



PROJET
Analyse de Données / Calcul Scientifique :
A quoi tu ressembles sans ton masque ?

DAI Guohao
DANTON LALOY Calliopé
DE ROCKER Tom

Département Sciences du Numérique - Première année
2020-2021

1 Séance 3

Question 3

Identification des paramètres de notre classifieur et recherche de valeurs optimales, avec ou sans masque.

Les paramètres de notre classifieur sont les suivants :

- le nombre de composantes principales (N), pouvant être fixé ou déterminé par le pourcentage d'informations souhaité (per)
- nombre de voisins considérés dans l'algorithme des kppv (k)

Après quelques expérimentations, on a choisi les paramètres suivants :

- $per = 0.9$
- $k = 3$

Question 5

Evaluation de la reconnaissance du classifieur. On utilise les images reconstruites correspondant à plusieurs requêtes. Imagination d'une évaluation de la reconstruction.

Il est assez peu évident de trouver un moyen d'évaluer "à quel point notre classifieur est bon", dans quelle mesure il aurait pu trouver un "meilleur résultat". Il est de plus à noter que la base d'apprentissage reste assez restreinte.

Nous avons remarqué que lors de la reconstruction, la posture considérée est parfois plus importante que la ressemblance des traits. La considération d'une bonne personne mais avec une posture très différente peut entraîner des reconstitutions que l'on peut considérer comme étant "moins bonne". Nous avons donc décidé d'évaluer notre classifieur en tenant compte de la bonne reconnaissance des postures.

Pour cela, nous avons fait une matrice de confusion avec sur les lignes la posture réelle de l'image considérée (6 labels), et sur les colonnes la posture de l'image la plus proche considérée dans la base d'apprentissage (4 labels). On remarque qu'il sera impossible d'obtenir une reconstitution avec la bonne posture dans certains cas, car certaines postures ne figurent pas dans la base de donnée. Cependant, certaines postures sont plus proches que d'autres.

D'un point de vue humain, nous avons considérés :

- que les postures 1, 3 et 5 sont très proches
- que les postures 2, 4 et 6 sont très proches

On construit la matrice de confusion dans le script *test_reconnaissance.m* :

Posture	1	2	3	4
1	16	0	14	2
2	3	8	4	17
3	11	0	19	2
4	2	4	4	22
5	11	0	17	4
6	2	5	5	20

TABLE 1 – Matrice de confusion réalisée sur l'ensemble de la base de données.

On refait la table en considérant les postures 1-3-5 comme "posture A", et les postures 2-4-6 comme "posture B" :

Au vu des postures trouvées, notre classification semble pertinente.

Posture	Posture A	Posture B	Taux erreur
1 (A)	30	2	0.0625
2 (B)	7	25	0.2188
3 (A)	30	2	0.0625
4 (B)	6	26	0.1875
5 (A)	28	4	0.125
6 (B)	7	25	0.2188

TABLE 2 – Matrice de confusion réalisée en regroupant les différentes postures en 2 labels.