A consistent framework for a statistical analysis of surfaces based on generalized stochastic processes

Benjamin Coulaud, Frédéric Jp Richard

A consistent framework for a statistical analysis of surfaces based on generalized stochastic processes

Benjamin Coulaud[1,*] **and Frédéric J.P. Richard**[1,**]

[1] *Aix Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373, Marseille, France.*
e-mail: [*] `benjamin.coulaud@univ-amu.fr`; [**] `frederic.richard@univ-amu.fr`

Abstract: The statistical analysis of surfaces is an important issue of Image Analysis, especially in Computational Anatomy. In 2005, Glaunès and Vaillant proposed to handle surfaces through some mathematical currents defined as linear forms on a space of mappings from $\mathbb{R}^3$ into itself. In this paper, we extend this deterministic representation of surfaces using some random linear forms inspired from generalized stochastic processes. Then, we set an observation model where observed surfaces are viewed as random variations of a mean representative of a population (called the template). This observation model accounts not only for the variability of surfaces within an homogeneous population but also for errors due to acquisition. Within this model, we construct an estimate of the template and establish its consistency.

Keywords and phrases: surface statistics, mathematical currents, generalized stochastic processes, template estimation, computational anatomy, image analysis.

1. Introduction

Since its emergence in engineering and research, the field of image analysis have always been influenced by statistical issues. Confronted with a set of images coming from homogeneous populations, it was often a question of using statistical models to extract, compare or validate some population properties by means of estimation and test techniques. Such issues still remain central to the so-called Computational Anatomy [15], an important discipline of image analysis which is mainly motivated by medical applications. Specific to this discipline, images from a same population are considered as variations of a mean image called a template. Variations result from both geometrical deformations of the template and additive intensity variations. The generic observation model can be formally stated as

$$I_k = \varphi_k \cdot \mu + B_k, \tag{1}$$

where I_k is a kth observed image in a population, $\varphi_k \cdot \mu$ a template μ deformed by the action of a deformation φ_k, and B_k a random additive term accounting for variations [2, 15]. In this context, the estimation of the template μ from a sample of observed images is a critical issue [2, 1, 20, 19, 22].

Computational Anatomy has been extensively developed for images [2, 3, 15]. But, it has always been concerned with primitives extracted from images such as curves [25], surfaces [6, 12, 21] or bundles [8, 7, 13, 14]. However, the definition of a model of the form (1) for these objects is problematic in many ways. In particular, the addition operation is not obviously defined since these objects do not belong to any Euclidean space. More importantly, it is still an open issue to give a mathematical meaning to random variations for these objects. In this paper, we focus on surfaces. We first propose a stochastic framework where Model (1) can be correctly interpreted when dealing with surfaces. In this framework, we then construct an estimate of the template and further establish its consistency in the case when they are not geometrical deformations.

In their seminal work [23], Glaunès and Vaillant have achieved a major breakthrough for the study of surfaces in the context of Computational Anatomy [11, 23, 12, 8, 7]. Rather than dealing directly with surfaces, they proposed to use surface representations inspired from mathematical currents. These representations are continuous linear forms defined on a Hilbert space V of smooth mappings from $\mathbb{R}^3$ into itself (see Section 2 for more details). Belonging to an Euclidean space, they can be manipulated with an addition and compared with a norm. Completed by an action of deformations [23], they can be used to compute templates and deformations from observation samples (template registration). However, they are purely deterministic and can not account for random variations within a model of the form (1).

To overcome this shortcoming, we propose to extend these representations in a stochastic way by defining random linear forms on the functional Hilbert space V. This extension is inspired from the generalized stochastic processes introduced by Gelfand and Vilenkin [10] and Itô [16, 17], and also studied in [9, 18, 24]. Using this extended representation, we build an observation model of the form of (1). In this model, observations are intended to be random linear forms associated to some meshed surfaces. These meshed surfaces are usually extracted from images by practitioners. Here, they are considered as discrete approximations of smooth real-world surfaces seen through an imaging device with a given precision. The model template stands for a mean representation of real-world surfaces over a population. Model variations account for two effects: an acquisition error due to the approximation of real-world surfaces by meshed surfaces, and a variability of real-world surfaces across a population.

After model definition, we investigate the issue of estimating the template. By construction of surface representations, the template belongs to an infinite dimensional space (the topological dual V^* of V). Hence, for its estimation, it has to be approximated in some finite-dimensional subspaces of V^*. In this paper, we adopt a Ritz-Galerkin approach to construct a series of embedded finite-dimensional subspaces tending to V^* [5]. Within each subspace, we then derive a maximum likelihood estimate of the template. We eventually show that the template estimates tend to the template as the approximation subspace tends to V^*, the number of samples to infinity, and the precision at which real-world surfaces are discretized to 0. Let us emphasize that this consistency result takes into account two main practical approximations resulting from surface discretizations and template approximations.

In Section 2, a surface representation akin to [23] is presented in a deterministic framework, and extended in a stochastic one. This section further includes the construction of an observation model. It ends up with a first asymptotic result ensuring the convergence of surface discretization errors to 0 as the precision tends to 0. Section 3 is devoted to the estimation of the template. The construction of the estimate is described into details together with the approximation procedure. The main consistency theorem is eventually established.

2. Surface modeling

2.1. Deterministic representations of surfaces

Let $\mathcal{D}$ be a closed and bounded cube of $\mathbb{R}^3$, and Θ be a set of compact and continuously differentiable surfaces included in $\mathcal{D}$. Let V_0 be a a Reproductive Kernel Hilbert Space (RKHS) [4] of $L^2(\mathcal{D},\mathbb{R})\cap C^1(\mathcal{D},\mathbb{R})$ equipped with an inner product $\langle\cdot,\cdot\rangle_{V_0}$ and having a reproductive kernel K_0 bounded on $\mathcal{D}\times\mathcal{D}$. Define the Hilbert space $V=V_0^3$ equipped with the inner product $\langle\cdot,\cdot\rangle_V$ defined, for any $v=(v_1,v_2,v_3)$ and $w=(w_1,w_2,w_3)$ in V, by $\langle v,w\rangle_V=\sum_{i=1}^{3}\langle v_i,w_i\rangle_{V_0}$, and its associated norm $\|\cdot\|_V$.

Following [23], any surface $\mathcal{S}$ of Θ can be represented by a linear form s defined, for any $w\in V$, by

$$s(w)=\int_{\mathcal{S}}\langle w(x),e_1(x)\wedge e_2(x)\rangle_{\mathbb{R}^3}d\sigma_{\mathcal{S}}(x), \tag{2}$$

where $\wedge$ stands for the cross product on $\mathbb{R}^3$, $\langle\cdot,\cdot\rangle_{\mathbb{R}^3}$ the Euclidean inner product on $\mathbb{R}^3$, $\sigma_{\mathcal{S}}$ the surface element on $\mathcal{S}$, and vectors $e_1(x),e_2(x)$ form a direct orthogonal basis of the tangent plane to $\mathcal{S}$ at the point x. The representative s is an element of the topological dual V^* of the Hilbert space V. Indeed, s is linear and, due to RKHS properties of V,

$$|s(w)|\leq\mathcal{A}(\mathcal{S})\|w\|_\infty\leq C\mathcal{A}(\mathcal{S})\|w\|_V,$$

where $\mathcal{A}$ denotes the area measure on surfaces. Hence, s is also continuous.

In practice, an observed surface is usually a meshed surface extracted from an image using some processing. It is a discrete approximation of a real-world surface seen through an imaging device. In our modeling, real-world surfaces $\mathcal{S}$ are in Θ. They are approximated by discrete surfaces F^ρ of the form

$$F^\rho=\bigcup_{m=1}^{M^\rho}f_m^\rho, \tag{3}$$

where, for $m\in\{1,\cdots,M^\rho\}$, f_m^ρ are triangular faces, and ρ stands for a precision parameter defined as the maximal area of these triangles.

The representative s^ρ of a discretized surface F^ρ as the linear form is defined, for any $w\in V$, by

$$s^\rho(w) = \sum_{s=1}^{M^\rho} \int_{f_m^\rho} \langle N_m^\rho, w(x)\rangle_{\mathbb{R}^3} dx, \tag{4}$$

where N_m^ρ denotes the unitary vector normal to the face f_m^ρ and external to the surface F^ρ. The error e^ρ of approximating a surface $\mathcal{S}$ by a discrete surface F^ρ can be expressed in terms of linear forms as $e^\rho = s - s^\rho$.

2.2. Stochastic extension of representations

Let $\mathcal{E} = (\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $L^2(\mathcal{E})$ the set of square integrable random variables defined on $\mathcal{E}$, equipped with the usual inner product $\langle \cdot, \cdot \rangle_{L^2(\mathcal{E})}$ and its associated norm $\| \cdot \|_{L^2(\mathcal{E})}$.

A random linear form S on V is a linear and continuous application mapping V into $L^2(\mathcal{E})$, such that

$$\mathbb{P}(\exists\, C > 0, \sup_{\|w\|_V \leq 1} |S(w)| \leq C). \tag{5}$$

Property (5) means that realizations of S, viewed as mappings from V to $\mathbb{R}$, are almost surely in V^*. The space of random linear forms will be denoted by $V^*_{\mathcal{E}}$. It can be equipped with a vectorial space structure and a norm $\|.\|_{V^*_{\mathcal{E}}}$ defined, for any $S \in V^*_{\mathcal{E}}$, as

$$\|S\|_{V^*_{\mathcal{E}}} = \sup_{\|w\|_V \leq 1} \|S(w)\|_{L^2(\mathcal{E})}. \tag{6}$$

The law of a random linear form S is defined as the set of distribution laws of random vectors $(S(v_1), \cdots, S(v_N))$ associated to any N-uplet $(v_1, \cdots, v_N)$ of elements of V. Two random linear forms S and T are said independent if, for any couple of N-uplets $(v_1, \cdots, v_N)$ and $(w_1, \cdots, w_N)$ of elements of V, random vectors $(S(v_1), \cdots, S(v_N))$ and $(T(w_1), \cdots, T(w_N))$ are independent.

The expectation $\mathbb{E}^*[S]$ of a random linear form S is defined as the mapping from V into $\mathbb{R}$ given, for any $w \in V$, by

$$\mathbb{E}^*[S](w) = \mathbb{E}[S(w)],$$

where $\mathbb{E}$ denotes the usual expectation of square-integrable random variables. A random linear form whose expectation is identically zero is said centered. The autocovariance γ_S of a random linear form S is the mapping from $V \times V$ into $\mathbb{R}$ given, for any $(v, w) \in V^2$, by

$$\gamma_S(v, w) = \mathrm{Cov}(S(v), S(w)), \tag{7}$$

where Cov stands for the usual covariance of square-integrable random variables. Due to RKHS properties, the expectation of any random linear form on V is an element of V^*. Similarly, its autocovariance is a continuous bilinear form on $V \times V$, which is an element of V^{**}. Therefore, we can define Riesz representatives for both of them.

A random linear form $S \in V_{\mathcal{E}}^*$ is Gaussian if, for any N-uplet $(v_1, \cdots, v_N)$ of V, the random vector $(S(v_1), \cdots, S(v_N))$ is Gaussian. Let us quote that a random Gaussian linear form S is entirely characterized by its expectation $\mathbb{E}^*[S]$ and autocovariance γ_S.

As mention earlier, the surface representation of this section is related to generalized stochastic processes [10, 16, 17]. These processes are also random linear forms but they are mapping the space of C^∞-functions with compact supports to a set of random variables which is more general than $L^2(\mathcal{E})$. The adaptation of definition spaces of these processes allowed us to obtain a space $V_{\mathcal{E}}^*$ with some required properties. In particular, we could define a norm on the space $V_{\mathcal{E}}^*$ and guarantee the existence and continuity of expectations and autocovariances of its elements.

2.3. Observation model

For $k \in \{1, \cdots, K\}$, let F_k^ρ be a meshed surface which is obtained through an imaging device and approximates a real-world surface $\mathcal{S}_k$ with a precision ρ. Let s_k^ρ and s_k be representatives in V^* of F_k^ρ and $\mathcal{S}_k$, respectively.

We consider representatives s_k^ρ and s_k as realizations of random linear forms S_k^ρ and S_k, respectively. We define the error of approximating S_k by S_k^ρ in terms of random linear forms as

$$\epsilon_k^\rho = S_k^\rho - S_k. \tag{8}$$

We consider that surfaces were acquired in same conditions and that approximation errors are independent from a surface acquisition to another. Hence, we assume that discretization errors ϵ_k^ρ are independent and identically distributed according to the law of a same centered Gaussian random linear form with autocovariance γ_ϵ^ρ.

We will also use the mean error defined, for any $K > 0$, as

$$\overline{\epsilon_K^\rho} = \frac{1}{K} \sum_{k=1}^{K} \epsilon_k^\rho. \tag{9}$$

Besides, we model variations of real-world surfaces through their associated random linear forms. More precisely, we assume that, for all $k \in \{1, \cdots, K\}$,

$$S_k = \mu + B_k, \tag{10}$$

where the template μ is a linear form of V^* and variations B_k are independent and identically distributed according to the law of a same centered Gaussian random linear form with autocovariance γ^B. In this model, the expectation of S_k is μ. Hence, the template can be seen as a mean element of the surface population.

Combining the Equations (8) and (10), we eventually obtain the observation model

$$S_k^\rho = \mu + B_k + \epsilon_k^\rho, \; k \in \{1, \cdots, K\}. \tag{11}$$

2.4. Convergence of discretization errors

In this part, we state some supplementary assumptions ensuring that the discretization errors ϵ_k^ρ of our model converge to 0 as the precision ρ tends to 0. Such a convergence is required to guarantee the coherence of the transition from continuous surfaces to their discrete approximations within the observation model.

Assumption 1. *For any $\rho > 0$ and $k > 0$, there exist a partition $(\mathcal{S}_{m,k}^\rho)_{m\in\{1,\cdots,M_k^\rho\}}$ of $\mathcal{S}_k$ and a set of diffeomorphisms $(H_{m,k}^\rho)_{m\in\{1,\cdots,M_k^\rho\}}$, each diffeomorphism $H_{m,k}^\rho$ mapping $\mathcal{S}_{m,k}^\rho$ onto $f_{m,k}^\rho$.*

There also exists a function a_1 satisfying $\lim_{\rho\to 0} a_1(\rho) = 0$ such that, for any $\rho > 0$ and $k > 0$,

$$\mathbb{E}\left[\sup_{m\in\{1,\cdots,M_k^\rho\}} (\|H_{m,k}^\rho - I\|_\infty + \|J(H_{m,k}^\rho) - I\|_\infty)^2\right] \leq a_1(\rho),$$

where I denotes the identity on $\mathbb{R}^3$ and $J(H_{m,k}^\rho)$ the Jacobian matrix of $H_{m,k}^\rho$.

Assumption 2. *There exists $C_0 > 0$ such that, for any $k > 0$,*

$$\mathbb{P}\left[\mathcal{A}(\mathcal{S}_k))^2 \leq C_0\right] = 1,$$

where $\mathcal{A}$ denotes the area measure of surfaces.

Theorem 1 (Convergence of the discretization error)**.** *Under Assumptions 1 and 2, one has*

$$\lim_{\rho\to 0} \lim_{K\to+\infty} \sup_{k\in\{1,\cdots,K\}} \sup_{\|w\|_V=1} \mathbb{E}\left[(\epsilon_k^\rho(w))^2\right] = 0, \tag{12}$$

and

$$\lim_{\rho\to 0} \lim_{K\to+\infty} \sup_{\|w\|_V=1} \mathbb{E}\left[\overline{\epsilon_K^\rho(w)}^2\right] = 0. \tag{13}$$

Proof. Let $k > 0$, $\rho > 0$ and $w \in V$ be arbitrary. We set

$$I_k^\rho(w) = \sum_{m=1}^{M_k^\rho} \int_{\mathcal{S}_{m,k}^\rho} |\langle N_k(u) - N_{m,k}^\rho, w(u)\rangle_{\mathbb{R}^3}| d\sigma(u),$$

$$\text{and } J_k^\rho(w) = \sum_{m=1}^{M_k^\rho} |\int_{f_{m,k}^\rho} \langle N_{m,k}^\rho, w(u)\rangle_{\mathbb{R}^3} du - \int_{\mathcal{S}_{m,k}^\rho} \langle N_{m,k}^\rho, w(u)\rangle_{\mathbb{R}^3} d\sigma(u)|.$$

Then, using a Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}\mathbb{E}\left[(\epsilon_k^\rho(w))^2\right] &\leq \mathbb{E}\left[(I_k^\rho(w))^2\right] + \mathbb{E}\left[(J_k^\rho(w))^2\right] + 2\mathbb{E}\left[I_k^\rho(w) J_k^\rho(w)\right], \\ &\leq \mathbb{E}\left[(I_k^\rho(w))^2\right] + \mathbb{E}\left[(J_k^\rho(w))^2\right] + 2\sqrt{\mathbb{E}\left[(I_k^\rho(w))^2\right]\mathbb{E}\left[(J_k^\rho(w))^2\right]}.\end{aligned}$$

On the one hand, since $|w(u)| \leq \|K_0(u,\cdot)\|_{V_0}\|w\|_V$ for any $u \in \mathcal{D}$, we get

$$|I_k^\rho(w)| \leq \|w\|_V \sup_{u\in\mathcal{D}}\|K_0(u,\cdot)\|_{V_0}\mathcal{A}(\mathcal{S}_k) \sup_{m\in\{1,\cdots,M_K^\rho\}} \sup_{u\in\mathcal{S}_k} \|N_k(u) - N_{m,k}^\rho\|_{\mathbb{R}^3}.$$

We now use the following proposition shown in Appendix.

Proposition 1. *Under Assumption 1, there exists a function a_2 satisfying $\lim_{\rho\to 0} a_2(\rho) = 0$ such that, for any $\rho > 0$ and $k > 0$,*

$$\mathbb{E}\left[\sup_{m\in\{1,\cdots,M_k^\rho\}} \sup_{u\in\mathcal{S}_{m,k}^\rho} \|N_k(u) - N_{m,k}^\rho\|_{\mathbb{R}^3}^2\right] \leq a_2(\rho),$$

where $N_k(u)$ and $N_{m,k}^\rho$ denote the normal unit vectors external to the surface $\mathcal{S}_k$ at point u and to the triangle $f_{m,k}^\rho$, respectively.

Due to this proposition, we have

$$\mathbb{E}\left[(I_k^\rho(u))^2\right] \leq \|w\|_V^2 \sup_{u\in\mathbb{D}}\|K_0(u,\cdot)\|_{V_0}^2 \mathbb{E}\left[\mathcal{A}(\mathcal{S}_k)^2\right] a_2(\rho).$$

Hence, using Assumption 2, we obtain

$$\mathbb{E}\left[(I_k^\rho(u))^2\right] \leq C_0\|w\|_V^2 \sup_{u\in\mathcal{D}}\|K_0(u,\cdot)\|_{V_0}^2 a_2(\rho).$$

Consequently, for some $C_1 > 0$,

$$\mathbb{E}\left[(I_k^\rho(w))^2\right] \leq C_1 a_2(\rho)\|w\|_V^2. \tag{14}$$

On the other hand,

$$\begin{aligned}
|J_k^\rho(w)| &\leq \sum_{m=1}^{M_k^\rho} \int_{\mathcal{S}_{m,k}^\rho} |\langle N_{m,k}^\rho, w(u) - \mathrm{Det}(J_{H_{m,k}^\rho}) w(H_{m,k}^\rho(u))\rangle_{\mathbb{R}^3}| d\sigma(u) \\
&\leq \sum_{m=1}^{M_k^\rho} (\int_{\mathcal{S}_{m,k}^\rho} |\langle N_{m,k}^\rho, w(u) - w(H_{m,k}^\rho(u))\rangle_{\mathbb{R}^3}| d\sigma(u) \\
&+ \int_{\mathcal{S}_{m,k}^\rho} |\langle N_{m,k}^\rho, w(H_{m,k}^\rho(u)) - \mathrm{Det}(J_{H_{m,k}^\rho}(u)) w(H_{m,k}^\rho(u))\rangle_{\mathbb{R}^3}| d\sigma(u)) \\
&\leq \sum_{m=1}^{M_k^\rho} (\int_{\mathcal{S}_{m,k}^\rho} \|w(u) - w(H_{m,k}^\rho(u))\|_{\mathbb{R}^3} d\sigma(u) \\
&+ \int_{\mathcal{S}_{m,k}^\rho} \|w(H_{m,k}^\rho(u)) - \mathrm{Det}(J_{H_{m,k}^\rho}(u)) w(H_{m,k}^\rho(u))\|_{\mathbb{R}^3} d\sigma(u)).
\end{aligned}$$

RKHS properties of V imply that

$$|J_k^\rho(w)| \leq C_2\|w\|_V \sum_{m=1}^{M_k^\rho} \int_{\mathcal{S}_{m,k}^\rho} \|K_0(u,\cdot) - K_0(H_{m,k}^\rho(u),\cdot)\|_{V_0} + |1 - \mathrm{Det}(J_{H_{m,k}^\rho}(u))| d\sigma(u),$$

for some $C_2 > 0$. We now apply two propositions shown in Appendix.

Proposition 2. *Under Assumption 1, there exists a function* a_3 *satisfying* $\lim_{\rho\to 0} a_3(\rho) = 0$ *such that, for any* $k > 0$,

$$\mathbb{E}\left[\sup_{m\in\{1,\cdots,M_k^\rho\}} \sup_{u\in\mathcal{S}_{m,k}^\rho} \|K_0(H_{m,k}^\rho(u),\cdot) - K_0(u,\cdot)\|_{V_0}^2\right] \le a_3(\rho).$$

Proposition 3. *Under Assumption 1, there exists a function* a_4 *satisfying* $\lim_{\rho\to 0} a_4(\rho) = 0$ *such that, for any* $k > 0$,

$$\mathbb{E}\left[\sup_{m\in\{1,\cdots,M_k^\rho\}} \sup_{u\in\mathcal{S}_{m,k}^\rho} |1 - \mathrm{Det}(J_{H_m^\rho}(u))|^2\right] \le a_4(\rho).$$

From these propositions and Assumption 2, we conclude that

$$\mathbb{E}\left[(J_k^\rho(w))^2\right] \le C_3\|w\|_V^2(a_3(\rho) + a_4(\rho)), \tag{15}$$

for some $C_3 > 0$. From (14) and (15), we then deduce that

$$\sup_{\|w\|_V\le 1} \mathbb{E}\left[(\epsilon_k^\rho(w))^2\right] \le a(\rho), \tag{16}$$

where $a(\rho) = C_1a_2(\rho) + C_3(a_3(\rho) + a_4(\rho)) + 2\sqrt{C_1C_3a_2(\rho)(a_3(\rho) + a_4(\rho))}$ is a function satisfying $\lim_{\rho\to 0} a(\rho) = 0$. This yields the convergence stated in Equation (12).

Moreover, since ϵ_k^ρ are *i.i.d.*, we have

$$\mathbb{E}\left[\overline{\epsilon_K^\rho(w)}^2\right] \le \frac{1}{K^2}\sum_{k=1}^{K}\mathbb{E}\left[(\epsilon_k^\rho(w))^2\right] + \frac{1}{K^2}\sum_{k\ne l}\mathbb{E}\left[\epsilon_k^\rho\right]\mathbb{E}\left[\epsilon_l^\rho\right].$$

Then, by using a Cauchy-Schwarz inequality, we get

$$\sup_{\|w\|_V\le 1} \mathbb{E}\left[\overline{\epsilon_K^\rho(w)}^2\right] \le \left(\frac{1}{K} + \frac{K-1}{K}\right) \sup_{k\in\{1,\cdots,K\}} \sup_{\|w\|_V\le 1} \mathbb{E}\left[(\epsilon_k^\rho(w))^2\right].$$

Due to Equation (16), it follows that

$$\sup_{\|w\|_V\le 1} \mathbb{E}\left[\overline{\epsilon_K^\rho(w)}^2\right] \le \left(\frac{1}{K} + \frac{K-1}{K}\right) a(\rho).$$

Hence, the convergence stated in Equation (13) holds.

□

We add a last assumption that will be required to establish the consistency of the template estimate.

Assumption 3. *There exists* $\rho_0 > 0$ *such that, for all* $K \in \mathbb{N}^*$, *one has almost surely:*

$$\forall\rho \in]0,\rho_0[, \forall k \in \{1,\cdots,K\}, \forall w \in V \text{ s.t. } \|w\|_V \le 1, |\epsilon_k^\rho(w)| \le 1.$$

3. Template estimation

This section is devoted to the estimation of the template μ of the observation model (11). As the template is defined in an infinite dimensional space, its estimation requires approximations which are presented first.

3.1. Approximations

Let $(\phi_i)_{i\in I}$ be a total and free system of V. Let $(I^N)_{N\in\mathbb{N}^*}$ an increasing sequence of embedded subsets of I, whose union is I. For $N \in \mathbb{N}^*$, we consider subspaces of V

$$V^N = \mathrm{Vect}(\{\phi_i, i \in I^N\}),$$

equipped with the inner product of V. These approximation subspaces are embedded and their reunion is dense in V. We also define subspaces $\widehat{V^{N,*}}$ of V^*

$$\widehat{V^{N,*}} = \{s = \langle v_S, \cdot\rangle, v_S \in V^N\}$$

formed by linear forms s whose Riesz representatives v_s are in V^N. For s in V^*, let

$$s^N = \underset{\nu\in\widehat{V^{N,*}}}{\mathrm{argmin}} \|s-\nu\|_{V^*}$$

be the orthogonal projection of s into $\widehat{V^{N,*}}$. Similarly, for all $v \in V$, let

$$v^N = \underset{w\in\widehat{V^N}}{\mathrm{argmin}} \|v-w\|_V$$

be the orthogonal projection of v into V^N. Notice that $v_{s^N} = v_s^N$.

Each element s of V^* can be approximated by an element of $\widehat{V^{N,*}}$ by using the orthogonal projection v_s^N of its Riesz representative v_s into V^N. This projection v_s^N satisfies

$$\langle v_s^N, w\rangle_V = s(w), \forall w \in V^N,$$

or, equivalently, $v_S^N = \sum_{i\in I^N} \alpha_i\phi_i$, where $(\alpha_i)_{i\in I^N}$ is the solution of the linear system

$$\sum_{j\in I^N} \alpha_i\langle\phi_i, \phi_j\rangle_V = \sum_{j\in I^N} s(\phi_j), \forall\, i \in I^N.$$

For any $i, j \in I^N$, let $h_{i,j}^N$ be the solution of

$$\sum_{k\in I^N} \langle\phi_i, \phi_j\rangle_V h_{k,j}^N = \delta_{i,j}. \tag{17}$$

We can also write

$$v_s^N = \sum_{i\in I^N} \left(\sum_{j\in I^N} h_{i,j}^N s(\phi_j) \right) \phi_i. \tag{18}$$

3.2. Estimates

Let us consider the observation model (11), and construct random linear forms that estimate projections of the template μ in successive approximation spaces $V^{N,*}$.

For $k \in \{1, \cdots, K\}$ and $i \in I$, let

$$\mathbf{S}^{\rho}_{k,i} = S^{\rho}_k(\phi_i), \;\; \mu_i = \mu(\phi_i), \;\text{ and }\; \eta^{\rho}_{k,i} = B_k(\phi_i) + \epsilon^{\rho}_k(\phi_i) \tag{19}$$

denote variables specifying terms of the observation model (11) for a basis element ϕ_i of V. Expressed in the basis $\{\phi_i, i \in I^N\}$ of V^N, the observation model becomes

$$\mathbf{S}^{N,\rho}_k = \mu^N + \eta^{N,\rho}_k, \forall k \in \{1, \cdots, K\}, \tag{20}$$

where $\mu^N = (\mu_i)_{i \in I^N}$ is a vector of size $|I^N|$, and $\mathbf{S}^{N,\rho}_k = (S^{\rho}_{k,i})_{i \in I^N}$ and $\eta^{N,\rho}_k = (\eta^{\rho}_{k,i})_{i \in I^N}$ are two Gaussian random vectors with a covariance matrix $\Gamma^{N,\rho}$ of size $|I^N| \times |I^N|$ whose terms are given by

$$\Gamma^{N,\rho}_{i,j} = \mathrm{Cov}(\mathbf{S}^N_{k,i}, \mathbf{S}^N_{k,j}) = \gamma^B(\phi_i, \phi_j) + \gamma^{\rho}_{\epsilon}(\phi_i, \phi_j).$$

Due to model assumptions, the expectation of $\mathbf{S}^{N,\rho}_k$ is μ^N. Within Model (20), its maximum-likelihood estimate is given by the empirical mean

$$\overline{\mathbf{S}^{N,\rho}_K} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{S}^{N,\rho}_k.$$

Now, let v^N_{μ} be the orthogonal projection of the Riesz representative v_{μ} of the template μ into V^N, and $\mu^N = \langle v^N_{\mu}, \cdot \rangle_V$ its corresponding linear form. According to Section 3.1, the projection v^N_{μ} may be estimated by

$$\widehat{v^{N,\rho}_K} = \sum_{m \in I^N} \left(\sum_{n \in I^N} h^N_{m,n} \overline{\mathbf{S}^{N,\rho}_{K,n}} \right) \phi_m, \tag{21}$$

where terms $\overline{\mathbf{S}^{N,\rho}_{K,n}}$ are components of the random vector $\overline{\mathbf{S}^{N,\rho}_K}$, and $h^N_{i,j}$ solutions of Equation (17). The linear form μ^N may be accordingly estimated by

$$\widehat{\mu^{N,\rho}_K}(w) = \langle \widehat{v^{N,\rho}_K}, w \rangle_V, \forall w \in V. \tag{22}$$

3.3. Consistency

Next, we establish the consistency of template estimates built in Section 3.2.

Theorem 2 (Consistency of template estimates)**.** *Under Assumptions 1, 2 and 3, the random linear forms $\widehat{\mu^{N,\rho}_k}$ converge to the template μ in the space $(V^*_{\mathcal{E}}, \|.\|_{V^*_{\mathcal{E}}})$ as N and K tend to $+\infty$ and ρ tends to 0, i.e.*

$$\lim_{\substack{\rho \to 0 \\ K,N \to \infty}} \sup_{\|w\|_V \leq 1} \sqrt{\mathbb{E}\left[|\widehat{\mu^{N,\rho}_k}(w) - \mu(w)|^2 \right]}.$$

Proof. Let us first introduce

$$\widetilde{v_K^N} = \sum_{m \in I^N} \left(\sum_{n \in I^N} h_{m,n}^N \overline{\mathbf{S}_{K,n}} \right) \phi_m, \tag{23}$$

where $\overline{\mathbf{S}_{K,n}}$ is the average over the sample of continuous surface representatives at ϕ_n

$$\overline{\mathbf{S}_{K,n}} = \frac{1}{K} \sum_{k=1}^{K} S_k(\phi_n),$$

and $h_{i,j}^N$ are solutions of Equation (18). This random element $\widetilde{v_K^N}$ would be an estimate of the Riesz representative v_μ of the template μ if continuous surface representatives S_k were directly observed. We also define the random linear form

$$\widetilde{\mu_K^N}(w) = \langle \widetilde{v_K^N}, w \rangle_V, \forall w \in V. \tag{24}$$

Using a triangular inequality, we obtain

$$\|\widehat{\mu_K^{N,\rho}} - \mu\|_{V_\mathcal{E}^*} \leq \|\widehat{\mu_K^{N,\rho}} - \widetilde{\mu_K^N}\|_{V_\mathcal{E}^*} + \|\widetilde{\mu_K^N} - \mu\|_{V_\mathcal{E}^*}. \tag{25}$$

We first show the convergence of the first term of the right-hand side of this inequality. We notice that

$$\|\widehat{\mu_K^{N,\rho}} - \widetilde{\mu_K^N}\|_{V_\mathcal{E}^*}^2 = \sup_{\|w\|_V \leq 1} \mathbb{E}\left[|\langle \widehat{v_K^{N,\rho}} - \widetilde{v_K^N}, w \rangle_V|^2 \right] = \sup_{\|w\|_V \leq 1} \mathbb{E}\left[|\langle \widehat{v_K^{N,\rho}} - \widetilde{v_K^N}, w^N \rangle_V|^2 \right].$$

Setting $w^N = \sum_{i \in I^N} \alpha_i^N \phi_i$, we have

$$\begin{aligned}
&\|\widehat{\mu_K^{N,\rho}} - \widetilde{\mu_K^N}\|_{V_\mathcal{E}^*}^2 \\
&= \sup_{\|w\|_V \leq 1} \sum_{s \in I^N} \sum_{t \in I^N} \alpha_s^N \alpha_t^N \sum_{j \in I^N} \sum_{m \in I^N} \mathbb{E}[\overline{\epsilon_K^\rho(\phi_m)}\, \overline{\epsilon_K^\rho(\phi_j)}] \sum_{i \in I^N} \langle \phi_s, \phi_i \rangle_V h_{i,j}^N \sum_{l \in I^N} \langle \phi_t, \phi_l \rangle_V h_{l,m}^N \\
&= \sup_{\|w\|_V \leq 1} \sum_{s \in I^N} \sum_{t \in I^N} \alpha_s^N \alpha_t^N \mathbb{E}[\overline{\epsilon_K^\rho(\phi_s)}\, \overline{\epsilon_K^\rho(\phi_t)}] = \sup_{\|w\|_V \leq 1} \mathbb{E}\left[\left(\overline{\epsilon_K^\rho(w^N)} \right)^2 \right].
\end{aligned}$$

Therefore,

$$\|\widehat{\mu_K^{N,\rho}} - \widetilde{\mu_K^N}\|_{V_\mathcal{E}^*}^2 \leq \sup_{\|w\|_V \leq 1} \mathbb{E}\left[\left(\overline{\epsilon_K^\rho(w)} \right)^2 \right].$$

Hence, the convergence of $\|\widehat{\mu_K^{N,\rho}} - \widetilde{\mu_K^N}\|_{V_\mathcal{E}^*}^2$ to 0 follows from Theorem 1.

We now show the convergence of the second term of the right-hand side of Equation (25). Using a triangular inequality, we have

$$\|\widetilde{\mu_K^N} - \mu\|_{V_\mathcal{E}^*} \leq \|\widetilde{\mu_K^N} - \mu^N\|_{V_\mathcal{E}^*} + \|\mu^N - \mu\|_V,$$

where μ^N is the orthogonal projection of μ into $\widehat{V^{N,*}}$. By definition of the projection, the second term of the right-hand side of the inequality converges to 0 as N tends to $+\infty$. Hence, it suffices to show the convergence of the first term.

Let be $K \in \mathbb{N}^*$. We have

$$\|\widetilde{\mu_K^N} - \mu^N\|_{V_{\mathcal{E}}^*} = \sqrt{\sup_{\|w\|_V \le 1} \mathbb{E}\left[(\widetilde{\mu_K^N}(w) - \mu^N(w))^2\right]} = \sqrt{\sup_{\|w\|_V \le 1} \mathrm{Var}\left(\widetilde{\mu_K^N}(w)\right)},$$

since $\mu^N(w) = \mathbb{E}\left[\widetilde{\mu_K^N}(w)\right], \forall w \in V$. Yet, as the sequence $(S_k), 1 \le k \le K$ is *i.i.d.*,

$$\mathrm{Var}\left(\widetilde{\mu_K^N}(w)\right) = \frac{1}{K}\mathrm{Var}\left(\widetilde{\mu_1^N}(w)\right) = \frac{1}{K}\mathbb{E}\left[(\widetilde{\mu_1^N}(w) - \mu^N(w))^2\right].$$

We now use the following proposition proved in Appendix.

Proposition 4. *For any $N > 0$,*

$$\sup_{\|w\|_V \le 1} \mathbb{E}\left[(\widetilde{\mu_1^N}(w) - \mu^N(w))^2\right] \le \|\gamma^B\|^2_{V^{**}},$$

where $\|\gamma^B\|^2_{V^{**}}$ is the norm of the bilinear form γ^B defined by

$$\|\gamma^B\|_{V^{**}} = \sup_{w \in V, |v|_V = 1} |\gamma^B(w,w)|.$$

Due to Proposition 4, we have

$$\|\widetilde{\mu_K^N} - \mu^N\|_{V_{\mathcal{E}}^*} = \sqrt{\frac{1}{K} \sup_{\|w\|_V \le 1} \mathbb{E}\left[(\widetilde{\mu_1^N}(w) - \mu^N(w))^2\right]} \le \frac{\|\gamma^B\|_{V^{**}}}{\sqrt{K}}.$$

Hence, $\|\widetilde{\mu_K^N} - \mu^N\|_{V_{\mathcal{E}}^*}$ converges to 0 as K tends to $+\infty$. This concludes the proof. □

4. Discussion

In the continuation of the work of Glaunès and Vaillant [23], we proposed a stochastic framework where surfaces are represented by some random linear forms defined on a Hilbert space of mappings from $\mathbb{R}^3$ into itself. We further constructed an observation model where observed surfaces are viewed as variations of a same template representating an average of surfaces over an homogeneous population. These variations accounted not only for a population variability but also for discretization errors due to acquisition. Using the observation model, we presented a methodology to estimate the template from a population sample. This methodology is based on an original procedure to approximate random linear forms in finite-dimensional subspaces. We eventually established a consistency result ensuring that the template can be estimated as closely as expected if the number of observed samples is sufficient and surface discretizations and approximations are accurate enough.

The approach we use to tackle the template construction departs from the ones proposed in [12, 21]. Rather than being a simple algorithmic solution to that issue, it is anchored in a stochastic framework which has a Bayesian interpretation. Furthermore, the method used to approximate the template is different. In [12, 21], the template is expanded in a basis of kernels centered at vertices of surface meshes. Hence, it depends on the discretization of observed surfaces. In particular, the template representation may become too large when observed surfaces are numerous or discretized in an accurate way. Moreover, kernels depend on a scale parameter that restrict scales at which surfaces can be analyzed. By contrast, our approximation is based on a functional basis which is independent on surface meshes and can be easily specified in a multiscale fashion.

The observation model we used did not account for possible geometric deformations of the template. When such deformations are included in the model, maximum-likelihood estimates of the template are not explicit anymore. However, as shown for images in [2, 1, 20], they can be approached using an EM-algorithm where likelihood terms are iteratively approximated with a stochastic procedure. In a future work, the present study could serve as a background to design such an EM-algorithm for surfaces and analyze its convergence.

In this paper, the estimation was focused on the template while the rest of the model was assumed to be known. Hence, we plan to extend the estimation to the autocovariance of model variations. In our framework, such an estimation could be obtained using a same maximum-likelihood approach as the one used for the template. It would enable to fully adapt the model to observed data. It would also path the way to the analysis of variation modes of a surface population.

References

[1] S. Allassonnière, Y. Amit, and A. Trouvé. Towards a coherent statistical framework for dense deformable template estimation. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 69(1):3–29, 2007.

[2] S. Allassonnière, J. Bigot, Glaunès, et al. Statistical models for deformable templates in image and shape analysis. *Ann. Math. Blaise Pascal*, 20(1):1–35, 2013.

[3] S. Allassonnière, E. Kuhn, A. Trouvé, et al. Construction of bayesian deformable models via a stochastic approximation algorithm: a convergence study. *Bernoulli*, 16(3):641–678, 2010.

[4] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68(3):337–404, 1950.

[5] P. Ciarlet. The finite element method for elliptic problems. *Classics in applied mathematics*, 40:1–511, 2002.

[6] C. Cury, J. Glaunes, and O. Colliot. Template estimation for large database: a diffeomorphic iterative centroid method using currents. In *Geometric Science of Information*, pages 103–111. Springer, 2013.

[7] S. Durrleman, P. Fillard, X. Pennec, et al. Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents. *Neuroimage*, 55(3):1073–109, 2009.

[8] S. Durrleman, X. Pennec, A. Trouvé, and N. Ayache. Statistical models of sets of curves and surfaces based on currents. *Med. Image Anal.*, 13(5):793–808, 2008.

[9] X. Fernique. Processus linéaires, processus généralisés. In *Ann. Inst. Fourier*, volume 17, pages 1–92, 1967.

[10] I. Gelfand and N. Vilenkin. *Generalized Functions*, volume 4. Academic Press, 1964.

[11] J. Glaunès. *Transport par Difféomorphismes de Points, de Mesures et de Courants pour la Comparaison des Formes et l'Anatomie Numérique*. PhD thesis, University Paris 13, 2005.

[12] J. Glaunès and S. Joshi. Template estimation from unlabeled point set data and surfaces for computational anatomy. In X. Pennec and S. Joshi, editors, *Proc. of the first MICCAI Workshop on the Mathematical Foundations of Computational Anatomy*, pages 29–39, 2006.

[13] P. Gori, 0. Colliot, Y. Worbe, et al. Bayesian atlas estimation for the variability analysis of shape complexes. In *MICCAI*, pages 267–274, 2013.

[14] P. Gori, O. Colliot, L. Marrakchi-Kacem, et al. Parsimonious approximation of streamline trajectories in white matter fiber bundles. *IEEE Trans. Med. Imaging*, 35(12):2609–2619, 2016.

[15] U. Grenander and M Miller. Computational anatomy: An emerging discipline. *Quart. App. Math.*, 56(4):617–694, 1998.

[16] K. Itô. Stationary random distributions. *Memoirs of the College of Science, University of Kyoto. Series A : Mathematics*, 28(3):209–223, 1954.

[17] K. Itô. Isotropic random current. In *Proc. of the third Berkeley Symposium on Mathematical Statistics and Probability*, pages 125–13, 1956.

[18] R. Meidan. On the connection between ordinary and generalized stochastic processes. *J. Math. Anal. Appl.*, 76(1):124–133, 1980.

[19] M. Miller and L. Younès. Bayesian template estimation in computational anatomy. *Neuroimage*, 42(1):252–261, 2008.

[20] F. Richard, A. Samson, and C.-A. Cuenod. A SAEM algorithm for the estimation of template and deformation parameters in medical image sequences. *Statistics and Computing*, 19(4):465–478, 2009.

[21] F. Tilotta, J. Glaunès, F. Richard, and Y. Rozenholc. A local technique based on vectorized surfaces for craniofacial reconstruction. *Forensic Sci. Int.*, 200(1):50–59, 2010.

[22] A. Trouvé and L. Younes. Local geometry of deformable templates. *SIAM J. Math. Anal.*, 37(1):17–59, 2005.

[23] M. Vaillant and J. Glaunès. Surface matching via currents. In G. Christensen and M. Sonka, editors, *Information Processing in Medical Imaging*, volume 3565. Springer Berlin / Heidelberg, 2005.

[24] A. Yaglom. *Correlation Theory of Stationary and Related Random Functions: Supplementary Notes and References.* Springer Science & Business Media, 2012.
[25] L. Younès. *Shapes and diffeomorphisms*, volume 171. Springer Science & Business Media, 2010.

Appendix A: Proof of supporting propositions

Proof of Proposition 1. We have the identity

$$\|N_k(u) - N^{\rho}_{m,k}\|_{\mathbb{R}^3} = \|N^{\rho}_{m,k}(J(H^{\rho}_{m,k}) - I)\|_{\mathbb{R}^3}.$$

This yields

$$\|N_k(u) - N^{\rho}_{m,k}\|_{\mathbb{R}^3} \leq \|J(H^{\rho}_{m,k}) - I\|_{\mathcal{L}(\mathbb{R}^3,\mathbb{R}^3)}.$$

We conclude the proof using Assumption 1. □

Proof of Proposition 2. This results from the uniform continuity of the function $y \to K_0(y, \cdot)$ on the compact $\mathcal{D}$ and Assumption 1. □

Proof of Proposition 3. Let P be a polynomial of order q defined on $\mathbb{R}^p$, and $\mathcal{K}$ a compact subset of $\mathbb{R}^p$. Then, there exists $C > 0$ such that, for all $x \in \mathcal{K}$, $\epsilon > 0$ with $x + \epsilon \in \mathcal{K}$,

$$|P(x+\epsilon) - P(x)| \leq C \sum_{k=1}^{q-1} \|\epsilon\|^k.$$

We get the result by applying this inequality to the determinant $\mathrm{Det}(J_{H^{\rho}_m}(u))$, which is a polynomial function of matrix coefficients. □

Proof of Proposition 4. Let w be in V and $w^N = \sum_{i \in I^N} \alpha_i^N \phi_i$ its orthogonal projection on V^N. One has

$$\mathbb{E}\left[\left(\widetilde{\mu_1^N}(w) - \mu^N(w)\right)^2\right] = \sum_{m \in I^N} \sum_{n \in I^N} \alpha_m^N \alpha_n^N C_{m,n},$$

where

$$C_{m,n} = \mathbb{E}\left[\left(\widetilde{\mu_1^N}(\phi_m) - \mu^N(\phi_m)\right)\left(\widetilde{\mu_1^N}(\phi_n) - \mu^N(\phi_n)\right)\right].$$

According to Model (10), $C_{m,n} = \mathbb{E}[B_1^N(\phi_m) B_1^N(\phi_n)]$, where B_1^N is the orthogonal projection of B_1 into V^N. Using the Riesz representative $\sum_{s \in I^N} \sum_{t \in I^N} h^N_{s,t} B_1(\phi_t)\phi_s$ of B_1^N, we thus obtain

$$C_{m,n} = \sum_{i \in I^N} \sum_{j \in I^N} \sum_{s \in I^N} \sum_{t \in I^N} \left(h^N_{i,j} h^N_{s,t} \langle \phi_i, \phi_m \rangle_V \langle \phi_s, \phi_n \rangle_V \mathbb{E}\left[B_1(\phi_j) B_1(\phi_t)\right]\right.$$

Since $\sum_{i \in I^N} h_{i,j}^N \langle \phi_i, \phi_m \rangle_V = \delta_{j,m}$, it follows that

$$C_{m,n} = \sum_{j \in I^N} \sum_{t \in I^N} \gamma_{j,t}^B \delta_{j,m} \delta_{t,n} = \gamma_{m,n}^B.$$

Therefore,

$$\mathbb{E}\left[\left(\widetilde{\mu_1^N}(w) - \mu^N(w)\right)^2\right] = \sum_{i \in I^N} \sum_{j \in I^N} \alpha_i^N \alpha_j^N \gamma_{i,j}^B.$$

Hence,

$$\mathbb{E}\left[\left(\widetilde{\mu_1^N}(w) - \mu^N(w)\right)^2\right] \leq \|w\|_V^2 \|\gamma^B\|_{V^{**}}.$$

$\square$