Research Article

Christophe Denis*

Classification in Postural Style Based on Stochastic Process Modeling

Abstract: We address the statistical challenge of classifying subjects as hemiplegic, vestibular or normal based on complex trajectories obtained through two experimental protocols designed to evaluate potential deficits in postural control. The classification procedure involves a dimension reduction step where the complex trajectories are summarized by finite-dimensional summary measures based on a stochastic process model for a real-valued trajectory. This allows us to retrieve from the trajectories information relative to their temporal dynamic. A leave-one-out evaluation yields a 79% performance of correct classification for a total of $n = 70$ subjects, with 22 hemiplegic (31%), 16 vestibular (23%) and 32 normal (46%) subjects.

Keywords: change point estimation; multiclass classification; cross-validation; postural maintenance; stochastic process modeling

DOI 10.1515/ijb-2012-0017

1 Introduction

This article contributes to the study of postural maintenance. Posture is fundamental for physical activity. A deficit in postural maintenance often results in falling, which is particularly hazardous in elderly people. The main objective of the research in postural maintenance is to adapt protocols for functional rehabilitation for people who display deficits in maintaining posture. We focus here on the issue of classifying subjects in terms of postural maintenance. A cohort of 70 subjects has been followed at the center for the study of sensorimotor functioning (CESEM) of Université Paris Descartes. The subjects who did not exhibit deficit in postural maintenance were labeled as normal. The others were hemiplegic and vestibular subjects and labeled accordingly. A hemiplegic subject suffers from a disorder of his proprioceptive system, which pertains to the sense of position, location, orientation of the body and its parts. A vestibular subject suffers from a deficit of his vestibular system, which is the system composed by the inner ear and the vestibular nerve and contributes to the sense of balance.

Each subject completed two experimental protocols designed to evaluate his/her potential deficits in maintaining posture. Each protocol is divided into three phases: a first phase of 15 s with no postural perturbation, a second phase of 35 s with postural perturbation followed by a last phase of 20 s without postural perturbation. During each protocol, measurements of the center-of-pressure of each foot are performed at discrete times, which results in temporal trajectories. Our objective is to classify subjects as hemiplegic, vestibular and normal based on these trajectories. This is a significantly more difficult extension of the problem considered by Chambaz and Denis [1], where only hemiplegic and normal subjects were to be classified (particularly because classifying into three classes is more difficult than classifying into two classes), and thus another step in the direction of clustering subjects in terms of their postural style. We refer to the bibliography in Chambaz and Denis [1] for a review on the analysis of postural control.

***Corresponding author: Christophe Denis,** Université Paris-Est Marne-la-Vallée, France, E-mail: Christophe.Denis@univ-mlv.fr

Our present study is related to the topic of functional data classification. There is a sizeable literature dedicated to this topic. A variety of methods have been proposed, relying for instance on linear discriminant analysis [2], principal component analysis [3], or a functional data version of the nearest-neighbors classification rule [4]. We refer to Ramsay and Silverman [5] for a general introduction to functional data analysis. Our approach here involves a dimension reduction step of our complex trajectories based on change points estimation and inference on a stochastic process model. Although the various statistical methods involved in the dimension reduction step were known, their combination for the sake of classifying in terms of postural style (three classes) is very orignal and performs quite well. This allows us to better use the data at hand than in Chambaz and Denis [1] in the sense that we manage to retrieve from the trajectories information relative to their temporal dynamic. In contrast, Chambaz and Denis [1] retrieve static information from the trajectories in the sense that the dimension reduction step relies on comparisons of basic statistics (such as the mean value of a segment of the trajectory, see Section 5.2) computed on arbitrarily chosen time intervals starting or ending where the perturbation phase starts or ends.

The dataset and its modeling are introduced in Section 2. We present our inference procedure in Section 3. We carry it out on real and simulated data, and summarize the results in the latter section. The classification procedure is presented in Section 4. The results of its application to the real dataset are exposed and commented on in Section 5, which concludes on some perspectives for future research.

2 Data and modeling

The dataset at hand is described in Section 2.1. We introduce and motivate its modeling in Section 2.2.

2.1 Original dataset

The dataset was collected at the center for the study of sensorimotor functioning (CESEM) of Université Paris Descartes. It is composed of a cohort of $n = 70$ subjects. Among the 70 subjects, 22 are *hemiplegic* (due to cerebrovascular accidents), 16 are *vestibular*, while the 32 remaining subjects are identified as *normal* based on an initial medical evaluation.

Each subject completed two protocols designed to evaluate potential deficits in maintaining posture. These protocols have been identified as the most informative among four similar protocols for classifying hemiplegic versus normal subjects in the earlier study [1]. Both protocols are divided into three phases: a first phase (lasting 15 s) with no postural perturbation, a second phase (lasting 35 s) with postural perturbations (either some muscular perturbations or a combination of muscular and visual perturbations), and a third phase (lasting 20 s) with no postural perturbation. We expect that the subject's postural sway changes around the beginning and the end of the perturbation phase (around 15 and 50 s).

For each protocol, the center-of-maximal-pressure exerted by each foot on a force-platform is recorded at equispaced discrete times. Thus each protocol results in a trajectory $(L_i, R_i)_{i \le N}$, where (L_i, R_i) is the observation at time $t_i = i\delta$ for $i = 1, \dots, N = 2,800$ and $\delta = 0.025$ s. For each $i = 1, \dots, N$, $L_i = \left(L_i^1, L_i^2\right) \in \mathbb{R}^2$ and $R_i = \left(R_i^1, R_i^2\right) \in \mathbb{R}^2$ respectively correspond to the left and right foot (Table 1).

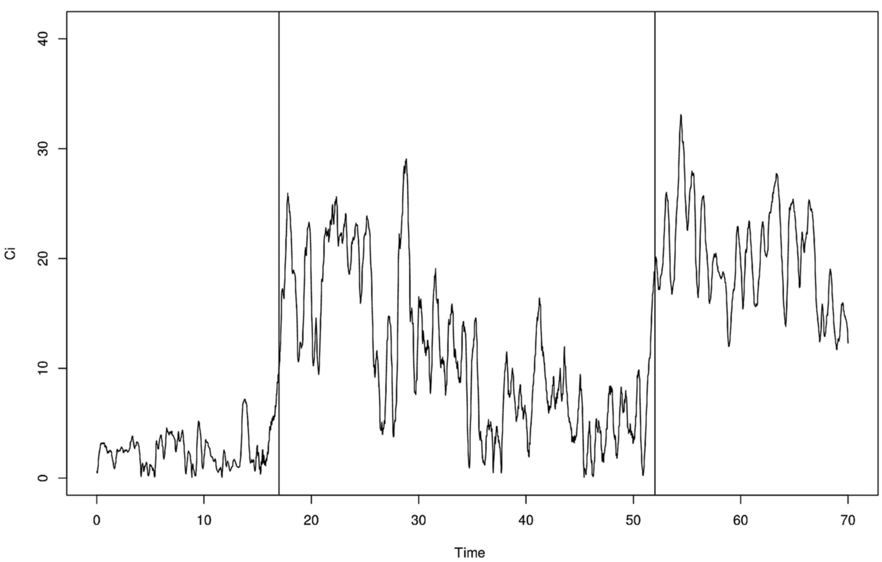
Following Chambaz and Denis [1], we derive from $(L_i, R_i)_{i \le N}$ the one-dimensional trajectory $C_{1:N}$ (sometimes we will denote a n-tuple $(x_1, \dots, x_n)$ by $x_{1:n}$) characterized by

$$C_i = \left\| \frac{L_i + R_i}{2} - \gamma \right\|_2,$$

where γ is defined as the component-wise median value of $(\frac{1}{2}(L_i + R_i))_{i \le 400}$ over the ten first seconds of the protocol. The process $C_{1:N}$ provides a relevant description of the sway of the body during the course of the protocol. In Figure 1, we display the trajectories $C_{1:N}$ corresponding to protocol 1 (left plot) and protocol 2

Table 1 Specifics of the two protocols considered in this study. A protocol is divided into three phases: a first phase with no postural perturbation is followed by a second phase with perturbations, which is itself followed by a third phase without perturbation. Different types of perturbations are considered. In protocol 1, the brain's processing of proprioceptive information (through muscular stimulation) is perturbed, whereas in protocol 2, the processing by the brain of both visual information (subjects must close their eyes) and proprioceptive information (through muscular stimulation) is perturbed

Protocol	1st phase (0→15 s)	2nd phase (15→50 s)	3rd phase (50→70 s)
1	No perturbation	Muscular stimulation	No perturbation
2		Eyes closed Muscular stimulation	



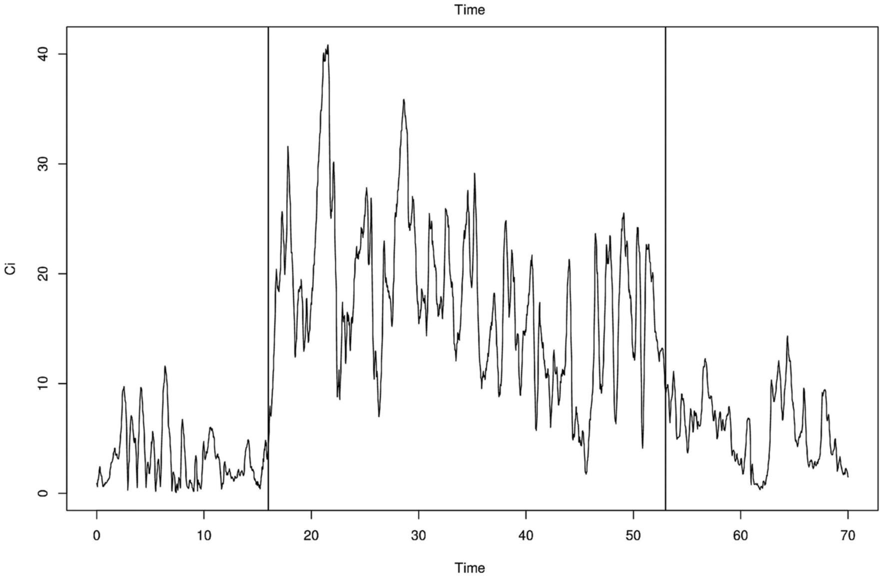


Figure 1 Two trajectories $C_{1:N}$ respectively corresponding to protocol 1 (left) and protocol 2 (right) undergone by a single hemiplegic subject. Vertical lines correspond to the estimated change points $(\hat{\tau}_1\delta, \hat{\tau}_2\delta)$ as obtained based on the two (real) observed trajectories

(right plot) as completed by a single hemiplegic subject. Figure 1 confirms the intuition that the subject's postural sway is not necessarily instantaneously affected by the start and end of perturbations.

2.2 Data modeling

We model the trajectory $C_{1:N}$ as the observation at discrete times of a stochastic process $(C(t))_{t\in[\delta,N\delta]}$ characterized by a stochastic differential equation.

We model the effects of the perturbations by possible changes in the volatility and drift functions at two unknown change points. In order to account for the fact that subjects react differently, we also assume that the change points differ among subjects. We denote by (T_1, T_2) the change points of a trajectory, with $T_0 = \tau_0\delta = \delta < T_1 = \tau_1\delta < T_2 = \tau_2\delta < T_3 = 70$ ($\tau_0 = 1$, τ_1 and τ_2 are integers). We assume that there exist two functions a, b, and two parameters (ϕ_1, ϕ_2, ϕ_3) and $(\sigma_1, \sigma_2, \sigma_3)$ such that, for all $t \in [T_{k-1}, T_k)$ and $k = 1, 2, 3$,

$$\begin{cases} \mathrm{d}C(t) & = a(C(t), \phi_k)\mathrm{d}t + b(C(t), \sigma_k)\mathrm{d}W(t) \\ C(T_{k-1}) = C_{\tau_{k-1}}, \end{cases}$$

where $(W(t))_{t\in[\delta,N\delta]}$ is a standard Wiener process.

We now specify the parametric forms of the volatility and drift functions a and b. Because Figure 1 indicates that the variance of $C_{1:N}$ is not constant over time and because, in addition, the process $C_{1:N}$ takes positive values, we decide to rely on the classical Cox–Ingersoll–Ross (CIR) process [6], by setting $a(x, \phi = (\lambda, \mu)) = \lambda(\mu - x)$ and $b(x, \sigma) = \sigma\sqrt{x}$ (for all $x \geq 0$).

In summary, we model the trajectory $C_{1:N}$ as $C_i = C(i\delta) = C(t_i)$ the observation at discrete times of the stochastic process $(C(t))_{t\in[\delta,N\delta]}$ being characterized, for $t \in [T_{k-1}, T_k)(k = 1, 2, 3)$, by

$$\begin{cases} \mathrm{d}C(t) & = \lambda_k(\mu_k - C(t))\mathrm{d}t + \sigma_k\sqrt{C(t)}\mathrm{d}W(t) \\ C(T_{k-1}) = C_{\tau_{k-1}}, \end{cases} \tag{1}$$

where the initial condition $C(T_{k-1})$ is the observation $C_{\tau_{k-1}}$ at time T_{k-1}, λ_k, and μ_k and σ_k are positive constants. It is known that if $2\mu_k\lambda_k/\sigma_k^2 \geq 1$ then the process remains positive and admits a stationary distribution, namely a Gamma distribution with shape parameter $2\mu_k\lambda_k/\sigma_k^2$ and scale parameter $\sigma_k^2/2\lambda_k$ [6]. We summarize all parameters as a single vector $(\tau_1, \tau_2, \theta_1, \theta_2, \theta_3)$ with $\theta_k = (\lambda_k, \mu_k, \sigma_k)$ for $k = 1, 2, 3$. *We assume that the changes, induced by the perturbation phase, necessarily affect the drift parameter through a change in μ.* In other words, we assume that $\mu_1 \neq \mu_2$ and $\mu_2 \neq \mu_3$. Changes can also affect the other parameters.

3 Inference

In this section, we address inference about the parameter $(\tau_1, \tau_2, \theta_1, \theta_2, \theta_3)$. We first describe the estimating procedure in Section 3.1, then present results from its application on the motivating dataset in Section 3.2. Finally, we summarize a simulation study performed to evaluate its performance in Section 3.3.

3.1 A two-step estimating procedure

We define a two-step estimating procedure: we first estimate the sequence of change points, and then estimate $(\theta_1, \theta_2, \theta_3)$ on each interval characterized by the previously estimated change points.

A great variety of methods have been proposed for the estimation of change points (among many others [7–9–11]. We choose to adopt a popular approach originally proposed by Bai [7]. The approach relies on a least-squares criterion and aims at detecting change points which affect the mean of a linear process. The estimator $(\hat{\tau}_1, \hat{\tau}_2)$ of the sequence of change points (τ_1, τ_2) is defined as:

$$(\hat{\tau}_1, \hat{\tau}_2) = \underset{(\tau_1,\tau_2)}{\arg\min} \frac{1}{N} \sum_{k=1}^{3} \sum_{i=\tau_{k-1}}^{\tau_k - 1} \left(C_i - \bar{C}_k\right)^2,$$

where $\bar{C}_k$ is the arithmetic mean of $C_{\tau_{k-1}:\tau_k-1}$. This makes sense because if $(C(t))_{t\in[T_{k-1},T_k)}$ reaches a stationary regime then the random variables $C_{\tau_k}, \ldots, C_{\tau_k-1}$ are identically distributed from a stationary distribution whose mean parameter is μ_k.

Fix $k \in \{1, 2, 3\}$. On each interval $[\hat{T}_{k-1}, \hat{T}_k) = [\hat{\tau}_1\delta, \hat{\tau}_2\delta)$, we estimate θ_k by minimizing a contrast function based on the log-likelihood of the approximated discrete-time Euler–Maruyama scheme (see Kessler [12], for instance). The latter scheme with step size δ guarantees that, for $i = \hat{\tau}_{k-1}, \ldots, \hat{\tau}_k - 1$,

$$C_{i+1} \approx (1 - \delta\lambda_k)C_i + \delta\lambda_k\mu_k + \sigma_k\sqrt{\delta}\sqrt{C_i}\,\eta_{i+1}, \tag{2}$$

where $(\eta_i)_{\hat{\tau}_{k-1} < i \leq \hat{\tau}_k}$ is a sequence of independent random variables with standard Gaussian distribution. By eq. (2), it is convenient to consider the following equivalent parametrization: $\theta_k = (\theta_{1,k}, \theta_{2,k}, \theta_{3,k})$, with $\theta_{1,k} = (1 - \delta\lambda_k)$, $\theta_{2,k} = \delta\lambda_k\mu_k$ and $\theta_{3,k} = \sigma_k\sqrt{\delta}$. The estimator $\hat{\theta}_k$ of parameter θ_k is defined as:

$$\hat{\theta}_k = \underset{\theta_k\in\mathbb{R}_+^3}{\arg\min} \sum_{i=\hat{\tau}_{k-1}}^{\hat{\tau}_k - 1} \frac{\left(C_{i+1} - \theta_{1,k}C_i - \theta_{2,k}\right)^2}{C_i\theta_{3,k}^2} + (\hat{\tau}_k - \hat{\tau}_{k-1})\log(\theta_{3,k}^2),$$

which actually yields closed-form expressions:

$$\hat{\theta}_{1,k} = \frac{(\hat{\tau}_k - \hat{\tau}_{k-1})\sum\frac{C_{i+1}}{C_i} - \sum C_{i+1}\sum\frac{1}{C_i}}{(\hat{\tau}_k - \hat{\tau}_{k-1})^2 - \sum C_i\sum\frac{1}{C_i}},$$

$$\hat{\theta}_{2,k} = \frac{\sum C_{i+1} - \hat{\theta}_{1,k}\sum C_i}{(\hat{\tau}_k - \hat{\tau}_{k-1})},$$

$$\hat{\theta}_{3,k} = \sqrt{\frac{\sum\frac{(C_{i+1} - \hat{\theta}_{1,k}C_i - \hat{\theta}_{2,k})^2)}{C_i}}{\hat{\tau}_k - \hat{\tau}_{k-1}}}$$

(the sums in the above expressions range over $[\hat{\tau}_{k-1}, \hat{\tau}_k - 1]$).

Under mild conditions, and if the process reaches the stationary regime, then $(\hat{\tau}_1, \hat{\tau}_2)$ consistently estimates (τ_1, τ_2) (see Lavielle [13], Lavielle and Ludeña [14], for instance). Furthermore, if the true change points (τ_1, τ_2) are known then, under another set of mild conditions, the estimators $(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$ consistently estimate $(\theta_1, \theta_2, \theta_3)$ (see Kessler [12], Theorem 1). To the best of our knowledge, there is no satisfactory result in the literature regarding the joint estimation of the change points (τ_1, τ_2) and the parameter $(\theta_1, \theta_2, \theta_3)$.

3.2 Application to the real dataset

We undertake a simulation study of the properties of the two-step estimating procedure presented in the previous section and summarize its results in the next section. Because we characterize our simulation scheme based on the results obtained when applying the latter procedure to the real dataset, we first present them here.

For each subject and protocol, we estimate $(\tau_1, \tau_2, \theta_1, \theta_2, \theta_3)$ from the corresponding (real) observed trajectory. The results pertaining to the estimation of (τ_1, τ_2) are summarized in Table 2 (the mean and standard deviation of the estimates $(\hat{\tau}_1\delta, \hat{\tau}_2\delta)$ computed over each group of subjects are provided) and illustrated in Figure 1. Results pertaining to the estimation of $(\theta_1, \theta_2, \theta_3)$ are summarized in Table 3 (the mean and standard deviation of the estimates $(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$ computed over each group of subjects are provided).

Table 2 For each subject and protocol, we estimate the change points (τ_1, τ_2). For each group of subjects, we compute over the group the mean and standard deviation (given between parentheses) of the estimates $(\hat{\tau}_1\delta, \hat{\tau}_2\delta)$

		Protocol 1		**Protocol 2**
Hemiplegic subjects	18.4 (7.4)	43.3 (12.2)	16.8 (1.0)	44.5 (12.4)
Vestibular subjects	20.9 (5.4)	49.9 (6.5)	19.5 (6.9)	50.3 (5.3)
Normal subjects	21.4 (8.2)	47.5 (10.7)	22.4 (8.1)	51.4 (2.7)

Table 3 For each subject and each protocol, we first estimate the change points $(\hat{\tau}_1\delta, \hat{\tau}_2\delta)$ and then compute the estimates $(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$ of $(\theta_1, \theta_2, \theta_3)$ over each of the three resulting intervals $[\delta, \hat{\tau}_1\delta)$, $[\hat{\tau}_1\delta, \hat{\tau}_2\delta)$ and $[\hat{\tau}_2\delta, N\delta]$, where $\theta_k = (\theta_{1,k}, \theta_{2,k}, \theta_{3,k})$ $(k = 1, 2, 3)$. For each group of subjects and each protocol, we compute over the group and for that protocol the mean and standard deviation (between parentheses) of the estimates $(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$

				Protocol 1			**Protocol 2**
		$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$
Hemiplegic subjects	$\hat{\theta}_{1,k}$	0.97 (0.02)	0.98 (0.02)	0.98 (0.01)	0.97 (0.02)	0.98 (0.01)	0.97 (0.02)
	$\hat{\theta}_{2,k}$	0.11 (0.08)	0.39 (0.48)	0.16 (0.11)	0.14 (0.09)	0.38 (0.33)	0.38 (0.62)
	$\hat{\theta}^2_{3,k}$	0.08 (0.05)	0.06 (0.06)	0.06 (0.04)	0.10 (0.06)	0.10 (0.10)	0.11 (0.09)
		$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$
Vestibular subjects	$\hat{\theta}_{1,k}$	0.99 (0.01)	0.99 (0.01)	0.99 (0.01)	0.98 (0.02)	0.98 (0.01)	0.98 (0.01)
	$\hat{\theta}_{2,k}$	0.06 (0.03)	0.23 (0.24)	0.09 (0.09)	0.14 (0.12)	0.56 (0.37)	0.19 (0.20)
	$\hat{\theta}^2_{3,k}$	0.05 (0.04)	0.07 (0.08)	0.07 (0.06)	0.10 (0.07)	0.12 (0.09)	0.09 (0.07)
		$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$
Normal subjects	$\hat{\theta}_{1,k}$	0.99 (0.01)	0.99 (0.01)	0.99 (0.01)	0.98 (0.02)	0.99 (0.01)	0.98 (0.01)
	$\hat{\theta}_{2,k}$	0.02 (0.22)	0.13 (0.15)	0.10 (0.09)	0.08 (0.08)	0.15 (0.29)	0.15 (0.14)
	$\hat{\theta}^2_{3,k}$	0.04 (0.04)	0.02 (0.03)	0.04 (0.05)	0.06 (0.05)	0.04 (0.04)	0.06 (0.06)

Three features of Table 2 are worth commenting on. First, one notes that there is no significant difference across groups of subjects (however, the means of $\hat{\tau}_1\delta$ and $\hat{\tau}_2\delta$ are slightly shifted to the left in the group of hemiplegic subjects relative to the two other groups). Second, the means of $\hat{\tau}_1\delta$ are close to the time of start of perturbations for all groups and protocols. As for the means of $\hat{\tau}_2\delta$, they are close to the end time in vestibular and normal subjects and both protocols. In hemiplegic subjects, one notes that the standard deviations are quite large and that the mean of $\hat{\tau}_2\delta$ is shifted to the left relative to the end time for perturbations. This is due to the fact that for each protocol, 30% of hemiplegic subjects feature an estimator $\hat{\tau}_2\delta$ close to 30 s.

Third, judging by the standard deviations, all hemiplegic subjects tend to react similarly by adjusting quickly to the perturbations undergone in protocol 2. Likewise, all normal subjects tend to react similarly by adjusting quickly to the end of perturbations undergone in protocol 2. In protocol 1, the large standard deviation associated to the mean value of $\hat{\tau}_2\delta$ computed over the group of normal subjects reflects the fact that 20% of these subjects feature an estimator $\hat{\tau}_2\delta$ close to 30 s.

Regarding Table 3, we consider in turn parameters $\theta_{1,k}$, $\theta_{2,k}$ and $\theta_{3,k}$ ($k = 1, 2, 3$). First, the estimates $\hat{\theta}_{1,k}$ behave similarly across groups of subjects, protocols and intervals $[\hat{T}_{k-1}, \hat{T}_k)$ for each $k = 1, 2, 3$. In contrast, for each $k = 1, 2, 3$, the estimates of $\hat{\theta}_{2,k}$ behave quite differently across groups of subjects, protocols and intervals $[\hat{T}_{k-1}, \hat{T}_k)$. This is a promising feature for the sake of classifying subjects by group, which is the problem at stake here. As for the estimates $\hat{\theta}_{3,k}$ ($k = 1, 2, 3$), for any given group of subjects and protocol, they behave quite similarly. However, for protocol 2, it seems that the estimates $\hat{\theta}_{3,k}$ ($k = 1, 2, 3$) in normal subjects behave differently from their counterpart in hemiplegic or vestibular subjects. This is another promising feature.

3.3 Simulation study

We carry out a simulation study to evaluate the performances of the two-step estimating procedure presented in Section 3.1. We directly simulate the trajectory $C_{1:N}$. Specifically,

(i) we set $(T_1, T_2) = (\tau_1\delta, \tau_2\delta) = (15, 50)$;

(ii) we rely on the Euler scheme (2) with step size $\delta/10$ to approximate the sampling of $(C(t))_{t\in[\delta,N\delta]}$ from the distribution characterized by eq. (1) where, for each $k \in \{1, 2, 3\}$, θ_k equals the mean of its 22 estimates based on the 22 real trajectories associated to the 22 hemiplegic subjects and protocol 1 (see Section 3.2 and Table 3);

(iii) we conclude by sub-sampling $(C(t))_{t\in[\delta,N\delta]}$ to derive $C_{1:N}$.

We sample $B = 100$ independent copies of $C_{1:N}$. For each copy, we estimate the parameters $(\tau_1, \tau_2, \theta_1, \theta_2, \theta_3)$. The means and standard deviations of the estimated change points and parameters computed over the 100 independent replications are reported in Table 4.

Table 4 For each of the $B = 100$ independently simulated datasets, we derive the estimates $(\hat{\tau}_1\delta, \hat{\tau}_2\delta, \hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$. We report here the true values and means and standard deviations (between parentheses) computed over the $B = 100$ replications

	$k = 1$	$k = 2$	$k = 3$
$\hat{\theta}_{1,k}$	0.975 (0.010)	0.975 (0.009)	0.973 (0.006)
$\theta_{1,k}$	0.970	0.980	0.980
$\hat{\theta}_{2,k}$	0.114 (0.033)	0.473 (0.121)	0.214 (0.076)
$\theta_{2,k}$	0.110	0.390	0.160
$\hat{\theta}^2_{3,k}$	0.078 (0.004)	0.058 (0.002)	0.060 (0.003)
$\theta^2_{3,k}$	0.080	0.060	0.060

$\hat{\tau}_1\delta$	16.2 (0.9)	$\hat{\tau}_2\delta$	50.8 (2.1)
$\tau_1\delta$	15.0	$\tau_2\delta$	50.0

Three comments on Table 4 are in order. First, regarding the estimation of (τ_1, τ_2), we note that the means of the estimated change points are very close to their respective true values. Moreover, the standard deviations are small. Second, regarding the estimation of $(\theta_{1,k}, \theta_{3,k})$ ($k = 1, 2, 3$), we note that the means of the estimates of $\theta_{1,k}$ and $\theta_{3,k}$ are very close to their respective true values and that the standard deviations are small. Third, regarding the estimation of $\theta_{2,k}$ ($k = 1, 2, 3$), we emphasize that the means are quite apart from their respective true values. Moreover, the standard deviations are not small. Overall this indicates a poorer estimation of $\theta_{2,k}$ ($k = 1, 2, 3$) than of $(\theta_{1,k}, \theta_{3,k})$ ($k = 1, 2, 3$). This is probably due to the fact that the time intervals $[T_{k-1}, T_k)$ ($k = 1, 2, 3$) are relatively narrow given the value of δ.

4 Classification

In Section 4.2, we describe our classification procedure of subjects as hemiplegic, vestibular or normal based on their trajectories obtained under the two protocols. It is built upon the previous section. Indeed, it does not rely on the trajectories themselves but rather on their finite-dimensional summary measures whose definition, given in Section 4.1, depends on the results of our two-step estimation procedure.

4.1 Summary measures

Most classification procedures based on trajectories involve a preliminary step of dimension reduction where the high-dimensional trajectories are summarized into a low-dimensional summary measure [4, 5]. Here we build a tailored finite-dimensional summary measure of every trajectory based on the estimates $(\hat{\tau}_1, \hat{\tau}_2, \hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$ derived from it via our two-step estimating procedure.

We argue in Section 3.2 that, overall, $\hat{\tau}_1$, $\hat{\tau}_2$, $\hat{\theta}_{2,k}$ and $\hat{\theta}_{3,k}$ $(k = 1, 2, 3)$ may be relevant for the sake of classifying subjects as hemiplegic, vestibular or normal. Because the ratio $\hat{\theta}_{2,k}/(1-\hat{\theta}_{1,k}) = \hat{\mu}_k$ is easier to interpret than $\hat{\theta}_{2,k}$ (and since $\hat{\theta}_{1,k}$ varies very little across subjects and protocols) and because $\hat{\theta}_{3,k} = \hat{\sigma}_k\sqrt{\delta}$, we choose to define our finite-dimensional summary measure as $(\hat{\tau}_1, \hat{\tau}_2, \hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma}_3)$. Hereafter, we denote X^1 the latter vector derived from the trajectory associated to protocol 1 and X^2 that derived from the trajectory associated to protocol 2.

4.2 Classification procedure

We actually construct two classifiers, $\hat{\phi}_1$ and $\hat{\phi}_2$, to classify subjects as hemiplegic, vestibular or normal based on either X^1 or X^2 for $\hat{\phi}_1$, and on both (X^1, X^2) for $\hat{\phi}_2$.

The generic observed data structure associated to a given subject is (X^1, X^2, Y), where $Y \in \{1, 2, 3\}$ indicates the subject's group (with convention $Y = 1$ for hemiplegic, $Y = 2$ for vestibular, and $Y = 3$ for normal). We denote P_0 the true distribution of (X^1, X^2, Y).

Let $\mathcal{S}$ be the set of all classifiers based on $X = (X^1, X^2)$. The misclassification risk associated to $S \in \mathcal{S}$ is $R^{(P_0)}(S) = E_{P_0}[1\{S(X) \neq Y\}]$. Denote by $\mathcal{R}^* = \min_{S \in \mathcal{S}} R^{(P_0)}(S)$ its minimum. It is achieved at the Bayes classifier $S^* \in \mathcal{S}$, characterized by $S^*(X) = \arg\max_{y \in \{1,2,3\}} P_0(Y = y|X)$. For $j = 1, 2$, we also introduce the Bayes classifier $S^{*j} \in \mathcal{S}$ which relies only on X^j: $S^{*j}(X) = \arg\max_{y \in \{1,2,3\}} P_0(Y = y|X^j)$.

We consider $D_n = \{(X_i, Y_i),\ i = 1, \ldots, n\}$, where $X_i = (X_i^1, X_i^2)$is the summary measure associated to the subject i and Y_i indicates his group. Our objective is to construct, based on D_n, $\hat{\phi}_1$ and $\hat{\phi}_2$ which respectively estimate the better classifier among $S^{1^*}$and S^{2^*}(i.e., $\arg\min_{S^{1^*}, S^{2^*}} R^{(P_0)}(S^{j*})$) and S^*. We choose to rely on the popular methodology of random forests [15], which have proved powerful in a variety of applications. This is made easy thanks to the `R` package `randomForest` (we used the default tuning). The construction of the estimators $\hat{S}^{1^*}$ and $\hat{S}^{2^*}$ of $S^{1^*}$and S^{2^*} and that of $\hat{\phi}_2$ is straightforward. We derive $\hat{\phi}_1$ from $(\hat{S}^{1^*}, \hat{S}^{2^*})$ by V-fold cross-validation, with $V = 10$.

5 Application

This section is devoted to the application of our classification procedure to the motivating dataset. In Section 5.1, we apply it exactly as it is described in Section 4. In Section 5.2, we consider a slightly enhanced version with improved performance. We provide concluding remarks in Section 5.3.

5.1 Performance of the classification procedure

We evaluate the performances of the classification procedure by the leave-one-out rule and the 0.632+ bootstrap method as described by Efron and Tibshirani [16]. We apply the 0.632 + bootstrap method in order to take into account the fact that the leave-one-out rule may result in overly optimistic error rates. Resorting to the leave-one-out rule is notably motivated by the relatively small sample size of our dataset. The results are reported in Table 5 (second row).

Table 5 Leave-one-out performances and .632 + bootstrap error rates of $\hat{\phi}_1$ and $\hat{\phi}_2$ on the real dataset for the sake of classifying subjects as hemiplegic, vestibular or normal. First row: classification procedure of Section 4.2 when imposing $(\hat{\tau}_1\delta, \hat{\tau}_2\delta) = (15, 50)$. Second row: classification procedure of Section 4.2. Third row: extended classification procedure of Section 5.2

		Performances		0.632 + error rate	
		$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_1$	$\hat{\phi}_2$
Procedure of Section	4.2 with $(\hat{\tau}_1\delta, \hat{\tau}_2\delta) = (15, 50)$	61%	64%	0.40	0.40
	4.2	74%	68%	0.35	0.36
	5.2	76%	79%	0.28	0.29

With its leave-one-out performance equal to 74% and its 0.632 + bootstrap error rate equal to 0.35, the best classifier is $\hat{\phi}_1$, which involves one protocol only. For curiosity, we also evaluate the performances of $\hat{\phi}_1$ and $\hat{\phi}_2$ when systematically replacing $(\hat{\tau}_1\delta, \hat{\tau}_2\delta)$ with $(15, 50)$ (the start and end times of the perturbation phase). We report the results in Table 5 (first row). Quite satisfactorily, we note that both $\hat{\phi}_1$ and $\hat{\phi}_2$ are not as good as the original versions: estimating the change points proves very relevant.

5.2 Performance of an extended classification procedure

It is tempting to extend our classification procedure by simply extending the definition of the summary measure which is at its core. Following Chambaz and Denis [1], we merely substitute (X^1, U^1) and (X^2, U^2) for X^1 and X^2 with U^j $(j = 1, 2)$ derived from $C_{1:N}$ as $U^j = (\bar{C}_1^+ - \bar{C}_1^-, \bar{C}_2^- - \bar{C}_1^+, \bar{C}_2^+ - \bar{C}_2^-)$ where

$$\bar{C}_1^- = \frac{\delta}{5} \sum_{i \in [10/\delta, 15/\delta[} C_i, \qquad \bar{C}_1^+ = \frac{\delta}{5} \sum_{i \in]15/\delta, 20/\delta]} C_i,$$

$$\bar{C}_2^- = \frac{\delta}{5} \sum_{i \in [45/\delta, 50/\delta[} C_i, \qquad \bar{C}_2^+ = \frac{\delta}{5} \sum_{i \in]50/\delta, 55/\delta]} C_i.$$

We refer to Chambaz and Denis [1] for a justification. Then we apply the extended classification procedure and report its performances in Table 5 (third row).

The performances achieve by both classifiers $\hat{\phi}_1$ and $\hat{\phi}_2$ are slightly better than the best performance obtained in Section 5.1: enriching the summary measures seems to provide relevant additional information for the sake of classifying subjects as hemiplegic, vestibular or normal.

5.3 Perspectives

We address the statistical challenge of classifying subjects as hemiplegic, vestibular or normal based on complex trajectories obtained through two experimental protocols which were designed to evaluate potential deficits in postural control. The classification procedure involves a dimension reduction step where the complex trajectories are summarized by finite-dimensional summary measures based on a stochastic process model for a real-valued trajectory. This allows us to retrieve from the trajectories information relative to their temporal dynamic. A leave-one-out evaluation yields a 79% performance of correct classification for a total of $n = 70$ subjects, with 22 hemiplegic (31%), 16 vestibular (23%) and 32 normal (46%) subjects.

In future work, we will extend the classification procedure by introducing finite-dimensional summary measures based on a stochastic process model for the original trajectories in $\mathbb{R}^4$. We will also draw advantage from our good understanding of the classification problem to tackle the closely related statistical challenge of clustering subjects according to postural style.

Acknowledgments: I wish to thank I. Bonan (Service de Médecine Physique et de réadaptation, CHU Rennes) and P-P. Vidal (CESEM, Université Paris Descartes), my PhD advisors A. Chambaz and A. Samson (MAP5, Université Paris Descartes), and my colleague V. Genon-Catalot (MAP5, Université Paris Descartes), for introducing me to this interesting problem and providing the dataset, for their attentive guidance and all the fruitful discussions.

References

1. Chambaz A, Denis C. Classification in postural style. Ann Appl Stat 2012;6(3):977–93.
2. James G, Hastie T. Functional linear discriminant analysis for irregularly sampled curves. J Roy Stat Soc Ser B 2001;63:533–50.
3. Hall P, Poskitt D, Presnell B. A functionnal data-analytic approach to signal discrimination. Technometrics 2001;43:1–9.
4. Biau G, Bunea F, Wegkamp M. Functional classification in Hilbert spaces. IEEE Trans Inform Theory 2005;51:2163–72.
5. Ramsay J, Silverman B. Functional data analysis. New York: Springer, 2005.
6. Kloeden P, Platen E. Numerical solution of stochastic differential equation, vol. 23. In: Stochastic modelling and applied probability. Berlin Heidelberg: Springer-Verlag, 1992.
7. Bai J. Least squares estimation of a shift in linear processes. J Time Ser Anal 1994;15:453–72.
8. Bardet JM, Kengne W, Witenberger O. Multiple breaks detection in general causal time series using penalized quasi-likelihood. Electron J Stat 2012;6:435–77.
9. Basseville M, Nikiforov IV. Detection of abrupt changes: theory and application. Prentice Hall Information and System Sciences Series. Englewood Cliffs, NJ: Prentice Hall Inc, 1993. ISBN: 0-13-126780-9.
10. Cuenod C, Favetto B, Genon-Catalot V, Rozenholc Y, Samson A. Parameter estimation and change-point detection from dynamic contrast enhanced MRI data using stochastic differential equations. Math Biosci 2011;233:68–76.
11. Lavielle M. Using penalized contrasts for the change-point problem. Signal Process 2005;85:1501–10.
12. Kessler M. Estimation of an ergodic diffusion from discrete observations. Scand J Statist 1997;24:211–29. ISSN: 0303-6898.
13. Lavielle M. Detection of multiple changes in a sequence of dependent variables. Stochastic Process Appl 1999;83:79–102. ISSN: 0304-4149.
14. Lavielle M, Ludeña C. The multiple change-points problem for the spectral distribution. Bernoulli 2000;6:845–69. ISSN: 1350-7265.
15. Breiman L. Random forests. Machine Learn 2001;45:5–32.
16. Efron B, Tibshirani R. Improvements on cross-validation: the .632 + bootstrap method. J Am Stat Assoc 1997;92:548–60.