

8. Visió artificial

Models d'intel·ligència artificial



Introducció

- La visió artificial és una de les àrees més antigues de la intel·ligència artificial.
- Els primers sistemes de visió artificial van ser desenvolupats a la dècada dels 60.
- Els sistemes de visió artificial són capaços d'analitzar imatges i vídeos per tal d'extreure'n informació.
- Veurem quins són els conceptes bàsics de la visió artificial i com s'apliquen en la pràctica.

Visió

- Procés de **percepció**, on el sistema visual és capaç de construir una representació (*imatge*) a partir de la informació captada per la retina.
- Aquest procés pot ser **actiu** (quan l'observador mou els ulls) o **passiu** (quan l'observador no mou els ulls).
- La visió artificial pura és un procés **passiu**, molts conceptes, però, com la **localització** o la **reconstrucció 3D** requereixen un procés **actiu**.

Enfocaments

- Hi ha dos enfocaments principals per a la visió artificial:
 - **Extracció de característiques:**
 - S'apliquen una serie de **transformacions** a la imatge per tal d'extreure característiques rellevants (*vores, textura, fluix òptic, segments*, entre d'altres).
 - **Basat en models:**
 - S'utilitzen models matemàtics (*geomètrics o estadístics*) per tal de representar la imatge.
- En la pràctica, sovint es combinen ambdós enfocaments.

El color (I)

- Propietat de la llum que depèn de la seva longitud d'ona.
- Els humans el percebem el a partir d'unes cèl·lules receptors de la retina: els **cons**.
 - Hi ha tres tipus de cons:
 - **L** (longitud d'ona llarga)
 - **M** (longitud d'ona mitjana)
 - **S** (longitud d'ona curta)
 - Cada tipus de cons és sensible a un rang de longituds d'ona i, per tant, a un rang de colors.

El color (II)

- **Principi de tricromia:** qualsevol color es pot representar com una combinació de tres colors primaris.
- **Colors primaris**
 - Aquells que no es poden descompondre en altres colors.
 - **blau, verd i vermell.**
- Espais de color: RGB, HSV, YUV, ...
- El més utilitzat en visió artificial és el **RGB (Red, Green, Blue).**

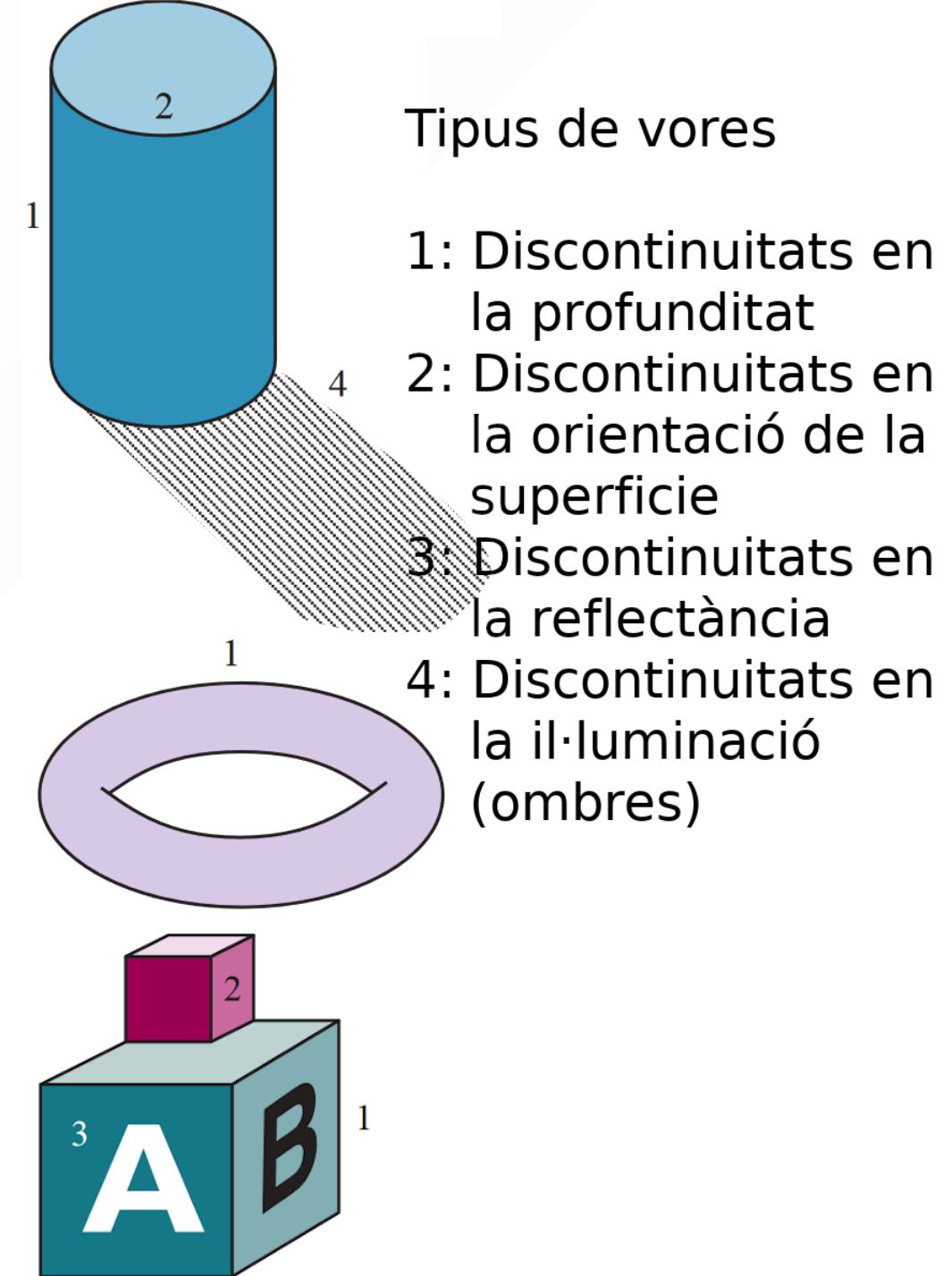
Característiques de les imatges

Definició

- En una imatge hi ha molta informació que no és rellevant.
- Per les tasques de visió artificial es solen utilitzar **característiques** de les imatges.
- Les característiques són aquelles parts de la imatge que són rellevants per a la tasca que es vol realitzar.
- Ens centrem en quatre característiques de les imatges quasi sempre rellevants.
 - *Vores, textura, fluix òptic i segmentació.*

Vores

- **Línees que separen regions de diferent intensitat.**
- Permeten **identificar objectes**.
- Simplifiquen la imatge i permeten **reduir la quantitat d'informació**.
- Passem d'una imatge molt gran a una matriu de vores: **matriu de vores**.



Tipus de vores

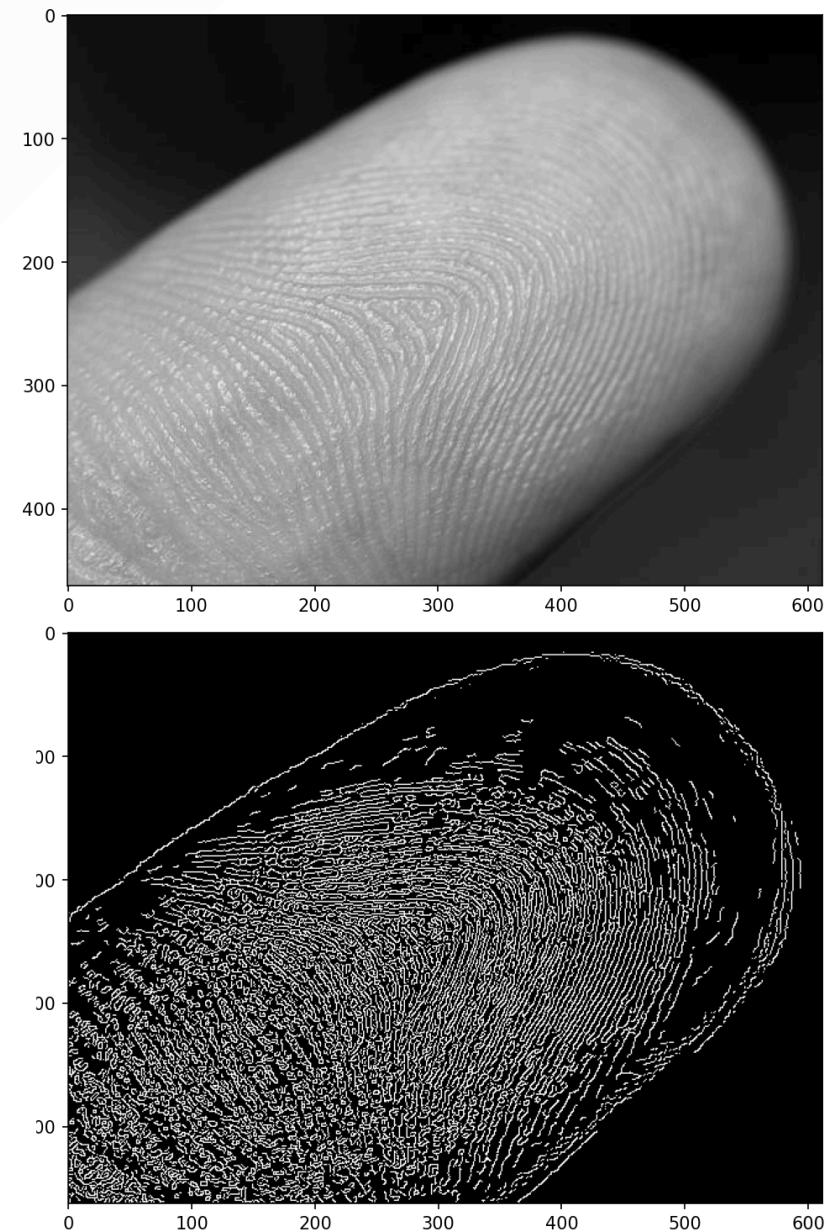
- 1: Discontinuitats en la profunditat
- 2: Discontinuitats en la orientació de la superficie
- 3: Discontinuitats en la reflectància
- 4: Discontinuitats en la il·luminació (ombres)

Detecció de vores

- Tasca de visió artificial que consisteix en detectar les vores d'una imatge.
- Hi ha molts algoritmes per detectar vores, però el més utilitzat és l'algoritme de **Canny**, per John F. Canny, que el va publicar el 1986.
- Objectius:
 - **Bona detecció**: detectar totes les vores.
 - **Bona localització**: les vores han de ser el més pròximes possible a les vores reals.
 - **Minimitzar les respostes falses**: no detectar vores on no n'hi ha.

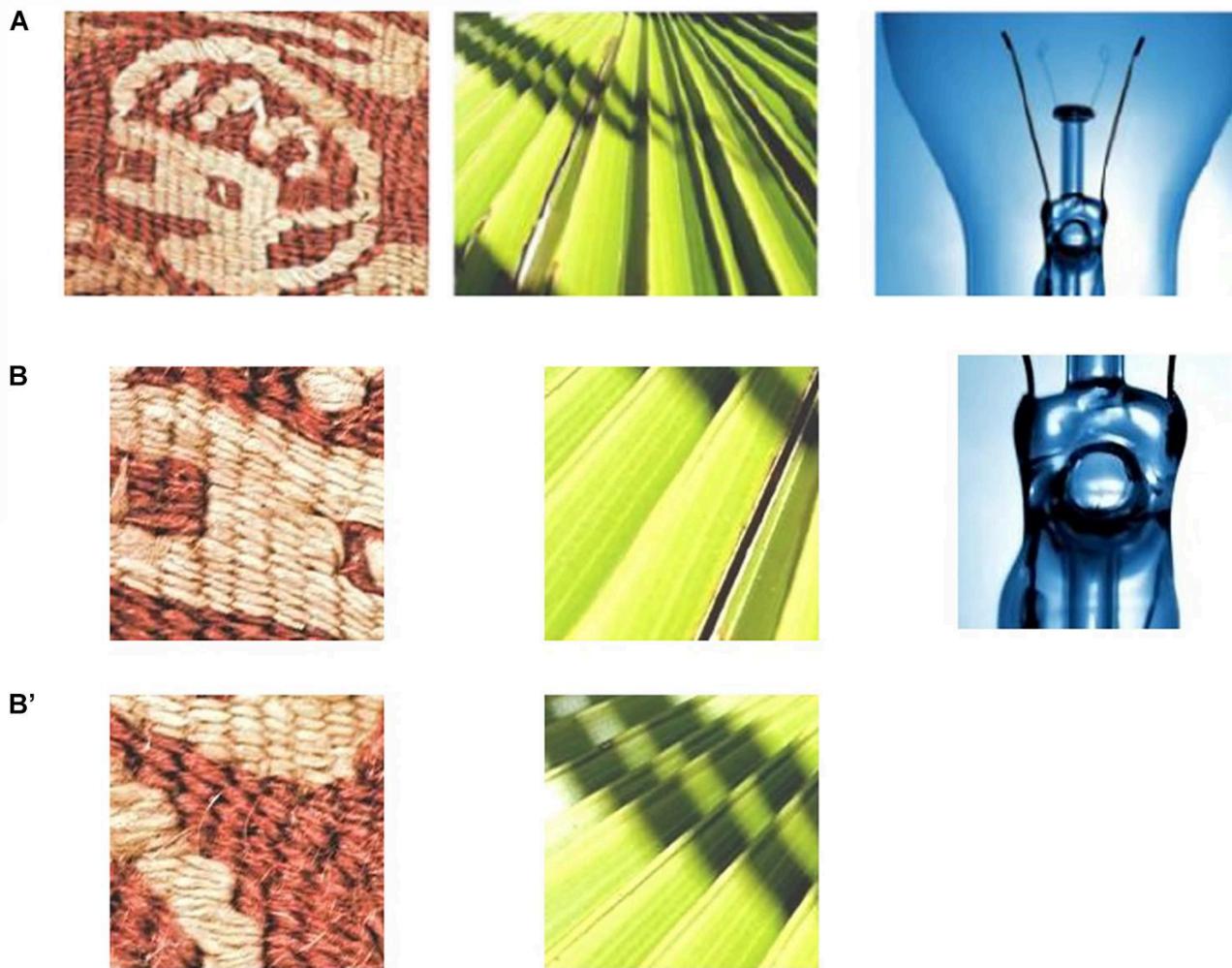
Algoritme de Canny

- Fa servir quatre passos per detectar les vores d'una imatge:
 1. Es redueix el soroll amb el **filtre de Gauss**.
 2. Calcula el gradient de la imatge amb el **filtre de Sobel**.
 3. Es detecten les vores amb el **mètode de supressió de no-màxims**.
 4. Es decideixen quines vores són vàlides amb el **mètode de la histèresi**.



Textura

- En visió artificial entenem com a textura un **patró de píxels** que es observable en una imatge.
 - Ex: Finestres en un edifici, taques en una vaca, etc.
- Ajuden, al igual que les vores, a **identificar objectes**.



Característiques de la textura

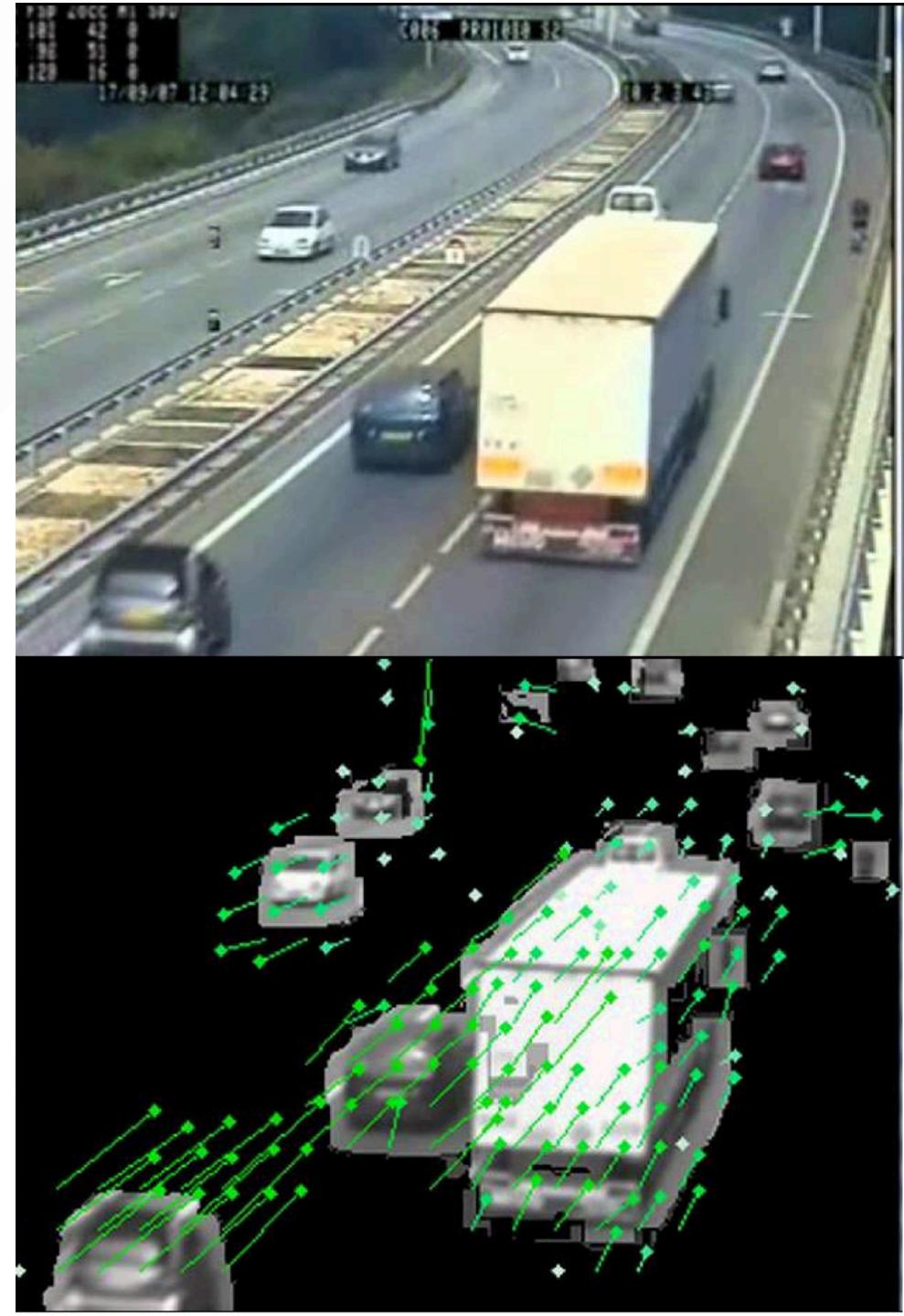
- La textura pot ser més o menys regular, per lo que es freqüent utilitzar un model de **tesel·les** per descriure-la. Vejam algunes característiques:
 - **Tesel·la**: patró que es repeteix en una imatge.
 - **Tesel·lació**: procés de cobrir una superfície amb tesel·les.
 - **Tipus**:
 - **Regulars**: es repeteixen sempre de la mateixa manera.
 - **Irregulars**: no hi ha un patró clar de repetició.
 - **Escala**: la textura pot ser més o menys gran.

Utilitats de la textura

- **Identificació:** permet identificar objectes. Ex: un cavall té una textura diferent a la d'una zebra.
- **Correspondència:** permet trobar zones corresponents en diferents imatges. Important en la reconstrucció 3D.
- **Segmentació:** permet separar la imatge en diferents regions.
- **Reconstrucció:** permet reconstruir la imatge a partir de les tesel·les.
- **Classificació:** permet classificar objectes.

Fluix òptic

- El **fluix òptic** és la **velocitat aparent** amb la que es mouen els objectes entre dues imatges.
- Els algoritmes de visió artificial són capaços de calcular el fluix òptic a partir de diferents imatges.
- Els algoritmes de càlcul de fluix òptic són molt útils per a la **reconstrucció 3D**, descriure el **moviment** i la descomposició de la imatge en **capes**.



Segments

- Anomenen **segments** a les **regions** de la imatge que tenen alguna propietat comuna (color, textura, forma, etc.).
- Per definit els segments hi ha dós enfocaments principals:
 - **Basat en límits**: es busquen els límits de les regions. Es pot entendre com un problema de *classificació* on cada pizel pertany o no a un segment i es soluciona amb tècniques de machine learning i models preentrenats.
 - **Basat en regions**: s'agrupen els pixels en regions segons alguna propietat comuna. Es pot entendre com un problema de b i es soluciona amb tècniques com k-means, etc.

Tipus de segmentació

- Tipus de segmentació:
 - **Segmentació binària:** es segmenta la imatge en dues regions: objecte i fons.
 - **Segmentació semàntica:** es segmenta en categories predefinides.
 - **Segmentació d'instàncies:** es segmenta en instàncies d'objectes.
 - **Segmentació panòptica:** es segmenta en categories predefinides, però també es segmenten les instàncies d'objectes.



(a) Image



(b) Semantic Segmentation



(c) Instance Segmentation



(d) Panoptic Segmentation



Tasques de visió artificial



Definició

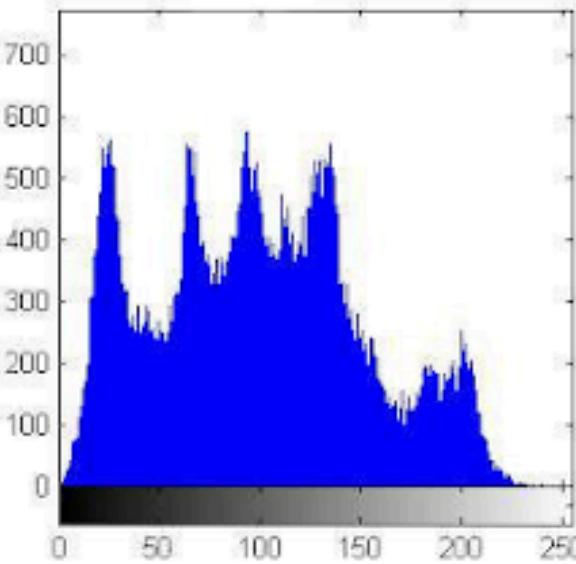
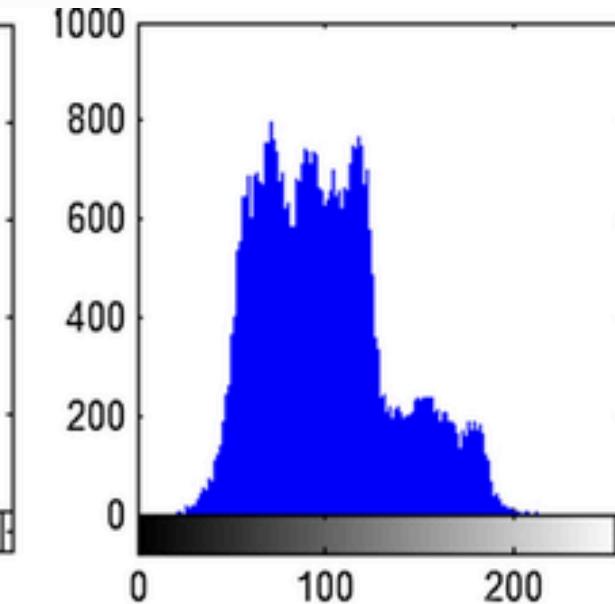
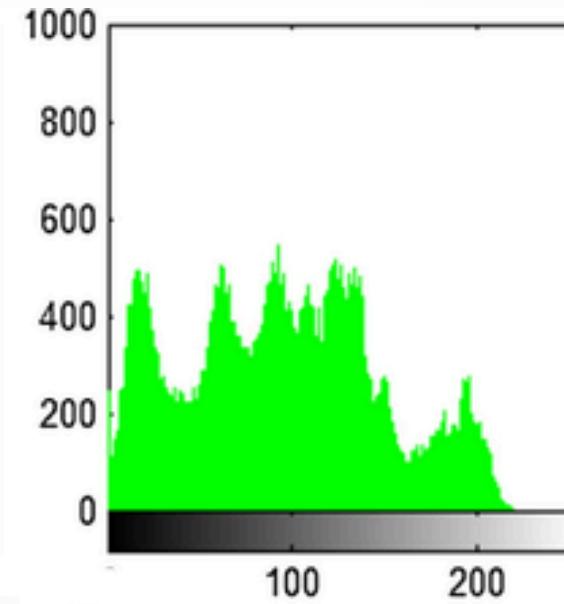
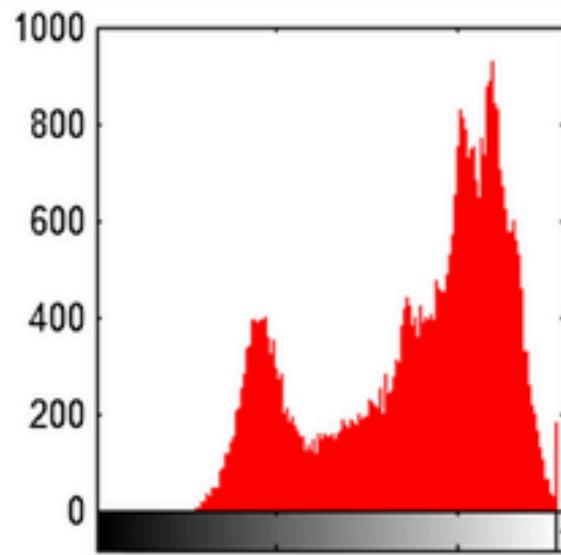
- Les tasques de visió artificial són aquelles que es poden realitzar a partir d'imatges.
- Veurem algunes de les més importants:
 - **Preprocessament d'imatges**
 - **Classificació d'imatges i Reconeixement d'objectes**
 - **Reconstrucció 3D**
 - **Localització**
 - **Segmentació**
 - **Reconstrucció**

Preprocessament d'imatges

- El **preprocessament** d'imatges és el conjunt de tècniques que s'apliquen a les imatges per tal de millorar-ne la qualitat o per tal d'extreure'n informació.
- Històricament, el preprocessament d'imatges era una tasca molt important, però amb l'aparició de les xarxes neuronals, aquesta tasca ha perdut importància.
- Tot i això, el preprocessament d'imatges segueix sent una tasca important en visió artificial.
- Veurem algunes de les tècniques més comunes.

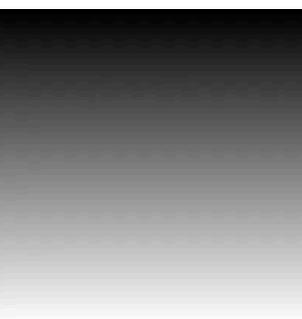
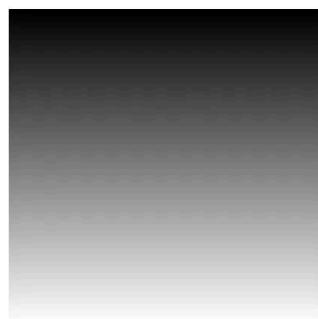
Histogrames

- El **histograma** d'una imatge és la representació gràfica de la distribució dels píxels en funció de la seva intensitat.
- Els histogrames són molt útils per entendre la distribució dels píxels en una imatge.
- Son molt utilitzats en el preprocessament d'imatges per tal de normalitzar-les.
- Els histogrames es poden calcular per cada canal de color (R, G, B) o per la imatge en escala de grisos.
- S'utilitzen molt en la **normalització** d'imatges.



Equalització de l'histograma

- L'**equalització de l'histograma** és una tècnica que es fa servir per tal de millorar el contrast d'una imatge.
- L'objectiu és que la distribució dels píxels sigui més uniforme.
- Es divideix l'histograma en *bins* i es redistribueixen els píxels de manera que la distribució sigui més uniforme.
- Els efectes moltes vegades no són realistes, però si sovint són útils per a tasques de visió artificial.



Filtratge

- El **filtratge** és una tècnica que es fa servir per tal de millorar la qualitat de la imatge.
- Hi ha molts tipus de filtres, però els més comuns són els filtres de **suavitzat** i els filtres de **realçament**.
- Els filtres de suavitzat són útils per tal de reduir el soroll de la imatge.
- Els filtres de realçament són útils per tal de millorar el contrast de la imatge.
- Els filtres es poden aplicar a tota la imatge o a una regió concreta.

Filtres de suavitzat

- El soroll és un problema comú en les imatges.
- Podem reduir el soroll de la imatge aplicant filters de suavitzat.
- Els filters més comuns són el **filtre de mitjana** i el **filtre de Gauss**.
- Filtre de mitjana: substitueix cada píxel per la mitjana dels píxels del seu entorn.
- Filtre de Gauss: substitueix cada píxel per la mitjana ponderada dels píxels del seu entorn.
 - Els píxels tenen un pes més gran com més propers estan al píxel central.

Filtres de realçament

- Els filters de realçament són útils per tal de millorar el contrast de la imatge. Molt utilitzats en la detecció de vores.
- Els filters més comuns són:
 - **Filtre de Sobel:**
 - Calcula el gradient de la imatge, és a dir, la intensitat de canvi de la imatge.
 - **Filtre de Laplace:**
 - calcula el laplacià de la imatge, és a dir, la segona derivada de la imatge.

Thresholding

- El **thresholding** és una tècnica que es fa servir per tal de binaritzar una imatge.
- El thresholding es fa aplicant un **llindar** a la imatge.
- Els píxels que tenen una intensitat superior al llindar es converteixen en blancs i els que tenen una intensitat inferior es converteixen en negres.
- És una forma simple de **segmentació**: es vol separar la imatge en dues regions: *objecte* i *fons*.
- Sol ser millor opció que ajustar la lluminositat i el contrast.



Global Thresholding



Adaptive Thresholding



- Les tècniques més comunes són:
 - Filtratge
 - Normalització
 - Correcció de distorsions
 - Reducció de soroll

- **Filtratge**: aplicar filtres a la imatge per tal de millorar-ne la qualitat.
- **Normalització**: ajustar la imatge per tal de que tingui una mitjana i una desviació estàndard determinades.
- **Correcció de distorsions**: corregir les distorsions de la imatge.
- **Reducció de soroll**: reduir el soroll de la imatge.

Classificació d'imatges

- Entenem per classificació d'imatges dues tasques:
 - **Classificació:** donada una imatge, es vol saber a quin objecte correspon.
 - **Reconeixement d'objectes:** donada una imatge, es vol saber quins objectes hi ha.
-

