

Formation et Analyse d'Images

James L. Crowley

ENSIMAG 3

Premier Séestre 2009/2010

Séance 10

4 janvier 2010

Analyse et Réconnaissance Statistique

Plan de la Séance :

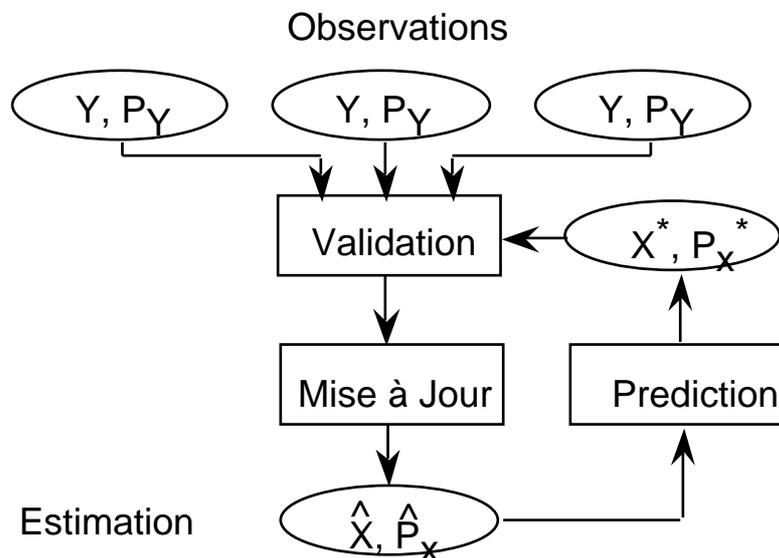
Détection et Suivi des Blobs Gaussiens.....	2
L'architecture d'un système de suivi robuste.....	2
Les Blobs Gaussiens.....	2
Détection : Classification et Groupement de pixels.....	3
Exemple : Détection de visage par ratio d'histogramme de couleur	4
Caractérisation d'un région par moments.....	6
Composantes principales.....	7
Histogrammes de Champs Réceptifs.....	8
La Classification Bayésienne	10
Fonctions de Discrimination	12

Notations

x	Un vecteur
X	Un vecteur aléatoire (non-prévisible).
D	Nombre de dimensions de X
T_k	La classe k
k	Indice d'une classe
K	Nombre de classes
M_k	Nombre d'exemples de la classe k .
M	Nombre totale d'exemples de toutes les classes
E_k	L'affirmation que l'événement $E = T_k$
$p_k = p(E = T_k)$	Probabilité de rencontrer un membre de la classe k .
X	Une observation (un vecteur aléatoire).
$P(X)$	Fonction de densité pour la Probabilité d'une observation X

Détection et Suivi des Blobs Gaussiens

L'architecture d'un système de suivi robuste



Les Blobs Gaussiens

Les entités suivies sont une ensemble connexes de pixels sont de cibles, parfois appelées des "blobs".

On peut décrire une blob par un vecteur de caractéristiques "invariantes" à l'orientation grâce aux "moments" d'est pixels sorties d'une procédure de détection appliqué à une "ROI" (Region of Interest).

Le ROI est une rectangle englobante représentée par quatre paramètres : (u, l, b, r)

- u - "up" - le premier ligne du ROI.
- l - "left" - la première colonne du ROI.
- b - "bottom" - le dernier ligne du ROI
- r - "right" - La dernière colonne du ROI

Les procédures de détection fréquemment utilisé sont :

- 1) Ratio d'Histogramme de couleurs.
- 2) Différences d'image avec un fond adaptatif.
- 3) Différence d'image temporelle
- 4) Détection probabiliste calculée avec les caractéristiques locale (e.g. Dérivées de Gaussiens).

Détection : Classification et Groupement de pixels.

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

Apprentissage :

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

Soit que dans chaque image, nous avons identifié "à la main" que un région de pixels d'objet, pour un total de M_o pixels de l'objet : $\{C^o_m\}$.

Il faut construire 2 histogrammes :

1) $h(C)$, l'histogramme de tous les M pixels.

$$m=1 \text{ to } M: h(C_m) = h(C_m) + 1$$

2) $h_o(C)$, l'histogramme des M_o pixels de l'objet "o".

$$m=1 \text{ to } M_o: h_o(C^o_m) = h_o(C^o_m) + 1$$

Ensuite :

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

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

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

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

Exemple : Détection de visage par ratio d'histogramme de couleur

On peut utiliser les histogrammes avec la règle de Bayes pour détecter les régions de peau. De plus, la chrominance de peau est invariante à l'orientation :

La composant "luminant" est déterminé par l'orientation de la surface.

La composant "chrominant" est déterminé par la composition de la spectre de la source et le spectre d'absorption des pigments de la surfaces. Si la spectre de la source est constante, la chrominance indique l'identité de l'objet

Pour détecter le peau, nous pouvons construire une histogramme pour le vecteur de chrominance $C = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$

La chrominance peut être défini par plusieurs codages. Par exemple :

$$L = R+G+B \quad c_1 = \frac{R}{R+G+B} \quad c_2 = \frac{G}{R+G+B}$$

R, G, B sont les entiers. Donc, c_1, c_2 sont issu d'une ensemble finit de valeurs dans l'intervalle $[0, 1]$.

On peut transformer c_1, c_2 en entier entre $[0, N-1]$, par

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

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)

Un histogramme des chrominance, $h(C)$ des M pixels dans les images de entrainement donne leurs fréquences d'occurrence.

Soit N images de $I \times J$ pixels : $M = I \times J \times N$

Soit une région de peau W_n dans chaque image pour un totalde M_0 pixels.

pour les M pixels dans les N images :

Pour chaque pixel $C = C(i, j)$ dans l'image, on incrémente la cellule de

l'histogramme qui correspond à C : $h(C) := h(C) + 1$

c'a-dire $h(c_1, c_2) := h(c_1, c_2) + 1$

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

Pour les M_o pixels dans les régions de peau W_n dans les N images :

$$(i,j) \quad W_n : h_o(C(i,j)) := h_o(C(i,j)) + 1$$

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_n sont dans les images, la probabilité de rencontrer un pixel de peau,

$$P(\text{peau}) = \frac{M_o}{M}$$

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

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



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 fenêtre (image) $w(i, j)$ 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 :

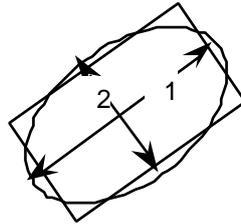
$$i^2 = \frac{1}{S} \sum_{i=1}^M \sum_{j=1}^N (w(i, j)) \cdot (i - \mu_i)^2$$

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

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

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.

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

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 ration 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}	$2^{30} = 1 \text{ Gi}$
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}	$2^{42} = 2 \text{ Ter}$
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 :

$$G_c = (G_x^L, G_x^C, G_x^2, G_{xx}^L, G_{xy}^L, G_{xx}^C, G_{xx}^2)$$

$$d = 7.$$

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.

Rappel :

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

La validité de l'approximation dépend du nombre de valeurs possible et de M .
En règle générale, on dit qu'il faut 10 exemples par cellule de l'histogramme.

On peut démontrer que l'écart type moyenne de l'erreur est en proportion avec la
racine de la somme des carrés de $h(x)$, N , sur le nombre d'échantillons, M .

$$\text{MSE} = \frac{1}{M} \sum_{x} h(x)^2$$

Que faire si la masse d'exemple est insuffisante : $M \ll N$?

Que faire si x n'est pas entier ? Il faut une fonction paramétrique pour $p(X)$.

La validité de l'approximation dépend du nombre de valeurs possible et de M .
En règle générale, on dit qu'il faut 10 exemples par cellule de l'histogramme.

Que faire si la masse d'exemple est insuffisante : $M < 10 (X_{\max} - X_{\min})$?

Que faire si x n'est pas entier ?

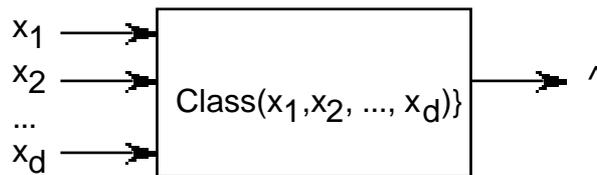
Dans ces cas, on peut faire appel à une fonction paramétrique pour $p(X)$.

La Classification Bayésienne

La technique Bayésienne de Classification repose sur une fonction de vérité probabiliste et le règle de Bayes.

Soit les événements E décrit par un vecteur de caractéristiques X : (E,X).
 Soit K classes d'événements $\{T_k\} = \{T_1, T_2, \dots, T_K\}$

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



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

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

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

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

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

Dans un système de vérité probabilisé, la valeur de vérité de la proposition une probabilité :

$$p(k) = p(E \quad k)$$

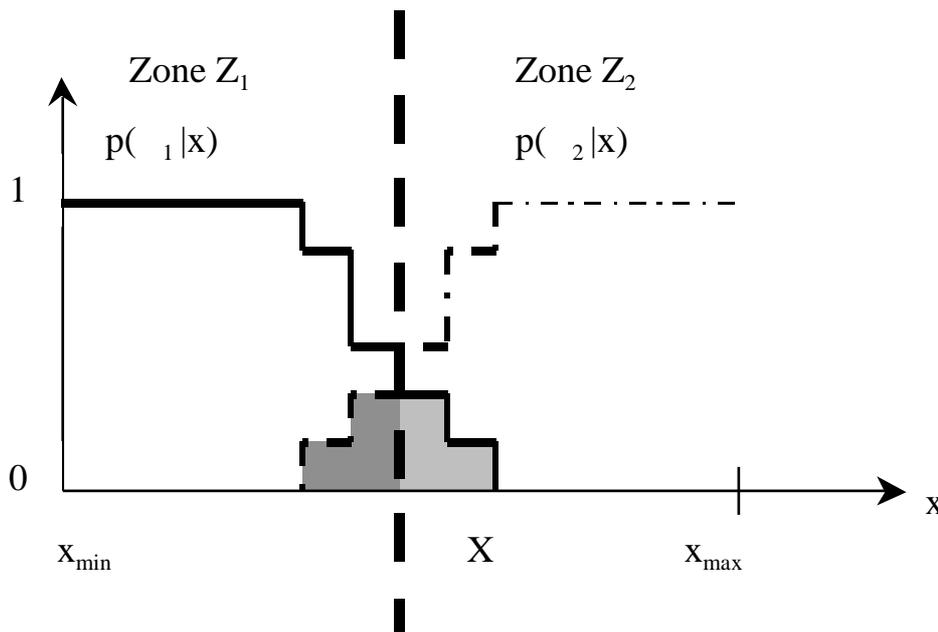
Le critère de décision est de minimiser le nombre d'erreur. Dans un système probabiliste, ca revient de minimiser la probabilité d'erreur. Ceci est équivalent à choisir la classe le plus probable.

$$\hat{k} = \text{Decider}(E \quad k) = \arg\text{-max}_k \{p(k | X)\}$$

Pour estimer la probabilité nous utilisons les caractéristiques, X, de l'événement.

Considère le cas D = 1 et K = 2. Dans ce cas, le domaine d'X est un axe. La classification est équivalente à une découpage du domaine d'X en deux zones : Z₁ et Z₂.

$$\hat{z}_1 \text{ si } X \in Z_1 \text{ et } \hat{z}_2 \text{ si } X \in Z_2$$



La probabilité d'erreur est la somme des probabilités de $p(z_2)$ en Z_1 et la somme de probabilité de $p(z_1 | X)$ en zone 2.

$$p(\text{erreur}) = \int_{Z_1} p(z_2 | X) + \int_{Z_2} p(z_1 | X)$$

La minimum est atteint quand $d(g_k(X)) = \arg\text{-max}_k \{p(z_k | X)\}$

Pour faire $g_k(X)$ nous allons utiliser :

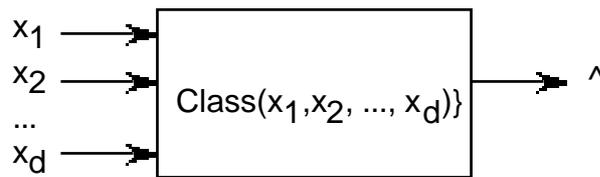
1) La règle de Bayes :

$$p(z_k | X) = \frac{P(X | z_k) p(z_k)}{P(X)}$$

2) La loi Normale :

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

Fonctions de Discrimination



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

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

$g(X)$: Une fonction de discrimination est une fonction $\mathbb{R}^D \rightarrow \mathbb{R}^K$
 $d()$: Une fonction de décision : $\mathbb{R}^K \rightarrow \{1, \dots, K\}$

$$g(X) = \begin{pmatrix} g_1(X) \\ g_2(X) \\ \dots \\ g_K(X) \end{pmatrix}$$

Etant donnée X , pour chaque k il existe une valeur de probabilité $p(k | X)$

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

Dans le cas général la probabilité minimum d'erreur est faite si k est choisi tel que

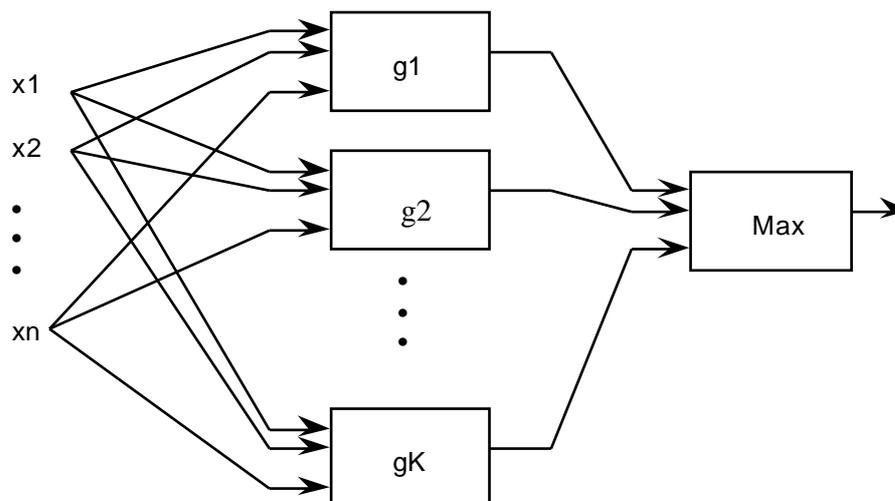
$$k = \arg\max_k \{g_k(X)\} = \arg\max_k \{p(k | X)\} = \arg\max_k \left\{ \frac{P(X | k) p(k)}{P(X)} \right\}$$

mais, comme $P(X)$ est constant pour tous k ,

$$k = \arg\max_k \{P(X | k) p(k)\}$$

Il suffit de l'évaluer $P(X | k)$, pour $X=x$

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



Fonction classique : Loi Normale

$$P(X|k) = \mathcal{N}(X; \mu_k, C_k)$$

ou encore par mélange de Gaussiennes

$$P(X|k) = \sum_{n=1}^N \alpha_n \mathcal{N}(X; \mu_{kn}, C_{kn})$$

par Exemple : Les deuxièmes moments sont "invariants" à l'orientation

μ_i est l'orientation de l'axe "majeur" et μ_j est le rapport entre la longueur et la largeur.

μ_j est une caractéristique invariante de la taille et de l'orientation.

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

