



Estimation of Landmarks with Occultation Using Statistical Models

Barbara Romaniuk, Michel Desvignes

► To cite this version:

Barbara Romaniuk, Michel Desvignes. Estimation of Landmarks with Occultation Using Statistical Models. Advanced Concepts for Intelligent Vision Systems 2004, 2004, Brussels, Belgium. 8 p. hal-00843482

HAL Id: hal-00843482

<https://hal.science/hal-00843482>

Submitted on 11 Jul 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MISSING DATA ESTIMATION WITH STATISTICAL MODELS

¹Barbara Romaniuk, ²Michel Desvignes

¹Barbara.Romaniuk@greyc.ismra.fr

¹GREYC CNRS UMR-6072, Image Team, 6, Bvd Mal Juin, 14050 Caen Cedex

²LIS-ENSERG, 961 rue de la Houille Blanche, BP 46, 38402 Saint Martin d'Hères Cedex, France

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

1.1. Cephalometry

The goal of cephalometry is the study of the skull growth of young children in order to improve orthodontic therapy. It is based on the landmarking of cephalometric points on tele-radiographs, two-dimensional X-ray images of the sagittal skull projection. These points are used for the computation of features, such as the length or the angles between lines. The interpretation of these features is used to diagnose the deviation of the patient form from an ideal one. It is also used to evaluate the results of different orthodontic treatments.

The cephalometry was initially introduced in the thirties. It consisted in manual landmarking directly on radiographs. The conclusion of these studies was that the location process is subject to a great subjectivity and that the landmarking quality depends on image quality. Three categories of approaches was proposed for landmarking automation: reproduction of expert landmarking; statistical modeling and finally appearance models.

The first landmarking automation method was proposed in 1986 by Lévy-Mandel *et al.* [17]. He used *a-priori* knowledge on anatomical structures lying near the anatomical points to localize them. Contours are identified in a predetermined order based on algorithms introduced to the system by *ad hoc* criteria. Parthasarathy *et al.* [20], Davis *et al.* [7] Ong *et al.* [19] presented similar knowledge based systems. However all these studies were carried out on different small sets of images. Moreover the training was made on the same images that tests. These methods

give good results for well defined and located on strong contours points. However they are subject to failure in images containing artifacts or low quality images and when anatomical structures variability is very high. For any new landmark that the practitioner need to introduce new rules have to be formulated, and then introduced in the system. Another type of approach was proposed simultaneously with our in 2000 by Hutton *et al.* [14]. Authors propose a method based on active contours. This method consists in learning a model composed of the mean observation and the variability authorized around the mean. This model is then positioned statistically on an image and is deformed in order to correspond as well as possible to reality.

The third type of approach appears in 1994 Cardillo and Sid-Ahmed [4] and propose an alternative solution. They develop an algorithm using the mathematical morphology based on grey levels. During training some structuring elements are identified and the probability density of each point position that is relative to the structuring element is evaluated. Recognition of the landmark is given by the intersection of these probability distributions.

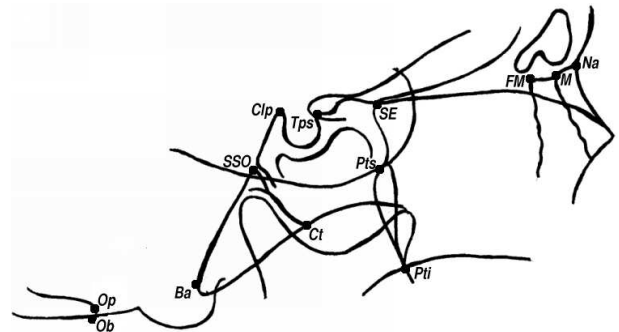


Figure 1: Cephalometric points position define relatively to anatomical structures position.

All these studies have a common point: the images on which these studies were done where all scanned radiographs. Most on them treat images with an insufficient resolution and the number of studied images did not allow to model anatomical variability. Only Hutton *et al.* work

on high resolution images, however they assume their localization is only an estimate. It is also important to notice that all these studies were done on different points especially located in the facial part where the landmarking is easiest than in the central part.

1.2. Shape modeling

The pioneers of shape analysis are Kendall [16] [10] and Bookstein [1]. In these works, shape is defined as the remaining information after alignment (rotation, translation, scaling) between two objects. In image analysis, Pentland [25] has defined modal analysis and a similar idea has been used by Cootes [5] in the Active Shape Model (ASM) and Active Appearance Model (AAM). They both involve a Principal Component Analysis (PCA) to build a statistical shape model. In this model, the mean object and the variation around this mean position are both represented.

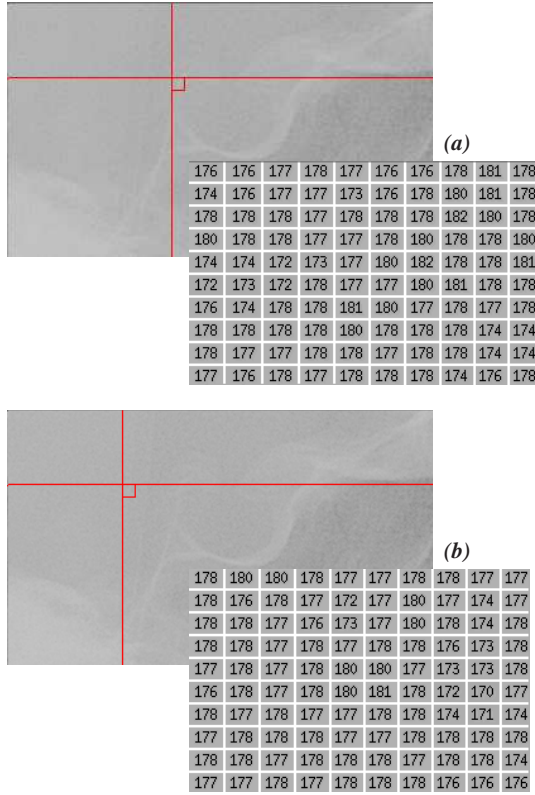


Figure 2: Greylevels in the vicinity of the point Clp . (a) are greylevels in a region centered on Clp . (b) are greylevels in a region lying in the neighborhood of the region (a).

Extension of ASM introduce a model of the underlying distribution of intensity around landmark; it becomes AAM.

AAM use the classical model of shape but adds a model of the profile intensity perpendicular to the object contour. Most of the work are based on the assumption of linear distribution [2] [8] [9] [26].

Some recent work introduces non linear model of gray intensity, based on a k-nearest neighbors classifier and on local selected features similar to texture analysis [12] [13]. The multi resolution approaches is a key feature of in this method. The non linear classification is also found in [3], with a shape particle filtering algorithm to fit the model on the image. Unfortunately, the key assumption under these approaches needs that the landmark is localized on a visible edge or contour. As we can see in figure 2, this is not the case in the cephalometric problem: our landmarks are not delimited by clear and visible contour. This is also the conclusion of Hutton [14]: AAM are not adequate for cephalometric landmarking. Of course, variability modeling with ASM or AAM is an important feature that will be used in our approach. We also retain the idea of non linear modeling, but it will be applied on the shape model itself.

1.3. Our approach

As we've seen previously, in cephalometry the anatomical definition of cephalometric point is difficult to apply on radiographs. The manual landmarking is hard and an important inter and intra expert variability is to be noticed. The manual location has a minimal variability of about 2 millimeters.

To reduce this variability, we will use statistical pattern recognition. We dispose of an expertise made by an expert that consist on landmarked radiographs (cephalometric points). In our approach we also use some *a priori* knowledge that consists in saying that the landmark position is relative to cranial shape¹. The cranial contour is then automatically extracted from the image. Our training set of points is composed of the cephalometric points and the sampled cranial contour. A statistical model is learned on this training set. To retrieve landmarks on a new image, the cranial contour is detected: the problem is now to retrieve the unknown points (cephalometric) of this image using the known points (cranial contour) and the model. Our goal is then to solve the problem of partial occultation of an observation, the partial occultation corresponding to the set of anatomical points and the visible part to the cranial contour.

Some works on Kernel PCA [18] are very close to our method. Briefly, Kernel PCA maps non-linearly the input data in a Feature Space (\mathcal{F} -Space). PCA is then performed in \mathcal{F} . The non linear principal components are given by the eigenvectors corresponding to the largest eigenvalues.

¹Cephalometric points are not necessarily localized on the cranial contour. The goal here is to model the unknown relationship existing between the cranial system and the studied anatomical points.

In a localization problem, mean shape in the F -space must be back-projected in the input space. The choice of the mapping and the back projection are difficult problems and are still open issues.

In this paper, we address the problem of retrieving the position of non visible parts of a partially visible object. The key idea is to use the authorized variation around the mean shape of the model to localize these parts. We present a comparison of three non-linear methods based on KPCA representation. They are applied to the cephalometric problem.

In this paper we first present kernel principal component analysis, then our method which goal is to estimate an observation in the input space supposing a part of it subject to an occultation. We propose a method using a explicit mapping function and then a second approach using the kernel trick. We finally explain our experiments, results and conclude.

2. KERNEL PRINCIPAL COMPONENT ANALYSIS

Kernel PCA consists in data mapping in a high dimensional feature space \mathcal{F} and then performs PCA in \mathcal{F} . The mapping function is non linear. This method extracts non linear characteristics from the dataset.

2.1. Mercer Kernels

A Mercer kernel is a function $k(X_i, X_j)_{(i,j) \in \{1, \dots, n\}^2}$, where X_i is an observation. This function is used to compute the kernel matrix K from all the observations. This matrix is a positive definite matrix. The use of this function can be used to rewrite data mapping in terms of a dot product. Let φ be the non linear mapping function. We can then write:

$$k(X_i, X_j) = \langle \varphi(X_i), \varphi(X_j) \rangle. \quad (1)$$

The three most popular kernels are:

- polynomial: $k(X, Y) = (\langle X, Y \rangle + c)^d$.
- gaussian: $k(X, Y) = \exp \frac{-\|X-Y\|}{2\sigma^2}$.
- sigmoid: $k(X, Y) = \tanh(\kappa \langle X, Y \rangle + \theta)$.

For most of kernels the non linear mapping functions is implicit.

2.2. PCA in the feature space \mathcal{F}

Let n be the number of observations in our dataset. Let $X_i \ i \in \{1, \dots, n\}$ be the observations. Let φ be the non linear mapping function such as:

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

We suppose that our mapped data is centered in F . Let C be the covariance matrix computed on that mapped data:

$$C = \frac{1}{n} \sum_{i=1}^n \varphi(X_i) \varphi(X_i)^t. \quad (2)$$

Let λ^k be the eigenvalues and V_k be the eigenvectors of the matrix C :

$$\lambda_k V^k = C V^k. \quad (3)$$

There exist coefficients α^k for $k \in \{1, \dots, n\}$ such that:

$$V^k = \sum_{i=1}^n \alpha_i^k \varphi(X_i). \quad (4)$$

We then obtain:

$$\begin{aligned} \lambda_k \sum_{i=1}^n \alpha_i^k \langle \varphi(X_k), \varphi(X_i) \rangle &= \\ \frac{1}{n} \sum_{i=1}^n \alpha_i^k \left\langle \varphi(X_k), \sum_{j=1}^n \varphi(X_j) \langle \varphi(X_j), \varphi(X_i) \rangle \right\rangle. \end{aligned} \quad (5)$$

We can then define the kernel matrix K :

$$K_{ij} = \langle \varphi(X_i), \varphi(X_j) \rangle.$$

Equation (5) becomes:

$$n \lambda_k K \alpha^k = K^2 \alpha^k, \quad (6)$$

that is equivalent to:

$$n \lambda_k \alpha^k = K \alpha^k. \quad (7)$$

As K is a positive definite matrix we know that its eigenvalues are positive. $n \lambda^k$ are the solutions of the equation (7).

We then preserve only the l eigenvalues that are not zero. For all $k \in \{1, \dots, l\}$ we have $\langle V^k, V^k \rangle = 1$. We normalize α^k respectively to their eigenvectors in F :

$$\begin{aligned} \langle V^k, V^k \rangle &= 1 \\ &= \sum_{i,j=1}^n \alpha_i^k \alpha_j^k \langle \varphi(X_i), \varphi(X_j) \rangle \\ &= \sum_{i,j=1}^n \alpha_i^k \alpha_j^k K_{ij} \\ &= \langle \alpha^k, K \alpha^k \rangle \\ &= \lambda_k \langle \alpha^k, \alpha^k \rangle \end{aligned}$$

It is then possible to compute the map of an observation on the k^{st} principal component:

$$\begin{aligned}\beta_k &= \langle V^k, \varphi(X) \rangle \\ &= \sum_{i=1}^n \alpha_i^k \langle \varphi(X_i), \varphi(X) \rangle \\ &= \sum_{i=1}^n \alpha_i^k k(X_i, X).\end{aligned}$$

The map associated to φ is:

$$P_l \varphi(X) = \sum_{i=1}^l \beta_i V^i. \quad (8)$$

2.3. Kernel PCA algorithm

Here we present the algorithm proposed by Schölkopf, it allows instead of computing the covariance matrix of the data to compute the Kernel matrix. The algorithm can be presented in three steps :

- 1. Compute the Kernel matrix: $K_{ij} = \langle \varphi(X_i), \varphi(X_j) \rangle$.
- 2. Compute the eigenvectors and eigenvalues of K . Normalize them: $\lambda_k \langle \alpha^k, \alpha^k \rangle = 1$.
- 3. Compute projections on the principal components: $\beta_k = \sum_{i=1}^n \alpha_i^k k(X_i, X)$.

Schölkopf algorithm stops here. To solve the input space mapping problem we need to define a function that we will minimize relatively to the function φ . This function is:

$$P_l \varphi(X) = \sum_{i=1}^l \beta_i V^i. \quad (9)$$

2.4. Centering

Previously we made the assumption that the observations were centered:

$$\sum_{i=1}^n \varphi(X_i) = 0. \quad (10)$$

This assumption is important for the computation of the matrix K . When observations are not centered the relation (7) is no more satisfied.

Observations centering is easy to achieve in the input space, but more difficult in the feature space \mathcal{F} , as we cannot explicitly compute the mean of the mapped observations in the space. There is a way to do it, computing Gram matrix

instead of the kernel matrix². The matrix to diagonalize is then:

$$G = K - 1_n K - K 1_n + 1_n K 1_n, \quad (11)$$

where $1_{nij} = \frac{1}{n}$ for all $\{i, j\} \in \{1, \dots, n\}^2$.

3. OBSERVATIONS ESTIMATION IN THE INPUT SPACE

Applied to the cephalometric problem, X is partially known: it is composed by the sampled cranial contour which is automatically extracted and well defined and the unknown cephalometric points. Then, the problem that we want to solve in this section is the reconstruction of partially unknown examples from the KPCA model and from the known data (sampled cranial contour). Let X be an example to reconstruct, with the first coordinates being known. We can see the statistical model as some variability parameters around a mean shape. Finding the unknown part of X is equivalent to find the shape belonging to the model (*i.e.* variability parameters) whose first coordinates are given by the known part of X .



Figure 3: Detection of the reference shape in the cephalometric problem : known part of the model

3.1. Explicit mapping function for cephalometry

3.1.1. Mapping Function

The first step consists in the detection of the reference shape. This shape is then sampled in p equidistant points $P_i, i \in \{1, \dots, p\}$. The non-linear feature space is defined by ratio of surfaces of triangles obtained from the previous

²If observations are centered in the feature space \mathcal{F} the kernel matrix is the Gram matrix.

sampling. The coordinates of an image point $M(x, y)$ are defined by β, γ et δ computed for each possible triangle:

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

where $\overline{P_i P_j P_k}$ is the algebraic area of the triangle $P_i P_j P_k$. This coordinates satisfy:

$$\beta \times \overrightarrow{MP_i} + \gamma \times \overrightarrow{MP_j} + \delta \times \overrightarrow{MP_k} = \vec{0}$$

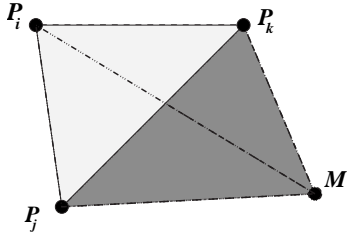


Figure 4: New coordinates of a point M

Let n be the number of triangles obtained from the set of points P_i . New coordinates of a point M are:

$$X' = [\beta_1 \ \gamma_1 \ \delta_1 \ \dots \ \beta_n \ \gamma_n \ \delta_n]^t = A' X,$$

where A' is the matrix used to project the data from the Cartesian to our new feature space.

3.1.2. Variability and Relationships Modeling.

Learning is done on a basis composed of N expertised images. For each image we detect the reference shape, and we sample it. For each image i , we have: a set of points $\{P_k^i\}_{k \in \{1, \dots, p\}}$, a matrix A'^i and the set of q coordinates of the characteristic points $\{X_j^i\}$. We compute the mean position of each characteristic point. Let ϑ^i be the vector representing a characteristic point of the image i in the new space. The mean position of this point is:

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

The variance $\hat{\sigma}$ of vectors ϑ^i is also computed. We deduce from this the weighting matrix P :

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

When the sampling of the reference shape gives an important number of points, only some of them are important. We then propose to apply a PCA on the covariance matrix of vectors ϑ^i . Only the most important d' components are retained. These components are those the eigenvalues of which are the highest, and they form the matrix Φ .

3.1.3. Points estimation

Let ϑ be the vector representing the characteristic point X in the new coordinate space. Landmarking X on a new image consists in resolving the system: $\hat{\vartheta} = A' X$, where $\hat{\vartheta}$ is the average learned vector, A' the matrix defined relatively to characteristics of the new image. We solve this problem using weighted least squares. The estimated position \tilde{X} of the characteristic point X on an unknown image is given by the equation:

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

3.2. Implicit mapping and pseudo-inverse using Kernel Trick

We are interested in solving the problem which goal is the estimation of an observation in the input space, using a model learned in the feature space \mathcal{F} . We suppose a part of the observation known. To solve this problem we use spatial relationships existing between the known part of the observation and the unknown one. Those relationships are also learned in our model.

There are two possible approaches to solve this problem. The first one use an explicit mapping function φ , the second one use Kernel PCA making φ implicit. In the first case estimation consists in computing the inverse of φ , in the second case the problem is much more complicate. Our model is trained on n observation, each composed of m characteristics. Our goal is to identify p characteristics of an observation, where $p < \frac{m}{2}$.

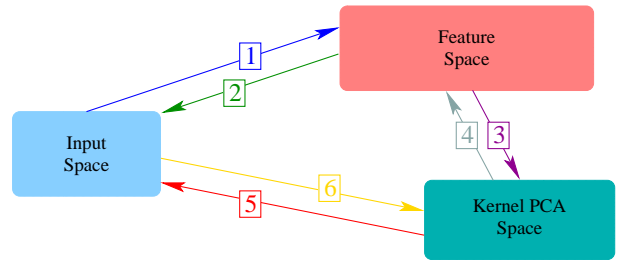


Figure 5: Three different observation spaces.

In this part we work in three different spaces (figure 5). The first one is the input space that is m dimensional (number of observation characteristics), the second one is the feature space \mathcal{F} of dimension $L > m$, and finally the third one is the Kernel PCA space which dimension is n . This type of scheme was already proposed in literature by Romdhani and al. [22]. We propose here to add one supplementary item: mapping from the Kernel PCA space into the input space (step 5, figure 5).

Let X be an observation in the input space, X_i the i^{st} observation from the training set. Let φ be the non linear

mapping function and β the coordinates of an observation in the Kernel PCA space. Let Z_{kpca} be the pre-image, that is the reconstructed observation X in the input space starting from the Kernel PCA space.

3.2.1. Pseudo-inverse using β

Using an observation X_{kpca} in the Kernel PCA space we can compute its image Z_{kpca} by minimizing:

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

with β the projections of X_{KPCA} on the first l principal components: $\beta = (\beta_1, \dots, \beta_l)$, $\beta_k = \sum_{i=1}^n \alpha_i^k k(X_i, X_{kpca})$. By developing this equation we obtain the pseudo-inverse using β :

$$\left\| \sum_{i=1}^n \alpha_i k(X_i, Z_{kpca}) - \beta \right\|^2. \quad (12)$$

3.2.2. Pseudo-inverse using Z_{kpca}

We can also rewrite the previous formula trying to express $k(X_i, Z_{kpca})$ relatively to β .

We know that:

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

We can rewrite it using matrix formalism:

$$\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} \alpha_1^1 & \dots & \alpha_n^1 \\ \vdots & \ddots & \vdots \\ \alpha_1^n & \dots & \alpha_n^n \end{bmatrix} \begin{bmatrix} k(X_1, Z_{kpca}) \\ \vdots \\ k(X_n, Z_{kpca}) \end{bmatrix}. \quad (13)$$

Let B be the matrix:

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

α^i are the pseudo-eigenvectors of the kernel matrix computed relatively to its eigenvectors V^i and eigenvalues λ^i :

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

We know that V^i form an orthonormal basis, i.e. $\langle V^i, V^j \rangle = 0$ for $i \neq j$ and $\langle V^i, V^j \rangle = 1$ for $i = j$. We then obtain:

$$\langle \alpha^i, \alpha^i \rangle = \frac{1}{\lambda^i}.$$

and :

$$B = \begin{bmatrix} \frac{1}{\lambda^1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\lambda^n} \end{bmatrix}.$$

We deduce B^{-1} :

$$B^{-1} = \begin{bmatrix} \lambda^1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda^n \end{bmatrix}.$$

We can then compute α^{-1} :

$$\begin{aligned} \alpha^{-1} \alpha &= Id \\ \alpha^{-1} \alpha \alpha^t &= \alpha^t \\ \alpha^{-1} B &= \alpha^t \\ \alpha^{-1} &= \alpha^t B^{-1}. \end{aligned}$$

Equation (13) and α^{-1} value leads to:

$$\begin{bmatrix} k(X_1, Z_{kpca}) \\ \vdots \\ k(X_n, Z_{kpca}) \end{bmatrix} = \begin{bmatrix} \alpha_1^1 & \dots & \alpha_n^1 \\ \vdots & \ddots & \vdots \\ \alpha_1^n & \dots & \alpha_n^n \end{bmatrix} \begin{bmatrix} \lambda^1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda^n \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}. \quad (14)$$

The pseudo-inverse relative to Z_{kpca} can finally be defined:

$$\left\| k(X_j, Z_{kpca}) - \sum_{i=1}^n \alpha_i^j \lambda^j \beta^j \right\|^2. \quad (15)$$

4. EXPERIMENTS

To test the feasibility of the algorithm, we have run several simulated and real world experiments. We have compared the distance between the computed points and the real points, using the different methods, with the gaussian kernel for Kernel PCA based method. The minimization in this case was done with Powell algorithm [11].

On the cephalometric problem, the cranial contour is approximated by 6 points, and we have tested the methods on the reconstruction of 14 cephalometric points. Models are built with 80 radiographs. We use a leave-one out approach to test the accuracy of the models.

Table 1 presents the results obtained with the explicit mapping function that makes data invariant to affine transformations, table 2 presents results obtained with the pseudo-inverse using β and the Kernel trick. Obtained results shows that the first method is better and near the inter-expert variability that is about 2 millimeters. The second method using kernels is less accurate, but the model is not invariant to affine transformations.

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
Mean	2.9	2.0	2.93	1.78

Table 1: Mean error and standard deviation (in millimeters) associated to the method with an explicit mapping function.

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
Mean	5.2	4.0	3.66	3.41

Table 2: Mean error and standard deviation (in millimeters) associated to the pseudo-inverse using β . Computation made with a gaussian kernel with $\sigma = 0.005$.

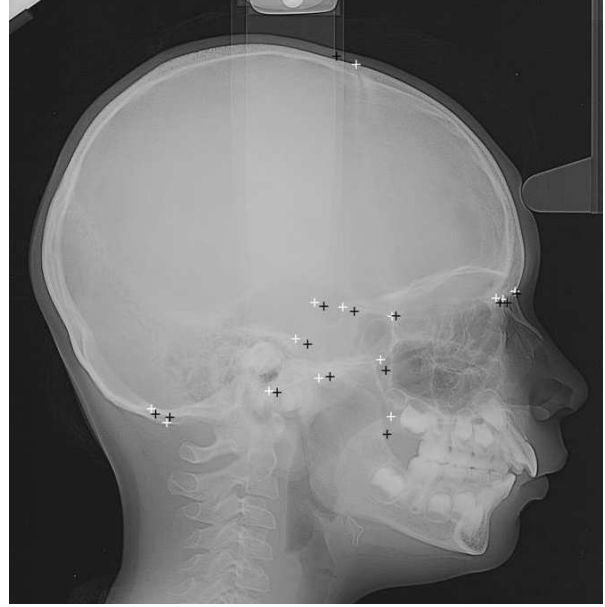


Figure 6: Cephalometric points estimation. In white expert position, in black our estimation.

5. CONCLUSION

In this paper, we have presented and compared two methods to reconstruct non visible parts of an object with a statistical model. The statistical framework offers an elegant way to solve this problem, using the variability authorized by the model. It seems that Kernel PCA models are interesting to solve this problem, even if an explicit mapping in the feature space \mathcal{F} gives better results, but is very hard to define.

6. ACKNOWLEDGMENTS

We are grateful to Dr Marie-Josèphe Deshayes and Telecrane Innovation society for providing the test image database, the expertise and all the *a priori* knowledge.

7. REFERENCES

- [1] F. L. Bookstein. *Principal Warps: Thin-Plate Splines and the Decomposition of Deformations*. IEEE PAMI, Vol. 11 (6), pp. 567-585, 1989.
- [2] M. Brejl and M. Sonka. *Automated Initialization and Automated Design of Border Detection Criteria in Edge-Based Image Segmentation*. 4th IEEE Southwest Symposium on Image Analysis and Interpretation, Austin, TX, USA, pp. 26-30, 2000.

- [3] M. de Bruijne and M. Nielsen. *Image Segmentation by Shape Particle Filtering*. ICPR, Cambridge, United Kingdom, 2004.
- [4] J. Cardillo, M.A. Sid-Ahmed. *An Image Processing System for Locating Craniofacial Landmarks*. IEEE Transactions on Medical Imaging, Vol. 13(2), pp. 275-289, 1994.
- [5] T.F. Cootes, G.J. Edwards, C.J. Taylor. *Active Appearance Models* IEEE PAMI, Vol. 23 (6), pp. 681-685, 2001.
- [6] T.F. Cootes, C.J. Taylor. *Statistical Models of Appearance for Computer Vision*. Internal report, 1999.
- [7] D.N. Davis, D. Forsyth. *Knowledge-based cephalometric analysis: A comparison with clinicians using interactive methods*. Computers and Biomedical Research 27, 1994.
- [8] N. Duta and M. Sonka. *An Improved Active Shape Model: Handling Occlusion and Outliers*. Proceedings of ICIAP '97 Conference, Lecture Notes in Computer Science, vol. 1310, pages 398-405, Springer-Verlag, 1997.
- [9] N. Duta and M. Sonka. *Segmentation and Interpretation of MR Brain Images: An Improved Active Shape Model*. IEEE Transactions on Medical Imaging, Vol. 17, No. 6, pp. 1049-1067.
- [10] I.L. Dryden. *General Shape and Registration Analysis*. In W. S. Kendall, O. Barndorff-Nielsen, and M. N. M. van Lieshout (Eds.), SEMSTAT 3, London. Chapman and Hall, 1997.
- [11] R. Fletcher, M.J.D. Powell. *A Rapidly Convergent Descent Method for Minimization*. Computer Journal, Vol. CJ-6, no. 2, pp. 163-168, 1963.
- [12] 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.
- [13] B. Van Ginneken, A.F. Frangi, J.J. Staal, B.M. ter Haar Romeny, M.A. Viergever. *A Non-Linear Gray-Level Appearance Model Improves Active Shape Model Segmentation*. IEEE workshop on Mathematical Models in Biomedical Image Analysis, Editor(s): L. Staib, A. Rangarajan, IEEE Soc. Press, p.205-212, 2001.
- [14] 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.
- [15] A. Jain, P. Duin, J. Mao. *Statistical pattern recognition: A review*. IEEE Transactions on PAMI, Vol. 22(1), pp. 4-37, 2000.
- [16] D.G. Kendall. *Shape manifold, Procrustean metrics and complex projective spaces*. Bull. London Math. Soc., Vol. 16, pp. 81-121, 1984.
- [17] A. Levy-Mandel, A.N. Venetsanopoulos, J.K. Tsotsos. *A Methodology for Knowledge-Based Landmarking of Cephalograms*. Computers and Biomedical Research, Vol. 19, pp. 282 - 309, 1986.
- [18] S. Mika, B. Schölkopf, A. Smola, K.R. Müller, M. Scholz, G. Ratsch. *Kernel PCA and de-noising in feature spaces*. Advances in Neural Information Processing Systems, Vol. 11, pp. 536-542, 1999.
- [19] S.H. Ong, X. Han, S. Ranganath, A.A. Kassim, K.F. Lim, K.W.C. Foong. *Identification of feature points in cephalograms*. ICARCV'98, Singapore, pp. 1056-1060, December 1998.
- [20] S. Parthasarathy, S.T. Nugent, P.H. Gregson, D.F. Fay. *Automatic landmarking of cephalograms*. Computers and Biomedical Research, Vol. 22, pp. 248-269, 1989.
- [21] B. Romaniuk, M. Desvignes, M. Revenu, M.J. Deshayes. *Linear and Non-Linear Model for Statistical Localization of Landmarks*. ICPR, Vol. 4, pp. 393-396, Québec, Canada, August 2002.
- [22] S. Romdhani, A. Psarrou, S. Gong. *Multi-View Non-linear Active Shape Model using Kernel PCA*. Proceedings of the 10th British Machine Vision Conference, September 1999, Nottingham, England.
- [23] B. Schölkopf, A. Smola, K.R. Müller. *Nonlinear Component Analysis as a Kernel Eigenvalue Problem*. Neural Computation, vol. 10, pp. 1299-1319, 1998.
- [24] B. Schölkopf, A. Smola. *Learning With Kernels*. Cambridge, MIT Press, 2002.
- [25] S. Sclaroff, A.P. Pentland. *Modal Matching for Correspondence and Recognition*. IEEE Transactions on PAMI, Vol. 17(6), pp. 545-561, 1995.
- [26] S. Sclaroff, J. Isirodo. *Active blobs: region-based deformable appearance models*. Computer Vision and Image understanding, vol. 89(2-3), pp. 197-225, 2003.