

Formation et Analyse d'Images

James L. Crowley

ENSIMAG 3

Premier Sémestre 2006/2007

Séance 8

17 novembre 2006

Analyse et Réconnaissance Statistique

Plan de la Séance :

Classification des pixels par ratio d'histogramme	2
Histogrammes	2
Détection par ratio d'histogramme	3
Histogrammes de Champs Réceptifs.....	5
Caractérisation d'un région par moments.....	6
Composantes principales.....	7
La Classification des Formes	8
Les Caractéristiques.....	9
Les Observations.....	10
Fonctions de Discrimination.....	11

Classification des Pixels par Ratio d'Histogramme

Histogrammes

Un histogramme est une table de fréquence. Il peut fournir une estimation d'une densité de probabilités. Pour x entier, tel que $x \in [x_{\min}, x_{\max}]$, on peut traiter chacun des valeurs possibles comme une classe d'événement.

Si les valeurs de x sont entières, tel que $x \in [x_{\min}, x_{\max}]$ on peut estimer la probabilité à partir de M observations de la valeur, $\{X_m\}$.

Pour estimer la probabilité d'une valeur on peut compter le nombre d'observation de chaque valeur, x , dans une table, $h(x)$.

L'existence des ordinateurs avec des centaines de megabytes rendre des tables de fréquence très pratique pour la mise en œuvre en temps réel des algorithmes de reconnaissance. Dans certains domaines, comme l'analyse d'images, par abus de langage, un tel table s'appelle une histogramme. Proprement dit, l'histogramme est une représentation graphique de $h(x)$

Ainsi la probabilité d'une valeur de $X \in [X_{\min}, X_{\max}]$ est la fréquence d'occurrence de la valeur. Avec M observations de la valeur, X , on peut faire une table, $h(x)$, de fréquence pour chacun des valeurs possibles. On observe M exemples de X , $\{X_m\}$.

Pour chaque observation on ajoute "1" à son entrée dans la table.

$$m=1, M : h(X_m) := h(X_m) + 1; M := M+1;$$

$h(x)$ est une table de fréquence pour chaque $x \in [x_{\min}, x_{\max}]$. Ainsi, on peut définir la probabilité d'une valeur x par sa fréquence :

$$p(X_m=x) = \lim_M \left\{ \frac{1}{M} h(x) \right\}$$

Quand M est fini, on peut faire appel à l'approximation.

$$P(X=x) \approx \frac{1}{M} h(x)$$

La validité de l'approximation dépend du nombre de valeurs possible et de M .

L'erreur moyenne entre $\frac{1}{M} h(C)$ et $P(C)$ est $E_{ms} \sim O\left(\frac{Q}{M}\right)$

Pour que l'estimation soit "raisonnable", il faut assuré que $M \gg Q = N^2$

En règle générale, on dit qu'il faut 10 exemples par cellule de l'histogramme.

Détection par ratio d'histogramme

On peut utiliser les histogrammes avec la règle de Bayes pour détecter les objets.

Par exemple, construisons une histogramme pour le vecteur de chrominance (r,v) .

La chrominance $C = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ est une signature pour l'objet.

La chrominance peut être définie par plusieurs codages. Par exemple, pour la détection du peau, il est fréquent de voir

$$c_1 = r = \frac{R}{R+V+B} \quad c_2 = v = \frac{V}{R+V+B}$$

Une autre codage fréquente est un codage en "couleur opposée", par exemple :

$$L = \frac{R+G+B}{3}, \quad c_1 = \frac{R-G}{2} \quad \text{et} \quad c_2 = \frac{R+G}{2} - B.$$

$$\begin{array}{rcccc} L & 0.33 & 0.33 & 0.33 & R \\ C_1 & = & 0.5 & -0.5 & 0 & G \\ C_2 & & 0.5 & 0.5 & -1 & B \end{array}$$

Chaque pixel est un vecteur (r, v) : $C(i, j) = \begin{pmatrix} r \\ v \end{pmatrix} (i, j)$

Supposons qu'on code c_1 et c_2 avec les entiers entre 0 et $N - 1$

$$c_1 = \text{Round} \left((N-1) \cdot \frac{R}{R+G+B} \right) \quad c_2 = \text{Round} \left((N-1) \cdot \frac{G}{R+G+B} \right)$$

On alloue un tableau 2D, $h(c_1, c_2)$, de taille $N \times N$ cellules.
(exemple $Q = 32 \times 32 = 1024$ cellules)

Pour chaque pixel $C = C(i, j)$ dans l'image, on incrémente la cellule de l'histogramme qui correspond à C : $h(C) := h(C) + 1$

$$\text{c'a-dire} \quad h(c_1, c_2) := h(c_1, c_2) + 1$$

Soit M Pixels dans l'image. Un histogramme des chrominance, $h(C)$, des M pixels dans une l'image donne leurs fréquences d'occurrence.

$$P(C) = \frac{1}{M} h(C)$$

Considère une région W de M_o pixels du même image correspondance à l'objet O .

$$(i,j) \in W : h_o(C(i,j)) := h_o(C(i,j)) + 1$$

$$\text{Ensuite: pour tout pixel } C(i, j) = \begin{matrix} r \\ v \end{matrix} (i, j) : p(C | \text{objet}) = \frac{1}{M_o} h_o(C)$$

Parce que W est dans l'image, la probabilité de rencontrer un pixel de W ,

$$P(W) = \frac{M_o}{M}$$

L histogramme permet d'utiliser la règle de Bayes afin de calculer la probabilité qu'un pixel corresponde à un objet.

$$\text{Pour chaque pixel } C(i, j) \quad p(\text{objet} | C) = p(C | \text{objet}) \frac{p(\text{objet})}{p(C)}$$

Soit M images de $I \times J$ pixels. Ceci fait $N = I \times J \times M$ Pixels.

Soit $h(r, v)$, l'histogramme de tous les N pixels.

Soit $h_o(r, v)$, l'histogramme des N_o pixels de l'objet "o".

$$p(\text{objet}) = \frac{M_o}{M}$$

$$p(C) = \frac{1}{M} h(C)$$

$$p(C | \text{objet}) = \frac{1}{M_o} h_o(C)$$

$$p(\text{objet} | C) = p(C | \text{objet}) \frac{p(\text{objet})}{p(C)} = \frac{1}{M_o} h_o(C) \frac{\frac{M_o}{M}}{\frac{1}{M} h(C)} = \frac{h_o(C)}{h(C)}$$

Par exemple, voici une image de la probabilité de peau fait par ratio d'histogramme de r, v



Histogrammes de Champs Réceptifs.

Cette méthode peut être généralisée en remplaçant la chrominance par un vecteur de champs réceptifs. Mais il faut bien gérer la relation Q/M !

Soit une image $p(i,j)$ et un vecteur de "d" champs réceptifs G

$V(i,j) = \langle G, p(i,j) \rangle$ est un vecteur de caractéristiques de d dimensions.

$h(V)$ aura $Q = N^d$

N= \ d=	1	2	3	4	5	6
2	2^1	2^2	2^3	2^4	2^5	2^6
4	2^2	2^4	2^6	2^8	$2^{10} = 1 \text{ Kilo}$	$2^{12} = 2 \text{ Kil}$
8	2^3	2^6	2^9	2^{12}	2^{15}	2^{18}
16	2^4	2^8	2^{12}	2^{16}	$2^{20} = 1 \text{ Meg}$	$2^{24} = 4 \text{ M}$
32	2^5	$2^{10} = 1 \text{ Kilo}$	2^{15}	$2^{20} = 1 \text{ Meg}$	2^{25}	$2^{30} = 1 \text{ G}$
64	2^6	2^{12}	2^{18}	2^{24}	$2^{30} = 1 \text{ Gig}$	2^{36}
128	2^7	2^{14}	$2^{21} = 2 \text{ Meg}$	2^{28}	2^{35}	$2^{42} = 2 \text{ Te}$
256	2^8	2^{16}	2^{24}	$2^{32} = 2 \text{ Gig}$	$2^{40} = 1 \text{ Tera}$	2^{48}

Soit les champs réceptifs achromatique

$$G = (G_x, G_y, G_{xx}, G_{xy}, G_{yy})$$

$$d = 5$$

ou chromatique avec normalisation de l'orientation et échelle :

$$G_c = (G_x^L, G_x^{C1}, G_x^{C2}, G_{xx}^L, G_{xy}^L, G_{xx}^{C1}, G_{xx}^{C2})$$

$$d = 7.$$

On peut faire

$$p(\text{objet}(i,j) | V(i,j)) = \frac{p(V(i,j) | \text{objet}(i,j)) p(\text{objet}(i,j))}{p(V(i,j))} \quad \frac{h_o(V(i,j))}{h_{\text{tot}}(V(i,j))}$$

sur condition de gérer M et Q.

Caractérisation d'un région par moments

Les ensemble connexes de pixels s'appelles les "blobs".

On peut décrire une blob par une vecteur de caractéristiques "invariantes" à l'orientation grâce aux "moments"

Les moments sont invariants aux transformations affines.

Pour une febre (imagerie) $w(i, j)$ de taille $N \times M$

$$\text{Somme des Pixels :} \quad S = \sum_{i=1}^M \sum_{j=1}^N w(i, j)$$

Premiers moments :

$$\mu_i = \frac{1}{S} \sum_{i=1}^M \sum_{j=1}^N w(i, j) \cdot i \quad \mu_j = \frac{1}{S} \sum_{i=1}^M \sum_{j=1}^N w(i, j) \cdot j$$

Le premier moment est le centre de gravité de la forme :

Deuxième moment :

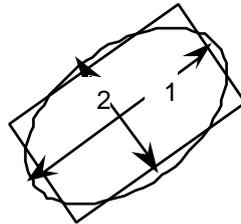
$$i^2 = \frac{1}{S} \sum_{i=1}^M \sum_{j=1}^N (w(i, j)) \cdot (i - \mu_i)^2$$

$$j^2 = \frac{1}{S} \sum_{i=1}^M \sum_{j=1}^N w(i, j) \cdot (j - \mu_j)^2$$

$$ij^2 = \frac{1}{S} \sum_{i=1}^M \sum_{j=1}^N w(i, j) \cdot (i - \mu_i)(j - \mu_j)$$

Ceci permet de définir les "axes", majeur, μ_1 et mineur, μ_2 , de la forme par analyse des composantes principales de la deuxième moment

$$C_o \cong \begin{pmatrix} i^2 & ij^2 \\ ij^2 & j^2 \end{pmatrix}$$

Composantes principales

Les deuxièmes moments sont "invariants" à l'orientation

Les axes sont calculés par une analyse en composantes principales de la matrice C . Il s'agit de trouver une rotation, Φ , dans l'espace de caractéristiques $\Phi C_P \Phi^T = \Lambda$ telles que Λ soit diagonale.

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{tel que } 1 > 2 \quad \Phi = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

$$\Phi C_P \Phi^T = \Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \Phi^T \Phi = \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Phi C_P \Phi^T \Phi = \Phi C_P = \Lambda \Phi = \begin{pmatrix} 1 & 0 & \cos(\theta) & \sin(\theta) \\ 0 & 2 & -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

Les lignes du Φ sont des vecteurs propres du C .

La longueur des axes majeurs et mineur est les valeurs propres de la matrice C .

θ est l'orientation de l'axe "majeur" et $1/2$ est le rapport entre la longueur et la largeur.

$1/2$ est une caractéristique invariante de la taille et de l'orientation.

La Classification des Formes

La classification est une capacité fondamentale de l'intelligence.

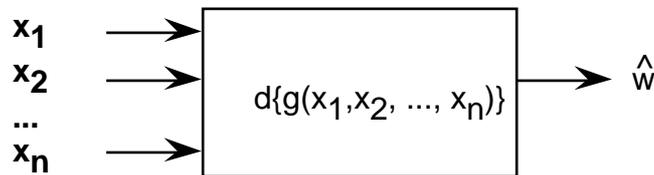
Classer : Reconnaître un membre d'une catégorie, ou d'une classe.

On peut distinguer "reconnaissance" et "identification".

Reconnaissance : Le fait de reconnaître, d'identifier un objet, un être comme tel.

Identifier : Reconnaître un individu

La classification est un processus d'association d'une observation à une classe par un teste d'appartenance.



Pour un vecteur de caractéristique il sort une estimation de la classe, \hat{w}

Les techniques de reconnaissance de formes statistiques fournissent une méthode pour induire des tests d'appartenance à partir d'un ensemble d'échantillons.

La classification se résume à une division de l'espace de caractéristique en partition disjoint. Cette division peut-être fait par estimation de fonctions paramétrique ou par une liste exhaustives des frontières.

Le critère est la probabilité d'appartenance.

Cette probabilité est fournie par la règle de Bayes.

$$p(k | X) = \frac{p(X | k) p(k)}{p(X)}$$

Les Caractéristiques

Forme n. f. : A. Apparence, aspect visible. 1) ... 2) apparence extérieure donnant à un objet ou à un être sa spécificité.

Les méthodes statistique de la reconnaissance de forme traite les observations sous forme de vecteur de caractéristiques.

Caractéristiques : (En anglais : Feature) Signes ou ensembles de signes distinctifs.
Une ensemble de propriétés. $\{ x_1, x_2 \dots x_D \}$.

En notation vectorielle :

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_D \end{pmatrix}$$

Par exemple, $X = \begin{pmatrix} \mu_i \\ \mu_j \\ 1 \\ 2 \end{pmatrix}$ est une vecteur de caractéristique pour les "blobs".

La formation des vrais objets physiques est sujette aux influences aléatoires. Pour les objets d'une classe, w_k , les propriétés des objets individuels sont, les valeurs aléatoires. On peut resume ceci par une somme d'une forme "intrinsèque" x plus ces influences aléatoires individuelles, B_i .

$$X = x + B_i$$

Les techniques probabiliste de RF suppose un bruit additif.

En notation vectorielle :

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} + \begin{pmatrix} B_1 \\ B_2 \\ \dots \\ B_n \end{pmatrix}$$

Les Observations

Une observation : une constatation attentive des phénomènes.

Pour des machines, des observations sont fournies par les capteurs.

Ceci donne une observation (un phénomène) sous forme d'une ensemble de caractéristiques : $\{Y_1, Y_2 \dots Y_D\}$.

$$Y = \begin{matrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{matrix}$$

Les observations sont corrompues par un bruit, B_o .

$$Y = y + B_o$$

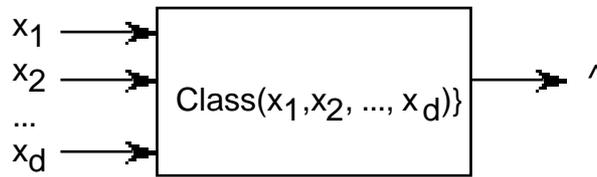
Le bruit est, par définition, imprévisible. Il est aléatoire.

Donc les caractéristiques observées sont des vecteurs aléatoires.

La corruption des observations par un bruit aléatoire est fondamentale aux capteurs physiques.

Fonctions de Discrimination

La classification est un processus d'estimation de l'appartenance d'un événement à une des classes A_k fondée sur les caractéristiques de l'événement, X .



$$\hat{k} = \text{Classifier}(E) = \text{Decider}(E, k)$$

\hat{k} est la proposition que $(E \in k)$.

La fonction de classification est composée de deux parties $d()$ et $g_k()$:

$$\hat{k} = d(g(X)).$$

$g(X)$: Une fonction de discrimination : $\mathbb{R}^D \rightarrow \mathbb{R}^K$

$d()$: Une fonction de décision : $\mathbb{R}^K \rightarrow \{1, \dots, K\}$

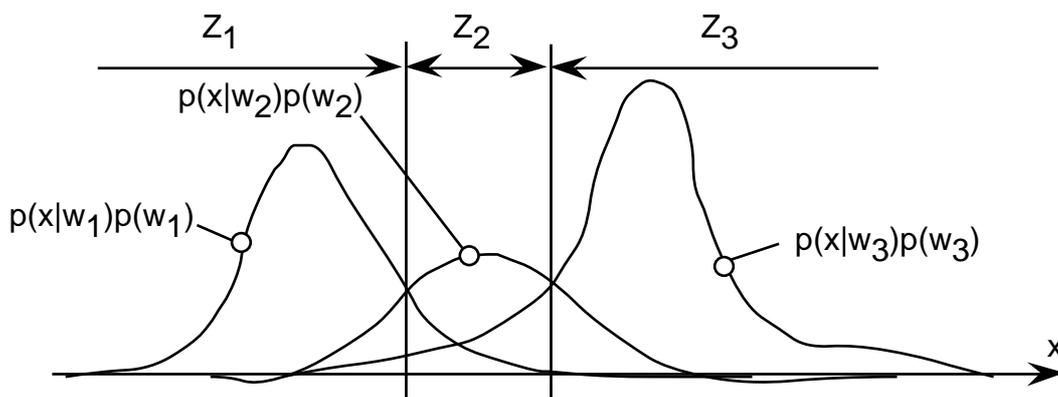
en générale $g(X) = \begin{pmatrix} g_1(X) \\ g_2(X) \\ \dots \\ g_K(X) \end{pmatrix}$ est un vecteur de K fonctions $g_k(X)$

Dans le cas général, $K > 2$, le nombre minimum d'erreurs est fait si k est choisi tel que :

$$k = \arg\text{-max}_k \{g_k(X)\} \quad \text{avec } g_k(X) = \frac{p(x | k)}{p(x)}$$

Les frontières entre régions i et j sont les valeurs pour lesquelles

$$g_i(X) = g_j(X)$$



Une fonction de discrimination partition l'espace de caractéristique en régions disjointes Z_1, \dots, Z_k pour chaque classe.

$$k = \underset{k}{\operatorname{arg-max}} \{g_k(X)\}$$

Mais comment calculer $g_k(X)$?

Les caractéristiques X de l'événement E sont aléatoires avec une dispersion due aux variations naturelles de sa classe.

Ceci est modélisé par une variable aléatoire B_k autour d'une valeur "type" x_k . La valeur type est spécifique à la classe.

$$X = x_k + B_k$$

Si $D=1$, les membres de la classe k auront les caractéristiques X tel que :

$$p(X=x | k) = \mathcal{N}(x; \mu_k, \sigma_k^2) = \frac{1}{\sqrt{2\pi} \sigma_k} e^{-\frac{(x-\mu_k)^2}{2\sigma_k^2}}$$

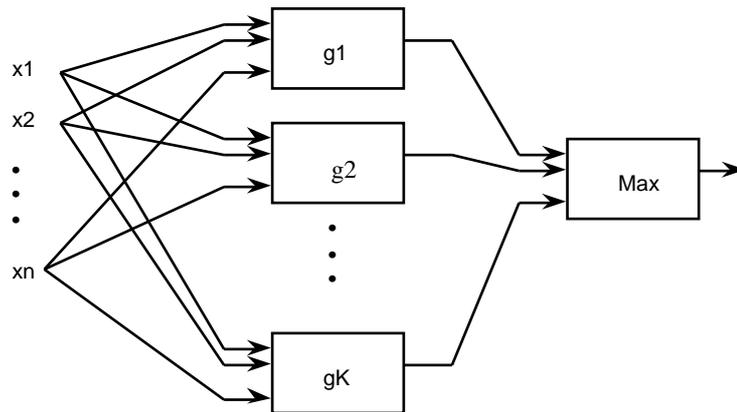
Donc notre fonction de discrimination devient :

$$g_k(X) = p(x | k) \frac{1}{\sqrt{2\pi} \sigma_k} e^{-\frac{(x-\mu_k)^2}{2\sigma_k^2}}$$

pour $D > 1$ il faut la fonction normales multi-variante

$$p(X) = \mathcal{N}(X; \mu, C_x) = \frac{1}{(2\pi)^{\frac{D}{2}} \det(C_x)^{\frac{1}{2}}} e^{-\frac{1}{2}(X - \mu)^T C_x^{-1} (X - \mu)}$$

La classificateur est une machine qui calcule K fonctions $g_k(x)$ suivie d'une sélection du maximum.



Soit $D = 1$. (une seul caractéristique).

On peut noter que $k = \arg\text{-max}_k \{g_k(X)\} = \arg\text{-max}_k \{\text{Log}\{g_k(X)\}\}$ parce que $\text{Log}\{\}$ est une fonction monotone.

$$k = \arg\text{-max}_k \left\{ \text{Log} \left\{ \frac{1}{\sqrt{2\pi} \sigma_k} e^{-\frac{(x-\mu_k)^2}{2\sigma_k^2}} \right\} + \text{Log}\{p(k)\} \right\}$$

ou $k = \arg\text{-max}_k \left\{ -\text{Log}\left\{ \frac{1}{\sqrt{2\pi} \sigma_k} e^{-\frac{(x-\mu_k)^2}{2\sigma_k^2}} \right\} + \text{Log}\{p(k)\} \right\}$

Dans le cas générale $D > 1$ La fonction de discrimination devient :

$$g_k(x) = -\frac{1}{2} \text{Log}\{\det(C_k)\} - \frac{1}{2}(X - \mu_k)^T C_k^{-1} (X - \mu_k) + \text{Log}\{p(k)\}$$

Ceci peut être traduit dans une forme canonique :

$$g_k(X) = X^T (D_k) X + d_k^T X + d_{ko}.$$

avec une terme 2^{ieme} ordre : $D_k = \frac{1}{2} C_k^{-1}$

une terme 1^{iere} ordre : $d_k = C_k^{-1} \mu_k$

Constant : $d_{ko} = -\frac{1}{2}(\mu_k^T C_k^{-1} \mu_k) - \frac{1}{2} \text{Log}\{\det(C_k)\} + \text{Log}\{p(k)\}$