

21739 - Percepció i Control per a Sistemes Encastats

Grau en Enginyeria Informàtica

Sistemes de Percepció: Pràctica 3

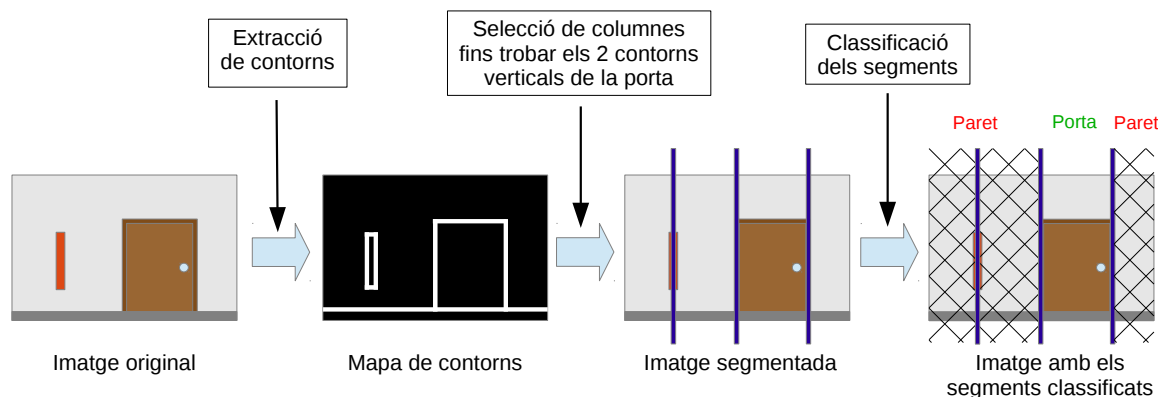
Xisco Bonnín

Universitat de les Illes Balears
31 de maig de 2022

Detecció de portes

Volem dissenyar un sistema que, donada una imatge on aparegui una porta, sigui capaç d'indicar on està situada. Partim del resultat de la pràctica anterior que consisteix en una segmentació de la imatge per les N columnes corresponents a les N rectes verticals més notòries de la imatge, incloent els dos contorns verticals de la porta.

A continuació es detalla la tasca completa per tal que puguis entendre millor on encaixa el que ja hem implementat i quina és la tasca que ens falta resoldre:



El que volem fer ara és classificar cada un dels segments de la imatge segons si conté o no la porta. Definim així dues classes: *porta* i *paret*. A continuació es detallen els passos a seguir per tal d'assolir el nostre objectiu.

1 Obtenció de les característiques

Es vol basar el procés de classificació en **tres** de les característiques del conjunt següent:

- To mitjà (en anglès *hue*): correspon al valor mitjà del primer canal de la imatge expressada en l'espai HSV.
- Saturació mitjana: correspon al valor mitjà del segon canal de la imatge expressada en l'espai HSV.
- Intensitat mitjana: correspon al valor mitjà del tercer canal de la imatge expressada en l'espai HSV.
- Energia: s'obté a partir de la *Gray-Level Co-occurrence Matrix* (GLCM) mitjançant la fórmula:

$$energy = \sqrt{\sum_{i=0}^{NL-1} \sum_{j=0}^{NL-1} p(i,j)^2}$$

on NL és el nombre de nivells de gris usats per generar la GLCM, i $p(i, j)$ són cada un dels elements de la GLCM.

- Homogeneïtat: també coneguda com a *Inverse Difference Moment*. S'obté a partir de la GLCM mitjançant la fórmula:

$$IDF = \sum_{i=0}^{NL-1} \sum_{j=0}^{NL-1} \frac{p(i, j)}{1 + (i - j)^2}$$

Selecciona tres de les característiques anteriors en funció del teu nombre de grup, tal com indica la taula següent:

	To	Saturació	Intensitat	Energia	Homogeneïtat
Grup 1	×	×		×	
Grup 2	×	×			×
Grup 3		×	×	×	
Grup 4		×		×	×

Per dur a terme l'aprenentatge, se't proporcionen dues col·leccions de fragments de **30 × 30 píxels** corresponents a imatges de cada una de les classes: porta i paret. Trobaràs les dues col·leccions a dins la carpeta **patches**. Per cada un dels fragments hauràs d'obtenir les tres característiques esmentades, generant la taula $data[m + n, 4]$:

	Característica 1	Característica 2	Característica 3	Classe
door_patch_1	$carac_1(dp_1)$	$carac_2(dp_1)$	$carac_3(dp_1)$	1
...	1
door_patch_m	$carac_1(dp_m)$	$carac_2(dp_m)$	$carac_3(dp_m)$	1
wall_patch_1	$carac_1(wp_1)$	$carac_2(wp_1)$	$carac_3(wp_1)$	2
...	2
wall_patch_n	$carac_1(wp_n)$	$carac_2(wp_n)$	$carac_3(wp_n)$	2

on m i n són el nombre de fragments de porta i de paret respectivament.

Notes:

1. Calcula la GLCM fent servir la funció `feature.greycomatrix` per a una distància de 1 píxel en horitzontal (angle = 0 radians), usant 256 nivells de gris i fent que retorni la matriu normalitzada. Veure API.
2. Calcula els valors de l'energia i/o l'homogeneïtat fent servir la funció `feature.greycoprops`. Veure API.

2 Aprenentatge

Un cop has obtingut la taula *data* implementa l'algorisme del **perceptró (variant de la butxaca)** per tal de separar les mostres corresponents a les dues classes. Tingues en compte que l'has d'implementar per un espai tridimensional i, per tant, la frontera de discriminació vindrà donada per un pla $w = [w1 \ w2 \ w3 \ w0]$.

Visualitza el resultat usant la funció `plot3D` (fent `from mpl_toolkits import mplot3d` i `ax = matplotlib.pyplot.axes(projection='3d')` prèviament) per tal de pintar les mostres de les dues classes (utilitza colors diferents per a cada classe) i pinta-hi a sobre el pla obtingut amb el perceptró utilitzant el codi següent:

```
# crea valor pels eixos x,y
x = y = numpy.arange(0.0, 1.0, 0.1)
[xx,yy]=numpy.meshgrid(x,y)
# calcula els valors de z
z = (-w1*xx - w2*yy - w0)/w3
# pinta el pla de separació
ax.plot_surface(xx,yy,z)
```

on $[w1 \ w2 \ w3 \ w0]$ són els coeficients del pla trobat amb el perceptró.

3 Classificació

Per tal de classificar cada un dels segments d'imatge obtinguts en la pràctica anterior segueix les passes següents:

1. Divideix el segment en fragments de 30×30 píxels (sense solapament entre ells).
2. Calcula els valors de les tres característiques per cada un dels fragments.
3. Classifica els fragments en les classes *porta* o *paret* usant el perceptró.
4. Classifica tot el segment de la manera següent:
 - Porta si $> 70\%$ dels fragments han sigut classificats com a porta
 - Paret si $> 60\%$ dels fragments han sigut classificats com a paret
 - Indeterminat si no es compleixen cap dels dos casos anteriors.

Un cop classificats tots els segments de la imatge, visualitza el resultat mostrant la imatge a color original on les parets estiguin pintades en negre i els espais indeterminats estiguin pintats en blanc.

Lliurament de la pràctica

La documentació a lliurar ha d'explicar de manera breu i entenedora el que heu fet per resoldre cada un dels apartats de la pràctica (punt per punt), s'ha d'incloure el codi Python que hàgiu fet servir, així com les imatges o figures que resultin de cada pas. Recorda incloure el resultat final obtingut per a totes i cada una de les imatges d'entrada. El lliurament es farà mitjançant *Aula Digital* i la data límit és el diumenge **12 de juny** a les 23:55.