EPJ manuscript No.
(will be inserted by the editor)

Methods of Information Geometry to model complex shapes

Angela De Sanctis[1] and Stefano Antonio Gattone[1]

University "G. d'Annunzio" of Chieti-Pescara, Italy

Abstract. In this paper, a new statistical method to model patterns emerging in complex systems is proposed. A framework for shape analysis of 2− dimensional landmark data is introduced, in which each landmark is represented by a bivariate Gaussian distribution. From Information Geometry we know that Fisher-Rao metric endows the statistical manifold of parameters of a family of probability distributions with a Riemannian metric. Thus this approach allows to reconstruct the intermediate steps in the evolution between observed shapes by computing the geodesic, with respect to the Fisher-Rao metric, between the corresponding distributions. Furthermore, the geodesic path can be used for shape predictions. As application, we study the evolution of the rat skull shape. A future application in Ophthalmology is introduced.

Shape analysis is a timely and interesting research field. It includes Imaging which is very important in medicine for the diagnosis and the study of the diseases. We consider in particular patterns emerging from complex systems as expression of the self-organization phenomenon among their interacting elements due to internal forces, [Bertuglia and Nagaoka, 2000; Nicolis, 1995]. The aim of the paper is the statistical modeling of a shape under the hypothesis that it is possible to extract from it a finite number of representing points, called landmarks.

We propose a framework for shape analysis of 2 dimensional landmark data following the idea of Peter and Rangarajan [Peter and Rangarajan, 2009] in which each landmark is represented by a bivariate normal distribution. However, Peter and Rangarajan restricted to family of distributions with the same variance, using only the means as parameters. This hypothesis does not take into account that variance provides information regarding the dispersion of the real points of the shape around their means. Numerical simulations prove that, in the case of a big variance, the image shows blots, similar to those of a photocopy from a damaged machine. In medical images this constitutes a warning sign of some problems threatening the involved organ, as we have already seen in the process of macular degeneration [Sanctis, 2012]. Therefore we consider variance as additional coordinate for each landmark, reflecting the uncertainty in the landmark's placement and the variability across a family of shapes, and allow the variances to vary among the landmarks and in time.

As mathematical tool we use Information Geometry which consider statistical models as riemannian manifolds with the Fisher-Rao metric. The methodology enables us to study the evolution in time of shapes, in particular geodesic paths allow the interpolation between observed shapes and also can be used to do predictions.

In the paper two applications in Biomedicine are showed.

1 Information Geometry

Differential manifolds are the object of study of Differential Geometry. They are locally euclidean topological spaces, like for example the M-sphere which is "locally equivalent" to the M-dimensional euclidean space. Although differential manifolds are mapped by local coordinates, it is well-known that all the consequent analysis is intrinsic, i.e. it does not depend on the choice of the coordinates. In particular, Information Geometry considers statistical manifolds, which are families of probability distributions or equivalently probability densities with their local coordinates defined by the model parameters [Amari and Nagaoka, 2000; Murray and Rice, 1984]. For example, the normal distribution:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{(x-\mu)^2}{2\sigma^2}\} \tag{1}$$

is univocally identifiable by two parameters, (μ, σ), where μ and $\sigma > 0$ represent respectively the mean and the standard deviation. Thus, we can identify the family of normal distributions with the upper half-plane:

$$\Theta = \{(\mu, \sigma) : \sigma > 0\}.$$

Let us recall that a Riemannian metric on a differential manifold is a metric "compatible" with the system of local coordinates. From Information Geometry we also know that the Fisher information matrix induces a Riemannian metric on the statistical manifold, called the Fisher-Rao metric, as follows

$$ds^2 = \sum_{i,j=1}^{M} g_{ij}(\theta)\theta^i\theta^j \tag{2}$$

with metric tensor:

$$g_{ij}(\theta) = \int p(x/\theta) \frac{\partial}{\partial \theta^i} \log p(x/\theta) \frac{\partial}{\partial \theta^j} \log p(x/\theta) dx. \tag{3}$$

Moreover, it is possible to prove that the above metric is the only one consistent with the theory of Statistical Inference.

In particular, for the family of the univariate normal distributions, simple calculations lead to: $g_{11}(\mu, \sigma) = \frac{1}{\sigma^2}$, $g_{22}(\mu, \sigma) = \frac{2}{\sigma^2}$ and $g_{12}(\mu, \sigma) = g_{21}(\mu, \sigma) = 0$, whence the metric (2) becomes $ds^2 = \left[(d\mu)^2 + 2(d\sigma)^2\right]/\sigma^2$.

It can be proved that one can remove the factor 2 in g_{22} in order to obtain a simpler expression. So we recognize that it is the metric which induces the hyperbolic geometry in the upper half-plane. It is also known that such a geometry is invariant under the special linear group $SL(2, \mathbb{R})$ of all the real unit determinant matrices which has, as its subgroups, the location scale group and the group $SO(2)$ of rotations.

More generally for the parametric family of the bivariate normal distributions with diagonal covariance matrix, an element can be represented by a single point with coordinates $\theta = (\mu_1, \mu_2, \sigma_1, \sigma_2)$ on a 4-dimensional manifold which is the Cartesian product $\Theta_1 \times \Theta_2$ of two half-planes.

Besides, on a Riemannian manifold, there exist curves, called geodesics locally minimizing the distance. We recall that one and only one geodesic is locally determined if:

i) if we are given two points through which the curve has to pass, or

ii) if we are given one point and a tangent vector at the same point, representing the "speed" with which the curve starts from that point.

For the Normal Family, the geodesics are the same as in hyperbolic geometry, that is they constitute the solutions $(\mu(t), \sigma(t))$, with $t \geq 0$, of the following ordinary differential equation:

$$\frac{\dot{\mu}^2}{\sigma^2} + 2\frac{\dot{\sigma}^2}{\sigma^2} = c$$

where $c \neq 0$. It is well known that these solutions are half-circles with their center on μ-axis and vertical lines.

In the product space of the bivariate normal distributions, a curve

$$(\mu_1(t), \mu_2(t), \sigma_1(t), \sigma_2(t))$$

is a geodesic if and only if $(\mu_i(t), \sigma_i(t))$ for every $i = 1, 2$, is a geodesic for the family of the univariate normal distributions.

2 Shape analysis

Consider a geometric object, as for example a triangle in the plane or a human head in the space. The shape of the object consists of all information invariant under similarity transformations that is translations, rotations and scalings [Dryden and Mardia, 1998]. One way to compare shapes of different objects is to register them on some common coordinates system using the similarity transformations [Bookstein, 1986; Kendall, 1984]. In the following, in order to simplify the analysis, we consider planar shapes in the Cartesian plane (x, y).

Data from a shape are often realized as a set of points. Many methods allow us to extract a finite number of points, which are representative of the shape and are called landmarks. Anatomical landmarks are points having a biological relevance (angle of an eye, a special point of the skull,etc) while mathematical landmarks are points having a mathematical relevance (point of maximum curvature, discontinuity points, etc.). Figure 1 shows a representation of the midsagittal section of a typical 15-day-old rat skull M.J. Baer and Ackerman [1983], where the landmarks represent various locations around its boundary. The choice of landmarks is crucial and different choices may lead to different results. In order to select them in a good way, in the applications, experts of the real problem can suggest where it is better to put them.

Suppose now that we have a collection of n planar shapes. We denote the coordinates of the K landmarks of the j-th shape S_j, $j = 1, ..., n$, with

$$\{\mu_{j_1}, \mu_{j_2}, \ldots, \mu_{j_K}\}$$

where the generic element is $\mu_{jk} = \{\mu_{j_{k1}}, \mu_{j_{k2}},\}$ for $k = 1, \ldots, K$.

For each landmark we can estimate the covariance matrix Σ_k by

$$\Sigma_k = \frac{1}{n}\sum_{j=1}^{n}(\mu_{jk} - \bar{\mu}_k)(\mu_{jk} - \bar{\mu}_k)'$$

with $\bar{\mu}_k$ representing the k-th landmark coordinates of the mean shape $\bar{\mu} = \frac{1}{n}\sum_{j=1}^{n}\mu_j$, where μ_j is the $K \times 2$ matrix with the coordinates of the shape S_j.

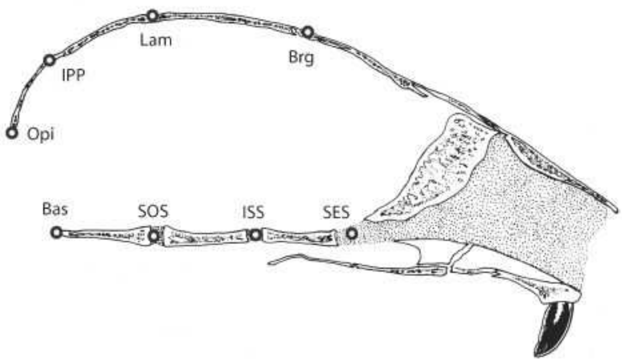

Fig. 1. Realistic drawing of a cross section of a rat cranium with the eight landmarks used in the analysis.

Peter and Rangarajan assume that each landmark of a shape of the population is modeled via a bivariate Gaussian, where the landmark coordinates are the means $\mu_k = (\mu_{k1}, \mu_{k2})$ and the covariance matrix is equal to $\Sigma_k = \Sigma = \sigma^2 \mathbf{I}_2$, $\sigma > 0$, for every landmark. Then the shape is represented by a K-components Gaussian mixture model (GMM)

$$p(\mathbf{x}, \mu, \sigma) = \frac{1}{2\pi\sigma^2 K} \sum_{k=1}^{K} \exp\{-\frac{\|\mathbf{x}-\mu_k\|^2}{2\sigma^2}\} \tag{4}$$

where $\mathbf{x}$ is a generic 2-dimensional vector.

In the absence of any a priori knowledge, it is acceptable to put in the model equal weight $1/K$ to every landmark. Peter and Rangarajan consider σ as a free parameter, which is isotropic across all components. Therefore, they only use the means of a (GMM) as coordinates of the statistical manifold.

On the contrary, our proposal is to relax the isotropic hypothesis by considering variances as additional coordinates for the landmarks of a complex shape, as it is compatible with Information Geometry theory. Thus we consider the following new model for the representation of a shape, where each landmark is modeled via a bivariate Gaussian and the shape is represented by a $2K$-components Gaussian model (GM).

Precisely the k-th landmark, $k = 1, \ldots, K$, of the shape is given by the following:

$$f(\mathbf{x}, \mu_k, \Sigma_k) = \frac{1}{2\pi} |\Sigma_k|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\mathbf{x} - \mu_k)' \Sigma_k^{-1} (\mathbf{x} - \mu_k)\} \tag{5}$$

under the condition

$$\Sigma_k = \sigma_k^2 \mathbf{I}_2 = \mathrm{diag}(\sigma_{k1}^2, \sigma_{k2}^2) \tag{6}$$

where $(\sigma_{k1}^2, \sigma_{k2}^2)$ is the vector of the variances of the k-th landmark coordinates of the shape, for $k = 1, \ldots, K$.

3 Evolution of complex shapes

Starting from a shape represented by $\theta_{ki}(0) = (\mu_{ki}(0), \sigma_{ki}(0))$, we are interested in its time evolution $\theta_{ki}(t) = (\mu_{ki}(t), \sigma_{ki}(t))$. Using the geodesics with respect to Fisher-Rao metric it is possible to predict in a short time the evolution of the shape. Analogously, if two shapes are represented by Gaussian models, we can construct a geodesic between them that will provide us the information about the intermediate shapes (landmarks and their variances). That intrinsic path will drive the reconstruction of the real intermediate shapes in the external space.

More precisely, let S(t) be the shape at time t with landmark coordinates $\mu_k(t) = \{\mu_{k1}(t), \mu_{k2}(t)\}$, $k = 1, \ldots, K$.

Since the same landmarks are collected at different times, the covariance structure may vary from one time to another. We assume the common principal components (CPC) model, where for each landmark k we have

$$\Gamma' \Delta_k(t) \Gamma = \Sigma_k(t) \tag{7}$$

with Γ an orthogonal 2×2-matrix and $\Delta_k(t)$ a diagonal 2×2-matrix.

In order to compute the geodesic path between two observations of the same shape at two different times t_1 and t_2, we propose the following procedure:

– given the covariance matrices $\Sigma_k(t_l)$ under the CPC model for $k = 1, \ldots, K$ and $l = 1, 2$, estimate Γ and $\Delta_k(t_l)$
– transform the coordinates of each landmark to $\nu_k(t_l) = \mu_k(t_l)\Gamma$ so to have covariance matrices equal to $\Delta_k(t_l)$ for $l = 1, 2$
– construct the geodesic path between two shapes

In this way, we allow the variances of the landmarks to vary not only among the landmarks but also between different times. In some cases, it is reasonable to assume isotropy in time, in particular when the change of the shape is due to external forces. Indeed, Peter and Rangarajan refer to the effect of deformation of the external space and unify representation and deformation. On the contrary, we are interested in the natural evolution of the shape induced by internal forces to the system. In this case, the model of Peter and Rangarajan induces a loss of information in the Fisher sense. Our approach, for example, may turn useful in medicine, for the diagnosis and the consequent therapy, to verify its efficacy but also in the screening program to identify diseases precociously.

4 Application: rat calvarial data

As an application, we will work with the rat calvarial data set presented in Bookstein [1991]. It corresponds to 8 landmarks digitized in two dimensions on the skull midsagittal section of 21 rats, which have been collected at ages of 7, 14, 21, 30, 40, 60, 90, and 150 days.

We apply the model in (5) and (6) calculating the covariance matrices $\Sigma_k(t_l)$ under the CPC model (7). For each landmark, the shape evolution between two observed times, say t_1 and t_2, is estimated by computing the geodesic path from time t_1 to t_2. An example is given in the right panel of figure 2, where the x and y coordinates of the mean shapes at times $t_1 = 14$ and $t_2 = 21$ days are plotted together with the geodesics connecting their landmarks.

Furthermore, it is possible to forecast the shape evolution at times close to t_2, following the geodesic path between t_1 and t_2 till a subsequent time t_3. For the rat skull, the predicted mean shapes are plotted in figure 3 together with the observed mean shapes in subsequent times to $t_2 = 14$ days.

The fit seems good except for landmarks 5 and 6 at earlier times. We should expect such a result since the dynamics of growth shows an initial strong development followed by a more stable shape change. More details are in [Sanctis and Gattone, 2015].

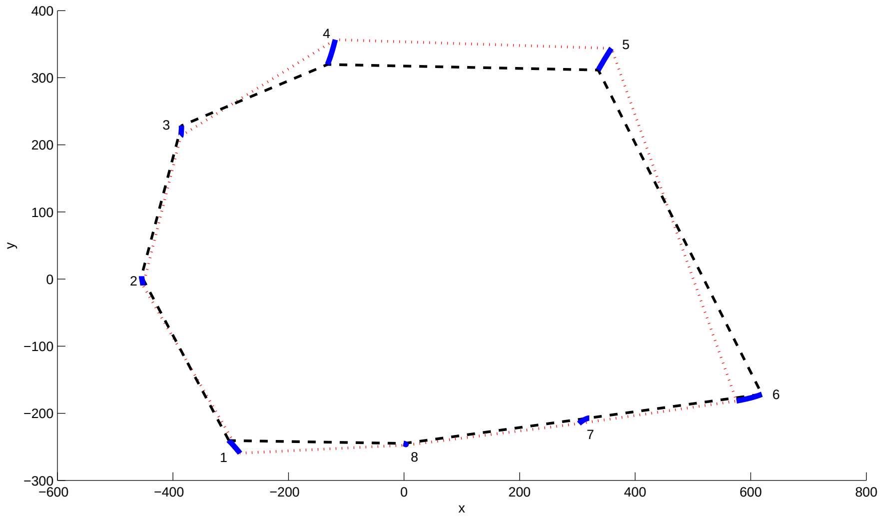

Fig. 2. Mean shape at time $t = 14$ days (dash-dotted line); mean shape at time $t = 21$ days (dashed line); geodesic paths (solid lines) from $t = 14$ to $t = 21$.

5 Future application: Keratokonus

Keratoconus, in Ophthalmology, means the degenerative disease of cornea due to the thinning and bulging of its connective tissue. Real images of a normal cornea and of a cornea with Keratokonus are showed in figure 4.

This pathology modifies the corneal shape, producing serious problems to the vision because the cornea is the first magnifier of the eye, which allows the passage of light. In a normal eye, the images are focused to the same point of the retina then processed correctly by the brain. In the people with Keratoconus this is not possible and causes astigmatism and visual aberrations. In order to detect the corneal shape, an imaging technique is the corneal topography. It is not like a geographic topography because it does not draw the corneal shape. It calculates the curvature in a lot of points of the corneal front then deduces the map of the whole surface. In order to calculate the curvature of the corneal surface, it is possible to use different algorithms: two for the vision power and one for the shape.The Axial algorithm is mostly used to evaluate the power of the central part because it measures the difference in curvature between the real cornea and a semi-sphere. The Tangential algorithm calculates the tangent in every point while the Altitudinal algorithm measures the altitude of the cornea. A complete study requires all the algorithms. The first two are important to set laser for refractive surgeries. The numerical data of the axial and tangential algorithms are transformed producing coloured maps, where every colour corresponds to a diopter range: for convention colder colours refer to more flat areas while warmer colours to more curved regions. In figures 5 and 6 we can see an example of numerical and colorimetric axial map of a patient who does not have Keratoconus.

Using the numerical data of the axial and tangential algorithms as x and y coordinates of a "non-geometrical" shape, we propose to apply our model in order to reconstruct the intermediate steps from an initial situation to the outbreak of the disease, as shown in figures 7 and 8.

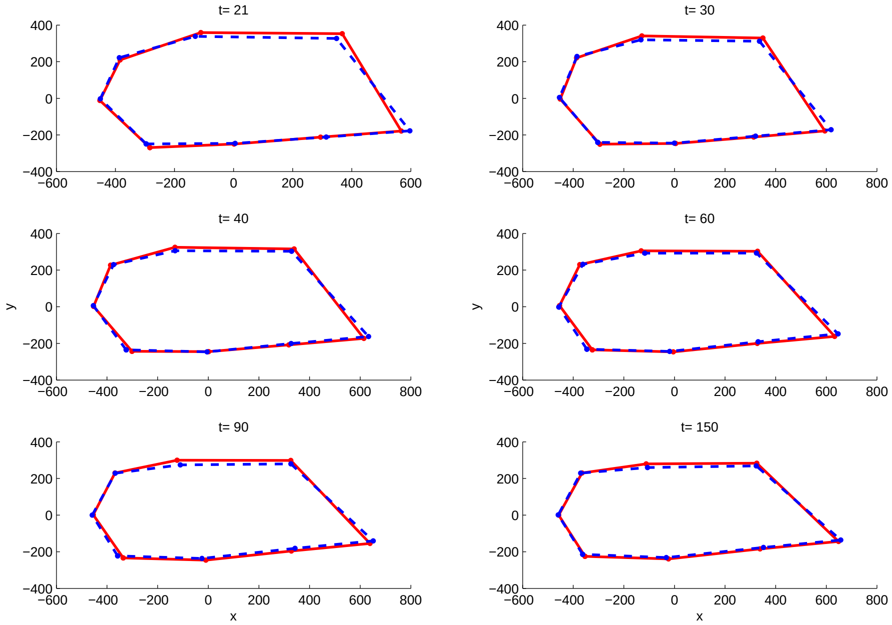

Fig. 3. Rat data set: observed (dashed lines) and estimated (solid lines) mean shapes from day 21 to day 150.

Fig. 4. Normal cornea (on the left) and cornea with Keratoconus (on the right)

Following the rat calvarial case, the model is also able to predict, in a short time, the evolution of the disease. Statistical evidence for this medical case will be the aim of our future work.

6 Conclusions

The most evident property of self-organization phenomena in complex systems is the formation of patterns which evolve in time. Our goal in this paper is to model such shapes statistically, using Information Geometry tools. This is very important mostly in biology when one wishes to study the spreading behavior of an organism but also in medicine for the precocious diagnosis and the analysis of the evolution of different

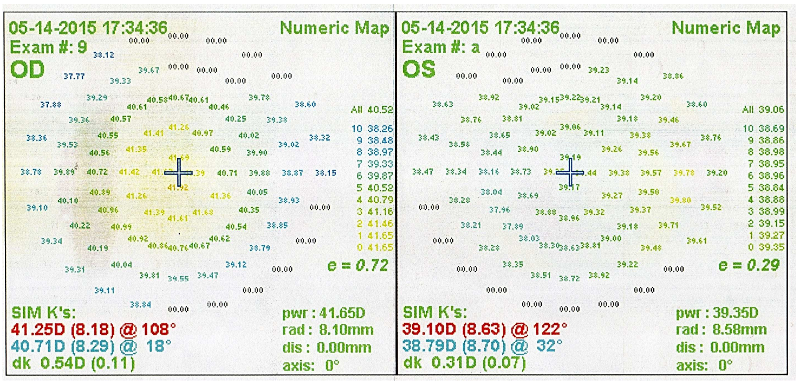

Fig. 5. Corneal topography of a normal patient: numeric axial map.

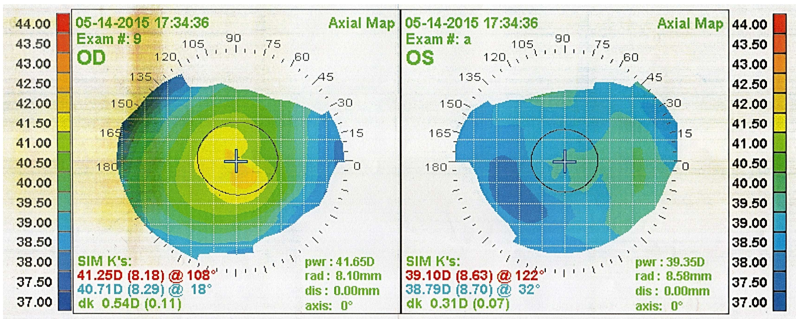

Fig. 6. Corneal topography of a normal patient: Colorimetric axial map.

diseases. Since many statistical methods allow us to extract some points, called landmarks, which are representative of a given shape, we propose to describe a pattern by a finite dimensional Gaussian model. For every landmark, such a model uses both the mean and variance as coordinates varying in time. In particular, the variance is recognized as being able to capture the trend in the self-organizing phenomenon and point out eventual crisis signals in the complex system. According to Information Geometry, we can obtain in this way a statistical manifold and use Fisher-Rao metric as Riemannian metric to determine geodesics locally minimizing distances in the Fisher information sense. Therefore, these curves can be used (a) to reconstruct the intermediate shapes from those known at two different times and also (b) to predict, for short times, the evolution of the pattern from its past. As an application, we have

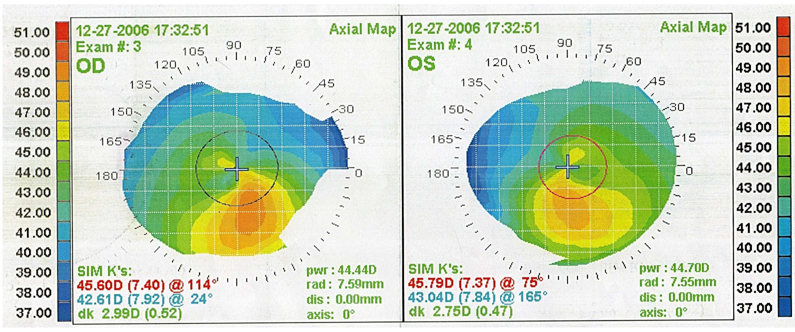

Fig. 7. Corneal topography of a patient with Keratokonus: initial condition of the disease.

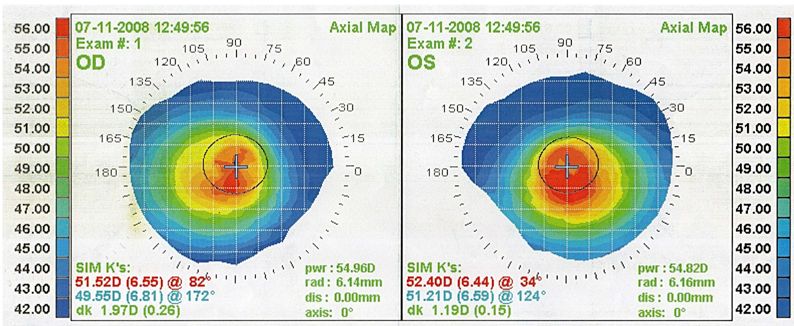

Fig. 8. Corneal topography of a patient with Keratokonus: outbreak of the disease.

considered a rat calvarial data set and proposed to use the same approach to medical images like those from corneal topography.

References

S. Amari and H. Nagaoka. *Methods of Information Geometry*, volume 191 of *Translations of mathematical monographs*. 2000.

C. S. Bertuglia and H. Nagaoka. *Nonlinearity, Chaos, and Complexity: The Dynamics of Natural and Social Systems*. Oxford University Press, 2000.

F. L. Bookstein. Size and shape spaces for landmark data in two dimensions. *Statistical Science*, 1:181–242, 1986.
F. L. Bookstein. *Morphometric Tools for Landmark Data: Geometry and Biology.* Cambridge University Press, 1991.
I. L. Dryden and K. V. Mardia. *Statistical Shape Analysis.* John Wiley & Sons, London, 1998.
D. G. Kendall. Shape manifolds, procrustean metrics and complex projective spaces. *Bulletin of the London Mathematical Society*, 16:81–121, 1984.
J. Bosma M.J. Baer and J. Ackerman. *The Postnatal Development of the Rat Skull.* University of Michigan Press, 1983.
M. K. Murray and J. W. Rice. *Differential Geometry and Statistics.* Chapman & Hall, 1984.
G. Nicolis. *Introduction to Nonlinear Science.* Cambridge University, 1995.
A. Peter and A. Rangarajan. Information geometry for landmark shape analysis: unifying shape representation and deformation. *IEEE Transactions on patter analysis and machine intelligence*, 31:337–350, 2009.
A. De Sanctis. Shape analysis for complex systems using information geometry tools. *Nonlinear Phenomena in Complex Systems*, 15:70–73, 2012.
A. De Sanctis and S.A. Gattone. A study of complex shapes using information geometry. *Nonlinear Phenomena in Complex Systems*, 18:70–80, 2015.