

Classification bayésienne

**TD 1 - Module Reconnaissance de Forme - Image 3
FISE 2.**

MANSSOUM MUSTAPHA, SABIR ILYASS.

2021 - 2022.

Table des matières

Liste de figures	3
Exercice 1 - Décision bayésienne	4
1	7
1.a	7
1.b	8
1.c	9
Remarque:	11
1.d	11
Commentaire:	17
2	17
2.a	17
2.b	18
2.c	18
Exercice 2 - Application à la détection de visages automatique	18
2.3 Partie 1 - Apprentissage du modèle de peau	19
2.3.1 Sélection des échantillons d'apprentissage	19
1	19
2	19
2.3.2 Choix des caractéristiques	19
1	20
2	20
3	20
4	21
2.3.3 Calcul des paramètres du modèle	21
1	22
2	22
3	22
4	22
2.3.4 Application du modèle à une image test	23
1	23
2	24
3	24
2.4 Partie 2 - Vers un classifieur bayésien	24
2.4.1 Classe "non-peau"	24
1	24
2	24
3	24
.	24
4	25
5	25
6	25

Liste de figures

Densités de probabilités des classes (1:rouge, 2:jaune)	7
Densité des probabilités (pxw1: rouge, pxw2: jaune, pw1x: noir, pw2x: vert, px: bleu)	9
Le graphe des densités de probabilités des 2 classes	12
le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori	12
Le graphe des densités de probabilités des 2 classes	13
le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori	13
Le graphe des densités de probabilités des 2 classes	14
le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori	14
Le graphe des densités de probabilités des 2 classes	15
le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori	15
Le graphe des densités de probabilités des 2 classes	16
le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori	16
le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori	18
Processus à mettre en oeuvre	19
.	20
L'histogramme des caractéristique de la distribution de chrominances Cb Cr,	20
résultat de la partie 2.3.4	24
résultat de la partie 2.4.1	25

Exercice 1 - Décision bayésienne

On considère un problème à 2 classes, modélisées par des lois gaussiennes de variances σ^2 et de moyennes respectives 0 et 1:

$$\forall x \in \mathbb{R}, p(x|\omega_1) = \frac{1}{\sqrt{\pi}} \exp(-x^2) \quad p(x|\omega_2) = \frac{1}{\sqrt{\pi}} \exp(-(x-1)^2)$$

Où ω_1, ω_2 sont deux réels positifs tel que $\omega_1 + \omega_2 = 1$

Question préliminaire: le calcul de la valeur de la variance des deux lois de probabilités définies précédemment:

La variance de $p(X|\omega_1)$:

On a

$$\text{Var}(p(X|\omega_1)) = \mathbb{E}(p(X|\omega_1)^2) - \mathbb{E}(p(X|\omega_1))^2$$

Commençant par le calcul de l'espérance

$$\begin{aligned} \mathbb{E}(p(X|\omega_1)) &= \int_{-\infty}^{+\infty} x p(x|\omega_1) dx \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x e^{-x^2} dx \\ &= 0 \end{aligned}$$

Car la fonction $x \mapsto x e^{-x^2}$ est impaire

D'autre part, on a

$$\begin{aligned} \mathbb{E}(p(X|\omega_1)^2) &= \int_{-\infty}^{+\infty} x^2 p(x|\omega_1) dx \\ &= \frac{2}{\sqrt{\pi}} \int_0^{+\infty} x^2 e^{-x^2} dx \\ &= \frac{1}{\sqrt{\pi}} \int_0^{+\infty} x \left(-\frac{e^{-x^2}}{2} \right)' dx \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sqrt{\pi}} \left[-x \frac{e^{-x^2}}{2} \right]_0^{+\infty} + \frac{1}{2\sqrt{\pi}} \int_0^{+\infty} e^{-x^2} dx \\
&= \frac{1}{2\sqrt{\pi}} \int_0^{+\infty} e^{-x^2} dx
\end{aligned}$$

Notons $I = \int_0^{+\infty} \exp(-x^2) dx$, on a par changement de variables en coordonnées cylindriques

$$\begin{aligned}
\iint_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy &= \int_{r=0}^{+\infty} \int_{\theta=0}^{2\pi} r \exp(-r^2) dr d\theta \\
&= 2\pi \int_{r=0}^{+\infty} r \exp(-r^2) dr \\
&= 2\pi \left[-\frac{r}{2} \exp(-r^2) \right]_0^{+\infty} \\
&= \pi
\end{aligned}$$

D'autre part d'après le théorème de Fubini on peut écrire

$$\begin{aligned}
\iint_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy &= \left(\int_0^{+\infty} \exp(-x^2) dx \right) \left(\int_0^{+\infty} \exp(-y^2) dy \right) \\
&= I^2
\end{aligned}$$

Donc $I = \sqrt{\pi}$

D'où

$$\mathbb{E}(p(X|\omega_1)^2) = \frac{1}{2}$$

Par suite

$$\text{Var}(p(X|\omega_1)) = \frac{1}{2}$$

La variance de $p(X|\omega_2)$:

Da la même façon, on a

$$\text{Var}(p(X|\omega_2)) = \mathbb{E}(p(X|\omega_1)^2) - \mathbb{E}((X|\omega_1))^2$$

Or,

$$\begin{aligned}
 \mathbb{E}(p(X|\omega_2)) &= \int_{-\infty}^{+\infty} x p(x|\omega_2) dx \\
 &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x \exp(-(x-1)^2) dx \\
 &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} (x+1) \exp(-x^2) dx \\
 &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \exp(-x^2) dx + \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x \exp(-x^2) dx \\
 &= \frac{1}{\sqrt{\pi}} \int_0^{+\infty} \exp(-x^2) dx \\
 &= \frac{1}{\sqrt{\pi}} I \\
 &= 1
 \end{aligned}$$

Et,

$$\begin{aligned}
 \mathbb{E}(p(X|\omega_2)^2) &= \int_{-\infty}^{+\infty} x^2 p(x|\omega_2) dx \\
 &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x^2 \exp(-(x-1)^2) dx \\
 &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} (x+1)^2 \exp(-x^2) dx \\
 &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x^2 \exp(-x^2) dx + \frac{2}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x \exp(-x^2) dx + \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \exp(-x^2) dx \\
 &= \frac{1}{2} + 0 + 1 \\
 &= \frac{3}{2}
 \end{aligned}$$

Ainsi

$$\text{Var}(p(X|\omega_2)) = \frac{3}{2} - 1 = \frac{1}{2}$$

1 Prenons dans ce cas $p(\omega_1) = p(\omega_2) = \frac{1}{2}$

1.a L'exécution du script Matlab suivant:

```
Max_x = 4;
Min_x = -4;
Pas = 0.01;
x = [Min_x:Pas:Max_x];
L = length(x);

% Considérons que 2 classes sont modélisées par les densités de
% probabilité gaussiennes suivantes :
pxw1 = exp(-(x.*x))./sqrt(pi);
mu2 = 1;
sigma2 = sqrt(0.5); % ecart-type
pxw2 = exp(-(x-mu2).*(x-mu2)/(2*sigma2^2))./sqrt(2*pi*sigma2^2);

%% 1. 1er cas: Pw1=Pw2=0,5
Pw1 = 0.5;
Pw2 = 1 - Pw1;

% Tracés des densités de probabilités des 2 classes
figure(1)
hold on
plot(x,pxw1,'color','red')
plot(x,pxw2,'color','yellow')
title('Densités de probabilités des classes (1:rouge, 2:jaune)')
```

l'exécution de ce code nous donne le graphe des deux densité de probabilités des deux classes



Figure 1. Densités de probabilités des classes (1:rouge, 2:jaune)

1.b Calcul des lois de probabilités a posteriori des 2 classes

Pour tout $x \in \mathbb{R}$, pour $i \in \{0, 1\}$, on a

$$\begin{aligned} p(\omega_i|x) &= \frac{p(x|\omega_i)p(w_i)}{p(x)} \\ &= \frac{p(x|\omega_i)p(w_i)}{p(x|\omega_1)p(w_1) + p(x|\omega_2)p(w_2)} \end{aligned} \quad (1)$$

Donc

$$\begin{aligned} p(\omega_1|x) &= \frac{\frac{1}{2} \times \frac{1}{\sqrt{\pi}} \exp(-x^2)}{\frac{1}{2} \times \frac{1}{\sqrt{\pi}} \exp(-x^2) + \frac{1}{2} \times \frac{1}{\sqrt{\pi}} \exp(-(x-1)^2)} \\ &= \frac{1}{1 + \exp(2x-1)} \end{aligned}$$

Et

$$\begin{aligned} p(\omega_2|x) &= \frac{\frac{1}{2} \times \frac{1}{\sqrt{\pi}} \exp(-(x-1)^2)}{\frac{1}{2} \times \frac{1}{\sqrt{\pi}} \exp(-x^2) + \frac{1}{2} \times \frac{1}{\sqrt{\pi}} \exp(-(x-1)^2)} \\ &= \frac{1}{1 + \exp(1-2x)} \end{aligned}$$

A partir de l'équation (1), on ajoute les lignes de code suivant pour tracer des graphes des probabilités a posteriori des 2 classes sur le même graphique

```
% Ajouter sur le même graphique les probabilités a posteriori
px = pxw1 * Pw1 + pxw2 * Pw2;
Pw1x = pxw1 * Pw1 ./ px;
Pw2x = pxw2 * Pw2 ./ px;

hold on;
plot(x, Pw1x, 'Color', 'black')
plot(x, Pw2x, 'Color', 'green')
plot(x, px, 'Color', 'blue'); title("densité des probabilités (pxw1: rouge, pxw2: jaune,
pw1x: noir, pw2x: vert, px: bleu)");
```

Après l'exécution du code, on obtient le graphe présenté dans la figure 2

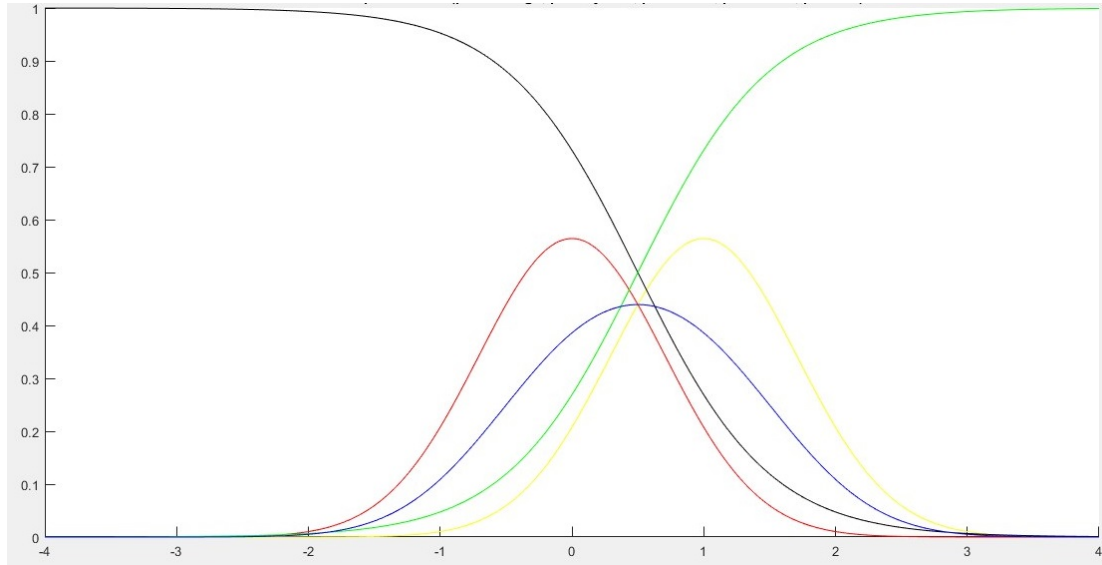


Figure 2. Densité des probabilités (pxw1: rouge, pxw2: jaune, pw1x: noir, pw2x: vert, px: bleu)

1.c Revenant au cas général, et notons $\mu_i = \mathbb{E}(p(X|\omega_i))$ et $\sigma_i^2 = \text{Var}(p(X|\omega_i))$, $i = 1, 2$
 Les lois des probabilités des deux classes sont alors données par:

$$\forall x \in \mathbb{R}, \forall i \in \{1, 2\} p(x|\omega_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(x - \mu_i)^2}{2\sigma_i^2}\right)$$

Et donc les lois de probabilités à posteriori des 2 classes sont données par

Pour $i \in \{1, 2\}$ et pour tout $x \in \mathbb{R}$:

$$\begin{aligned} p(\omega_i|x) &= \frac{p(x|\omega_i)p(w_i)}{p(x)} \\ &= \frac{p(x|\omega_i)p(w_i)}{p(x|\omega_1)p(w_1) + p(x|\omega_2)p(w_2)} \\ &= \frac{\frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(x - \mu_i)^2}{2\sigma_i^2}\right) p(w_i)}{\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma_1^2}\right) + \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(x - \mu_2)^2}{2\sigma_2^2}\right)} \end{aligned}$$

Le seuil conduisant au taux d'erreur de classification minimum selon la théorie bayésienne se définit par le réel x_B qui vérifie

$$p(\omega_1|x_B) = p(\omega_2|x_B)$$

Donc nous cherchons x_B qui vérifie l'équation:

$$p(x_B|\omega_1)p(w_1) = p(x_B|\omega_2)p(w_2)$$

Ce qui est équivalent à écrire

$$\log(p(x_B|\omega_1)) + \log(p(w_1)) = \log(p(x_B|\omega_2)) + \log(p(w_2))$$

Ainsi

$$\frac{1}{2} \left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2} \right) x_B^2 + \left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2} \right) x_B + \frac{\mu_2^2}{2\sigma_2^2} - \frac{\mu_1^2}{2\sigma_1^2} + \log \left(\frac{\sigma_2 p(\omega_1)}{\sigma_1 p(\omega_2)} \right) = 0$$

Par suite

$$(\sigma_1^2 - \sigma_2^2) x_B^2 + 2(\mu_1 \sigma_2^2 - \mu_2 \sigma_1^2) x_B + \mu_2^2 \sigma_1^2 - \mu_1^2 \sigma_2^2 + 2\sigma_1^2 \sigma_2^2 \log \left(\frac{\sigma_2 p(\omega_1)}{\sigma_1 p(\omega_2)} \right) = 0$$

Posons

$$\begin{cases} a = \sigma_1^2 - \sigma_2^2 \\ b = 2(\mu_1 \sigma_2^2 - \mu_2 \sigma_1^2) \\ c = \mu_2^2 \sigma_1^2 - \mu_1^2 \sigma_2^2 + 2\sigma_1^2 \sigma_2^2 \log \left(\frac{\sigma_2 p(\omega_1)}{\sigma_1 p(\omega_2)} \right) \end{cases}$$

En a une seul solution de l'équation au cas où $a = 0$ (i.e $\sigma_1 = \sigma_2$). Dans ce cas on a:

$$x_B = \frac{-c}{b}$$

Si $a \neq 0$, dans ce cas l'équation trinômiale admet deux solution:

$$x_{B,1} = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \text{ et } x_{B,2} = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

A partir de ces calculs, on écrit les lignes de code suivants pour calculer le(s) seuil(s) au taux d'erreur de classification minimum selon la théorie bayésienne

```
% En déduire le seuil de classification optimale.
mu1 = 0; sigma1 = sqrt(0.5);
a = sigma1^2 - sigma2^2;
b = 2 * (mu1 * sigma2^2 - mu2 * sigma1^2);
c = (mu2^2) * (sigma1^2) - (mu1^2) * (sigma2^2) + 2 * (sigma1^2) * (sigma2^2)
* log((sigma2 * Pw1) / (sigma1 * Pw2));
if a == 0
    seuil = -c / b
else
    deltaSqrt = sqrt(b^2 - 4 * a * c);
    seuil1 = (-b + deltaSqrt) / (2 * a)
    seuil2 = (-b - deltaSqrt) / (2 * a)
end
```

On obtient comme résultat:

seuil = 0,5

Remarque:

Le seuil qui minimise l'erreur correspond bien au réel x_B qui vérifie l'équation

$$p(x_B|\omega_1)p(w_1) = p(x_B|\omega_2)p(w_2)$$

Démonstration. Dans le cas où on a un seul seuil (le deuxième cas se traite de la même façon)

Notons E : la fonction de probabilité d'erreur de classification, on a

$$E(x) = \int_{-\infty}^x p(t|\omega_2)p(w_2)dt + \int_x^{+\infty} p(t|\omega_1)p(w_1)dt$$

E est une fonction dérivable sur \mathbb{R} , et au minimum (en x_B) on a $\left. \frac{dE}{dx} \right|_{x=x_B} = 0$

Donc

$$p(x_B|\omega_2)p(w_2) - p(x_B|\omega_1)p(w_1) = 0$$

D'où le résultat. □

1.d Dans le cas général et avec les mêmes notations de la question précédente. Notons $E(\mu_1, \mu_2, \sigma_1, \sigma_2)$: la probabilité d'erreur de classification (qui dépend de $\mu_1, \mu_2, \sigma_1, \sigma_2$)

Dans le cas général, le positionnement relatif des graphes des probabilités à posteriori peut être quelconque (dépend de $\mu_1, \mu_2, \sigma_1, \sigma_2$), mais toujours l'erreur de classification veut l'intégrale du minimum des probabilités à posteriori des deux classes

$$E(\mu_1, \mu_2, \sigma_1, \sigma_2) = \int_{-\infty}^{+\infty} \min(p(x|\omega_1)p(w_1), p(x|\omega_2)p(w_2))dx$$

pour calculer cette erreur, on ajoute à notre script Matlab précédente les lignes de codes suivants:

```
% calcul de probabilité d'erreur de classification
PErrreur = Pas * sum(min(pwx1 * Pw1, pwx2 * Pw2))
```

On obtient le résultat suivant:

PErrreur = 0,2398

On refait les mêmes questions en faisant varier les paramètres de la 2ème loi de la manière suivante (on changeant seulement **mu2** et **sigma2** dans le code)

✓ $\mu_2 = 0.5$:

Le graphe des densités de probabilités des 2 classes:

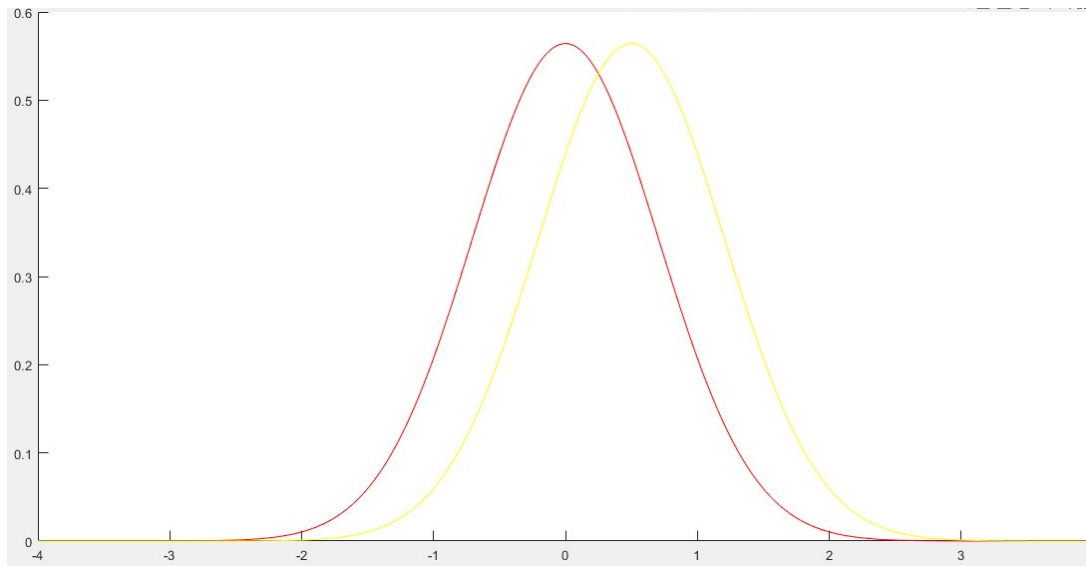


Figure 3. Le graphe des densités de probabilités des 2 classes

Le graphe des densités de probabilités des lois de probabilités a posteriori des 2 classes

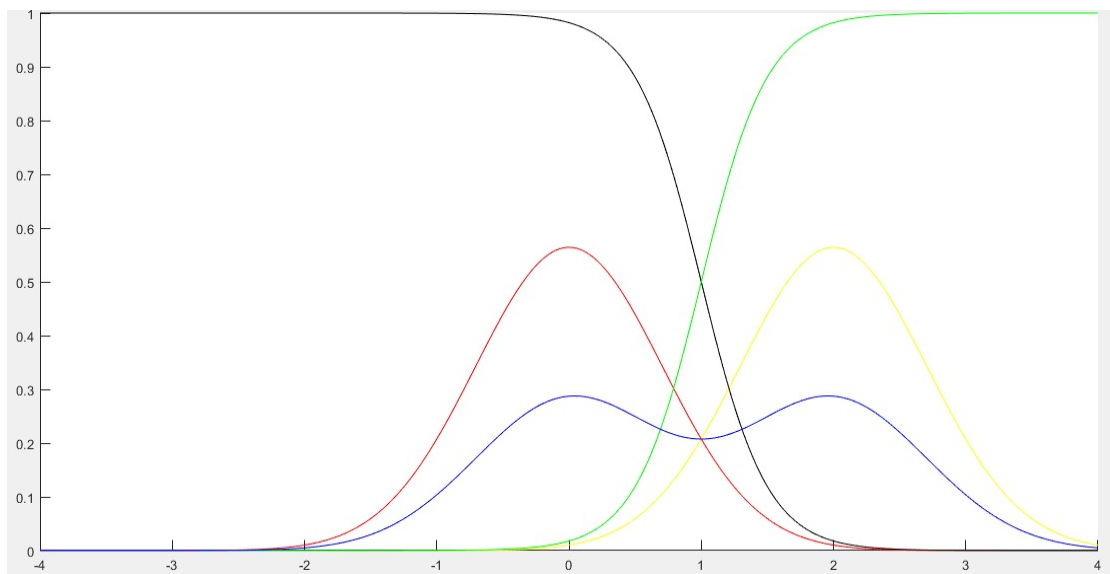


Figure 4. le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori

Le seuil conduisant au taux d'erreur de classification minimum selon la théorie bayésienne:

$$\text{seuil} = 0.25$$

La probabilité d'erreur de classification obtenue:

$$P_{\text{Erreur}} = 0.3618$$

$$\checkmark \mu_2 = 2:$$

Le graphe des densités de probabilités des 2 classes:



Figure 5. Le graphe des densités de probabilités des 2 classes

Le graphe des densités de probabilités des lois de probabilités a posteriori des 2 classes

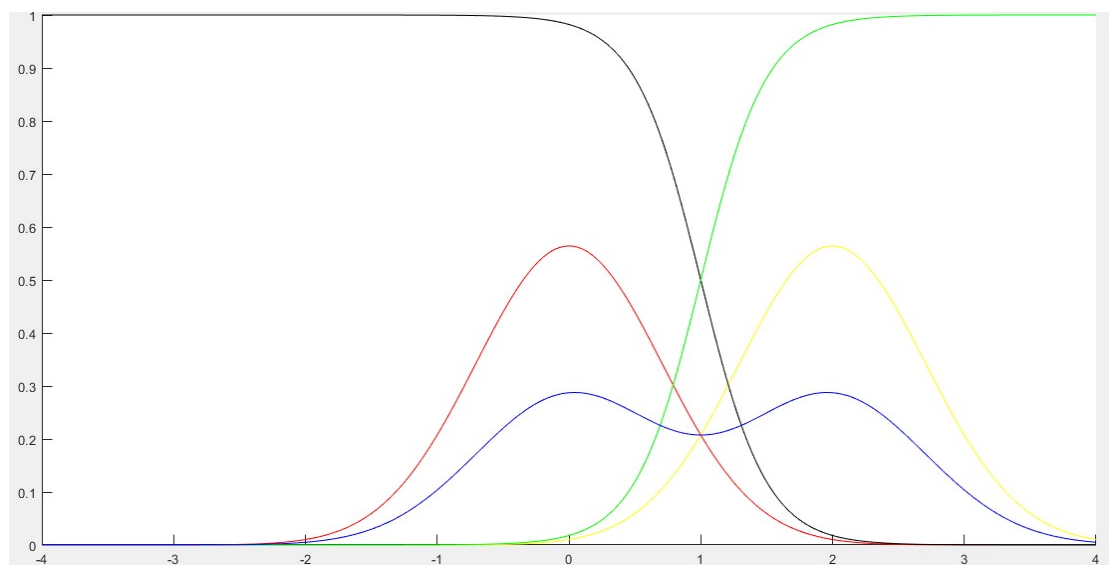


Figure 6. le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori

Le seuil conduisant au taux d'erreur de classification minimum selon la théorie bayésienne:

$$\text{seuil} = 1$$

La probabilité d'erreur de classification obtenue:

$$P_{\text{Erreur}} = 0.0787$$

$$\checkmark \mu_2 = 1, \sigma_2 = 0.25:$$

Le graphe des densités de probabilités des 2 classes:

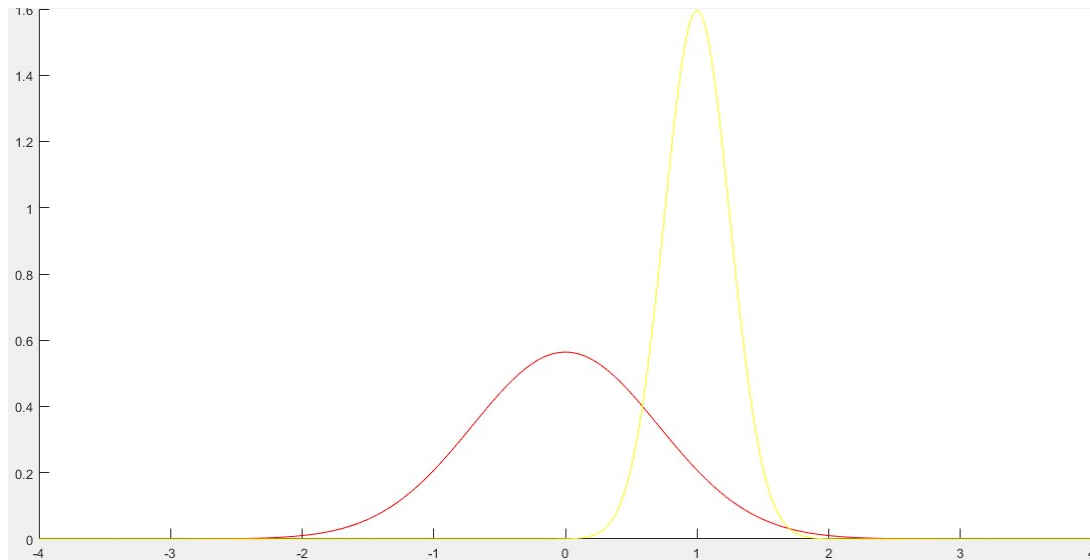


Figure 7. Le graphe des densités de probabilités des 2 classes

Le graphe des densités de probabilités des lois de probabilités a posteriori des 2 classes

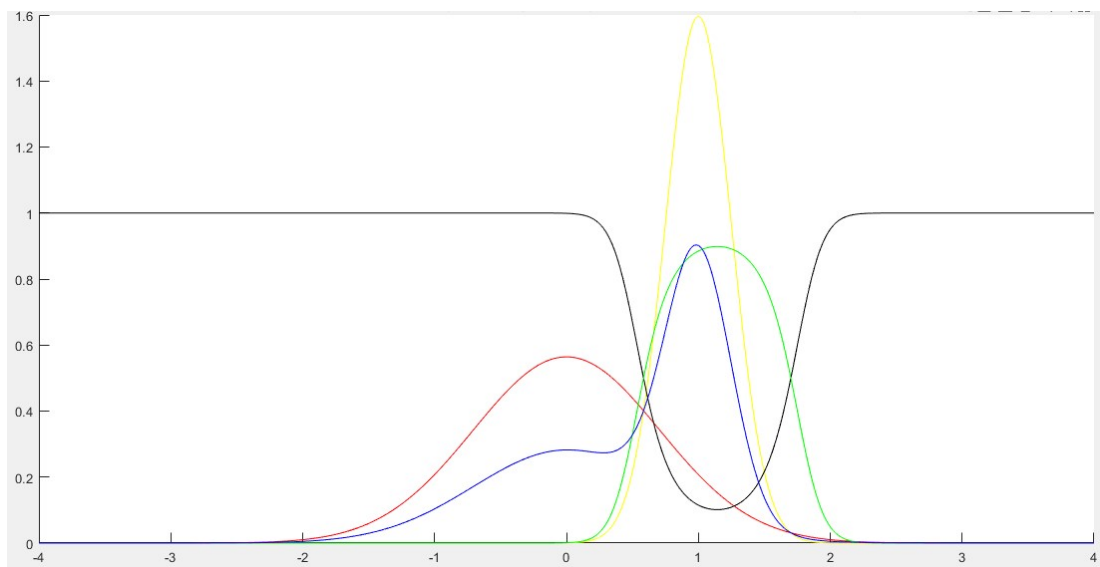


Figure 8. le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori

Les seuils conduisant au taux d'erreur de classification minimum selon la théorie bayésienne:

$$\text{seuil1} = 1.7012 \text{ et } \text{seuil2} = 0.5845$$

La probabilité d'erreur de classification obtenue:

$$P_{\text{Erreur}} = 0.1235$$

$\mu_2 = 1, \sigma_2 = 1$:

Le graphe des densités de probabilités des 2 classes:



Figure 9. Le graphe des densités de probabilités des 2 classes

Le graphe des densités de probabilités des lois de probabilités a posteriori des 2 classes



Figure 10. le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori

Les seuils conduisant au taux d'erreur de classification minimum selon la théorie bayésienne:

$$\text{seuil1} = -2.6411 \text{ et } \text{seuil2} = 0.6411$$

La probabilité d'erreur de classification obtenue:

$$PErreur = 0.2710$$

$\sqrt{\mu_2 = 1, \sigma_2 = 2}$:

Le graphe des densités de probabilités des 2 classes:

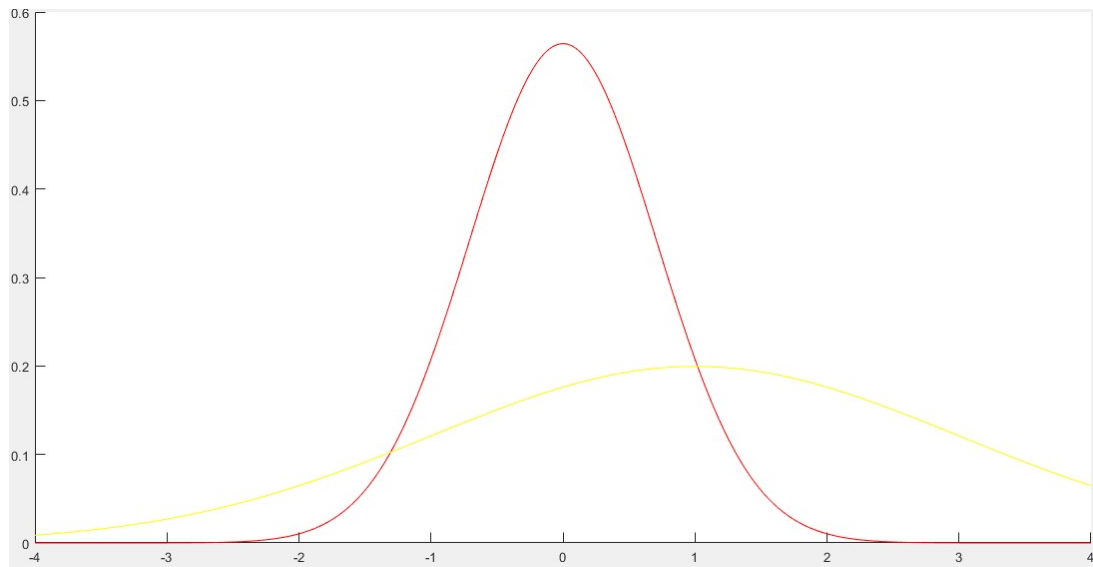


Figure 11. Le graphe des densités de probabilités des 2 classes

Le graphe des densités de probabilités des lois de probabilités a posteriori des 2 classes



Figure 12. le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori

Les seuils conduisant au taux d'erreur de classification minimum selon la théorie bayésienne:

$$\text{seuil1} = -1.3054 \text{ et } \text{seuil2} = 1.0197$$

La probabilité d'erreur de classification obtenue:

$$P_{\text{Erreur}} = 0.2432$$

Commentaire:

On remarque que plus $|\sigma_2 - \sigma_1|$ et $|\mu_1 - \mu_2|$ sont grands plus que l'erreur de classification est petit, et c'est tout à fait logique, puisque plus $|\sigma_2 - \sigma_1|$ et $|\mu_1 - \mu_2|$ sont grands implique que les probabilités des 2 classes sont plus distinguables.

De manière générale, dans la théorie bayésienne, plus les probabilités des classes sont distinguables, plus la classification par décision bayésienne est plus efficace. Dans le cas des distributions gaussiennes, plus $|\sigma_i - \sigma_j|$ et $|\mu_i - \mu_j| \forall i \neq j$ sont grands plus on a moins d'erreur de classifications, où σ_i μ_i désigne respectivement l'écart-type et la moyenne du probabilité de la classe $i, \forall i$

2 Prenons maintenant le cas $P(\omega_1) = 0,9$ et $P(\omega_2) = 0,1$

2.a On ajoutons les lignes de codes suivants, on obtient le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori.

Plus qu'une classe à une probabilité à priori grand, plus que l'erreur de décision diminue

```
%% 2. Modification des valeurs de probabilité à priori
Pw1 = 9/10;
Pw2 = 1/10;

% Tracés des densités de probabilités des 2 classes
figure(2);
hold on;
plot(x,pxw1,'color','red')
plot(x,pxw2,'color','yellow')
title('Densités de probabilités des classes (1:rouge, 2:jaune)')

% Ajouter sur le même graphique les probabilités à posteriori
px = pxw1 * Pw1 + pxw2 * Pw2;
Pw1x = pxw1 * Pw1./px;
Pw2x = pxw2 * Pw2./px;

hold on;
plot(x, Pw1x, 'Color', 'black')
plot(x, Pw2x, 'Color', 'green')
plot(x, px, 'Color', 'blue'); title('densité des probabilités (pxw1: rouge, pxw2: jaune,
pw1x: noir, pw2x: vert, px: bleu)');
```

Après l'exécution, on obtient le graphe suivant:



Figure 13. le graphe des deux densité de probabilités des deux classes, et des probabilités à posteriori

La probabilité d'erreur de classification obtenue:

$$P_{\text{Erreur}} = 0.0908$$

2.b Le seuil optimal conduisant au taux d'erreur de classification minimum selon la théorie bayésienne:

$$\text{seuil} = 1.5986$$

2.c On décide de ne prendre en compte que le numérateur de la probabilité à posteriori, le seuil optimal ne change pas (seuil = 1,5986), en effet le dénominateur est une constante positive pour toutes les classes, sa valeur n'influence pas sur le maximum à posteriori.

Exercice 2 - Application à la détection de visages automatique

On souhaite mettre en place un système pour détecter automatiquement la présence de visages sur des images (Figure 14). Ce système exploitera un modèle de couleur de peau gaussien multivarié (dimension supérieure ou égal à 2) et fonctionnera sur le principe d'un classifieur bayésien optimal, avec modèles de gaussiennes.

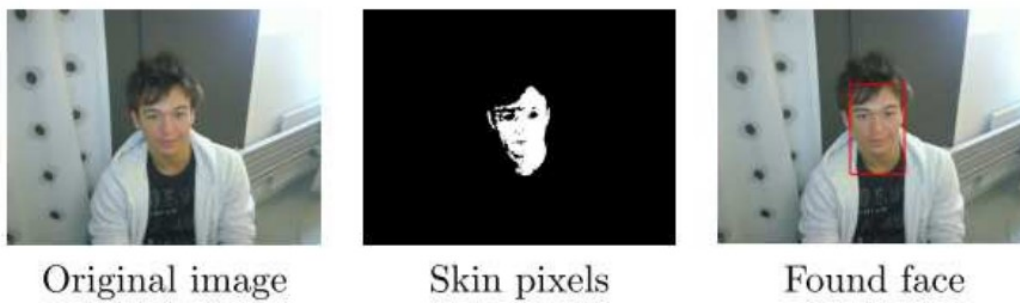


Figure 14. Processus à mettre en oeuvre

2.3 Partie 1 - Apprentissage du modèle de peau

2.3.1 Sélection des échantillons d'apprentissage

Dans ce TD, on souhaite modéliser la couleur de peau par un modèle gaussien multivarié. Soit \mathbf{x} le vecteur qui contient les données couleur d'un pixel (rgb par exemple). On cherche à déterminer le modèle suivant ($\omega_1 = \text{peau}$) :

$$p(\mathbf{x}|\omega_1) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_1|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mu_1)^T \Sigma_1^{-1} (\mathbf{x}-\mu_1)}$$

avec $|\Sigma_1|$ le déterminant de la matrice Σ_1 . Les paramètres (μ_1, Σ_1) sont inconnus. Ils doivent être appris à partir de données d'apprentissage.

Il est donc nécessaire de disposer de données d'apprentissage. Par chance, de nombreuses personnes ont constitué des bases de données d'images couleur de visage et les ont rendu accessibles. Nous allons utiliser les images fournies à l'adresse : <http://vis-www.cs.umass.edu/lfw/> La base de données complètes contient un très grand nombre d'images, largement plus qu'il nous en faut pour définir notre modèle de peau. Par conséquent, vous ne devez télécharger qu'une petite proportion d'images à : <http://viswww.cs.umass.edu/lfw/lfw-bush.zip>. On se propose d'utiliser ces images pour apprendre le modèle de couleur de peau (il n'est pas utile d'utiliser toutes les images pour apprendre le modèle).

1 On observant les images, on peut utiliser un approche couleur pour extraire les les échantillons de l'image.

2 En regardant le code du fichier Matlab *SelectPixelsCentre.m*, on remarque que le rôle de cette fonction est de selection une sous image centrée de l'image d'entrée, p est un réel qui doit être entre $[0, 1]$, plus p est proche de 1 plus la dimension de sous image de l'image est faible, l'avantage de cette approche qu'on selectionne seulement les zones d'intérêts dans les images d'apprentissage, cette approche est plus adaptées pour les images où le visage est situé au centre de l'image, mais l'inconvénient de cette approche apparait pour une base de données d'apprentissage où les visages se positionnent de manière quelconque sur l'image.

2.3.2 Choix des caractéristiques

Maintenant que l'on dispose des échantillons d'apprentissage, une question essentielle porte sur le choix des caractéristiques. On propose d'utiliser l'information chromatique de la peau, qui représente des caractéristiques importantes

1 Sur un visage quelconque, on constate que les valeurs prises par les pixels de peau ont une distribution d'allure gaussienne dans les espaces couleurs, ce qui incite naturellement à vouloir les décrire grâce à des mélanges de gaussiennes.

2 On préfère de se placer dans l'espace couleur **YCbCr** (pour ces excellentes résultats selon les travaux de recherche).

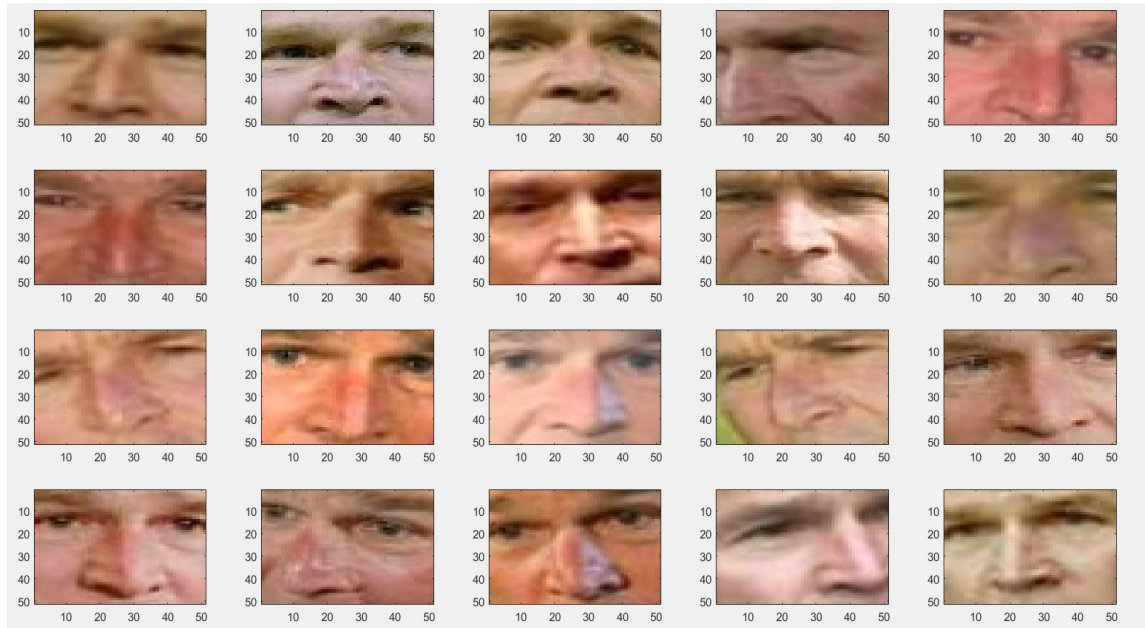


Figure 15.

3 L'histogramme représente les caractéristiques de la distribution de chrominances Cb Cr, selon les différents échantillons, cette histogramme peut être modélisé par une distribution gaussienne

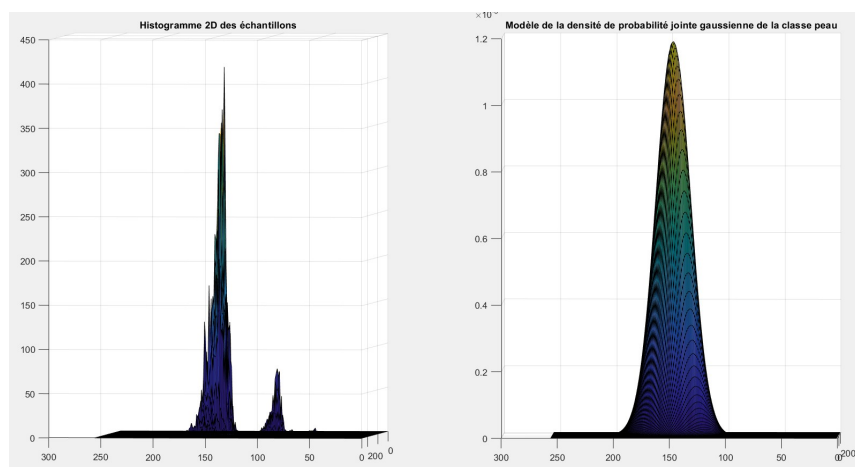


Figure 16. L'histogramme des caractéristiques de la distribution de chrominances Cb Cr,

- 4 Selon l'histogramme obtenue à la question précédente, le modèle de probabilité théorique le plus adapté pour modéliser les classes est le modèle des distributions gaussiennes. le code des questions précédentes:

```
%% Apprentissage d'un modèle de peau
database_dir='George_W_Bush';

fnames = dir(fullfile(database_dir, '*.jpg'));
num_files = size(fnames,1);
p=0.2;

% Image de test
%inputfile='George_W_Bush/George_W_Bush_0024.jpg';
inputfile='George_W_Bush/George_W_Bush_0027.jpg';
I_test=imread(inputfile);

figure;imshow(I_test);
imycbcr = rgb2ycbcr(I_test);

all_data=[];
for f = 1:20 %num_files to learn from all images
    im=strcat(database_dir,'/',fnames(f).name);
    % extraction d'une zone centrale de l'image
    cim=SelectPixelsCentre(im, p);
    % sortie : cim : composé de 3 matrices correspondant aux 3 plans couleurs
    % r=cim(:,1), g=cim(:,2), b=cim(:,3)
    subplot(4,5,f)
    imagesc(cim)

    % extraction des composantes chromatiques cb, cr de la zone précédemment extraite
    [cb, cr] = get_cbcr(cim);
    cb_data=reshape(cb,[size(cb,1)*size(cb,2),1]);
    cr_data=reshape(cr,[size(cr,1)*size(cr,2),1]);
    crcb_data=[cr_data cb_data];
    clear cb; clear cr; clear cb_data; clear cr_data;
    % on obtient un vecteur de taille size(cim,1)*size(cim,2) lignes * 2
    % colonnes (col1 = cr; col2 = cb)
    % chaque ligne de la matrice correspond à un pixel (un échantillon/individu de type
    % peau) caractérisé par 2 valeurs cr, cb filtrées.
    all_data=[all_data;crcb_data];
    % figure;imshow(cim);
end

%% Représentation des échantillons dans l'espace des paramètres
figure;
plot_hist2d(all_data(:,1), all_data(:,2));
```

2.3.3 Calcul des paramètres du modèle

A partir des échantillons d'apprentissage précédents, on estime maintenant les paramètres du modèle de la classe peau.

- 1 L'expression mathématique du modèle: (dans ce cas $d = 2$)

$$\forall x \in \mathbb{R}, p(\mathbf{x}|\omega_1) = \frac{1}{2\pi|\boldsymbol{\Sigma}_1|} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_1^{-1}(\mathbf{x}-\boldsymbol{\mu}_1)}$$

- 2 calcul des paramètres et de densité

```
%% Modélisation des données par une gaussienne
%estimation des paramètres statistiques du modèle à partir des échantillons
mu1 = mean(all_data)';
C1 = cov(all_data);
C1_inv = C1^(-1);
dC1 = det(C1);
%calcul du modèle dans un plan 2D
modelegaussien = zeros(256);
for r = 0:255
    for b = 0:255
        x = [r; b];
        % calcul de la vraisemblance de chaque pixel
        modelegaussien(r+1,b+1) = GaussLikelihood(x,mu1,C1_inv,dC1);%likelihood=vrai-
semblance
    end
end
```

- 3 La matrice de covariance obtenue est de taille 2×2 :

$$\begin{pmatrix} 243.171 & 115.2879 \\ 115.2879 & 128.8945 \end{pmatrix}$$

La matrice de covariance montre que les deux variables aléatoires des deux classes sont corrélés.

- 4

```
%représentation de l'histogramme 2D
subplot(1,2,1);
plot_hist2d(all_data(:,1), all_data(:,2))
title('Histogramme 2D des échantillons')
%représentation du modèle obtenu à partir de l'histogramme 2D
subplot(1,2,2);
surf(modelegaussien);
title('Modèle de la densité de probabilité jointe gaussienne de la classe peau')
```

2.3.4 Application du modèle à une image test

On souhaite étudier la pertinence du modèle appris précédemment en l'appliquant sur des images d'entrée. Pour une image d'entrée (non utilisée pour l'apprentissage), on souhaite regarder pour chaque pixel s la valeur de la probabilité $p(\mathbf{x}_s|\omega_1)$ en utilisant la loi estimée précédemment.

1

```
%% Application du modèle à des images test

inputfile='George_W_Bush/George_W_Bush_0024.jpg';
%inputfile='George_W_Bush/George_W_Bush_0027.jpg';
I_test=imread(inputfile);

figure;imshow(I_test);
imycbcr = rgb2ycbcr(I_test);

[m,n,l] = size(I_test);
%initialisation d'une matrice 2D de la taille de l'image à traiter
pxnskin = zeros(m,n);
for i = 1:m
    for j = 1:n
        %extraction des caractéristiques d'un échantillon (=pixel)
        cr = double(imycbcr(i,j,3));
        cb = double(imycbcr(i,j,2));
        x=[cr; cb];
        % calcul de la vraisemblance de chaque pixel
        pxnskin(i,j)=GaussLikelihood(x,mu2,C2_inv,dC2);
    end
end

%filtrage moyen pour lisser les valeurs
lpf= 1/9*ones(3);
pxnskin = filter2(lpf,pxnskin);
%normalisation des valeurs de vraisemblance par la valeur max
pxnskin = pxnskin./max(max(pxnskin));

%affichage de l'image résultat
figure;
subplot(1,2,1);
imshow(I_test, [0 1]);
title('Image RGB initiale')
subplot(1,2,2);
imshow(pxnskin, [0 1]);
title('Image p(x/nskin)')
```

la densité de probabilité des pixels de la peau est élevée par rapport à la densité de la deuxième classe.

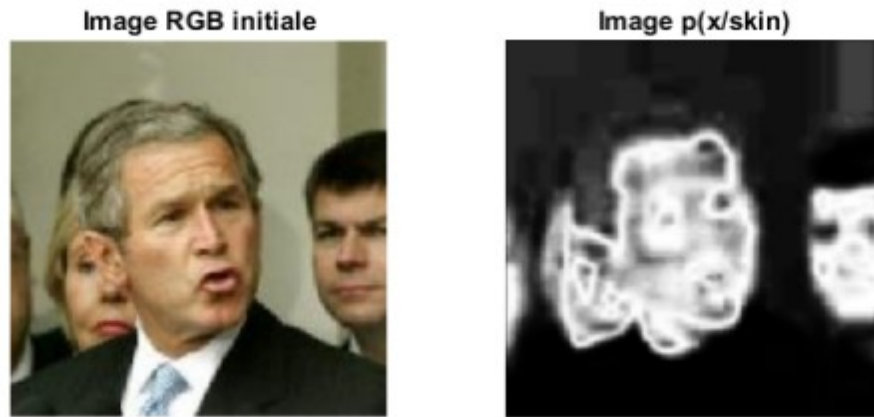


Figure 17. résultat de la partie 2.3.4

2 La distribution des pixels de la peau est bien séparable à la distribution des pixels de « non-peau », donc on peut considérer ce modèle de probabilité comme pertinent.

3 Par seuillage avec un seuil adapté à la distribution des pixels de la peau, le calcul du ce seuil est basé sur la partie d'apprentissage du modèle.

2.4 Partie 2 - Vers un classifieur bayésien

2.4.1 Classe “non-peau”

Dans la partie 1, nous avons établi un modèle gaussien $p(\mathbf{x}|\omega_1)$, avec $\omega_1 = \text{peau}$ pour modéliser la couleur de peau à partir des données d'apprentissage. Le traitement d'une image consiste à décider pour un pixel donné s'il s'agit d'un pixel peau ou non. Dans le cas d'un classifieur bayésien optimal, cela revient à décider la classe “peau” si $p(\omega_1|\mathbf{x}) > p(\omega_2|\mathbf{x})$, avec $\omega_2 = \text{non-peau}$. On choisit la classe “non-peau” dans le cas contraire.

1 On a

$$\forall x \in \mathbb{R} p(\omega_1|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_1)p(\omega_1)}{p(\mathbf{x}|\omega_1)p(\omega_1) + p(\mathbf{x}|\omega_2)p(\omega_2)}$$

2 Pour $p(\omega_1) = p(\omega_2) = \frac{1}{2}$, l'expression précédente devient

$$\forall x \in \mathbb{R} p(\omega_1|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_1) + p(\mathbf{x}|\omega_2)}$$

3 On propose de partir de même modélisation faite dans la partie 1, Pour agrandir les dimensions de sous image extraits on reprend une valeur de p proche de 1, pour tenir compte aussi de la partie qui ne correspond pas au visage (peau).

4 On s'inspire de la première partie, on choisie de modéliser $p(\omega_2|\mathbf{x})$ par un modèle gaussienne définie par:

$$\forall x \in \mathbb{R} p(\omega_2|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_2)}{p(\mathbf{x}|\omega_1) + p(\mathbf{x}|\omega_2)}$$

5

Le code de cette partie:

```
%% Classification bayésienne
% Probabilités a priori P1 = 0.5 ;
P2 = 1 - P1;
im_class = zeros(m,n);
for i = 1:m
for j = 1:n
%extraction des caractéristiques d'un échantillon (=pixel)
cr = double(imycbcr(i,j,3));
cb = double(imycbcr(i,j,2));
x=[cr; cb];
% calcul de la fonction discriminante
g1 = GaussLikelihood(x,mu1,C1_inv,dC2)*P1; % Fonction discriminante classe 1
g2 = GaussLikelihood(x,mu2,C2_inv,dC2)*P2; % Fonction discriminante classe 2
if(g1 > g2)
im_class(i,j) = 1;
end
end
end
figure
subplot(1,2,1); imshow(I_test, [0 1]); title('Image RGB initiale');
subplot(1,2,2); imshow(im_class, [0 1]); title('Image classification');
```

6 Après l'exécution du code précédent on obtient:

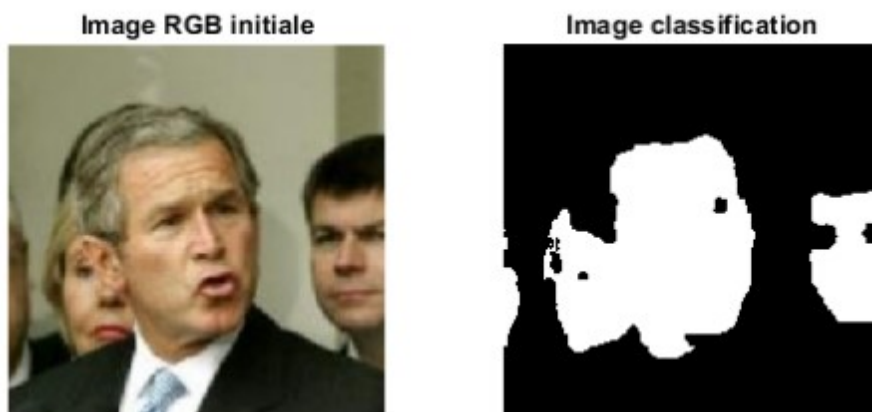


Figure 18. résultat de la partie 2.4.1

La considération du deuxième classe « non-peau » nous a permis de segmenter les masques des visages de manière pertinente.