TD TERI

Master Informatique - IAD et IMA

22 octobre 2010

1 Reconnaissance des formes

1.1 Classification bayésienne

On s'intéresse à la détection des cultures avec le satellite SPOT et le satellite Landsat.

Deux cultures nous intéressent : le blé d'une part (classe C_1) qui couvre 60 % des terres de la région d'intérêt, les légumes (classe C_2) qui couvrent 10% (le reste est occupé par des cultures non identifiées et des aménagements du sol).

1. On a détecté une zone $\mathcal{Z}$ très grande et très homogène dont le niveau moyen dans l'image SPOT est $n = 80$.
   Après apprentissage des probabilités conditionnelles des niveaux de gris en fonction de la classe ($P(n|C_1)$ et $P(n|C_2)$), on a déterminé que, pour le satellite SPOT, les probabilités du niveau de gris pour le blé et les légumes suivent des lois gaussiennes de paramètres :
   blé : moyenne $= m_1 = 100$ écart type $= \sigma_1 = 20$
   légumes : moyenne $= m_2 = 85$ écart type $= \sigma_2 = 5$
   Si l'on prend une décision selon le critère du maximum a posteriori (MAP), à quelle classe doit-on attribuer la zone ?
2. En fait on n'est pas très sûr de la probabilité a priori des légumes qui selon les années varie de 8% à 20% des terres. Est-ce que cela peut changer notre décision ?
3. Quelle classification obtient-on si on utilise un critère de maximum de vraisemblance (MV) ?

1.2 Classification automatique

On dispose d'une image radar du port d'Osaka et l'on souhaite la classifier à partir de la valeur de niveau de gris des pixels. A cause de la nature de l'image (bruit de speckle), on va choisir deux paramètres pour chaque pixel : le premier est calculé à partir de la valeur moyenne et de la valeur maximale à l'intérieur d'un voisinage, le second dépend de la variance locale.

Afin de restreindre le problème, on va se focaliser sur 6 "prototypes" : les points A, B, C, D, E, F qui ont pour coordonnées dans le plan des paramètres :

$$A = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \quad B = \begin{pmatrix} 4 \\ 3 \end{pmatrix} \quad C = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \quad D = \begin{pmatrix} 3 \\ 4 \end{pmatrix} \quad E = \begin{pmatrix} 7 \\ 3 \end{pmatrix} \quad F = \begin{pmatrix} 3 \\ 7,0001 \end{pmatrix}$$

Sur chacun de ces 6 points, on place 1 échantillon, sauf en C où on en place 2 et en A où on en place 4. On a donc au total 10 échantillons.

On cherche à construire un classifieur de ces 10 échantillons.

1. En utilisant l'algorithme des k-moyennes, et en supposant que l'on cherche un classifieur en 2 classes, montrer que l'algorithme peut converger vers (au moins) deux solutions. Pour cela, on suggère de traiter les deux cas suivants :
   - initialiser la classe 1 par A et la classe 2 par F
   - initialiser la classe 1 par E et la classe 2 par F

   Pour simplifier les calculs, on peut choisir la norme L^1 $(d(x, y) = |x - y|)$.
2. Quel classifieur choisirez vous en pratique ? Sur quel principe fondamental vous justifierez vous ?

2 Modèle statistique

La figure 1 représente une image en niveaux de gris. Le visage de Mickey est symétrique et les valeurs de niveaux de gris sont indiquées par des flèches. Les dimensions totales de l'image sont de 32×32 pixels carrés.

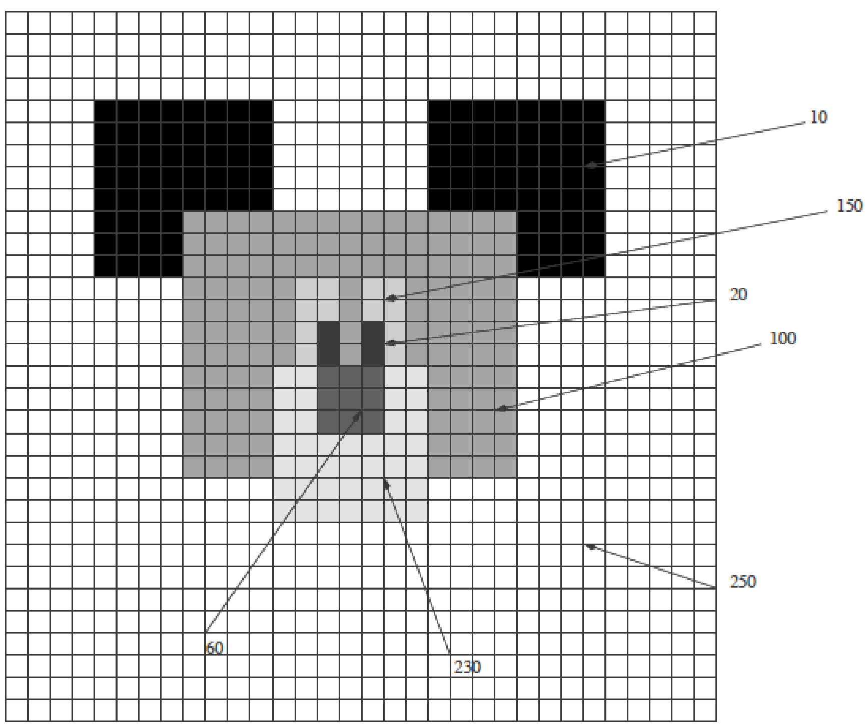

FIGURE 1 – Mickey Mouse.

1. Tracer l'histogramme de l'image. Combien y a-t-il de niveaux de gris ? Quelle est l'entropie de l'image ?
2. Calculer une normalisation d'histogramme sur cette image. Calculer une égalisation d'histogramme sur cette image. Que devient l'entropie dans les nouvelles images ?

3 Normalisation d'histogrammes

Soit I une image numérique de largeur w et de hauteur h, à valeur dans $\{0, 255\}$.

1. Écrivez (pseudo-code) une fonction qui réalise une modification des valeurs de I correspondant à une normalisation d'histogramme, c'est-à-dire une fonction prenant I en entrée, fournissant l'image J en sortie, et réalisant une transformation affine des valeurs de niveaux de gris, telle que la valeur minimum de J soit 0 et la valeur maximum de J soit 255.
2. Modifiez la fonction précédente, en ajoutant comme argument d'entrée un réel $p \in [0, 1]$, de telle sorte que tout pixel de I ayant une valeur k telle que la proportion dans I des pixels ayant une valeur inférieure à k soit inférieure à p (resp. supérieure à p) prenne la valeur 0 (resp. la valeur 255) dans J.
3. Quel effet sur l'image auront les fonctions précédentes ? Quel est l'intérêt de la deuxième fonction par rapport à la première ?

4 Opérations sur les histogrammes

Soit une image f et son histogramme H_f . Quel effet ont sur l'histogramme les opérations suivantes (décrivez H_g par rapport à H_f) ?

1. Ajout d'une constante : $g(x) = f(x) + c$.
2. Inversion : $g(x) = N_{\max} - f(x)$.
3. Quantification : $g(x) = E\left(\frac{f(x)}{N}\right)$ (E désigne la partie entière).
4. Lissage par un noyau de convolution h de somme égale à 1 : $g(x) = (f \star h)(x)$

5 Filtre de moyenne et filtre médian

Dans cet exercice on ne s'intéressera qu'au carré 3×3 du centre de l'image ci-dessous.

10	10	10	10	10
10	10	10	10	10
10	10	0	0	0
10	10	0	0	0
10	10	0	0	0

1. Décrivez le filtre médian sur une fenêtre 3×3. Quel est son avantage sur le filtre par moyenne ?
2. Donnez le résultat du filtrage médian 3×3 et du filtrage de moyenne 3×3 sur l'image ci-dessus.
3. Que concluez-vous sur l'effet d'un filtrage médian ?

6 Transformée de Hough

La transformée de Hough permet de détecter des formes spécifiques dans les images. La contrainte imposée par cette méthode est que la forme recherchée soit paramétrisable, ce qui rend la transformée de Hough intéressante pour la détection des formes simples telles que des droites, cercles ou ellipses. L'avantage majeur de la transformée de Hough est qu'elle tolère bien les discontinuités dans les contours des formes recherchées, aussi bien que le bruit des images.

Par la suite, on s'intéresse à la détection de droites en utilisant la transformée de Hough. Pour la paramétrisation des droites, on adopte la solution suivante :

$$x \cos \theta + y \sin \theta = r$$

oà r est la distance entre l'origine et la droite, et θ donne l'orientation de r par raport à l'axe Ox (voir Fig. 2). Pour un point donné, x et y sont les constantes, alors que (r, θ) sont les variables associées à une des droites qui passe par (x, y).

1. Dans une image, quels sont les points intéressants pour la détection de droites ?
2. Dans l'espace des variables (r, θ), qu'on appelle accumulateur, quelle est la forme générée par toutes les droites qui passent par un point (x, y) donné ? Donnez un exemple du domaine de définition pour ces variables.
3. Considérons deux points (x_1, y_1) et (x_2, y_2). Dans l'espace (r, θ) quelle est la représentation de la droite qui les relie ? Même question s'il s'agit de n points colinéaires.
4. À partir d'une paire (r, θ), comment peut-on dessiner dans l'image initiale la droite qui lui correspond ?
5. Interprétez l'image de la figure 3 et son accumulateur correspondant, figure 4, (θ se trouve sur l'axe horizontal).
6. Écrire un algorithme qui prend en compte les remarques antérieures afin de détecter les droites dans une image donnée.
7. Quels changements sont nécessaires pour que l'algorithme détecte cette fois-ci des cercles ? Des ellipses ?
8. Si on considère que chaque paramètre de l'accumulateur peut prendre 100 valeurs différentes, quelle sera la taille de la mémoire demandée pour une détection de droites ? de cercles ? d'ellipses ?

7 Segmentation par Quadtree

1. Rappeler l'algorithme de segmentation par Quadtree.
2. Illustrer le fonctionnement de l'algorithme sur l'image 64×64 suivante (chaque élément de la grille représente un carré 8×8). On construira au fur et à mesure l'arbre 4-aire.
3. Quel prédicat de découpage peut-on utiliser pour segmenter correctement cette image ?
4. Calculer les moyennes et variances de chaque feuille de l'arbre (les niveaux de gris du fond, du carré et du rectangle sont respectivement à 150, 50 et 100).

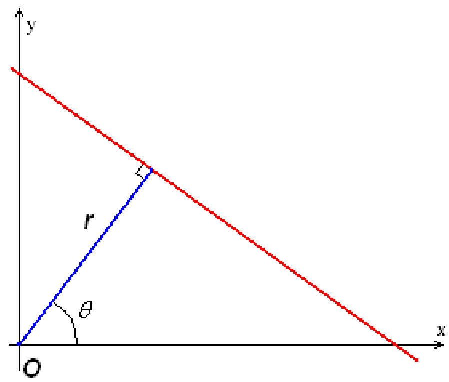

Figure 2 – Droite en coordonnées polaires (r, θ)

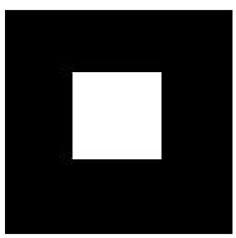

Figure 3 – Image initiale

Figure 4 – Accumulateur