

1) Un enginyer de telecomunicacions vol utilitzar l'algorisme de Harris per a obtenir uns bons descriptors de keypoints. Raona si et sembla una bona estratègia.

No ho és. Harris és una tècnica per a detectar vèrtexos, no per a descriure keypoints

2) Per a segmentar una imatge utilitzant la tècnica de watershed, no es treballa habitualment sobre la imatge original, sino sobre el mòdul del gradient. Per quin motiu?

Per aplicar el procés d'inundació de watershed cal, interpretant la imatge com una superfície topogràfica, que les regions es corresponguin amb pous i les fronteres amb muntanyes. El mòdul del gradient compleix aquests requisits

3) Quin significat té el pixel (i,j) d'una imatge integral ?

És la suma de tots els píxels de la imatge situats en coordenades $\leq (i,j)$

4) Els HoG, com d'invariants són a la rotació? Raona la resposta

Les rotacions de l'objecte es tradueixen en shiftats de l'histograma

5) Com podriem reduir la dimensionalitat d'un histograma en l'espai RGB, fent-lo alhoramés invariant a canvis d'il.luminació?

Podem fer un histograma 2D: $[r,g]$
On $r = R/(R+G+B)$, $g = G/(R+G+B)$

6) Hem utilitzat un classificador per a autoritzar l'accés a persones autoritzades mitjançant imatges del seu iris. Realitzem un experiment amb un conjunt molt gran d'imatges de test, i obtenim la tasa de error, de falsos positius, de falsos negatius, i el Fscore. Si ja hem obtingut aquests paràmetres de resultat, per quin motiu hauriem de calcular la corba ROC?

La corba ROC ens aporta alguna informació addicional? Els classificadors en poden es poden entrenar per a obtenir diferents tases de falsos positius/negatius, és per aquest motiu que ens cal una gràfica per a mostrar els diferents resultats (positius certes.Falsos positius).

7) Descriu, pas per pas, com detectaries circumferències en una imatge usant la transformada Hough.

1. Crear una matriu 3D d'acumuladors $[a,b,r]$
2. Per a cada píxel de contorn, incrementar aquelles posicions de la matriu que satisfan l'equació: $(x-a)^2 + (y-b)^2 = r^2$
3. Buscar els màxims de la matriu

Com reduiríes la dimensionalitat de la taula usant la informació de la direcció del gradient ?

Equació en coordenades polars: $x = x_0 + r \cos \alpha$ y $y = y_0 + r \sin \alpha$ Coneixent la direcció del gradient α , fixem un grau de llibertat i passem d'una taula de Hough 3D a una 2D.

9) (1.5p) Donat un conjunt d'aparellaments putatius entre píxels de les imatges A i B, descriu com usaries la transformada de Hough per a trobar la matriu de transformació afí entre les imatges A i B

Agafem el primer aparellament: $(X_1, Y_1) \rightarrow (X'_1, Y'_1)$

Apliquem la matriu de transformació:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \begin{bmatrix} x'_1 \\ y'_1 \\ 1 \end{bmatrix}$$

Plantegem les 2 equacions: $aX_1 + bY_1 + c = X'_1$
 $dX_1 + eY_1 + f = Y'_1$

En una matriu 3D, incrementem en 1 totes aquelles posicions (a,b,c) que satisfan l'equació 1.

En una altra matriu 3D, incrementem en 1 totes aquelles posicions (d,e,f) que satisfan l'equació 2.

Repetim el procés per a la resta d'aparellaments

Busquem les posicions (a,b,c) i (d,e,f) que presenten un valor màxim en ambdues matrius d'acumuladors.

8) Volem descriure la forma d'un objecte. Enumera els avantatges i inconvenients d'una signatura(ϕ) (distància vs. angle), respecte a una funció slope density.

La signatura $r(\phi)$ té problemes amb les concavitats. La dimensió de la slope density no és constant, depèn del nombre de píxels de la silueta de l'objecte

1) Un enginyer de telecomunicacions presenta el següent pseudocodi:

```
ee = genera_element_estructurant ('disc',17)
```

```
im_oberta_1 = open (im_orig, ee)
```

```
im_oberta_2 = open (im_oberta_1, ee)
```

Quin resultat conté la imatge im_oberta_1? Quin resultat conté la imatge im_oberta_2?

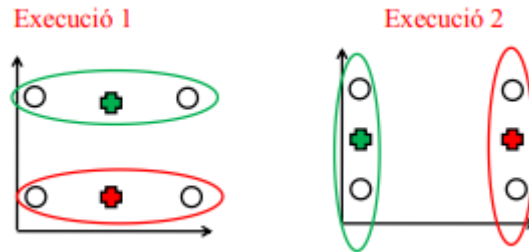
Filtrarà totes les components blanques de mida menor que l'element estructurant
El mateix. L'open és una operació idempotent

3) Com afecta la saturació d'un píxel en la precisió en el càlcul del seu hue?

Quan la saturació és molt baixa, el píxel és pràcticament gris, i l'error en el càlcul del hue és molt elevat. En el cas extrem, si la saturació és 0, el hue no està definit.

2) Volem segmentar una imatge usant l'algorisme k-means. Les coordenades inicials dels kclusterses generen aleatòriament. El resultat final, pot dependre de la posició inicial dels clusters? Raona la resposta amb un exemple gràfic.

Si, el resultat final depèn de la posició inicial dels clusters



4) Per a calcular el gradient d'una imatge greylevel, volem usar un dels kernels de convolució següents: $Gx1 = [-1, 0, 1]$ o $Gx2 = [-1, 1]$ Quins avantatges i inconvenients té usar $Gx1$ respecte $Gx2$?

$Gx2$ implementa una cosa més semblant a una derivada, doncs els píxels a restar estan més propers. En canvi, presenta el problema que el imatge gradient resultant estarà desplaçada un píxel.

5) Volem realçar els contorns d'una imatge mitjançant convolucions, però només disposem filtres integratius. Explica com ho faries

$Im_Realçada = Im_Orig - filtratIntegratiu (Im_Orig)$

6) Explica com funciona la tècnica de binaritzat 'moving averages'.

El nivell de gris del píxel a binaritzar, es compara amb el promig dels nivells de tots els píxels inclosos en una finestra lliscant.

7) Explica com implementar un closing morfològic a base de dilatacions, sense usar erosions.

$Im_close = complementa (dilata (complementa (dilata (im_orig))))$

8) És el 4-tree la estructura de dades idònea per a implementar l'algorisme 'split &merge' ? Raona la resposta.

No. Funcionaria molt be per a l'etapa de split, però per a fer el merge ens caldria un graf d'adjacència.

2) En quin moment aturarem les iteracions d'un algorisme k-means ?

En el moment en que els centres dels clusters ja no variïn més la seva posició.

1) Un enginyer de telecomunicacions presenta el següent pseudocodi per a visualitzar la imatge resultat d'unes operacions morfològiques:

```
ee = [ 1, 1, 1]
im_dil = dilata (im_orig, ee)
im_ero = erosiona (im_orig, ee)
im_dif = im_dil - im_ero
im_res = (im_dif < 0)
display (im_res)
```

Quina imatge es veurà en pantalla ? Justifica la resposta

Sempre es veurà una imatge totalment negra. El resultat de la dilatació sempre serà més gran que el de l'erosió.

3) Calculem l'histograma del hue d'una imatge color i ens dona un histograma pla. Quin és el hue promig (en radians) de la imatge ?

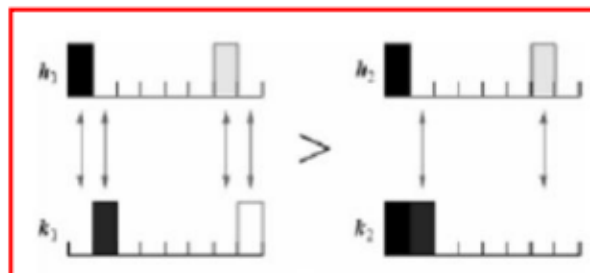
És indeterminat. El hue és un angle, i un promig cíclic de totes les orientacions no té sentit.

7) Donat un blob de N pixels en una imatge, què volen dir els moments $M00$, $M10/N$, i $M01/N$?

$M00$ = Area ($M10/N$, $M01/N$) = centre de masses

4) Mesurem la similitud entre els histogrames h_1 i h_2 usant $\text{distHistos} = \text{SumaElements}(|h_1 - h_2|)$. Quin inconvenient pot tenir fer-ho d'aquesta manera? Il·lustra-ho amb un exemple.

Histogrames molt similars poden donar distàncies molt grans per causa de la discretització. Les Mostres d'histogrames similars poden no caure exactament en els mateixos bins, sino en inmoltproppers.

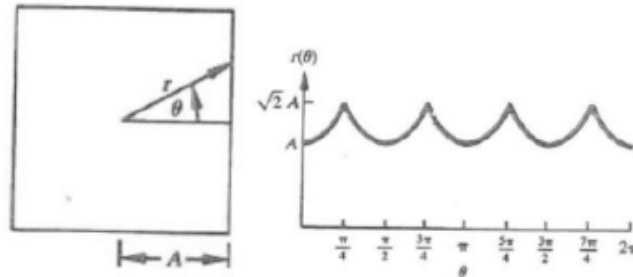


6) Com es calcula la direcció del gradient dels píxels d'una imatge greylevel ?

$\arctg (Gy/Gx)$. Essent Gy i Gx les components (y,x) de gradient.

5) Què és una transformació $r(\varphi)$, també anomenada signatura ? Il·lustra-ho amb una gràfica.

Representa la distància dels píxels de contorn al centroid de l'objecte en funció de l'angle φ .



9) Explica com funciona el binaritzat amb histèresi de l'algorisme de Canny.

Si $f(p) > TH$ llavors $out(p) = 1$

Si $f(p) < TL$ llavors $out(p) = 0$

Per a les regions tals que $TL \leq p \leq TH$

Si regio i connectada a 1 llavors $out(p \in regioi) = 1$

Altrament $out(p \in regioi) = 0$

7. Quina expressió matemàtica s'intenta minimitzar amb l'algoritme

$$\sum_{\text{clusters } i} \sum_{\text{points } p \text{ in cluster } i} \|p - c_i\|^2$$

10) Proposa un algorisme per a implementar la transformada de distància a base d'erosions.

$Im_dist = Im_orig;$ // Im_orig és una imatge binària amb valors (0/1)

Mentre ($Im_orig > 0$) {

$Im_orig = \text{erosiona}(Im_orig)$

$Im_dist = Im_dist + Im_orig$

} // Im_dist és una imatge d'enters que conté el resultat

11) Què és la corba ROC (Receiver Operating Characteristic) i que representen els seus eixos?

Representa el rati de vertaders positius enfront del rati de falsos positius. La sensibilitat enfront de la especificitat

8. Existeix algun color expressat en HSV que no pugui ser convertit a RGB? I a l'inrevés?

Els nivells de gris no tenen representació en HSV (de fet dona divisió per zero al calcular H). Per contra, tot color expressat en HSV pot ser convertit a RGB.

1) En una imatge binària, explica com separaries els blobs que es toquen entre ells.

Fent un watershed sobre la transformada de distància de la imatge.

4) Al fer un open d'una imatge binària, usant un element estructurant de tipus disc de radi 20, observem que els blobs del resultat tenen una forma octogonal. Què en podem deduir d'aquesta implementació de la funció open?

Amb l'objectiu de reduir temps de computació, s'ha aproximat la forma forma circular de l'element estructurant a una forma poligonal. Un element estructurant de forma poligonal es pot descomposar en elements estructurants més petits.

7) Segmentem una imatge color usant l'algorisme k-means en el pla (Hue , Saturació). Sabem que els objectes de la imatge son de 4 colors fàcilment diferenciables, i el fons és fosc. Per aquest motiu fixem k=5. Observem que el resultat ens fragmenta els objectes de color vermell. A que pot ser degut aquest efecte?

La implementació del k-means que s'està usant no té en compte la ciclicitat del hue. El color vermell a quedat partit en 2 clusters, al principi i final del pla.

9) Per a trobar vèrtexos usant l'algorisme de Harris, no cal calcular vectors propis, n'hi ha prou calculant només valors propis. Quin és el motiu ?

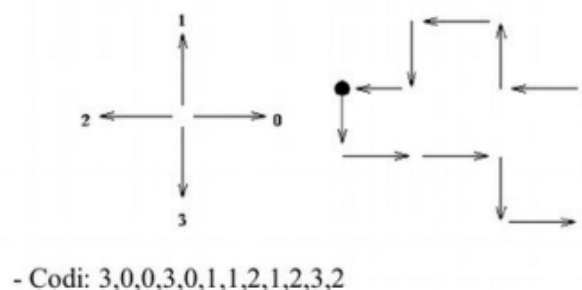
N'hi ha prou amb saber si hi ha una variació important en les dues direccions principals. Aquesta informació ja ens la donen els valors propis. No cal conèixer quines són aquestes direccions principals.

1) Què és el SKIZ (Skeleton by Influence Zone) d'una imatge binària?

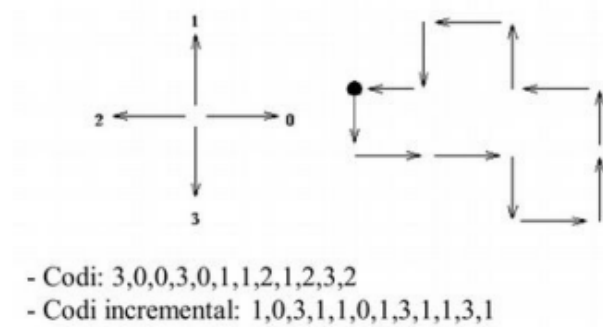
És el conjunt de píxels del fons que equidisten de dues (o més) formes. Ho implementaria mitjançant un watershed sobre la transformada de distància.

2) Com es pot descriure el contorn d'una figura usant codis de cadena?

Assignant un codi numèric a cada segment del contorn en funció de la seva orientació



Fent una codificació incremental en funció de la variació de la orientació



3) Cita els inconvenients de reconèixer objectes mitjançant template matching.

No és invariant a la rotació, ni al canvi d'escala, ni a canvis d'il·luminació, ni de perspectiva, ni a oclusions. Probar tots els templates resultants de totes les possibles transformacions és inviable.

4) Avantatges i inconvenients dels histogrames de color enfront dels HoGs per a descriure un objecte.

Els histogrames de color funcionen bé si les imatges tenen constància de color. Són molt sensibles a la il.luminació. Els histogrames d'orientació són més robustos a canvis d'il.luminació, però no funcionaran bé si l'objecte pateix deformacions.

Que és més convenient: dimensionar els histogrammes d'orientació entre -180° i 180° , o entre 0° i 180° ?. Proposa dos exemples que mostrin quan convé fer-ho d'una forma o de l'altra.

Depèn de l'aplicació. Un gradient amb orientació α , i una gradient amb orientació $\alpha+180^\circ$ tenen la mateixa direcció. Un representa una transició clar--->fosc, mentre que l'altre representa una transició fosc-->clar. Per tant, si desconeixem la lluminositat de l'objecte respecte del fons, és absurd treballar amb signe, doncs el mateix objecte es descriuria diferent sobre un fons clar que sobre un fons fosc.

5) Per a calcular una certa matriu de transformació T , volem usar l'algorisme RANSAC. Calculem el número d'iteracions necessàries usant la fórmula:
 I obtenim que calen $N = 10000$ iteracions per a trobar la matriu T amb els paràmetres corresponents. Malgrat això, executem RANSAC i l'algorisme s'atura a la iteració $N=37$. Com s'explica això?

$$N = \log(1 - p) / \log(1 - (1 - e)^s)$$

Perquè l'algorisme RANSAC recalcula N en cada iteració en funció del nº d'inliers detectats.

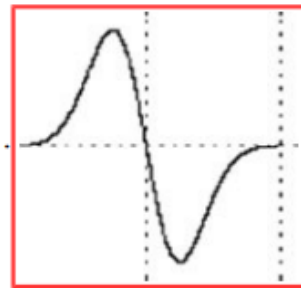
Quina seguretat tenim de que la T trobada sigui la matriu de transformació correcta?

T és la matriu de transformació correcta amb una probabilitat p .

2) (1p.) Què és un Bag of Features? quins avantatges i inconvenients té?

És un descriptor d'una classe. Consisteix en un vector on s'acumulen el número d'ocurrències de cada característica (o conjunt de característiques). Un histograma de color és, per exemple, un BoF. L'avantatge és que em permet comparar ràpidament si dues imatges tenen característiques comunes. L'inconvenient és que es perd la informació de la relació estructural entre aquestes característiques.

3) (1p.) Què és un filtre Derivative of Gaussian? Explica-ho mitjançant un dibuix 1D del filtre. Descriu l'avantatge d'usar aquest filtre en comptes de fer un filtratge gaussià pur seguit d'un filtratge derivatiu pur

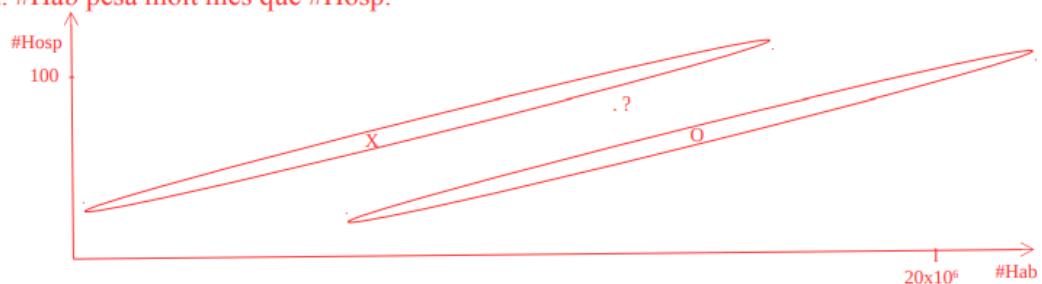


Usar un filtre gaussià pur, i al resultat passar-li un filtre derivatiu suposaria haver de convolucionar la imatge dues vegades. Gracies a la propietat associativa, el resultat és equivalent a convolucionar la imatge un sol cop usant un filtre Derivative of Gaussian.

5) (1'5p.) En un espai de 2 característiques, tenim mostres etiquetades de les classes 'X' i 'O'. Volem classificar una nova mostra sense etiquetar '?', com a 'X' o 'O', en funció de la seva distància al centre d'ambdues classes en aquest espai. Podria donar-se el cas de que la mostra '?' estigués més propera a la classe 'X' usant distància euclídea, però més propera a la classe 'O' usant distància euclídea normalitzada? Respon inventant un exemple (amb característiques creïbles). Dibuixa l'espai de característiques 2D il·lustrant l'exemple.

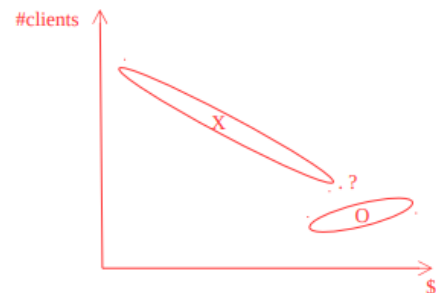
Si. En el cas de que una de les característiques tingui una variància significativament més gran que l'altra. P.ex: volem classificar ciutats en funció del seu número d'habitants i hospitals.

La mostra '?' seria classificada com a 'X' si no normalitzem la distància, i com a 'O' si la normalitzem. #Hab pesa molt més que #Hosp.



Podria donar-se el cas de que la mostra '?' estigués més propera a la classe 'X' usant distància euclídea normalitzada, però més propera a la classe 'O' usant distància de Mahalanobis ? Respon inventant un exemple (amb característiques creïbles). Dibuixa l'espai de característiques 2D il·lustrant l'exemple.

Si. En el cas de que les característiques de la classe estiguin molt correlades. P.ex: volem classificar 2 productes en funció del preu i el numero de compradors. La classe 'X' representa un producte que segueix la llei oferta-demanda, mentre que la classe 'O' és un producte de luxe que quan més car és, més es ven.



7) (1p.) Per a dividir per 4 la dimensió d'una imatge, un enginyer de telecomunicacions decideix quedar-se amb els píxels de les files i columnes parells, i descartar els de les senars. Justifica si et sembla una bona solució, o si podria portar problemes en algun cas.

No. S'està vulnerant el criteri de Nyquist, i això provocarà aliasing. Abans de submostrejar, s'ha de filtrar la imatge.

8) (1p.) Defineix que volen dir les components hue i saturation de l'espai HLS. Quins inconvenients té l'espai HLS respecte l'espai RGB?

Hue i saturació són les components cromàtiques de la imatge. El hue és el matís del color. La saturació és la puresa d'aquest color.

Un possible inconvenient és la indefinició del hue quan la saturació val zero, (o l'error en el hue si la saturació és molt baixa). Un altre possible inconvenient és que el hue és un angle, per tant, cal treballar sempre usant operacions cícliques.

1) En un projecte de reconeixement mitjançant visió per computador, es dona el cas de que els punts d'interès trobats en una imatge tenen uns descriptors SIFT que són molt similars als del model de l'objecte que estem buscant. S'en poden aparellar molts amb un grau molt alt de similitut. Si es donés aquest cas, podríem afirmar que l'objecte de la imatge és el del model sense utilitzar el RANSAC? (utilitzant només l'alt nombre de keypoints aparellats)

Només podem afirmar que tots dos objectes tenen moltes característiques similars

En quins casos podríem tenir resultats diferents fent-ho amb RANSAC o fent-ho directament per aparellament?

El reconeixement per simple aparellament no funciona quan és important l'estructura entre els keypoints trobats. En aquests casos cal usar RANSAC o algún algorisme de matching global. Per exemple, les expressions «VIOLENCE» i «NICE LOVE» tenen totes les lletres en comú, però volen dir just el contrari.

2) Si el vector de característiques te dimensionalitat 1000000, el numero de vectors propis de la matriu de covariancia és també 1000000. És a dir, la dimensionalitat és la mateixa. Llavors, quina és la utilitat de l'anàlisi de components principals? De quina manera permet reduir la dimensionalitat de l'espai de característiques?

Si les dades presenten alguna correlació, molts d'aquests vectors propis trobats seran irrelevants, doncs la variació de les dades en aquesta direcció serà mínima. Podrem eliminar tots aquells vectors propis amb valors propis petits, i quedar-nos només amb aquells més grans que són els que expliquen millor la variació de les dades.

3) Cita un descriptor de regions que sigui invariant a la

traslació: àrea

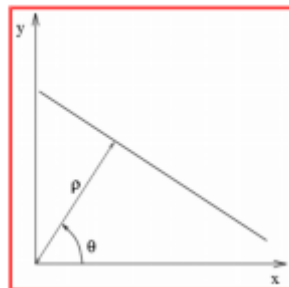
rotació: àrea

escala: compacitat

oclusió d'un 50%: Les oclusions distorsionen la descripció global. Cal usar descriptors locals, per exemple, SIFT

4) Per determinar si un conjunt de píxels estan aliniats, podem usar la transformada de Hough mapejant les coordenades (x,y) de cada pixel en la taula (m,n). Les posicions de la taula (m,n) venen donades per l'equació de la recta $y = mx + n$. Malgrat això, els programadors s'estimen més usar l'equació de la recta en forma polar $r = x \cos \alpha + y \sin \alpha$, i treballen amb la taula (r, α). Per quin motiu?

m i n no estan afitades. La pendent, per exemple, pot valdre infinit. Això complica dimensionar la taula.



5) Un estudiant calcula l'histograma de color d'un objecte model per a buscar-lo en una seqüència d'imatges. En comptes d'implementar-lo usant una matriu tri-dimensional RGB de dimensió $N \times N \times N$, ho implementa amb 3 vectors uni-dimensionals R, G, i B (de dimensió N cadascun). Fent-ho d'aquesta manera, ha de manejar una estructura de dades més senzilla, estalviant espai de memòria. Quin inconvenient pot presentar aquesta solució? Il·lustra-ho amb un exemple on la solució proposada per l'estudiant detecti un objecte incorrecte.

Al incrementar els vectors R, G, i B per separat, perdem la informació relativa a la relació entre les 3 components de color. Aquest estudiant, amb la solució proposada, confondria la bandera d'Arzebadjan amb la samarreta del Madrid