

# Un viatge al Món de les Imatges

DISS4BTES DE LES  
M4TEM4TIQUES 2019

Bartomeu Coll  
Universitat de les Illes Balears

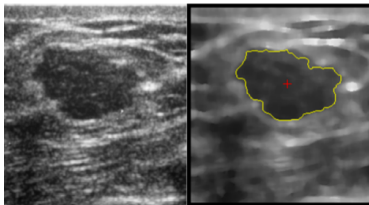
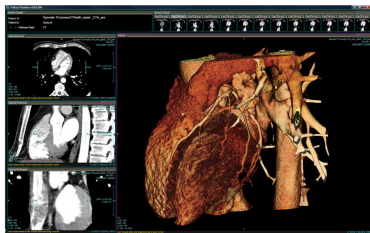


# Exemples

El nostre món real està ple d'imatges

- ▶ fotografies (mòbils, càmeres digitals, ...)
- ▶ imatges video (video comercial, video-vigilància, etc)
- ▶ Imatge satèl.lit (metereologia, terra, sol, altres)
- ▶ Imatge mèdica (Radiografia, TAC, Resonància magnètica, ...)
- ▶ Spots comercials (fotografia professional)
- ▶ etc, etc

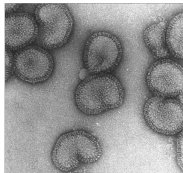
## Exemples: imatge mèdica



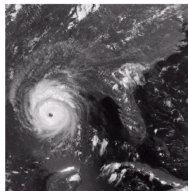
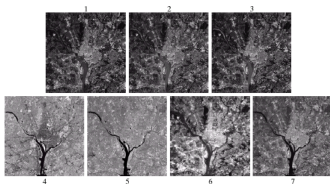
## Exemples: tecnologia



NASA, Telescopio Hubble



Rob Ruigrok, UVHCI





## Exemple: imatge satèl·lit (Pleiades)



## Exemples: fotografia digital

Fotografia feta amb una Nikon D70 18-55mm ISO 200



## Exemples d'aplicacions



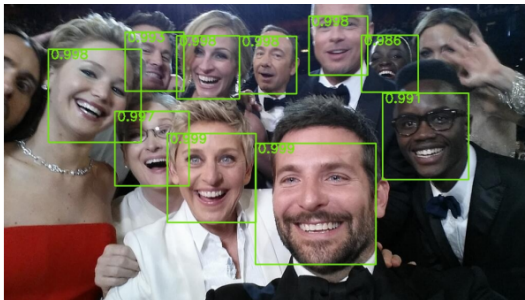
## Exemples d'aplicacions



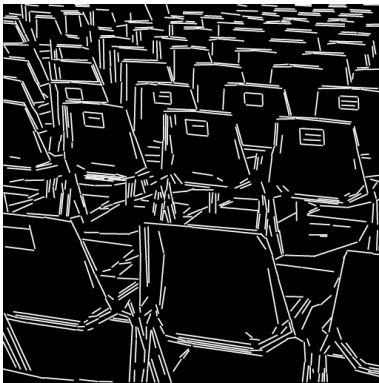
## Exemples d'aplicacions



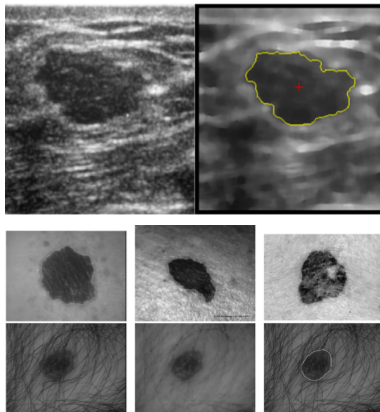
## Exemples d'aplicacions



## Exemples d'aplicacions



## Exemples d'aplicacions





## Exemples d'aplicacions (i)



## Exemples d'aplicacions (ii)



# Outline

Que és una imatge digital?

Contrast i color

Com millorar la qualitat d'una imatge

La detecció d'objectes en una imatge digital

La visió estèreo i la reconstrucció 3D

# Outline

Que és una imatge digital?

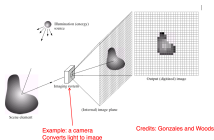
Contrast i color

Com millorar la qualitat d'una imatge

La detecció d'objectes en una imatge digital

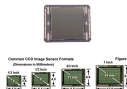
La visió estèreo i la reconstrucció 3D

# Com es forma una imatge digital



- ▶ La llum reflectida per l'objecte (fotons), arriba al sensor de la càmera.
- ▶ Aquest sensor (o matriu anomenat CCD) està format per multitud de receptors (elements de la matriu).

L'element més important de la càmera digital és una targeta anomenada CCD (couple charged device).



- ▶ La llum incident genera un petit senyal elèctric en cada receptor
- ▶ Aquest senyal es transforma en dades digitals (zeros i uns, anomenats dígit binaris).
- ▶ Finalment, aquestes dades de la matriu de sensors dona lloc als píxels.

# Pixels i resolució de la imatge digital



$$M = \begin{bmatrix} 160 & 162 & \dots & 166 & 154 \\ 162 & 158 & \dots & 122 & 69 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 60 & 55 & \dots & 79 & 94 \\ 58 & 55 & \dots & 99 & 109 \end{bmatrix}$$

- Three components

–  $M = \{R, G, B\}$

$$R = \begin{bmatrix} 73 & \dots & 87 \\ \vdots & \ddots & \vdots \\ 27 & \dots & 17 \end{bmatrix}, G = \begin{bmatrix} 66 & \dots & 98 \\ \vdots & \ddots & \vdots \\ 36 & \dots & 13 \end{bmatrix}, B = \begin{bmatrix} 31 & \dots & 61 \\ \vdots & \ddots & \vdots \\ 36 & \dots & 14 \end{bmatrix}$$



R



G

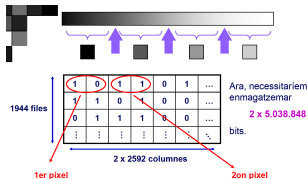
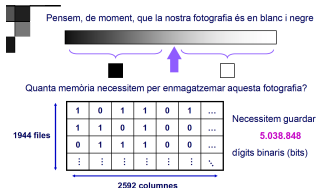


B



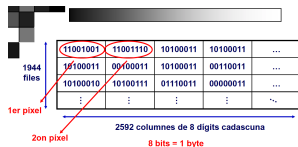
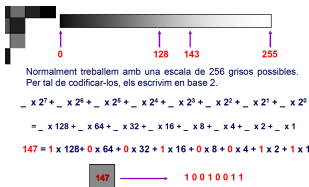
- ▶ Un pixel ens dona la informació del color a cada element  $a_{i,j}$  de la matriu CCD (un nombre enter).
- ▶ Resolució espacial d'una imatge és la quantitat de pixels.
- ▶ També s'empra amb dispositius relacionats amb pantalles d'ordinador, televisió, impresores, escàners, etc.

# Com emmagatzemar aquesta imatge? (I)



- ▶ En una imatge de dos colors (blanc i negre), cada píxel és un bit, 0 o 1 (imatge binària).
- ▶ En una imatge de quatre nivell de grisos, cada píxel són dos bits, codificats en base 2.

## Com emmagatzemar aquesta imatge? (II)



Per tant, per al nostre cas, haurem de guardar 5.038.848 de bytes, és a dir la fotografia ocuparà 5 Megabytes.

- ▶ Si els valors dels píxels estan dins  $[0, 255]$ , cada píxel són 8 bits codificats en base 2.
- ▶ En una imatge, 8 bits equival a 1 byte.



# Outline

Que és una imatge digital?

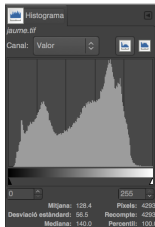
Contrast i color

Com millorar la qualitat d'una imatge

La detecció d'objectes en una imatge digital

La visió estèreo i la reconstrucció 3D

# Contrast en les imatges digitals



# Transformacions sobre l'histograma

## definició de l'histograma

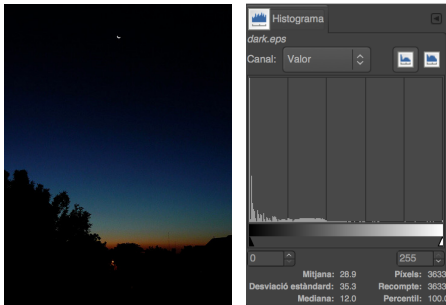
L'histograma de una imatge digital de dimensions  $n_x \times n_y$ , amb  $L$  el nombre total de nivells en el rang  $[min, max]$  es defineix com la funció discreta

$$h(l_k) = n_k,$$

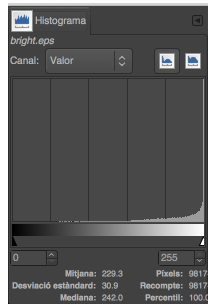
on  $l_k$  és el nivell d'intensitat  $k$  a l'interval  $[min, max]$  i  $n_k$  és el nombre de píxels en la imatge que tenen com a nivell d'intensitat  $l_k$ .

L'histograma ens dóna una descripció global de la imatge.

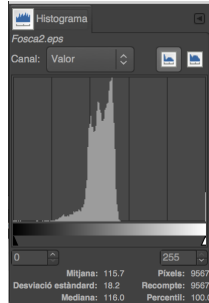
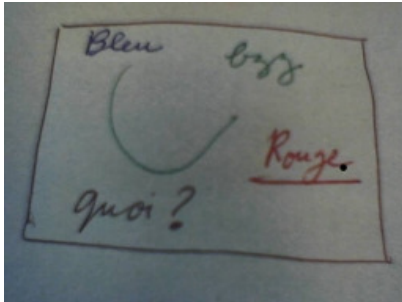
## Exemple d'imatge fosca



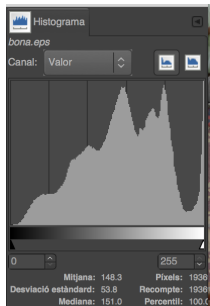
## Exemple d'imatge amb molta llum



## Exemple d'imatge amb baix contrast



## Exemple d'imatge ben contrastada



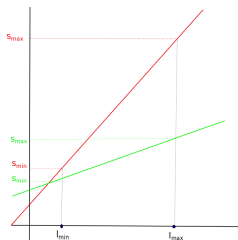
# Transformació de la intensitat

## Com fer-ho

$$T(I) = \frac{s_{max} - s_{min}}{I_{max} - I_{min}}(I - I_{min}) + s_{min}$$

on  $I_{max}$  i  $I_{min}$  són els valors màxims i mínims dels valors d'intensitat en la imatge,  $s_{max}$  and  $s_{min}$  són els valors màxims i mínims que desitjam a l'histograma.

- ▶ si el quocient  $\frac{s_{max} - s_{min}}{I_{max} - I_{min}}$  és més petit que 1 la transformació contreu l'histograma
- ▶ si la proporció és més gran que 1, la transformació estira l'histograma
- ▶ si  $s_{max} = 255$  i  $s_{min} = 0$  la transformació estén, tant com sigui, l'histograma.



$$y(i) = A(I - I_{min}) + s_{min}$$

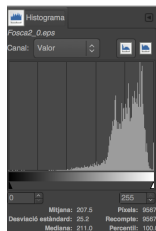
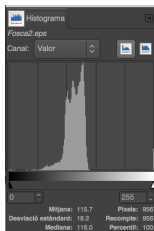
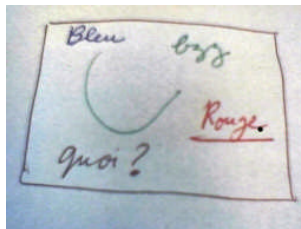
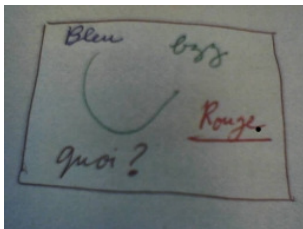
$$A = (s_{max} - s_{min}) / (I_{max} - I_{min})$$

—  $A > 1$ , estira l'histograma

—  $A < 1$ , contreu l'histograma



## Exemple de transformació. Resultat bo



## Exemple de transformació. Resultat insuficient



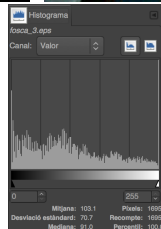
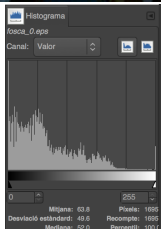
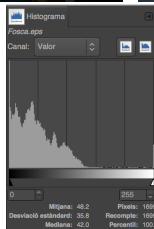
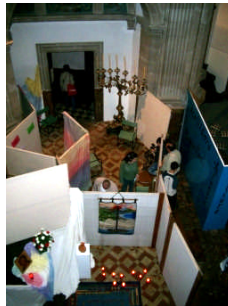
L'histograma de la imatge ocupa tot el rang  $[0, 255]$  degut als llums de les espelmes.

## Millora del resultat

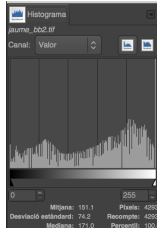
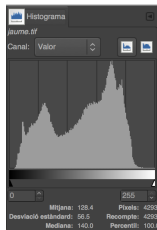
Les passes:

- ▶ saturar un petit percentatge dels píxels amb els valors més alts a 255 i un petit percentatge dels píxels amb els valors més baixos a 0
- ▶ aplicar la transformació lineal

# Exemples



# Exemples



Saturació del 3%

# Outline

Que és una imatge digital?

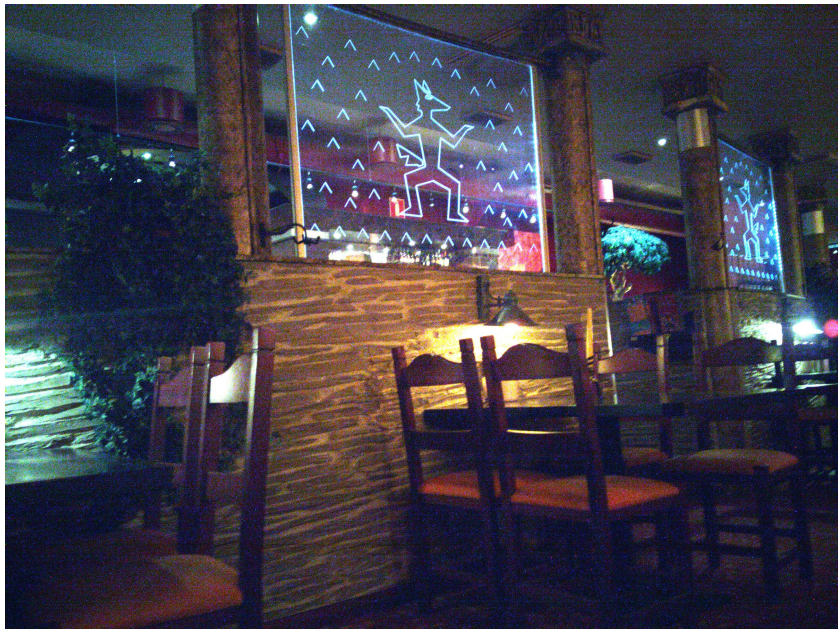
Contrast i color

Com millorar la qualitat d'una imatge

La detecció d'objectes en una imatge digital

La visió estèreo i la reconstrucció 3D

## Aplicació a la millora de la imatge



## Millora de la imatge





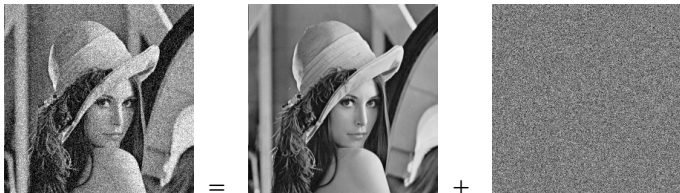
# Que ha passat en aquestes fotografies de mòbil?

Algunes preguntes sobre aquestes fotografies ...

- (i) Que són aquestes pertorbacions que es veuen a la fotografia? Aquest granulat?
- (ii) A que són degudes?
  - ▶ Si no hi ha prou il·luminació de l'escena ...
  - ▶ si el material de l'objecte que fotografiam no reflecteix la llum (absorció), ...
  - ▶ si el temps d'exposició de la càmera és llarg, si ....
  - ▶ Llavors no hi ha prou fotons que es reflecteixin a la matriu de sensors
  - ▶ i per tant no es genera el senyal electric en condicions que dona lloc a la digitalització.
  - ▶ Aquesta "aleatorietat" dels fotons o energia que es crea, fa que el resultat final, la fotografia, estigui pertorbada per aquest fet

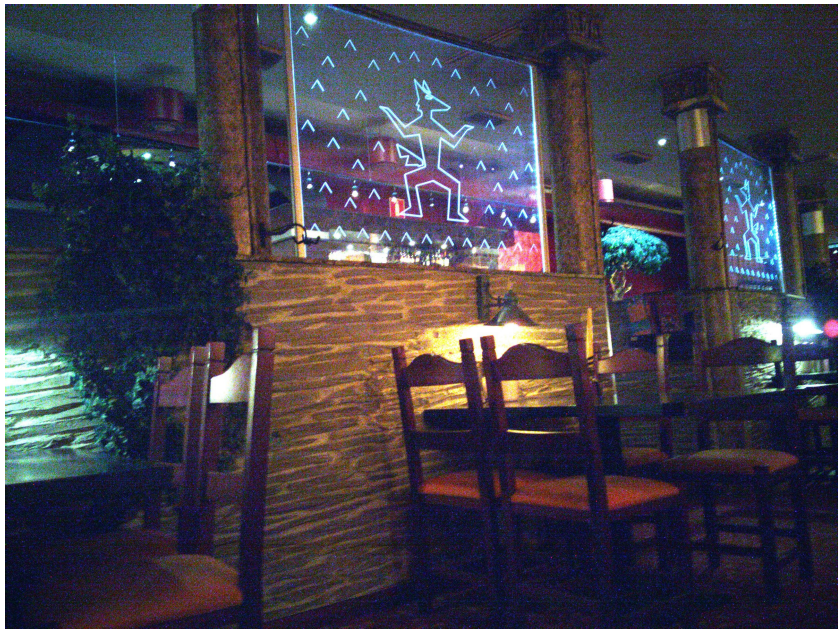
Es defineix el **soroll** associat a la imatge com aquestes pertorbacions aleatòries que s'afegeixen a la imatge original

## Exemple d'un soroll aditiu



**Problema:** Com podem recuperar la fotografia original a partir de la imatge amb soroll?

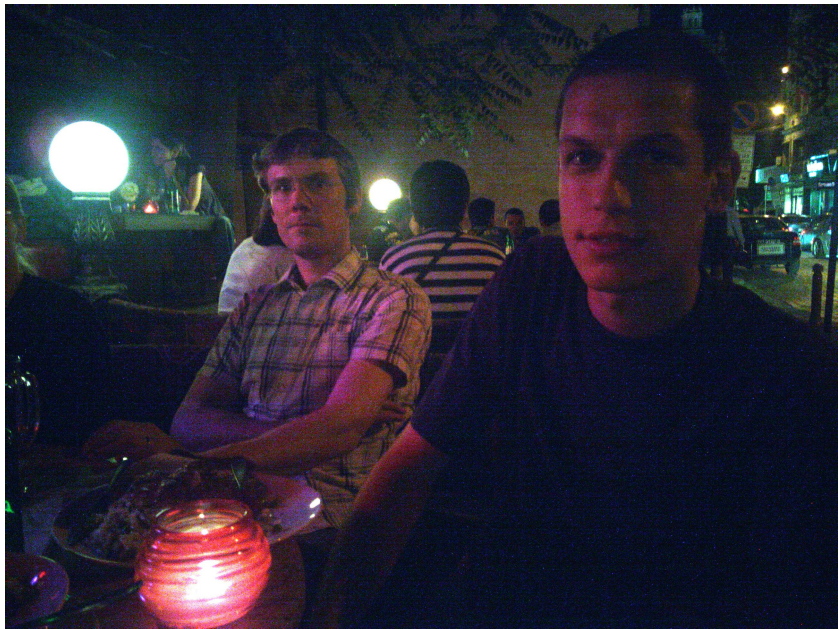
## Exemples



## Exemples



## Exemples



## Exemples





## Com podem eliminar el soroll? Un petit experiment

Agafem una successió de  $N$  nombres aleatoris qualsevols,  $\{a_i\}$ , on  $i$  indica el lloc que ocupa el nombre dins la successió.

Per exemple, agafem  $N = 10$  amb

$$\{1, 5, -2, 0, 6, -1, 3, -2, 4, 7\}, \quad \min = -2, \quad \max = 7$$

Definim el següent algorisme: per a cada nombre de la successió

1. fem la mitjana  $a_i^1 = \frac{a_{i-1} + a_i + a_{i+1}}{3}$ ,  $a_1^1 = \frac{2a_1 + a_2}{3}$  i  $a_N^1 = \frac{a_{N-1} + 2a_N}{3}$

2. primera iteració

$$\{7/3, 4/3, 1, 4/3, 5/3, 8/3, 0, 5/3, 3, 6\},$$

$$\min = 0, \max = 6, \max - \min = 6$$

3. segona iteració

$$\{3, 14/6, 11/6, 2, 17/6, 13/6, 13/6, 14/6, 32/6, 5\},$$

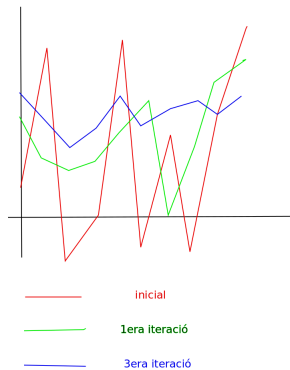
$$\min = 11/6 \approx 1,83, \max = 32/6 \approx 5,33, \max - \min = 3,5$$

4. tercera iteració

$$\{68/18, 43/18, 37/18, 40/18, 42/18, 43/18, 40/18, 59/18, 76/18, 92/18\},$$

$$\min = 37/18 \approx 2,06, \max = 92/18 \approx 5,11, \max - \min = 3,05$$

## comprovació gràfica





# Un petit experiment

## Conclusions:

- ▶ a cada iteració el mínim creix i el màxim decreix
- ▶ quan el nombre d'iteracions creix, el mínim s'acosta cap al màxim
- ▶ en el límit,  
$$\text{mínim} \approx \text{màxim}$$
- ▶ haguessim pogut fer mitjanes entre els 5 més propers en lloc dels 3

## Exemple d'una successió amb un "salt"

Agafem la successió  $\{1, 1, 1, 1, 1, 0, 0, 0, 0, 0\}$

- fem la mitjana  $a_i^1 = \frac{a_{i-1} + a_i + a_{i+1}}{3}$  i obtenim la nova successió

$$\{1, 1, 1, 1, 2/3, 1/3, 0, 0, 0, 0\}$$

- segona iteració

$$\{1, 1, 1, 8/3^2, 6/3^2, 3/3^3, 1/3^2, 0, 0, 0\}$$

- tercera iteració

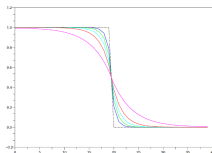
$$\{1, 1, 26/3^3, 23/3^3, 17/3^3, 10/3^3, 4/3^3, 1/3^3, 0, 0\}$$

- cinquena iteració

$$\{242/3^5, 237/3^5, 222/3^5, 192/3^5, 147/3^5, 96/3^5, 51/3^5, 21/3^5, 6/3^5, 1/3^5\}$$

- $n$ -èssima iteració

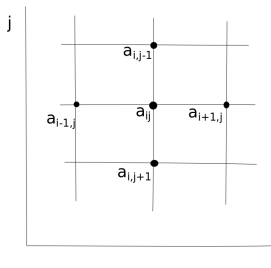
...



## L'experiment en el cas 2D

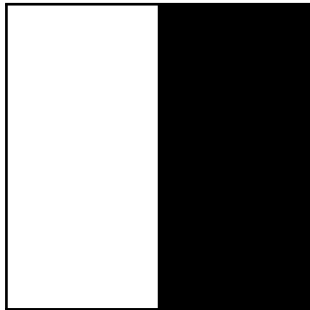
En el cas 2D, una imatge, la matriu de píxels és de la forma  $(a_{i,j})$ .

Es pot implementar el mateix algorisme que l'exposat en el cas 1D



$$a_{ij}^1 = (a_{ij} + a_{i-1,j} + a_{i+1,j} + a_{i,j-1} + a_{i,j+1})/5$$

Un exemple en una imatge sintètica



Un exemple en una imatge sintètica



Un exemple en una imatge sintètica



## Exemple de les mitjanes 2D



## Exemple de les mitjanes 2D





## Exemple de les mitjanes 2D



## Exemple de les mitjanes 2D



## Conclusions en l'aplicació de l'algorisme de mitjanes

- ▶ Hi ha un efecte de "difuminació" o regularització.
- ▶ En el cas de la fotografia (mitjanes 2D):
  - ▶ Quan iterem, la fotografia es torna més difusa: perdrem precisió dels contorns dels objectes, els colors es barrejen, ...
  - ▶ Hem netejat la fotografia, hem eliminat una part de les pertorbacions del soroll però també perdem detalls a cada iteració

**Llavors, com podem limitar la difuminació de la frontera dels objectes?**

# Una modificació de l'algorisme 1D de les mitjanes

**Idea:** en lloc de fer mitjanes, posar una condició sobre els nombres amb els quals es fan les mitjanes

Recordem que en el cas 1D, d'una successió de nombres  $\{a_i\}$  construïem una altra successió en la forma  $a_i^1 = \frac{a_{i-1} + a_i + a_{i+1}}{3}$

**Condició que imposam:** Si  $N_i = \{a_{i-1}, a_i, a_{i+1}\}$ , per al nombre  $a_i$ , només farem mitjanes dels nombres  $x$  de  $N_i$  tals que  $|a_i - x| < L$ , per un cert llindar  $L$ .

D'aquesta manera, **no barrejam valors de píxels que estan veïns** i que donen lloc a aquesta difuminació

## El mateix experiment amb l'algorisme modificat

Agafem  $N = 10$  i la successió

$$\{1, 1, 1, 1, 1, 0, 0, 0, 0, 0\}$$

Agafem la condició que per fer la mitjana

$$|a_i - a_{i-1}| < 0.5 \quad \text{i} \quad |a_i - a_{i+1}| < 0.5$$

Notem que  $a_5 = 1$  i  $a_6 = 0$ .

En aquest cas,  $a_5^1 = (a_4 + a_5)/2 = (1 + 1)/2 = 1$  i

$$a_6^1 = (a_6 + a_7)/2 = (0 + 0)/2 = 0$$

- Després de la primera iteració, obtenim la successió  $\{a_i^1\}$  en la forma

$$\{1, 1, 1, 1, 1, 0, 0, 0, 0, 0\}$$

- Si tornam iterar, ens sortirà el mateix i així per totes les iteracions

**Conclusió: la successió inicial no s'ha modificat!!**

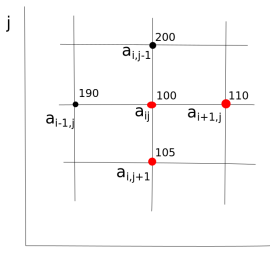
## L'algorisme modificat en el cas 2D

Definim per al píxel  $a_{i,j}$  el conjunt a

$$N_i = \{a_{i,j}, a_{i+1,j}, a_{i-1,j}, a_{i,j+1}, a_{i,j-1}\}$$

Com en el cas 1D, contruïm la iteració a partir de la següent condició:

**Idea bàsica:** per fer les mitjanes, només agafarem els píxels  $x$  de  $N_i$  tals que  $|a_{i,j} - x| < L$ , on  $L$  és un llindar.



$$\begin{aligned} a_{ij}^1 &= (a_{ij} + a_{i-1,j} + a_{i+1,j} + a_{i,j-1} + a_{i,j+1})/5 \\ &= (100 + 110 + 105)/3 \end{aligned}$$

on hem agafat  $L=50$

## Exemple 2D



## Exemple 2D de l'algorithme de les mitjanes modificat





# Outline

Que és una imatge digital?

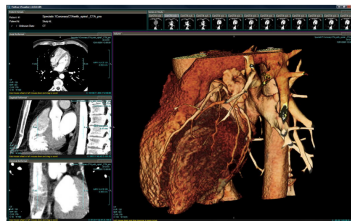
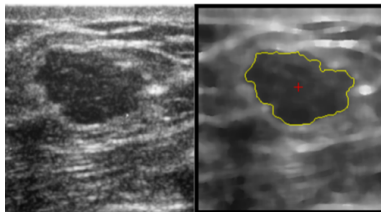
Contrast i color

Com millorar la qualitat d'una imatge

La detecció d'objectes en una imatge digital

La visió estèreo i la reconstrucció 3D

El problema de la segmentació d'estructures 2D i 3D

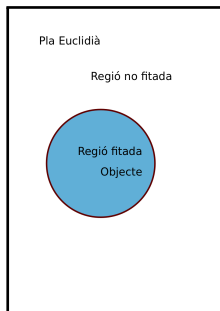


**Problema general: com detectar la frontera d'un objecte?**

# Que és la frontera d'un objecte?

Idealment, una imatge amb un objecte

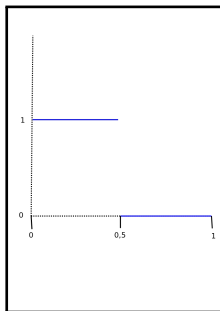
- ▶ L'objecte és una regió de la imatge tancada que té un nivell de gris o color més o menys homogeni
- ▶ L'objecte té una frontera que separa dues regions (interior i exterior)
- ▶ Els punts de la frontera o **contorns** són punts on hi ha un canvi abrupte dels valors de nivell de gris o color (contrast fort).



# Com podem definir un contorn digital

Vegem com podem formalitzar aquesta idea de contorn

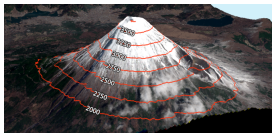
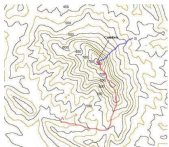
- ▶ Un punt  $x$  és un punt de frontera o contorn, si al seu voltant hi ha un canvi brusc del nivell de grisos
- ▶ en el cas d'un senyal 1D, això voldria dir que la funció senyal té un pendent alt en  $x$  (derivada gran)
- ▶ Inconvenient: els punts on el senyal oscil·la poden tenir derivada gran
- ▶ en un senyal binari, el punt de contorn seria un punt de discontinuïtat



## I en el cas d'una imatge 2D?

Extrapolació del concepte de derivada al cas 2D

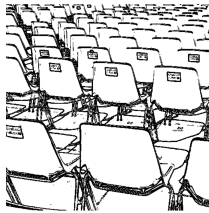
- ▶ una imatge contínua  $I$  és una funció  $(x, y) \rightarrow I(x, y)$
- ▶  $I(x, y)$  denota el nivell de gris al punt/píxel  $(x, y)$
- ▶ podem calcular la variació local de la funció respecte a la variable  $x$  i respecte la variable  $y$  al voltant d'un punt
- ▶ aquest vector ho denotem com a  $\nabla I(x, y) = (I_x, I_y)$ , anomenat gradient de  $I$  al punt  $(x, y)$
- ▶ la direcció del gradient ens dona la direcció del pla al voltant del punt  $(x, y)$  on la funció varia més ràpidament



# Un model per detectar contorns dins la imatge

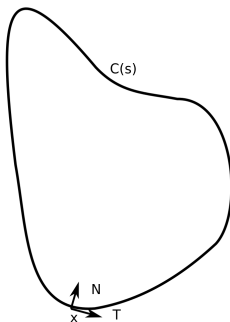
Esquema:

- ▶ aplicar un filtre de mitjanes a la imatge  $I$
- ▶ calcular el gradient de la imatge  $\nabla I(x, y)$  a cada punt
- ▶ calcular el mòdul del gradient  $\|\nabla I(x, y)\|$
- ▶ si  $\|\nabla I(x, y)\| > L$  llavors  $(x, y)$  és un punt de contorn



## Un mètode per detectar la frontera d'un objecte (I)

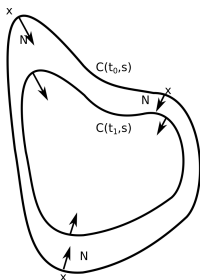
- ▶ suposem que tenim una corba tancada plana (per exemple, moviment d'una partícula al pla)
- ▶ la funció corba vendrà definida per una funció que depen d'un paràmetre  $s$  (per exemple el temps)
- ▶ llavors la corba  $C(s)$  ens donarà per a cada  $s$  la seva posició al pla
- ▶ podem definir el vector velocitat o tangent  $T$  a cada punt de la corba, així com el vector normal  $N$ , perpendicular a  $T$  (normal interior)
- ▶ podem agafar la norma de  $N$  igual a 1,  $\|N\| = 1$ .
- ▶ pensem amb un moviment circular dins la física ( $N$  seria l'acceleració centrípeta)



## Un mètode per detectar la frontera d'un objecte (II)

- ▶ anem a donar-li una dinàmica a aquesta corba, per això introduïm la variable  $t$
- ▶ llavors  $C(t, s)$
- ▶ definim el vector velocitat de la corba respecte a aquesta dinàmica  $t$  com el vector derivada de la corba respecte  $t$ ,  $\frac{\partial C(t, s)}{\partial t}$
- ▶ Notis que la velocitat de la corba és la norma del vector,  $\|\frac{\partial C(t, s)}{\partial t}\|$
- ▶ volem que aquest vector evolucioni a cada punt  $x$  segons la direcció de  $N$ ,

$$\frac{\partial C(t, s)}{\partial t} \approx \vec{N}$$



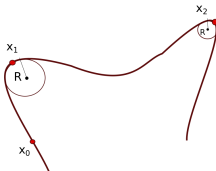


## Un mètode per detectar la frontera d'un objecte (III)

Localment la corba és diferent a cada punt (més oberta, més tancada, ....)

- ▶ donat un punt de la corba  $x$ , aproximam la corba al voltant de  $x$  pel millor cercle que l'aproxima
- ▶ aquest cercle té radi  $R$
- ▶ definim la curvatura  $k$  de  $C$  al punt  $x$  com a  $k(x) = \frac{1}{R}$
- ▶ definim l'evolució de la corba  $C$  com a

$$\frac{\partial C(t, s)}{\partial t} = (k + \alpha) \vec{N}, \quad \text{i la velocitat} \quad \left\| \frac{\partial C(t, s)}{\partial t} \right\| = |k| + \alpha, \quad \alpha > 0 \text{ constant.}$$



curvatura  $\kappa = 1/R$

En  $x_2$  la curvatura és més gran  
que en  $x_1$

En  $x_0$  la curvatura és zero

## Un mètode per detectar la frontera d'un objecte (IV)

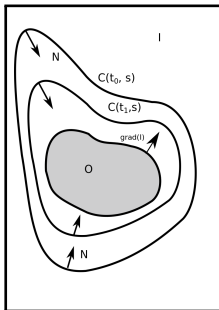
Com podem aplicar aquest model per detectar la frontera d'un objecte?

- ▶  $I(x, y)$  ens dóna el nivell de gris de la imatge al punt  $(x, y)$
- ▶ als punts de la frontera, el vector gradient o "derivada" 2D,  $\nabla I$ , el mòdul  $\|\nabla I\|$  agafa un valor alt
- ▶ hem de posar un terme que quan arribi a la frontera de  $O$ , s'aturi l'evolució
- ▶ si  $g$  és aquest terme, llavors si  $L$  és un llindar,

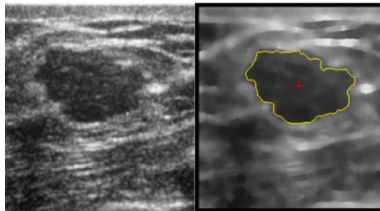
$$g(x) = \begin{cases} 1 & \text{si } \|\nabla I(x)\| < L, \\ 0 & \text{si } \|\nabla I(x)\| > L, \end{cases}$$

- ▶ finalment, l'equació final d'evolució serà

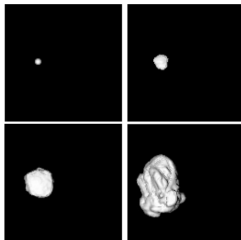
$$\frac{\partial C(t, s)}{\partial t} = g(x)(k(x) + \alpha)\vec{N}$$



## Exemples d'aplicació



## Exemples d'aplicació



# Outline

Que és una imatge digital?

Contrast i color

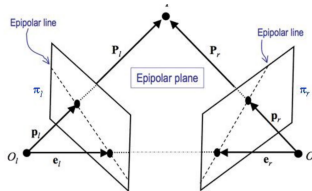
Com millorar la qualitat d'una imatge

La detecció d'objectes en una imatge digital

La visió estèreo i la reconstrucció 3D

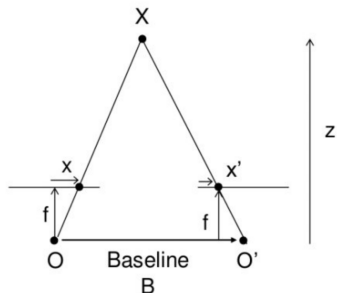
## Visió estèreo: aplicació al càlcul de la profunditat (I)

- ▶ Imatges estèreo: dues imatges per calcular un punt de l'espai 3D
- ▶ Els píxels d'una imatge es comparen amb els píxels de la segona i la geometria epipolar permet trobar aquest punt a l'espai
- ▶ el resultat és un mapa de punts 3D, amb informació de la profunditat



## Visió estèreo: aplicació al càlcul de la profunditat (II)

- ▶ El mapa de la disparitat es refereix a la diferència aparent de píxels o moviment entre un parell d'imatges estèreo.
- ▶  $f$  és la distància focal de la càmera i  $B$  la distància entre les càmeres
- ▶  $\text{disparitat} = x - x' = \frac{Bf}{Z}$
- ▶ llavors la profunditat  $Z$  d'un punt de l'escena és inversament proporcional a la diferència de distància dels punts de les imatges corresponents



## Visió estèreo: aplicació al càlcul de la profunditat (III)





## Visió estèreo: aplicació al càlcul de la profunditat (III)



## Visió estèreo: aplicació al càlcul de la profunditat (III)



i per últim ....



Si escollim una fotografia d'una gran ciutat, podem tenir una aplicació que ens digui de quina ciutat és?

la resposta està en machine learning, deep learning, big data, ....

MOLTES GRÀCIES!!

