



Resum Primer Parcial - 2017

Visió per Computador (Universitat Autònoma de Barcelona)

Tema 1: Formació de la imatge

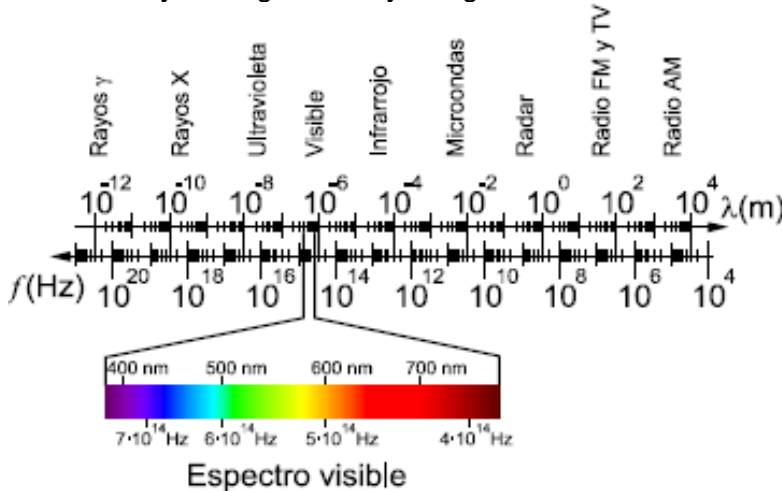
Espectre electromagnètic: conjunt de totes les ones (menor freqüència = ones radio i major freqüència=ones gamma). Només en percebem una petita porció.

Espectre visible: entre 400 nm i 800 nm

Longitud d'ona: distància entre 2 màxims

Més energia = més freqüència

Menys energia = menys longitud d'ona



$$E = h \nu$$

$$c = \lambda \nu$$

relació de Planck

h : constant de Planck

c : velocitat de la llum en el buit

Intensitat lluminosa: quantitat de flux lluminós que incideix sobre una superfície. La intensitat emesa o reflectida per un objecte per totes les longituds d'ona defineix el seu color.

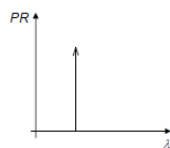
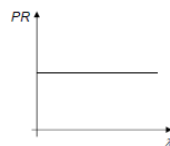
Color: sensació visual produïda per una determinada distribució espectral que incideix en la retina i es processa pel SVH.

-**Objecte difús (superfície no polida)**: No importa on està l'espectador, molta dispersió. MATE

-**Objecte especular**: Tota la llum es reflecteix en una direcció. BRILLANT

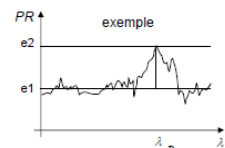
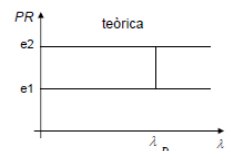
Llum

- Llum blanca: espectre pla (equienergètic). Per obtenir una sensació de llum blanca no és necessari que l'espectre sigui pla.
- Llum monocromàtica: espectre concentrat en una única longitud d'ona. Representa un color espectral pur.

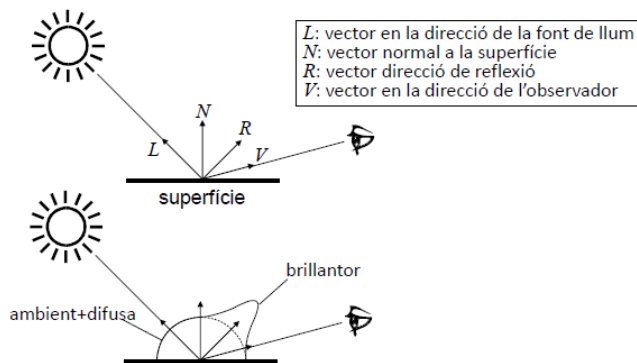


Paràmetres espectrals / color

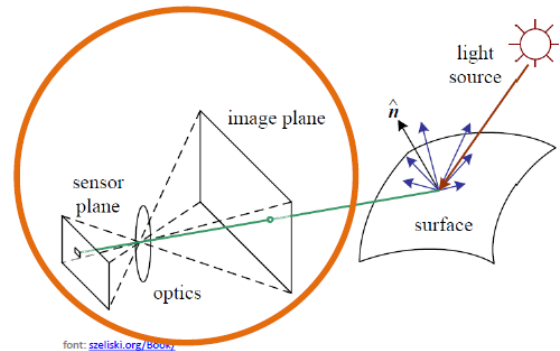
- **Longitud d'ona dominant / λ_D**
longitud d'ona amb més energia. Pic del espectre (e_2).
- **Puresa / saturació**
relació entre l'energia de la longitud d'ona dominant (e_2) i l'energia de la llum blanca (e_1). Si $e_1=e_2$ la puresa és zero i si e_1 és zero la puresa és 1 (màxima).
- **Luminància / Luminositat**
proporcional a l'àrea sota la corba de la distribució espectral.



Geometria



Elements implicats



Led: Campana de gauss d'un cert marge al voltant d'un color

Laser: llum monocromàtica

Ombrejat de Phong: Utilitzat en modelació 3D

Llum ambient: no tota la llum ve de la font de llum

Generació 3D: Possibilitat de crear models 3D a partir de 2 fotografies planes.

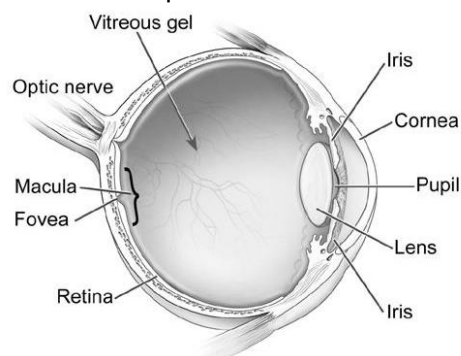
SVH – Sistema Visual Humà

-Estructura

-La **còrnia**, el cristal·lí i l'humor aquós s'encarreguen de projectar una imatge sobre la retina. Lents/òptiques

-L'**iris** és un diafragma que regula la quantitat de llum que entra en l'ull. Diafragma

-La **retina** és una capa de cèl·lules nervioses que transforma l'energia lluminosa en impulsos elèctrics. Sensor



-Receptors:

-**Bastons:** Saturats durant el dia, serveixen per veure durant la nit (intensitat baixa). Millor resolució al voltant de la fòbia, no al centre.

-**Cons:** S'utilitzen per veure durant el dia (sensibles a nivells alts). Tenim diferents que poden veure al violeta, verd i groc. Veiem bé a la fòbia, centre de la retina

Mesures de color

Qualsevol color es pot posar en base de x , y i z

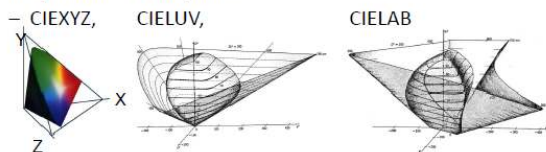
- Els tres primaris són X , Y , Z i són imaginaris (no es poden donar a la natura).
- Locus: Colors de l'arc iris, colors molt saturats del límit del sòlid de color

Espais de color

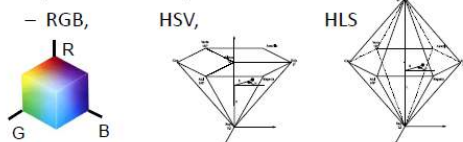
- Independents de dispositiu:
 - CIEXYZ: Útil per mesurar
 - CIELUV: Uniforme. Utilitzat per emissors
 - CIELAB: Uniforme. Utilitzat per pintures i colors reflectits.
- Dependents de dispositiu
 - RGB: monitors, càmeres, imatges. Molt utilitzat.
 - CMY/CMYK: impressió
 - Y'U'V' i Y'I'Q': Televisió, vídeo
 - HSV: Imiten espais perceptuals

Espais de color

- Independents de dispositiu

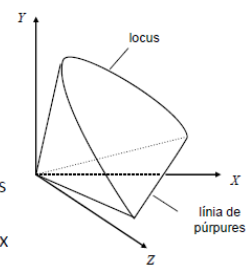


- Dependents de dispositiu



Espai XYZ

- Els valors XYZ dels possibles colors defineixen una mena de con "sòlid de color". Tots els possibles colors són dins d'ell.
- Els eixos de l'espai XYZ no cauen a dins del sòlid (colors imaginaris).
- L'origen es correspon amb el negre
- La frontera representa els colors espectrals purs (locus), els púrpures que no hi són a l'espectre estan representats per una línia que uneix vermell i violeta.

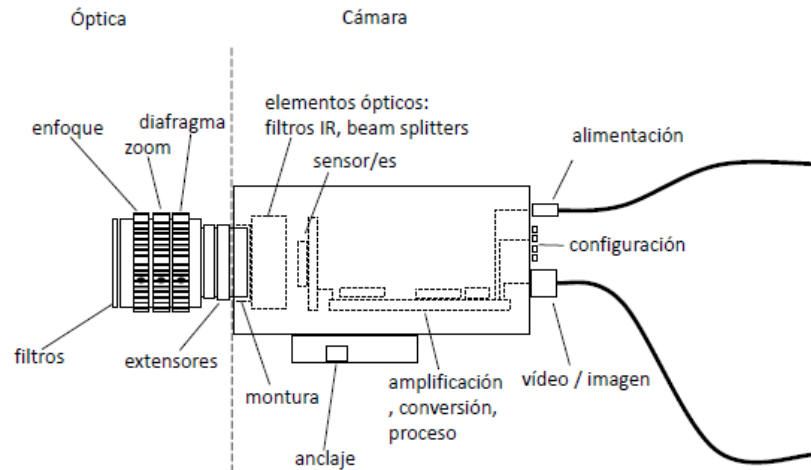


Càmeres

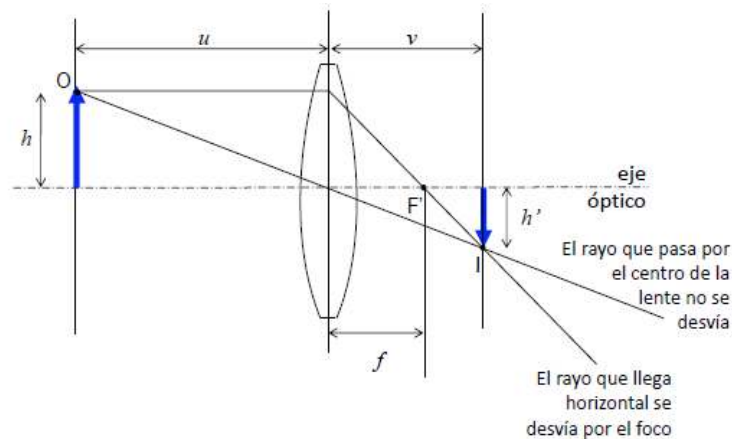
Pinhole: És un model simple de càmera, quan amb un forat es projecta una imatge inversa en una paret

Calibratge:

- Intrínsecs: Focal, punt principal, distorsions. Ens pot servir per generar 3D
- Extrínsecs: Rotació, translació



Òptica



Objectiu

Lent simple amb una determinada focal

Distància focal:

Els objectius estan formats per grups òptics en els que la suposició de la lent prima no es dona.

El terme "distància focal" fa referència a la **EFL (distància focal efectiva)**

-Altres mesures útils són la BFD (back focal distance) i la FFD (Flange focal distance)

La distància focal està relacionada amb l'angle que engloba la imatge

-Menor distància focal, major camp angular

Obertura

Relativa: Relacionat amb la lluminositat que tindrà la imatge

-**Diametre de l'objectiu:** Llum que entra

-**Distància focal:** Mida de la imatge

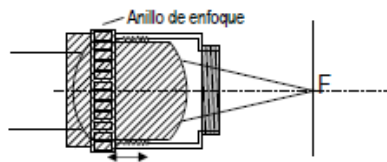
Enfoc

-Els objectius ajusten la distància lent-sensor amb l'enfoc, aconseguint una imatge nítida sobre el sensor

-El posicionament amb rosca de pas fi (5mm)

-Quan la lent està el més aprop possible del sensor, aquest es situa en el pla focal del sistema (objectes distants enfocats)

-Amb l'anell d'enfoc al màxim tenim la mínima distància entre objecte i objectiu. Es calcula amb la fórmula de Gauss



Ejemplo: para un objetivo de $f=25\text{ mm}$ con 5 mm de recorrido de enfoque tenemos un rango

$$\frac{1}{25\text{ mm}} = \frac{1}{u} + \frac{1}{25\text{ mm} + 5\text{ mm}}; \quad u = \text{MOD} = 150\text{ mm}$$

Nota: en aplicaciones con condiciones adversas conviene fijar el enfoque con un tornillo.

Camp de vista

FOV (Field of View): àrea de l'escena, a la distància de treball u , que forma la seva imatge en el sensor. Depèn de tres factors:

Montura

Distància entre el pla posterior de l'òptica (sense contar la rosca) u el sensor (plà focal). Els sensors de gran format fan servir altres montures (42 mm) i les de baix cost (12 mm) o menors.

Profunditat de camp

Cercle de confusió: l'ull humà o una ltre sensor és incapaç de distingir detalls més petits que l'element utilitzat per captar-ho. Depenent del sensor que utilitzem tindrem un o altre creclre de confusió. Amb les càmeres està entre 10 microm i 30 microm

- Tipo C (indústria i ciència)
- Tipo CS (vigilància)

Sensor

- Los sensores convierten la energía radiante en una señal eléctrica.
- La luz está "formada" por fotones. El sensor "cuenta" fotones.



QE (Quantum Efficiency):

Número de electrones generados por cada fotón incidente.

WC (Well Capacity):

Número de electrones que puede almacenar la celda sensora.

Transferencia

Problemas asociados al proceso de transferencia:

- **Blooming:** desbordamiento de carga.
- **Smear:** La luz llega en la etapa de lectura. La carga se propaga a lo largo de toda la columna.

Possible soluciones:

Tiempo de integración menor \Rightarrow blooming menor
Tapar el sensor en el tiempo de lectura.

Color

Bayer pattern (1 único sensor)

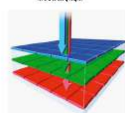
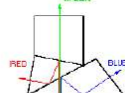
- El sistema más utilizado
- No tenemos R,G,B en ningún pixel. Se necesita interpolar (demosaiing)

3 sensores independientes

- Prismas y filtros dividen el haz de luz en 3 componentes que iluminan 3 sensores monocromáticos.

Píxeles apilados

- La luz dependiendo de su energía (color) llega a diferentes profundidades.



Cámaras Analógicas

- Ventajas
 - Longitud del cable prácticamente ilimitada ($\sim 100\text{ metros}$)
 - Bajo precio de las cámaras
- Inconvenientes
 - Típicamente, menor resolución
 - Requieren del uso de un *frame grabber*
 - Los problemas de sincronismo cuando se usan múltiples cámaras requieren cables específicos.
- Barrido
 - Entrelazado: proviene de la TV. CCIR (25 Hz), RS-170/EIA (30 Hz)
 - Progresivo: no estándar

Cámaras Digitales

- Al no derivar de un estándar de televisión, se pueden encontrar modelos de mayor resolución, mayor velocidad, ventana variable.
- Varios sistemas:
 - Cámaras Firewire (IEEE1394a, IEEE1394b)
 - Cámaras CameraLink
 - Otros: USB 2.0, USB 3.0, GigaEthernet , Cameralink, ...

Tema 2 – Processat d'imatges

Una imatge és una funció bidimensional que a cada punt es fa correspondre un valor d'intensitat lumínica. Una imatge digital és una estructura de dades (píxels) que conté informació discreta.

- Mostreig: Discretització en l'espai. Generat en captar una senyal continua del món real amb un sensor discret
- Quantització: Discretització en amplitud. Emmagatzemar una senyal en format digital.

Classes d'imatges:

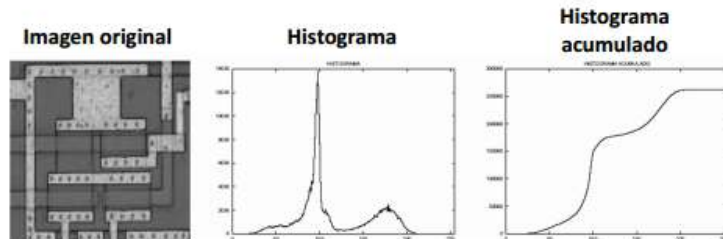
- Binària: Blanc i negre.
- Monocromàtica: Únic component espectral, matisat per diferents components d'intensitat.
- Color: Tres components cromàtiques: RGB
- Vídeo: Seqüència d'imatges en un instant de temps
- 3D: Imatges apilades en z
- Estéreo: Dos imatges de la mateixa escena en el mateix instant des de dos punts de vista diferents.

Matlab: Matriu 2D amb valors uint8 (enters de 0 a 255) o double (valors de 0 a 1)

```
Im = imread ();      %Llegir una imatge
imshow();            %Mostrar una imatge
```

Histograma: Vector de longitud igual al número de nivells d'intensitat. Cada posició reflexa el número de píxels de la imatge amb cada intensitat (densitat de probabilitat).

```
imhist(im) %Histograma
cumsum(imhist(im)) %Histograma acumulat
```



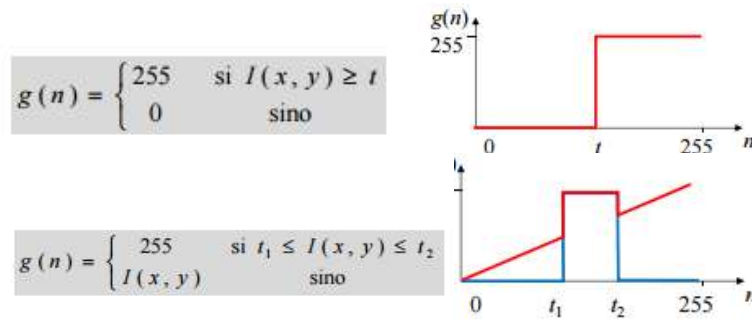
Transformacions puntuals simples

- Umbral simple i doble
- LUT
- Transformacions lineals
 - Imatge negativa
 - Maximització del contrast
- Transformacions no lineals
 - Logaritme
 - Exponencial
- Operacions lògiques

Transformacions puntuals simples

Treballen amb píxels com elements individuals.

- Umbralització (threshold, binarització)
 - Umbral simple: Els punts amb un nivell de intensitat major o igual que un cert Umbral t reben el màxim i la resta el mínim.
 - Umbral doble



LUT (look-up table)

Es fa servir per transformar les dades entrants en un format de sortida més desitjable mitjançant un vector d'indexació.

-Més lent si la substitució és simple.

Transformacions lineals

Transformació d'una imatge en una nova amb diferent rang de valors, a partir d'una funció lineal.

-Imatge negativa: S'aplica un contrast invers

$$g(n) = 255 - I(x, y)$$

-Maximització del contrast: Aplicar transformació lineal que fa que la imatge ocupi tot el rang possible. "S'estira l'histograma"

$$g(n) = 255 \left[\frac{I(x, y) - \min(I(x, y))}{\max(I(x, y)) - \min(I(x, y))} \right]$$

Transformacions no lineals

Les transformacions lineals modifiquen el rang de la imatge. En canvi les no lineals modifiquen el contrast d'una imatge dins del rang d'aquesta.

-Logarítmiques: millora píxels obscurs sense saturar els brillants. Serveix per visualitzar imatges amb rang dinàmic ampli (transformada de Fourier)

$$g(I(x, y)) = \log(I(x, y) + 1)$$

-Exponencials: millora dels brillants sense saturar els obscurs

$$g(I(x, y)) = e^{I(x, y)}$$

Transformacions aritmètiques

-Suma – Adició: El soroll es va reduint

-Resta – Substracció: Serveix per diferenciar coses, observar els canvis entre dos imatges (background)

-Producte: Regions d'interès. Posem 1 al que volem deixar passar i 0 per no deixar passar en una imatge. Reducció d'efectes de cantonades.

La convolució en l'espai equival al producte (Fourier)

-Divisió: Serveix per eliminar un factor en una imatge (per exemple llum transmesa)

Pseudocolor i fals color

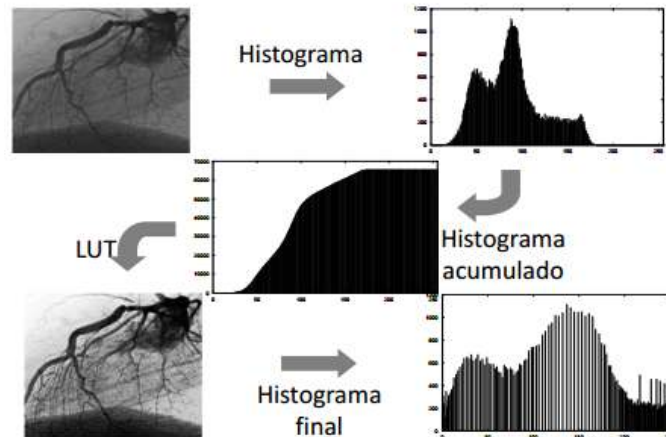
Visualitzar en color una imatge de tons grisos. Assignant a cada nivell de gris un color

Fals color: modificar els canals de color per ressaltar algun aspecte de la imatge.

Equalització

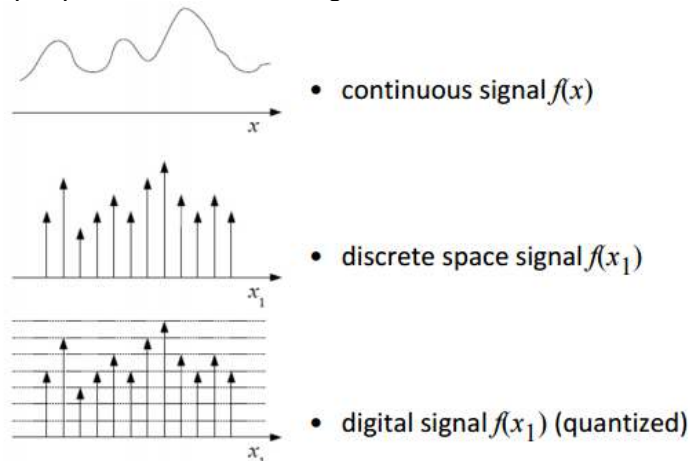
Busca que tots els nivells de gris tinguin la mateixa freqüència d'aparició (histograma pla). Pot aproximar-se per una LUT que consisteix en un histograma acumulat portat al rang de la imatge d'entrada.

- Equalització local: per cada píxel es calcula l'histograma en un veïnat
- Equalització logarítmica: s'eleva l'histograma acumulat a una constant



Tema 3 – Filtrat lineal

Diferents senyals que provenen d'una imatge. El seu domini està dintre dels enters



Tractament de senyal: Representar una funció com una suma d'uns coeficients amb pesos. Volem explicar el filtratge

Sistema: funció que donada una imatge i una funció i obtenim una nova imatge.

Sistemes lineals:

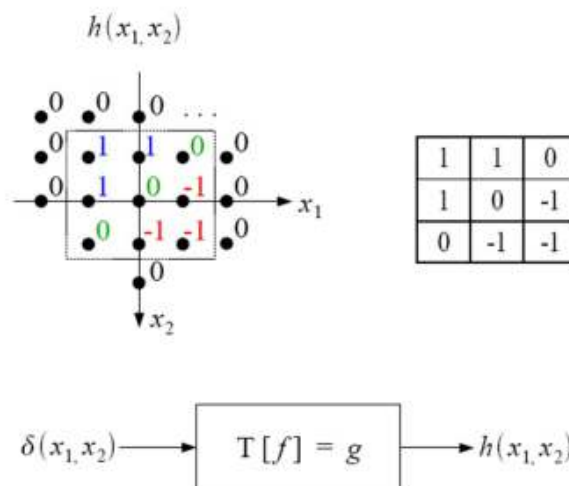
- Simples
- Representen transformacions interessants
- Modelin coses que passen a la realitat

Convulsió: Filtratge que opera una màscara per tota la imatge

- Modela mitjançant una màscara (h)

A les cantonades utilitzem imatges simètriques.

Màscara de convulsió ($h(x_1, x_2)$)



Exemple:

$$Im = [1, 2, 3, 1]$$

$$h = [-1, 1]$$

$$h(-k) = [1, -1]$$

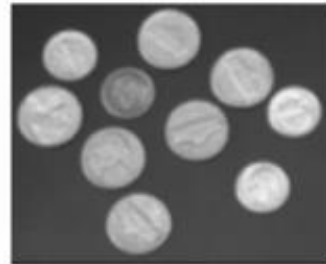
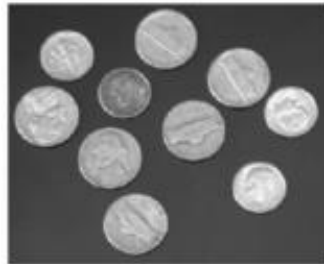
$$im * h = [-1, -1, -1, 2, 1] \text{ o } [-1, -1, 2] \quad \text{S'utilitza la } h(-k)$$

Correlació

Són iguals amb la convolució si la màscara és simètrica

Filtres lineals

- **Smoothing:** Redueix la intensitat de variació local, com un desenfoc, i normalment porta soroll



$M=2, N=2$

- **Detectors de cantonades (Edge detectors)**

Aplicar derivades en diferents direccions per detectar límits (contorns: canvis bruscs de la derivada en una direcció).

- Vector gradient: On calculem les derivades
- Magnitud: Calcular el mòdul del gradient
- Orientació: Podem obtenir la fase, direcció del salt que em trobat



$f * D_x^s$

$f * D_y^s$

-Roberts: Aproximar el gradient (pendent) d'una imatge

- Vores ben definides
- Fons sense soroll
- Intensitat de les vores = ull humà

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

-Sobel: El resultat és un vector gradient i la norma d'aquest. Calcula el gradient de la intensitat d'una imatge.

Derivada en X:

$$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

Derivada en Y:

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

-Prewitt: Calcula el gradient de la intensitat d'una imatge. Donant la direcció de l'increment més llarg de clar a obscur.

$$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

-Laplacian of Gaussian: No està basat en càlculs de màxims del vector gradient.

-Canny: És el més complex però el que més s'utilitza. Utilitza un algoritme de múltiples capes. Màscara de 5x5.

- **Pattern Matching**

Buscar patrons o estructures en una imatge. S'utilitza la correlació (Normalized Cross-Correlation), on es va normalitzant punt a punt.

- **Transformada de Fourier**

Transformació que fa convolucions molt ràpidament. Canvia una imatge de domini espacial (x,y) a una de domini freqüencial (u,v).

Propietats:

- Simètrica, periodicitat, escalat
- Rotació: Si rotem la imatge la transformada també rota

- **Derivades Gaussians**

Les derivades són sensibles al soroll. Si apliquem smoothing amb Gaussian minimitzen el soroll

- **Laplacian of Gaussian**

És la segona derivada espacial d'una imatge. La laplaciana ressaltava els canvis d'intensitat:

- S'utilitza per la detecció de vores
- Normalment primer es suavitzen les imatges

Filtres de Gabor: Filtre lineal, funcions passa banda. Estan localitzades tant en domini espacial com freqüencial. Es poden sintonitzar en freqüència i en orientació.

- Utilitzades per escorça visual

Watershed: Estudia les imatges millor que Fourier (no assumeix que siguin repetitius, al contrari que Fourier). Representació freqüencial d'una senyal (representa freqüència i posició).

- Separa zones d'influència (conques)
- En un mapa topogràfic les fronteres es correspondrien amb les crestes de les muntanyes
- Té sobresegmentació, doncs hi ha molts mínims locals (zones d'influència)

Tema 4 – Filtratge no lineal

Median filters: Un filtre de mediana s'implementa desplaçant una màscara per tota la imatge i en els veïns ordenem els valors i ens quedem amb el del mig. Elimina molt malament el soroll impulsional.

Homomorphic filters: Sabent que la il·luminació varia lentament, estarà en la senyal de baixa freqüència. Aplicant Fourier deixarem passar mitja i alta freqüència. Aconsegüim més contrast a les imatges.

Mathematical Morphology: Conceptualment operacions simples que combinades defineixen operacions més complexes. Solen ser operacions geomètriques.

- Binary Morphology
 - Erosion
 - Dilation
 - Open
 - Close
- Grey-level Morphology
 - Umbra
 - Watershed

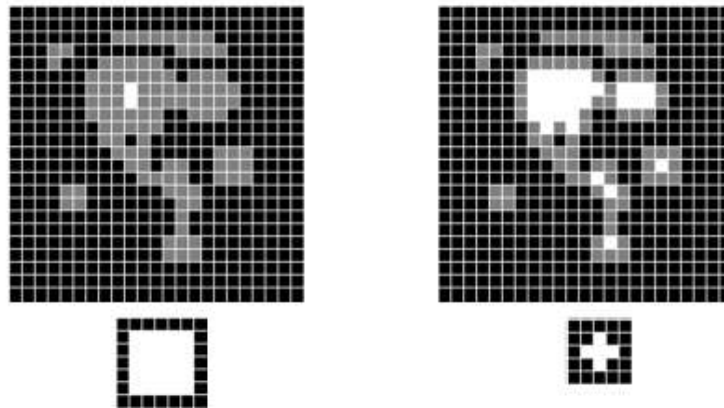
Morfologia Binària

- **Erosió**

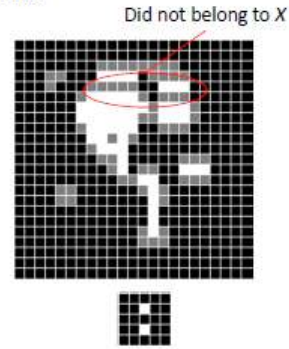
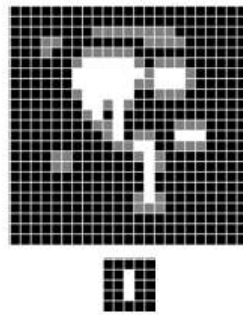
Passem la màscara i només queden els píxels centrats que contenen la màscara, si està completament dins la màscara. El resultat és la subtracció d'elements.

Resta de Minkowski: Intersecció

Erosion with other structuring elements



Erosion with other structuring elements

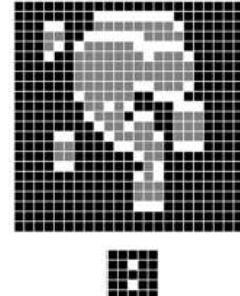
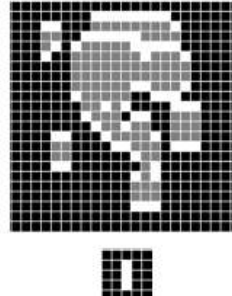
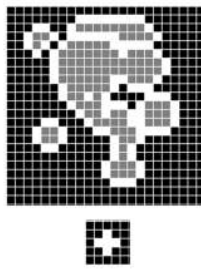
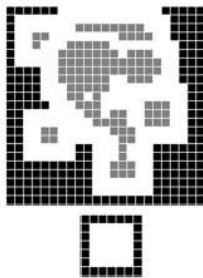


- **Dilatació**

A cada píxel se li passa la màscara i el resultat serà cada píxel més la màscara.
La dilatació no és la suma de Minkowski.

Diferència entre dilatació i suma de minkowski: que l'element estructurant ha d'estar girat.

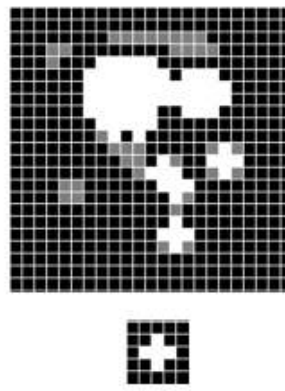
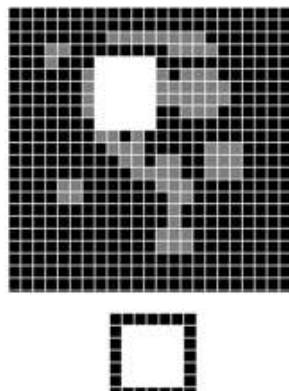
Dilation with other structuring elements



- **Open**

El resultat és tot píxel que pertanyi en algun moment a la màscara i la màscara d'aquest. Primer erosionem (treiem imatge) i després dilatem (afegim imatge).
Suprimeix: els petites illes, petites unions

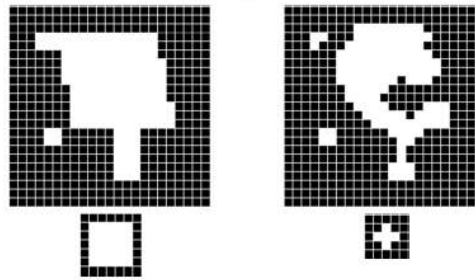
Opening with other structuring elements



- **Close**

Suprimeix: els petits forats, petites separacions

Closing with other structuring elements

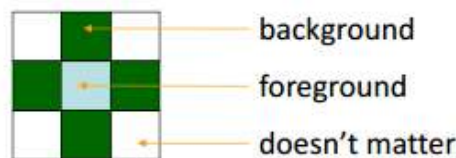


Dilation i closing són extensives
Erosion i Opening són anti-extensives

Si apliquem dos cops opening o closing són idempotents

Hit-or-miss: Detecta una determinada configuració (patró) en una imatge binària. Tenint dos elements estructurants diferents apliquem primer un a la part blanca, després un a la part negra i després fem la intersecció.

-S'aplica erode i elements estructurants disjunts



Thinning, Thickening:

Useful transforms (S'utilitza el hit-or-miss)

-**Contour:** Mascarà que només deixa el contorn de les figures

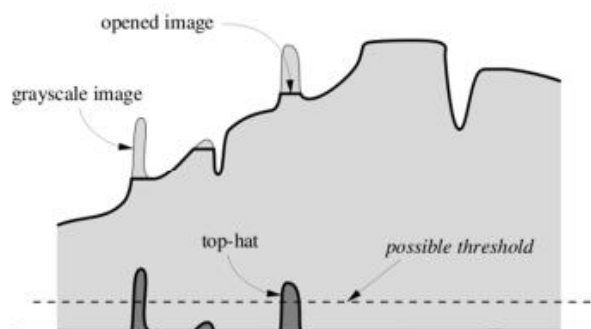
-**Convex-hull:** unió de engruiximents (thickenings), cadascun fins a idempotència. Serveix per simplificar un objecte ramificat

-**Skeleton:** Erosió d'un objecte. Unió de centres del màxim de discs. Són sorollosos però serveix per observar la forma.

Morfologia de nivells de gris

-**Umbra:** Part obscura d'una ombra, on la font de llum està completament bloquejada per un cos opac.

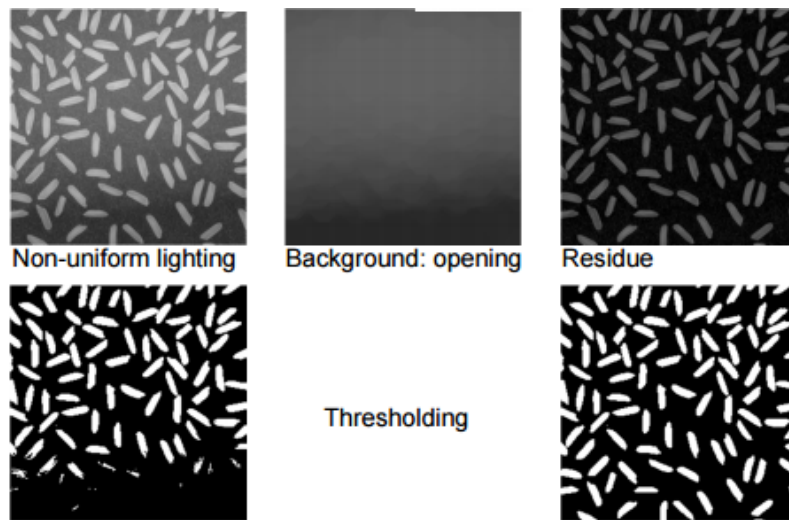
-**Opening Residue:** Erosió + Dilatació



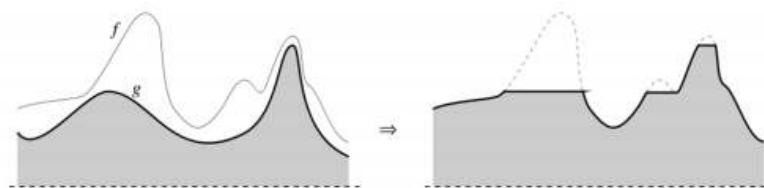
$$f - (f \circ B) = f - ((f \ominus B) \oplus B)$$

Keeps narrow peaks (bright points and bright thin lines).

Amb threshold no podem extreure bé doncs la il·luminació és en degradat



Reconstrucció: Dilatem. També serveix per treure soroll de fons



Reconstruction of g conditioned to $f, f > g$: successively dilate g , and at each step condition (intersect, min) to f , until g doesn't change.

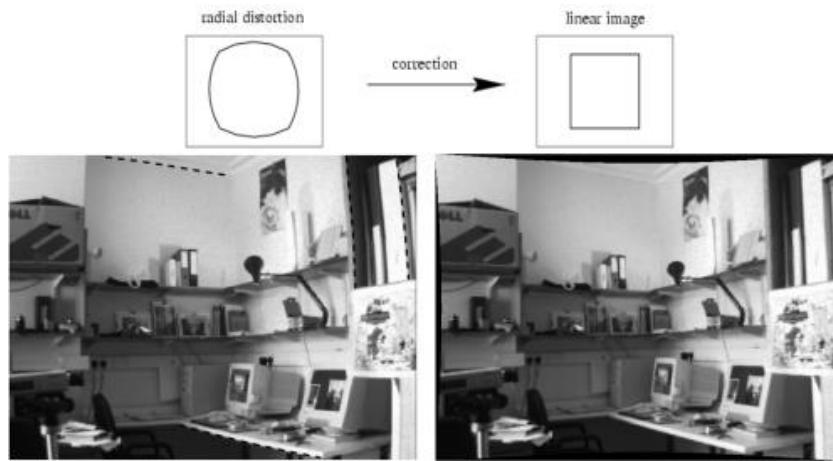
We'll use it later to remove local minima other than those desired.

Watershed

Es busquen les zones d'influència, o les crestes de les muntanyes, que separen les diferents muntanyes (imatges topogràfiques). Són interessants per la segmentació. Utilitza algoritme "flood filling" que consisteix en omplir amb aigua les parts de mínims.

Tema 5 – Transformacions Geomètriques

Les transformacions geomètriques canvia les coordenades, no el valor dels píxels. Serveixen per corregir distorsions causades pels dispositius de captura o per comparar imatges.



- **Càlcul de la cartografia:** Trobar l'equació que relaciona els sistemes de coordenades de l'original i les imatges resultants.
- **El remostreig o interpolació:** Assignar valors als píxels de la imatge resultant dels de l'entrada imatge

Mapping Computation: Buscar coeficients dels polinomis perquè a partir de les coordenades i, j trobar la correspondència.

- Punts de control semi-automàtics entre imatge d'entrada i sortida (similitud amb correlació, covariància, ...)
- Automàtic: Utilitza RANSAC per correspondre característiques locals

RANSAC (Random Sample Consensus): És un mètode iteratiu que estima parametres d'un model matemàtic que conté dades observades .

- Selecciona a l'atzar el mínim nombre de punts requerits i determina el model
- Determina quants punts del set concorden amb una certa tolerància
- Si inliers supera un llindar tornar a re-estimar el model
- Repeteix els passos anteriors un màxim establert

No assegura un resultat estable, doncs és aleatori. El temps pot variar si repetim el procés diferents vegades.

Resampling (interpolació)

Aconseguir mides majors a la imatge original reomplint els píxels minvants amb algorismes específics:

- Per aproximació: promig de valors de 2 píxels pròxims
- Bicúbic: Fer el promig de 16 píxels veïns
- S-Spline: Determina el color basantse en la totalitat de colors

Aplicacions:

- Alineació geomètrica per comparar o fusionar píxels
- Imatges mèdiques i sensors remots
- Detector de diferències temporals o combinar diferents sensors

Homografia

Transformació projectiva que determina la correspondència entre figures geomètriques. S'utilitza DLT.