

Introduction à la vision par ordinateur

TP1 : Détection de visages

Romain Catajar - romain.catajar@student.ecp.fr

20 mars 2016

Résumé

Présentation d'implémentations de différentes méthodes de détection de peau ou de visage et analyse des résultats. Le code a été réalisé avec l'aide de deux autres étudiants : Léo Hoummady et Thibault Kaspi.

4 Approche 1 : Detection de la peau

4.1 Avant de commencer

1. Que pensez vous de ce type d'approches, que l'on pourra qualifier d'approches colorimétriques, pour la détection de visages ? La détection de la peau est-elle suffisante ? Quelles en sont les limites ?

Cette approche me semble être limitée. En effet, même si on imagine pouvoir détecter parfaitement les pixels de peau, cela ne suffit pas à détecter le visage (le visage n'est pas forcément la seule peau apparaissant sur une image). De plus, je ne suis pas convaincu qu'une approche colorimétrique est suffisante pour détecter efficacement les pixels de peau. En effet la couleur des pixels de peau peut varier grandement d'une image à une autre en fonction de plusieurs critères (exposition et luminosité de la photo, ethnicité de la personne sur la photo, ...). Il me semble qu'une approche efficace devrait prendre en compte le contexte autour des pixels pour dire s'il s'agit de peau ou non (par exemple, en regardant le gradient de l'image, la couleur des pixels de peau devant a priori varier peu).

2. Dans la mise en oeuvre de ce type d'approches, quelles sont les principales questions à se poser ?
 - Qualité et variété des images utilisées (différents éclairages et couleurs de peau par exemple)
 - Choix du modèle de séparation entre peau et non peau
 - Choix de l'espace de couleur

4.2 A vous de jouer

4.2.1 Base d'images

J'ai construit une base de 78 images a partir du dataset *Pratheepan Dataset*. Ces images sont ensuite séparées en deux groupe, un groupe "d'entraînement" de 52 images et un groupe de test de 26 images.

4.2.2 Une première méthode simple

Les résultats obtenues sont présenté dans le tableau ci dessous.

Dataset	Taux de bonne détection		Taux de mauvaise détection	
	<i>Moyenne</i>	<i>Ecart type</i>	<i>Moyenne</i>	<i>Ecart type</i>
<i>Entrainement</i>	0.87	0.14	0.28	0.20
<i>Test</i>	0.91	0.09	0.16	0.15

TABLE 1 – Résultats de la méthode de Peer et Al.

Globalement le taux de bonne détection est acceptable, mais celui de mauvaise détection est trop élevée (son ecart type également). La règle semble être trop permissive.

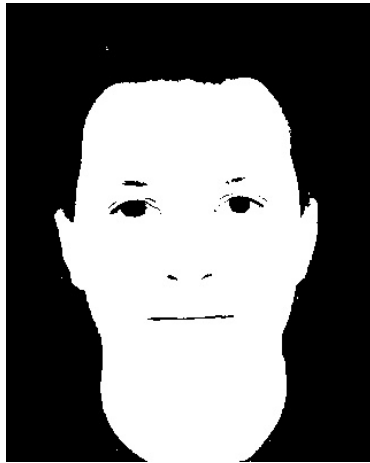


(a) Résultat



(b) Original

FIGURE 1 – Méthode de Peer et Al. : Exemple de mauvaise détection où l'arrière plan est détecté comme de la peau



(a) Résultat



(b) Original

FIGURE 2 – Méthode de Peer et Al. : Exemple de bonne détection



(a) Résultat

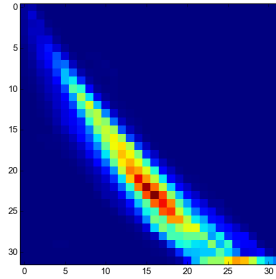


(b) Original

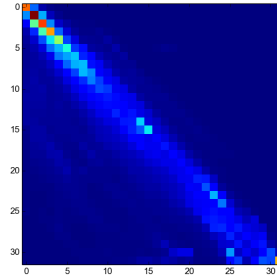
FIGURE 3 – Méthode de Peer et Al. : Exemple de mauvaise détection

4.2.3 Approche non paramétrique : Histogramme et modèle de peau

J'ai construit les histogramme a partir du dataset d'entrainement pour les espaces RGB, LAB et HSV.

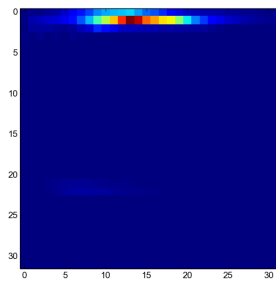


(a) peau

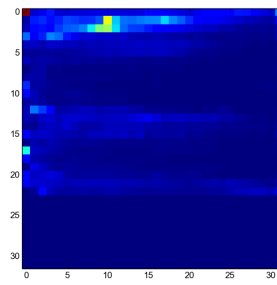


(b) non peau

FIGURE 4 – Histogramme RGB (sur R et G)

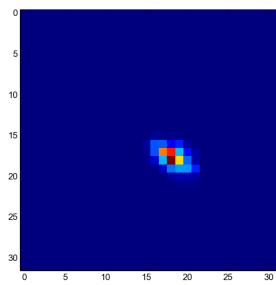


(a) peau

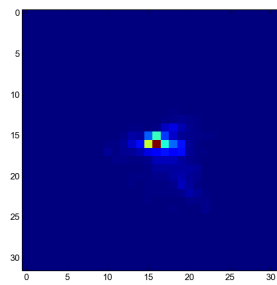


(b) non peau

FIGURE 5 – Histogramme HSV (sur H et S)



(a) peau



(b) non peau

FIGURE 6 – Histogramme LAB (sur a et b)

Les résultats sont présentés ci dessous.

Dataset	Taux de bonne détection		Taux de mauvaise détection	
	<i>Moyenne</i>	<i>Ecart type</i>	<i>Moyenne</i>	<i>Ecart type</i>
<i>Entrainement</i>	0.85	0.19	0.19	0.15
<i>Test</i>	0.86	0.20	0.13	0.14

TABLE 2 – Résultats avec les histogrammes RGB

Dataset	Taux de bonne détection		Taux de mauvaise détection	
	<i>Moyenne</i>	<i>Ecart type</i>	<i>Moyenne</i>	<i>Ecart type</i>
<i>Entrainement</i>	0.73	0.27	0.11	0.11
<i>Test</i>	0.73	0.28	0.08	0.10

TABLE 3 – Résultats avec les histogrammes LAB

Dataset	Taux de bonne détection		Taux de mauvaise détection	
	<i>Moyenne</i>	<i>Ecart type</i>	<i>Moyenne</i>	<i>Ecart type</i>
<i>Entrainement</i>	0.86	0.17	0.25	0.11
<i>Test</i>	0.89	0.15	0.21	0.15

TABLE 4 – Résultats avec les histogrammes HSV

Les résultats montre que le choix de l'espace de couleur influe sur la qualité du detecteur. L'espace HSV donne le meilleur taux de bonne detection, mais le moins bon taux de mauvaise detection. Globalement, la qualité des résultats est similaire à la méthode précédente et montre bien les différences entre espace de couleurs.

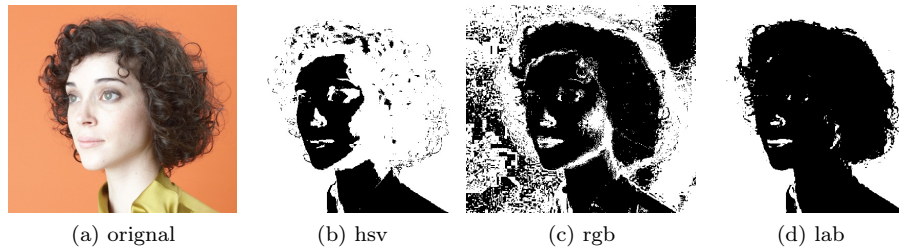


FIGURE 7 – Résultats avec histogrammes

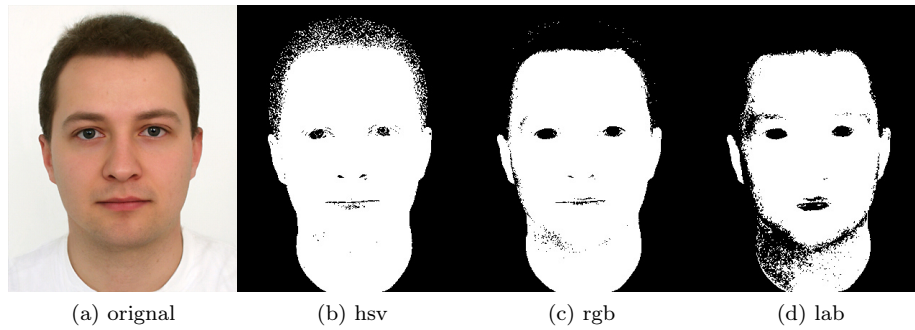


FIGURE 8 – Résultats avec histogrammes

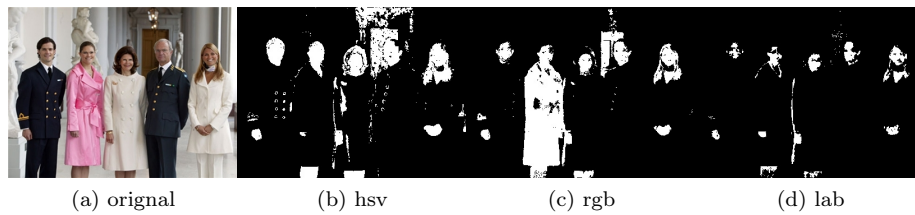


FIGURE 9 – Résultats avec histogrammes

4.2.4 Méthode de Bayes

J'utilise les histogrammes de la partie précédente pour mon modèle de Bayes. Pour déterminer le meilleur seuil, je cherche à maximiser le taux de bonne détection tout en minimisant celui de mauvaise détection. Pour cela, j'ai fait le choix de tester différentes valeurs de seuil en cherchant à maximiser leur différence. J'obtiens les valeurs de seuil suivante :

— *RGB* : 0.2

— *LAB* : 0.15

— *HSV* : 0.15

Des seuils plus faibles permettent d'augmenter le taux de bonne détection, mais rapidement le taux de mauvaise détection grimpe au dessus des 20%. Les résultats présentés ci dessous sont basés sur ces seuils.

Dataset	Taux de bonne détection		Taux de mauvaise détection	
	<i>Moyenne</i>	<i>Ecart type</i>	<i>Moyenne</i>	<i>Ecart type</i>
<i>Entrainement</i>	0.82	0.20	0.16	0.12
<i>Test</i>	0.82	0.21	0.12	0.13

TABLE 5 – Résultats avec la méthode de Bayes (RGB, seuil 0.2)

Dataset	Taux de bonne détection		Taux de mauvaise détection	
	<i>Moyenne</i>	<i>Ecart type</i>	<i>Moyenne</i>	<i>Ecart type</i>
<i>Entrainement</i>	0.74	0.26	0.12	0.11
<i>Test</i>	0.73	0.28	0.08	0.10

TABLE 6 – Résultats avec la méthode de Bayes (LAB, seuil 0.15)

Dataset	Taux de bonne détection		Taux de mauvaise détection	
	<i>Moyenne</i>	<i>Ecart type</i>	<i>Moyenne</i>	<i>Ecart type</i>
<i>Entrainement</i>	0.88	0.15	0.26	0.21
<i>Test</i>	0.91	0.11	0.22	0.16

TABLE 7 – Résultats avec la méthode de Bayes (HSV, seuil 0.15)

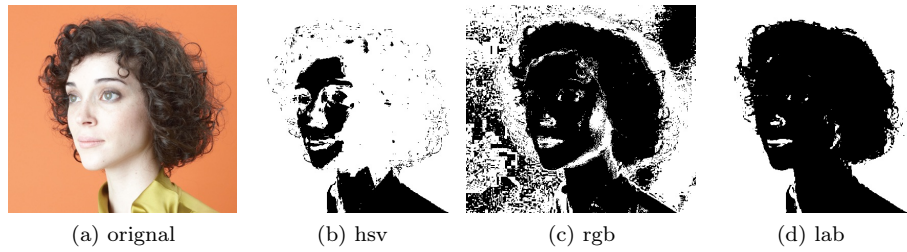


FIGURE 10 – Résultats avec la méthode de Bayes

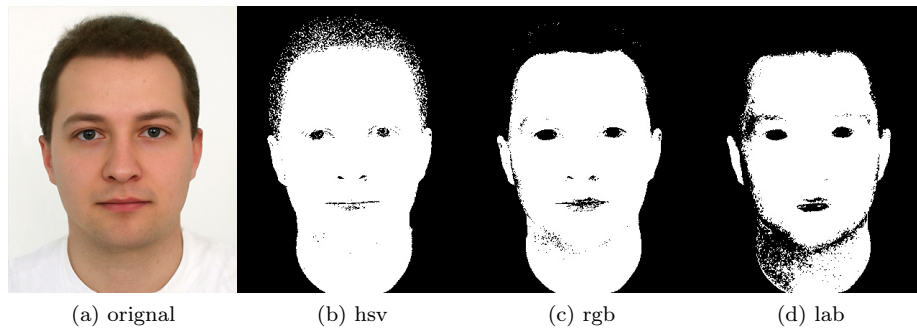


FIGURE 11 – Résultats avec la méthode de Bayes

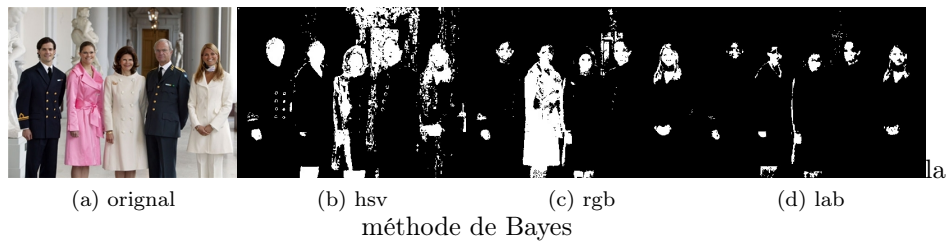


FIGURE 12 – Résultats avec la méthode de Bayes

4.2.6 Discussion

Globalement les différents modèles ont des résultats similaires. Je pense qu'il serait possible de les affiner en utilisant plus d'images différentes pour créer les histogrammes. On pourrait également affiner la qualité de la détection en combinant la détection à une recherche de contour pour détecter les visages.

5 Détection de visages par l'approche de Viola Jones

5.1 Etude de la methode

La méthode est implémenté avec les `haar_cascade` de `opencv`. J'ai testé différents paramètres pour le détecteur. Finalement j'obtiens les meilleurs résultats avec un léger scaling de l'image (10%), une taille minimale de détection de 30 pixels par 30 pixels et un nombre minimal de voisin pour que la détection soit valide de 5.

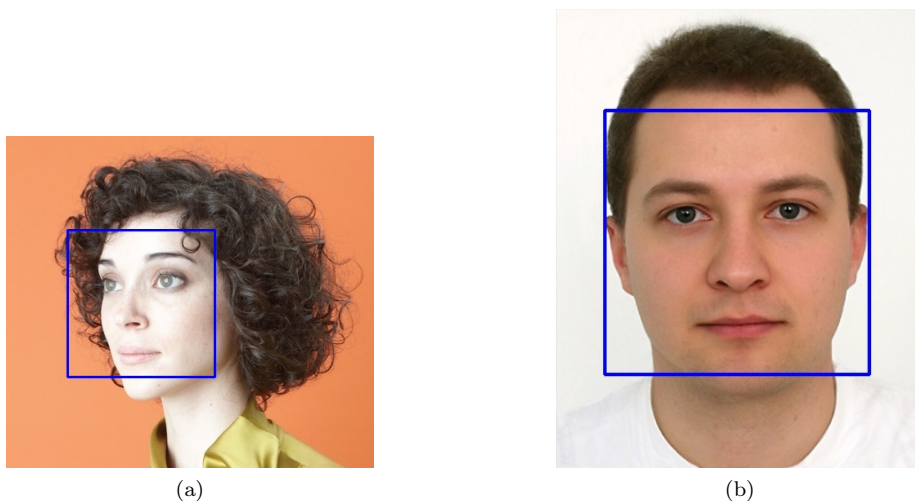
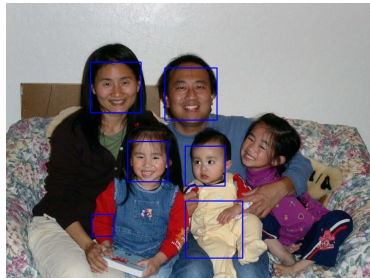


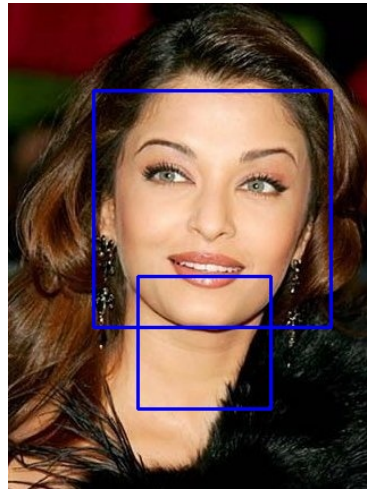
FIGURE 13 – Méthode de Viola Jones : exemple de bonne détection



FIGURE 14 – Méthode de Viola Jones : exemple de bonne détection sur photo de groupes



(a)



(b)

FIGURE 15 – Méthode de Viola Jones : exemple de mauvaise détection
Globalement la méthode donne de très bon résultats. Cependant comme on peut le voir dans les exemples, la détection est parfois un peu trop permissive et détecte des faux visages.

5.2 Discussion : pour aller plus loin !

Différents articles récents présentent des modèles de détection de visage utilisant des réseaux de neurones profonds (*deep learning*) et obtiennent de très bons résultats.

- Deep Learning Face Representation by Joint Identification-Verification - *Y Sun, Y Chen, X Wang, X Tang* (2014)
- Deep Convolutional Network Cascade for Facial Point Detection - *Yi Sun, Xiaogang Wang, Xiaoou Tang* (2013)