# Statistical Shape Model of Variability and Spatial Relationships

Barbara Romaniuk, Michel Desvignes, Marinette Revenu, Marie-Josèphe Deshayes

# Statistical Shape Model of Variability and Spatial Relationships

B. Romaniuk*, M. Desvignes**, M. Revenu*, MJ. Deshayes ***

** GREYC-CNRS 6072 (équipe Image)*
*ENSICAEN, Bd Maréchal Juin*
*14050 Caen Cedex, France*
*{Barbara.Romaniuk, Marinette.Revenu}@greyc.ismra.fr*

*** LIS-ENSERG*
*961, rue de la Houille Blanche*
*Domaine Universitaire - BP 46*
*38 402 Saint Martin d'Hères Cedex*
*michel.desvignes@lis.inpg.fr*

**** Société TCI*
*http://www.cranexplo.com*
*tci@cranexplo.net*

## Abstract

***In this paper we model shape objects variability and spatial relationships between characteristics of this object. What is an object? We consider that an object is a set of shapes and points. We establish here a formalism unifying shapes and points by approximating shape contours by composite Bézier curves. These are equivalent to their control points. We then propose a non linear invariant statistical model and we learn the average object, it's variability and relationships. We finally evaluate our methodology in cephalometry, by modeling anatomical structures and points.***

## 1 Introduction

In this paper we are interested in modeling the variability of points and shapes for pattern recognition. Our first goal is to find a formalism that allows us to consider simultaneously points and shapes. We propose to approximate shapes by composite Bézier curves, that are equivalent to their control points. Our second goal is to build a non linear model, invariant up to affine transformations, including the average object (points and shapes), variability allowed around this object, and finally spatial relationships between all the characteristics of the studied object. Evaluation of our methodology is done in cephalometry, that consists in landmarking anatomical points on radiographs. Those points are defined relatively to anatomical structures, that we modeled with composite Bézier curves. We then learned simultaneously anatomical points and relative structures position.

## 2 Shape representation

We are interested in finding an optimal representation for shapes. Shape learning requires a compact and efficient computer oriented representation, allowing a correct modeling of a population of shapes. This representation must also translate the non-rigidity of studied shapes. The appearance of an object can change when its geometric properties change. Most of the authors treating the problem of recognition of a set of shape characteristic points propose to represent the outline of the shape by a set of points [1][3][10]. The simplest idea is to sample the contour of the shape. The problem is that sampling is not equivalent to the shape.

### 2.1 Bézier curves

We propose in this paper to chose Bézier curves [7] to approximate shape contours. This proposal is due to the properties of these curves. Three major aspects can be taken to the fore:

- Equivalence between control points of the curve and the curve;
- Analytic definition of the curve and its derivatives;

- Invariance of the curve up to affine transformations.

This formalism allows us to unify two different notions: the point notion and shape notion.

Let $(P_k)_{k\in\{0,\ldots,n\}}$ be an ordered set of point in 2D. These points, called control points, define a polygonal line $\Pi$. We can coligate to this line a parametric curve (Bézier curve) defined by the equation:

$$\forall u \in [0,1] \qquad P(u) = \sum_{k=0}^{n} P_k B_{k,n}(u), \tag{1}$$

$B_{k,n}(u)$ are the Bernstein polynomials and $n$ is the degree of the curve.

There are two different ways to approximate a set of ordered points by a Bézier curve: the first one is interpolation, the second smoothing. Both of them imply parameterization of the curve. This step consists in finding a good repartition of the parameter $u$ (*cf.* equation (1)). Two different methods exist. The first type of parameterization gives a uniform repartition of the parameter, the second one gives a repartition proportional to the distance between the points that we want to approximate.

### 2.1.1 Interpolation

Let $(Q_i)_{i\in\{0,\ldots,r\}}$ be a set of points that we want to interpolate. The problem is to find a curve going exactly thought this set of points. This problem can be formulated by the system of equations:

$$Q_i = \sum_{k=0}^{n} P_k B_{k,n}(u_i), \quad i \in \{0,\ldots,r\},\ u_i \in [0,1]. \tag{2}$$

We want to find here the position of the points $P_k$. At first we must define the degree $n$ of the curve. For this method the degree is fixed: $n = r$. We must then solve the system of linear equations:

$$\begin{bmatrix} P_0 \\ \vdots \\ P_n \end{bmatrix} = \begin{bmatrix} B_{0,n}(u_0) & \ldots & B_{n,n}(u_0) \\ \vdots & \ddots & \vdots \\ B_{0,n}(u_r) & \ldots & B_{n,n}(u_r) \end{bmatrix}^{-1} \begin{bmatrix} Q_0 \\ \vdots \\ Q_r \end{bmatrix}. \tag{3}$$

This method is interesting when the degree $n$ is small. A high degree will cause spurious oscillations.

### 2.1.2 Smoothing

Let $\left(Q'_j\right)_{j\in\{0,\ldots,s\}}$ be a set of points that we want to smooth. The problem in this case is to find a curve $P(u')$ in the nearest neighborhood of the set of point $\left(Q'_j\right)$. The classical method consists in minimizing distances between points $Q'_j$ and the curve $P(u')$, defined as:

$$\varepsilon_j = P\left(u'_j\right) - Q'_j, \quad j \in \{0,\ldots,s\}. \tag{4}$$

Control points $P_k$ can be computed by solving the system:

$$\begin{bmatrix} P_0 \\ \vdots \\ P_n \end{bmatrix} = \left([B_{n,s}]^T [B_{n,s}]\right)^{-1} [B_{n,s}]^T \begin{bmatrix} Q'_0 \\ \vdots \\ Q'_s \end{bmatrix}, \tag{5}$$

where:

$$[B_{n,s}] = \begin{bmatrix} B_{0,n}(u'_0) & \ldots & B_{n,n}(u'_s) \\ \vdots & \ddots & \vdots \\ B_{0,n}(u'_s) & \ldots & B_{n,n}(u'_s) \end{bmatrix}.$$

The choice of the degree must verify the constraint: $n < s$. If $n = s$ and all distances $\varepsilon_j$ are reduced to zero, the solution is the same as the one obtained by interpolation.

#### 2.1.3 Interpolation and Smoothing

It is also possible to mix both of the methods. Let $(Q_i)_{i\in\{0,\dots,r\}}$ be passage constraints, and $\left(Q'_j\right)_{j\in\{0,\dots,s\}}$ points to smooth. Control points $P_k$ procure $(s+1)$ equations for smoothing. The degree of the curve is defined as: $n = r+s+1$. The equation to solve is:

$$\begin{bmatrix} P_0 \\ \vdots \\ P_n \end{bmatrix} = \begin{bmatrix} A \\ CD \end{bmatrix}^{-1} \begin{bmatrix} Id & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} Q \\ Q' \end{bmatrix} \tag{6}$$

where $Id$ is a $(r+1)\times(r+1)$ identity matrix, and $A$, $C$ and $D$ defined as:

$$A = \begin{bmatrix} B_{0,n}(u_0) & \dots & B_{n,n}(u_0) \\ \vdots & \ddots & \vdots \\ B_{0,n}(u_r) & \dots & B_{n,n}(u_r) \end{bmatrix}, C = \begin{bmatrix} B_{0,n}(u'_0) & \dots & B_{0,n}(u'_s) \\ \vdots & \ddots & \vdots \\ B_{s,n}(u'_0) & \dots & B_{s,n}(u'_s) \end{bmatrix} \text{ and } D = \begin{bmatrix} B_{0,n}(u'_0) & \dots & B_{n,n}(u'_0) \\ \vdots & \ddots & \vdots \\ B_{0,n}(u'_s) & \dots & B_{n,n}(u'_s) \end{bmatrix}.$$

## 2.2 Composite Bézier curves: constraints used to join two curves $P^{(a)}\left(u^{(a)}\right)$ and $P^{(b)}\left(u^{(b)}\right)$

Approximating a shape by a Bézier curve implies a compromise between fidelity ($n$ hight) and robustness ($\varepsilon_j$ height). It is possible to solve this problem using composite Bézier curve. It consists in separating the set of points defining the shape to approximate several subsets, each of them defining one Bézier curve. All of them will have to verify joining constraints. Let $n^{(a)}$ and $n^{(b)}$ be respective degrees of curves $P^{(a)}\left(u^{(a)}\right)$ and $P^{(b)}\left(u^{(b)}\right)$. Associate sets of control points are $\left\{P^{(a)}_{k^{(a)}}\right\}$ and $\left\{P^{(b)}_{k^{(b)}}\right\}$. These two polynoms:

1. are continuous when their extremities coninside (last point of $P^{(a)}\left(u^{(a)}\right)$ is also the first of $P^{(b)}\left(u^{(b)}\right)$):

$$P^{(a)}(1) = P^{(b)}(0) \iff P^{(a)}_{n^{(a)}} = P^{(b)}_0, \tag{7}$$

2. have continuous derivatives at the junction points. This implies the continuity of the vector tangent at those points. We have the proportionality between first derivatives. The relation between $P(u_1)$ and $P'(u_2)$ can be written as:

$$n^{(a)}\left(P^{(a)}_{n^{(a)}-1} - P_{n^{(a)}}\right) = \alpha n^{(b)}\left(P^{(b)}_1 - P^{(b)}_0\right) \tag{8}$$

if the connected extremities are $P^{(a)}(1)$ and $P^{(b)}(0)$.

3. have a constant curvature radius at the junction. This implies the continuity of the second derivative of the curves. For the continuity of the osculator plane must be verified at the junction:

$$\left(P^{(a)}_{n^{(a)}-2} - P^{(a)}_{n^{(a)}-1}\right) \wedge \left(P^{(a)}_{n^{(a)}-1} - P^{(a)}_{n^{(a)}}\right) = \gamma\left[\left(P^{(b)}_2 - P^{(b)}_1\right) \wedge \left(P^{(b)}_1 - P^{(b)}_0\right)\right], \ \gamma \in \mathbb{R}^*. \tag{9}$$

The equality of the curvature radius is assumed by the relation:

$$\frac{n^{(a)}\left\|P^{(a)}_{n^{(a)}-1} - P^{(a)}_{n^{(a)}}\right\|^3}{\left(n^{(a)}-1\right)\left\|\left(P^{(a)}_{n^{(a)}-2} - P^{(a)}_{n^{(a)}-1}\right) \wedge \left(P^{(a)}_{n^{(a)}-1} - P^{(a)}_{n^{(a)}}\right)\right\|} = \frac{n^{(b)}\left\|P^{(b)}_1 - P^{(b)}_0\right\|^3}{\left(n^{(b)}-1\right)\left\|\left(P^{(b)}_2 - P^{(b)}_1\right) \wedge \left(P^{(b)}_1 - P^{(b)}_0\right)\right\|}. \tag{10}$$

Parameters $\alpha$ and $\gamma$ define the shape of each curve. Different situations are presented in figure 1. Parameters $\alpha$ and $\gamma$ are generally fixed to 1 or $-1$.

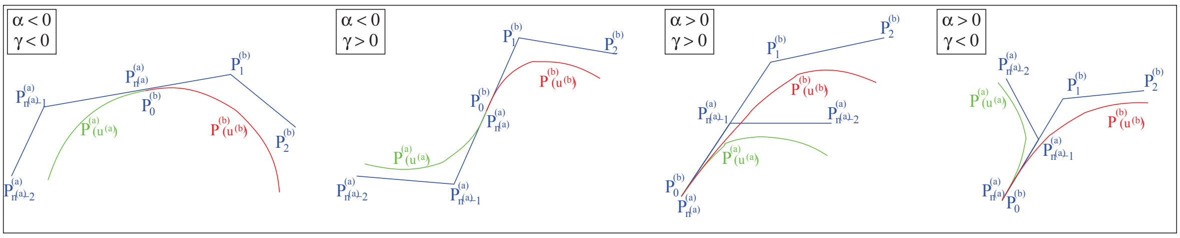



Figure 1: Influence of parameters $\alpha$ and $\gamma$ on the junction appearance between two curves

### 2.3 Optimization and generalization of the representation

We formulate here the problem of joining $N$ composite Bézier curves and determining all control points. We consider here the case $\alpha = -1$ and $\gamma = -1$. The junction problem of $N$ curves with this same criterion implies the resolution of the system:

$$\begin{bmatrix} \mathcal{A}^{(1)} & \mathcal{B} & & \cdots & 0 \\ \mathcal{C} & \mathcal{A}^{(2)} & \mathcal{B} & & \\ \vdots & & & \ddots & \vdots \\ 0 & & \cdots & \mathcal{C} & \mathcal{A}^{(N)} \end{bmatrix} \begin{bmatrix} \mathcal{P}^{(1)} \\ \mathcal{P}^{(2)} \\ \vdots \\ \mathcal{P}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathcal{D}^{(1)} & & \cdots & 0 \\ & \mathcal{D}^{(2)} & & \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & \mathcal{D}^{(N)} \end{bmatrix} \begin{bmatrix} \mathcal{Q}^{(1)} \\ \mathcal{Q}^{(2)} \\ \vdots \\ \mathcal{Q}^{(N)} \end{bmatrix}, \tag{11}$$

with:

$$\mathcal{A}^{(1)} = \begin{bmatrix} 1 & 0 & \cdots & & 0 \\ & & C^{(1)}D^{(1)} & & \\ 0 & & \cdots & 0 & 1 \\ 0 & & \cdots & 1 & -2 \end{bmatrix}, \mathcal{A}^{(N)} = \begin{bmatrix} 0 & 0 & 1 & \cdots & 0 \\ & & C^{(N)}D^{(N)} & & \\ 0 & & \cdots & 0 & 1 \end{bmatrix}, \mathcal{A}^{(i)} = \begin{bmatrix} 0 & 0 & 1 & \cdots & 0 \\ & & C^{(i)}D^{(i)} & & \\ 0 & & \cdots & 0 & 1 \\ 0 & & \cdots & 1 & -2 \end{bmatrix},$$

$$\mathcal{B}^{(i)} = \begin{bmatrix} 0 & & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \end{bmatrix}, \mathcal{C} = \begin{bmatrix} 0 & \cdots & -1 & 2 & -2 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 \end{bmatrix}, \mathcal{P}^{(i)} = \begin{bmatrix} P_j^{(i)} \end{bmatrix}, \mathcal{Q}^{(i)} = \begin{bmatrix} \begin{bmatrix} Q_{k(1)}^{(i)} \end{bmatrix} \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\mathcal{D}^{(1)} = \left[\begin{array}{ccc|ccc} 1 & 0 & \cdots & & & \\ & & C^{(1)} & 0 & 0 & 0 \\ \hline & & 0 & & & \\ & & 0 & & 0 & \\ & & 0 & & & \end{array}\right], \mathcal{D}^{(N)} = \begin{bmatrix} & C^{(N)} & & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \text{ and } \mathcal{D}^{(i)} = \left[\begin{array}{c|ccc} C^{(i)} & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & 0 & \\ 0 & & & \end{array}\right].$$

Only two control points belong to the shape. We model the passage of the curve by the points $Q_0^{(0)}$ and $Q_{p^{(N)}}^{(N)}$, where $p^{(N)}$ is the number of points that the curve $P^{(N)}(u^{(N)})$ has to approximate. We can then conclude that $s^{(i)} = n^{(i)} - a$, where $a = 1$ when $i \in \{1, n^{(i)}\}$ and $a = 2$ in other cases.

## 3 Modeling variability and spatial relationships

Once we have defined how to represent a shape by a set of points, we can now see how to model shape variability. Our model will enclose three different aspects. The first one is the average shape, the second one the variability allowed around this average shape, and finally the last one correspond to the existing relationships between all features of the shape.

### 3.1 Existing methods

The shape variability definition that we used is the one proposed by [4]: shape variability relative to a model is the shape difference that we obtain after alignment of this shape on the model. The alignment consists in eliminating similarity group transformations.

We then have to define shape representation, the alignment and invariance, and finally variability. Previous section in the paper treats on the representation on the shape using composite Bézier curves. We represent a shape by a finite set of points that we call characteristic points of the shape. The alignment can be assumed by the analysis of Procrustes [1][10]. Data is then projected in the shape space, that is invariant to similarity group transformations. This space called manifold of Proscustes is not Euclidean. When shape variability is not important, it's more interesting to work in the tangent space to the shape space. Euclidean properties of the tangent space are often statistically more appropriate than non-Euclidean ones of the shape space[6]. Methods used for features extraction consist in finding an $n'$ dimensional subspace of the original $n$ dimensional space ($n' \leq n$). We can use for this two types of methods: linear methods or non-linear methods.

The most popular linear method is the Principal Components Analysis (PCA). PCA gives principal axis of clouds of dots defined by the data. These principal axis approximate all the points issued from the learning set using the defined model. PCA can be resumed by three steps: compute centered data; compute their covariance matrix $S$; compute the eigenvectors $\phi_i$ and the eigenvalues $\lambda_i$ of the matrix $S$, where $\lambda_i \geq \lambda_{i+1}$. Let $\Phi$ be the matrix composed of $t$ eigenvectors $\phi_i$. A shape $X$ from the training set can be approximate by $X = \bar{X} + \Phi b$, where $b$ is the vector of variability parameters of shape.

There are many non-linear methods for feature extraction. Kernel PCA [2][9] is issued from PCA. The main idea of this method is to transfer data in a new feature space $F$ using a non linear function $\varphi$. In this new space is then applied a PCA. $F$ is often a space of high dimensionality $n'$. Defining the function $\varphi$ is problematic. Many authors propose to use Mercer kernel, that converts the projection problem to: $K(X,Y) = \varphi(X).\varphi(Y)$. Kernel PCA is used for classification. Back-projection in the original space is not simple, kernels are the not often used for shape variability modeling.

### 3.2 Non linear model of variability in pattern recognition

In pattern recognition, a shape is often represented by its average, implying Gaussian modeling. Recent works insert a second parameter : variability. We will present here the model that we adopted. Four major axes define the modelization:
- invariance up to affine transformations using a non-linear space;
- representation of the variability and relationships between features of the shape;
- general formalism mixing points and shapes;
- possibility to use the model for pattern recognition.

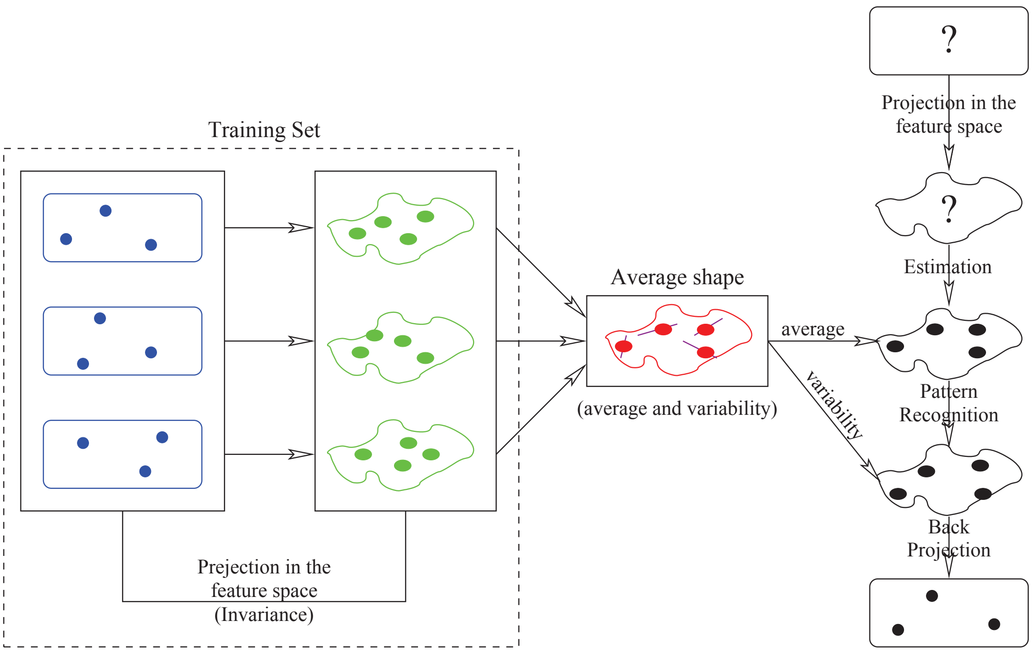



Figure 2: General view : modeling shape variability

Figure 2 summarizes our approach. We start with a training set composed of expertised data. This data is then projected in a feature space. A statistical model provides the average shape and its variability. The average shape can be an estimation of this shape on an unknown image, and can also be the initial shape for an iterative process of shape recognition. In this case, each deformation must verify learned constraints on authorized variability around the average shape. In this paper we suppose that every image contains a reference shape that provides a basis for a common feature space. Three different models were build: a linear one and two non-linears [8]. We will only present the last one.

### 3.2.1 Feature space definition

The first step consists in the detection of the reference shape. This shape is then sampled in $p$ equi-distant points. The non-linear feature space is defined by ratio of surfaces of triangles obtained from the previous sampling. The coordinates of an image point $M(x, y)$ are defined by $\beta$, $\gamma$ et $\delta$ computed for each possible triangle:

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

where $\overline{P_i P_j P_k}$ is the algebraic area of the triangle $P_i P_j P_k$. This coordinates satisfy: $\beta \times \overrightarrow{MP_i} + \gamma \times \overrightarrow{MP_j} + \delta \times \overrightarrow{MP_k} = \overrightarrow{0}$

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.

### 3.2.2 Model: Variability and Relationships

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 $\vartheta^i$ 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 $\vartheta^i$ is also computed. We deduce from this the weighting matrix $P$:

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

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 $\vartheta^i$. Only the most important $d'$ components are retained. These components are those the eigenvalues of which are the highest, and they form the matrix $\Phi$.

### 3.2.3 Pattern recognition: localization of points

Let $\vartheta$ 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:

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

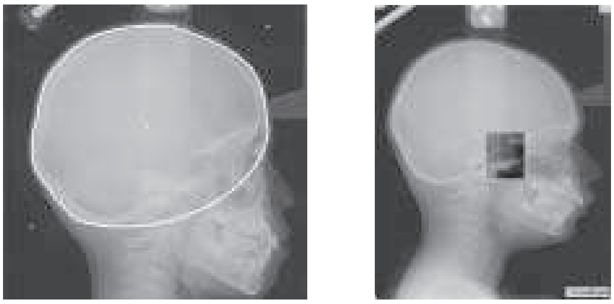
# 4 Results: Evaluation of the method in cephalometry

## 4.1 Goals of cephalometry

The evaluation of our methodology was made in cephalometry. What is cephalometry? It is a discipline used in orthodontics for predicting dental dysharmonies of the patient. This prediction is done on cranial radiographs on which the practitioner landmarks anatomical points. He then obtain a set of distances and angles, and obtain from this the diagnosis. In our case we focalize on landmarking. One of our goals in this project was the automation of landmarking of cephalometric points on a consequent database. All the points that we have to localize have anatomical definitions: their position is defined relatively to anatomical structures such as bones and sutures.

## 4.2 *A priori* knowledge

The first step consists in defining the set of structures that we have to identify. This was possible with the *a priori* knowledge of the practitioner: position of each cephalometric point depends on skull biodynamics. We conclude that our reference shape should be the external contour of the skull. A fully automated method [5] provides us with this shape (figure 3 (a)).

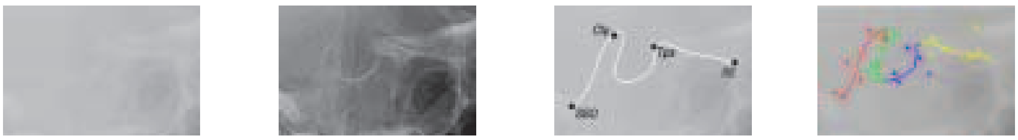


(a) Reference shape (b) Sella turcica region

Figure 3: Position of the region of sella turcica relatively to the detected reference shape

## 4.3 Approximation and learning of the sella turcica

The practitioner provides the expertise of the anatomical structures to identify. In this paper we will focalize on the region of sella turcica (figure 3 (b)) and cephalometric points TPS and CLP (figure 4 (c)).



(a) Original Image (b) Contrasted Image (c) Expertise (d) Approximation

Figure 4: Expertise of the sella turcica (cephalometric points TPS and CLP)

Figure 4 presents the expertise of the sella turcica. Images (a) and (b) present the region of the structure, image (c) presents the expertise of the sella and the points positioned on it, and finally image (d) is the result of the approximation of the sella with four composites Bézier curves. We obtained for this structures 28 control points.

Once we have obtained the approximation of the structure, we can learn in the same model the structures and cephalometric points.

### 4.4 Results

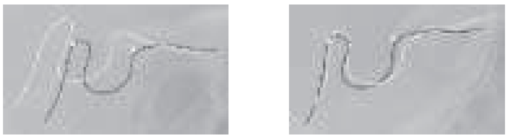


Figure 5: Position of the average shape (black) relatively to the expertise (white) on two images

This study was evaluated on 80 expertised images. Figure 5 presents results obtained with our model. We can see on these two images the position of the average shape relatively to the expertise. It's interesting to notice that our model is adaptive. Average shapes are not the same on images, even if they are in the feature space.

## 5 Conclusion

In this paper, we treated the problem of modeling the variability of a shape. We decompose our study in two parts. The first one concerns the representation of the shape by a set of points, the second one concerns the non linear model that we build. Our shape is represented by a set of points. We approximate the contour of the shape using composite Bézier curves. This approximation is invariant, and provides us with a set of representative control points equivalent to the analytic equation of the polynom. The adopted formalism allows to learn simultaneously points and shapes. Our second goal was to build a non linear model of variability and spatial relationships. We propose here a model based on the detection of a reference shape that defines the projection in a feature space, in which we learn the model. Our methodology was validated in cephalometry and presents interesting results. This model can be used to identify partially occulted data, using learned spatial relationships and visible part of the shape.

## 6 Acknowledgments

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

## References

[1] F.L. Bookstein. *Landmark methods for forms without landmarks: Morphometrics of group differences in outline shape.* Medical Image Analysis, pp. 225-244, 1997.

[2] C. Campbell. *Kernel Methods: a Survey of Current Techniques.* 2000.

[3] T.F. Cootes, C.J. Taylor. *Statistical Models of Appearance for Computer Vision.* Report, 2001.

[4] J. T. Kent, K.V. Mardia. *Shape, Procrustes tangent projections and bilateral symmetry.* Biometrika, Vol. 88(2), pp. 469-485, 2001.

[5] M. Desvignes, B. Romaniuk, J. Robiaille, M. Revenu, M.J. Deshayes. *Détection et modélisation du contour crânien sur des céphalogrammes* GRETSI 2001, Toulouse, France, 2001.

[6] I. L. Dryden. *General Shape and Registration Analysis.* In W. S. Kendall, O. BarndorffNielsen, and M. N. M. van Lieshout (Eds.), SEMSTAT 3, London. Chapman and Hall, 1997.

[7] J.C. Léon. *Modélisation et construction de surfaces pour la CFAO.* Hérmes, Paris, 1991.

[8] B. Romaniuk, M. Desvignes, J. Robiaille, M. Revenu, M.J. Deshayes. *Une approche non linéaire pour la reconnaissance des formes statistique: application à la céphalométrie.* RFIA 13e Congrès Reconnaissance des formes et intelligence artificielle, Angers, Vol. III, pp. 829-838, Janvier 2002.

[9] B. Scholkopf, S. Mika, C. Burges, P. Knirsch, K.-R. Muller, G. Ratsch, A. Smola. *Input space vs. feature space in kernel-based methods.* IEEE Trans. Neural Networks, Vol. 10(5), pp. 1000-1017, 1999.

[10] S. Sclaroff, A.P. Pentland. *Modal Matching for Correspondence and Recognition.* IEEE Transactions on PAMI, Vol. 17(6), pp. 545-561, 1995.