

Eigenfaces

Bouarah Romain

Langdorph Matthieu
Souffan Nathan

Ketels Lucas

24 avril 2020

Première partie

Partie Mathématiques

1 Calcul des Eigenfaces

1.1 Travail dans $\mathbb{R}^{N \times N}$

Considérons une image de visage comme une matrice $N \times N$ dont le coefficient (i, j) est égal au niveau de gris du pixel (i, j) (l'origine se situant dans le coin haut gauche). On transforme ensuite cette matrice comme un vecteur de $\mathbb{R}^{N \times N}$ en juxtaposant les colonnes l'une en dessous de l'autre, par exemple.

$$\begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,N} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N,1} & p_{N,2} & \cdots & p_{N,N} \end{pmatrix} \rightarrow \begin{pmatrix} p_{1,1} \\ p_{2,1} \\ \vdots \\ p_{N,1} \\ \vdots \\ p_{1,N} \\ \vdots \\ p_{N,N} \end{pmatrix}$$

1.2 Calcul des valeurs propres et des vecteurs propres de la matrice de covariance

Les images des visages sont globalement similaires, donc ces images ne seront pas distribuées aléatoirement dans notre espace $\mathbb{R}^{N \times N}$. On peut donc décrire notre espace des visages de manière plus fine (*i.e.* avec moins de dimensions).

1.2.1 Matrice de covariance

Avant de commencer la réduction de l'espace de travail, nous allons d'abord analyser cette dispersion en calculant la matrice de covariance de manière empirique.

Définition 1 (Matrice de Covariance). La matrice de covariance d'un vecteur de p variables aléatoires $\vec{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}$ dont chacune possède une variance, est la matrice carrée dont le terme générique est donné par $a_{i,j} = \text{Cov}(X_i, X_j)$.

Définition 2 (Matrice de Covariance). La matrice de covariance, notée parfois Σ , est définie par :

$$\text{Var}(\vec{X}) = E[(\vec{X} - E(\vec{X}))(\vec{X} - E(\vec{X}))^T]$$

Définition 3 (Estimation de la Matrice de Covariance). En partant d'un échantillon de réalisations indépendantes d'un vecteur aléatoire, un estimateur non-biaisé de la matrice de covariance est donné par :

$$\text{Var}(\vec{X}) = \frac{1}{n-1} \sum_{i=1}^n (\vec{X}_i - \vec{\mu})(\vec{X}_i - \vec{\mu})^T$$

où $\vec{\mu} = \frac{1}{n} \sum_{i=1}^n \vec{X}_i$ est le vecteur des moyennes empiriques.

Application à notre cas Supposons que nous avons M images de visage. On note $I = [I_1, I_2, \dots, I_M]$ la matrice de taille $N^2 \times M$ de l'ensemble de nos images.

1. On calcule le visage moyen $\Psi = \frac{1}{M} \sum_{i=1}^M I_i$.
2. On retire le visage moyen à chacun de nos visages, en effet nous nous intéressons uniquement aux particularités.
Donc, chaque visage diffère de la moyenne par le vecteur $\Phi_i = I_i - \Psi$.
3. On calcule la matrice de covariance

$$\begin{aligned} C &= \frac{1}{M-1} \sum_{i=1}^M \Phi_i \Phi_i^T \\ &= \frac{1}{M-1} A A^T \end{aligned}$$

où $A = [\Phi_1, \Phi_2, \dots, \Phi_M]$ est la matrice $N^2 \times M$ de nos visages centrés.

Cette matrice encode la dispersion de nos images de visages dans \mathbb{R}^{N^2} . Les coefficients de la diagonale sont les variances selon les axes e_1, e_2, \dots, e_{N^2} . Les autres coefficients sont la covariance entre deux axes.

Remarque 1. C est symétrique réelle donc diagonalisable dans une base orthonormée. C est également définie semi-positive, c'est à dire que toutes ses valeurs propres sont positives

1.2.2 Méthode 1 : Analyse en composantes principales

L'analyse en composantes principales consiste à transformer des variables liées entre elles (dites "corrélées" en statistique) en nouvelles variables décorréliées les unes des autres.

Ces nouvelles variables sont nommées "composantes principales", ou axes principaux, dans notre cas elles sont appelées *eigenfaces*.

L'ACP nous permet de réduire le nombre de variables et de rendre l'information moins redondante. En effet, nous cherchons les meilleurs axes ou *eigenfaces* décrivant le mieux notre espace de visages.

De plus, le nombre de *eigenfaces* est toujours inférieurs au nombre d'images de visages.

Finalement, nous cherchons le vecteur $u \in \mathbb{R}^{N^2}$ tel que la projection des images des visages sur u ait une variance maximale. Cette projection s'écrit :

$$p_u(A) = Au$$

La variance empirique de $p_u(A)$ vaut donc $p_u(A)^T p_u(A) = u^T C u$

Comme C est symétrique réelle, elle est donc diagonalisable dans une base orthonormée.

Notons P le changement de base associé et $D = \text{Diag}(\lambda_1, \dots, \lambda_L)$ la matrice diagonale formée de son spectre rangé en ordre décroissant, on a :

$$p_u(A)^T p_u(A) = u^T P^T D P u = (P u)^T D \underbrace{(P u)}_v$$

Le vecteur unitaire u qui maximise $v^T D v$ est un vecteur propre de C associé à la valeur propre λ_1 , on a alors :

$$v^T D v = \lambda_1$$

La valeur propre λ_1 est la variance empirique sur le premier axe de l'ACP. On continue la recherche du deuxième axe de projection w sur le même principe en imposant qu'il soit orthogonal à u .

1.2.3 Méthode 2 : Décomposition en valeurs singulières

La décomposition en valeurs singulières permet de factoriser des matrices carrées ou rectangulaires réels ou complexes, on s'intéressera ici au cas réels. Énoncé : Soit M une matrice $m \times n$, alors il existe une décomposition de la forme :

$$M = U\Sigma V^t$$

Avec U une matrice unitaire $m \times m$, Σ une matrice $m \times n$ où les coefficients diagonaux sont des réels positifs ou nuls et tous les autres sont nuls, et V est une matrice unitaire $n \times n$. On appelle ainsi cette factorisation la décomposition en valeurs singulières de M .

- La matrice V contient un ensemble de vecteurs de base orthonormés de \mathbb{R}^n d'entrée
- La matrice U contient un ensemble de vecteurs de base orthonormés de \mathbb{R}^m de sortie
- La matrice Σ contient dans ses coefficients diagonaux les valeurs singulières de la matrice M . Elles correspondent aux racines des valeurs propres de $M^t M$

On appelle ainsi valeur singulière de M toute racine carrée d'une valeur propre de $M^t M$, autrement dit tout réel positif λ tel qu'il existe un vecteur unitaire u dans \mathbb{R}^m et un vecteur unitaire v vérifiant dans \mathbb{R}^n :

$$M^t u = \lambda v \text{ et } Mv = \lambda u$$

Dans le cas d'une matrice carrée symétrique définie semi-positive (ce qui est le cas ici pour C) les valeurs singulières et vecteurs singuliers correspondent aux valeurs propres et vecteurs propres de M .

Il existe plusieurs façon de calculer une décomposition en valeurs singulière. Un algorithme courant consiste en :

- Effectuer une décomposition QR si la matrice possède plus de lignes que de colonnes
- Réduire le facteur R sous forme bidiagonale, (on pourra notamment utiliser des transformations de Householder alternativement sur les colonnes et sur les lignes de la matrice).

- Les valeurs singulières et vecteurs singuliers sont alors trouvés en effectuant une itération de type QR bidiagonale avec la procédure DBD-SQR

2 Utilisation des eigenfaces pour classer une image de visage

*Définir M' comme étant le nombre de visage du training set
Définir la base de l'espace des eigenfaces*

2.1 Projection dans l'espace des visages

Soit Γ une nouvelle image de visage, on la projette dans l'espace des visages par :

$$\omega_k = u_k^T(\Gamma - \Psi)$$

pour $k = 1, \dots, M'$, on obtient ainsi un vecteur Ω tel que :

$$\Omega = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_{M'} \end{pmatrix}$$

Ω décrit la contribution de chacun des eigenfaces pour l'image en question.

2.2 Analyse de la projection

On peut maintenant utiliser ce vecteur pour reconnaître si ce vecteur correspond à un visage déjà connue étant dans la base ou si c'est un visage inconnu. Pour ce faire, on cherche la classe k qui minimise la distance euclidienne $\epsilon_k = \|(\Omega - \Omega_k)\|^2$ où Ω_k est le vecteur décrivant la k^{ieme} classe de visage. Les classes de visages Ω_i sont calculés en faisant la moyenne de plusieurs ou une image du visage de chaque individu.

On considère ensuite qu'un visage appartient à une certaine classe de visage k si le minimum ϵ_k est en dessous un certain seuil Θ . Si $\forall k, \epsilon_k > \Theta$, alors le visage est inconnu et on a alors une nouvelle classe de visage.

Il y a finalement 4 possibilités pour une image :

1. L'image est proche de l'espace des visages et proche d'une classe de visage en particulier, c'est alors un visage connu

2. L'image est proche de l'espace des visages mais n'est proche d'aucune classe de visage, c'est alors un visage inconnu
3. L'image est distante de l'espace des visages mais proche d'une classe de visage, *EXPLICATION!!!* C'est la plupart du temps un faux positif.
4. L'image est distante de l'espace des images et également distante de toutes les classes de visages, on peut alors en conclure que ce n'est pas un visage

Deuxième partie

Partie pratique

3 Autres techniques utilisés pour la reconnaissance faciale

La reconnaissance faciale par les Eigenfaces n'est pas forcément la technique la plus utilisée. D'autres techniques sont utilisés dans le domaine de la reconnaissance faciale et visuelle. Notamment on va voir qu'il y a le *Embedded Hidden Markov Model* (Embedded HMM) et les réseaux neuronaux convolutifs.

3.1 Embedded HMM

En français, nous parlons de modèle de Markov caché, ou encore de automate de Markov à états cachés. Nous sommes dans le cadre des probabilités conditionnelles ici. Cette technique est plus utilisé pour de la reconnaissance visuelle, reconnaître des formes, des objets. Posons la notion de processus de Markov : une fonction aléatoire vérifiant que la distribution conditionnelle de probabilité de l'états futur de l'automate, ne dépend que de l'état présent, et non pas des états passés (propriété de Markov). Soit pour E^{n+2} un ensemble de $n+2$ états :

$$\forall n \geq 0, \forall (i_0, \dots, i_{n-1}, i, j) \in E^{n+2},$$

$$P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_n = i) = P(X_{n+1} = j | X_n = i)$$

Une chaîne de Markov est un processus de Markov si X est une variable discrète muni par P , la fonction de probabilité. Au contraire, dans un modèle de Markov caché, nous sommes dans un modèle markovien dont les états d'une exécution sont inconnus (cachés). Ce modèle est donc utile pour modéliser un système physique probabiliste, utilisé ici dans le cadre d'un algorithme d'optimisation (espérance-maximisation). Dans le cadre de la reconnaissance d'objets, c'est par l'étude des formes que ce modèle probabiliste propose un objet correspondant à une image reçue.

3.2 Réseau de neurones convolutifs

Nous sommes dans le cadre d'un réseau de neurones convolutifs. Le réseau de neurones convolutifs est une branche de l'intelligence artificielle, le Deep Learning. (rappelons que l'analyse d'image est un domaine où le deep learning est utilisé) Le fonctionnement d'un réseau de neurones est similaire aux réseaux de neurones chez l'Homme. Chaque neurone a plusieurs entrées et une seule sortie, et des actions entre les deux. Les réseaux de neurones utilisent de l'algèbre linéaire (multiplication de matrices, dérivées partielles). Un réseau de neurones convolutifs est un réseau de neurones où les motifs de connexions entre les neurones est inspiré par le cortex visuel des humains. Chaque propriété d'une image analysée va occuper une place dans sa reconnaissance. L'analyse d'image ici passe par plusieurs "couches de neurones". Par exemple en premier la luminosité des pixels va être analysée, puis les corrélations entre des pixels voisins, à partir de ces corrélations identifier des lignes directrices, puis regrouper ces lignes directrices pour former des structures, puis identification des structures (par exemple des yeux), puis mettre en lien les identifications pour trouver l'identification de l'image. Une matrice de convolution est utilisée dans le traitement d'images pour le floutage, l'amélioration de la netteté de l'image, le gaufrage, la détection de contours, etc... Chaque neurone du réseau analysera une matrice de convolution en fonction du pixel. Ainsi sont identifiés les éléments d'une image, ou un visage sur une image amenant à la reconnaissance faciale. Ces réseaux de convolutions sont aussi utiles dans la compréhension de texte, en médecine, et même pour le jeu de go (soit lorsqu'il y a une partielle corrélation globale, relative invariance spatiale ou temporelle (vidéo)).

4 Applications de reconnaissances faciales par toutes les techniques vues

Il existe de nombreuses applications traitant de la reconnaissance faciale. On peut en trouver pour l'industrie, pour de la recherche, pour un particulier ou pour un organisme plus grand. Par sécurité, on peut équiper notre maison d'un système de vidéo-surveillance, pour se protéger d'éventuels intrusions d'individus dans celle-ci. Supposons que cela se passe, le propriétaire aura alors accès à une banque d'image des individus. Il pourra alors les faire analyser par la police pour trouver les coupables. La reconnaissance faciale peut également être utile comme droit d'entrer dans une entreprise, ou via une reconnaissance de l'iris, au même niveau que la reconnaissance digitale ou vocale. La reconnaissance faciale est donc très utile dans la sécurité.

4.1 Reconnaissance faciale sur smartphone

L'évolution des appareils électroniques est exponentielle depuis plusieurs années, et les nouveaux smartphones sont à la pointe de la technologie. Ils sont équipés d'un système de sécurité de reconnaissance faciale. Le propriétaire du téléphone peut déverrouiller celui-ci en le tenant naturellement, objectifs de l'appareil photo facial orienté vers son visage. Il sera également possible de faire des achats par son téléphone par reconnaissance faciale. Tout le travail de la confection de ce système est d'empêcher le système de se faire duper par une photo, masque, etc... Il est donc intéressant d'avoir un système de reconnaissance faciale 3D, pour optimiser la sécurité, sans compter la reconnaissance vocale et digitale pouvant l'accompagner. Par exemple Apple a ajouté ce système sur son Iphone X, le "Face ID", ici nous sommes dans le cas de la reconnaissance faciale 3D. Pour l'utiliser il suffit juste d'enregistrer des images de son visage lors de la configuration des paramètres de l'appareil. Le système vise à la sécurité de l'utilisateur, et en fluidifiant son utilisation du smartphone. Le système est très robuste, et il paraît très compliqué de le berner.

4.2 Des robots intelligents dans l'industrie

Dans le milieu de l'industrie, plutôt que de parler de reconnaissance faciale, on aura tendance à parler de reconnaissance visuelle. Diverses entre-

prises utilisent dans leurs usines des robots pour vérifier que leurs produits n'ont pas de défauts de fabrication. Pour cela il faut apprendre au robot à reconnaître des objets visuels, et lui faire associer aussi des mots, afin qu'il explicite les éventuels défauts. Par exemple, Spot est un robot quadrupède travaillant dans une entreprise norvégienne (Aker BP, entreprise pétrolière), son rôle au sein de l'entreprise est d'inspecter les produits, enregistrer des données puis transmettre des rapports. La phase d'inspection est donc celle où l'utilisation de la reconnaissance visuelle est utilisée, pour récupérer ainsi des données et transmettre ses rapports. La reconnaissance faciale/visuelle est également utilisée dans le milieu du transport. Les caméras de recul peuvent être équipées d'un système intelligent pouvant trouver au mieux la trajectoire de la voiture et les obstacles. Les voitures autonomes sont équipées d'un système de lasers, mais aussi de caméras, pour modéliser l'environnement dans laquelle l'automobile se trouve en trois dimensions, et pour identifier les éléments le composant (piétons, panneaux de signalisation, poteaux, voitures). À partir de cela, des commandes sont lancées par un système d'intelligence artificielle, pour que la voiture agisse comme il le faut.

4.3 Des enjeux politiques

La démocratisation de la reconnaissance faciale dans la vie sociétale d'un pays implique un choix de société. C'est pourquoi le débat de l'utilisation de la reconnaissance faciale à des fins nationales ou plus larges encore doit être connu par tous. Certains aspects de l'utilisation par reconnaissance faciale lors d'événements par exemple pour cibler certains individus etc... peut poser des problèmes d'un point de vue éthique. Même si d'un point de vue sécurité cela peut être une grande avancée, cela amènerait à une atteinte à l'anonymat dans l'espace public. Une personne ne pourrait plus se déplacer anonymement dans sa vie quotidienne quand il se déplace dans des espaces publics. Cela toucherait donc à l'anonymat et l'intimité de tous. C'est pourquoi il est important de discuter de cela en société pour ne pas s'embarquer dans quelque chose sans en connaître les bienfaits et les risques, comme l'explique bien le CNIL.