

Estimation des données occultées : modèles statistiques à noyau

Barbara ROMANIUK¹, Michel DESVIGNES², Maxime BÉRAR².

¹Laboratoire CReSTIC LERI, URCA
Rue des Crayères, BP 1035, 51687 Reims Cedex 2, France

²Laboratoire LIS
961, Rue de la houille blanche, BP 46 St Martin d'Herès, France

Barbara.Romaniuk@leri.univ-reims.fr, michel.desvignes@lis.inpg.fr, berar@lis.inpg.fr

Résumé – Dans cet article nous proposons d'utiliser des modèles statistiques non linéaires dans le cadre de la reconnaissance de formes. Les objets que nous visons à reconnaître sont définis par des ensemble ordonnés de points. Nous présentons ici deux modèles : le premier repose sur la définition d'une fonction de projection explicite, le second est inspiré de l'analyse en composantes principales à noyau (Kernel Principal Component Analysis). Ces travaux ont pour objectif l'estimation de la position d'objets partiellement visibles. Les deux modèles ont été validés en céphalométrie avec de bons résultats.

Abstract – In this paper, we deal with the pattern recognition problem using non-linear statistical models based on Kernel Principal Component Analysis. Objects that we try to recognize are defined by ordered sets of points. We present here two types of models: the first one uses an explicit projection function, the second one uses the Kernel trick. The present work attempts to estimate the localization of partially visible objects. Both are applied to the cephalometric problem with good results.

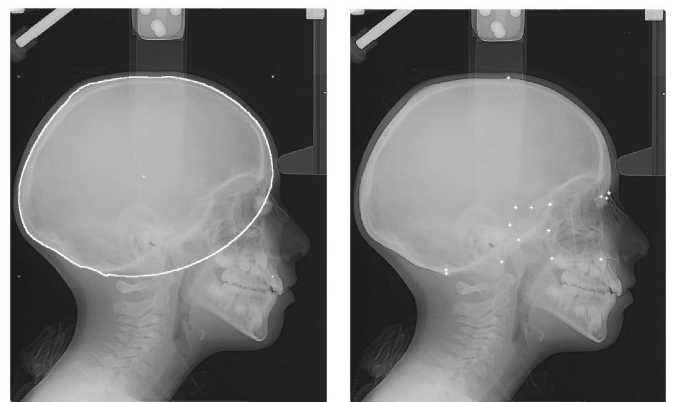
1 Introduction

Notre approche s'inscrit dans le cadre de la reconnaissance de formes statistique. L'un des pionniers en analyse de formes est Kendall[5]. Il assimile une forme à l'information restante après alignement de deux objets. Cette définition est réutilisée par Cootes[2] dans les modèles de formes et d'apparence. Composés d'une forme moyenne et de la variabilité autorisée autour de cette forme moyenne, ces deux types de modèles utilisent l'Analyse en Composantes Principales (ACP). Un modèle de profil d'intensité intervient en plus dans le modèle d'apparence. Ce dernier reste valable sous l'hypothèse d'une distribution linéaire. Quelques travaux récents proposent des modèles d'intensité non linéaires basés sur des classificateurs de type K plus proches voisins et sur la sélection de caractéristiques locales similaires à l'analyse de texture[3] ou sur le recalage sur l'image d'un modèle inspiré de l'algorithme de filtrage particulaire[1]. Ces derniers restent cependant applicables si les caractéristiques recherchées sont localisées sur des contours bien marqués dans l'image.

Dans cet article nous nous intéressons plus particulièrement à l'application en céphalométrie. Son objectif est de localiser des points anatomiques sur les radiographies crâniennes en vue d'un diagnostic médical. La définition anatomique des points céphalométriques est cependant difficilement applicable sur les radiographies. Le repérage s'avère difficile et est sujet à une variabilité d'environ 2 millimètres. Les hypothèses nécessaires à l'utilisation des modèles d'apparence n'y sont pas vérifiées[4].

Dans notre problème, nous disposons d'une base d'images radiographiques expertisées (points anatomiques appelés points céphalométriques) et de connaissances *a priori* (la position des points recherchés est relative à la forme du crâne). Le contour crânien est détecté automatiquement dans chaque image. L'ensemble d'apprentissage est composé des points céphalométri-

ques et des points issus de l'échantillonnage du contour crânien (*c.f.* figure 1).



Contour crânien détecté automatiquement.

Points anatomiques (expertise humaine).

FIG. 1: Base d'apprentissage en céphalométrie

Nous proposons d'aborder le problème d'estimation de points caractéristiques par une approche statistique. Notre modèle doit être capable de retrouver les points anatomiques sur une image inconnue. Sachant que le contour crânien peut être détecté sur toutes les images, le problème que nous posons est celui d'identifier l'occultation partielle d'une observation à partir de la partie visible. L'idée majeure consiste à utiliser la variabilité, l'observation moyenne ainsi que les relations spatiales existantes entre les caractéristiques de l'observation. Nous présentons ici la comparaison de deux modèles non linéaires. Les deux modèles sont appris dans un espace caractéristique non linéaire. Le premier modèle est basé sur l'ACP à noyau (Kernel PCA) et repose sur une projection implicite, le second sur une projection explicite.

2 Méthodologie

En céphalométrie une observation est partiellement connue. Elle est composée des coordonnées des points issues de l'échantillonnage du contour crânien détecté automatiquement ainsi que de celles des points céphalométriques dont la position est supposée inconnue. Le problème que nous visons à résoudre est de déterminer la partie inconnue de l'observation en utilisant un modèle statistique non linéaire et la partie connue de l'observation. Soit X l'observation à reconstruire (les données connues sont placées en tête de l'observation). Le modèle statistique permet alors de déterminer les paramètres de variabilité de l'observation étudiée par rapport à une observation moyenne. Trouver la partie inconnue de X est équivalent à trouver la forme appartenant au modèle dont les premières coordonnées sont données par la partie connue de X .

Nous proposons dans ce cadre deux modèles non linéaires différents. Le premier est un modèle implicite basé sur l'ACP à noyau. Le second en revanche est un modèle explicite basé sur les rapports d'aires algébriques de triangles.

2.1 Modèle implicite basé sur l'ACP à noyau

2.1.1 L'ACP à noyau - Kernel PCA

L'ACP à noyau consiste en une projection des données dans un espace caractéristique de grande dimensionnalité \mathcal{F} , puis en une ACP dans \mathcal{F} . La fonction de projection est non linéaire. Dans un problème de localisation, la forme moyenne explicitée dans \mathcal{F} doit être projetée dans l'espace d'origine. Le choix de la fonction de projection inverse s'avère problématique.

Soit n le nombre d'observations présentes dans l'ensemble d'apprentissage. Soient X_i $i \in \{1, \dots, n\}$ les observations. Soit φ la fonction de projection non linéaire telle que :

$$\begin{aligned} \varphi : \mathbb{R}^n &\mapsto \mathcal{F} \\ X &\mapsto \varphi(X) \end{aligned}$$

Soit $k(X_i, X_j)_{(i,j) \in \{1, \dots, n\}^2}$ un noyau de Mercer. Ce noyau peut être utilisé pour calculer une matrice K définie positive. Il permet de reformuler une projection en terme de produit scalaire : $k(X_i, X_j) = \langle \varphi(X_i), \varphi(X_j) \rangle$.

Le noyau que nous utiliserons est le noyau gaussien : $k(X, Y) = \exp\left(-\frac{\|X-Y\|^2}{2\sigma^2}\right)$.

Dans son algorithme Schlkopf remplace le calcul de la matrice de variance-covariance dans \mathcal{F} (ACP) par celui de la matrice K (matrice de Gram). Cet algorithme se résume à trois étapes :

- 1. Calcul de la matrice K : $K_{ij} = \langle \varphi(X_i), \varphi(X_j) \rangle$.
- 2. Calcul des valeurs propres et vecteurs propres de K . Leur normalisation : $\lambda_k \langle \alpha^k, \alpha^k \rangle = 1$.
- 3. Calcul des projections sur les composantes principales : $\beta_k = \sum_{i=1}^n \alpha_i^k k(X_i, X)$.

L'algorithme de Schölkopf s'arrête là. Afin de résoudre le problème de retroprojection dans l'espace d'origine il est intéressant de définir la fonction que nous serons amenés à minimiser par rapport à φ . Cette fonction est :

$$P_l \varphi(X) = \sum_{i=1}^l \beta_i V^i.$$

Cet algorithme est valable pour des observations centrées :

$$\sum_{i=1}^n \varphi(X_i) = 0.$$

Le centrage des données est facile à assurer dans l'espace d'origine, beaucoup plus difficile dans l'espace \mathcal{F} , puisque l'on ne peut pas calculer explicitement la moyenne des observations projetées dans cet espace. Il existe cependant une solution à ce problème. Il s'agit de calculer la matrice de Gram à la place de la matrice K . La matrice à diagonaliser devient alors :

$$G = K - 1_n K - K 1_n + 1_n K 1_n,$$

où $1_{nij} = \frac{1}{n}$ pour tout $\{i, j\} \in \{1, \dots, n\}^2$.

2.1.2 Projection implicite et pseudo-inverses utilisant l'astuce noyau

Chaque observation est composée de m coordonnées points caractéristiques (contour crânien échantillonné et points anatomiques). Notre objectif est d'identifier p coordonnées inconnues d'une observation quelconque, où $p < \frac{m}{2}$. Pour ceci nous sommes amenés à travailler dans trois espaces différents (figure 2). Le premier est l'espace image (espace d'origine) de dimension m (nombre de coordonnées des observations), le second est l'espace caractéristique \mathcal{F} de dimension $L > m$, le troisième est l'espace Kernel PCA de dimension n . Ce type de schéma a été proposé dans [6]. Nous proposons ici d'étudier la projection directe entre l'espace Kernel PCA et l'espace image (projection 5, figure 2), projection non étudiée jusqu'à présent. Cette étude nous conduit à expliciter des pseudo-inverses en effectuant des minimisations multidimensionnelles.

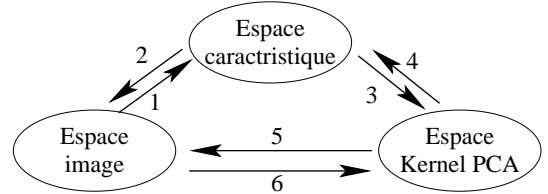


FIG. 2: Trois différents espaces.

Soient X une observation dans l'espace image et X_i la $i^{\text{ème}}$ observation dans l'ensemble d'apprentissage. Soient φ la fonction non linéaire de projection entre l'espace image et l'espace Kernel PCA et β les coordonnées d'une observation dans l'espace Kernel PCA. Soit Z_{kpca} la pre-image, qui correspond à l'observation X reconstruite dans l'espace image à partir de l'espace Kernel.

Pseudo-inverse utilisant β

Considérant une observation X_{kpca} dans l'espace Kernel PCA nous pouvons calculer son image Z_{kpca} en minimisant :

$$\|X_{kpca} - \beta\|^2,$$

avec $\beta = (\beta_1, \dots, \beta_l)$ les projections de X_{kpca} sur l composantes principales : $\beta_k = \sum_{i=1}^n \alpha_i^k k(X_i, X_{kpca})$. En développant cette équation nous obtenons la pseudo-inverse utilisant β :

$$\left\| \sum_{i=1}^n \alpha_i k(X_i, Z_{kpca}) - \beta \right\|^2. \quad (1)$$

Pseudo-inverse utilisant Z_{kpca}

Il est possible de reformuler la pseudo-inverse précédente en exprimant $k(X_i, Z_{kpca})$ en fonction de β . Nous savons que :

$$\beta_k = \sum_{i=1}^n \alpha_i^k k(X_i, Z_{kpca}).$$

En réécrivant les β_k en utilisant un formalisme matriciel on obtient :

$$\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} \alpha_1^1 & \cdots & \alpha_n^1 \\ \vdots & \ddots & \vdots \\ \alpha_1^n & \cdots & \alpha_n^n \end{bmatrix} \begin{bmatrix} k(X_1, Z_{kpca}) \\ \vdots \\ k(X_n, Z_{kpca}) \end{bmatrix}. \quad (2)$$

Soit B la matrice :

$$B_{ij} = \langle \alpha^i, \alpha^j \rangle.$$

α^i sont les pseudo-vecteurs propres de la matrice K calculés par rapport à ses vecteurs propres V^i et ses valeurs propres λ^i :

$$\begin{aligned} \alpha^i &= \frac{V^i}{\sqrt{\lambda^i}}, \\ \langle \alpha^i, \alpha^j \rangle &= \frac{\langle V^i, V^j \rangle}{\sqrt{\lambda^i \lambda^j}}. \end{aligned}$$

Nous savons que les V^i forment une base orthonormée. Nous obtenons alors :

$$\langle \alpha^i, \alpha^i \rangle = \frac{1}{\lambda^i} \quad \text{et} \quad B = \begin{bmatrix} \frac{1}{\lambda^1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\lambda^n} \end{bmatrix}.$$

Nous en déduisons B^{-1} . Nous pouvons alors calculer α^{-1} :

$$\alpha^{-1} = \alpha^t B^{-1}.$$

L'équation (2) devient :

$$\begin{bmatrix} \alpha_1^1 & \cdots & \alpha_n^1 \\ \vdots & \ddots & \vdots \\ \alpha_1^n & \cdots & \alpha_n^n \end{bmatrix} \begin{bmatrix} \lambda^1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda^n \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} k(X_1, Z_{kpca}) \\ \vdots \\ k(X_n, Z_{kpca}) \end{bmatrix}. \quad (3)$$

La pseudo-inverse associée à Z_{kpca} est alors définie par :

$$\left\| k(X_j, Z_{kpca}) - \sum_{i=1}^n \alpha_i^j \lambda^j \beta^i \right\|^2. \quad (4)$$

2.2 Modèle explicite

Le deuxième modèle est explicite. Il est basé sur l'élaboration d'une fonction de projection explicite entre l'espace image et l'espace de caractéristique dans lequel la variabilité est modélisée.

Fonction de projection

La première étape consiste en la détection de la forme de référence (contour crânien). Cette forme est ensuite échantillonnée en p points équidistants $P_i, i \in \{1, \dots, p\}$.

L'espace caractéristique non linéaire est alors défini par les rapports de surfaces de triangles obtenus à partir de l'échantillonnage précédent. Les coordonnées d'un point de l'image $M(x, y)$ sont β, γ et δ calculés pour chaque triangle possible :

$$\beta = \frac{\overline{P_j M P_k}}{\overline{P_i P_j P_k}} \quad \gamma = \frac{\overline{P_k M P_i}}{\overline{P_i P_j P_k}} \quad \delta = \frac{\overline{P_i M P_j}}{\overline{P_i P_j P_k}}$$

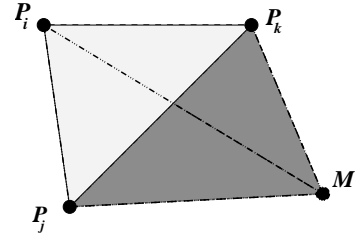


FIG. 3: Nouvelles coordonnées d'un point M

où $\overline{P_i P_j P_k}$ est l'aire algébrique du triangle $P_i P_j P_k$.

Ces coordonnées vérifient :

$$\beta \times \overline{M P_i} + \gamma \times \overline{M P_j} + \delta \times \overline{M P_k} = \overline{0}.$$

Soit l le nombre de triangles obtenus à partir de l'ensemble des points P_i . Les nouvelles coordonnées de M sont :

$$\vartheta = [\beta_1 \ \gamma_1 \ \delta_1 \ \dots \ \beta_l \ \gamma_l \ \delta_l]^t = A' M,$$

où A' est une matrice utilisée pour projeter les données de l'espace cartésien dans notre nouvel espace caractéristique.

Le point fort de cette projection est le fait que les données sur lesquelles nous travaillons ne doivent pas être nécessairement centrées. Cette projection assure leur invariance vis à vis des transformations affines, propriété très intéressante lorsqu'on travaille dans le cadre d'une application médicale.

Modélisation de la variabilité et des relations spatiales

L'apprentissage est effectué sur une base composée de N images expertisées. Pour chaque image nous détectons la forme de référence et on l'échantillonne. Pour chaque image i nous disposons d'un ensemble de points $\{P_k\}_{k \in \{1, \dots, p\}}$ d'une matrice A'^i et d'un ensemble de $m' = \frac{m}{2}$ coordonnées de points anatomiques $\{C'^i_j\}$. On calcule la position moyenne de chaque point caractéristique. Soit ϑ^i le vecteur représentant les coordonnées dans de nouvel espace d'un point caractéristique dans l'image i . La position moyenne de ce point est donnée par :

$$\hat{\vartheta} = \frac{1}{N} \sum_{i=1}^N \vartheta^i.$$

La variance $\hat{\sigma}$ des vecteurs ϑ^i est aussi calculée. Nous en déduisons une matrice de pondération P :

$$P = \begin{bmatrix} \frac{1}{\hat{\sigma}_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\hat{\sigma}_{3l}} \end{bmatrix}.$$

La matrice P permet de donner plus d'importance aux coordonnées robustes : Lorsque l'échantillonnage de la forme de référence donne un nombre important de points, seuls certains sont importants. On propose alors d'appliquer une ACP sur la matrice de covariance des vecteurs ϑ^i . Seuls les d' composantes sont retenues. Ces composantes correspondent aux valeurs propres les plus élevées, elle forment la matrice Φ .

Estimation des points

Soit ϑ le vecteur représentant les points caractéristique C dans le nouvel espace. Estimer C dans une nouvelle image consiste en la résolution du système : $\hat{\vartheta} = A' C$, où $\hat{\vartheta}$ est le vecteur moyen appris et A' la matrice définie relativement aux caractéristiques de la nouvelle image. On résout ce problème en

utilisant les moindres carrés pondérés. La position estimée \tilde{C} du point C dans une image inconnues est donnée par l'équation :

$$\tilde{C} = (A^t P^t \Phi \Phi^t P A')^{-1} A^t P^t \Phi \hat{\vartheta}.$$

3 Résultats

En céphalométrie, le contour crânien est échantillonné en 6 points, soit 12 coordonnées. Nous avons testé les méthode pour l'estimation de 14 points céphalométriques sur 80 images et ceci avec l'approche *leave-one-out*. Dans la méthode utilisant les noyaux on a utilisé la méthode de Powell pour la minimisation.

pts	E_x	E_y	σ_x	σ_y
NA	3.8	4.2	2.93	3.37
M	3.9	4.2	2.83	3.44
FM	4.1	4.0	3.10	3.36
SE	4.6	3.5	3.37	3.63
TPS	4.8	3.8	3.77	3.54
CLP	4.9	4.3	3.45	3.79
SSO	4.8	3.9	3.70	3.47
BA	5.8	3.8	3.76	3.50
CT	5.4	3.3	3.52	2.90
OP	7.2	4.8	4.78	3.65
OB	6.5	4.5	4.56	3.38
PTS	4.6	3.3	3.23	3.25
PTI	4.5	3.2	3.42	2.51
BR	7.4	5.1	4.87	3.88
Moyenne	5.2	4.0	3.66	3.41
(2)				

TAB. 1: Erreur moyenne et écart type (en millimètres) associés à la méthode basée sur la pseudo-inverse utilisant β (calcul effectué avec un noyau gaussien et $\sigma = 0.005$).

pts	E_x	E_y	σ_x	σ_y
NA	0.6	1.6	0.73	1.83
M	1.1	1.7	0.94	1.84
FM	1.3	1.6	1.00	1.78
SE	2.1	2.3	1.68	1.91
TPS	2.3	2.2	1.87	1.63
CLP	2.5	2.3	1.85	1.70
SSO	2.8	2.5	2.17	1.94
BA	4.1	2.3	3.22	1.82
CT	3.4	2.0	2.67	1.61
OP	4.9	2.2	4.37	1.87
OB	4.6	2.1	4.04	1.73
PTS	2.5	2.2	2.09	1.77
PTI	3.4	2.3	2.83	1.82
BR	4.5	1.2	3.70	0.90
Moyenne	2.9	2.0	2.93	1.78
(1)				

TAB. 2: Erreur moyenne et écart type (en millimètres) associés à la méthode basée sur une fonction de projection explicite.

Le tableau 1 présente les résultats obtenus en utilisant la pseudo-inverse utilisant β et l'astuce noyau. Le tableau 2 présente

les résultats obtenus en utilisant la fonction de projection explicite (données invariantes par transformations affines). Les résultats obtenus montrent que la deuxième méthode donne de meilleurs résultats approchant la variabilité inter-experte. La méthode utilisant les noyaux est moins performante du fait de la minimisation numérique qui ne converge pas toujours dans le cadre général. Des améliorations utilisant les expressions analytiques des noyaux dans ce formalisme sont en cours de réalisation. L'avantage de la méthode basée sur les noyaux réside dans le fait que la fonction de projection n'a pas à être explicite et calculé.

4 Conclusion

Dans cet article, nous avons présenté et comparé deux méthodes de reconstruction des parties invisibles d'un objet en utilisant une approche statistique. Le cadre statistique offre un moyen élégant pour résoudre ce problème, en utilisant aussi bien la variabilité autorisée par le modèle que les relations spatiales existantes entre les différentes caractéristiques de l'objet modélisé. Il apparait que les modèles basés sur les Kernel PCA sont intéressants pour résoudre ce problème, même si une fonction de projection explicite dans \mathcal{F} donne de meilleurs résultats. La définition d'une telle fonction est un problème difficile.

Remerciements

Nous remercions le Docteur Marie-Josèphe Deshayes et la société Télécran Innovation pour la base d'images, l'expertise et les connaissances *a priori*.

Références

- [1] M. de Bruijne and M. Nielsen. *Image Segmentation by Shape Particle Filtering*. ICPR, Cambridge, United Kingdom, 2004.
- [2] T.F.Cootes, G.J. Edwards, C.J.Taylor. *Active Appearance Models* IEEE PAMI, Vol. 23 (6), pp. 681-685, 2001.
- [3] B. van Ginneken, A.F. Frangi, J.J. Staal, B.M. ter Haar Romeny, M.A. Viergever. *Active Shape Model segmentation with optimal features*. IEEE Transactions on Medical Imaging, vol. 21, p.924-933, 2002.
- [4] T.J. Hutton, S.Cunningham, P. Hammond. *An Evaluation of Active Shape Models for the Automatic Identification of Cephalometric Landmarks*. European Journal of Orthodontics, Vol. 22(5), pp. 499-508, 2000.
- [5] D.G. Kendall. *Shape manifold, Procrustean metrics and complex projective spaces*. Bull. London Math. Soc., Vol. 16, pp. 81-121, 1984.
- [6] S. Romdhani, A. Psarrou, S. Gong. *Multi-View Nonlinear Active Shape Model using Kernel PCA*. Proceedings of the 10th British Machine Vision Conference, September 1999, Nottingham, England.
- [7] B. Schölkopf, A. Smola. *Learning With Kernels*. Cambridge, MIT Press, 2002.