Pattern Recognition 1992 - Proceedings of the 16th OAGM Conference. Vienna, May 6-8, 1992
ISBN 3-85403-062-2

Deformable Templates: a Statistical Setting

Marco Bonetti Mauro Piccioni

Facolta' di Scienze Statistiche
Universita' di Roma La Sapienza
00185 Roma, Italy
and
Dipartimento di Matematica Pura e Applicata
Universita' di L'Aquila
63100 L'Aquila, Italy

Abstract: *In this paper the approach to pattern thery developed by Ulf Grenander is used to construct models of randomly deformable templates. A probability measure on the deformations is constructed through an effort function, which induces a statistical model on the observed objects. The invariance of such model with respect to similarities is discussed. Two examples are given in order to clarify the relevance of the theme and the technical difficulties.*

1 Introduction

In this paper some of the ideas proposed by Ulf Grenander (1976) in his foundations of pattern theory are developed in order to set up a statistical theory of deformable templates. Grenander's is to model the variabilty of a class of objects or shapes through a probability distribution. This allows to check the adequacy of the model by Monte–Carlo sampling and to assess parameter values by estimation from real data.

Randomly deformable templates are random objects which are generated by applying a random deformation to a reference object (the template or prototype), embodying their main desired features. Models of this kind have been recently used in image processing by :Chow, Grenander and Keenan (1991) :and :Amit, :Grenander and Piccioni (1991) .

It is often desirable that family of probability laws on objects should be invariant with respect to some natural transformations of the object space, e.g. the group of shape invariance (translations, rotations and uniform changes of scale), see Kendall (1989). For this reason the theory of invariant statistical models should have a more relevant impact on pattern recognition models.

In the development of deformable template models the developments in the field of continuum mechanics have been a source of ideas (see Kass, Witkin and Terzopoulos, 1987). :In

the deterministic setting a cost :function (which we call effort, conformally to the already mentioned work of Grenander (1976) and Pavel (1989)) is minimized on the object space. Usually such a cost function is the sum of a data–fitting term and a regularity or template–fitting term, which reminds to statisticians the additive decomposition of the logarithm of the posterior density as a sum of the log–likelihood and the logarithm of the prior density in the Bayes' formula. Under suitable assumption this can be made precise clarifying the relation between the principles of least effort and maximum likelihood.

In the paper we focus on the simple problem of classification without degradation due to noise, that is assigning an observed object to one of a finite class of pre–specified possible templates. This can be of interest in practical situations like the detection of abnormalities in biological shapes (see also Grenander and Manbeck, 1992).

The organization of the paper is as follows. The next two sections contain the theoretical part. The remaining section will be devoted to two particular examples: the first (pattern of n points on the circle) is an instance in which the compactness assumptions underlying our approach hold true, whereas in the second (convex sets in the plane) they fail and only the special nature :of the selected deformations, that is a Gaussian model on a commutative group, still allows to :analyze :the problem in a similar fashion.

2 Invariant classification problems

Let $\mathcal{X}$ be the class of observed objects and Θ the class of reference ones, the idea being that observed objects are produced by some reference object $\vartheta \in \Theta$ according to a probability P_ϑ on a σ–algebra of subsets of $\mathcal{X}$.

Let $\mathcal{S}$ be a group acting measurably on $\mathcal{X}$ and on Θ, with the property that

$$P_\vartheta s^{-1} = P_{s\vartheta}, \quad \forall \vartheta \in \Theta, \quad \forall s \in \mathcal{S}. \tag{1}$$

Then we say that the statistical model $\{P_\vartheta, \vartheta \in \Theta\}$ is $\mathcal{S}$–invariant. We refer to $\mathcal{S}$ in general as the group of similarities, so that the $\mathcal{S}$–invariance of the model means in practice that if the observed object coming from the reference ϑ is transformed by a similarity, its probability distribution is the same as if it comes from the "similar" reference object $s\vartheta$.

The easiest way to construct such invariant models is when there is a σ–finite measure λ on $\mathcal{X}$ which is $\mathcal{S}$–invariant, that is $\lambda s^{-1} = \lambda \; \forall s \in \mathcal{S}$. In this case, if we assign P_ϑ through its density w.r.t. to λ, that is

$$P_\vartheta(A) = \int_A p(x|\vartheta)\lambda(dx) \tag{2}$$

then the invariance is obtained provided

$$p(x|\vartheta) = p(sx|s\vartheta) \quad \forall s \in \mathcal{S}. \tag{3}$$

The construction of λ for our class of models will be dealt with in the next section. However, for this to be possible, we assume from now on that $\mathcal{X}$ is a locally compact Hausdorff topological space, on which the compact topological group $\mathcal{S}$ acts topologically. For technical details about topological groups the reader is referred to Lenz (1990).

Now let us examine the following classification problem. Having observed some object $x \in \mathcal{X}$ decide whether it comes from one of the exclusive classes Θ_i, $i = 0, \ldots, k$, where $\Theta_i \cap \Theta_j = \phi$ for $i \neq j$, and $\Theta_0 \cup \ldots \cup \Theta_k = \Theta$. If each of the Θ_i's is invariant under each $s \in \mathcal{S}$, the classification problem is said to be invariant. When $k = 1$ this is the classical testing hypotheses problem, Θ_0 and Θ_1 being the null and the alternative hypothesis, respectively. The simplest situation is when both Θ_0 and Θ_1 are made by a single equivalence class in $\Theta/\mathcal{S}$, which we call from now on a template. Formally this means that for $i = 0, 1$ and $\vartheta \in \Theta_i$, $\vartheta' \in \Theta_i$ there exists $s \in \mathcal{S}$ such that $\vartheta' = s\vartheta$.

When dealing with an invariant classification problem it is reasonable to restrict our considerations to invariant classification rules. A classification rule is a mapping c from $\mathcal{X}$ to $\{0, \ldots, k\}$ specifying for each observed object the class of templates from which it is believed to come from. It is invariant whenever $c(x) = c(sx)$, $\forall\, s \in \mathcal{S}$. The main property of invariant classification rules is that

$$P_\vartheta(c(x) = i) = P_\vartheta(c(sx) = i) = P_{s\vartheta}(c(x) = i) \tag{4}$$

hence the statistical properties of such rules are invariant over each template. In particular for testing problems of hypothesis made by a single template, the features of each rule are given by the two probabilities $P_{\vartheta_0}(A)$ and $P_{\vartheta_1}(A)$, A being the critical region $(c(x) = 1)$ and ϑ_0 and ϑ_1 being arbitrary reference objects belonging to Θ_0 and Θ_1, respectively. It is obvious that any classification rule which is based on the values of invariant statistics (i.e $T(x) = T(sx)$, $\forall\, s \in \mathcal{S}$) will be invariant.

In particular we mention the maximum likelihood classifier

$$T_i'(x) = \sup_{\theta \in \Theta_i} p(x|\vartheta) \tag{5}$$

and the Bayes classifier

$$T_i''(x) = \int_p (x|\sigma\vartheta_i)\nu(d\sigma) \tag{6}$$

where $\vartheta_i \in \Theta_i$ and ν is a left invariant measure on $\mathcal{S}$.

3 Deformations and effort functions

In this section we discuss in an abstract setting an approach to the construction of invariant statistical models for pattern recognition problems. For our source of ideas the reader is again referred to Grenander (1976) and Pavel (1989). In practice the effectiveness of the construction will depend on the different choices which have to be made.

The first and more delicate one is to set up the deformations of the reference objects. These are transformations of Θ into $\mathcal{X}$, which can be thought as explanations of the way a reference

object is appearing in a different form. For mathematical convenience we require deformations to be a locally compact topological group acting topologically on $\mathcal{X}$. This implicitly assumes $\Theta \subset \mathcal{X}$. Also, a right invariant Haar measure μ exists on $\mathcal{D}$ and it is unique up to a factor. Unfortunately often we expect that admissible deformations will vary with the reference object. The final example will show it. This is actually the main problem with this kind of approach.

Our purpose is to construct a probability measure on the group $\mathcal{D}$ of deformations and, by evaluating a deformation on a reference object ϑ, to get the induced probability measure $\{P_\vartheta\}$ describing the observed object as coming from the reference ϑ.

Before moving further let us relate $\mathcal{S}$ with $\mathcal{D}$. There is no need to assume that $\mathcal{S}$ is a subgroup of $\mathcal{D}$ as done by Pavel: the less restrictive requirement is that the transformation $s(d) = s^{-1}ds$ on $\mathcal{X}$ still belongs to $\mathcal{D}$ for any $s \in \mathcal{S}$, $d \in \mathcal{D}$. In order to give a meaning to this we need to ensure that different defermations act differently at least on one object. Moreover we need to ensure that the action of $\mathcal{S}$ so defined on $\mathcal{D}$ is measurable and the Haar measure μ is invariant w.r.t. such an action.

Now the effort function can be introduced. This is a real valued continuous mapping E of $\mathcal{D}$. We do not impose any property of E, except its $\mathcal{S}$–invariance

$$E(d) = E(s^{-1}ds) \quad \forall\, s \in \mathcal{S} \tag{7}$$

and that for some $\beta > 0$

$$Z_\beta = \int_{\mathcal{D}} \exp(-\beta E(\delta))\mu(d\delta) < +\infty\ , \tag{8}$$

which allows to define the probability measure

$$\pi_\beta(A) = Z_\beta^{-1} \int_A \exp(-\beta E(\delta))\mu(d\delta)\ . \tag{9}$$

By the assumptions made so far π_β is $\mathcal{S}$–invariant

$$\pi_\beta(A) = \pi_\beta(s^{-1}As) \quad \forall\, s \in \mathcal{S}\ . \tag{10}$$

It is clear that, being μ a sort of uniform measure on $\mathcal{D}$, E controls the probability around each deformation, which increases as the effort decreases. Moreover as β increases such a behaviour is emphasized. The value of β must thus be selected depending on of the variability which is allowed for the observed objects around the reference ones: hence in general β will very with ϑ, even if it must be constant over similarity classes to keep the model invariant. In the sequel β is keep fixed for notational simplicity.

The final step for the construction of the statistical model is to compute the probability measure $P_\vartheta = \pi_\beta e_\vartheta^{-1}$, where $e_\vartheta(d) = d(\vartheta)$ is continuous, hence measurable. Since π_β has density w.r.t. μ, P_ϑ has a density w.r.t. μe_ϑ^{-1}. Assume finally that $\mathcal{D}$ acts transitively on $\mathcal{X}$, that is for each $\vartheta, x \in \mathcal{X}$ there exists a deformation such that $d(\vartheta) = x$. Then $\lambda = \mu e_\vartheta^{-1}$ does not really depend on ϑ, since $\mu e_\vartheta^{-1} = \tilde{\mu}_d e_{d\vartheta}^{-1}$ and

$$\tilde{\mu}_d(A) = \mu(Ad^{-1}) = \mu(A) \tag{11}$$

since μ is right-invariant. The density $p(x|\vartheta)$ of P_ϑ w.r.t. λ is then the conditional expectation

$$Z_\beta^{-1} E_\mu(e^{-\beta E(d)} | d(\vartheta) = x) \tag{12}$$

which can be seen to obey (3) by virtue of our assumptions. If, for each $\vartheta \in \Theta$, $x \in \mathcal{X}$ there exists a unique $e_(^{-1} x)\vartheta \in D$ which sends ϑ into x, depending measurably on x, then

$$p(x|\vartheta) = Z_\beta^{-1} \exp(-\beta E(e_\vartheta^{-1}(x)) \tag{13}$$

which means that the probability of producing the observed object x from the reference ϑ is locally controlled by the effort of the deformation needed to accomplish the job. The easiest situation to think about is when $\mathcal{X}$ is itself a topological group so that $\mathcal{D} = \mathcal{X}$ acting through the group operation.

The design of deformations must usually be preceeded by a description of objects in term of more convenient mathematical representers which are invariant at least under a smaller closed normal subgroup $\mathcal{N}$ of $\mathcal{S}$. A description $W : \mathcal{X} \to \mathcal{X}'$ has the property that

$$W(x) = W(x') \Leftrightarrow \exists n \in \mathcal{N}, x = nx' \quad .$$

Thus $\mathcal{X}'$ can be made homeomorphic to $\mathcal{X}/\mathcal{N}$ and will serve as a new object space, whereas the new reference objects become the elements of $\Theta' = \Theta/\mathcal{N}$. By the normality of $\mathcal{N}$, the equivalence classes of $\mathcal{S}$

$$[s] = \{n_1 s n_2 : n_1, n_2 \in \mathcal{N}\}$$

became a compact group $\mathcal{S}' = \mathcal{S}/\mathcal{N}$ by inheritin g the group structure of $\mathcal{S}$.

Finally let us try to make a comparison between the maximum likelihood and the Bayes classifiers (5) and (6), respectively. If we restrict our consideration to invariant classification rules we are actually considering only the projection of the observed object X on the quotient space $\mathcal{X}/\mathcal{S}$: call ρ such a projection. By invariance the law of $\rho(X)$ depends on ϑ only through $\rho(\vartheta)$.

Since X has density $p(x|\vartheta)$ w.r.t the $\mathcal{S}$-invariant reference measure λ, $\rho(X)$ has density

$$E(p(X|\vartheta)|\rho(X))$$

w.r.t. to the image of λ under ρ. Being λ $\mathcal{S}$-invariant it is possible to show (Eaton, 1989) that such a conditional expectation is equal to the Bayes classifier

$$\int p(\sigma x|\vartheta)\nu(d\sigma)$$

which is trivially seen to depend on ϑ only through its equivalence class, since by (3) and the invariance of ν, for $t \in \mathcal{S}$

$$\int p(\sigma x|t\vartheta)\nu(d\sigma) = \int p(t^{-1}\sigma x|\vartheta)\nu(d\sigma) = \int p(\sigma x|\vartheta)\nu(d\sigma).$$

At least when the number of classes $k = 2$ and both the classes are constituted by a single template, the reduction to invariant rules converts the problem into the testing of two point

hypotheses in $\Theta/\mathcal{S}$. Hence, by the Neyman–Pearson lemma, the optimal tests should be based on the statistic

$$\frac{T_1(x)}{T_0(x)} = \frac{\int p(\sigma x|\vartheta_1)\nu(d\sigma)}{\int p(\sigma x|\vartheta_0)\nu(d\sigma)} \tag{14}$$

where ϑ_1 and ϑ_0 are arbitrary representers of the two templates and $p(x|\vartheta)$ is given by (12). At least in this case, the classifier (6) is preferable to (5). It is worth to note that the choice of the threshold corresponding to a given significance level, for discriminating between the two templates can be done through Monte–Carlo simulations.

Under suitable conditions the above results can be generalized to the case in which $\mathcal{S}$ is only locally compact, see Eaton (1989).

4 Two examples of application

Example 1. Pattern of points on the circle. We imagine to observe a pattern of n points on the unit circle. For problems of this kind in biological science see Batschelet (1981). We do not distinguish between patterns which are rotated of an arbitrary angle or reflected along a diameter of the circle. Moreover, even if we number the points for convenience, permutation of the order does not affect the pattern.

Mathematically, this means that $\mathcal{X} = (S^1)^n$ and the group of similarities is made by $\mathcal{S} = \mathcal{P}_n \times S^1 \times \{-1, 1\}$, where $\mathcal{P}_n$ is the group of permutations of n elements. The group operation on $\mathcal{S}$ is

$$(\tilde{\pi}, \tilde{\vartheta}, \tilde{j}) \cdot (\pi, \vartheta, j) = (\pi\tilde{\pi}, \tilde{j}\vartheta + \tilde{\vartheta}(\text{mod}2\pi), \tilde{j}j) \tag{15}$$

and the action of $\mathcal{S}$ on $\mathcal{X}$ is

$$(\pi, \vartheta, j)(x_1, \ldots, x_n)_i = j x_{\pi(i)} + \vartheta(\text{mod}2\pi) \; . \tag{16}$$

It is clear that $\mathcal{S}$ is a compact group and the Haar measure is the product of uniform measures on $\mathcal{P}_n$, S^1 and $\{-1, 1\}$, respectively.

It is quite natural to try to factor out at least the effect of rotations by putting the origin into the first point automatically. This means using the descriptions

$$W : (S^1)^n \to (S^1)^{n-1} \tag{17}$$

$$W(x_1, \ldots, x_n) = (x_2 - x_1, \ldots, x_n - x_1) \tag{18}$$

from which

$$W(\underline{x}) = W(\underline{y}) \Leftrightarrow x_i = y_i + \vartheta, \; i = 1, \ldots, n$$

hence $(S^1)^{n-1}$ is homeomorphic to $(S^1)^n/\mathcal{N}$, where $\mathcal{N}$ is the subgroup of rotations. It is straightforward to see that $\mathcal{N}$ is closed and normal and $Y/\mathcal{N}$ is homeomorphic to $\mathcal{P}_{n-1} \times \{-1, 1\}$ acting on $(S^1)^{n-1}$ by

$$(\pi, j)(z_1, \ldots, z_{n-1})_i = j z_{\pi(i)} \; . \tag{19}$$

Next we model deformations on $(S^1)^{n-1}$ simply by its product group operation, that is $\underline{d} \in (S^1)^{n-1}$ operates on $\underline{z} \in (S^1)^{n-1}$ by

$$\underline{d}(\underline{z})_i = d_i + z_i(\text{mod}2\pi)\ . \tag{20}$$

The Haar measure μ on $\mathcal{D}$ is then simply the product of of n times the Lebesgue measure on the circle, which is clearly invariant under permutation of coordinates and reflections around π. Hence the reference measure λ on $\mathcal{D}$ coincides with μ and is $\mathcal{S}$-invariant. In order for the density to satisfy (3) we must impose (7) on the effort, which means:

$$E(d_1, \dots, d_{n-1}) = E(d_{\pi(1)}, \dots, d_{\pi(n-1)}), \pi \in \mathcal{P}_{n-1} \tag{21}$$

$$E(d_1, \dots, d_{n-1}) = E(2\pi - d_1, \dots, 2\pi - d_{n-1}). \tag{22}$$

A possible choice is then

$$-E(d_1, \dots, d_{n-1}) = \sum_{i=1}^{n-1} \cos d_i\ .$$

Then (9) is the distribution of $n-1$ independent von Mises variables (Batschelet, 1981).

Finally, let us remark that given any description of an observed object $(z_1, \dots, z_{n-1})$ and a description of a reference object $(\vartheta_1, \dots, \vartheta_{n-1})$ there is a unique deformation

$$d_i = z_i - \vartheta_i \quad (\text{mod}2\pi) \tag{23}$$

sending $\underline{\vartheta}$ into $\underline{z}$. By consequence

$$p(\underline{z}|\underline{\vartheta}) \alpha \exp(-\beta E(\underline{z} - \underline{\vartheta})). \tag{24}$$

In order to appreciate with a practical example the difference between the various classifiers let us consider an instance of a classification problem with two hypotheses, with $n = 3$. Given a specific pattern of three points we must decide if they are just deformations of a pattern of three points on the same location, or there is one of them coming from an antipodal position. By using the description above, the parametric hypotheses are thus

$$\Theta_0 = \{(0,0)\} = \{\vartheta_1\}$$

$$\Theta_1 = \{(0,\pi),(\pi,0)\} = \{\vartheta_2, \vartheta_3\}\ .$$

Thus, given the observed object (z_1, z_2), the deformation needed for ϑ_1 to produce it is just (z_1, z_2), whereas for ϑ_2 and ϑ_3 is $(z_1, (z_2 - \pi)\text{mod}2\pi)$ and $(z\text{mod}2\pi, z_2)$, respectively. Again by the symmetry of the effort function z_1 and z_2 around π so that $0 \le z_1 \le z_2 \le \pi$. Then the three efforts are

$$E(z_1, z_2) = -(\cos z_1 + \cos z_2)$$

$$E(z_1, \pi - z_2) = -(\cos z_1 + \cos(\pi - z_2)$$

$$E(\pi - z_1, z_2) = -(\cos z_2 + \cos(2\pi - z_1))\ .$$

It is clear that

$$E(z_1, \pi - z_2) \le E(\pi - z_1, z_2)$$

with the equality holding only when $z_1 = z_2$. By consequence the maximum likehooc classifier

$$e^{\beta(E(z_1,z_2) - E(z_1, \pi - z_2))} = e^{\beta\{\cos(\pi - z_2) - \cos z_2\}} \tag{25}$$

is based only on z_2. At the contrary the Bayes classifier involves both z_1 and z_2, being proportional to

$$e^{\beta(E(z_1,z_2) - E(z_1,\pi - z_2))} + e^{\beta(E(z_1,z_2) - E(\pi - z_1, z_2))} = e^{\beta\{\cos(\pi - z_2) - \cos(z_2)\}} + e^{\beta\{\log(\pi - z_1) - \cos z_1\}}\ . \tag{26}$$

This is in agreement with our former remark about the lack of optimality of the maximum likelihood (minimum effort) classifier, which ignores part of the available data.

Example 2. Convex sets in the plane. The essentials of this example are contained in Grenander (1976). Suppose we need to classify bounded closed convex sets in the plane up to shape invariance (translations, rotations and uniform scale changes). In order to avoid unnecessary technicalities we consider strictly convex sets with C^2 boundary. It is clear that the boundary of such sets can be parametrized by the angle ψ formed by the exterior normal to the boundary and the x-axis. Then by

$$\dot{x}(\psi) = -c \sin \psi R(\psi) \tag{27}$$

$$\dot{y}(\psi) = c \cos \psi R(\psi) \tag{28}$$

where $R(\psi)$ is a positive bounded continuous function on S^1 normalized to become a density. In geometrical terms $cR(\psi)$ is the radius ψ of curvature at ψ, $c = \frac{2\pi}{L}$, L being the length of the boundary. Hence

$$x(\psi) = x(0) - c \int_0^\psi \sin \lambda R(\lambda) d\lambda \tag{29}$$

$$y(\psi) = y(0) + c \int_0^\psi \cos \lambda R(\lambda) d\lambda \tag{30}$$

from which it is seen that R is common to sets which are obtained from the original one through translation and uniform scale change. On the other hand any positive normalized continuous function on S^1 defines through (29-30), for any choice of $(x(0), y(0))$ and $c > 0$ the boundary of a strictly convex set provided

$$\int_0^{2\pi} \sin \lambda R(\lambda) d\lambda = \int_0^{2\pi} \cos \lambda R(\lambda) d\lambda = 0\ . \tag{31}$$

Thus the normalized radius of curvature is a description of the above class in terms of the set $\mathcal{X}$ of densities in $C(S^1)$ which satisfy (29). In computer vision such a description has been introduced by Slansky and Nahin (1972). Sets with the same description are related through an element of the group $\mathcal{N}$ of translations and uniform scale changes which is a closed normal subgroup of the shape invariance group $\mathcal{S}$. The action of $\mathcal{S}/\mathcal{N}$ (rotations) on X is through a shift of the domain

$$\varphi(R)(\lambda) = R(\lambda - \varphi(\text{mod}2\pi))\ , \quad \lambda \in S^1\ . \tag{32}$$

The situation seem to fit our scheme but since $\mathcal{X}$ is not locally compact, there cannot be a locally compact group $\mathcal{D}$ acting transitively on $\mathcal{X}$. This rules out the possibility of applying the previous theory.

However, if we are willing to forget the positivity constraint on R, the cone $\mathcal{D}$ of functions in $C(S^1)$ with zero mean and satisfying (31) is a commutative group acting on $\mathcal{X}$ through addition. Moreover given the observed R and the reference R_0, $\delta = R - R_0$ is the unique deformation sending the latter into the former. It is clear that we cannot model deformations by means of its density w.r.t. an invariant reference measure, since there is no measure with this feature. However we can define a density at least on a dense set of "regular" objects δ by comparing the ratio of the probability of a ball around δ with the probability of a ball around zero and then taking the limit as the radius decrease to zero. This construction works for example in the Gaussian case, in which the formal density is $\exp\left(-\frac{\beta}{2}E(\delta)\right)$, where E is the quadratic form associated to the inverse of the covariance operator. For example, the effort

$$E(\delta) = \int_0^{2\pi} |\dot{\delta}(\psi)|^2 d\psi \tag{33}$$

has this interpretation for $\delta \in H^2(S^1)$. It formally corresponds to model the additive deformation as a periodic Brownian motion. This is very convenient since the constant, sine and cosine are eigenfunctions of the covariance operator, hence the constraints (31) and the zero mean can be taken into account by simply deleting the corresponding terms in the orthogonal series expansion. For example the periodic Brownian motion without the first three Fourier terms is written as

$$X_t = \frac{1}{\sqrt{\pi\beta}} \sum_{m=2}^{\infty} m^{-1}(\xi_m \cos mt + \eta_m \sin mt) \tag{34}$$

where $\{\xi_i, \eta_i\}$ are independent $N(0,1)$ random variables.

The effort (33) is very sensitive to local irregularities of the boundary and give rise to a process (34) with very irregular sample paths. Grenander (1976) considered series expansion of this type with a general choice of coefficients c_m replacing m^{-1}, which allow also to control the smoothness of the sample paths of the process. The corresponding effort Q is then easily written only in the Fourier domain.

The reader should also have noticed that the choice of additive deformations d is such that the shift conjugate deformation $\varphi^{-1} d\varphi$ by $\varphi \in (0, 2\pi)$ is nothing but the additive deformation by

$$\tilde{d}(\lambda) = d((\lambda + \varphi)(\mathrm{mod} 2\pi)) \ . \tag{35}$$

The $\mathcal{S}$-invariance of the effort thus requires the effort to be the same no matter which is the origin chosen on the circle, which is fulfilled by any random Fourier expansion of the type considered above.

To summarize in this case from an effort function E we are able to construct an $\mathcal{S}$-invariant stochastic model for the deformations, having a formal density proportional to $\exp\left(-\frac{\beta}{2}E\right)$.

The maximum likelihood classifier

$$\exp\left(-\frac{\beta}{2}(E(R - R_1) - E(R - R_0))\right)$$

and the Bayes one

$$\frac{\int_0^{2\pi} \exp\left(-\frac{\beta}{2}E(R - \varphi(R_1)\right) d\varphi}{\int_0^{2\pi} \exp\left(-\frac{\beta}{2}E(R - \varphi(R_0)\right) d\varphi}$$

are then well defined $\mathcal{S}$-invariant statistics, where R is the observed boundary curve and R_0 and R_1 those of representers of the two templates Θ_0 and Θ_1, all supposed to be regular enough.

5 References

Y. :Amit, U. Grenander, M. Piccioni (1991), Structural image :restoration : through deformable templates, Jour. Amer. :Stat. :Ass. 86, 376-387.

E. :Batschelet (1981), Circular Statistics in :Biology, :Academic Press, London.

Y. :Chow, :U. Grenander, D. M. Keenan (1991), :Hands: :a :Pattern Theoretical Study of Biological Shape, Springer-Verlag, New York, 1991.

L. D. Cohen (1991), On active contour models and balloons, :Comp. Vis. Graph. and Image Proc., 53, 211-218.

M. L. Eaton (1989), Group Invariance Applications in Statistics, CBMS, Reg. Conf. Series in Prob. and Stat., IMS, Hayward.

U. Grenander (1976), Pattern Analysis, Lectures in Pattern Theory Vol. 1, Springer-Verlag, New York.

U. Grenander, K. Manbeck (1992), A stochastic shape and color model : for defect detection in potatoes, Rep. Patt. Th. :156, :Div. Appl. Math. Brown Univ., Providence.

D. G. Kendall (1989), A survey of the statistical theory of :shape, Statistical Science, 4, 87-120.

M. Kass, A. Witkin, D. Terzopoulos (1987), Snakes: active contour models, Int. Jour. Comp. Vis. 1, 321-331.

R. Lenz (1990), Group Theoretical Methods in Image Processing, Lect. Notes in Comp. Sc. 413, Springer-Verlag, Berlin.

M. Pavel (1989), Foundations of Pattern Recognition, Dekker, :New York.

J. Slansky, P. J. Nahin (1972), A parallel mechanism for :describing silhouettes, IEEE Trans. Comp., 11, 1233-1239.