Curves Similarity Based on Higher Order Derivatives

Florence Nicol, Stéphane Puechmorel

Curves Similarity Based on Higher Order Derivatives

Florence Nicol†, Stephane Puechmorel‡

Ecole Nationale de l'Aviation Civile
Email: †nicol@recherche.enac.fr, ‡stephane.puechmore@enac.fr

***Abstract*—In many applications, data originate from the observation of a phenomenon depending on time. Trajectories of mobiles fall within this category and receive an increasing attention as many connected objects have the ability to broadcast their positions. When the raw location is the value of interest, several statistical procedures exist to deal with analysis of trajectories. Depending on whether the geometrical shape or the time to position relation is relevant, one will use a parametrization invariant distance or a simple L^2 metric to assess the similarity between any two trajectories. However, it is sometimes advisable to use higher order information like velocity or acceleration, while retaining some kind of geometrical invariance. The purpose of the present work is to introduce a framework especially adapted to such a situation.**

***keywords*—bundle metric, curve manifold, shape space.**

I. Introduction

In many applications, the data of interest are measured values through time of a system in evolution. It is often the result of the observation of a physical phenomenon, obeying an underlying dynamics that may be unknown. As an extension, images can be modeled pretty much the same way, using two coordinates instead of a single time axis. Most algorithms designed to deal with such data rely on a sampled representation that is amenable to multivariate statistics.

Another approach was taken in the functional statistics framework, where the basic objects are mappings from a time interval to a given state space. Unfortunately, very little is known about probabilities in infinite dimension spaces, and one has to revert to finite dimensional representations during the implementation phase.

Several works were dedicated to the extension of classical multivariate algorithms to sample paths of Hilbert processes. As a starting point, data is first expanded on a truncated Hilbert basis [1], then, the vectors of expansion coefficients enter a standard finite dimensional analysis. A clever choice of the representation space and basis allows to take into account the prior knowledge about the studied process. Unfortunately, the dimension of the samples produced that way may be high, and varies with the geometric features of the sample paths. In particular, the presence of high curvature values will increase the number of expansion coefficients needed to keep a good approximation of the original function. In [2], an expectation maximization (EM) functional clustering algorithm is presented with an adaptive basis in each group, yielding an efficient numerical method to deal with this issue.

Another class of methods relies on a non-parametric approach [3][4]. A recent work [5] pertaining to this approach presents a hierarchical clustering principle, with application to electric power consumption.

Finally, some algorithms use a shape manifold [6] representation in order to derive a metric between sample paths. This last class of methods has distinguished benefits, like the ability to focus only on the shape of the curves and forget about a specific parametrization. The major drawback is a high computational cost, that may preclude its use on large data sets. While basically designed for using the first derivatives, it is possible to consider higher order information, although the notion of parametrization invariance becomes less intuitive.

The purpose of the present work is to introduce a framework in which the higher order derivatives are explicitly taken into account, but with a well controlled notion of invariance. It was motivated by an application in civil aviation, that is the assessment of runway adherence using only radar tracks of landing and taxiing aircraft. In this context, a full parametrization invariance is not advisable, as it will remove an important part of the relevant information. On the other side, a raw Sobolev distance between curves will be fooled by the diversity of aircraft and will induce false alarms.

In section (II), a general approach to the question of curve similarity is addressed. The main idea is to split between the observation and the geometrical object on which it lies, yielding a metric that gather in a controlled way the contribution of the underlying shape and the extra information borne by the measured data. Starting with a model of curves based on differential geometry, the data can be thought as a section of a vector bundle with base space the curve. Using a riemanian metric on it will be the key to obtain a measure of similarity between any two samples, in the spirit of the works dedicated to distances between shapes [6].

In section (III), the application of the general framework to the case of landing aircraft tracks will be discussed. Finally, a conclusion will be drawn, with a view towards future work on application on real and simulated data prior to an operational use.

II. Distance between bundles

A. Problem statement

The purpose of this section is to introduce a suitable state space for representing data that has both geometrical and cinematic features. It was motivated by the application that will be described later, where tracks of landing and taxiing aircraft must be clustered into homogeneous groups distinguished by

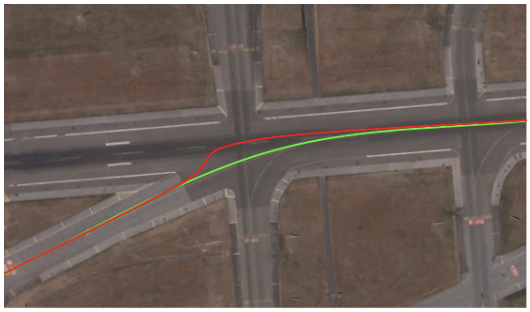

Figure 1. Runway clearing trajectories

the adherence of the runway. An example of a nominal (lower) and an abnormal (upper) track is given in Figure 1. The shape of the upper track is clearly different from the nominal one, and will not belong to the same cluster if one use a algorithm based for example on curvature. However, the divergence from the nominal curve may be due to low adherence condition or to a late turn. Only the first case must trigger a corrective action, that for the present situation is quite constraining: the runway has to be closed for the duration needed to perform an on-site adherence measurement. Using tangential acceleration will add a very discriminating feature since for the same shape, a high value means a high braking force, thus rejecting the hypothesis of low adherence. However, comparing raw velocity and acceleration induces two new issues:

- The aircraft type and airline procedures are varying, so that a shape comparison must be performed to minimize this effect;
- The time span of the trajectories is not the same. A registration procedure is needed, and it is a quite difficult problem to solve.

To summarize, neither parametrization invariant nor time dependent distances are fully satisfying. It is thus advisable to split the similarity computation into a purely geometrical part and a remainder that is not explained by shape variations.

All curves will be assumed to be smooth, that is indefinitely differentiable. This is not a real constraint in practical applications.

B. Curves as manifolds

In almost all applications, a curve is understood as a mapping γ from a real interval $[a, b]$ to a state space, generally $\mathbb{R}^p$. Furthermore, only curves with nowhere vanishing first derivative will be considered to avoid possible singularities. In practice, this condition is never an issue. This is the standard framework in which functional data statistics takes place, and is well suited to problems where the information is contained in the mapping. As an example, if one wants to analyze the delays occurring in road or air traffic, the time to position mapping is the most relevant data. However, it is quite common to encounter cases where the shape of the image $\gamma([a, b])$ holds the discriminating features. It obviously the case in image recognition algorithms, but also in spectrometric data [3], in biometric measurements (electroencephalogram, electrocardiogram) and generally speaking in all areas where the information will not change if the parametrization of the curves is changed. In the sequel, such a situation will be referred to as geometric data analysis (GDA).

The easiest approach to GDA is to let all curves be parametrized by arclength, which is defined as:

$$s\colon t \in [a, b] \to \int_a^t \|\gamma'(t)\| dt$$

The arclength has domain $[0, l(\gamma)]$ with $l(\gamma)$ the length of the curve. It comes at once that:

$$\frac{ds}{dt} = \|\gamma'(t)\|$$

so that taking the derivative with respect to s of a function of t can be done easily using the formula:

$$D_s = \frac{1}{\|\gamma'(t)\|} D_t$$

The arclength is related only to the shape of the image of γ. In fact, let $u \in [c, d]$ such that $t = \phi(u)$ with ϕ a strictly increasing smooth diffeomorphism from $[c, d]$ to $[a, b]$, it comes:

$$s(u) = \int_c^u \|D_u \gamma(t(u))\| du = \int_c^u \|\gamma'(t(u))\| \frac{dt}{du} du = \int_a^t \|\gamma'(t)\| dt \quad (1)$$

Many important features of the curve are naturally expressed with arclength: $D_s\gamma(s)$ yields the unit tangent vector $T_\gamma(s)$ while $\|D_{ss}\gamma(s)\|$ is the curvature at s.

For curve similarity computations, the drawback of the arclength is that its domain depends on the length of the curve and varies with the curves. A scaled version η with domain $[0, 1]$ is more convenient:

$$\eta = s/l(\gamma)$$

An obvious benefit of using η as a reference parametrization is that it solves the so-called registration problem, that consists of finding a common domain for all the functional samples [7]. It worth notice that in a sampled context, it is equivalent to have evenly spaced landmarks on the image of γ, as in [8].

It is worth noticing that the geometry of smooth curves with values in $\mathbb{R}^p$ is entirely defined by the so-called Frenet frame $(u_1, \dots u_p)$ and its associated curvatures $(\kappa_1, \dots, \kappa_{p-1})$. For the sake of completeness, the procedure for finding them is given by

$$u_1(t) = \gamma'(t)/\|\gamma'(t)\| \quad (2)$$

$$\tilde{u}_i(t) = \gamma^{(i)}(t) - \sum_{j=1}^{i-1} \langle \gamma^{(i)}(t), u_j(t) \rangle u_j(t), i = 2 \dots p \quad (3)$$

$$u_i(t) = \tilde{u}_i(t)/\|\tilde{u}_i\|, i = 2 \dots p \quad (4)$$

$$\kappa_i(t) = \frac{\langle u_i'(t), u_{i+1}(t) \rangle}{\|\gamma'(t)\|}, i = 1 \dots p-1 \quad (5)$$

It is clear from the construction that changing the curve parametrization will keep the Frenet frame and the curvatures invariant: all information related to velocity, tangential acceleration and so on will be lost. In many applications, this is a major issue.

From a mathematical standpoint, one possible geometric model for a curve is a one dimensional riemanian manifold. It will be assumed in the sequel that the reader is familiar with the basic concepts of differential geometry that may be found in any textbook on the subject [9]. For shape analysis and recognition, one deals almost always with closed curves, that comply with the usual notion of manifold. However, when dealing with paths with distinct endpoints, the right model is a manifold with boundary. In the sequel, all curves will be represented that way.

Any real interval $[a, b]$ has the structure of a trivial one dimensional manifold with boundary $\{a, b\}$. Tangent vectors are couples (t, u) where t is the basepoint in $[a, b]$ and u is a one dimensional vector, that can be represented as a real number. One can think of a tangent vector (t, u) as the velocity at t of a point moving along the interval $[a, b]$. Boundary conditions impose $u(a) > 0, u(b) < 0$. A vector field defined on it is just a smooth mapping $u\colon t \in [a, b] \mapsto \mathbb{R}$ such that all derivatives admit a limit to the right (resp. to the left) at a (resp. b). These limits will define the field at the boundary, and are not required to comply with the conditions satisfied by tangent vectors.

A curve $\gamma\colon [a, b] \to \mathbb{R}^p$ with nowhere vanishing derivative is an immersion from the manifold $]a, b[$ to $\mathbb{R}^p$ that can be extended to $[a, b]$. As such, it inherits a metric from the local embedding in $\mathbb{R}^p$ by letting, for any $t \in]a, b[$ and real numbers u, v, interpreted as tangent vectors:

$$g_t(u, v) = \|\gamma'(t)\|^2 uv$$

The Levi-Civita connection on $]a, b[$ is given by:

$$\nabla_{\partial_t} u = \partial_t u + \frac{1}{\|\gamma'(t)\|^2} \langle \gamma'(t), \gamma''(t) \rangle u$$

The first term corresponds to the intrinsic variation with respect to the parameter t, while the second is the part of the tangential acceleration coming from the geometry of the immersion. When the immersion parameter is chosen to be the arclength, the second term in the right hand vanishes as $D_s\gamma$ is the unit tangent vector at s and $D_{ss}\gamma$ is orthogonal to it. The Levi-Civita connection reduces thus to derivative with respect to arclength. The same applies for the scaled arclength η.

While curvature is a very important information for curves, it is a characteristic of the immersion and not an intrinsic feature of the trivial manifold $[a, b]$. It can be recovered from the immersed normal bundle $\mathcal{N}$, whose elements are couples (t, v) with $t \in]a, b[$ and $v \perp \gamma'(t)$. It is a vector bundle with base manifold $[a, b]$ and typical fiber $\mathbb{R}^{p-1}$, and its sections can be easily recovered using the Frenet frame $u_1(t), \ldots, u_p(t)$ introduced earlier, as any vector normal to $\gamma'(t)$ lies in span$(u_2(t), \ldots, u_p(t)$. A section of $\mathcal{N}$ is then a smooth mapping $s\colon [a, b] \to R^{p-1}$, with $s(t)$ the coordinates on the frame $(u_2(t), \ldots, u_p(t))$.

The immersed normal bundle is the prototype of objects bearing vector information along a curve, and fulfilling the requirement of partial geometrical invariance mentioned earlier. For a single curve $\gamma\colon [a, b] \to R^p$, one can attach a vector $v(\theta)$ for each $\theta \in [a, b]$ that is interpreted as a sample at position $\gamma(\theta)$. Going back to the case of landing tracks, one can think of this vector information as the couple velocity/acceleration $v(\theta) = (u(\gamma(\theta)), D_t u(\gamma(\theta))$. It is important to note that the shape parameter θ is not related to time t, that is used to compute velocity and acceleration: at a given θ, $u(\theta), D_t(\theta)$ are the respective velocity and acceleration sampled at position $\gamma(\theta)$ on the trajectory. Different aircraft following the same curve but with different braking profiles will have different samples $v(\theta)$. Within the immersed bundle, they will be described by different sections. However, the geometric object underlying the sections, that is the base immersion γ will remain the same. When we let it vary, it appears that data can be compared at two well defined levels: the first one is the geometry and arises from the difference between the base immersions and the second is related to the sections themselves.

In a general setting, a immersed vector bundle will be defined as a vector bundle with base manifold $[a, b]$, typical fiber $\mathbb{R}^n$ and whose sections are of the form $s(t) = v(\gamma(t)), t \in [a, b]$, with $\gamma\colon [a, b] \to \mathbb{R}^p$ an immersion. As above, the geometry is encoded in γ, while the non-geometric information is described by the section.

The ability to deal with similarity between such objects relies on a notion of distance between them, that will be now introduced.

C. Distance between bundle sections

In order to simplify the derivations, all curves will be assumed to be immersions from $[0, 1]$ to $\mathbb{R}^p$, using the η parameter. The derivation of a distance between immersed bundles sections will closely follow the principle underlying the construction of geometric distances between curves, as presented in [10]. Let $\mathcal{E}_0, \mathcal{E}_1$ be immersed vector bundles on respective immersions γ_0, γ_1 with values in $\mathbb{R}^p$ and with typical fibers $\mathbb{R}^n$. Let s_0, s_1 be sections respectively of $\mathcal{E}_0, \mathcal{E}_1$, that will represent the vector samples along the respective curves γ_0, γ_1. An immersed path between s_0 and s_1 is a smooth mapping $\phi\colon [0, 1] \times [0, 1] \to \mathbb{R}^p \times \mathbb{R}^n$ such that:

- For all $s \in [0, 1]$, the mapping $t \in [0, 1] \mapsto \phi(s, t)$ is a smooth section the trivial bundle $\mathbb{R}^p \times R^n \mapsto^\pi \mathbb{R}^p$;
- For all $s \in [0, 1]$, $\pi \circ \phi(s, \bullet)$ is a smooth immersion in $\mathbb{R}^p$;
- For all $t \in [0, 1]$, $\phi(0, t) = s_0(t), \phi(1, t) = s_1(t)$.

For a given $s \in [0, 1]$, the mapping $t \in [0, 1] \to \pi \circ \phi(s, t)$ defines an immersion from $[0, 1] \to \mathbb{R}^p$ that interpolates between γ_1, γ_2 in the sense of [10][6]. This immersion defines in turn an immersed bundle, with typical fiber $\mathbb{R}^n$. It will be referred to as $\mathcal{E}_s$ in the sequel. Finally, if a metric g_s is

available on each of the member of the family $\mathcal{E}_t$, it is possible to compute for a given ϕ a path length:

$$l(\phi) = \int_0^1 \int_0^1 g_s \left(\frac{\partial \phi}{\partial s}, \frac{\partial \phi}{\partial s} \right)^{1/2} dsdt \tag{6}$$

Having this measure at hand, the distance between s_0 and s_1 is defined the infimum of the values $l(\phi)$ over all the admissible paths ϕ.

For clustering applications, it may be convenient to use the energy of a path ϕ that is defined to be:

$$E(\phi) = \int_0^1 \int_0^1 g_s \left(\frac{\partial \phi}{\partial s}, \frac{\partial \phi}{\partial s} \right) dsdt$$

Paths minimizing the energy are the same as those minimizing the length. Since it saves a square root in the computation, the minimization is easier. However, if a distance is really needed at the end, it is enough to compute $l(\phi)$ on the minimizing path obtained. The way the interpolating bundles $\mathcal{E}_s$ can be constructed will be deferred to a future work; however, normal bundles (or closely related ones) are quite natural in applications. It was the choice made for the classification of landing aircraft tracks.

A second question is the choice of the metrics g_s. While out of the scope of the present paper, a requirement is that the family be smooth and natural in the sense of [11]. For the application, an ad-hoc metric will be derived in the next section.

III. A WORKED EXAMPLE: SKID DETECTION

A. Problem statement

The problem of early detection of runway bad adherence has a great importance in airport operational management. When the runway or the taxiways are in bad condition, the only reliable procedure used nowadays is a direct measurement using an especially designed vehicle. Unfortunately, the runway has to be closed for the duration of the operation, which has a high cost both from the economic and traffic management point of view. Attempts where made to infer adherence from clustering of landing trajectories obtained by the surface surveillance means (radars), but due to the diversity of aircraft and airline procedures, a registration must be applied to curves, which is quite awkward to design. Since both the geometry of the landing trajectories and the deceleration law are important to make the right decision, the above theoretical framework seems to be ideally suited. As the phenomenon of interest stems from contact mechanics, it is worth starting with the underlying physics to derive a suited bundle metric.

B. From contact mechanics to curve similarity

Landing aircraft may experience slip during deceleration phase when the runway is in degraded conditions. It may result from icing, snow, bad runway surface state but also from pilot's actions, namely a too strong braking action or sharp turn. In this last case, it is not related to runway condition and must not trigger a maintenance action from the airport services.

Slip can be detected on-board by comparing wheel rotation rate with aircraft velocity and computing the so-called wheel slip factor:

$$\lambda = \frac{\omega_w - \omega_a}{\max(\omega_w, \omega_a)} \tag{7}$$

where ω_w is the wheel angular velocity and $\omega_a = V_a/R_w$ is the expected angular velocity that can be computed as the ratio of the aircraft velocity to the wheel radius. Please note that on the real vehicle, several wheels are used, and the λ coefficient has to be understood as a mean value. Furthermore, due to tire elasticity, λ is not zero even if there is no actual slip: this is due to the fact that when a traction or a braking force is applied, the rubber will stretch, resulting in the tire outer part actually traveling more or less than expected from rigid body dynamics. This information is not yet downlinked in real time to ground centers and thus cannot be used in the intended application. From the ground standpoint, λ cannot be observed without on-board information, but some aspects of the landing or taxiing aircraft behavior may still be inferred. It is assumed in the sequel that Coulomb's law for friction [12] is applicable, so that the contact force F_c depends only on aircraft weight and tire/runway conditions:

$$F_c \leq \mu g M \tag{8}$$

with M the aircraft mass, g the gravity of the Earth and μ the adhesion coefficient. Without slip, μ is equal to the static friction coefficient μ_s and F_c can be increased until it reaches the upper bound in (8). At that point, slip occurs and μ drops to the value of the dynamic friction coefficient μ_d. F_c remains constant until it falls below $\mu_d g M$. In real world experiments, this simple behavior is no longer valid and one has to express μ as a function of λ, which can be found in [13]. Within this frame, the expression of the contact force is $F_c = \mu(\lambda) g M$, which is valid for both non-slip and slip case. Furthermore, in the case of aircraft, aerodynamics forces are exerted, with a net result of a braking force F_a that adds to the actual brakes action, but does not contribute to the friction analysis. Putting things together, the equation of motion along the aircraft trajectory γ can be expressed as:

$$\ddot{\gamma}(t) = \frac{F_a(t)}{M} + \mu(\lambda(t)) g \vec{u} \tag{9}$$

where $\vec{u}$ is a unit vector in the direction of the contact force F_c. Without making additional assumptions, it is not possible to use (9) for slip detection. However, if actions taken are assumed to be optimal, then F_a and $\vec{u}$ will be collinear so as to maximize the net braking effect. The expression of the aircraft dynamics becomes:

$$\ddot{\gamma}(t) = \left(K(t) + \mu(\lambda(t) g \right) \vec{u} \tag{10}$$

where the coefficient $K(t)$ accounts for the aerodynamic braking force intensity. As aircraft must loose speed fast, μ will be close to the maximum at least during the landing and the beginning of taxi. The same applies for K, as it will

not impair adherence. It can then be deduced that aircraft will try to keep the ratio between longitudinal and normal acceleration as high as possible. An observable measurement of slip condition will then be given by:

$$\theta = \arctan\left(\frac{\kappa\|D_t\gamma\|^3}{\langle D_{tt}\gamma, D_t\gamma\rangle}\right) \tag{11}$$

where γ is the aircraft trajectory and κ its curvature. It can further reduced to:

$$\theta = \arctan\left(\frac{\det\left(D_t\gamma, D_{tt}\gamma\right)}{\langle D_{tt}\gamma, D_t\gamma\rangle}\right) \tag{12}$$

using the well known property $\kappa = \det\left(D_t\gamma, D_{tt}\gamma\right)/\|D_t\gamma\|^3$.

In good runway conditions, the longitudinal acceleration will be high and nearly constant, at least in the first part of the landing trajectory. As a consequence, one can expect θ to be relatively small and be proportional to $\det\left(D_t\gamma, D_{tt}\gamma\right)$. Reciprocally, under slip conditions, a trade off has to be made between path following and deceleration: the angle θ will thus increase towards the limiting value $\pm\pi/2$.

From the above discussion, it appears that θ makes sense as a weighting factor for curve comparisons.

C. An adapted metric

In the sequel, the symbol D_t will stand for the partial derivative with respect to variable t. Higher order derivatives with respect to variables $t_1 \dots t_N$ will be written similarly as $D_{t_1 t_2 \dots t_N}$. Please note that the same variable may occur several times.

A smooth planar curve $\gamma\colon [0,1] \to \mathbb{R}^2$ will be an immersion when the derivative $D_s\gamma$ is everywhere non vanishing in $]0,1[$. The set of such curves will be denoted by $\mathbf{Imm}([0,1],\mathbb{R}^2)$. Taking $\gamma \in \mathbf{Imm}([0,1],\mathbb{R}^2)$, its length is:

$$l(\gamma) = \int_0^1 \|D_t\gamma(t)\|dt \tag{13}$$

It is invariant under parametrization change $\gamma \to \gamma \circ \phi$ with $\phi\colon [0,1] \to [0,1]$ a smooth diffeomorphism. Given a path $\Phi\colon [-\epsilon,\epsilon] \to \mathbf{Imm}([0,1],\mathbb{R}^2)$ that can be seen as a smooth mapping $\Phi\colon [-\epsilon,\epsilon]\times[0,1] \to \mathbb{R}^2$, the variation of $l(\Phi(0,\bullet))$ can be computed :

$$D_s|_{s=0} l(\Phi(s,\bullet)) = \langle D_s\Phi(0,1), T(1)\rangle - \langle D_s\Phi(0,0), T(0)\rangle \tag{14}$$

$$+ \int_0^1 \langle D_s\Phi(0,t), \kappa(t)\|D_t\Phi(0,t)\|^2 N(t)\rangle dt \tag{15}$$

with $T(t), N(t)$ the respective unit tangent and normal vectors to the curve $t \mapsto \Phi(0,t)$ and $\kappa(t)$ its unsigned curvature. The extension to more general immersions is quite straightforward [10]. In the same reference, the variation formula (14) is used to derive a riemanian metric on the quotient space $\mathbf{Imm}(\mathbb{S}^1,\mathbb{R}^2)/\mathbf{Diff}([0,1],\mathbb{R}^2)$.

In the present work, a similar approach will be taken. However, due to the fact that the slip condition must come into play as a weighting factor, it is not possible to keep invariance under change of parametrization. Furthermore, curves with vanishing second derivative must be excluded since the slip angle θ 11 is not defined at points where $D_t t\gamma(t) = 0$. The last condition boils down to the requirement that the curve $t \in [0,1] \mapsto (\gamma, D_t\gamma)$ be an immersion. The space of such objects will be denoted by $\mathbf{Imm}([0,1],\mathbb{R}^4)$.

As an arc tangent appears in the definition of θ, it is more convenient to use instead $sin(\theta)$ that has a simpler expression but otherwise similar behavior:

$$\sin(\theta(t)) = \frac{\kappa(t)\|D_t\gamma(t)\|^2}{\|D_{tt}\gamma(t)\|} = \frac{\det\left(D_t\gamma(t), D_{tt}\gamma(s)\right)}{\|D_t\gamma(t)\|\|D_{tt}\gamma\|} \tag{16}$$

Definition 1. Let γ be a smooth curve. An admissible variation of γ is a smooth mapping $\Phi\colon]-\epsilon,\epsilon[\to \mathbb{R}^2, \epsilon > 0$, such that $\Phi(0,\bullet) = \gamma(\bullet)$ and $\forall s \in]-\epsilon,\epsilon[, \Phi(s,0) = \gamma(0), \Phi(s,1) = \gamma(1)$.

An admissible variation defines a tangent vector at γ: it is the smooth vector field $t \in [0,1] \mapsto D_s\Phi(0,t)$.

The expression (16) has a nice variational interpretation as indicated in the next lemma.

Lemma 1. *Let $\gamma\colon [0,1] \to \mathbb{R}^2$ be a smooth path and Φ an admissible variation of it. Let ϕ be a smooth path such that $\phi(0) = \gamma(1)$, $\phi(1) = \gamma(0)$. Then:*

$$D_s A(0) = \int_0^1 \det\left(D_s\Phi(0,t), D_t(0,t)\right) dt \tag{17}$$

where $A(s)$ is the net area enclosed by the loop $\Phi(s,\bullet), \phi$ for $s \in]-\epsilon,\epsilon[$.

Using lemma 1, the integral :

$$\int_0^1 \frac{|\det\left(D_t\gamma(t), D_{tt}\gamma(s)\right)|}{\|D_t\gamma(t)\|\|D_{tt}\gamma\|}\|D_t\gamma(t)\|dt \tag{18}$$

may be interpreted as the total infinitesimal area swept by the curve γ when moved in the direction $D_{tt}\gamma$. This quantity is global indication of the slipping experienced along the path γ.

Turning back to bundles, if $\gamma\colon [0,1] \to \mathbb{R}^2$ is the immersion describing the geometry of the problem (with the η parametrization) and $v(\eta), v'(\eta)$ the respective velocity and acceleration at position $\gamma(\eta)$, the above metric can be adapted to yield a bundle metric. Without going into derivation detail, it can be expressed as:

$$g((u(\eta),u'(\eta)),(v(\eta),v'(\eta)) = \langle u(\eta)_{\mathcal{N}}, v(\eta)_{\mathcal{N}}\rangle\left(1+\kappa^2(\eta)\right) + \det\left(D_\eta\gamma(t), D_{\eta\eta}\gamma(s)\right) \tag{19}$$

Looking at the expression (19) reveals a sum of a geometric distance between immersions in the sense of Mumford-Michor and a proper vector variation. This bundle metric is injected in the procedure for computing distances between immersed bundles sections (6) to yield the desired similarity measure that discriminates skid conditions.

IV. Conclusion and future work

While quite developed from the theoretical standpoint, the work is still ongoing to assess performance in operational environment on real data. Unfortunately, there is no access to data correlated to adherence condition, as this information can only be obtained from direct measurements and is not communicated by airports authorities. The option taken was to develop a realistic taxi and landing simulator, that has been recently completed and will be released in open source. With the help of this tool, slipping and non-slipping trajectories can be simulated and the performance of the classification procedure assessed. On available landing data, and using an upper bound estimate of the distance based on a linear homotopy as an admissible path, promising results have been obtained. However, they cannot be matched against adherence conditions due to the aforementioned data availability issue. It is nevertheless expected, based on this experiment, that even the simplest linear homotopy procedure will outperform all the state-of-the-art algorithms investigated so far.

References

[1] J. Ramsay and B. Silverman, *Functional Data Analysis*, ser. Springer Series in Statistics. Springer, 2005.

[2] C. Bouveyron and J. Jacques, "Model-based clustering of time series in group-specific functional subspaces," *Advances in Data Analysis and Classification*, vol. 5, no. 4, pp. 281–300, 2011.

[3] F. Ferraty and P. Vieu, *Nonparametric Functional Data Analysis: Theory and Practice*, ser. Springer Series in Statistics. Springer New York, 2006.

[4] A. Delaigle and P. Hall, "Defining probability density for a distribution of random functions," *The Annals of Statistics*, vol. 38, no. 2, pp. 1171–1193, 2010.

[5] M. Boullé, R. Guigourès, and F. Rossi, *Advances in Knowledge Discovery and Management: Volume 4*. Cham: Springer International Publishing, 2014, ch. Nonparametric Hierarchical Clustering of Functional Data, pp. 15–35.

[6] D. Mumford, *Colloquium De Giorgi 2009*. Pisa: Scuola Normale Superiore, 2012, ch. The geometry and curvature of shape spaces, pp. 43–53.

[7] L. Sangalli, P. Secchi, S. Vantini, and V. Vitelli, "Functional clustering and alignment methods with applications," *Communications in Applied and Industrial Mathematics*, vol. 1, no. 1, pp. 205–224, 2010.

[8] I. Dryden and K. Mardia, *Statistical Shape Analysis*, ser. Wiley Series in Probability & Statistics. Wiley, 1998.

[9] F. Flaherty and M. do Carmo, *Riemannian Geometry*, ser. Mathematics: Theory & Applications. Birkhäuser Boston, 2013.

[10] P. W. Michor and D. Mumford, "Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms." *Documenta Mathematica*, vol. 10, pp. 217–245, 2005.

[11] O. Kowalski and M. Sekizawa, "Natural transformations of riemanian metrics on manifolds to metrics on tangent bundles," *Bull. Tokyo Gagukei University*, vol. 40, no. 4, pp. 1–29, 1988.

[12] V. L. Popov, *Contact Mechanics and Friction: Physical Principles and Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, ch. Coulomb's Law of Friction, pp. 133–154.

[13] R. Rajamani, *Vehicle Dynamics and Control*, ser. Mechanical Engineering Series. Springer US, 2011.