

# **Rapport TP reconnaissance de formes : méthode structurelle**

Application : Interprétation des gestes de la main

Mohamed Ameziane TOUIL

Abdeslam DJEMAI

22 novembre 2020

## **A. Introduction :**

La reconnaissance des gestes de la main est une technologie qui devient de plus en plus pertinente, étant donné la croissance et la popularité récentes des technologies de réalité virtuelle et augmentée. C'est un aspect essentiel de l'interaction Homme Machine (IHM), qui permet une interaction bidirectionnelle dans les espaces virtuels.

Le geste est un symbole du comportement physique ou de l'expression émotionnelle d'une personne. Le geste peut être de deux types : le geste du corps et le geste de la main. En outre, les gestes peuvent être statiques et dynamiques. Pour ces derniers, chaque mouvement représente une signification. Les gestes de la main sont particulièrement importants. Les humains utilisent leurs doigts main gestes plus souvent pour pointer quelque chose, pour compter, pour exprimer ou souligner un point. Les mains d'une personne partagent des caractéristiques similaires à celles d'autres personnes. C'est le cas de leurs doigts la forme, le comptage et le mouvement.

Toutefois, de nombreux cas d'interaction de ce type sont actuellement limités à des utilisations spécialisées ou à des dispositifs plus coûteux tels que le Kinect et l'Oculus Rift. Dans ce rapport, nous exposons une méthode de reconnaissance des gestes de la main à l'aide d'un appareil plus courant, la caméra web pour ordinateur portable.

Plusieurs méthodes et techniques sont utilisées pour réaliser cette tâche, des méthodes telles que l'utilisation d'un réseaux de neurones artificiel pour entraîner un modèle statistique qui permet de reconnaître le geste de la main, cette méthode est parfois, surtout dans notre cas, est coûteuse de telle sorte qu'elle nécessite un ensemble de données d'entraînement qui est une tâche pas facile à réaliser, vu que c'est délicat de trouver un dataset étiqueté qui

répond à ce besoin. De plus elle nécessite une configuration matérielle coûteuse.

Dans le cadre de notre projet nous avons opter pour une méthode structurelle pour la réalisation de notre produit final que nous allons détailler plus tard.

## **B. Objectifs :**

Dans le cadre de ce projet nous allons nous familiariser avec les différentes notions nécessaires pour réaliser une solution informatique qui répond à un besoin, qui peuvent être cités ci-après :

B.A. Cadrage du problème et de ses spécifications.

B.B. Faire une recherche bibliographique sur les travaux existants et adaptation du contenu à notre besoin.

B.C. Proposition d'une solution relative à l'analyse des images et de vidéos.

B.D.Utilisation des bibliothèques et code sources techniques accessibles.

B.E. Réalisation d'un programme informatique exécutable sur des plateformes et systèmes recommandés.

B.F. Faire une perspective du travail réalisé et proposition d'une généralisation de la méthode proposée.

## **C. Implémentation :**

Pour l'implémentation de notre solution nous avons opté pour le langage de programmation Python qui est un langage de programmation interprété, multi-paradigme et multiplateformes, placé sur une license libre et fonctionne sur la plupart des plates-formes Informatiques, c'est un langage où la syntaxe est facile et séparée des mécanismes de bas niveau. Il est doté d'une énorme

base de bibliothèques spécialisées dans plusieurs domaines d'études, telle que OpenCV qui est une bibliothèque graphique libre, initialement développée par Intel, spécialisée dans le traitement d'images en temps réel, et Numpy qui est une extension de ce langage, destinée à manipuler des matrices ou tableaux multidimensionnels ainsi que des fonctions mathématiques opérant sur ces tableaux. Nous avons opté pour l'IDE Jupyter notebook qui est une application web utilisée pour programmer dans plus de 40 langages de programmation, dont Python, Julia, Ruby, R, ou encore Scala. Jupyter est une évolution du projet IPython. Jupyter permet de réaliser des calepins ou notebooks, c'est-à-dire des programmes contenant à la fois du texte en markdown et du code en Julia, Python, R... Ces calepins sont utilisés en science des données pour explorer et analyser des données.

#### **D. Conception de la solution :**

Dans cette section nous allons expliquer les différentes méthodes utilisées pour l'interprétation des gestes, passant par la récupération de donnée, la segmentation de l'image récupérée et enfin l'interprétation du geste.

Pour l'interprétation des gestes on a opté à la méthode des Enveloppes convexes qu'on détaillera plus tard.

#### D.A. Architecture :

Récupération des données :

- Utilisation de la caméra de PC pour capturer les images



Interprétation des gestes :

- Utilisation des Enveloppes convexes
- Calcul de « convex defects »
- Utilisation des « convex defects » pour le calcul d'angle entre les doigts



Segmentation de l'image :

- Transformation de l'image en niveau de gris
- Lissage de l'image avec le filtre gaussien
- Binarisation de l'image avec la méthode d'Otsu
- Récupération des contours depuis l'image seuillée

### **D.B. Segmentation :**

Pour la segmentation de la main depuis les images récupérées par la caméra on passe par les étapes suivantes :

- Convertir l'image en niveau de gris.
- Lissage de l'image avec le filtre gaussien pour l'élimination du bruit.
- Seuillage de l'image avec la méthode d'Otsu pour bien séparer la main du fond.
- Récupération des points de contours de la surface la plus grand dans notre image (dans notre cas on suppose que notre main sera la surface la plus grande de notre image).

### **Bruit Gaussien :**

La fonction de flou gaussien est obtenue en brouillant (lissant) une image à l'aide d'une fonction gaussienne pour réduire le niveau de bruit. Il peut être considéré comme un filtre passe-bas non uniforme qui préserve les basses fréquences spatiales et réduit le bruit et les détails négligeables d'une image. Il est généralement obtenu par la convolution d'une image avec un noyau gaussien. Le noyau gaussien rend l'image floue pour réduire la quantité de bruit et éliminer les taches dans l'image. Il est important d'éliminer les composantes à très haute fréquence qui dépassent celles associées au filtre de gradient utilisé, sinon, elles peuvent provoquer la détection de faux bords.

### **Ots'u:**

Un problème avec le seuillage simple est que vous devez spécifier manuellement la valeur du seuil. Nous pouvons vérifier manuellement la qualité d'un seuil en essayant différentes valeurs, mais c'est fastidieux et cela risque de se rompre dans le monde réel. Il nous faut donc un moyen de déterminer

automatiquement le seuil. La méthode de seuillage d'Otsu consiste à itérer toutes les valeurs de seuil possibles et à calculer une mesure de la dispersion des niveaux de pixels de chaque côté du seuil, c'est-à-dire les pixels qui se trouvent soit au premier plan soit à l'arrière-plan. L'objectif est de trouver la valeur seuil où la somme des écarts entre l'avant-plan et l'arrière-plan est à son minimum.

Les avantages d'une telle approche sont une segmentation précise en temps réel des caractéristiques et des performances éprouvées par rapport aux méthodes de segmentation par seuil.

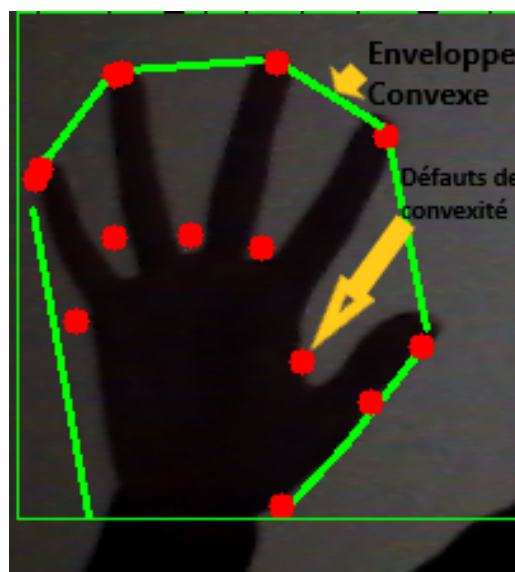
## **D.C. Interprétation de geste :**

### **D.C.A. Enveloppe Convexe**

L'enveloppe convexe peut être considérée comme une enveloppe dynamique et extensible qui s'enroule autour de l'objet d'intérêt. En géométrie, convexe d'une forme est le plus petit ensemble convexe qui la contient.

### **D.C.B. Défauts de convexité**

On parle de défaut de convexité si les courbes de convexes sont bombées à l'intérieur et dans ce cas on calcule les écarts maximums locaux de l'enveloppe convexe par rapport aux contours.



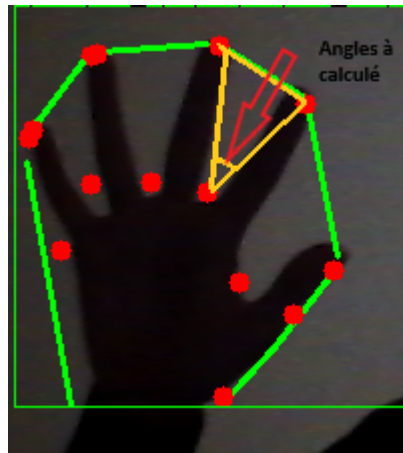
**Enveloppe convexe et défauts de convexité**

### D.C.C. Interprétation des gestes

Pour l'interprétation des différents gestes on a opté à la méthode des enveloppes convexes. Le choix de cette méthode est dû à sa performance pour l'interprétation des gestes de la main approuvée par les différents articles trouvés dans la littérature.

L'étape suivante résume l'algorithme de calcul de nombre de doigts implémenté dans notre solution :

- Calcule de l'enveloppe convexe en utilisant les points de contours.
- Calcule des défauts de convexités en utilisant l'enveloppe convexe et le contour de la main.
- Calcule des angles créés par les points de l'enveloppe convexe et les points des défauts de convexité.



#### Calcul des angles

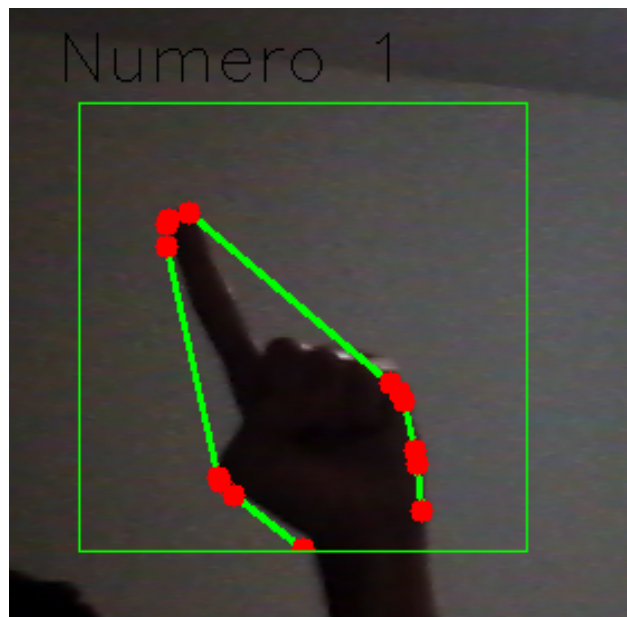
- Calcule du nombre d'angle qui sont inférieur à  $90^\circ$ .
- Nombre de doigts est égale alors au nombre d'angle inférieur à  $90^\circ + 1$ .



### E. Tests et Résultats :

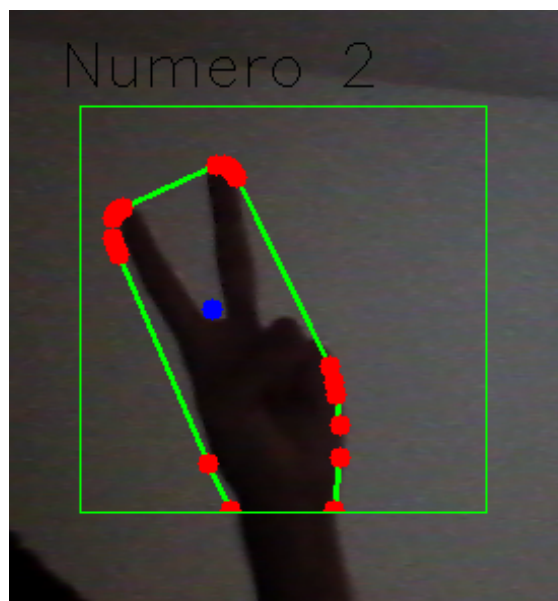
Pour la phase des tests de notre algorithme implémenté nous avons essayé les différents gestes de la main pour interpréter les chiffres de 1 jusqu'à 5 :

- Pour le chiffre 1 on n'aura pas d'angle de degré moins de  $90^\circ$  et la distance qui sépare le convexe et le point qui forme l'angle est supérieure



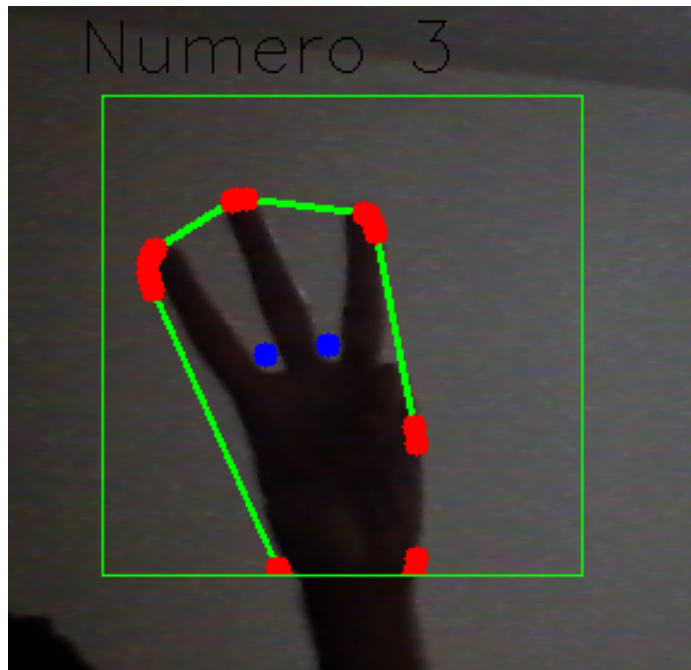
**Interprétation du numéro 1**

- Pour le chiffre 2 on aura un angle de degré moins de  $90^\circ$



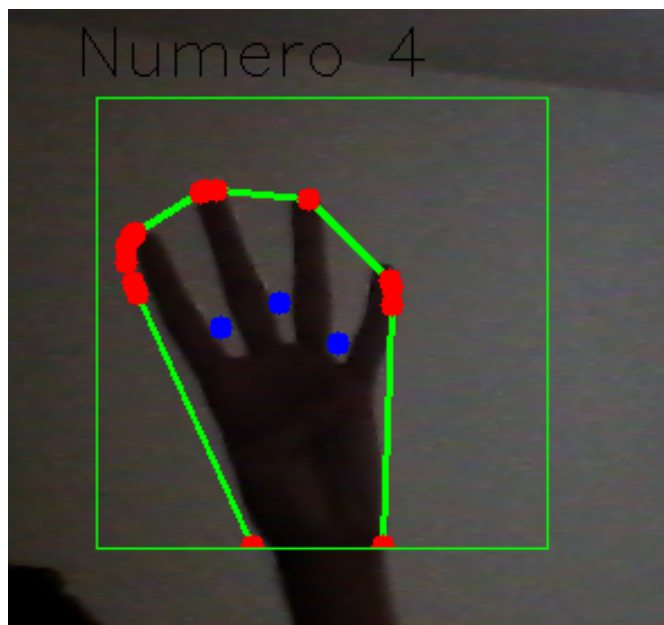
**Interprétation du numéro 2**

- Pour le chiffre 3 on aura deux angles de degré d'angle moins de  $90^\circ$



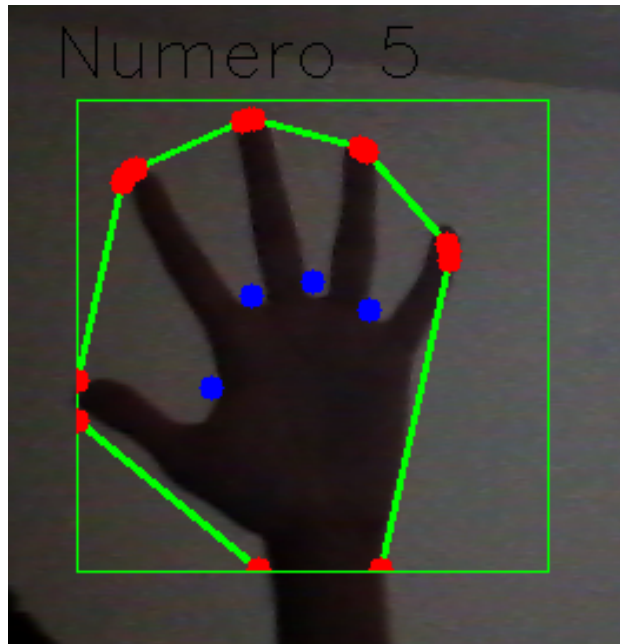
**Interprétation du numéro 3**

- Pour le chiffre 4 on aura trois angles de degré d'angle moins de  $90^\circ$



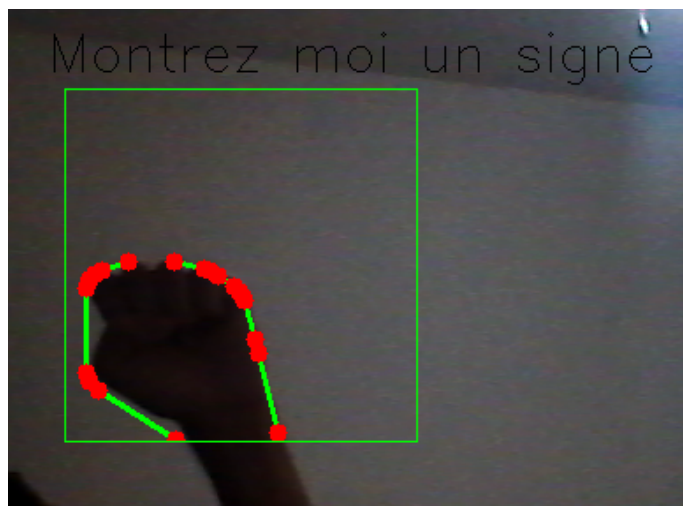
**Interprétation du numéro 4**

- Pour le chiffre 5 on aura quatre angles de degré d'angle moins de  $90^\circ$



### **Interprétation du numéro 5**

- Et dans le cas ou nous ne trouvons pas la présence d'un doigt et les angles détectés sont petits alors on demande à l'utilisateur de montrer un geste.



### **Interprétation de la non présence de chiffre**

## **F. Conclusion**

A la fin de ce projet nous avons pu nous familiariser avec les différents concepts de traitement des données visuelles et interprétations des signes de la main. Nous avons implémentée une méthode structurelle pour traduire les signes de la main à des nombres significatifs. Cette méthode pourra être généralisée de telle sorte à ce qu'on détecte la région d'intérêts (la main) sur n'importe quel fond, de plus trouver une approche qui permet de faire la traduction du langage de signes.

## G. Bibliographies

1. [JF16] MA Jayaram and Hasan Fleyeh. Convex hulls in image processing: A scoping review. American Journal of Intelligent Systems, 6(2) :48-58, 2016.
2. [Yeo] Zi Xian Justin Yeo. Hand recognition and gesture control using a laptop web-camera.
3. [https://docs.opencv.org/master/dd/d49/tutorial\\_py\\_contour\\_features.html](https://docs.opencv.org/master/dd/d49/tutorial_py_contour_features.html)
4. [https://docs.opencv.org/master/d3/dc0/group\\_imgproc\\_shape.html#ga014b28e56cb8854c0de4a211cb2be656](https://docs.opencv.org/master/d3/dc0/group_imgproc_shape.html#ga014b28e56cb8854c0de4a211cb2be656)
5. <http://vipulsharma20.blogspot.com/2015/03/gesture-recognition-using-opencv-python.html>
6. <https://gogul.dev/software/hand-gesture-recognition-p2>
7. [https://en.wikipedia.org/wiki/Convex\\_hull](https://en.wikipedia.org/wiki/Convex_hull)