



Application de l'ACP : les «Eigenfaces»

Calcul efficace des vecteurs propres principaux

Le travail de M. Turk et A. Pentland intitulé «Eigenfaces for Recognition»¹ et publié en 1991 dans la revue «Journal of Cognitive Neuroscience» (vol. 3, pp 71-86) constitue une des applications les plus intéressantes et populaires de l'ACP au domaine de la reconnaissance de forme. Il s'agit tout simplement d'appliquer l'ACP à partir de données de très grandes tailles : des images de visages.

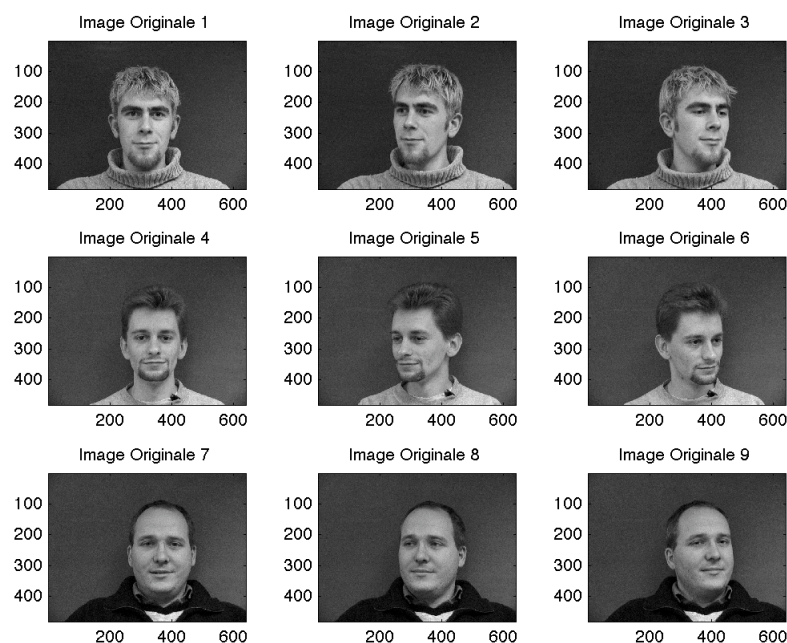


FIGURE 1 – Une base de visages

1 Données manipulées

La figure 1 montre une base de $n = 9$ visages présentant 3 personnes dans 3 positions différentes relativement à la caméra. Dans cet exemple, chaque image est de taille : 480 lignes \times 640 colonnes. Il est immédiat, par empilement des colonnes, de se ramener à la situation où chaque image est un vecteur $\mathbf{x}_i \in \mathbb{R}^p$ avec $p = 480 \times 640 = 307200$ composantes (niveaux de gris / pixels).

1. une recherche internet rapide permet de trouver le papier original

En adoptant les notations usuelles pour l'ACP, la base d'images forme un tableau de données $\mathbf{X} \in \mathcal{M}_{\mathbb{R}}(n = 9, p = 307200)$ avec, en ligne i , le vecteur $\mathbf{x}_i^T \in \mathbb{R}^p$ qui représente la i^{ieme} image de la base.

L'individu moyen $\bar{\mathbf{x}} \in \mathbb{R}^p$ est le vecteur des **moyennes arithmétiques des variables** : il représente alors, tout simplement, la version «vectorisée» de l'image moyenne $\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i$ des images de la base. Cette image moyenne est montrée en haut et à gauche de la figure 3. Il est alors possible de centrer les données (c'est-à-dire les images de la base) pour former le tableau centré $\mathbf{X}^c = \mathbf{X} - \mathbf{1}_n \bar{\mathbf{x}}^T$.

2 Reconnaissance de visages

On s'intéresse alors à un mécanisme de prédiction visant la reconnaissance de visages. L'idée est de donner une image requête $\mathbf{x} \in \mathbb{R}^{307200}$ en entrée d'un prédicteur dont l'objectif est de prédire l'image (ou les images) de la base la (ou les) plus similaire(s) à cette requête selon le schéma de boîte noire suivant :

$$\mathbf{x} \in \mathbb{R}^{307200} \rightarrow \boxed{\text{h}} \rightarrow h(\mathbf{x})$$

Une requête simple est illustrée par la figure 2 suivante. L'image requête est celle d'un des trois individus présents dans la base précédente mais dans une nouvelle position (bouche ouverte). En d'autres termes, l'image requête n'appartient pas à la base de visage mais est similaire à trois des images de cette base. L'objectif est donc d'identifier ces similarités pour, éventuellement, reconnaître l'individu. Dans la figure 2, les trois images du même individu présentes dans la base sont renvoyées comme résultat (correct) de la requête. Ce résultat est ordonné : la première image trouvée (choix 1) est mathématiquement (mais pas visuellement) la plus similaire à la requête.

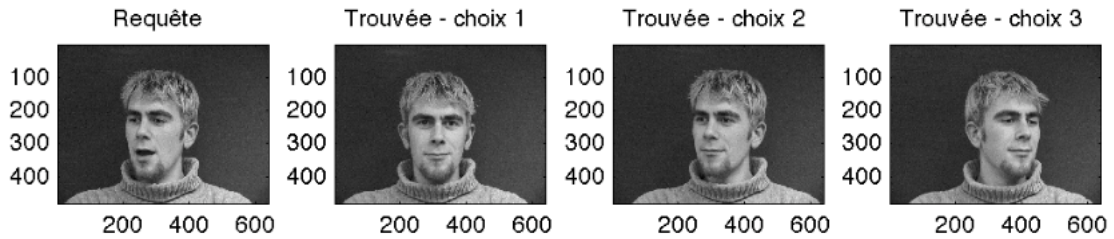


FIGURE 2 – Résultat d'une requête sur une base de visages.

Il est inutile et inopportun d'utiliser les $p = 307200$ niveaux de gris pour comparer l'image requête avec chacune des images de la base. L'ACP est donc un prétraitement qui consiste à réduire la dimension des données selon le schéma suivant où le prédicteur prend en entrée la sortie de l'ACP :

$$\mathbf{x} \in \mathbb{R}^{307200} \rightarrow \boxed{\text{ACP}} \rightarrow \mathbf{c}(\mathbf{x}) \in \mathbb{R}^{q \ll 307200} \rightarrow \boxed{\text{h}} \rightarrow h(\mathbf{x})$$

Le vecteur caractéristique $\mathbf{c}(\mathbf{x})$ est tout simplement formé des q premières composantes principales résultant de la projection de l'image requête \mathbf{x} sur les q premiers vecteurs propres unitaires

$$\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_q\} \subseteq \mathbb{R}^p \text{ de } \Sigma = \frac{1}{n} (\mathbf{X}^c)^T \mathbf{X}^c.$$

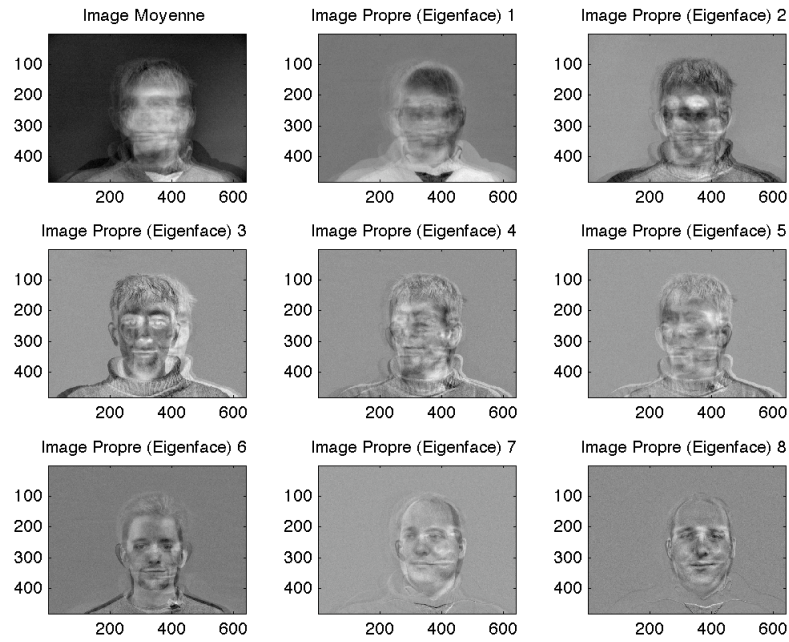


FIGURE 3 – Les «eigenfaces»

Ces vecteurs propres sont des vecteurs de \mathbb{R}^{307200} est donc également des images que l'on appelle «eigenfaces». Ces images sont présentées dans la figure 3. On a ainsi :

$$\mathbf{c}(\mathbf{x}) \stackrel{\text{déf}}{=} \begin{bmatrix} \mathbf{w}_1^\top \mathbf{x} \\ \mathbf{w}_2^\top \mathbf{x} \\ \vdots \\ \mathbf{w}_q^\top \mathbf{x} \end{bmatrix}$$

On dit que $\mathbf{c}(\mathbf{x})$ est une représentation compacte préservant au mieux l'information contenue dans \mathbf{x} avec seulement q composantes principales. La prédiction h peut alors être définie de diverses manières. Parmi les possibilités les plus simples, on peut chercher les images de la base dont les représentations compactes sont les plus proches (au sens d'une distance d bien choisie) de celle associée à l'image requête :

$$h(\mathbf{x}) = \arg \min_{i \in 1 \dots n} d(\mathbf{c}(\mathbf{x}_i), \mathbf{c}(\mathbf{x}))$$

C'est en choisissant $q = 3$ et $d(\mathbf{c}(\mathbf{x}_i), \mathbf{c}(\mathbf{x})) = \|\mathbf{c}(\mathbf{x}_i) - \mathbf{c}(\mathbf{x})\|_2$ que les résultats des figures précédentes ont été produits.

Cette présentation occulte toutefois un problème délicat : la grande taille de $\Sigma = \frac{1}{n}(\mathbf{X}^c)^\top \mathbf{X}^c$ ($p \times p = 307200 \times 307200$) ne permet pas de calculer ses vecteurs propres de manière aisée. Ce sera un des premiers problèmes auquel nous nous intéresserons lors de ce projet.

3 Projet Calcul Scientifique et Analyse de Données

Le projet que nous vous proposons aura pour but de réaliser cette reconnaissance d'image de la manière la plus efficace. Nous souhaitons vous faire développer un code efficace, interfaçant *Matlab* pour les pré et post-traitements et *FORTRAN* pour le calcul des vecteurs propres. Ce développement sera accompagné d'expérimentations et de tests de performance devant vous faire réfléchir aux limites de la solution demandée.

Pour vous accompagner dans ce projet, plusieurs étapes vous sont proposées avec des points de rendez-vous sous forme de séances de projet, des rendus de «délivrables» réguliers, ainsi qu'une auto-formation au langage *FORTRAN*.



Attention : Le projet est réalisé en groupe/équipe de 3 étudiant·e·s appartenant **au même groupe de TD**.

La date limite de constitution des équipes est le **vendredi 6 mars 2020** avant de partir en vacances. Vous donnerez l'information par mail à votre enseignant de projet.

Programme / Dates à retenir

Voir Moodle