **nivells singulars**.

Demostració. El teorema és conseqüència dels teoremes de la funció implícita i de Sard⁴.

Observació. La hipòtesi de continuïtat de u i de les seves derivades és perfectament aplicable ja que, tal com s'ha argumentat en el Tema 4, una imatge digital s'ha d'interpretar com la versió discreta d'una funció contínua subjacent, d'amplada de banda limitada, la qual és interpolable a partir dels valors discrets mitjançant la interpolació de Shannon (Tema 2), i per tant, de classe C^∞ .

Definició. Anomenam **mapa topogràfic** d'una imatge $u : \Omega \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ al conjunt de corbes de Jordan obtingudes en calcular les línies de nivell per a tots els valors de λ no singulars de u .

Teorema de les corbes de Jordan [Rudin]. Una corba tancada de Jordan C divideix el pla en dues components connexes, una acotada i l'altra no. Anomenam **interior** de C , i ho denotam $\text{Int}(C)$, a la component connexa acotada. Anomenam **exterior** de la corba a la component connexa no acotada.

Una conseqüència del teorema anterior és que podem associar a cada corba del mapa topogràfic una **orientació**, segons si el vector $\nabla u(x, y)$ està dirigit cap a l'interior o l'exterior de la línia de nivell en tots els seus punts. En el primer cas parlem d'orientació **positiva** i en el segon **negativa**.

3.4. Organització del mapa topogràfic en una estructura d'arbre

Propietat. Les corbes del mapa topogràfic mai es creuen. O bé estan incloses unes dins les altres, o bé els seus interiors són disjunts (veure la Figura 8).

Demostració. (Exercici).

Aquesta propietat ens permet organitzar les línies de nivell del mapa topogràfic dins una estructura d'arbre. La Figura 9 mostra un exemple de mapa topogràfic per a una imatge senzilla i l'estructura d'arbre que se'n deriva.

3.5. Mapa topogràfic i conjunts de nivell

Teorema. Si u és de classe C^1 , la línia de nivell $u^{-1}(\lambda)$ del mapa topogràfic és la frontera del conjunt de nivell $\mathcal{X}_\lambda u$:

$$u^{-1}(\lambda) = \partial \mathcal{X}_\lambda u$$

La Figura 10 mostra un de conjunt de nivell i la seva línia de nivell associada.

³Una corba contínua es diu de Jordan si no té cap autointersecció, excepte possiblement en els seus punts extrems. Més formalment, C és una corba de Jordan si podem trobar una aplicació contínua

$$\mathbf{x} : [a, b] \subseteq \mathbb{R} \longrightarrow \mathbb{R}^2$$

tal que \mathbf{x} és bijectiva en (a, b) i $g([a, b]) = C$. Si $\mathbf{x}(a) = \mathbf{x}(b)$ llavors es diu que la corba de Jordan és tancada.

⁴**Teorema de Sard**[?]. Sigui u una funció real C^1 de finida en un rectangle R de \mathbb{R}^2 , llavors per a quasi tot $\lambda \in \mathbb{R}$ el conjunt $u^{-1}(\lambda)$ és no singular, és a dir, $\forall (x, y) \in u^{-1}(\lambda)$, $|\nabla u(x, y)| \neq 0$.

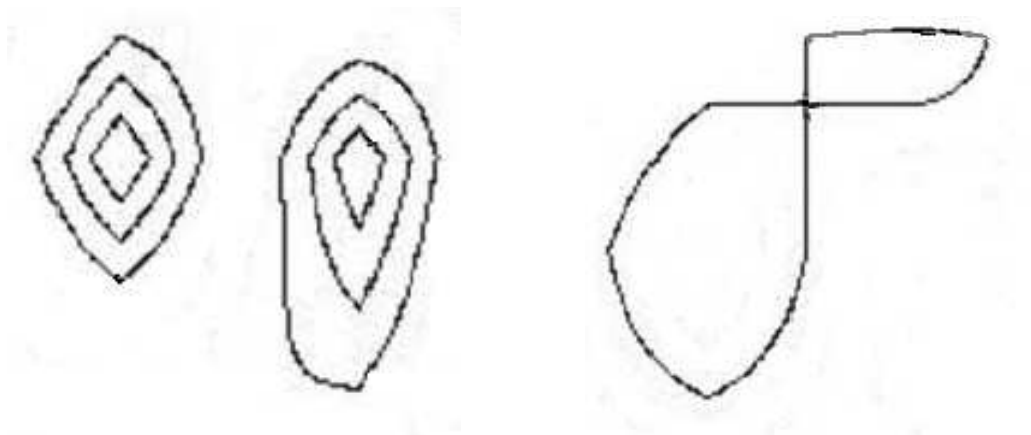


Figura 8: La primera imatge mostra alguns exemples de corbes del mapa topogràfic. Les corbes estan incloses unes dins les altres o bé els seus interiors són disjunts. La segona imatge mostra una corba que no pot pertànyer al mapa topogràfic, ja que es creua amb ella mateixa.

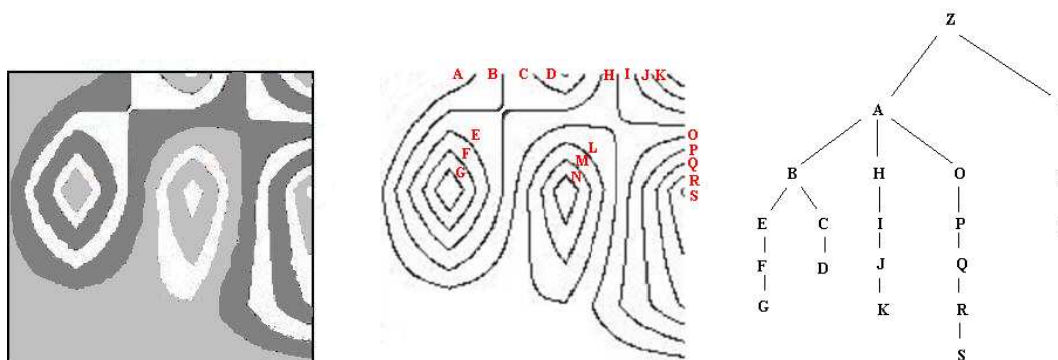


Figura 9: D'esquerra a dreta, imatge original, algunes línies del seu mapa topogràfic i estructura d'arbre del mapa topogràfic. La corba etiquetada com Z representa el contorn de la imatge.

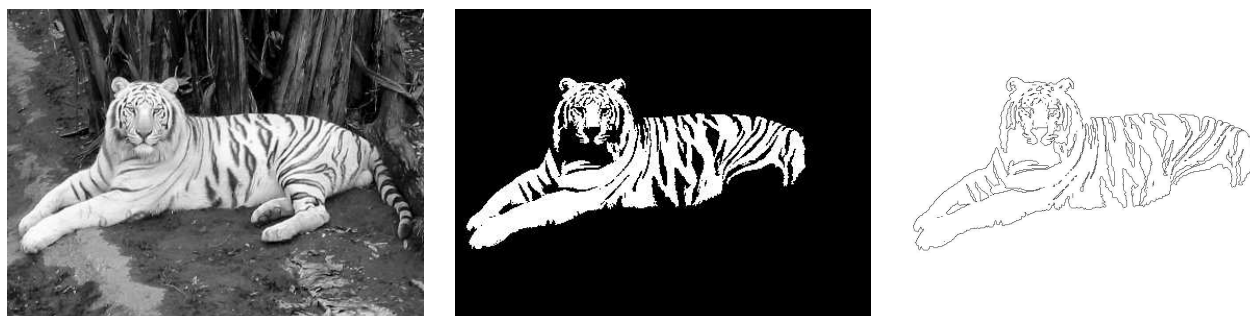


Figura 10: D'esquerra a dreta: imatge original, un dels seus conjunts de nivell i la línia de nivell del conjunt.

El conjunt de nivell $\mathcal{X}_\lambda u$ es pot obtenir a partir de la seva frontera. Si anomenem c_i les corbes de $u^{-1}(\lambda)$ orientades positivament i c_{ij} les corbes contingudes en l'interior de c_i orientades negativament, llavors

$$\mathcal{X}_\lambda = \cup_i (\text{Int}(c_i) - (\cup_{c_{ij} \subset \text{Int}(c_i)} \text{Int}(c_{ij})))$$

Intuitivament, l'anterior expressió equival a 'omplir' primer les corbes c_i per llevar després els seus forats.

Com que el conjunt de λ 's no singulars per als quals hem calculat el mapa topogràfic és dens (pel teorema de Sard, el conjunt de valors singulars de λ és de mesura nula), les anteriors propietats ens asseguren que podrem trobar els conjunts de nivell de la imatge per a quasi tots els possibles λ . A partir d'aquests conjunts de nivell podrem reconstruir la imatge original mitjançant la fórmula de reconstrucció (6). En conclusió:

Proposició. Sigui u una funció C^1 definida en un conjunt tancat de \mathbb{R}^2 . Llavors u es pot recuperar a partir del seu mapa topogràfic.

Conclusions. El conjunt de línies de nivell d'una imatge (el mapa topogràfic) hereta les propietats d'invariància als canvis de contrast dels conjunts de nivell. A més, el mapa topogràfic representa una estructura més simple que la dels conjunts de nivell, a partir de la qual és sempre possible reconstruir aquests conjunts de nivell. Per aquest motiu **el mapa topogràfic és el model que utilitzarem per representar les imatges de manera invariant per canvi de contrast**. També direm que el mapa topogràfic conté la **informació geomètrica o morfològica** de la imatge, en el sentit que ens permet descriure la *forma* dels objectes que apareixen dins l'escena, independentment del seu nivell de gris o color.

3.6. Mapa topogràfic d'una imatge digital

Una manera pràctica de calcular el mapa topogràfic d'una imatge digital és a partir de la seva versió contínua obtinguda per interpolació bilineal a trossos.

Per a cada grup de quatre píxels (i, j) , $(i+1, j)$, $(i, j+1)$, $(i+1, j+1)$ amb nivells de gris respectius λ_1 , λ_2 , λ_3 i λ_4 , es defineix una funció $u_{ij}(x, y) = axy + bx + cy + d$ tal que $u(0, 0) = \lambda_1$, $u(1, 0) = \lambda_2$, $u(0, 1) = \lambda_3$ i $u(1, 1) = \lambda_4$. La següent figura mostra un exemple d'aquesta interpolació.

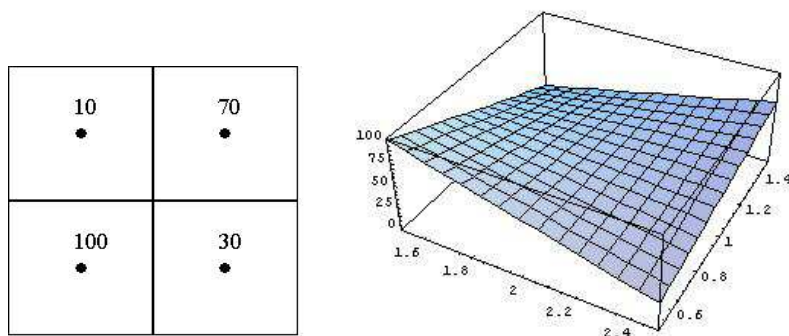


Figura 11: Exemple d'interpolació bilineal dels valors de quatre píxels.

El conjunt de funcions u_{ij} obtingudes per a cada grup de quatre píxels de la imatge dona lloc a una funció bilineal a trossos com la mostrada en la Figura 12.

L'avantatge d'aquest mètode és que és fàcil de calcular i no crea nous màxims o mínims en la imatge, ja que, per a cada grup de píxels, els valors de la funció interpolada estan compresos entre els valors màxim i mínim dels píxels. Les línies de nivell obtingudes d'aquesta manera són contínues i estan formades per la concatenació d'hipèrboles. A més, les propietats d'inclusió comentades en les seccions anteriors es continuen

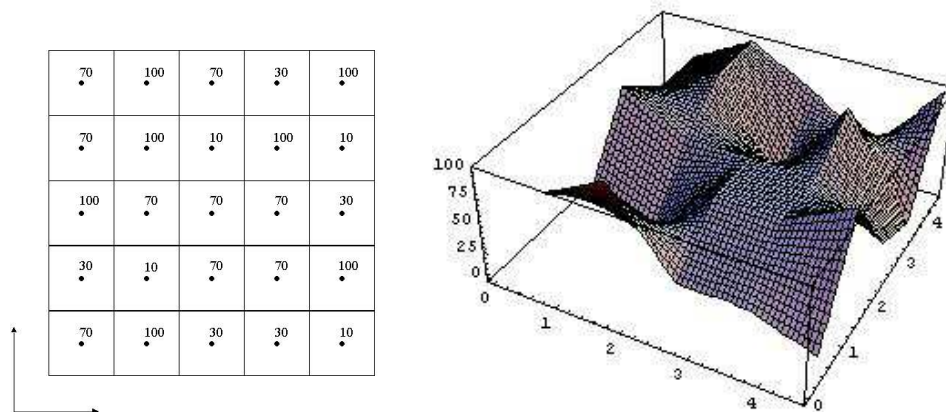


Figura 12: D'esquerra a dreta, imatge original (per a cada pixel s'indica el seu valor de nivell de gris) i imatge obtinguda per interpolació bilineal.

cumplint. Les línies de nivell es calculen per a un conjunt discret de valors no singulars λ i s'organitzen en una estructura d'arbre.

La Figura 13-esquerra mostra un exemple d'algunes de les línies de nivell obtingudes per a la imatge de la Figura 12. Finalment, comentar que les línies que arriben al contorn de la imatge es tanquen seguint el camí més curt damunt d'aquest contorn, tal com mostra la Figura 13-dreta.

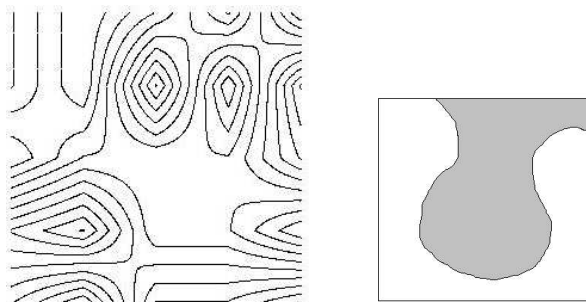


Figura 13: Esquerra, mapa topogràfic de la imatge de la Figura 12. Dreta, criteri per al tancament de les corbes que arriben al contorn de la imatge.

3.7. Aplicacions del mapa topogràfic

3.7.1. Comparació morfològica d'imatges

Siguin dues imatges u_1 i u_2 de la mateixa escena tals que, per a tots els pixels (i, j) , els valors $u_1(i, j)$ i $u_2(i, j)$ corresponen mesuren la intensitat de la llum en el mateix punt de l'escena. En tal cas es diu que les imatges estan perfectament **posades en correspondència** (*registered*). Volem detectar els canvis en l'escena deguts a la presència de nous objectes, però no els deguts a canvis en la il·luminació. Aquesta situació és típica en la comparació d'imatges de satèl·lit preses en dates diferents i en les quals volem detectar la construcció o destrucció d'edificis, la variació dels camps de cultiu, etc.

Les tècniques de comparació basades en variacions de nivell de gris no permetran la detecció de tals canvis, ja que entre dates diferents les condicions d'il·luminació poden variar bastant. En aquests casos, la comparació dels mapes topogràfics és un mètode eficaç per trobar els canvis. La següent propietat ens permet dissenyar un algorisme ràpid detecció:

Propietat. Per un teorema clàssic de l'anàlisi de funcions de vàries variables podem afirmar que les línies de nivell són perpendiculars a la direcció del vector gradient en tots els punts en què aquest gradient existeix.

La figura 14 il·lustra la propietat anterior. La diferència de l'angle del vector gradient entre dos punts de correspondència de les imatges ens dóna una mesura de la variació local del mapa topogràfic en els punts.

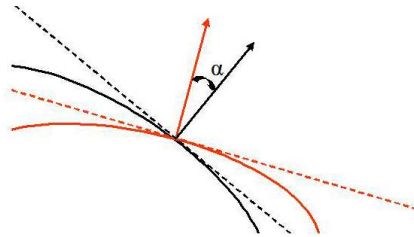


Figura 14: La figura mostra, en diferents colors, dues línies de nivell corresponents a dues imatges de la mateixa escena. Suposam que les imatges estan perfectament posades en correspondència, de manera que el pixel (x, y) de la primera imatge es correspon amb el pixel (x, y) de la segona. La variació local del mapa topogràfic entre les imatges es pot quantificar calculant, per a cada pixel, l'angle entre vectors gradients en cada imatge.

Aprofitant aquesta propietat s'ha dissenyat el següent algorisme:

1. Calcular el gradient en cada punt de les imatges (utilitzar els mètodes de discretització del valor del gradient descrits en el Tema 5).
2. Calcular el mòdul i l'angle del gradient (respecte a la direcció horitzontal) en cada punt de les imatges. Anomenam $M_1(i, j)$ i $\alpha_1(i, j)$ al mòdul i l'angle del gradient del pixel (i, j) de la primera imatge; respectivament, $M_2(i, j)$ i $\alpha_2(i, j)$ representen el mòdul i l'angle del gradient del pixel (i, j) de la segona imatge.
3. Si $M_1(i, j)$ i $M_2(i, j)$ són majors que un cert llindar (típicament 5), calculam la diferència entre els seus angles, $\Delta A(i, j) = |\alpha_1(i, j) - \alpha_2(i, j)|$.
4. La diferència anterior representa la variació del mapa topogràfic en cada punt de la imatge. Representam amb el següent codi de colors la imatge 'diferència': blanc, els pixels amb $\Delta A(i, j) < \delta$; els pixels amb

$\Delta A(i, j) \geq \delta$; i gris els píxels tals que $M_1(i, j)$ o $M_2(i, j)$ és inferior al llindar de gradient. δ és un llindar que determina a partir de quina diferència d'angle consideram diferents les orientacions del gradient (típicament 20°).

5. La imatge resultant ens dona una idea gràfica del grau de variació dels mapes topogràfics de les imatges: com més punts blacs, més s'assemblen les imatges.

Les següents figures mostren els resultats de la comparació per a dues imatges d'una ciutat i per un llindar de comparació de 20° i un gradient mínim de 5.

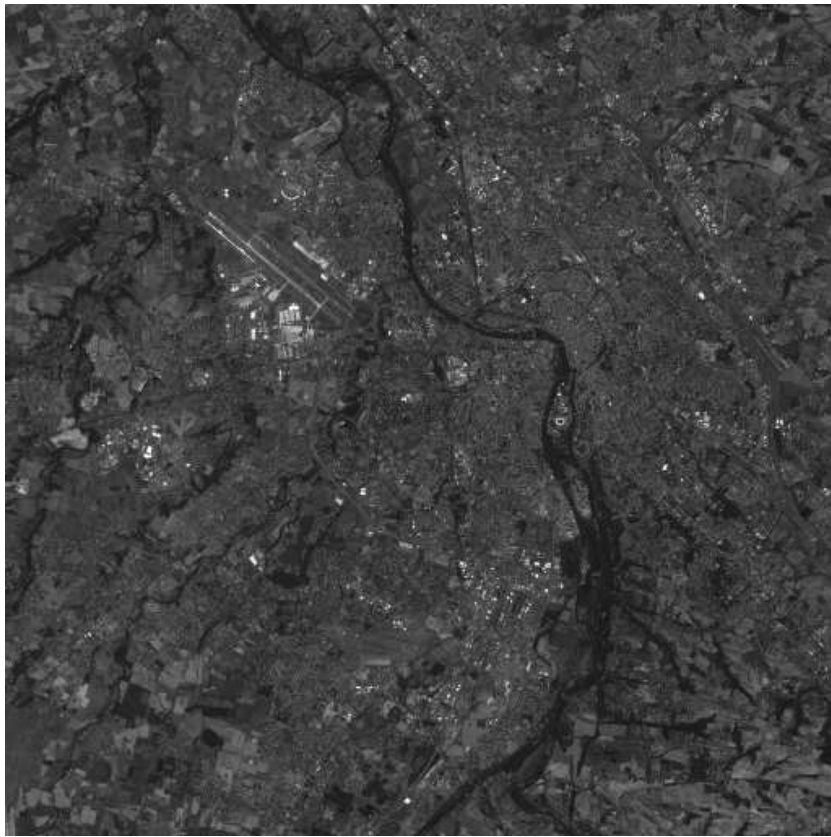


Figura 15: Primera de les imatges a comparar.

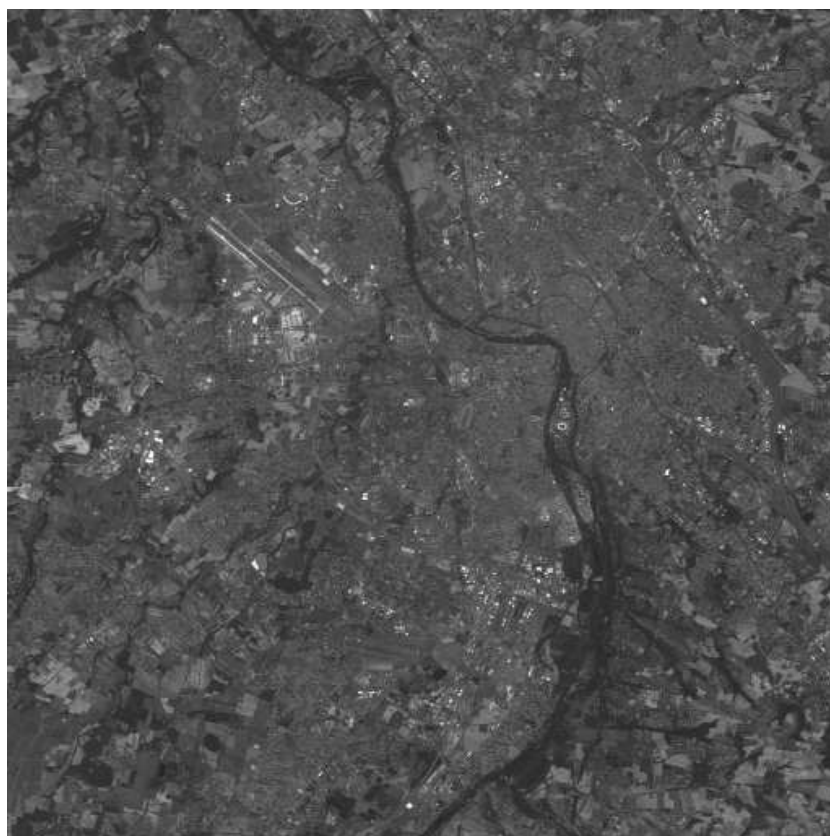


Figura 16: Segona de les imatges a comparar. Les imatges han estat preses en dates diferents.

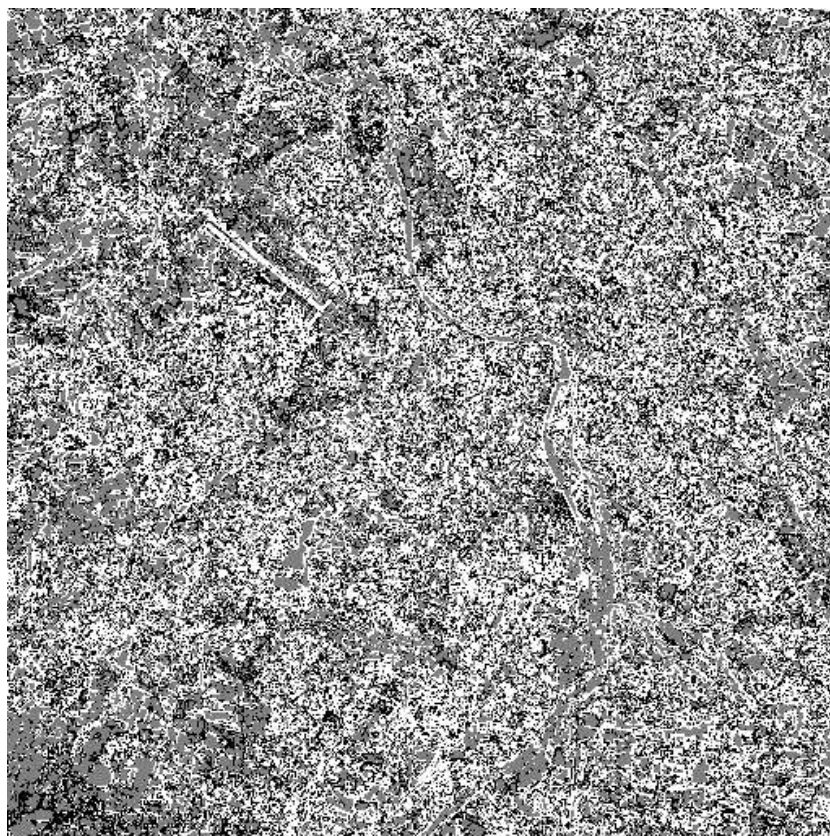


Figura 17: Resultat de la comparació de les imatges anteriors. S'ha mesurat el canvi en la direcció dels seus gradients amb llindars 5 i 20° per al mòdul i la diferència d'angle dels gradients, respectivament. Els punts grisos representen píxels amb mòdul de gradient insuficient en alguna de les imatges; els punts blaus són punts on l'orientació del gradient no canvia i els negres on l'orientació canvia.

3.7.2. Filtratge: filtre de gra [GuichardMonasse]

En general, les formes petites de la imatge (moltes vegades degudes al renou) són difícilment percebudes. Podem aprofitar aquesta característica de la visió per simplificar les imatges i eliminar-ne el renou.

L'organització en forma d'arbre del mapa topogràfic resulta especialment indicada per eliminar les formes petites de la imatge. Una forma petita però ben contrastada en la imatge donarà lloc a una component connexa per a algun conjunt de nivell de la imatge; el contorn d'aquesta component connexa serà una corba del mapa topogràfic. Ja que la forma és petita, la corba també ho serà i es trobarà, generalment, en les 'fulles' de l'estructura d'arbre del mapa topogràfic. Per tant un algoritme que 'podí' l'arbre aconseguirà eliminar aquestes formes. Resta el problema d'assignar un nou nivell de gris als pixels pertanyents a les formes eliminades. El criteri seguit és assignar-los el nivell de gris de la corba de la mateixa orientació més petita en la què estava contingut el contorn de la forma. La Figura 18 mostra un exemple senzill de simplificació del mapa topogràfic.

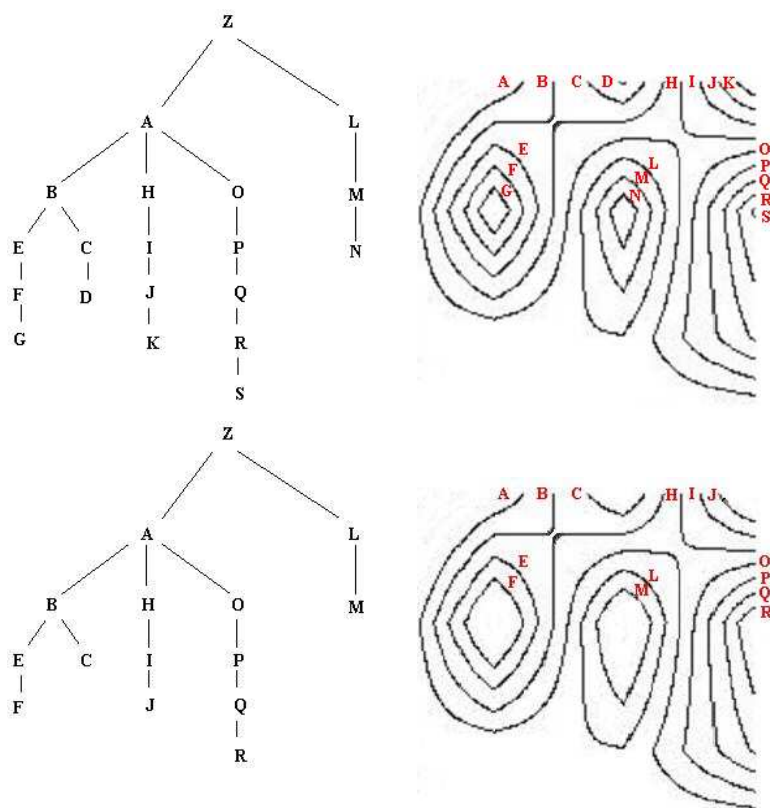


Figura 18: Exemple de simplificació del mapa topogràfic mitjançant el filtre de gra. Les imatges de la part superior mostren un mapa topogràfic i la seva estructura d'arbre. El contorn de la imatge, que constitueix l'arrel de l'arbre, s'ha marcat amb la lletra Z. A més, per a aquest exemple suposam que l'orientació de totes les corbes és la mateixa. Donat un paràmetre del filtre γ , si l'interior de les corbes G, N, S, D i K té un àrea inferior a γ , llavors aquestes corbes s'eliminen de l'arbre. Les imatges de la part inferior mostren el resultat del filtratge. Els nous nivells de gris assignats als pixels de l'interior de les corbes eliminades són els de les corbes més petites que els contenen. Així, als pixels continguts en la corba G s'els assigna el nivell de la corba F, als de la corba N els de la corba M, etc.

La següent figura mostra l'efecte d'eliminar les formes de tamany (en nombre de pixels) 5, 10, 20 i 40 d'una imatge.



Figura 19: D'esquerra a dreta i de dalt a baix, imatge original i imatges resultants d'eliminar les formes de tamany 5, 10, 20 i 40. Observar com el nivell de detall de la imatge es va reduint.

3.7.3. Posada en correspondència d'imatges (*registration*)

Posar en correspondència (*register*) dues imatges significa calcular la transformació geomètrica que aplica els pixels d'una imatge en els de l'altra imatge. Aixó significa que, una vegada posades en correspondència, el pixel (i, j) de cada imatge representa el nivell d'il·luminació procedent del mateix punt de l'espai tridimensional.

S'han desenvolupat dos mètodes de *registration* basats en el mapa topogràfic de les imatges.

Registration basat en formes [Monasse]

Definim una *forma* de la imatge com una component connexa d'algun dels seus conjunts de nivell. Per tant, una forma no és més que un subconjunt de \mathbb{R}^2 i pot ésser caracteritzada per una sèrie de valors: àrea, posició del centre de masses, moments de segon ordre, etc. La similitud entre dues formes pot ésser llavors mesurada en termes de les diferències entre aquests valors.

Donades dues imatges, si som capaços de trobar les correspondències entre les formes d'una i l'altra, podrem estimar com es mouen els objectes d'una imatge respecte als de l'altra i, si la majoria dels objectes es desplacen amb el mateix moviment, serà llavors possible estimar el moviment global entre les imatges.

L'algoritme de recalatge basat en les formes és el següent:

1. Calcular el mapa topogràfic de les dues imatges
2. Per a cada corba tancada del mapa, considerar el seu interior, el qual és una component connexa (*forma*) d'algun conjunt de nivell.
3. Per a cada *forma* calcular els següents moments: àrea (moment d'ordre 0), centre de masses (moments d'ordre 1) i moments centrats de segon ordre.

En una imatge digital, el moment d'ordre ij d'una forma A es calcula com

$$m_{ij} = \sum_{(x,y) \in A} x^i y^j$$

i el moment centrat d'ordre ij és

$$\mu_{ij} = \sum_{(x,y) \in A} (x - m_{10})^i (y - m_{01})^j$$

Els exponents i i j en les expressions anteriors fan que el càlcul de moments d'ordres superiors a 2 sigui molt sensible a petites variacions en els valor de (x, y) , degudes al renou.

4. Calcular els *invariants* de cada forma. La informació proporcionada pels moments no és suficient per comparar dues formes sotmeses a un moviment arbitrari. Per exemple, el valor de l'àrea no varia quan la forma es sotmet a una translació o rotació, però si quan hi ha un canvi d'escala (zoom). Per aquest motiu es cerquen combinacions de moments que siguin el més invariants possibles. Els següents valors són invariants a translacions, rotacions i canvis d'escala:

$$s_1 = \frac{\mu_{20} + \mu_{02}}{m_{00}^2}$$

$$s_2 = \frac{\mu_{20}\mu_{02} - \mu_{11}^2}{m_{00}^2}$$

Per trobar invariants a altres tipus de moviment (com, per exemple, transformacions afins) hauriem d'utilitzar moments d'ordre superior que, com ja hem dit, són poc fiables per ésser molt sensitius al re-nou. De manera que el mètode servirà únicament per registrar imatges sotmeses a una transformació de similitud (translació, rotació i canvi d'escala).

5. Comparar les formes. Es defineixen uns llindars de tolerància ϵ_1 i ϵ_2 tals que, donades dues formes A i B , si $|s_1^A - s_1^B| < \epsilon_1$ i $|s_2^A - s_2^B| < \epsilon_2$, llavors consideram que A i B són la mateixa forma.

Per a cada forma de la primera imatge obtindrem una llista de formes similars en la segona imatge.

6. Calcular el moviment entre parells de formes. Els moments calculats en els punts anteriors no ens permeten calcular el moviment entre les formes d'una i altra imatge. Com a màxim, si coneixem que la forma A de la primera imatge es correspon amb la forma A' de la segona imatge, podríem calcular la translació entre els seus centres de masses. Per calcular la rotació i el canvi d'escala entre les formes feim el següent:
 - a) Consideram les formes A i B en la primera imatge i les formes A' i B' de la segona, tals que A' és similar a A i B' és similar a B (per als llindars establerts en el punt anterior).
 - b) Formam els vectors \vec{AB} i $\vec{A'B'}$ i calculam la transformació (translació, rotació i zoom) que aplica un vector damunt l'altre.
7. Calcular el moviment global entre les imatges. Per a cada parell de formes de la primera imatge repetim el procés anterior. Enmagatzemam els valors de rotació, translació i zoom trobats en un histograma i calculam el màxim pic de l'histograma. Aquest pic ens dóna el moviment dominant entre les imatges.
8. Posar en correspondència les imatges amb el valor de moviment trobat.

Les següents figures il.lustren els diferents passos de l'algoritme:

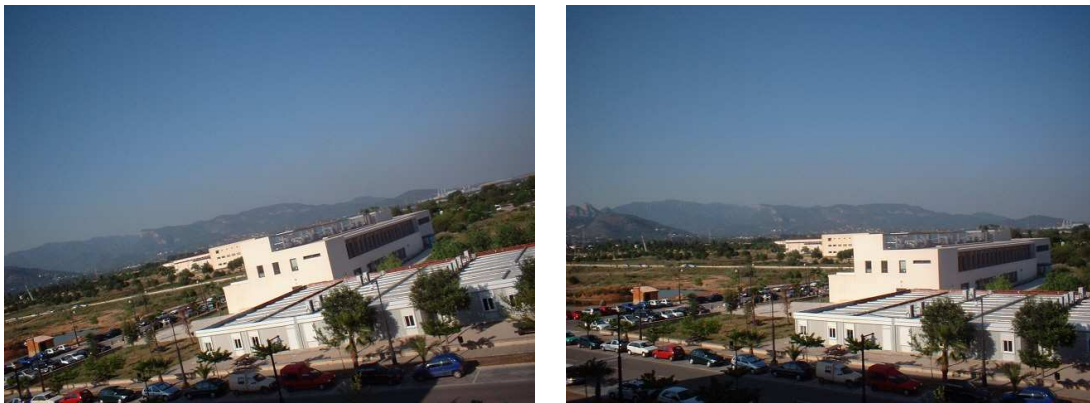


Figura 20: Imatges originals. Observar l'existència d'una part comú entre les imatges.



Figura 21: Exemple d'algunes de les formes que es corresponen entre les imatges.



Figura 22: Exemple d'un dels parells de formes utilitzats per calcular el moviment entre les imatges.

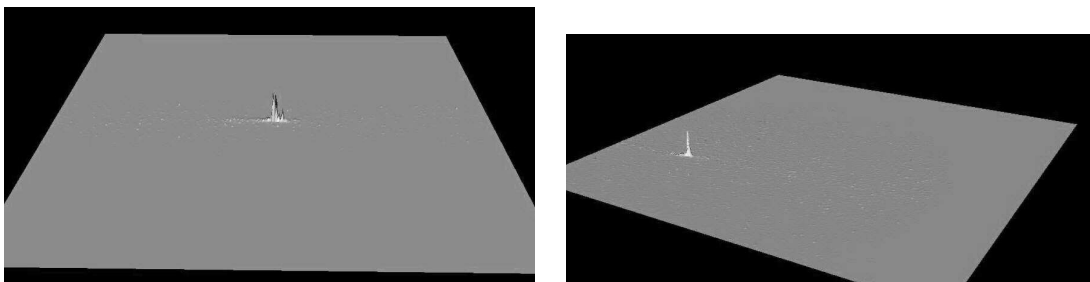


Figura 23: Histogrames (representats en tres dimensions) de rotació-zoom i translació horitzontal-vertical que ens permetran calcular el moviment global entre les imatges.



Figura 24: Mosaic: resultat de aplicar el moviment trobat a la primera imatge.

Registration basat en peces de línies de nivell [LMMM]

El mètode de posada en correspondència basat en les formes presenta dos inconvenients principals:

1. No permet calcular transformacions afins entre les imatges.
2. És sensible a oclusions parcials.

Per il·lustrar aquest darrer inconvenient consideram les imatges de la Figura 25.

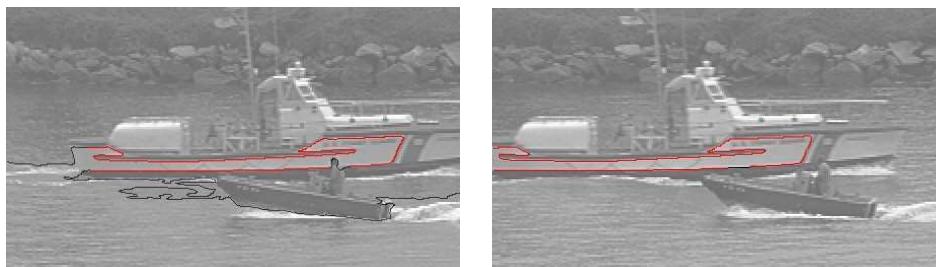


Figura 25: Dues formes entre les quals hi ha una oclusió parcial. La part comúna entre les línies de nivell s'ha marcat en color vermell.

Observam com les formes mostrades en les imatges corresponen als mateixos objectes, no obstant la segona d'elles presenta una oclusió que impedeix compararla amb la primera amb el mètode dels moments. Aquesta comparació serà possible únicament si som capaços de codificar de manera local el contorn de la forma.

L'observació anterior ens duu a desenvolupar un nou mètode de *registration* basat en la codificació local de les línies de nivell. Aquesta codificació es fa de forma invariant a transformacions afins del pla de les

imatges i permet comparar les peces de línies de nivell d'una i altra imatge. La informació de moviment calculada per a cada peça de corba s'enmagatzema en un histograma i es calcula el moviment global entre les imatges de manera similar al mètode anterior.

Les següents figures mostren alguns exemples de l'aplicació del mètode.

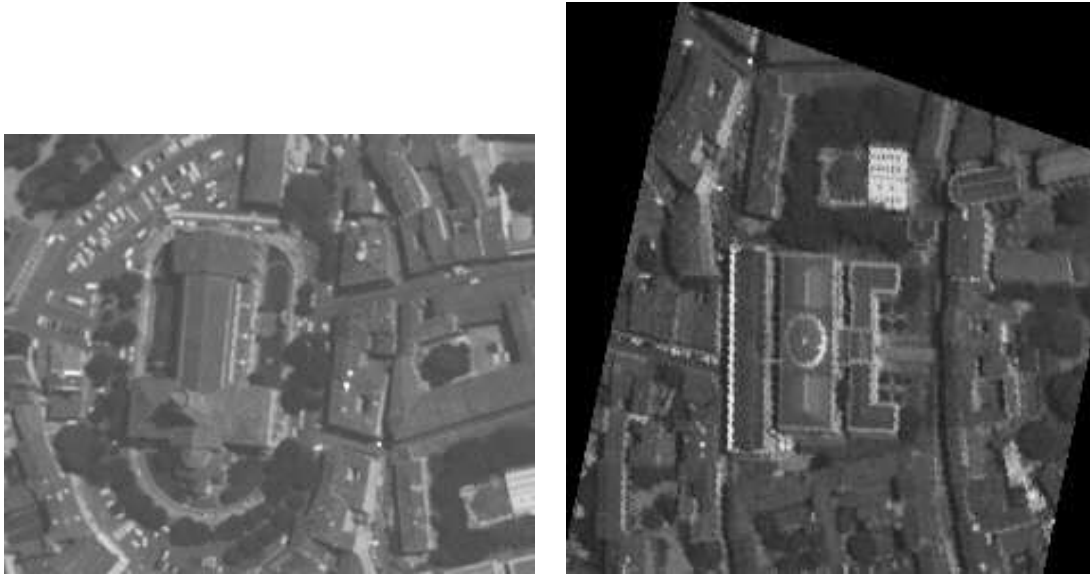


Figura 26: Imatges originals. L'imatge de la dreta presenta una part comú amb la imatge de l'esquerra (cantó superior esquerre). Les imatges pertanyen a canals espectrals diferents (blau i vermell) i existeix una transformació afí entre elles.



Figura 27: Peces de línies de nivell comunes entre les imatges i mosaic de les imatges anteriors obtingut a partir de la informació de moviment proporcionada pel nou mètode de posada en correspondència.



Figura 28: Imatges originals.

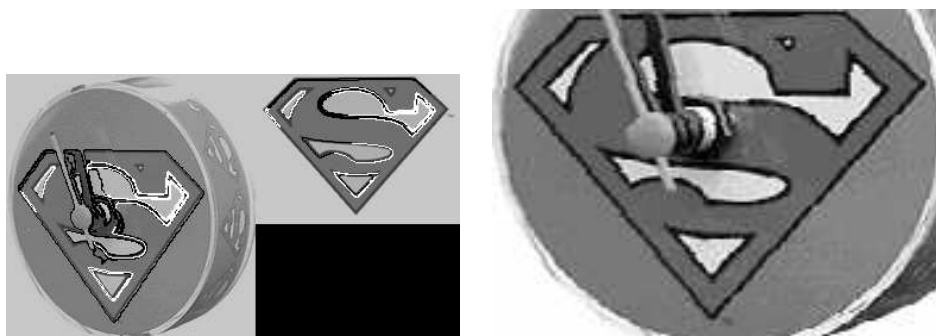


Figura 29: Línies de nivell comunes i resultat d'aplicar el moviment estimat a una de les imatges.

4. Referències

- [GuichardMorel] F. Guichard, J.M. Morel. *Iterative smoothing and PDE's*. Llibre en preparació, 2002.
- [Rudin] W. Rudin. *Real and complex analysis*, 2^a edició. McGraw-Hill, 1974.
- [GuichardMonasse] F. Guichard, P. Monasse. *Fast computation of a contrast-invariant image representation*. IEEE Transactions on Image Processing, vol. 20, gener, 1999.
- [Monasse] P. Monasse. *Contrast invariant image registration*. Proc. Int. Conference on Acoustics, Speech and Signal Processing, vol. 6, pp. 3221-3224, 1999.
- [LMMM] J.L. Lisani, L. Moisan, P. Monasse, J.M. Morel. *On the theory of planar shape*. SIAM Multiscale, Model. Simul. Vol. 1, No. 1, pp. 1-24, Gener, 2003.