



Master Informatique

Reconnaissance des mouvements de la main

Rapport

en vue de la validation de l'UE Initiation à la recherche

Étudiants : Victor DALLÉ
Claire KURTH

Encadrante : Madame BOLTCHEVA

Décharge de responsabilité

L'Université de Lorraine n'entend donner ni approbation ni improbation aux opinions émises dans ce rapport, ces opinions devant être considérées comme propres à leurs auteurs.

Remerciements

Table des matières

| | |
|---|----------|
| Introduction | 1 |
| 1 Rappel du sujet et encadrement | 2 |
| 1.1 Rappel du sujet | 2 |
| 1.2 Encadrement | 2 |
| 2 État de l’art | 3 |
| 2.1 Article de départ | 3 |
| 2.2 Médiapipe | 3 |
| 2.3 Classificateur classique Haar-cascade | 3 |
| 3 Reconnaissance de la main sans classificateur | 4 |
| 3.1 Méthodologie | 4 |
| 3.2 Expériences et résultats | 4 |
| 3.3 Conclusion | 4 |
| 4 Reconnaissance de la main avec classificateur classique Haar-cascade | 5 |
| 4.1 Entraîner un classificateur Haar-cascade | 5 |
| Conclusion | 5 |
| Annexes | 6 |
| Bibliographie | 6 |

Introduction

De nos jours, la vision par ordinateur est un domaine en plein essor. La reconnaissance de gestes fait partie intégrante de ce domaine et à ce titre, incarne une révolution dans la manière dont les utilisateurs interagissent avec les systèmes informatiques. Cette technologie est en effet en train de transformer la façon dont nous interagissons avec les machines. Cette avancée offre des opportunités novatrices dans des domaines tels que l'interaction entre l'homme et la machine, la réalité augmentée ou encore l'accessibilité numérique. De nombreuses techniques existent déjà pour permettre la détection des mains. MediaPipe de Google [src] utilise le machine learning pour entraîner un modèle qui détecte les segments composant la main. D'autres travaux ont été réalisés comme ceux de l'équipe du professeur Kalpana Joshi [src] qui utilise les angles formés entre 2 doigts pour détecter la forme de la main (nombre de doigts, main ouverte ou fermée).

Contrairement aux interfaces traditionnelles basées sur le clavier et la souris, la reconnaissance des gestes permet aux utilisateurs de communiquer plus simplement avec les ordinateurs. Ils peuvent désormais avoir recours à leurs mains ou à leur corps pour contrôler les applications, ou encore naviguer dans des environnements virtuels. Cette approche favorise une expérience utilisateur plus immersive et ergonomique, ouvrant ainsi de nouvelles perspectives dans des domaines variés tels que le divertissement interactif, l'éducation, ou encore la médecine. Comme dit précédemment, la reconnaissance des mouvements joue un rôle crucial dans l'accessibilité numérique en permettant à des personnes porteuses d'un handicap physique ou moteur de pouvoir communiquer et d'interagir avec des outils numériques plus facilement. En effet, cela permet de passer outre les obstacles liés à l'utilisation des outils traditionnels (tels que le clavier, la souris, la télécommande . . .) grâce à la simple utilisation de mouvements du corps. Cette nouvelle manière d'interagir avec un système numérique est déjà utilisée dans plusieurs domaines notamment le sport avec des applications de coaching personnel qui permettent de suivre les mouvements de l'utilisateur et ainsi lui donner des conseils pour améliorer sa technique, ou encore sa posture. Ce nouveau concept d'interaction permet également de pouvoir contrôler des appareils tels que des téléviseurs où par un simple geste, nous pouvons par exemple gérer le son ou changer de chaîne.

Dans ce contexte, ce projet vise à se questionner vis-à-vis d'un système de reconnaissance des mouvements de la main à l'aide d'un classificateur classique Haar-cascade et de le comparer avec un système de reconnaissance de la main sans classificateur. L'objectif principal est de concevoir un système capable de détecter et de classifier différents gestes de la main effectués par l'utilisateur, tels que le poing fermé, ou alors la main ouverte avec un certains nombre de doigts levés. Ces gestes seront ensuite associés à différentes actions telles que le lancement d'applications ou encore l'ouverture de sites web. Pour réaliser ce projet nous utiliserons principalement la bibliothèque OpenCV. Nous parlerons dans un premier temps plus en détail des techniques de reconnaissance de gestes actuellement utilisées, puis nous expliciterons les notions techniques ainsi que les méthodes que nous exploiterons. Nous verrons ensuite comment nous avons implémenté la reconnaissance de la main sans le classificateur classique Haar-Cascade puis avec ce classificateur.

1 Rappel du sujet et encadrement

1.1 Rappel du sujet

Le but de ce projet est d'implémenter un système de reconnaissance des mouvements de la main à l'aide d'un classificateur classique Haar-cascade. Le système doit reconnaître le geste de la main de l'utilisateur (poing, un doigt, deux, trois, quatre,...) et le mapper à différentes tâches telles que le lancement d'applications comme le bloc-notes, la peinture, et l'ouverture de sites web. Le système doit être mis en œuvre avec l'aide de la librairie de "Computer Vision" - OpenCV, comme dans l'article [src]. Une extension possible serait l'implémentation d'un système de détection des mouvements de la tête ou du corps, tout entier.

1.2 Encadrement

2 État de l’art

2.1 Article de départ

2.2 Médiapipe

Médiapipe est un framework open-source développé par Google permettant de construire des pipelines de traitement de données multimédia. Il propose des solutions pour la détection de la main, du visage ou encore de la pose. Il est basé sur des modèles de machine learning notamment grâce à de l’apprentissage via des réseaux de neurones.

[Image de la main détectée par Médiapipe]

2.3 Classificateur classique Haar-cascade

Le classificateur Haar-cascade est une méthode de détection d’objets dans une image introduit par Paul Viola et Michael Jones en 2001. Il est basé sur l’utilisation de caractéristiques de type Haar. Ces caractéristiques sont des fenêtres de taille fixe qui sont déplacées sur l’image et qui permettent de calculer la différence de luminosité entre les pixels de la fenêtre. Ces caractéristiques sont ensuite utilisées pour entraîner un classificateur qui permet de détecter des objets dans une image.

Les classificateurs Haar-cascade sont utilisés pour la détection de visages, de voitures, de plaques d’immatriculation, de piétons, de mains ou de tout autres objets. Ils sont très utilisés dans le domaine de la vision par ordinateur et sont très efficaces pour la détection d’objets dans une image.

3 Reconnaissance de la main sans classificateur

3.1 Méthodologie

3.2 Expériences et résultats

3.3 Conclusion

4 Reconnaissance de la main avec classificateur classique Haar-cascade

4.1 Entraîner un classificateur Haar-cascade

Pour entraîner un classificateur Haar-cascade, il faut tout d'abord collecter des images positives et négatives. Les images positives sont des images contenant l'objet que l'on souhaite détecter, tandis que les images négatives sont des images ne contenant pas l'objet. Il faut ensuite générer des fichiers de descriptions des images positives et négatives. Ces fichiers contiennent les coordonnées des objets à détecter dans les images positives. Enfin, il faut entraîner le classificateur à l'aide de ces fichiers de descriptions.

L'entraînement du classificateurs en lui-même se fait grâce à ce que l'on appelle des "features". Ces features sont des caractéristiques de l'objet que l'on souhaite détecter, ce sont des patterns de pixels qui permettent de distinguer l'objet des autres éléments de l'image. Il existe différents types de patterns :

- Les "edges" : ce sont des patterns qui permettent de détecter les contours de l'objet.
- Les "lines" : ce sont des patterns qui permettent de détecter les lignes de l'objet.
- Les "center-surrounder" : ce sont des patterns qui permettent de détecter les changements d'intensité entre le centre d'une région rectangulaire et le reste de la région. Cela permet de détecter des objets de forme particulière.

[Image de features Haar]

Ces features sont ensuite utilisées pour entraîner un classificateur qui permet de détecter l'objet dans une image. Le classificateur est entraîné à l'aide d'un algorithme de machine learning tel que AdaBoost qui permet de déterminer les features les plus pertinentes pour la détection de l'objet.

AdaBoost est un algorithme de machine learning qui permet de construire un classificateur fort à partir de plusieurs classificateurs faibles. Il fonctionne en entraînant itérativement des classificateurs faibles sur des sous-ensembles de données pondérés. À chaque itération, il met à jour les poids des données d'entraînement pour donner plus d'importance aux données mal classées. Il combine ensuite les classificateurs faibles pour former un classificateur fort qui est capable de détecter l'objet dans une image.

Conclusion

Annexes

Bibliographie

Glossaire

Déclaration sur l'honneur contre le plagiat

Résumé