



Projet Calcul Scientifique et Analyse de Données

A quoi tu ressembles sans ton masque ?

Grégoire Honvault, Rachida Oussakel

Département Sciences du Numérique - Première année
2021-2022

Table des matières

1	Introduction	3
1.1	Contexte du projet	3
1.2	Objectif de cette première partie	3
2	Analyse en composantes principales des données sans masque	3
2.1	Eigenfaces	4
2.2	Projection des données sur les eigenfaces	5
3	Analyse en composantes principales de données avec masque	6
3.1	Eigenfaces	7
3.2	Projection des données sur les eigenfaces	7
4	Algorithme de la puissance itérée	8
5	Conclusion	9

Table des figures

1	Une base de visages	3
2	Eigenfaces	4
3	Visages à la première et dernière itération	5
4	Evolution du RMSE au fil des itérations	6
5	Les visages masqués	6
6	Eigenfaces associées aux visages masqués	7
7	Visages masqués à la première et dernière itération	8
8	Evolution du RMSE au fil des itérations sur les visages masqués	8

1 Introduction

1.1 Contexte du projet

Avec la pandémie de la Covid-19, nous avons appris à porter des masques quotidiennement. Désormais la moitié de notre visage est cachée nous laissant le regard pour seul moyen d'expression visible. Cette opération de cacher le visage s'apparente à un domaine en traitement d'images et de vidéos appelé "inpainting". En effet, les pixels correspondant à la zone du masque (modélisé par un rectangle) sont mis à 0. Et les eigenfaces permettent, entre autres, de restaurer la zone dégradée.

1.2 Objectif de cette première partie

Nous allons ici nous baser sur la méthode ACP (analyse en composantes principales) pour réduire la dimension des images et créer un système de reconnaissance grâce aux composantes principales.

2 Analyse en composantes principales des données sans masque

La figure 1 présente 16 images de 4 personnes dans des postures faciales différentes, avec 3 émotions.

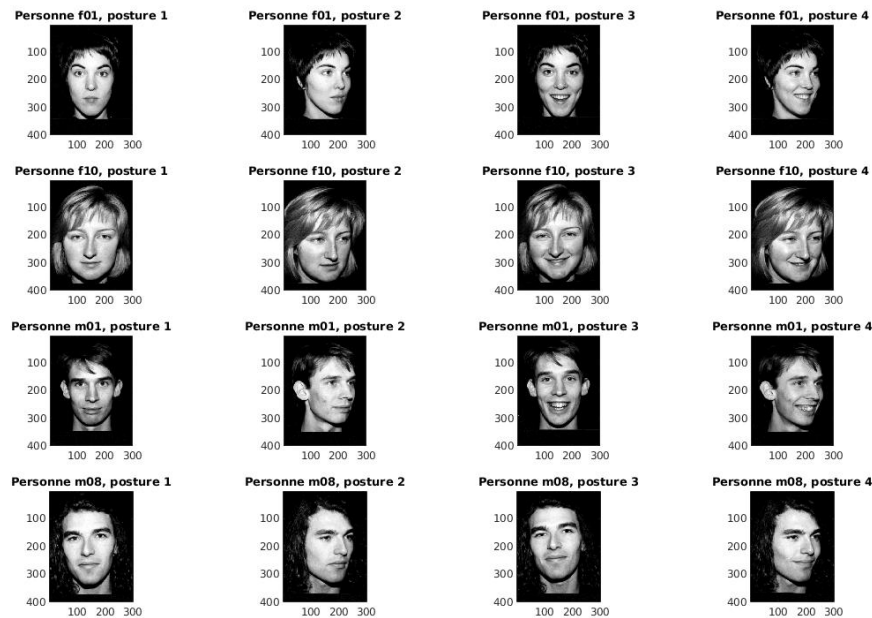


FIGURE 1 – Une base de visages

2.1 Eigenfaces

A partir de la base de données de visages, nous avons calculé l'individu moyen, puis les couples propres de la matrice de variance et covariance pour obtenir les axes principaux : les eigenfaces. Nous avons ensuite trié ces eigenfaces en se basant sur les valeurs propres, calculées avec la fonction eig de matlab, par ordre décroissant.

La figure 2 présente l'individu moyen et les eigenfaces récupérés.

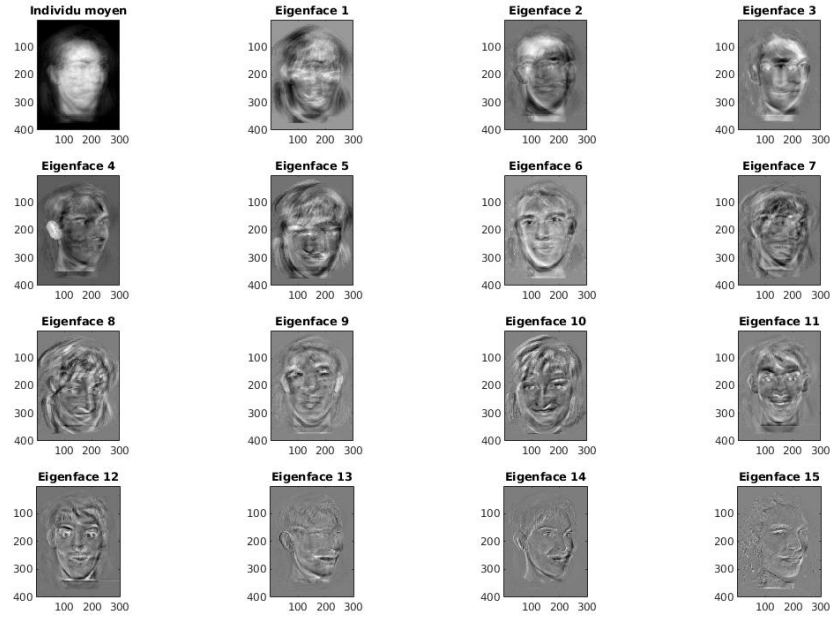


FIGURE 2 – Eigenfaces

2.2 Projection des données sur les eigenfaces

Une fois connues les eigenfaces, on peut calculer les composantes principales. On cherche ici à reconstruire les images d'apprentissage. On procède par itération. A l'itération n°q, on fait la reconstruction à l'aide des q premières eigenfaces et des q premières composantes principales, en partant de l'individu moyen.



FIGURE 3 – Visages à la première et dernière itération

On affiche également sur la figure 4 l'évolution, en fonction de q, de la racine carrée de l'erreur quadratique moyenne (Root Mean Square Error, ou RMSE) entre les images originales et les images reconstruites. Ainsi, on remarque que cet écart décroît au fil des itérations jusqu'à s'annuler pour $q=15$, où on retrouve bien les visages de départ.

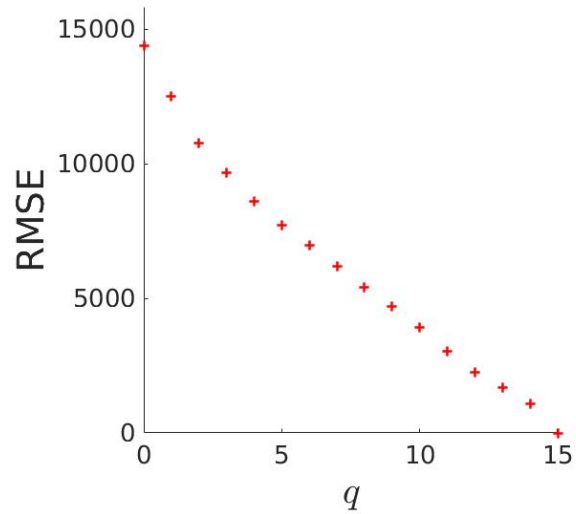


FIGURE 4 – Evolution du RMSE au fil des itérations

3 Analyse en composantes principales de données avec masque

La figure 5 présente les même images de personnes, avec des rectangles noirs représentant les masques sur leur visage.

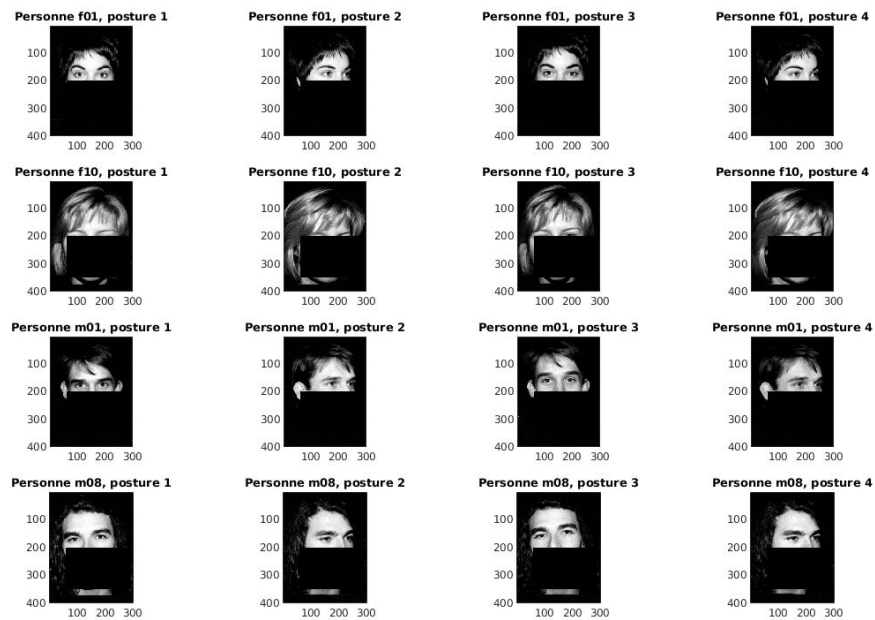


FIGURE 5 – Les visages masqués

3.1 Eigenfaces

En appliquant le même principe que 2.1 aux visages masqués, on obtient la figure 6

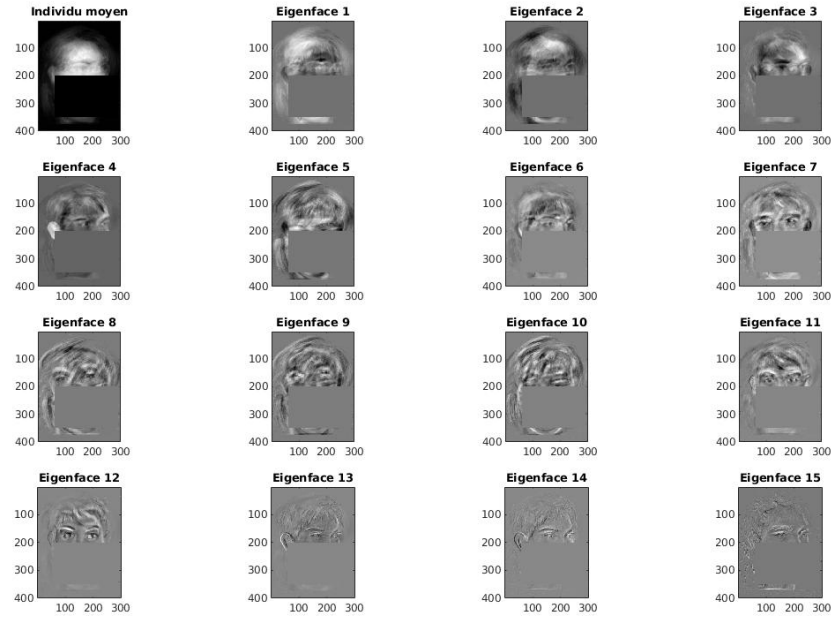
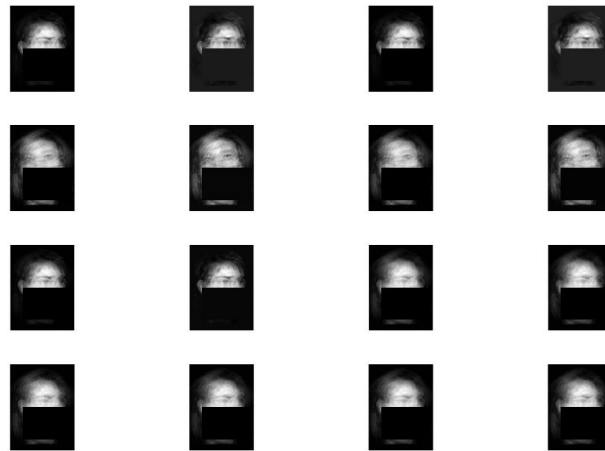


FIGURE 6 – Eigenfaces associées aux visages masqués

3.2 Projection des données sur les eigenfaces

De même pour la projection des visages masqués sur les eigenfaces.



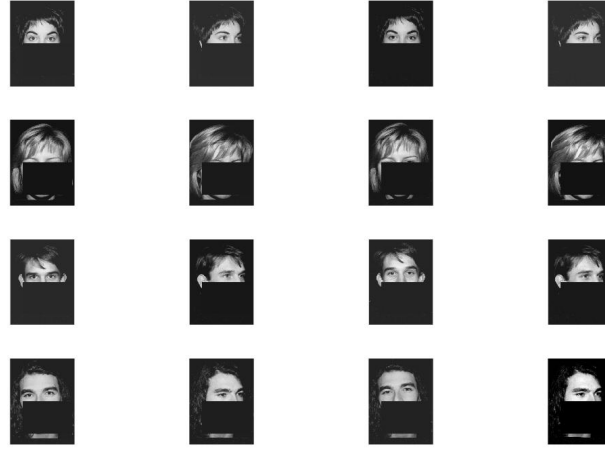


FIGURE 7 – Visages masqués à la première et dernière itération

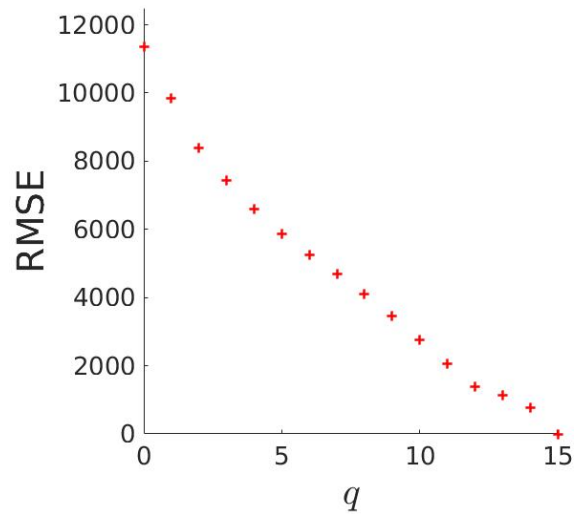


FIGURE 8 – Evolution du RMSE au fil des itérations sur les visages masqués

4 Algorithme de la puissance itérée

Dans cette partie au lieu d'utiliser la fonction eig de matlab, on implémente l'algorithme de la puissance itérée avec déflation qui permet d'obtenir le couple propre dominant.

Question 4 : Sens direct :

Soit λ une valeur propre non nulle de $H^\top H$ et X un vecteur propre associé ($X \neq 0$).

Alors $H^\top HX = \lambda X \Rightarrow HH^\top HX = \lambda HX$

On pose maintenant $Y = HX$ et procédons par l'absurde :

$Y = 0 \Rightarrow HX = 0 \Rightarrow \lambda X = 0 \Rightarrow X = 0$ (car $\lambda \neq 0$), absurde car $X \neq 0$. On a bien $HH^\top Y = \lambda Y$ et $Y \neq 0$. D'où Y vecteur propre de HH^\top de valeur propre associée λ .

Reciproquement :

Soit λ une valeur propre non nulle de HH^\top et X un vecteur propre associé ($X \neq 0$).

Alors $HH^\top X = \lambda X \Rightarrow H^\top HH^\top X = \lambda H^\top HX$

On pose $Y = H^\top X$. De la même façon :

$Y = 0 \Rightarrow H^\top X = 0 \Rightarrow \lambda X = 0 \Rightarrow X = 0$ (car $\lambda \neq 0$), absurde car $X \neq 0$. On a bien $HH^\top Y = \lambda Y$ et $Y \neq 0$. D'où Y vecteur propre de HH^\top de valeur propre associée λ .

Finalement, on a montré que $H^\top H$ et HH^\top ont les mêmes valeurs propres non nulles.

Question 6 : Il est plus utile en théorie d'utiliser la méthode de la puissance itérée pour calculer les éléments propres de \sum que la fonction « Eig ». En effet, eig retourne plus que les éléments propres nécessaires pour l'ACP (puisque'un des couples propres est éliminé par la suite),et nécessite de les trier dans l'ordre décroissant. La méthode de puissance itérée, elle, retourne les couples propres directement dans l'ordre voulu, et un par un, ce qui permet de se limiter aux éléments voulus. La fonction "eig" de matlab est donc une perte de ressources par rapport à la puissance itérée.

Question 7 : D'après les temps d'itération calculés,on en déduit que si on choisit d'utiliser la méthode de la puissance itérée avec déflation pour calculer les éléments propres de \sum ,il vaut mieux appliquer la méthode sur $A^\top A$ (la matrice de dimension minimale) pour minimiser le temps de calcul et la mémoire utilisée.

5 Conclusion

Dans cette partie du projet,nous avons appliqué l'ACP à partir de données (des images de visages avec/sans masques) de très grande taille pour pouvoir réduire la dimension des images .Ce qui va nous permettre par la suite de créer un système de reconnaissance grâce aux composantes principales pour retrouver le visage entier dans la base d'apprentissage le plus similaire au visage masqué et de permettre une reconstruction de la zone du masque.