

# Examen Parcial de Visió Artificial

29.10.2012

Nom:..... Cognoms:.....

DNI:..... NIUB:.....

Cada pregunta té una única resposta correcta. (Si hi ha més d'una resposta marcada, no se'n considerarà cap). Cada resposta correcta suma 0.5 punts. Les respostes incorrectes resten 0.125 punts.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
a															
b															
c															
d															

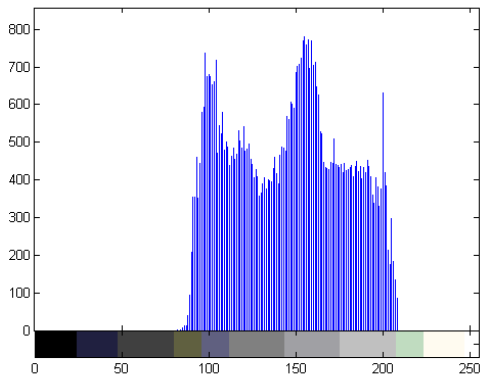
## 1. Determinar la resposta correcta:

- a) El model de formació de la imatge ens demostra que a un punt de la imatge correspon un únic punt a la realitat però a un punt de l'escena 3D poden correspondre varis punts a la imatge.
- b) El model de formació de la imatge ens permet entendre la relació entre el món físic i els píxels de la imatge. Donada la projecció perspectiva si sabem les coordenades a la imatge d'un punt i la distància de la imatge al sistema de coordenades, podem reconstruir el punt corresponent al Mon real.
- c) Un contorn a la imatge és degut a una discontinuïtat a la fórmula  $i = \frac{IR \cos(\theta)}{d^2}$ . Si hi ha un canvi brusc a la intensitat degut a discontinuïtat a la il·luminació, estem parlant de contorn tipus ombra.
- d) Cap de les anteriors.

## 2. La màscara [1, 2, 1; 0,0,0; -1,-2,-1]

- a) Implementa la màscara de Sobel.
- b) Permet trobar contorns horitzontals.
- c) Permet derivar la imatge verticalment.
- d) Totes les anteriors.

**3. Donat el següent histograma d'una imatge,**



- a) Una forma per obtenir aquest histograma d'una imatge en Matlab és utilitzant la comanda `histogram()`.
- b) L'histograma mostra que la imatge està poc contrastada. Per augmentar el contrast s'ha d'assegurar que el mínim valor de gris sigui 0, per tant s'ha de restar 80 de la imatge.
- c) Si s'aplica:  $(im(x,y) - \min(\min(im))) / (\max(\max(im)) - \min(\min(im))) * 255$ , on `im` és la imatge, s'aconseguirà augmentar el contrast de la imatge assegurant que l'histograma estigui cobrint el rang de valors entre 0 i 255.
- d) Cap de les anteriors.

**4. El detector de Canny:**

- a) Ens assegura que sempre els contorns dels objectes formaran corbes tancades.
- b) Està basat en el Laplacà de la Gaussiana més l'efecte histèresi després d'aprimar els contorns,
- c) És millor que els detectors basats en la convolució amb màscares de Sobel o Laplacà de la Gaussiana ja que assegura contorns primis i connectats,
- d) Histèresi consisteix en trobar els contorns més dèbils i connectar-los amb contorns forts.

**5. El programa:**

```
maskara=[1,4,1;4,-20,4;1,4,1];
```

```
im2=conv2(double(im),maskara);
```

```
imshow(abs(im2)<10,[]);
```

- a) Com els contorns donen resposta creuant per zero convolucionant amb el Laplacà, buscant els valors petits de la imatge `im2`, donarà on els contorns són més grans que 20,
- b) Donarà un error d'execució ja que intenta fer una binarització de la imatge, la comanda correcta a Matlab és `im2bw()`.
- c) Convoluciona una imatge amb una màscara que implementa el Laplacà, és a dir la suma de les segones derivades,
- d) Cap de les anteriors.

## 6. Per analitzar les textures:

a) Els textons són característiques visuals de la imatge que s'utilitzen per estudiar i explicar la percepció de les textures. Aspectes principals per detectar-los i descriure'ls són: el número d'extrems, la co-linearitat, els creuaments, etc.

b) Una imatge binària (amb valors de 0 i 1) té una única matriu de co-ocurrència de tamany 2x2.

c) L'energia és un descriptor de textura que s'extreu a partir de les matrius de co-ocurrència. Donada una imatge en nivell de gris, es calcula tants descriptors d'energia com píxels hi ha a la imatge i després es calcula el seu promig.

d) Cap de les anteriors.

## 7. Respecte els descriptors de textura:

a) Les matrius de co-ocurrència i els patrons locals binaris són mètodes d'anàlisi de textures que quantifiquen les variacions de la intensitat dels píxels i les relacions dels píxels dintre d'una veïnat tenint en compte el seu valor de gris. Aquesta veïnat correspon als veïns més propers del píxel i per tant és de mida 3 per 3.

b) Les dues subimatges:

40 41 50      50 56 56

43 56 89   i   49 66 99

39 99 75      58 87 96

tenen el mateix patró local binari.

c) Els patrons locals binaris no distingeixen entre textures verticals i horitzontals ja que sumen els canvis del veïns respecte el píxel central.

d) Cap de les anteriors.

## 8. Respecte la segmentació:

a) És un procés de dividir la imatge en regions homogènies i així assegurar que es trobin els diferents objectes de l'escena.

b) Existeixen dues estratègies per segmentar: guiar-se per un model i buscar-lo a la imatge (nominat top-down) o extreure regions homogènies a partir de trobar clústers de píxels amb el mateix color, textura, etc. (nominat bottom-up). El mètode k-means és top-down mentre el mètode mean-shift és bottom-up.

c) El mètode mean-shift està basat en el càlcul del Gradient de la funció de densitat probabilística. El gran avantatge és que té en compte la distribució probabilística de les dades, encara que no calcula explícitament la distribució de les dades sinó el seu Gradient per trobar el valor de gris més probable de la regió (el mode de la distribució probabilística).

d) Totes les anteriors.

**9. Respecte la convolució d'una imatge,**

- a) Si es convoluciona una imatge amb la màscara [1,1,1] tindrà un efecte de suavització i difuminació dels contorns
- b) Convolucionant amb la suma de les primeres derivades del Gradient, donarà zeros justament als llocs dels contorns de la imatge,
- c) Si es convoluciona una imatge amb una màscara [0,0,0;0,1,0;0,0,0], torna una còpia de la imatge original.

d) Totes les anteriors.

**10. Convolucionant una imatge amb la màscara: [1,3,5,7,9;2,4,6,8,10;3,5,7,9,11]**

- a) no es pot fer perquè no implementa cap derivada
- b) no es pot fer perquè la màscara no és quadrada

c) la convolució està basada en la següent operació: per cada píxel de la imatge centra la màscara al píxel i multiplica els píxels de tot el veïnat amb els números de la màscara corresponent i fa la suma d'aquestes multiplicacions,

d) cap de les anteriors.

**11. El paràmetre "sigma" en una funció Gaussiana, determina la mida del nucli (kernel)**

- a) Determina els contorns que extraurem: horitzontals o verticals.
- b) El mínim sigma de la Gaussiana ha de ser igual a 1 que seria corresponent a utilitzar veïns a distància 1.
- c) Disminuint sigma aconseguim que només els veïns més propers afectin el càlcul centrat al píxel.
- d) Cap de les anteriors.

**12. A Matlab, donada la matriu a = [ 1 1 1; 2 2 2; 3 3 3; 4 4 4], el resultat deb = reshape(a,1,12) és:**

- a) b = [1 1 1 2 2 2 3 3 3 4 4 4];
- b) b = [1 2 3 4 1 2 3 4 1 2 3 4];
- c) b = [1; 1; 1; 2; 2; 2; 3; 3; 3; 4; 4; 4];
- d) Cap de les anteriors.

**13. Donada la matriu a = uint8(ones (256, 256)), per visualitzar una imatge blanca utilitzarem la instrucció:**

- a) imshow (a, [0 256])
- b) imshow (a)
- c) imshow (a, [0 1])
- d) imshow (a, [1 max (max(a))]).

**14. L'histograma H = imhist(A,nBins) de la següent imatge A = 100\*ones(100,100):**

- a) És un vector H = [0 0 100 100], per nBins = 4
- b) És un vector H = [0 0 0 10000], per nBins = 4
- c) És un vector H = [0 100], per nBins = 2
- d) Cap de les anteriors.

**15. Donats els vectors:**

a = [10 13 15 17];

b = [0 0 0 0];

Amb quina instrucció s'aconsegueix b =[0 0 0 170]?

- a) b(4) = a (4) \*10
- b) b(a>16) = a(a>16)\*10
- c) b(a == 17) = a (a ==17) .\*10
- d) Totes les anteriors.

## Problemes

1. (1 punts)Donada la matriu A

1 2 3 4

5 6 7 8

8 7 6 5

4 3 2 1

Realitzar les següents operacions, amb codi MatLab, sense la utilització de bucles:

- 1.1. Assigna el valor 1 a tots els elements de la tercera fila.
- 1.2. Assigna el valor 10 a tots els elements de la matriu majors que 6
- 1.3. Crea una matriu A3 de 3 canals, de 4 files i 4 columnes de tipus uint8. Assigna al primer canal una matriu de zeros, al segon canal una matriu d'uns i al tercer, la matriu A.
- 1.4. Crea i mostra l'histograma corresponent a la matriu A, utilitzant 4 bins.

Respostes 1:

- 1.1 A(3,:) = 1;
- 1.2 A(A>6) = 10;
- 1.3 A3 = zeros (4, 4, 3, 'uint8');  
A3(:,:,1) = 0;  
A3(:,:,2) = 1;  
A3(:,:,3) = A;
- 1.4 hist = imhist(A, 4);  
bar(hist);

2. (1.5 punts)La càmera Kinect, gràcies als seus 2 sensors, pot capturar dos streams de vídeo, un corresponent a la imatge en color de l'escena i l'atre a la profunditat dels elements. Amb MatLab, mitjançant les llibreries kinect-mex, podem accedir als frames capturats per la càmera.

- 2.1. Aplica kmeans (4 clusters) sobre la imatge rgb, tenint en compte que tindrem tants punts com píxels té la imatge i 3 variables, corresponents als canals R, G i B.
- 2.2. Què representa cada píxel de la imatge de profunditat? Perquè no podem utilitzar una matriu de tipus uint8?
- 2.3. Què signifiquen i en quines ocasions apareixen píxels amb valor zero a la imatge de profunditat?

- 2.4. Supposem que fem una captura amb la Kinect de la imatge rgb i la guardem com a rgb.tif i fem el mateix amb la imatge de profunditat, guardada com a depth.tif
- 2.4.1. Carrega les dues imatges amb MatLab.
- 2.4.2. A partir de la matriu de profunditat, crea una matriu binària on els elements valdran 1 en cas que la profunditat sigui major que 200 i 0 en cas contrari.

Respostes 2:

#### 2.1

```
kmeansData = reshape (rgb, size(rgb,1)*size(rgb,2), 3);
```

```
IDX = kmeans (double(kmeansData), 4);
```

2.2 Cada píxel de la imatge de profunditat representa la distància a la que es troba el punt respecte el pla de la càmera, mesurat en mil·límetres. La matriu no pot ser de tipus uint8 perquè d'aquesta forma no podríem representar objectes situats a més de 255 mil·límetres. La imatge és de tipus uint16.

2.3 Significa que la càmera no ha pogut determinar la distància. Això succeeix quan el sensor d'infrarojos no troba el patró projectat pel làser, degut a oclusions, al tipus de superfície o a interferències amb la llum solar.

#### 2.4.1

```
depth = imread('depth.tif');
```

```
rgb = imread('rgb.tif');
```

#### 2.4.2

```
bin = zeros (size(depth));
```

```
bin (depth > 200) = 1;
```

\*Ajuda: helpkmeans.

KMEANS K-meansclustering.

IDX = KMEANS(X, K) partitionsthepoints in the N-by-P data matrix X into K clusters. Thispartitionminimizesthesum, over all clusters, of thewithin-clustersums of point-to-cluster-centroiddistances. Rows of X correspond to points, columnscorrespond to variables. KMEANS returnsan N-by-1 vector IDX containingtheclusterindices of eachpoint. Bydefault, KMEANS uses squaredEuclidean distances.

KMEANS treatsNaNs as missing data, and ignores any rows of X thatcontainNaNs.

[IDX, C] = KMEANS(X, K) returnsthe K clustercentroidlocations in the K-by-P matrix C.