

Nom i Cognoms: _____ **Una possible solució** _____

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, podriem 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 podriem 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 expresions «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 covariància 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 píxel en la taula (m,n) . Les posicions de la taula (m,n) venen donades per l'equació de la recta $y = mx + n$.

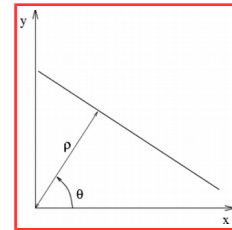
Malgrat això, els programadors s'estimen més usar l'equació de la recta en forma polar

$r = x \cos \alpha + y \sin \alpha$, 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.

Que volen dir r i α ? (il·lustra-ho gràficament)



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

