PREAMBULE

Copier les données des TPs

1. créez un répertoire `VO` dans votre homedir
   ```
   cd ~/
   mkdir VO
   cd VO
   ```

2. copier le dossier contenant toutes les données pour les TPs
   `http://people.rennes.inria.fr/Eric.Marchand`

Décompressez l'archive : `unzip att_faces.zip`

Installation de la librairie VISP et création du projet

La librairie VISP est utilisée. Elle doit être installée sur vos postes de travail. Les sources, ainsi que les APIs de cette librairie sont disponibles à l'adresse suivante :
`http://www.irisa.fr/lagadic/visp/visp.html`

Vous devez créer un projet `CMAKE` intitulé `TP-VO` vous permettant de compiler le TP.

Documents à rendre

Le compte rendu de TP seront envoyés par mail à :

- `eric.marchand@irisa.fr`

sous la forme d'une archive contenant votre rapport ainsi que vos fichiers sources :

`tar -cvzf tpintreigenfaces_Eleve1_Eleve2.tar.gz`

Les rapports doivent répondre aux questions posées dans les sujets et être illustrés avec des images issues des traitements que vous avez réalisé.

Eigenfaces

1 Objectifs du projet

Etudier les propriétés de la méthode de reconnaissance de visages par eigenfaces, et son applicabilité pratique. Développer un système capable de : reconnaître un visage à partir d'une base de donnée de visages de référence

On appliquera les outils développés sur la base de face. Dans un deuxième temps, on pourra intégrer les photos du trombinoscope de l'ESIR ou des photos personnelles.

2 Principe géneral

Le problème de la reconnaissance de visages est défini comme suit : étant donnée une image de visage, on souhaite déterminer l'identité de la personne correspondante.

Pour ce faire, il est nécessaire d'avoir des images de référence, sous la forme d'une base de données de visages de toutes les personnes connues par le système. A chaque visage est associé un vecteur de caractéristiques. Ces caractéristiques sont supposées être invariantes pour une même personne, et différentes d'une personne à l'autre. La reconnaissance consiste alors à comparer le vecteur de caractéristiques du visage à reconnaître avec celui de chacun des visages de la base. Ceci permet de retrouver la personne ayant le visage le plus ressemblant, qui est celui dont le vecteur est le plus similaire. Il existe plusieurs types de méthodes, qui se distinguent par le type de caractéristiques utilisées :

- Les approches par modèles de visage procèdent à une analyse biométrique des visages pour déterminer des mesures telles que la distance entre les yeux, la longueur du nez, la forme du menton...
- Les approches "image" comparent au contraire directement les visages, en les considérant comme des images, pour lesquelles des mesures de similarité pré-attentives (sans modèle a priori) sont définies.
- Des approches hybrides utilisent les notions de similarité entre images, mais en rajoutant des connaissances a priori sur la structure d'un visage.

3 Analyse par eigenface

La reconnaissance de visages par eigenfaces est une approche de type " image ". Chaque image de visage est considérée comme un vecteur dans un espace ayant autant de dimensions

que de pixels dans l'image. Les caractéristiques de l'image sont extraites par une méthode mathématique de réduction de dimensionnalité basée sur l'analyse en composante principales (ACP).

Cette approche a été originellement proposée par Turk et Pentland en 1991.

Dans la suite, nous utiliserons la notation italique pour désigner les scalaires (m, K) ainsi que le gras pour les vecteurs $(\mathbf{I}, \mathbf{J}, \mathbf{u})$ et les matrices $(\mathbf{A}, \mathbf{W})$.

On note $\mathbf{I}_k \quad (k = 1..m)$ l'image du visage de référence numéro k, et $\mathbf{I}_k(i) \quad (i = 1..n)$ le pixel numéro i de cette image. Chaque visage peut donc être considéré comme un vecteur à n composantes : l'ensemble des visages forme un nuage de points dans un espace $\mathbf{R^n}$.

On note $\mathbf{\Psi}$ la moyenne de ce nuage de points, ou "visage moyen". Le principe de la méthode des eigenfaces est de modéliser la différence d'un visage quelconque par rapport à ce visage moyen par un ensemble limité d'images uh, appelées eigenfaces. Une image de visage $\mathbf{J} \in bfR^n$ est donc exprimée comme le visage moyen auquel s'ajoute une combinaison linéaire d'eigenfaces :

$$\mathbf{J} = \mathbf{\Psi} + \sum_h w_h \mathbf{u_h} + \varepsilon$$

où w_h représente le poids de l'eigenface d'indice h dans le visage $\mathbf{J}$, et ε représente l'erreur entre $\mathbf{J}$ et son approximation par les eigenfaces. Les coefficients w_h jouent un rôle très important pour la reconnaissance des visages, car ils correspondent aux coordonnées du visage $\mathbf{J}$ dans le sous-espace des visages.

La méthodes des eigenfaces repose sur le fait que le nombre d'eigenfaces considérées est bien inférieur à la dimension totale de l'espace, ce que l'on appelle "réduction de dimensionnalité". Les images sont donc analysées dans un sous-espace de dimension réduite, qui représente plus spécifiquement les visages, parmi tous les types d'images possibles.

Le visage moyen étant toujours le même pour une base de référence fixée, nous considèrerons dans la suite systématiquement chaque visage sous sa forme centrée, c'est-à-dire après soustraction du visage moyen.

4 Mise en pratique

4.1 Base de données de visages

Récupérez la base d'image "faces" mise à disposition.

Dans la suite, nous manipulerons toujours les images de visage sous la forme de vecteurs, et un ensemble de visages sous la forme d'une matrice dont chaque colonne est un visage. Les images seront stockées dans des matrices de réels (double).

En particulier, les images de la base seront stockées dans une matrice $\mathbf{I}$ de taille $n \times m$:

$$\mathbf{I} = [\mathbf{I_1}...\mathbf{I_m}]$$

Question 1 *Calculez la taille mémoire nécessaire pour stocker la matrice* $\mathbf{I}$*, sous la forme d'une matrice de réels au format double.*

Chargez cette matrice à partir des fichiers image. On normalisera l'intensite de chaque pixel entre 0 et 1.

Question 2 *Calculez le visage moyen* $\mathbf{\Psi}$*.*

Question 3 *Affichez le visage moyen, ainsi que quelques visages accompagnés des visages centrés associés.*

5 Analyse en Composantes Principales

La méthode développée par Turk et Pentland définit les eigenfaces comme les axes principaux obtenus en effectuant l'Analyse en Composantes Principales des vecteurs associés aux visages de référence.

Les eigenfaces sont ainsi les vecteurs propres de la matrice de covariance $\mathbf{A}\mathbf{A}^T$, de taille $n \times n$, où la matrice $\mathbf{A}$ de même taille que $\mathbf{I}$ représente l'ensemble des visages centrés :

$$\mathbf{A} = [\mathbf{I_1} - \mathbf{\Psi} \ldots \mathbf{I_m} - \mathbf{\Psi}]$$

Chaque ligne de $\mathbf{A}$ correspond à un pixel p, chaque colonne de $\mathbf{A}$ correspond à un visage de référence de numéro k.

Remarque Plutôt que d'utiliser la décomposition en valeurs propres, nous utiliserons la décomposition en valeurs singulières (SVD, ou singular value decomposition).

La SVD décompose la matrice $\mathbf{A}$ de taille $n \times m$ en :

$$\mathbf{A} = \mathbf{U}.\mathbf{S}.\mathbf{V^T}$$

où $\mathbf{U}$ et $\mathbf{V}$ sont des matrices orthonormales ($\mathbf{U}.\mathbf{U^T} = \mathbf{U^T}.\mathbf{U} = \mathbf{I_n}$ et $\mathbf{V}.\mathbf{V^T} = \mathbf{V^T}.\mathbf{V} = \mathbf{I_m}$ de tailles respectives $n \times n$ et $m \times m$, et $\mathbf{S}$ est une matrice $n \times m$ nulle partout sauf sur la diagonale principale.

Cette décomposition présente les propriétés suivantes :

- les colonnes de $\mathbf{V}$ sont les vecteurs propres de $\mathbf{A^T}.\mathbf{A}$,
- les colonnes de $\mathbf{U}$ sont les vecteurs propres de $\mathbf{A}.\mathbf{A^T}$,
- les carrés des valeurs singulières s_k de $\mathbf{S}$ sont les valeurs propres λ_k de $\mathbf{A^T}.\mathbf{A}$ et $\mathbf{A}.\mathbf{A^T}$.

Question 4 *Que represente la matrice* $\mathbf{U}$
Question 5 *Calculez la matrice des eigenfaces.*
Question 6 *Affichez le visage moyen ainsi que les premières eigenfaces et les valeurs propres associées.*

Dans la suite, nous utiliserons un nombre réduit d'eigenfaces afin de modéliser l'espace de visages représenté sous la forme de la base $\mathbf{U'_K}$ des K premiers vecteurs propres :

$$\mathbf{U'_K} = [\mathbf{u_1} ... \mathbf{u_K}]$$

Notez que les colonnes forment un ensemble orthogonal, donc $\mathbf{U'^T}.\mathbf{U'} = \mathbf{I_K}$.

6 Projection dans le sous-espace des visages

La projection d'une image dans le sous-espace des visages se fait simplement en soustrayant le visage moyen et en effectuant le produit scalaire de l'image obtenue avec chaque eigenface. Ceci donne les coordonnées de l'image test dans le sous-espace des visages, qui est de dimension K. Chaque visage possède donc plusieurs représentations :

- Son image d'origine, représenté par un vecteur $\mathbf{J}$ général dans $\mathbf{R^n}$
- Son image projetée dans le sous-espace des visages, $\mathbf{J_p}$
- Les coordonnées de l'image projetée dans la base des eigenfaces, W_k

Compte tenu des propriétés d'une base orthonormée, le vecteur $\mathbf{W_k}$ s'obtient par simple produit scalaire entre l'image centrée et chacune des eigenfaces :

$$W_k = \mathbf{u}_k^T.(\mathbf{J} - \mathbf{\Psi})$$

Il permet de calculer l'image projetée:

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

Cette image projetée est aussi appelée "image reconstruite", pour mettre en avant qu'elle est issue de la recomposition d'une image à partir du visage moyen et des eigenfaces.

L'erreur de reconstruction (ou de projection) est définie comme la distance entre une image et l'image reconstruite (projetée) associée:

$$E(\mathbf{J}) = \parallel \mathbf{J} - \mathbf{J_p} \parallel$$

Question 7 *Calculez les coordonnées d'un visage dans le sous-espace des visages, et calculez l'image reconstruite correspondante, pour plusieurs valeurs de K*

Question 8 *Affichez le résultat pour plusieurs visages, de référence et de test, et plusieurs valeurs de K. Y a-t-il une différence entre les visages de référence et les visages de test ? Pourquoi ? Remarque : Attention à ajouter/soustraire l'image moyenne correctement dans les calculs et pour l'affichage.*

Question 9 *Vérifiez qu'il n'y a pas d'erreur de reconstruction pour les visages de référence lorsque K=m : chaque visage de référence est-il bien identique à sa reconstruction ?*

Question 10 *Normalisez les valeurs propres afin que leur somme soit égale à 100%, puis tracez la courbe de leur somme cumulée, afin de voir combien de variation est capturée par les K premières eigenfaces. Combien d'eigenfaces sont nécessaires pour obtenir une "bonne" reconstruction ?*

Question 11 *Tracez l'évolution de la moyenne de l'erreur de reconstruction des visages de test lorsque K varie de 1 à m. Cette évolution est-elle cohérente avec la somme cumulée ?*

7 Identification

A chaque visage de référence $\mathbf{I^k}$ on associera une identité, sous la forme d'un numéro $id(k)$. Nous chercherons dans cette partie à identifier un visage test à partir des visages de référence.

La méthode la plus simple consiste à comparer la projection $\mathbf{J_p}$ du visage test $\mathbf{J}$ avec la projection $\mathbf{J_p^k}$ de chaque image de référence $\mathbf{I^k}$. La dissimilitude entre les deux est quantifiée par la distance dans le sous-espace $E_k(J)$:

$$E(\mathbf{J}) = \| \mathbf{J} - \mathbf{J_p^k} \|$$

En évaluant cette distance pour chaque visage de référence, on peut déterminer le visage de référence $\mathbf{I^k}$ le plus proche du visage test $\mathbf{J}$. On appelle alors erreur d'identification la distance correspondante dans le sous-espace. Si l'erreur d'identification est inférieure à un seuil θ, on peut estimer qu'il s'agit d'un visage reconnu d'identité, sinon le visage est classifié comme inconnu.

Etant donné un corpus de visages de référence et de visages de test, le taux d'identification est défini comme le ratio entre le nombre de visages correctement identifiés et le nombre de visages testés. Il vaut 1 lorsque tous les visages testés sont correctement identifiés.

Question 12 *Montrez que l'on peut estimer la distance $E_k(J)$ dans le sous-espace. Quel est l'intérêt de procéder ainsi ?*

Question 13 *Pour l'instant on prend K=20. Calculez pour chaque visage du jeu de référence sa distance dans le sous-espace par rapport à l'ensemble des visages de référence. On pourra par exemple visualiser le résultat sous la forme d'une image de matrice. Que constate-t-on ?*

Question 14 *Quelles sont les distances min et max entre deux visages de la même classe (même personne) ? entre deux visages de classes différentes ? Quelles conclusions peut-on en tirer concernant θ*

Question 15 *Tracez la courbe du nombre de visages reconnus en fonction de K sans tenir compte de θ Quelle valeur de K peut-on prendre pour avoir une bonne reconnaissance et un temps de calcul faible ?*