

# Formation et Analyse d'Images

James L. Crowley

ENSIMAG 3

Premier Sémestre 2007/2008

Séance 7

23 novembre 2007

## Analyse et Réconnaissance d'Images

### Plan de la Séance :

La Reconnaissance.....	2
La Probabilité d'un Evénement.....	4
Définition Fréquentielle.....	4
Définition Axiomatique.....	4
La probabilité de la valeur d'une variable aléatoire.....	5
La Règle de Bayes.....	6
Classification des Pixels par Ratio d'Histogramme .....	8
Histogrammes.....	8
Exemple : Détection d'object par ratio d'histogramme de couleur.....	8
Histogrammes de Champs Réceptifs.....	10

## La Reconnaissance

La reconnaissance est une capacité fondamentale de l'intelligence et même de la vie. Pour la survie, il faut savoir reconnaître les amis, les ennemies et la nourriture.

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

Identifier : Reconnaître un entité comme un individu

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

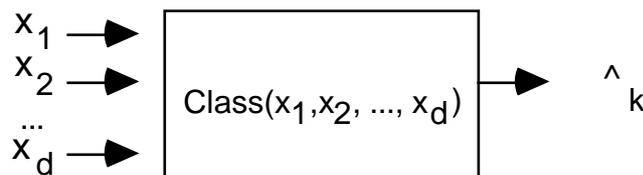
Un ensemble est défini par un test d'appartenance.

La classification est un processus d'association un entité (un événement) à une classe. L'entité est décrit par un vecteur des caractéristiques, produit par une observation. L'affectation de l'entité à une classe est fait par un test, calculer sur le vecteur de caractéristiques.

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

En notation vectorielle :  $X = \begin{matrix} x_1 \\ x_2 \\ \dots \\ x_D \end{matrix}$

Pour un vecteur de caractéristique,  $X$ , un processus de classification propose une estimation de la classe,  $\hat{k}$



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

Les classes peuvent êtres définis par  
 extension : une liste complète des membres  
 intention : une conjonction de prédicats.

par extension : Une comparaison d'une observation avec des membres connus de la classe (des prototypes). Ceci correspond (grosso modo) à des méthode dite "générative" de reconnaissance.

Dans ce cas, la classification peut être faite par comparaison avec les membres de la classe.

$$\hat{k} = \arg\text{-max}_k \{ \text{Sim}(Y, X_m^k) \} \quad \text{pour tout } k, m \quad \text{ou bien}$$

$$\hat{k} = \arg\text{-min}_k \{ \|Y - X_m^k\| \} \quad \text{pour tout } k, m$$

par intention : Conjonction de prédicats définis sur les propriétés observées. Ceci correspond (grosso modo) à des méthodes dites "discriminatives" de reconnaissance.

Génératives : Les techniques fondées sur un modèle.

Discriminatives : Les techniques fondées sur des tests quelconques.

La teste d'appartenance est une forme de partition de l'espace de caractéristiques.

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

Le critère est la probabilité conditionnelle d'appartenance.

$$p(k) = \Pr(E = T_k) \quad \begin{array}{l} \text{Proposition que l'événement } E \text{ est la classe } k \\ \text{Probabilité que } E \text{ est un membre de la classe } k. \end{array}$$

Ayant une observation,  $X$ , le critère de partition est la probabilité.

$$p(k | X) = \Pr(E = T_k \text{ étant donnée l'observation de } X)$$

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

Cette probabilité est fournie par la règle de Bayes.

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

## La Probabilité d'un Evénement.

La sémantique (ou "sens") de la probabilité d'un événement peut être fourni par sa fréquence d'occurrence ou par un système d'axiomes. L'approche fréquentielle a l'avantage d'être facile à comprendre. Par contre, elle peut entraîner les difficultés dans l'analyse. La définition axiomatique favorise les analyses mathématiques.

Dans le deux cas, la probabilité est une fonction numérique,  $\Pr() \in [0, 1]$ .  
Le domaine de la fonction  $\Pr()$  est les événements,  $E$ .

### Définition Fréquentielle.

Une définition "Fréquentielle" de la probabilité sera suffisante pour la plupart des techniques vues dans ce cours.

Soit  $M$  observations des événement aléatoire dont  $M_k$  appartiennent à la classe  $A_k$ .  
La Probabilité d'observer un événement  $E$  de la classe  $A_k$  est

$$p(E = A_k) = \lim_M \left\{ \frac{M_k}{M} \right\}$$

Pour le cas pratique où  $M$  est fini,  $p(E = A_k) \approx \frac{M_k}{M}$

La validité (ou précision) de l'approximation dépend du nombre d'échantillons  $M$ .

### Définition Axiomatique.

Une définition axiomatique permet d'appliquer certaines techniques d'analyse de systèmes probabilistes. Trois postulats sont suffisants :

Postulat 1 :  $A_k \in S : p(E = A_k) \geq 0$

Postulat 2 :  $p(E \in S) = 1$

Postulat 3 :

$$A_i, A_j \in S \text{ tel que } A_i \cap A_j = \emptyset : p(E \in A_i \cup A_j) = p(E \in A_i) + p(E \in A_j)$$

## La probabilité de la valeur d'une variable aléatoire

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

## La Règle de Bayes

Soit un événement "E". Soit deux tribus d'événements A et B tel que certains événements sont communs à A et à B.

E peut appartenir à A ∩ B ou à  $\bar{A} \cap B$  ou à  $A \cap \bar{B}$  ou à  $\bar{A} \cap \bar{B}$

Soit deux propositions p et q.

donc  $P(p) = \Pr\{E \in A\}$  et  $P(q) = \Pr\{E \in B\}$ .

Par axiome 2 de la définition des systèmes de probabilités :

$$P(q) + P(\bar{q}) = 1.$$

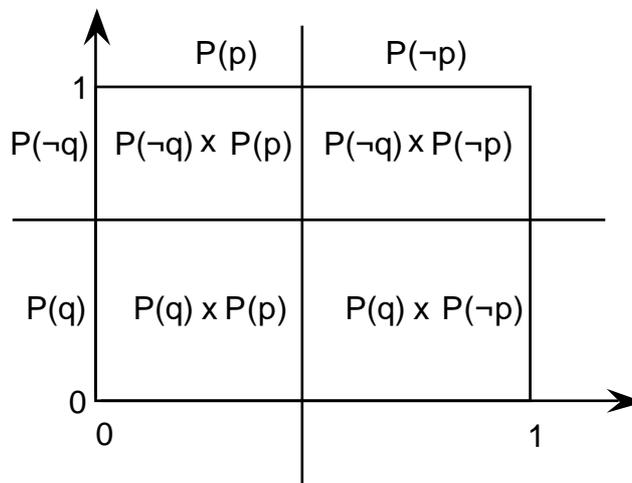
$P(p \cap q)$  est la probabilité "conjointe" de p et q.

Si p et q sont indépendantes

$$P(p \cap q) = P(p) \cdot P(q),$$

$$P(p \cup q) = P(p) + P(q).$$

On peut voir ça d'une manière graphique :



$$P(p \cap q) + P(p \cap \bar{q}) + P(\bar{p} \cap q) + P(\bar{p} \cap \bar{q}) = 1$$

Dans ce cas, les probabilités marginales sont

$$P(p) = P(p \cap q) + P(p \cap \bar{q})$$

$$P(q) = P(p \cap q) + P(\bar{p} \cap q)$$

La probabilité conditionnelle de q étant donnée p s'écrit  $P(q | p)$

$$P(q | p) = \frac{P(p \wedge q)}{P(p)} = \frac{P(p \wedge q)}{P(p \wedge q) + P(p \wedge \neg q)}$$

de la même manière :

$$P(p | q) = \frac{P(p \wedge q)}{P(q)} = \frac{P(p \wedge q)}{P(p \wedge q) + P(\neg p \wedge q)}$$

Par algèbre on déduit :

$$P(q | p) P(p) = P(p \wedge q) = P(p | q) P(q)$$

d'où

$$P(q | p) P(p) = P(p | q) P(q)$$

Ceci est une forme de règle de Bayes. On peut écrire :

$$P(q | p) = \frac{P(p | q) P(q)}{P(p)}$$

$P(q | p)$  est la probabilité "conditionnelle" ou "postérieur"

## Classification des Pixels par Ratio d'Histogramme

### Histogrammes

La probabilité d'une valeur  $x$  est sa fréquence d'occurrence.

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

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 ratio entre le nombre de Cellules,  $Q = N^d$ , de  $h(x)$  et le nombre d'échantillons,  $M$ .

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  $M \gg Q = N^d$   
En règle générale, on dit qu'il faut 10 exemples par cellule de l'histogramme.

### Exemple : Détection d'objet par ratio d'histogramme de couleur

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

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

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  $C$  :  $h(C) := h(C) + 1$

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

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  $O$ .

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

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.

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

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 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,  $M$ .

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

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  $p(X)$ .