

Compte-rendu projet individuel 96 : les yeux nous trahissent ?

Auteurs : Elliot VANEGUE et Gaëtan DEFLANDRE

Encadrant : Marius BILASCO, Benjamin ALLAERT et José MENNESSON

19 mars 2015

Remerciement

Résumé

Abstract

Table des matières

1	Introduction	6
1.1	Contexte	6
1.2	Problématique de l'existant	7
1.3	Objectif du projet	7
2	Application existante	8
2.1	Architecture	8
2.2	Reconnaissance du visage : Viola et Jones	8
2.3	Suivi des yeux	8
3	Recherche de solution	9
3.1	L'algorithme de Canny	9
3.2	Normalisation et égalisation d'histogramme	10
3.3	Les modèles Colorimétrique	11
3.4	Détection des contours et calcul du barycentre de la forme	13
3.5	L'algorithme de Gabor	13
4	Implémentation de la solution	14
	Conclusion	15
	Annexes	16
	Références	17

Table des figures

1	Exemple de caractéristiques pseudo-Haar utilisé pour l'algorithme Viola et Jones	8
2	Image de test et résultat de l'algorithme de Canny avec une moyenne des pixels .	9
3	Image de test et résultat de l'algorithme de Canny avec division de l'image en 16	9
4	Image de test et résultat de l'algorithme de Canny avec dilatation puis érosion des sous-images	10
5	Algorithme de Canny après normalisation de l'histogramme de l'image de gris . .	10
6	Algorithme de Canny après égalisation de l'histogramme de l'image de gris . . .	10
7	Décomposition du modèle HSV	11
8	Résultat du masque après application du masque	11
9	Décomposition du modèle HSV avec une couleur de peau plus foncée	11
10	Décomposition du modèle YUV	12
11	Décomposition du modèle YUV avec une couleur de peau plus foncée	12
12	Résultats après traitement sur la chrominance 1	12
13	Première localisation du centre de l'oeil avec traitement sur le modèle YUV . . .	12

1 Introduction

Durant nos études de master informatique à l'université de Lille 1, nous avons l'occasion de participer à un projet proposé par une équipe de recherche. Cette expérience a pour but de nous faire découvrir le milieu de la recherche.

L'application sur laquelle nous allons travailler permet de localiser le visage d'un individu au travers d'un flux vidéo afin de reconnaître sur celui-ci différente émotion comme la fatigue ou encore l'intérêt de la personne pour quelque chose. Ce procédé se démocratise et il est de plus en plus utilisé dans des applications dans le domaine :

- de la sécurité, ce qui permettra de détecter des comportements inhabituel d'un individu.
- du loisir, pour les jeux vidéo ou encore pour une interaction plus intuitif avec un ordinateur.
- de la publicité, pour que celle-ci corresponde au besoin des individu.

Ce type d'application utilise différent type de technologie, mais celle-ci sont souvent très intrusifs et demande à l'utilisateur de mettre un casque ou tout autre appareil du même type. Les applications plus moderne effectue des algorithmes de reconnaissance de forme, tel que l'algorithme de viola et Jones, dans le but de détecter un visage à partir de simple caméra. Cela permet de mieu intégrer ce type d'application afin qu'il n'y est aucune contrainte pour l'utilisateur.

L'objectif de l'application sur lequel nous travaillons est de détecter les mouvements du visage de l'utilisateur afin d'effectuer de la reconnaissance d'émotions lors d'application du type e-learning¹. Ce procédé permettra de détecter si un cours intéresse ou non les élèves afin de pouvoir l'améliorer. Si on voit sur une partie du cour que de nombreux élèves ont présentés des signes de fatigue ou qu'ils n'écoutaient plus, il sera alors possible aux enseignant de modifier cette partie.

1.1 Contexte

Pour la réalisation de ce projet, nous travaillons avec l'équipe FOX qui étudie l'analyse du mouvement à partir de vidéos. Leurs recherches portent sur l'extraction du comportement humain depuis les flux vidéo. Leurs travaux sont divisés en quatre grands domaines : le regard, qui est la partie sur laquelle nous travaillons, l'événement, l'émotion et la reconnaissance de personnes. La grande majorité de leurs travaux sont des applications temps réel, ce qui permet d'avoir un niveau de réactivité très élevé.

Le projet sur lequel nous travaillons est basé sur les travaux d'anciens étudiants qui se sont concentrés sur la détection de visage et des yeux. Une première approche de la reconnaissance d'émotion sur une vidéo a été réalisé par les étudiant précédent, cependant elle n'est pas finalisé. Le projet est une application temps réel, dont les flux vidéo peuvent provenir de vidéos enregistrées ou d'une caméra type webcam. Nos travaux pourront être utilisé par l'équipe dans d'autre application de reconnaissance d'émotions sur lesquels ils effectue leur recherche.

1. formation en ligne

1.2 Problématique de l'existant

Les algorithmes de localisation du visage utilisés dans le projet sont très performant. Cependant, ils ont quelques limitations lorsqu'ils sont utilisés sur des applications temps réel. De nombreux cas ne sont pas traités dans ce type d'algorithme comme par exemple l'orientation du visage.

Actuellement l'application n'arrive pas à suivre un visage dont les yeux sont fermés. De plus, les points représentant le centre des yeux ne sont pas parfaitement stables, ce qui gêne le bon déroulement de l'application.

1.3 Objectif du projet

Pour corriger ce défaut, nous cherchons à extraire les yeux du reste du visage, afin de retrouver des points fixes, nous permettant de recentrer les points calculés auparavant. Ensuite, l'objectif est de stabiliser les points pris en compte pour la localisation du visage avec les informations que nous avons récupérées. Une fois que ces régions seront stabilisées, le reste de l'application normalisera le visage et cela permettra d'avoir des résultats beaucoup plus fiables lors de la reconnaissance d'émotions.

2 Application existante

2.1 Architecture

2.2 Reconnaissance du visage : Viola et Jones

L'application est divisée en deux parties. La première recherche le visage grâce à l'algorithme de Viola et Jones et la seconde recherche les yeux dans la région délimitée précédemment.

L'algorithme de Viola et Jones est une méthode qui a été créée pour la reconnaissance de visage dans une image. Cette méthode s'est par la suite généralisée à toutes sortes d'objets. L'algorithme nécessite une base de connaissances composée des caractéristiques de l'objet recherché. L'algorithme de Viola et Jones nécessite un apprentissage supervisé, c'est à dire que l'algorithme a besoin de données représentant l'objet à détecter pour classifier les caractéristiques de celui-ci.

Cette algorithme est basé sur des caractéristiques pseudo-Haar qui crée des masques rectangulaires et adjacentes dans différente zone de l'image. Chaque masque calcule l'intensité des pixels qu'il contient, puis l'algorithme fait la différence entre les masque blanc et les masque noir. Cette méthode va permettre de détecter des contours ou des changements de texture.

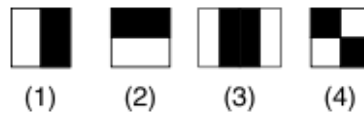


FIGURE 1 – Exemple de caractéristiques pseudo-Haar utilisé pour l'algorithme Viola et Jones

Pour améliorer les performances de leur algorithme, Viola et Jones utilisent la méthode Adaboost. Son principe est de sélectionner les caractéristiques les plus performantes pour la détection de l'objet.

2.3 Suivi des yeux

3 Recherche de solution

3.1 L'algorithme de Canny

3.1.1 Version de base

Le filtre de Canny a été créé en 1986 dans le but d'améliorer les résultat du filtre de Sobel. Le principe du filtre est d'utiliser deux filtres, un filtre haut et un filtre bas. L'algorithme commence par sélectionner les pixels supérieur au seuil haut, puis recherche à partir de chaque pixels au dessus du seuil haut les pixels qui sont au dessus du seuil bas. Ainsi on voit que cet algorithme prend en compte deux caractéristiques, l'intensité et la direction des gradients.

Nous utilisons cette algorithme dans les régions autour des yeux afin de délimiter le contour de chaque oeil. Le résultat permet de faire ressortir l'oeil, mais aussi les sourcils et certaines ombres présentent dans le creu de l'oeil. Ce filtre permet donc de détecter les contours des yeux, mais le résultat comporte soit beaucoup de bruit, soit trop peu de détails. Dans un premier temps, afin d'obtenir un seuil adaptatif, nous prenons la moyenne des pixels.

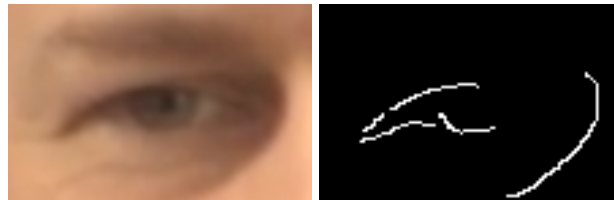


FIGURE 2 – Image de test et résultat de l'algorithme de Canny avec une moyenne des pixels

On voit que cette méthode, n'est pas optimal et supprime beaucoup trop d'information.

3.1.2 Avec une moyenne de pixels sur des parties d'image

Plutôt, que de prendre la moyenne sur l'image de la zone périoculaire, nous segmentons l'image en plusieurs zones dans lequel nous calculons la moyenne et ou nous effectuons l'algorithme de Canny.

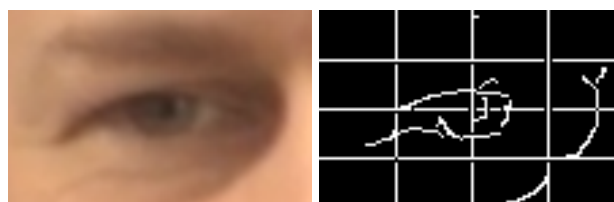


FIGURE 3 – Image de test et résultat de l'algorithme de Canny avec division de l'image en 16

Avec cette méthode, nous obtenons une image bruité avec les bords de chaque rectangle utilisé pour le calcul précédent, car l'algorithme de Canny ne prend pas en compte les bords de l'image qu'il traite.

3.1.3 Ouverture des sous-images

Afin de supprimer le bruit présent dans le résultat précédent nous effectuons une ouverture de chaque sous-image.

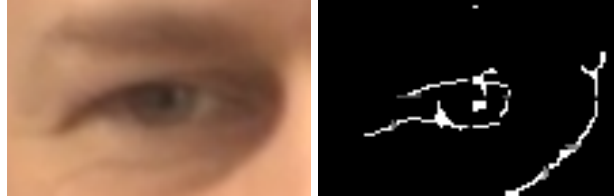


FIGURE 4 – Image de test et résultat de l'algorithme de Canny avec dilatation puis érosion des sous-images

Grâce à ce procédé, nous obtenons une image sans la grille que nous avons précédemment. Cependant, le bruit dû aux ombres et au ride ne permet pas encore d'effectuer un algorithme de détection de forme sur l'œil.

3.2 Normalisation et égalisation d'histogramme



FIGURE 5 – Algorithme de Canny après normalisation de l'histogramme de l'image de gris



FIGURE 6 – Algorithme de Canny après égalisation de l'histogramme de l'image de gris

Quelque soit les traitements effectués au préalable sur l'image avant l'utilisation de l'algorithme de Canny sur une image de gris, nous n'obtenons aucun résultat. Pour supprimer les ombres présentes sur les images que nous traitons, nous décidons d'utiliser un autre modèle colorimétrique.

3.3 Les modèles Colorimétrique

3.3.1 Le modèle HSV

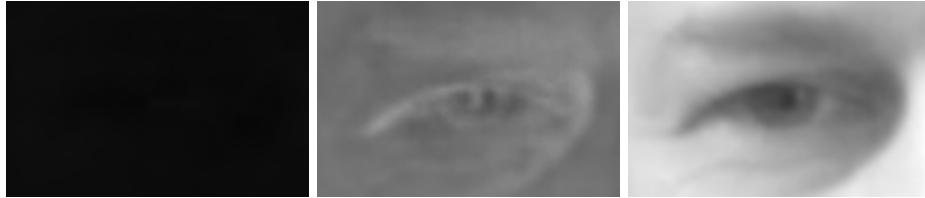


FIGURE 7 – Décomposition du modèle HSV

solution -> calcul d'un masque pour réduire l'image pris en compte par canny étape de construction du masque :

- égalisation des trois composante de l'image de départ
- passage de l'image en hsv
- égalisation de la composante saturation
- érosion de l'image



FIGURE 8 – Résultat du masque après application du masque

saturation -> probleme sur la couleur de peau.

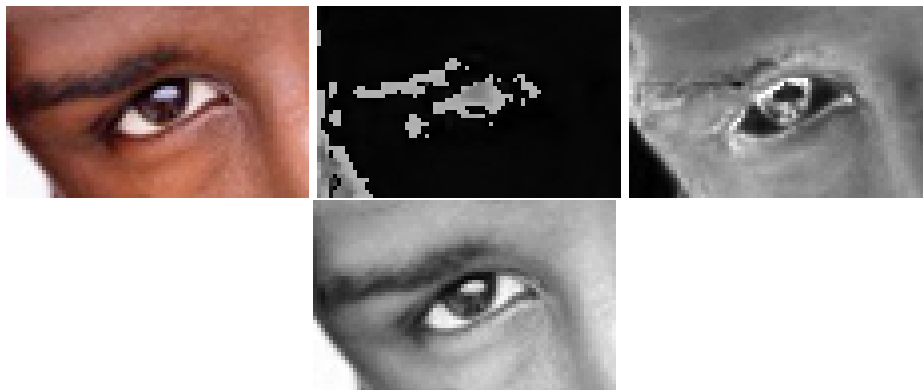


FIGURE 9 – Décomposition du modèle HSV avec une couleur de peau plus foncée

3.3.2 Le modèle YUV

calcule RGB vers YUV

$$Y = 0.299R + 0.587G + 0.114B$$

$$U = -0.147R - 0.289G + 0.436B = 0.492(B - Y)$$

$$V = 0.615R - 0.515G - 0.100B = 0.877(R - Y)$$



FIGURE 10 – Décomposition du modèle YUV

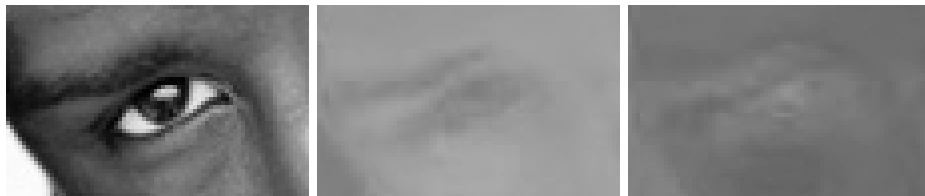


FIGURE 11 – Décomposition du modèle YUV avec une couleur de peau plus foncée



FIGURE 12 – Résulata après traitement sur la chominance 1

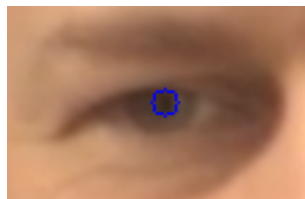


FIGURE 13 – Première localisation du centre de l'oeil avec traitement sur le modèle YUV

problème -> instable et ne fonctionnement pas sur toutes les personnes

3.4 Détection des contours et calcul du barycentre de la forme

3.5 L'algorithme de Gabor

4 Implémentation de la solution

Conclusion

Annexes

Références