ÉCOLE POLYTECHNIQUE MONTRÉAL

Applications des mathématiques:

Reconnaissance des visages

Mathématiques Appliquées et Génie Industriel

Résumé On montre comment il est possible de reconnaître des visages à l'aide d'une matrice utilisant les vecteurs propres.

Domaines du génie Logiciel, Informatique

Notions mathématiques Vecteurs Propres, Matrice Symétrique, Matrice de Covariance

Cours pertinents Algèbre linéaire, Probabilités et Statistiques

Auteur(es) M.Laforest, I. Jalliffier-Verne

Sommaire

1 Introduction

La prochaine frontière de l'informatique consiste à construire des machines qui se comportent comme des humains. Bien qu'il soit très facile pour un humain de reconnaître que les dix images de la figure 1 proviennent de la même personne, s'assurer qu'un ordinateur en soit capable est encore un défi scientifique de taille. Si un jour nous somme capables de créer un tel logiciel informatique, alors on pourra avoir des applications immédiates à la reconnaissance :

- des visages de criminels ou de terroristes ;
- des employés autorisés dans un laboratoire de haute sécurité ;
- de textes et formules ;
- ou même des objets à partir de photos aériennes.

Dans cet exemple, nous expliquerons comment on peut réduire le problème de reconnaissance d'un visage à celui d'un calcul de vecteurs propres.

2 Modélisation

Chaque image d'une personne est formée d'un nombre p de pixels où chaque pixel peut prendre une valeur comprise entre 0 et 255 représentant un ton de gris. Pour chaque image on a donc un vecteur colonne avec p composantes :

$$\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m.$$

Figure 1: Les 10 images de la 1ère personne.

À partir de ces vecteurs, on peut calculer la valeur moyenne de chaque pixel :

$$\overline{\mathbf{u}} = \frac{1}{m} \sum_{j=1}^{m} \mathbf{u}_j$$

et construire une matrice $p \times m$ dont les colonnes sont les images :

$$U = (\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m).$$

Soit la matrice de déviation par rapport à la moyenne $\overline{\mathbf{u}}$:

$$E = \big((\mathbf{u}_1 - \overline{\mathbf{u}}), (\mathbf{u}_2 - \overline{\mathbf{u}}), \cdots, (\mathbf{u}_n - \overline{\mathbf{u}})\big).$$

On sait que pour une variable aléatoire X, sachant que $\mu_X = \mathsf{E}[X]$, sa variance est :

$$\sigma_X^2 = \mathsf{E}[(X - \mu_X)^2].$$

Avec 2 variables X et Y, la généralisation de cette notion est la covariance :

$$\sigma_{XY} = \mathsf{E}[(X - \mu_X)(Y - \mu_Y)].$$

Cette covariance est grande si la variance par rapport à leur moyenne est identique.Si l'on cherche la corrélation entre le $i$ème et le $j$ème pixels, il faut regarder la covariance :

$$\sigma_{ij} = \frac{1}{m} \sum_{k=1}^{M} e_{ik} e_{jk}.$$

On écrit e_{ij} pour la composante de E sur la $i$ème ligne et la $j$ème colonne.

Alors on définit la matrice de corrélation :

$$EE^T = \begin{pmatrix} e_{11} & e_{21} & \cdots \\ e_{12} & e_{22} & \cdots \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} e_{11} & e_{12} & \cdots \\ e_{21} & e_{22} & \cdots \\ \vdots & \vdots & \vdots \end{pmatrix} = \begin{pmatrix} e_{11}^2 + e_{21}^2 + e_{12}^2 & e_{22}^2 & e_{11}e_{12} + e_{21}e_{22} + \cdots & \cdots \\ e_{11}e_{12} + e_{21}e_{22} + \cdots & \ddots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{pmatrix}$$

Les éléments de la matrice EE^T s'écrivent ainsi :

$$(EE^T)_{ij} = \sum_{k=1}^{m} e_{ik} e_{jk}$$

Intuitivement, la valeur de corrélation sera maximale si le $j$ème pixel est toujours proportionnelle au $i$ème pixel.

Or, $A = EE^T$ est une matrice symétrique car

$$A^T = (EE^T)^T = (E^T)^T E^T = EE^T = A$$

et A est définie positive car, pour tout vecteur $\mathbf{u}$:

$$\mathbf{u}^T A \mathbf{u} = \mathbf{u}^T E E^T \mathbf{u} = (E^T \mathbf{u})^T E^T \mathbf{u} > 0.$$

Selon un théorème en algèbre linéaire, pour toute matrice symétrique, il existe une matrice orthogonale P qui diagonalise A.

$$\begin{aligned} D &= P^T A P \\ &= P^T (EE^T) P \\ &= (E^T P)^T E^T P. \end{aligned}$$

La matrice D devient donc une matrice de covariances pour la matrice $E^T P$.

Dans cette nouvelle base de vecteurs propres, la matrice D est diagonale c'est à dire que les images ne sont donc plus corrélées entre elles car chaque vecteur qui la représente est indépendant des autres et, donc, chacune des images contient de l'information différente de toute autre image. L'image dont le vecteur a le plus grand nombre d'informations correspondra alors au vecteur propre (la colonne de P) avec la plus grande valeur propre.

En géométrie, lorsque deux vecteurs sont proches, leur produit scalaire est élevé. C'est en utilisant la même technique que l'on peut savoir qui est qui.

Pour reconnaître une image, il suffit ainsi d'extraire les vecteurs propres présents dans la matrice P, et de multiplier le vecteur colonne représentant l'image par l'un des vecteurs propres. En comparant le résultat, avec les vecteurs propres des autres images dans la base donnée, on peut en déduire quelle est la personne.

3 Résolution

Dans notre exemple, l'image a 92×112 pixels donc notre vecteur colonne pour chaque image sera long de 10304 lignes.

Regardons tout d'abord deux personnes et le rendu des trois premiers vecteurs propres :

Figure 2: Les 10 images de la 1ère personne.

Figure 3: Les 3 premiers vecteurs propres de la 1ère personne.

Figure 4: Les 10 images de la 2ème personne.

Figure 5: Les 3 premiers vecteurs propres de la 2ème personne.

On observe que les vecteurs propres nous donnent un rendu relativement proche de l'image, nous pourrions presque y reconnaître la personne.

On commence avec une base de données de 9 images pour la personne que l'on cherche à reconnaître.

Figure 6: Les 9 images de la personne à reconnaître.

La caméra reçoit deux images :

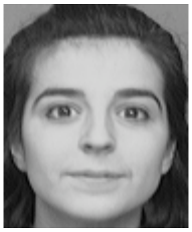

Figure 7: Nouvelle image de la personne à reconnaître.

Figure 8: Image d'un inconnu.

Les vecteurs propres utilisés sont :

Figure 9: Les vecteurs propres $\mathbf{v}_1$,$\mathbf{v}_2$ et $\mathbf{v}_3$.

Les résultats des produits scalaires sont :

	$\mathbf{u} \cdot \mathbf{v}_1$	$\mathbf{u} \cdot \mathbf{v}_2$	$\mathbf{u} \cdot \mathbf{v}_3$
Personne à reconnaître	21 586 548	20 687 999	23 170 296
Personne inconnue	21 204 953	18 246 567	22 044 073

On remarque que, dans les 3 cas, le produit scalaire est plus élevé pour la figure 7 que pour la figure 8. Ainsi, nous avons pu reconnaître 3 fois la bonne image. On observe tout de même que les valeurs sont relativement proches et que le seuil numérique de reconnaissance pourrait être problématique.

4 Conclusion

L'ensemble des données devant être utilisé pour pouvoir reconnaître de manière certaine les personnes est beaucoup trop grand pour créer un système de reconnaissance faciale efficace pour une grande population. Cependant, cela permet dès aujourd'hui de reconnaître des délinquants déjà identifiés dans les dossiers de police ou encore il permet, dans un autre registre, de reconnaître les lettres écrites sur un PDA ou des commandes auditives.

Les vecteurs propres ainsi que le produit scalaire sont des opérations mathématiques relativement simples qui permettront peut être, un jour, de reconnaître les personnes déambulant dans une ville pour en assurer la sécurité, entre autre.

Références

[1] KIM, K., Face Recognition using Principle Component Analysis, University of Maryland, http ://www.umiacs.umd.edu/∼knkim/KG_VISA/PCA/FaceRecog_PCA_Kim.pdf. Page consultée le 24 juillet 2009.

[2] Images. Laboratoire de recherche Olivetti, Université de Cambridge, Royaume Uni.