

TP 2 : EigenFaces

L'objectif de ce TP est d'étudier les propriétés de la méthode de reconnaissance de visages par eigenfaces. **Au cours de ce TP, nous n'avons pas normalisé les valeurs des images, cependant les résultats sont les mêmes, seules les valeurs propres et les distances se retrouvent nettement supérieures.**

Chaque visage détient un vecteur de caractéristique censé être invariant pour une même personne. Il est donc possible de reconnaître ou différencier des personnes selon leur vecteur associé. Ainsi si l'on compare ce vecteur avec une base de données, on peut en déduire le visage le plus ressemblant à celui testé.

Plusieurs méthodes existent pour définir les caractéristiques d'un visage. Dans notre cas, nous analyserons une méthode de type "image" qui n'est autre que l'analyse par eigenface.

On extrait les caractéristiques de l'image par une méthode mathématique basée sur l'analyse en composante principale.

Principe de l'analyse par eigenfaces : Nous pouvons obtenir un visage spécifique à partir d'un visage moyen et d'une combinaison linéaire d'eigenfaces. Ces eigenfaces représentent un ensemble limité d'images représentant les caractéristiques spécifiques du visage d'une personne.

Ainsi une image de visage J s'exprime sous la forme suivante : $J = \Psi + \sum_h w_h u_h + \varepsilon$ où Ψ représente le visage moyen, u l'ensemble des eigenfaces et w les poids qui leur sont associés.

Question 1:

Chaque image est considérée comme un vecteur de taille égale au nombre de pixels. Chacune des images à disposition fait partie intégrante d'une matrice I afin de constituer la base de données. Sachant qu'un pixel est codé sur 8 bits et que la matrice I est de taille $n \times m$, avec n le nombre de pixels dans une image et m le nombre d'images dans notre base de données, la mémoire nécessaire pour stocker cette base de données sera de $8nm$ bits.

Question 2:

Comme nous l'avons abordé lors de l'introduction, le principe des eigenfaces repose en partie sur le visage moyen. Pour le calculer nous effectuons une simple moyenne de tous les visages mis à disposition dans notre base de données et nous obtenons le résultat suivant :



Visage moyen

Comme nous pouvons nous y attendre, l'image résultante prend la forme d'un visage flouté.

Question 3:

Par conséquent si l'on soustrait ce visage moyen à un visage de la base de données, nous devrions obtenir l'ensemble des visages centrés avec leur poids associés (plus ou moins l'erreur d'approximation).



Visages centrés du sujet 1

Comme nous pouvons le voir le visage moyen assombrit seulement le visage des sujets dans une zone relativement centrée. Il n'apporte pas ou peu de modification sur la forme du visage.

Comment obtenir les eigenfaces ?

Le principe de l'analyse en composante principales réside dans les vecteurs propres de la matrice de covariance AA^T où A représente l'ensemble des visages centrées de notre base de données.

Afin d'obtenir les vecteurs propres nous réalisons une Décomposition en valeurs singulières afin d'obtenir la matrice orthonormale U contenant les vecteurs propres et la matrice diagonale S contenant les valeurs propres associées.

Question 4:

La matrice U aura son importance dans la suite du tp puisque ces colonnes représentent les vecteurs propres de la matrice de covariance AA^T , soit les eigenfaces de la base de données. Ci-dessous un tableau contenant les valeurs propres associées aux premiers eigenfaces (ou vecteurs propres).

1.12674e+09	8.25826e+08	4.37721e+08	3.56966e+08	3.26956e+08
2.1515e+08	1.56583e+08	1.49152e+08	1.25551e+08	1.1535e+08

Tableau des premiers eigenfaces pour une base de données de 400 images

Comme on peut le voir, grâce aux valeurs propres ci-dessus, les eigenfaces sont triés au sein de la matrice U issu de la SVD. Plus les valeurs sont élevées, plus l'eigenface sera descriptif pour l'ensemble des visages de la base de données.



10 premiers eigenfaces pour une base de données de 400 images

Maintenant que nous avons les eigenfaces et le visage moyen, il nous faut calculer les coordonnées de l'image test dans notre sous-espace des visages de dimension K , qui est en d'autres mots notre base d'eigenfaces limité à K eigenfaces.

Pour un visage donné, nous calculons donc le produit scalaire de chaque eigenfaces avec le visage centré de celui-ci afin d'obtenir W les coordonnées de notre image test dans le sous-espace des visages. Grâce à celles-ci, nous pouvons calculer la projection de l'image test dans le sous-espace des visages par la formule suivante :

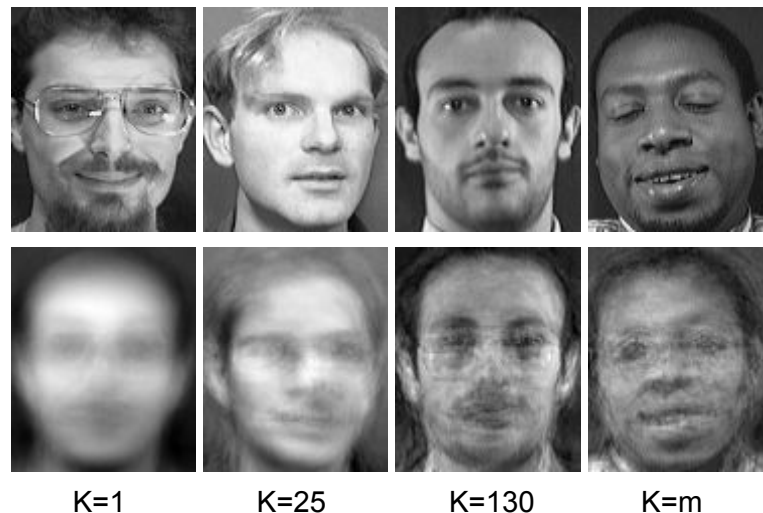
$$\mathbf{J_p} = \Psi + \sum_{k=1}^K W_k \mathbf{u}_k$$

L'image projetée représente l'image issue de la recombinaison d'une image à partir du visage moyen et des eigenfaces. Nous définissons une erreur de reconstruction comme la distance entre l'image projetée et un visage.

Question 8:

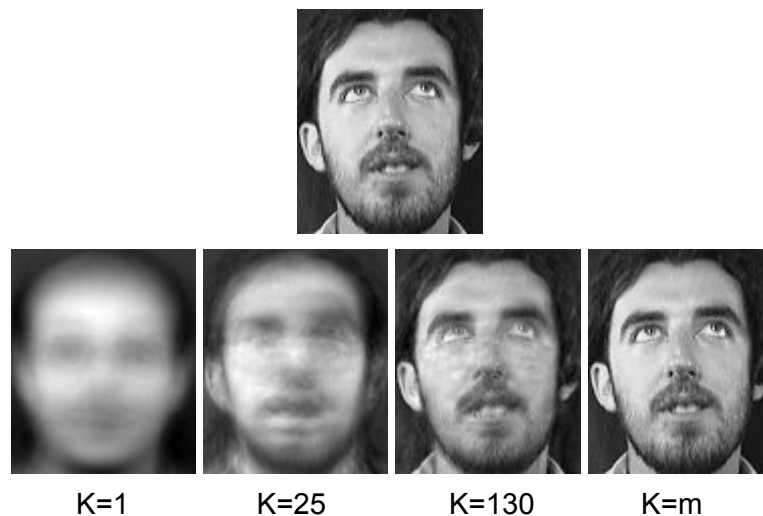
Ainsi pour diverses valeurs de K et une base de données composées d'environ 250 visages référents et 150 images tests, on obtient les images projetées suivantes :

Les 2 bases de données étant construite de manière aléatoire, les visages ne sont pas les même pour chaque test.



Nous pouvons nous apercevoir ci-dessus que même avec $K = m$, le visage ne peut être reconstruit parfaitement ce qui est tout à fait normal puisque la base de données ne détient pas notre image test.

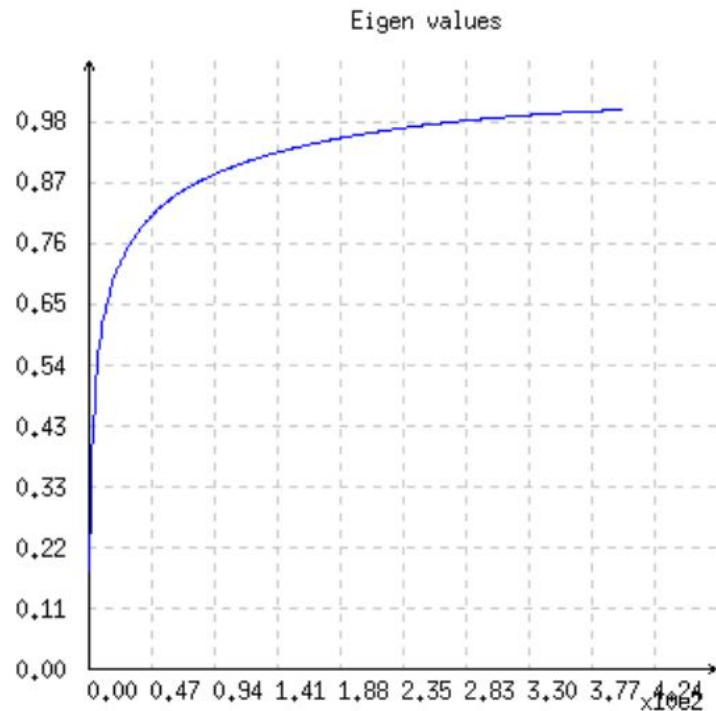
A l'inverse, pour une base de données rempli que d'image références (400 images), le résultat est différent. Etant donné que la matrice U contient toutes les eigenfaces de tous les visages référent, et donc a été construite avec l'image que l'on test, il est normal d'obtenir une image parfaitement reconstruite (aucune erreur de reconstruction) lorsque l'on utilise $K = m$ pour les visages de références.



Question 9-11:

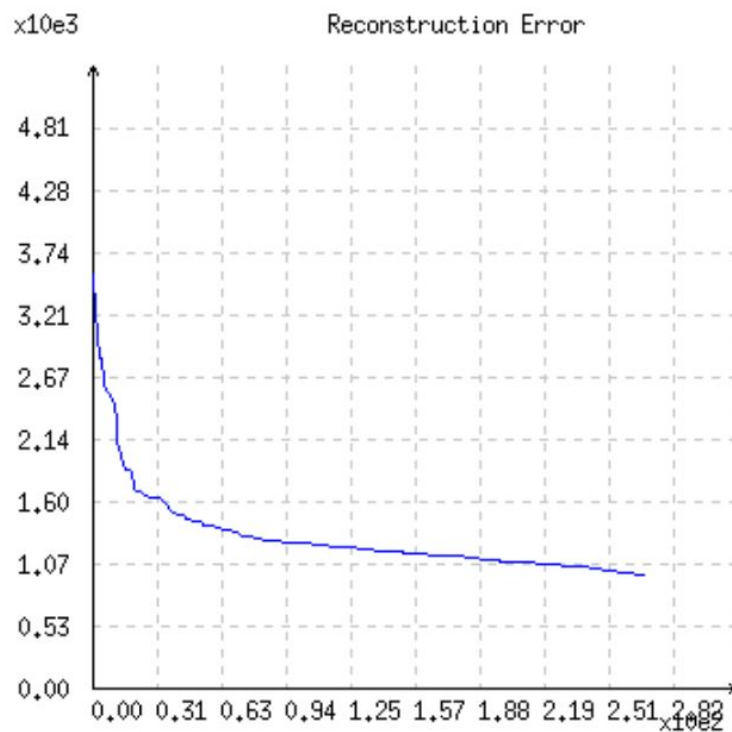
Si l'on effectue la reconstruction d'un visage étant présent dans la base de données et que l'on prend en compte toutes les eigenfaces ($K=m$), l'erreur entre l'image projetée et le visage testé devrait être nulle comme constaté ci-dessus.

Si l'on normalise les vecteurs propres et que l'on trace la courbe de leur somme cumulée on peut apercevoir les variations capturées par les K premières eigenfaces. Ainsi on se rend compte avec le résultat ci-dessous, qu'une bonne reconstruction (90%) est définie lorsque l'on considère au minimum une centaine d'eigenfaces.

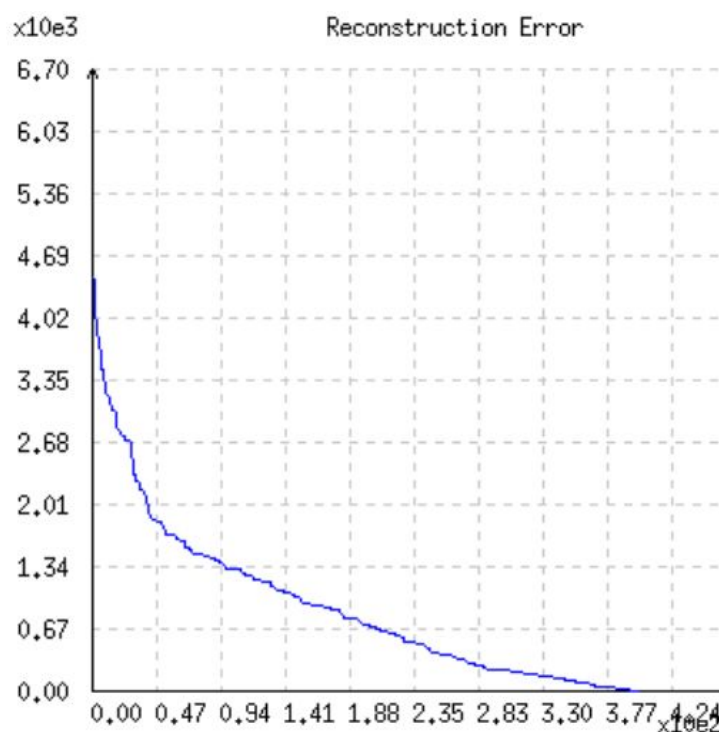


Eigen values pour une base de données de 400 visages

Ces observations sont cohérentes avec l'évolution de l'erreur de reconstruction par rapport à K .



Erreur de reconstruction pour 150 images tests/ 250 images références



Erreur de reconstruction pour la base de données de 400 images références

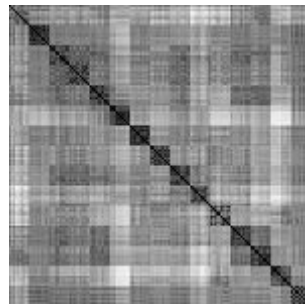
Question 12:

L'étape d'identification d'un visage test parmi un ensemble de visage de référence est donc une simple comparaison entre l'image projetée de l'image test et les images projetées des images de la base de données. Afin d'effectuer cette comparaison il est possible de le faire directement en comparant pixel à pixel les images projetés mais nous

pouvons aussi simplement comparer les coordonnées de l'image test avec les coordonnées des eigenfaces dans le sous espace des visages ce qui permettrait de réduire fortement le coût de calcul en évitant de calculer l'image projetée de chaque visage testé et de devoir comparer pixel à pixel les images qui peuvent être de résolution plus ou moins grande.

Question 13:

Pour 400 images dans la base de données référente, nous définissons une matrice représentant les distances dans le sous-espace par rapport à l'ensemble des visages de références.



Matrice de distance

Comme on pouvait s'y attendre lorsque l'on compare le même visage on obtient une valeur nulle. Cependant ce qui est réellement intéressant c'est les petits carrés qui se dessinent autour de la diagonale, ces derniers représentent chaque classe composant la base de données de l'image (une classe est l'ensemble de visage représentant la même personne). Ils sont perceptibles car la distance entre des visages similaires est plus faible que lorsque l'on compare 2 visages de personnes différentes.

On peut aussi noter la présence d'une symétrie de la matrice par rapport à la diagonale, ce qui est tout à fait logique puisque les comparaisons sont les mêmes (A-B et B-A).

Question 14:

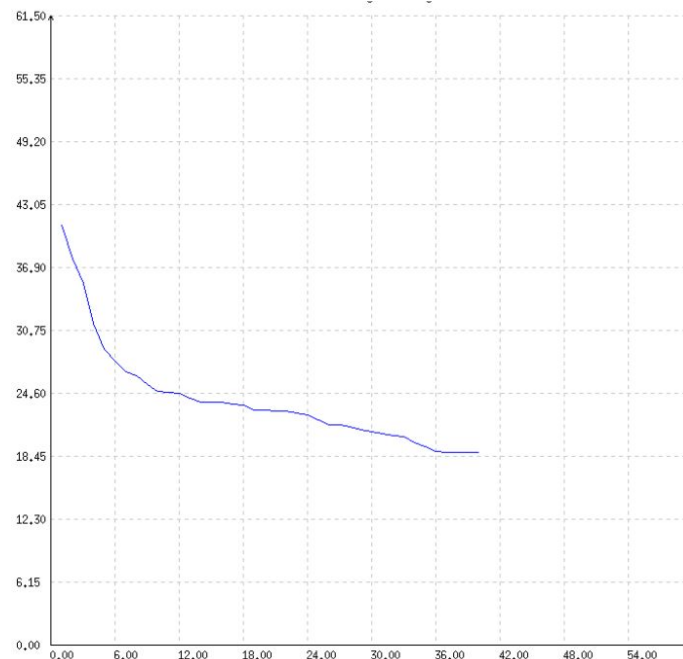
Si l'on regarde les distances minimales et maximales inter et intra classes (voir ci-dessous), on se rend compte que la distance minimale intra classes est bien plus faible que son homologue, ce qui s'explique car nous comparons deux fois la même personne. En revanche, il arrive que l'erreur au sein d'une même classe soit relativement grande, et donc non-distinguable des erreurs inter-classes. Par conséquent, nous faisons preuve d'incertitude pour certains visages concernant l'identification. Il est donc difficile de choisir un bon seuil θ permettant de distinguer si oui ou non le visage est reconnu avec un visage déjà existant dans la base de données.

distance euclidienne	Inter-classes	Intra-classes
minimale	3214.47	398.663
maximale	6637.84	6241.91

Question 15:

Néanmoins, fixons arbitrairement un seuil afin d'observer le nombre d'eigenfaces nécessaire pour une bonne reconnaissance pour un temps de calcul faible.

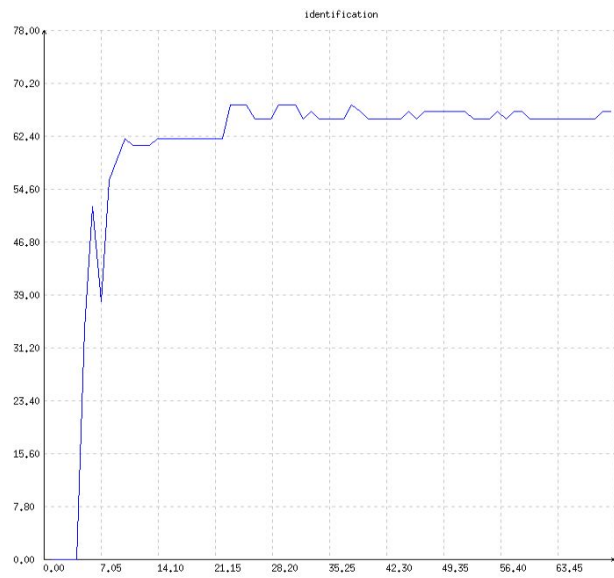
Par soucis de temps de calcul (VM et non-parallélisation du calcul) nous avons effectué les tests avec une trentaine de visages tests pour 40 visages de références.



Cette courbe représente le nombre de visages reconnus en fonction du nombre d'eigenfaces. De premier abord on pourrait constater que plus l'on prend d'eigenfaces plus l'identification devient mauvaise mais ce n'est pas le cas. Lorsque l'on prend peu d'eigenfaces nous effectuons des erreurs d'identification. Par conséquent certains visages ont été identifiés dans la base de visages référents alors qu'il n'aurait pas dû.

Ainsi pour obtenir une bonne identification il faut observer le moment où la courbe se stabilise. La première valeur obtenue représentera un temps de calcul faible pour une bonne qualité. Par manque de données on assumera arbitrairement une bonne identification pour un temps faible à 36 eigenfaces.

Suite à une erreur de compréhension d'énoncé, le résultat obtenu ci-dessus n'est pas celui attendu. En comparant toutes les distances associées aux visages référents, on considère la plus petite distance étant l'image référence similaire à l'image testée. Si elles sont de mêmes classe alors le visage testé est identifié. Ainsi la courbe obtenue est la suivante :



Pour un K proche de 20 on obtient un bon ratio entre bonne identification et temps de calcul.

Conclusion :

Nous avons implémenté lors de ce tp, une méthode d'identification de visage par image utilisant les eigenfaces. Dans notre cas, nous nous sommes intéressé à des images de faible résolution détenant seulement le visage centré dans l'image afin de nous faciliter la tâche. Nous avons pu effectuer divers tests et obtenu de bon résultats. Cependant, il faut savoir que cette méthode est la première méthode de reconnaissance faciale qui fonctionne correctement mais qui peut être assez coûteuse, notamment lorsque l'on utilise des images de grande résolution. Nous pouvons ainsi constater que, de nos jours, les méthodes de reconnaissances faciales sont bien plus performantes en utilisant des réseaux de neurones.