

COLLIARD Valentin

TRAORE Madina

COMPTE-RENDU

TME 9-10

4I600

Master d'Informatique

Sorbonne Université



Année universitaire 2018-2019

L'objectif de ce dernier TME est d'étudier les propriétés de la méthode de reconnaissance de visages « eigenfaces » en développant un système capable d'identifier un visage parmi une base de données de visages de référence, de déterminer si une image contient un visage présent dans la base de données et de reconnaître si une image représente un visage ou non.

Exercice 1 : Chargement de la base et centrage des visages

Nous écrivons les fonctions **visageMoyen** et **centre** afin de pouvoir visualiser la moyenne x_moy des visages de référence (appelé visage moyen) ainsi que les visages sous leur forme centrée (visage auquel on soustrait le visage moyen).

Le visage moyen est le suivant :

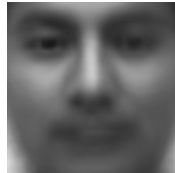


FIGURE 1 – Visage moyen

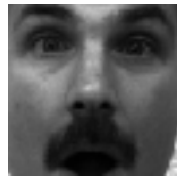
Quelques exemples de visages de la base et des visages centrés correspondant :



(a) Visage 6



(b) Visage 6 centré



(c) Visage 12



(d) Visage 12 centré



(e) Visage 18



(f) Visage 18 centré



(g) Visage 24



(h) Visage 24 centré



(i) Visage 40



(j) Visage 40 centré

FIGURE 2 – Exemples de visages centrés

Exercice 2 : Calcul des Eigenfaces (ACP)

Quinze premières eigenfaces et leurs valeurs propres associées :

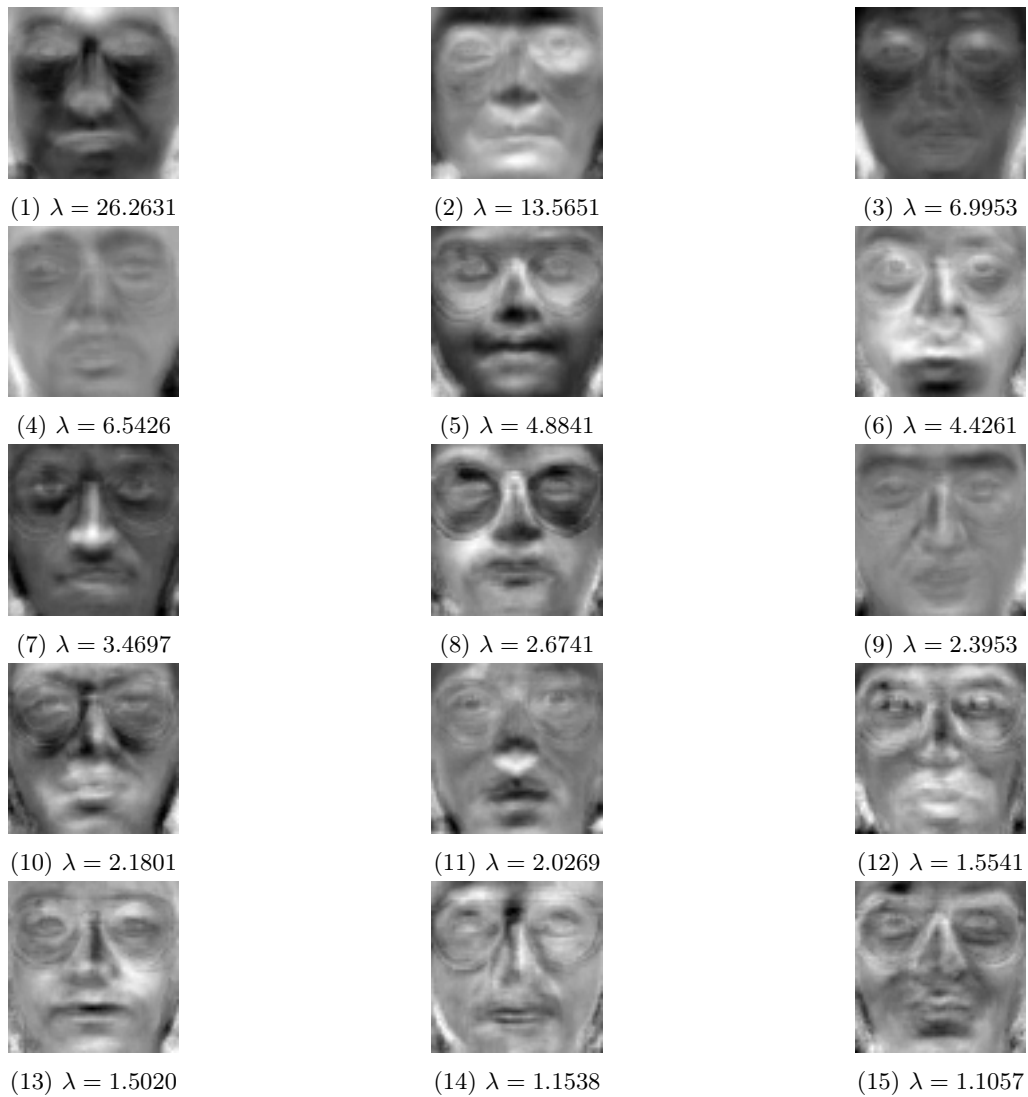


FIGURE 3 – 15 premières eigenfaces

Les eigenfaces correspondent aux axes principaux obtenus en effectuant l'Analyse en Composantes Principales (ACP). Autrement dit, ils mettent en avant les caractéristiques principales des visages.

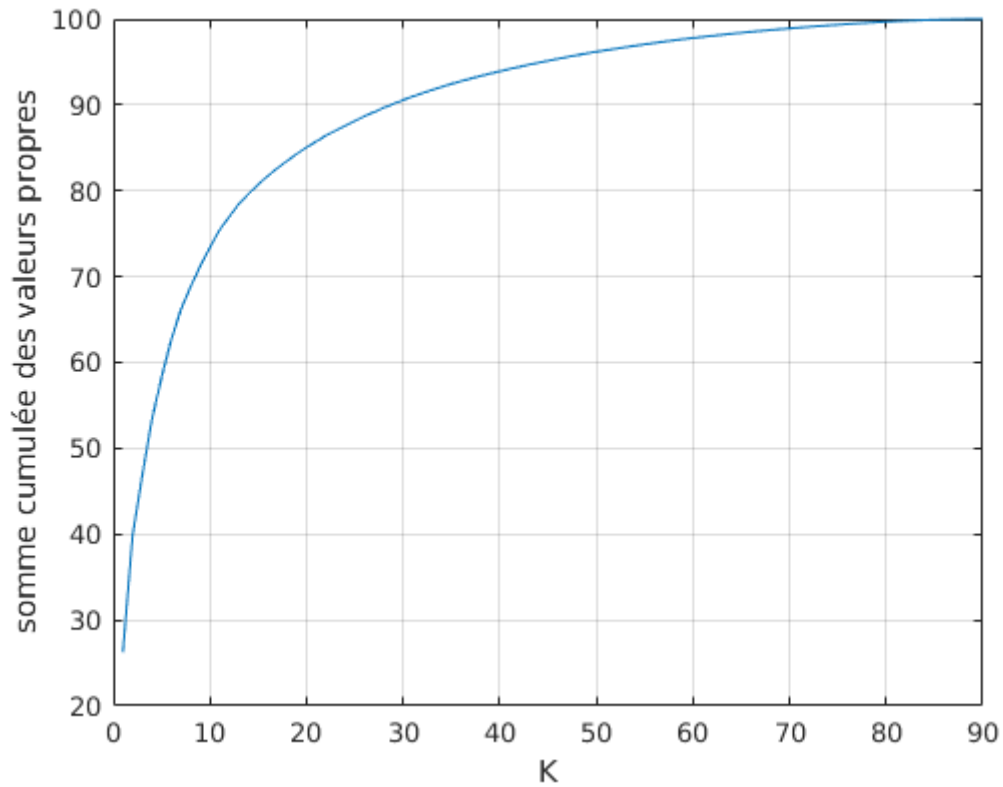


FIGURE 4 – Somme cumulée des valeurs propres normalisées

Pour $K = 50$ on obtient une reconstruction d'environ 95%. On peut donc considérer que les 50 premières eigenfaces suffisent à obtenir une reconstruction correcte.

Exercice 3 : Projection dans le sous-espace des visages

Notons e_r l'erreur de reconstruction.

Voici un exemple de reconstruction d'un visage de la base d'apprentissage et d'un visage de la base de test ainsi que l'erreur de reconstruction associée :

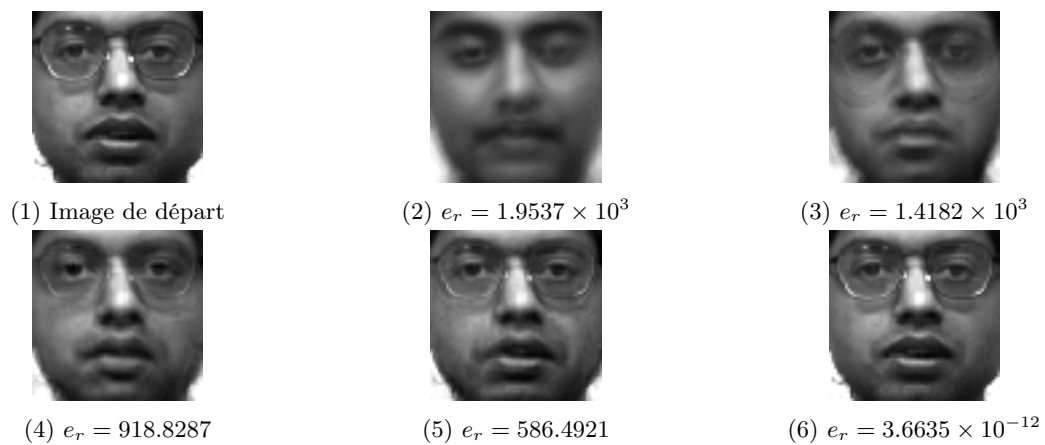


FIGURE 5 – Reconstruction de l'image 55 de la base d'apprentissage

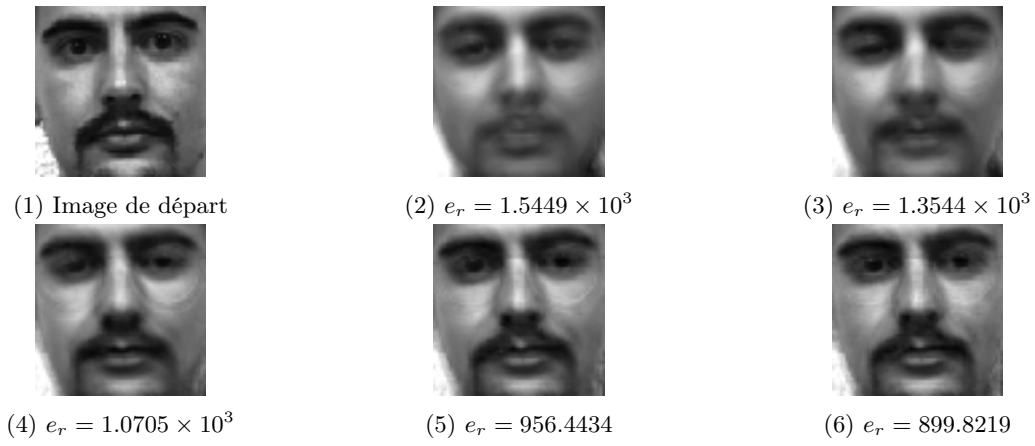


FIGURE 6 – Reconstruction de l'image 17 de la base de test

On observe une différence entre les visages issus de la base d'apprentissage et ceux issus de la base de test en terme de reconstruction. En effet, dans l'exemple présent à la Figure 5 (visage de la base d'apprentissage), le visage reconstruit semble identique au visage de départ et l'erreur de reconstruction pour $K = 90$ avoisine 0 ($e_r = 3.6635e - 12$) tandis que dans l'exemple de la Figure 6 (visage de la base de test) l'image reconstruite est beaucoup moins nette et l'erreur de reconstruction est bien plus élevée ($e_r = 899.8219$).

Question bonus :

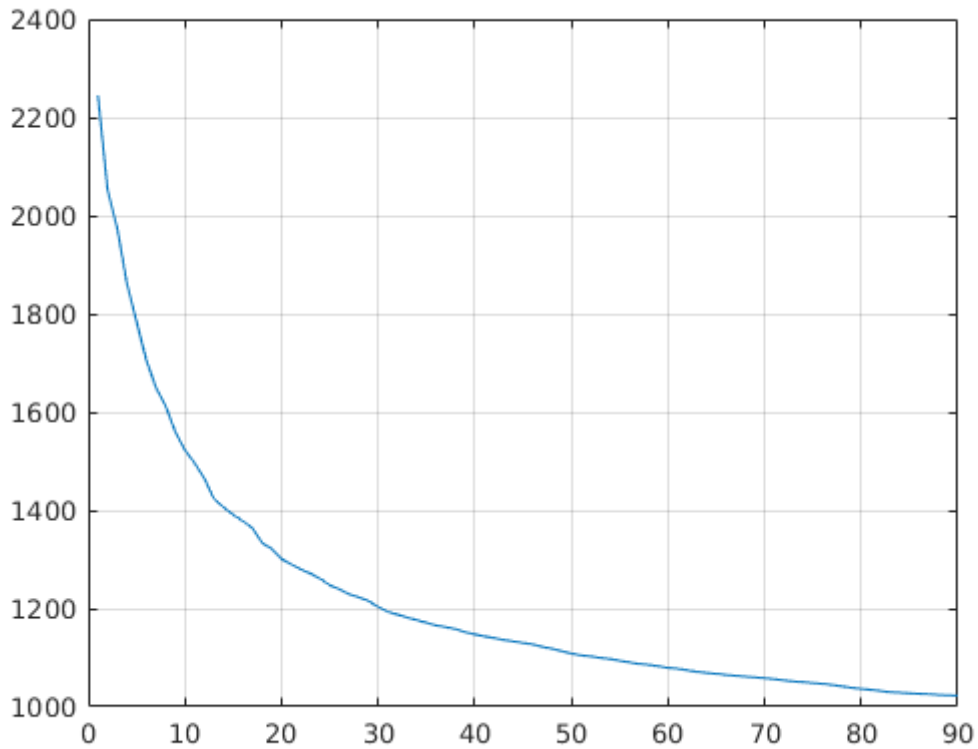


FIGURE 7 – Courbe de l'erreur moyenne de reconstruction

La courbe obtenue est cohérente avec celle de la somme cumulée calculée précédemment (Figure 4) puisqu'elle témoigne d'une évolution inversement proportionnelle à celle-ci.

Exercice 4 : Reconnaissance de visage : identification

Travailler dans le sous-espace des visages plutôt que dans l'espace de départ permet au modèle d'être moins sensible aux variations de poses, d'illumination et d'expressions de visages et ainsi ne pas être trop spécifique de la base d'apprentissage : cela assure un meilleur taux de reconnaissance pour de nouvelles images données.

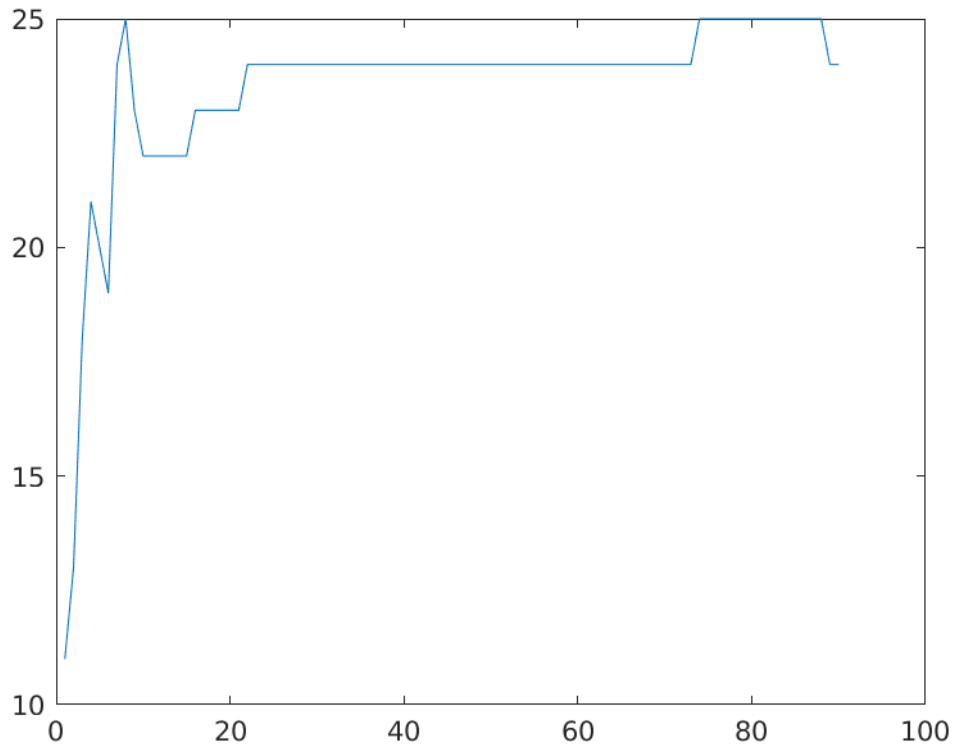


FIGURE 8 – Courbe du nombre de visages reconnus en fonction de K

On pourrait s'attendre à ce que la courbe soit strictement croissante (plus on a d'informations sur les visages de la base d'apprentissage mieux on sait reconnaître les visages de la base de test) cependant, comme dit précédemment, être trop spécifique de la base d'apprentissage peut être handicapant tout comme ne pas l'être assez. La "bonne" valeur de K ne peut donc qu'être trouvée de manière expérimentale.

Pour $K = 30$ on obtient un taux d'identification de 80% tandis que pour $K = 8$ le taux d'identification est d'environ 83%. Ainsi, pour avoir une bonne reconnaissance et un temps de calcul faible, on peut choisir $K = 8$.

Questions bonus :

La matrice de distance entre les visages de la base d'entraînement est la suivante :

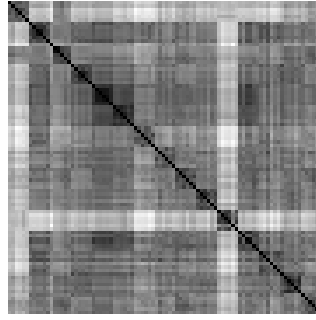


FIGURE 9 – Matrice de distance entre les visages de la base d'entraînement (plus les pixels sont clairs plus la distance entre les deux visages est grande)

On note $\min_{D_intra-classes}$ et $\max_{D_intra-classes}$ les distances min et max entre deux visages d'une même personne, $\min_{D_inter-classes}$ et $\max_{D_inter-classes}$ les distances min et max entre deux visages de la base d'entraînement.

On obtient les valeurs suivantes :

- $\min_{D_intra-classes} = 760.2438$;
- $\max_{D_intra-classes} = 4.4267 \times 10^3$;
- $\min_{D_inter-classes} = 270.8728$;
- $\max_{D_inter-classes} = 6.7306 \times 10^3$.

Ainsi, si l'on souhaite mettre en place un seuil θ pour détecter la présence d'un visage n'étant pas répertorié dans la base d'apprentissage on pourra choisir $\theta = \max_{D_inter-classes} = 6.7306 \times 10^3$. Toute image dont la distance minimum aux images de la base d'entraînement dépassera la valeur de θ sera alors rejetée.

Exercice 5 : Classification visage/non-visage

On obtient les courbes d'erreur suivantes pour $K = 30$:

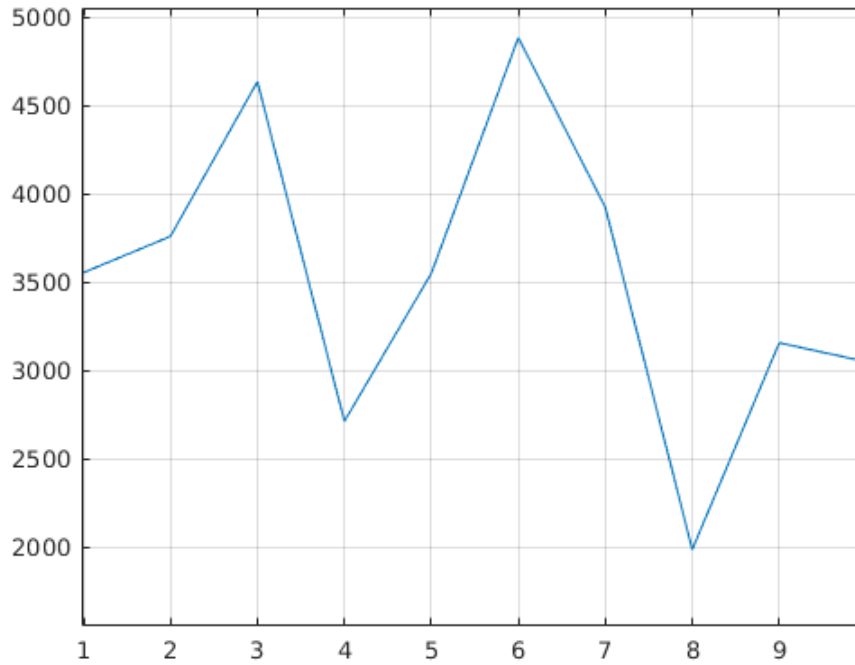


FIGURE 10 – Erreur de reconstruction des images de la base des "non-visages"

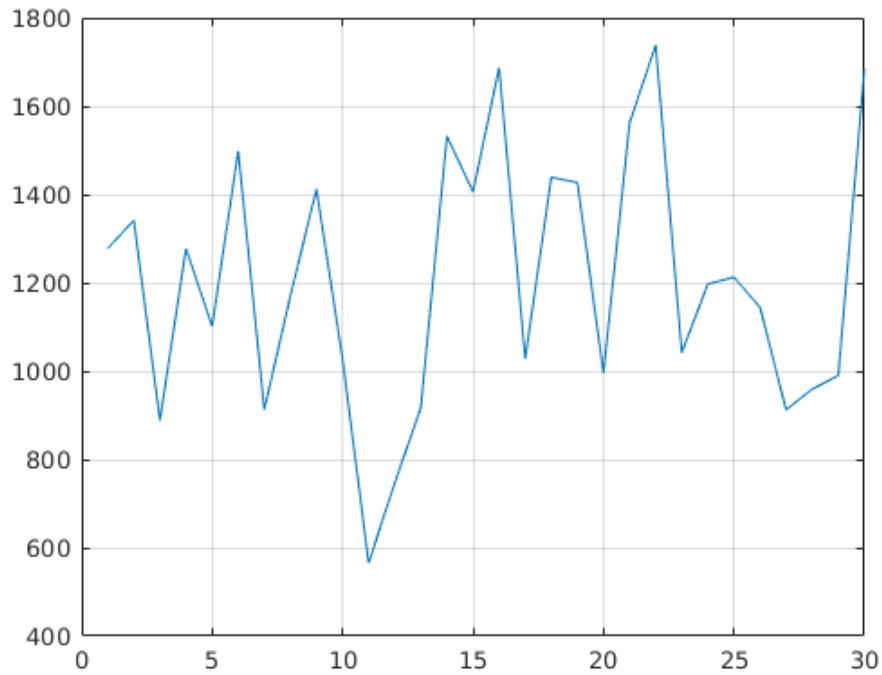


FIGURE 11 – Erreur de reconstruction des images de la base des visages

On note $moy_{e_r, visages}$ et $max_{e_r, visages}$ l'erreur moyenne et l'erreur maximale de reconstruction des images

de la base des visages, $moy_{e_r, non-visages}$ et $min_{e_r, non-visages}$, l'erreur moyenne et l'erreur minimale de reconstruction des images de la base des non-visages.

On obtient les valeurs suivantes :

- $moy_{e_r, visages} = 1.2031 \times 10^3$;
- $moy_{e_r, non-visages} = 3.5193 \times 10^3$;
- $min_{e_r, non-visages} = 1.9848 \times 10^3$;
- $max_{e_r, visages} = 1.7384 \times 10^3$.

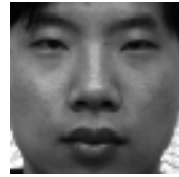
Comme prévu on constate que l'erreur moyenne de reconstruction des images de la base des non-visages est bien plus élevée (environ 3 fois plus) que celle des images de la base des visages. Par ailleurs, l'erreur minimale de reconstruction des images de la base des non-visages est supérieure à l'erreur maximale de reconstruction des images de la base des visages. On peut utiliser cette information pour mettre en place un seuil permettant de détecter si une image est un visage ou non (i.e. toute image dont l'erreur de reconstruction dépassera la valeur de $max_{e_r, visages}$ sera considérée comme n'étant pas un visage). Visualisons l'erreur de reconstruction pour quelques images des bases de visages et de non-visages :



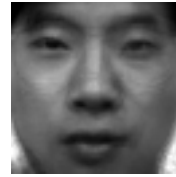
(1) Visage 6



(2) $e_r = 1.4985 \times 10^3$



(3) Visage 7



(4) $e_r = 912.5368$



(5) Visage 8



(6) $e_r = 1.1703 \times 10^3$



(7) Visage 9



(8) $e_r = 1.4124 \times 10^3$



(9) Visage 10

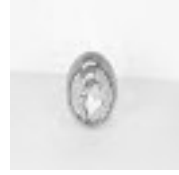


(10) $e_r = 1.0272 \times 10^3$

FIGURE 12 – Exemples de reconstruction de visages de la base de test



(1) Image 1



(3) Image 2



(5) Image 3



(7) Image 4



(9) Image 5



(2) $e_r = 3.5525 \times 10^3$



(4) $e_r = 3.7574 \times 10^3$



(6) $e_r = 4.6330 \times 10^3$



(8) $e_r = 2.7103 \times 10^3$



(10) $e_r = 3.5483 \times 10^3$

FIGURE 13 – Exemples de reconstruction d'images de la base de "non-visages"

Naturellement, on observe visuellement une différence très importante entre les images de la base des non-visages et leur reconstruction. En effet, seule la luminosité et les contrastes de l'image semblent être bien reproduites. Les images de la base de visages ont eux été reproduits de manière satisfaisante.