

# Pràctiques de Visió Artificial

---

## Pràctica 3: Agrupació i segmentació.

### (Part I)

---

Els temes principals d'aquesta pràctica són:

- 1) Agrupació
- 2) Segmentació d'imatges

Per completar la pràctica és necessari conèixer els següents conceptes de teoria: Agrupació per K-Means i Mean Shift.

Els següents apartats del llibre “Computer Vision: Algorithms and Applications” de Richard Szeliski proporcionen informació addicional a la vista en les classes de teoria.

Chapter 5: Segmentation.

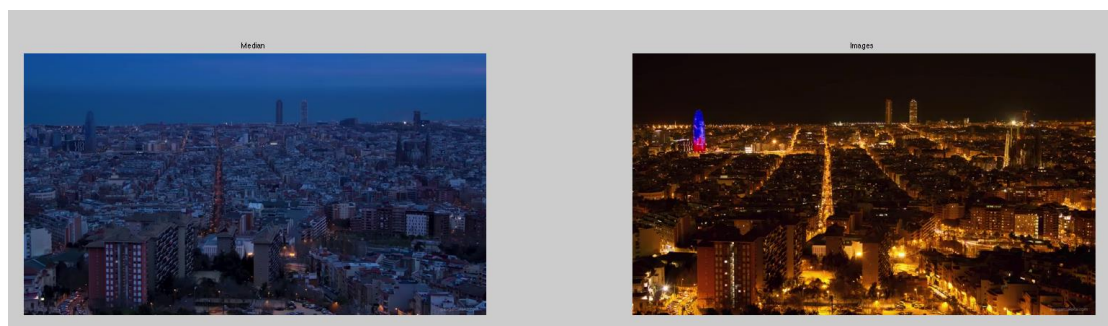
**3.1: Mètodes de “background subtraction”:** donat el vídeo: Barcelona.mp4, que conté imatges adquirides amb càmeres estàtiques, eliminar tots els artefactes relacionades amb el moviment extraient les imatges del fon.

Nota: Una de les aplicacions d'aquesta opció es pot relacionar amb el 'botó elimina turistes' inventat per a les càmeres de fotos: per exemple, Adobe acaba de presentar el 'Monument mode' que suprimeix automàticament la gent que passa davant la càmera.

Per realitzar l'algorisme, es necessita:

- a) trobar on s'acaba una escena i on comença l'altra (aquests frames es denominen els shots del vídeo). Quina mesura de les imatges pots utilitzar per distingir les escenes?
- b) aplicar un algorisme de background subtraction (consultar material de teoria).

Visualitzar per cada segment delimitat per dos shots del vídeo la imatge estàtica extraïda i les imatges a partir dels quals l'has obtingut, per exemple:



Comenta en detalls la implementació. Què succeeix si els shots no estan correctament extrets? Què succeeix si trobes massa shots del vídeo? Comenta que representen les imatges estàtiques obtingudes. En quines situacions l'algorisme funciona i quan no?

Comenta si veus altres aplicacions que es poden treure d'aquest algorisme.

**Opcional:** aplica'l per algun més dels següents vídeos o vídeos que tu has trobat on l'algorisme té sentit i aplicació:

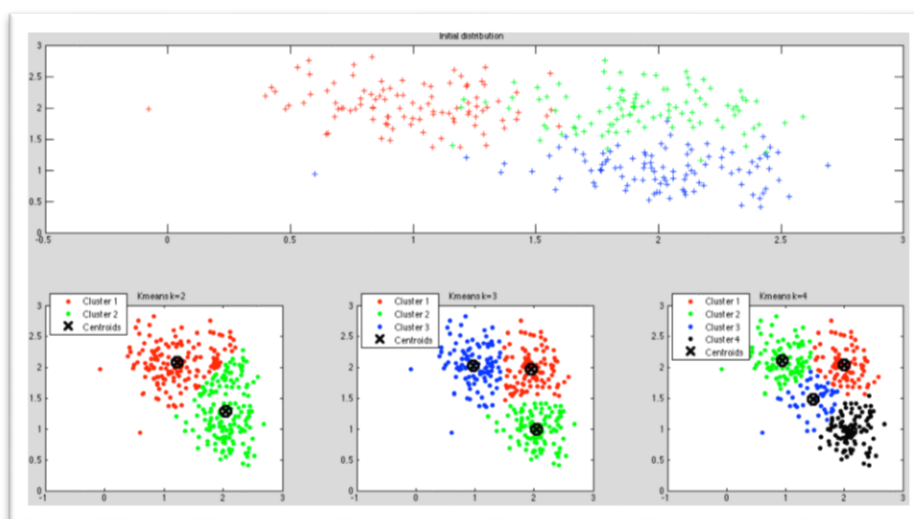
- [Hokkey:](#)
- [Red light running:](#)
- [Waynedale Memorial Parade](#)
- [Bilbao Timelapse](#)
- [Barcelona airport](#)
- [Boqueria \(25:00\), Colon \(27:00\), Sagrada família \(41:10\).](#)
- [Paseo de Gracia \(2:59\), Museo de Historia \(3:12\)](#)
- [Removing tourists in Barcelona pictures \(4:11\)](#)

**Opcional:** Que succeeix si restes de la imatge original la imatge del fons? Visualitza'l.

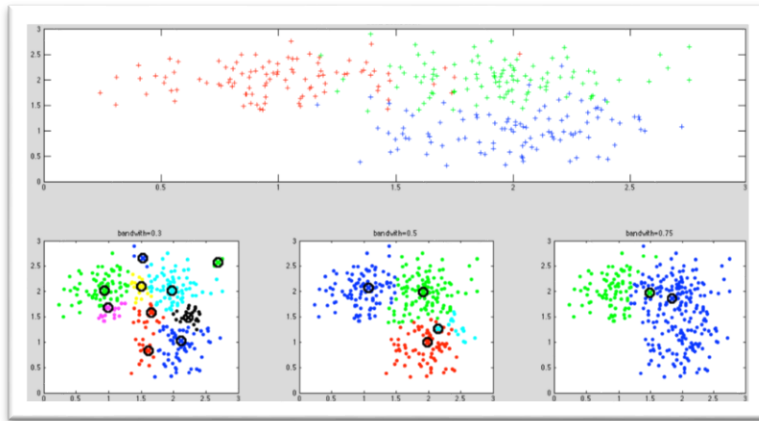
### 3.2 Mètodes d'agrupació de dades numèriques

Crea un fitxer `ex32.m` que implementi els següents passos:

- a) La funció `gaussRandom` (`mu`, `sigma`, `numSamples`) proporcionada amb l'enunciat permet generar núvols de punts amb una distribució gaussiana amb matriu de covariança diagonal, utilitzant com a centre les coordenades de la `mu`. Genera tres nubols de 100 punts amb centre `[1 2]`, `[2 2]` i `[2 1]`. En els tres casos utilitza una desviació estàndard de 0.1 en tots els eixos. Visualitza els punts generats en l'apartat anterior (Ayuda: `plot`).
- b) Utilitza el mètode `kmeans` per agrupar les dades anteriors. Visualitza en un mateix plot (utilitzant `subplot`) una primera fila amb les dades originals, i el resultat (incloent els centres) de les agrupacions amb 2, 3 i 4 centres respectivament en la segona fila del `subplot`, utilitzant diferents colors.



- c) Utilitza el mètode *MeanShiftCluster* per agrupar les dades generades en el primer apartat. Mostra els resultats en el mateix format que l'apartat anterior per valors de finestra (bandwidth) de 0.3, 0.5 i 0.75



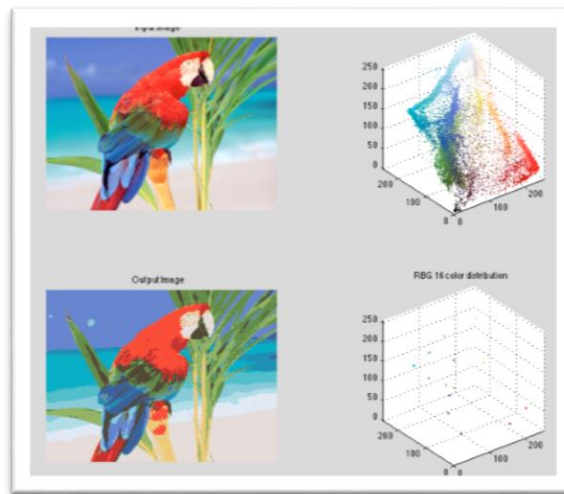
respectivament.

- d) Comenta els resultats que has trobat, valorant el nombre d'agrupacions que has trobat en cada cas i la similitud amb les dades inicials.

**Nota:** Fixeu-vos que en aquests exercicis les característiques per les que s'aplica el k-means (i que defineix l'espai de característiques) són la posició dels punts, és dir, les coordenades dels punts.

### 3.3 Mètodes d'agrupació: Segmentació en l'espai RGB

- a) Carrega la imatge loro.png. Converteix-la en nivells de gris i aplica la segmentació amb kmeans. Prova diferents k per trobar la millor segmentació.  
Ajuda: Fixa't com s'han de convertir els píxels a característiques i en quin format s'han de passar al kmeans. Nota: El kmeans funciona millor quan les característiques són com a mínim 2D. Per això, en el cas 1D, crea una matriu de 2 columnes on les dues columnes coincideixen.  
Visualitza la imatge segmentada utilitzant el nivell de gris promig trobat pel mètode de la segmentació. A què correspon?
- b) Afegeix com a característiques les coordenades dels píxels i comprova si millora el resultat de la segmentació.
- c) Visualitza-la en una figura juntament amb la distribució dels seus colors. Pots utilitzar el mètode *plotPixelDistribution*.
- d) A partir de la imatge d'entrada, crea una matriu que contingui a cada fila la tripleta RGB d'un píxel de la imatge. Tindrem tantes files com píxels a la imatge.
- e) Utilitza el mètode *kmeans* per reduir el nombre de colors de la imatge a només 16 colors diferents.
- f) Visualitza en una mateixa figura les imatges del primer apartat, i la imatge amb 16 colors juntament amb la seva distribució de colors.



### 3.4 (Opcional) Mètodes d'agrupació: Segmentació en l'espai RGB amb Mean-shift

Repeteix els mateixos passos del exercici 3.4 utilitzan el mètode Mean-shift en lloc del kmeans.

Amb l'enunciat se us proporciona el mètode *meanShiftPixCluster*. Al principi del mètode trobareu un codi comentat. Utilitzeu aquest codi per veure quin és l'efecte de canviar els paràmetres *hs* i *hr*. Expliqueu les vostres conclusions comparant els dos mètodes (kmeans i mean-shift).

**Data de lliurament (conjuntament amb la II part de la pràctica 3): 17 de Novembre a les 23:00h. El document ha de contenir un document pdf amb els comentaris de les implementacions dels algorismes, les observacions i les respostes de les preguntes de l'enunciat i tres fitxers .m amb les implementacions dels tres exercicis.**