Electronic Journal of Statistics
Vol. 8 (2014) 1848–1855
ISSN: 1935-7524
DOI: 10.1214/14-EJS937E

Removing phase variability to extract a mean shape for juggling trajectories*

Nicolas J.-B. Brunel

ENSIIE & Laboratoire Statistique et Génome, Université d'Evry, France
e-mail: nicolas.brunel@ensiie.fr

and

Juhyun Park†

Department of Mathematics and Statistics, Lancaster University, U.K.
e-mail: juhyun.park@lancaster.ac.uk

Abstract: One of the purposes of the curve alignment has been to recover a *structural mean* of the curves by taking into account the common structural information or *shape*. Borrowing ideas from shape analysis, we introduce the Frenet-Serret framework to remove phase variation and to define a mean shape for three dimensional curves. Our method effectively regularizes the estimation of the geometry through curvature and torsion, and does not require curve alignment to define a mean. The method is demonstrated with the juggling data set.

MSC 2010 subject classifications: Primary 62G05, 65D10; secondary 49K15.
Keywords and phrases: Curve alignment, multi-dimensional curves, Frenet-Serret frame, Fréchet mean.

Received August 2013.

1. Introduction

Many existing curve registration methods are developed specifically for one dimensional curve. For one dimensional curve the concept of phase and amplitude variability is easily understood as horizontal and vertical variability, respectively. Consequently, visual impression of the variability of the curves and the degree of alignment of salient features are often considered sufficient to judge the need or success of curve registration. However such notion of variability cannot be directly extended to higher dimensional curves such as the Juggling data set. An easy way to get around such problem would be to summarize the curves into one dimensional features such as first or second derivatives and apply one dimensional curve registration methods. This may be sufficient, as shown in the initial analysis provided by J. Ramsay for the Juggling data set, although this may highly depend on the context of the problem. An alternative strategy would

*Main article 10.1214/14-EJS937.
†Corresponding author.

be to simultaneously apply a one dimensional curve registration method to all marginal curves, as used in the analysis of three dimensional vascular geometries by Sangalli et al. (2009).

In this article we attempt to develop a new framework to directly deal with higher dimensional curves. As the Juggling data set is our main interest, we focus on the analysis of three dimensional curves. For each trial $j = 1, \dots, J$, denote the population of cycles of curves by f_i, $i = 1, \dots, N_j$. Each curve is then a function f_i from $[0, T_i]$ to $\mathbb{R}^3$. These trajectories can roughly be considered as pseudo-periodic curves in $\mathbb{R}^3$. We are interested in analyzing the variations of the cycles $t \mapsto f_i(t)$, $i = 1, \dots, N_j$ within a trial j (where $j = 1, \dots, , J$). Our objective is to distinguish the variations in *phase* from those in *shape*. Note that we denote the first three derivatives of the function f with respect to t by $\dot{f}(t), \ddot{f}(t)$ and $\dddot{f}(t)$.

To formulate the problem, we borrow ideas from differential geometry. Differential geometry (of curves) aims at dealing with intrinsic properties, independent of the parametrization of the object. In other words, for any diffeomorphism (warping function) $h : [0, T'] \longrightarrow [0, T]$, differential geometry looks for properties shared by any function defined as $t' \mapsto f(h(t'))$. Basically, it will give the shape of the curve, independently of the way we move along the curve (coordinate systems). Shape should also be invariant under the action of Euclidean isometries (group of rotation-translation). The warping function underlying phase variation can still be understood as a time transformation, in the form of acceleration or deceleration, required to transform one shape to another. Defining variation in shape enforces us to recognize the three dimensional curves as a unit of the analysis. In particular, the description of shape variation needs to be much more elaborated, because the shape variation cannot be restricted to *amplitude* variation.

We analyze the shape by means of the Frenet-Serret representation of the curves f_i, which provides a flexible coordinate system driven by the geometric features of the curve itself. Our view in taking this approach is that geometry of the shape is an important (structural) property, moreover that there is an interplay between phase and geometry. Indeed the warping function for the juggling data can be readily understood through the tangential speed of a ball moving in one dimensional manifold. Hence it is clear that the variation in the warping functions is related to the change in the geometry of the curve.

Derivatives are informative in identifying features or landmarks for one dimensional curve registration. This is still the case for the alignment of multidimensional curves. These functional data are related to specific human movements, which are subject to some underlying law of physics. In particular, it is suspected that human movements tend to optimize some criterion, for instance the jerk, the rate of change in acceleration, i.e. the third derivative of the function in time (Todorov and Jordan, 1998). An earlier analysis of the Juggling data set by Ramsay (ch.12, Ramsay and Silverman (2002)) focused on the modeling of the third derivative $\dddot{f}(t)$, with a time varying coefficient linear differential equation as a function of the second derivative $\ddot{f}(t)$ and the first derivative $\dot{f}(t)$–but the jerk is not really minimized in this analysis. As will be seen later, the Frenet-Serret frame is a natural framework to incorporate information on derivatives.

In a very first step, we are interested in defining a mean trajectory, hence we introduce the way to compute a mean shape.

2. Frenet-Serret representation of the curves in $\mathbb{R}^3$

In order to apply the Frenet-Serret representation of the curves, we require that the curves be regular. A curve f from $[0,T]$ to $\mathbb{R}^3$ is said to be regular if f is C^3 on $[0,T]$ (piecewise C^3 would be sufficient) and the derivative does not vanish. In our case, the curves f are not closed (i.e $f(0) \neq f(T)$ as the trajectories slightly fluctuate from cycle to cycle) and they are simple i.e f is injective (the curve does not cross itself). Let $\mathcal{F}_I = \{f : I \longrightarrow \mathbb{R}^3 \mid f \in C^3, \text{regular}, \text{simple}\}$ be the space of curves that we consider in this article.

2.1. Arclength parametrization

The first geometrical information of a curve f is its length defined as $L = L[f] = \int_0^T \|\dot{f}(t)\| dt$, where $\|\cdot\|$ is the usual Euclidean norm in $\mathbb{R}^3$. A fundamental result of the differential geometry of curves states that the arc-length $s : t \mapsto s(t) = \int_0^t \|\dot{f}(u)\| du$ is an admissible parametrization of curves in $\mathcal{F}_I$, i.e., $f(t) = \gamma(s(t))$, where $\gamma : [0,L] \longrightarrow \mathbb{R}^3$ is the shape or the arc-length parametrized curve. It is then useful to distinguish in notation the derivation with respect to (w.r.t.) time t, and the derivation w.r.t. arclength (or space) s and the latter is denoted by $\frac{d}{ds} \triangleq '$. A direct consequence of the arc-length parametrization is that the derivative of the shape $\gamma'(s) \triangleq T(s)$ is the normalized tangent to the curve at point s, and the chain rule formula gives $\dot{f}(t) = \dot{s}(t)T\big(s(t)\big)$. This classical relationship shows that $\dot{s}(t) = \|\dot{f}(t)\|$, where $\dot{s}$ is called the tangential speed. This gives a simple decomposition of the tangent vector into its direction $(\frac{\dot{f}(t)}{\|\dot{f}(t)\|})$ and norm $(\|\dot{f}(t)\|)$. It is worth noting that all the shape information about the curve is contained in the direction of the tangent vector.

2.2. Frenet-Serret frames

The normalized curve γ can be completely characterized by introducing the Normal vector $N(s) = T'(s)/\|T'(s)\|$, and the Bi-Normal vector $B(s) = T(s) \times N(s)$. The triplet $(T(s), N(s), B(s))$ forms an orthonormal frame, i.e., the matrix $R(s) = (T(s) \mid N(s) \mid B(s))$ is a rotation matrix in $SO(3)$. The Frenet-Serret frame satisfies the following Ordinary Differential Equation (ODE)

$$\begin{cases} T'(s) = & \kappa(s)N(s) \\ N'(s) = & -\kappa(s)T(s) + \tau(s)B(s) \\ B'(s) = & -\tau(s)N(s) \end{cases} \tag{2.1}$$

where $\kappa, \tau : [0, L] \longrightarrow \mathbb{R}$. The function κ is the curvature and is positive, and τ is the torsion and can be of either sign. This equation is completely specified by an initial condition $T(0) = T_0$, $N(0) = N_0$ and $B(0) = B_0$, and the shape γ is obtained by integration: $\gamma(s) = \gamma(0) + \int_0^s T(u)du$.

The triplet $(T(s), N(s), B(s))$ is the Frenet-Serret frame, also known as TNB frame and represents the local change in geometry of the curve. The shape γ can be retrieved from the evolution of the frame $R(s)$ (w.r.t a reference frame, usually the canonical one in $\mathbb{R}^3$). Hence, two curves having the same curvature and torsion are the same (they have the same shape), modulo a rigid transformation, i.e., a rotation and a translation.

3. Analysis of the juggling data

For the juggling cycles within a fixed trial, the trajectories f_i can have different lengths $L_i = L\,[f_i]$. We can also define a population of shapes γ_i as well as that of arc-length parametrizations $s_i : t \mapsto s_i(t)$. Our basic model for the juggling data set is of the form

$$\forall t \in [0, T_i], \quad f_i(t) = \gamma_i(s_i(t)).$$

The objective is to obtain a mean shape $\bar{\gamma}$ in order to evaluate the variability around the mean. The information about the shape can be easily accessed by computing the derivatives frame $s \mapsto D\,[\gamma_i]\,(s)$ and deriving then the Frenet-Serret frame curve $s \mapsto R_i(s)$. For ease of comparison, we consider the normalized shapes $\tilde{\gamma}_i(s) \triangleq \frac{1}{L_i}\gamma_i\,(sL_i)$ such that $L\,[\tilde{\gamma}_i] = 1$ for all $i = 1, \dots, N$. The tangential speed is changed into $\tilde{s}_i(t) = \frac{1}{L_i}s_i(L_i t)$.

3.1. Estimation of curvature

The Frenet-Serret representation of γ is a function of (κ, τ). In order to estimate γ we could proceed by focusing on estimating (κ, τ) from the derivatives of the individual curve. A standard approach to the estimation of curvature and torsion is based on the classical expression of curvature and torsion as a function of the derivatives:

$$\begin{cases} \kappa(t) = & \frac{\left\|\ddot{f}(t)\times\dot{f}(t)\right\|}{\left\|\dot{f}(t)\right\|^3} \\ \tau(t) = & \frac{\left\langle \dot{f}(t)\times\ddot{f}(t),\, \dddot{f}(t)\right\rangle}{\left\|\ddot{f}(t)\times\dot{f}(t)\right\|^2} \end{cases} \tag{3.1}$$

However, it is well known that formulas (3.1) are inappropriate for computation, as these are numerically unstable (Younes, 2010). Hence, some regularization is necessary.

Alternatively, if the Frenet-Serret frames $R(s)$ are available on a fine grid $s_j = jh$, $j = 0, \dots, N$ (and $h > 0$ in the step size), the Euler approximation to the Frenet-Serret ODE gives

$$\begin{cases} T(s_j + h) - T(s_j) \approx & h\kappa(s_j)N(s_j), \\ N(s_j + h) - N(s_j) \approx & -h\kappa(s_j)T(s_j) + h\tau(s_j)B(s_j), \end{cases} \tag{3.2}$$

and this implies that

$$\begin{cases} \langle T(s_j+h), N(s_j)\rangle = & h\kappa(s_j), \\ \langle N(s_j+h), B(s_j)\rangle = & h\tau(s_j). \end{cases} \tag{3.3}$$

We use the second method to compute the curvature and torsion in our analysis.

3.2. Estimation of a mean shape

Our strategy for estimation of the mean shape is summarized below.

1. Compute $t \mapsto s_i(t)$ and L_i. Normalize the curve so that $\tilde{\gamma}_i(s) = \frac{1}{L_i}\gamma_i(sL_i)$. Estimate $(\tilde{\gamma}_i'(s), \tilde{\gamma}_i''(s), \tilde{\gamma}_i'''(s))$ (by local polynomial).
2. Compute the TNB frame by Gram-Schmidt (i.e QR factorization) of the frame $(\tilde{\gamma}_i'(s), \tilde{\gamma}_i''(s), \tilde{\gamma}_i'''(s))$.
3. Compute the discretized curvature and torsion with equation (3.3).
4. Compute the mean shape (TNB curve) $s \mapsto \bar{R}(s)$: Assuming that Frenet-Serret frames $R(s)$ are available on a find grid $s_j = jh, j = 0, \ldots, N$, compute the Fréchet mean of the Frenet-Seret frames (represented as an orthogonal matrix $R_i(s) \in SO(3)$):

$$\forall j,\ \bar{R}(jh) = \arg\min_{R\in SO(3)} \sum_{i=1}^{N} \|R_i(jh) - R\|_F^2 .$$

   This defines a (discrete) path $s \mapsto \bar{R}(s)$. The mean orthogonal matrix $\bar{R}(jh)$ is computed from the polar decomposition of the mean matrix $\widetilde{R_N}(jh) = \frac{1}{N}\sum_{i=1}^{N} R_i(jh)$; this means that $\widetilde{R_N}(jh) = \bar{R}(jh)U(jh)$. The matrix $U(jh)$ is symmetric and positive definite, and is an indicator for dispersion of the data.
5. By using the Euler approximation (3.2), the mean curvature and torsion $\bar{\kappa}, \bar{\tau}$ are computed by minimizing the (squared) prediction error between time s and $s+h$, hence the discrete path $s \mapsto \bar{R}(s)$ is approximately a solution of the ODE:

$$\dot{\bar{R}}(s) = \begin{pmatrix} 0 & \bar{\kappa}(s) & 0 \\ -\bar{\kappa}(s) & 0 & \bar{\tau}(s) \\ 0 & -\bar{\tau}(s) & 0 \end{pmatrix} \bar{R}(s)$$

6. The mean shape is computed by integrating the mean tangent $\bar{\gamma}(s) = \int_0^s \bar{T}(u)du$.
7. For each curve, we can compute $\bar{f}_i(t) = L_i\bar{\gamma}(\frac{s_i(t)}{L_i})$ (whereas $f_i(t) = L_i\tilde{\gamma}_i(\frac{s_i(t)}{L_i})$). The curvilinear coordinate $s_i(t)$ is the reference warping function to the mean shape. But we can find for each function γ_i the optimal warping to the mean function.

We applied this procedure to cycles from each trial to estimate a mean shape. As an example, Figure 1 displays a three-dimensional view of the estimate of the

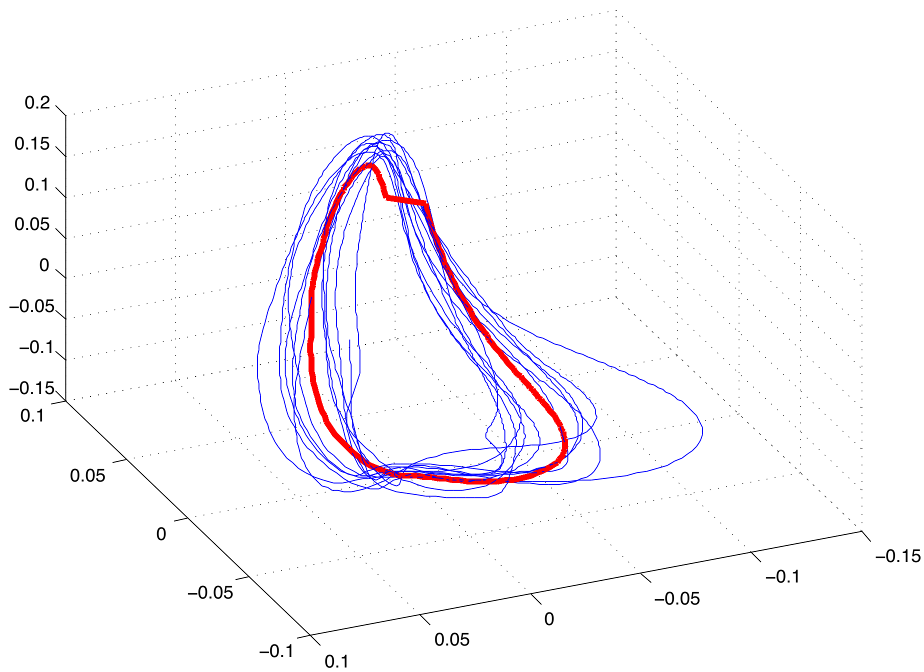

Fig 1. *Estimate of the mean shape for the cycles in trial 10 (red).*

mean curve for the cycles in trial 10. As in statistical shape analysis, the right positioning and scaling of the mean curve has been obtained by Procrustes Analysis (that gives the proper translation, rotation and scaling factor). The abrupt change in the curve around the top can be seen as an artefact of the cycle splitting of a trial, partly magnified by discretization error. Nevertheless, it emphasizes that cycles are not closed curves, and that the osculating plane (generated by Tangent and Normal) fluctuates during a trial; overall, the mean shape appears to capture the underlying common structure reasonably well. A similar conclusion could be drawn from all other trials.

An example of estimates of the mean curvature and torsion is shown in Figure 2, superimposed on the individual estimates. Although it suggests that some additional smoothing would be beneficial, the relatively stable estimates compared to the individual estimates are due to the regularization obtained by the Fréchet mean of Frenet-Serret frames. Investigation into further regularization techniques and estimation of the optimal warping function is the topic of ongoing research. Nevertheless, we see that the peak in curvature occurs exactly in the middle (0.5), at the bottom of the curve, with small variation across the cycles. Interestingly, there is a prominent peak in torsion just before the maximum curvature occurs and possibly a minor one after then, otherwise close to zero. A possible explanation would be that the hand goes forward at high rate before reaching the bottom and a small correction occurs when the hand goes

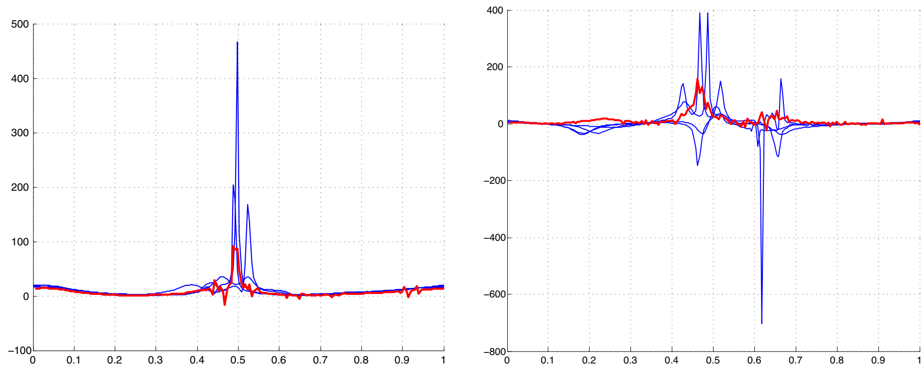

Fig 2. *Estimate of the mean curvature and torsion for the cycles in trial 1 (red).*

up, though with big variability. Consequently, the shape drifts away from the original plane. Going back to Figure 1, we see almost planar curves until the halfway, and then a bulge towards the bottom corresponding to the peak in torsion, which results in the abrupt change in the top due to the drift in shape.

4. Conclusion

One of the purposes of the curve alignment has been to recover a *structural mean* of the curves by taking into account the common structural information or *shape*. Having this in mind, we have demonstrated through the juggling data set how this could be achieved in three dimensional curves. Borrowing ideas from shape analysis, we have introduced the Frenet-Serret framework to remove phase variability and to define a mean shape. Our method effectively regularizes the estimation of the geometry through curvature and torsion, and does not require curve alignment to define a *structural mean*. On the other hand it is desirable to be able to characterize the phase variation, and to understand the link between the phase and shape variability. In particular for the juggling data, the tangential speed is directly influenced by curvature and geometry, and therefore, identifying the interplay between the phase and shape variability warrants further investigation.

Acknowledgements

The second author is grateful to MBI Mathematical Biosciences Institute for its financial support and also acknowledges the funding from ENSIIE in France to allow her to complete this work in collaboration.

References

Ramsay, J. O. and Silverman, B. W. (2002). *Applied functional data analysis: Methods and case studies* **77**. Springer New York. MR1910407

Sangalli, L. M., Secchi, P., Vantini, S. and Veneziani, A. (2009). A case study in exploratory functional data analysis: Geometrical features of the internal carotid artery. *Journal of the American Statistical Association* **104** 37–48. MR2663032

Todorov, E. and Jordan, M. I. (1998). Smoothness maximization along a predefined path accurately predicts the speed profiles of complex arm movements. *Journal of Neurophysiology* **80** 696–714.

Younes, L. (2010). *Shapes and Diffeomorphisms*. *Applied Mathematical Sciences*. Springer Berlin Heidelberg. MR2656312