

# Rapport de TP de Traitement d'image

## Baye Malick GNING

### 1 Introduction

On effectue dans ce TP une classification de formes avec en utilisant les descripteurs de Fourier et des paramètres géométriques sans avoir à extraire le contour de l'image. On utilisera les méthodes de classification des K-moyennes et des K plus proches voisins (KMeans et KNeighborsClassifier de scikit-learn) de Python.

### 2 Génération d'une base d'image

Une base d'image est générée à partir des images de la figure 1. On effectue une translation, une rotation et une homothétie aléatoires. On génère ainsi 100 images avec cette transformation affine.

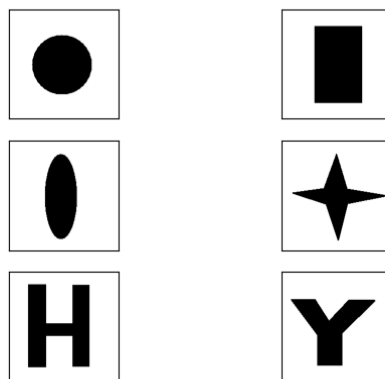


FIGURE 1 – Formes de référence

### 3 Obtention d'une image binaire et Extraction du contour de la forme

L'histogramme est la représentation de la distribution des pixels de l'image selon l'intensité. Ainsi on peut en déduire le seuil  $s$  à appliquer afin d'obtenir notre image binaire.

Soit une image  $f$  de dimension  $N \times M$ ,  $f(i, j)$  étant la valeur d'intensité d'un pixel de la ligne  $i$  et de la colonne  $j$ . Alors l'image binaire est obtenu comme suite :

$$g(i, j) = \begin{cases} 255 & \text{si } f(i, j) \geq s \\ 0 & \text{sinon} \end{cases}$$

ON calcule les coefficients descriptifs de Fourier et la fonction inverse qui permet d'approximer le contour. En effet, on représente le contour par une suite ordonnée de  $N$  points  $(x_m, y_m)$ . On construit ensuite une série de nombre complexe  $z_m = x_m + iy_m$ . Les descripteurs de Fourier sont :

$$a_k = \frac{1}{N} \sum_{m=0}^{N-1} (z_m - \bar{z}) \exp\left(-\frac{2\pi j k m}{N}\right)$$

avec  $\bar{z}$  la moyenne des nombres complexes.

La figure 2 contient l'image binaire de la première image de la base de données avec son histogramme pour avoir une idée du seuil à choisir, le contour détecté, le tracé des modules de Fourier, le contour tracé avec les descripteurs de Fourier et enfin une approximation avec la fonction inverse qui donne les nombres complexes.

On voit avec la figure 3 que lorsque le paramètre  $cm_{ax}$  augmente la forme approximée avec les nombres complexes inverses est moins étirée. Ainsi avec plus de descripteurs de Fourier on approxime une forme avec une homothétie amoindrie.

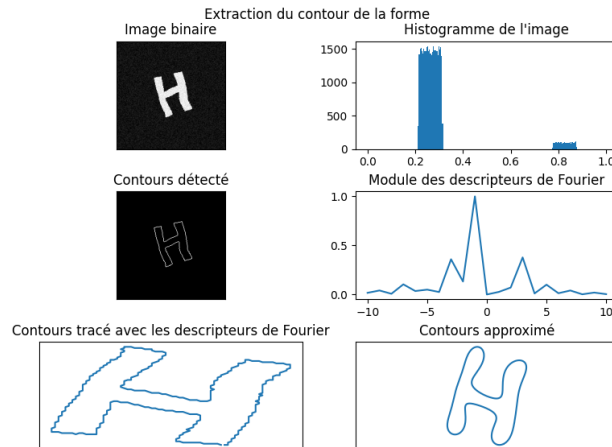


FIGURE 2 – Contours et descripteurs de Fourier

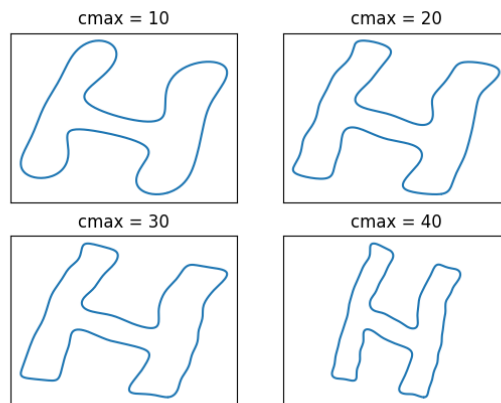


FIGURE 3 – Influence du paramètre  $cmax$

## 4 Autres paramètres géométriques

On utilise ici `regionprops` sur l'image binaire pour obtenir des paramètres sur les formes sans avoir à extraire un contour. On s'intéresse particulièrement aux moments Hu de l'image binaire.

Les moments de l'image sont une moyenne pondérée des intensités des pixels de l'image. Prenons un exemple simple pour comprendre la déclaration précédente.

Pour simplifier, considérons une image binaire à un seul canal  $I$ . L'intensité des pixels à l'emplacement  $(x, y)$  est donnée par  $I(x, y)$ . Remarque pour une image binaire  $I(x, y)$  peut prendre

la valeur 0 ou 1.

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y)$$

Ces moments sont souvent appelés moments bruts pour les distinguer des moments centraux.

Le centroïde d'un blob binaire est simplement son centre de masse. Le centre de gravité  $(\bar{x}, \bar{y})$  est calculé à l'aide de la formule suivante.

$$\bar{x} = \frac{M_{10}}{M_{00}}$$

$$\bar{y} = \frac{M_{01}}{M_{00}}$$

Les moments centraux sont très similaires aux moments d'image brute, sauf qu'on soustraie le centroïde de x et y dans la formule des moments.

$$\mu_{ij} = \sum_x \sum_y (x - \bar{x})^i (y - \bar{y})^j I(x, y)$$

Les moments centraux sont invariants par translation. On introduit les moments centraux normalisés, comme indiqué ci-dessous.

$$\eta_{ij} = \frac{\mu_{ij}}{\mu_{00}^{(i+j)/2+1}}$$

Les moments Hu (ou les invariants de moment Hu) sont un ensemble de 7 nombres calculés à l'aide de moments centraux invariants aux transformations d'image. Les 6 premiers moments se sont invariants à la translation, à l'échelle, à la rotation et à la réflexion. Alors que le signe du 7e moment change pour la réflexion de l'image. Les moments calculés sont :

$$h_0 = \mu_{20} + \mu_{02}$$

$$h_1 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2$$

$$h_2 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2$$

$$h_3 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2$$

$$h_4 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$h_5 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$h_6 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

Ces paramètres sont obtenues grâce à la fonction *regionprops* du module scikit-learn. On les utilise pour faire notre classification.

## 5 Analyse en composante principale

On réalise une analyse en composante principale sur les descripteurs de Fourier afin de réduire les éventuelles redondances d'informations avant de prendre leur module. L'analyse en composante principale des contours détectés pour chaque image réduit la taille du contours à deux points. Ici on aura  $cmx = 1$  et on verra que cette méthode d'analyse en composante principale ne donne pas de très bon résultats pour la classification.

## 6 Méthodes des K-Moyennes

La méthode des K-Moyennes recherche des clusters qui regroupe les données en essayant de séparer les échantillons en  $k$  groupes de variance égale ( $k$  clusters disjoints), en minimisant un critère connu sous le nom d'inertie ou somme des carrés intra-cluster. L'inertie est :

$$\sum_{i=0}^n \min_{\mu_j \in C} (\|x_i - \mu_j\|)$$

L'inertie n'est pas une métrique normalisée : on a simplement que les valeurs inférieures sont meilleures et que zéro est optimal. Mais dans les espaces de très grande dimension, les distances euclidiennes ont tendance à gonfler. L'exécution d'un algorithme de réduction de dimensionnalité tel que l'analyse en composantes principales (ACP) avant le clustering k-means peut atténuer ce problème et accélérer les calculs.

On applique l'algorithme des K-Moyennes pour les 7 paramètres de moments de chaque image calculés grâce à la méthode de *regionprops* et pour les modules des descripteurs de Fourier après avoir appliqué la l'analyse en composante principale aux contours.

Tout d'abord on introduit la méthode d'Elbow pour avoir une idée du nombre de clusters même si nous savons que nous devons entrer 6 comme paramètre pour le  $n$  de clusters. On calcule l'inertie obtenue après le clustering pour chaque nombre de cluster (ici entre 1 et 10). Il s'agit ici de la somme des carrés intra-cluster (WCSS : Within-Cluster Sum of Square), c'est-à-dire la somme de la distance carrée entre les points d'un cluster et le centroïde du cluster. Au vu des figures 4 et 5, la somme des carrés intra-cluster est bien plus faible pour les descripteurs de Fourier mais les résultats obtenus sont moins satisfaisants qu'avec les moments Hu. En effet un seul cluster regroupe la lettre H, l'étoile et même le rectangle.

Pour ce qui est des résultats avec les moments Hu, les résultats sont assez satisfaisants. Seuls quelques clusters ont une forme en plus. La classification est très bonne avec :

- cluster des rectangles : uniquement de 9 rectangles
- cluster des Y : uniquement 18 Y
- cluster des ellipses : uniquement 20 ellipses
- cluster des H : 18 H et 2 étoiles
- cluster des cercles : uniquement 15 cercles
- cluster des étoiles : 16 étoiles, 1 H et 1 Y

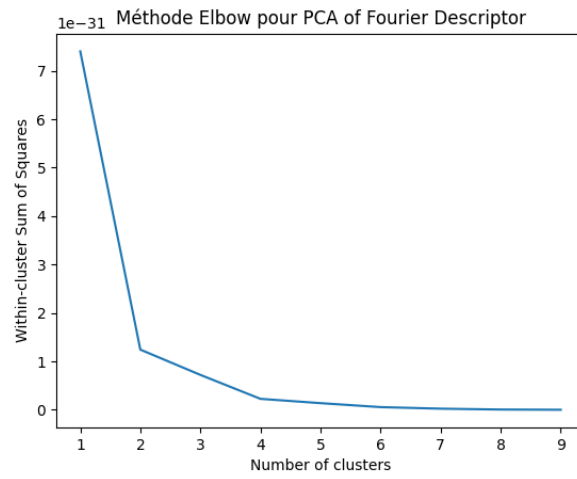


FIGURE 4 – Méthode Elbow pour descripteurs de Fourier

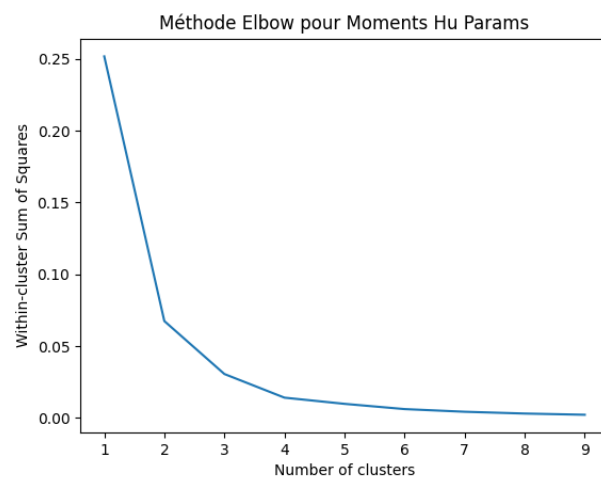


FIGURE 5 – Méthode Elbow pour Moments Hu Params

## 7 Méthode des K plus proches voisins

On passe à l'utilisation de la méthode des K plus proches voisins dont la classification est basée sur les voisins qui stocke simplement des instances des données d'apprentissage. La classification est calculée à partir d'un vote à la majorité simple des plus proches voisins de chaque point : un point de requête se voit attribuer la classe de données qui a le plus de représentants parmi les plus proches voisins du point.

Ici les résultats seront sous formes de matrices de confusion. Pour mettre en évidence les meilleurs résultats avec les moments Hu, on fait en plus la classification pour les descripteurs de Fourier. Chaque image a été labélisée pour l'apprentissage.

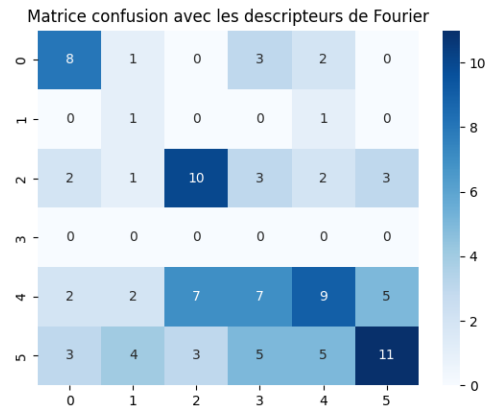


FIGURE 6 – Matrice de confusion pour les descripteurs de Fourier

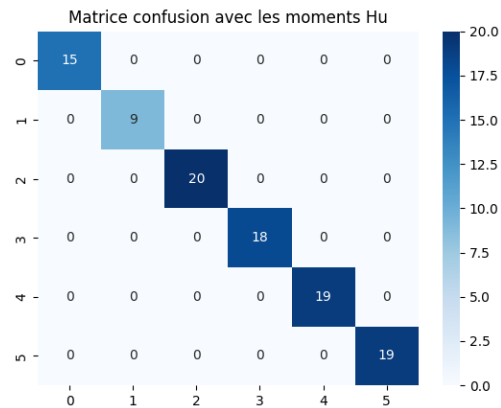


FIGURE 7 – Matrice de confusion pour les Moments Hu

## 8 Evaluation des performances

Un test a été effectué avec de nouveaux paramètres :  $N = 100$ ,  $\sigma = 0.07$ ,  $\Delta X = \Delta Y = 30$  et  $NX = NY = 15$ . On utilise le même classificateur qui a été entraîné sur la base d'entraînement et on le teste sur une nouvelle base de données. Dans ce cas nous voulons des résultats performants alors jusqu'à la fin du TP les paramètres utilisés seront les moments Hu pour la prédiction.

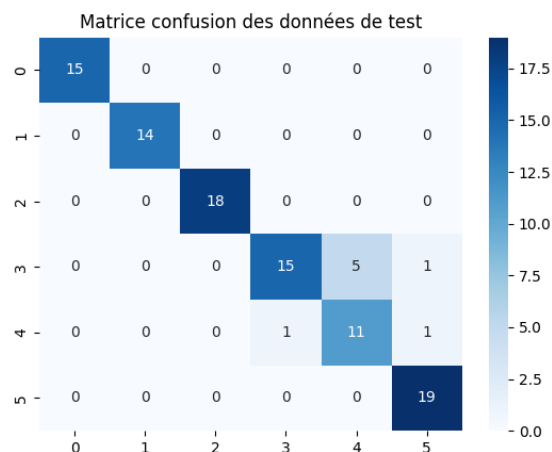


FIGURE 8 – Evaluation des performances

Les résultats sont satisfaisants. Pour quatre labels les prédictions sont parfaites. Il y a le clusters 4 qui a été prédits en 5 fois en 3 : il s'agit sûrement de l'étoile et de la lettre H.

## 9 Généralisation à d'autres formes : les chiffres

La généralisation est effectuée avec des chiffres comme images de référence. Les images sont obtenues avec l'application Procreate sur iPad ; les chiffres sont dessinés à la main. Voici en figure 9 nos nouvelles formes :

Ensuite on constitue une nouvelle base de données composée de 500 images dont la translation, la rotation et l'homothétie sont aléatoires. Il s'agit du même programme qu'en début de TP. Et enfin, on constitue 100 images de test pour tester notre classificateur. La figure 10 donne les résultats. La méthode des K plus proches voisins est très performante au vu de la matrice de confusion sur les données de test.



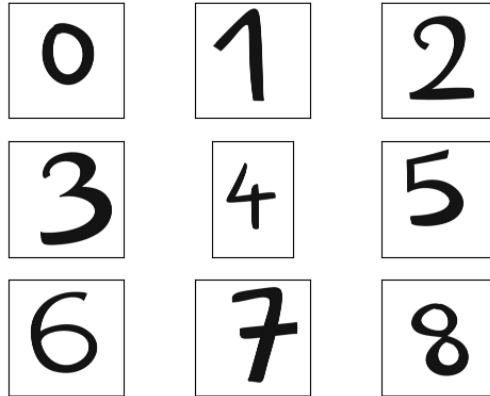


FIGURE 9 – Nouvelles formes

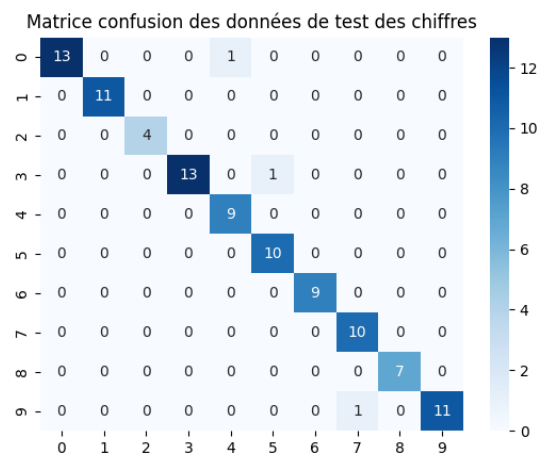


FIGURE 10 – Matrice de confusion pour les données test de chiffres

## 10 Conclusion

On a pu faire de la classification de formes avec les méthodes des K-Moyennes et des K plus proches voisins. La deuxième méthode de classification est très performante même si la première reste quand même bonne pour faire de la classification. Cependant elle reste limitée car l'inertie est assez élevée du fait de la distance euclidienne lorsque les paramètres sont nombreux. Notons également que nos meilleurs résultats ont été obtenus avec les moments Hu des images. Les descripteurs de Fourier sont limités du fait que on a que deux points après l'analyse en composante principale du contours détecté. En conclusion, les moments Hu sont bien meilleurs à prendre comme paramètres car invariants aux transformations des images et surtout ne dépendent pas des contours.