

ANNALES DE LA FACULTÉ DES SCIENCES DE TOULOUSE Mathématiques

STÉPHANE PUECHMOREL

Geometry of curves with application to aircraft trajectories analysis

Tome XXIV, n° 3 (2015), p. 483-504.

http://afst.cedram.org/item?id=AFST_2015_6_24_3_483_0

© Université Paul Sabatier, Toulouse, 2015, tous droits réservés.

L'accès aux articles de la revue « Annales de la faculté des sciences de Toulouse Mathématiques » (<http://afst.cedram.org/>), implique l'accord avec les conditions générales d'utilisation (<http://afst.cedram.org/legal/>). Toute reproduction en tout ou partie de cet article sous quelque forme que ce soit pour tout usage autre que l'utilisation à fin strictement personnelle du copiste est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

cedram

Article mis en ligne dans le cadre du
Centre de diffusion des revues académiques de mathématiques
<http://www.cedram.org/>

Geometry of curves with application to aircraft trajectories analysis

STÉPHANE PUECHMOREL⁽¹⁾

RÉSUMÉ. — Cet article fait un état de l'art des métriques pouvant être utilisées sur les espaces de courbes que celles-ci soient définies à partir de points de référence ou comme immersions ou plongements. Dans ce dernier cas, l'espace final est obtenu en quotientant par l'action d'un groupe de difféomorphismes afin d'assurer l'invariance par changement de paramétrage. La détermination de la métrique adéquate pour une classe de problèmes est un sujet de recherche actif, spécialement dans les domaines de la vision par ordinateur ou de la reconnaissance de formes. Des questions similaires se posent dans l'analyse des trajectoires d'avions dans le cadre de la gestion du trafic. En dépit de son importance, peu d'études ont été menées sur ce sujet, en grande partie par absence d'un cadre théorique adapté. L'utilisation des espaces de courbes ou de formes pour représenter les vols ainsi qu'un exemple d'application à la classification des trajectoires seront présentés en seconde partie de l'article.

ABSTRACT. — This article presents a survey of some metrics that can be used on geometric curve spaces which can be defined using samples points, known as landmarks, or by taking a space of immersions or embeddings and quotienting out by a group of diffeomorphisms in order to get rid of the influence of the parametrization. Finding the right metric for a class of problems is an active topic of research, with a special emphasis on applications related to computer vision or shape recognition. Similar problems arise in the field of air traffic management where the analysis of aircraft trajectories is one of the most basic issues. Despite its importance, only a few studies have been conducted on the subject, mainly due to the lack of suitable frameworks. The use of some of the shape spaces for representing aircraft flight paths, along with an example of trajectory classification will be given in the second part of the article.

(*) Reçu le 04/01/2015, accepté le 24/03/2015

⁽¹⁾ ENAC, 7, Avenue Edouard Belin, 31055 Toulouse France.
stephane.puechmorel@enac.fr

Article proposé par Jean-Pierre Otal.

1. Introduction

Based on recent studies [4], traffic in Europe is expected to grow on an average yearly rate of 2.6%, yielding a net increase of 2 million flights per year at the 2020 horizon. Long term forecast gives a two fold increase in 2050 over the current traffic, pointing out the need for a paradigm change in the way flights are managed. Two major framework programs, SESAR (Single European Sky Air traffic management Research) in Europe and Nextgen in the US have been launched in order to first investigate potential solutions and then to deploy them in a second phase. One of the main changes that the air traffic management (ATM) system will undergo is a switch from airspace based to trajectory based operations with a delegation of the separation task to the crews. Within this framework, trajectories become the basic object of ATM, changing the way air traffic controllers will be working. At the same time, present performance indicators will be partly obsolete: as an example, aircraft density, used today to dimension controllers teams, is a point measure and will become irrelevant in a trajectory based system.

The purpose of the article is to present an application of the geometry of curves to the context of trajectory based ATM. In order to get a feeling of the complexity of aircraft trajectories, radar tracks recorded on one day (12th February 2011) over France is pictured on figure 6 at the end of the text. 15030 flights took place on this day, that is quite a standard one, with half of them being commercial airliners (the other are general aviation aircraft, generally used for leisure or aerial work and fly at comparatively low altitude). The traffic pattern undergoes seasonal changes due to people travelling to holidays destinations and long term reshapes when companies create or drop routes, but major flows are quite stable and are reproduced from day to day. Understanding the way the traffic is organized is a key point in airspace structure and will also be critical when switching to trajectory based operations.

A second important application within the frame of ATM is trajectory prediction, that forecast aircraft positions based on past observations and intended flight path. The major interest in knowing where aircraft will be in the future is for early detection of situations where the distance between mobiles will fall below the minimum allowed (such an encounter is referred to as a conflict, not to be confused with a collision). When a potential conflict is detected, avoidance maneuvers are undertaken, that divert the aircraft from their original routes. Since this kind of action has an adverse effect on controllers workload and on fuel consumption, it must be decided only when the probability of real conflict occurrence is high. In the context of nowadays airspace based design, routes are predefined and well known by controllers.

Furthermore, flight management systems (FMSs) are efficient enough to limit cross-track deviation to tight bounds, letting only the longitudinal uncertainty (that is time of arrival at a given point) remaining. The classical way humans are acting in such a situation is to let things evolve as long as possible, in order to reduce uncertainties before starting conflict resolution maneuvers. In many cases, the anticipated conflict will in fact not take place. When switching to trajectory based operations, aircraft will no longer be bound to predefined routes, only to a reference business trajectory (RBT) that is published before taking off and may be revised during flight (although this is expected to be a rare event). In nominal situations, the FMS will do its job keeping the aircraft on the RBT, the trajectory prediction in such a case boiling down to a trajectory broadcast. However, meteorological conditions are not known accurately at the time RBT is designed and it may be advisable or even mandatory for the aircraft to update its initial route. Although the FMS will broadcast its new flight path, the pertinence of the information becomes lower. In such a case, an efficient trajectory predictor will help ATC controllers or even automated systems to anticipate conflicts with a sufficient level of confidence.

After a survey on the approaches introduced to compute geodesic distances, a tool dedicated to the clustering of recorded aircraft flight paths will be presented. The results obtained in this context are the best known to date. The problem of trajectory prediction starts with a good clustering, and will benefit from new means of performing it. Using spaces of curves to get a sound prediction is a promising axis of research, that is expected to be investigated in a near future.

2. The shape manifold

One of the main issues arising for aircraft trajectory analysis is how to define a notion of similarity or distance between flight paths. A recent work [3] conducted for the US Federal Aviation Administration (FAA) by the MITRE corporation addresses the question of clustering arrivals and between trajectories. A conclusion of the study is the fact that the performance of the algorithm depends critically on the distance chosen, and there is some clues indicating that the usual L^2 norm is not adequate in many situations.

Starting with this remark, more relevant distances have been investigated. A quite overlooked fact is that the shape of the trajectory is more important than the velocity law along it. It is thus advisable to use frameworks related to the geometry of flight paths.

2.1. Definition and first properties

One of the first attempts to give a mathematically sound description of the notion of shape is due to Kendall [7] with an approach based on landmarks, which are sampled points on a curve. As his primary goal was to classify shapes of similar appearance, two paths that are related by scaling, translation or rotation must be considered equivalent. Taking \mathbb{R}^p as the ambient space for shapes, a path $\gamma: [0, 1] \rightarrow \mathbb{R}^p$ will be represented as a finite sequence of n points $(\gamma(t_i)), 0 = t_1 < t_2 < t_n = 1$, where the $t_i, i = 1 \dots n$ are assumed to be fixed once for all. It is convenient to use a matrix notation for the sampled curve:

$$\Gamma = \left(\gamma(t_1), \dots, \gamma(t_n) \right).$$

with $\Gamma \in \mathbb{R}^{n \times p}$ viewed as representing a linear mapping through the natural identification:

$$\mathbb{R}^{n \times p} \simeq \text{hom}(\mathbb{R}^n, \mathbb{R}^p).$$

Two transformations are applied to Γ in order to get rid of the effects of translations and scalings. First of all, the center of mass is placed at the origin. In matrix notation:

$$\Gamma \leftarrow \Gamma T.$$

with:

$$T = \left(Id - \frac{1}{n} \mathbf{1} \right).$$

and $\mathbf{1}$ the matrix with all elements equal to 1. It is easily checked that $T^2 = T, T^t = T$. Let $e^{k,l}$ be the elementary $n \times n$ matrix with entries $e_{i,j}^{k,l} = \delta_{k,i} \delta_{l,j}$. The matrix $Id + e^{k,l}, k \neq l$ is invertible and:

$$\Gamma T (Id + e^{2,1}) \dots (Id + e^{(n-1),1}) = \Gamma \begin{pmatrix} 0 & -\frac{1}{n} & -\frac{1}{n} & \dots & -\frac{1}{n} \\ 0 & \frac{n-1}{n} & -\frac{1}{n} & \dots & -\frac{1}{n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & -\frac{1}{n} & \dots & -\frac{1}{n} & \frac{n-1}{n} \end{pmatrix}.$$

has vanishing first column. Keeping only the remaining ones gives a matrix Γ_0 , that is non zero except in the degenerate case when the points $\gamma(t_i), i = 1 \dots n$ are all located at the same place. Such a situation is not representative of a shape and will not be taken into account in the sequel. Γ_0 is further normalized to get the so-called pre-shape matrix $\tilde{\Gamma}$:

$$\tilde{\Gamma} = \Gamma_0 \left(\text{tr} (\Gamma_0 \Gamma_0^t) \right)^{-1/2}.$$

The space of the pre-shape matrices obtained after the previous two normalization stages is referred to as the pre-shape space and can be identified with $\mathbb{S}^{p(n-1)-1}$. Finally, taking the quotient by the left action of the rotation group $SO(p)$ yields the shape space Σ_p^n and the quotient map $\pi_p^n : \mathbb{S}^{p(n-1)-1} \rightarrow \Sigma_p^n$. When $p = 1$, it is clear that:

$$\Sigma_1^n = \mathbb{S}^{n-2}.$$

The $p = 2$ case can be dealt with using complex representation of points. The matrix $\tilde{\Gamma}$ will be thus written as a sequence of $n - 1$ complex numbers:

$$\tilde{\Gamma} = (z_1, \dots, z_{n-1}).$$

with $\sum_{i=1}^{n-1} |z_i|^2 = 1$. The action of $\mathbf{SO}(2)$ is just a product element-wise by a unit complex number, thus is proper and free, showing that all the spaces Σ_2^n are smooth manifolds (with the special case Σ_2^2 reduced to a point). Proceeding a little bit further, starting with a pre-shape:

$$(z_1, \dots, z_{n-1}).$$

one can take a non-zero component z_k and apply the rotation $u = z_k^{-1}|z_k|$ yielding an equivalent shape:

$$(uz_1, \dots, |z_k|, \dots, uz_{n-1}).$$

Up to a scaling transform that can be absorbed in the pre-shape construction, it comes that:

$$\Sigma_2^p \sim \mathbb{CP}^{n-2}.$$

When $p \leq 3$, rotations will fix points on the unit sphere, and the Σ_p^k will no longer be smooth manifolds. This phenomenon occurs when the matrix Γ has a rank less than $p - 1$, meaning that samples all belong to a low dimensional subspace. Away from this set of singular points, the map π_p^n is a submersion.

A riemannian metric may be defined on the non-singular part of Σ_p^n (globally when $p = 1, 2$). The pre-shape space may be endowed with the ambient inner product $\langle \tilde{\Gamma}_1, \tilde{\Gamma}_2 \rangle = \mathbf{tr}(\tilde{\Gamma}_1 \tilde{\Gamma}_2^t)$, that induces in turn a metric on Σ_p^n . The pre-shape space is identified with $\mathbb{S}^{p(n-1)-1}$, for which the aforementioned metric yields the geodesic great circle distance:

$$d(\tilde{\Gamma}_1, \tilde{\Gamma}_2) = \arccos \mathbf{tr}(\tilde{\Gamma}_1 \tilde{\Gamma}_2^t).$$

passing to the quotient, a classical argument [2] shows that on the shape space Σ_p^n :

$$d(\pi_p^n \tilde{\Gamma}_1, \pi_p^n \tilde{\Gamma}_2) = \inf_{U \in \mathbf{SO}(p)} \{d(U\tilde{\Gamma}_1, \tilde{\Gamma}_2)\}.$$

This special distance is referred to as the “procrustes distance” in [15, 7] , a term chosen with the famous (or infamous) Greek character of the same name in mind. An explicit computation of the distance is possible using the singular value decomposition. Let two pre-shapes $\tilde{\Gamma}_1, \tilde{\Gamma}_2$ be given. The singular value decomposition of $\tilde{\Gamma}_1 \tilde{\Gamma}_2^t$ is written as $V\Lambda W^t$ with Λ a diagonal matrix and $V, W \in \mathbf{SO}(p)$. Please note that it is different from the usual definition where the matrices V, W are assumed in $\mathbf{O}(p)$. As a consequence, the diagonal λ_p of $\Lambda = \mathbf{diag}(\lambda_1, \dots, \lambda_p)$ will be negative if $\mathbf{det}(\tilde{\Gamma}_1 \tilde{\Gamma}_2^t) < 0$. Since the diagonal elements of Λ are the eigenvalues of $\tilde{\Gamma}_1 \tilde{\Gamma}_2^t$, they will be preserved if $\tilde{\Gamma}_1$ (resp. $\tilde{\Gamma}_2$) is multiplied to the left by a rotation matrix, showing that Λ depends only on the equivalence classes $\pi_p^n \tilde{\Gamma}_1, \pi_p^n \tilde{\Gamma}_2$.

For any $U \in \mathbf{SO}(p)$:

$$d(U\tilde{\Gamma}_1, \tilde{\Gamma}_2) = \arccos \mathbf{tr}(UW\Lambda V^t).$$

One can find the minimum of the distance (dropping the arccos) with respect to U using the lagrangian:

$$\mathcal{L}(U, \Theta) = \mathbf{tr}(UW\Lambda V^t) + \mathbf{tr}(\Theta(UU^t - \mathbf{Id})).$$

where the multiplier Θ is symmetric. Taking the derivative with respect to U and applying the first order optimality condition yields:

$$W\Lambda V^t + U^t\Theta = 0.$$

or equivalently:

$$\Lambda + (UW)^t\Theta V = 0.$$

using the fact that Θ is real symmetric, it admits a orthonormal basis of eigenvectors and thus:

$$UW = V \Leftrightarrow U = VW^t.$$

As a byproduct, the Lagrange multiplier appears to be similar to $-\Lambda$. The minimal distance is thus:

$$d(\pi_p^n \tilde{\Gamma}_1, \pi_p^n \tilde{\Gamma}_2) = \arccos \left(\sum_{i=1}^p \lambda_i \right).$$

where the $\lambda_i, i = 1 \dots p$ are the diagonals entries of Λ . The optimal matrix U may not be unique when the rank of $\tilde{\Gamma}_1 \tilde{\Gamma}_2^t$ is strictly less than $p - 1$. In fact, it is proved in [9] that the cut locus on Σ_p^n for a non singular $\pi_p^n \tilde{\Gamma}_1$ is the set:

$$\{\pi_p^n \tilde{\Gamma}, \mathit{rank}(\tilde{\Gamma}_1 \tilde{\Gamma}^t) < p - 1 \text{ and } \lambda_p = -\lambda_{p-1}\}.$$

A nice property of such a distance, that makes it so interesting for shape recognition applications, is that some usual properties of the standard euclidean distance can be readily transposed to shape spaces. As an example, defining a mean shape $\pi_p^n \bar{\Gamma}$ from a finite set of shapes $(\pi_p^n \tilde{\Gamma}_1, \dots, \pi_p^n \tilde{\Gamma}_N)$ can be done following as:

$$\pi_p^n \bar{\Gamma} = \operatorname{argmax}_{\pi_p^n X} \sum_{i=1}^N d(\pi_p^n \tilde{\Gamma}_i, \pi_p^n X).$$

$\pi_p^n \bar{\Gamma}$ is called the Karcher mean of the set $(\pi_p^n \tilde{\Gamma}_1, \dots, \pi_p^n \tilde{\Gamma}_N)$.

Using the procedure described in [8] for sphere "centroids" gives another way to compute centroids of shapes on Σ_p^n , allowing in turn clustering algorithms like k -means or mean shift to be applied.

For aircraft trajectory applications, the setting must be adapted as the translation and scaling invariance is not relevant. Instead, the centroid move and normalization scaling will be incorporated into the computation of the distance, along with the $\mathbf{SO}(2)$ part. Namely, this amounts to split the metric into orthogonal parts, the first one being related to trajectory registration and the second one to shape deformation. The corresponding terms that need to be added to shape distance are the euclidean distance between centroids, the absolute value of the difference between the logarithms of the scaling factors and the absolute value of the angles between the angles representing the $\mathbf{SO}(2)$ rotation part. Geodesics produced that way are obtained as a combination of four independent moves:

- Translation at constant speed of the centroids.
- Scaling with an exponential factor of the form $\exp(at)$ with a a fixed real number.
- Rotation at constant speed around the centroid.
- Geodesic move in shape space.

2.2. Implementation

The shape space based approach works fine on synthetic examples but cannot be applied as is to real traffic due to its high computational cost. Since a SVD is needed for each pair of trajectories and the size of the matrix involved is $n \times n$ with n the number of landmarks used on each curve, the amount of elementary operations needed scales as $o(n^3)$ for the SVD

part and $o(N^2)$ for the pairwise distance, with N the number of curves. A trajectory is well represented using around 20 landmarks, apart from some exceptional cases already mentioned in the introduction where the flight path is self crossing. A practical limit for a crude implementation will be around 100 trajectories, depending on the hardware available.

A better approach is to split the task into two parts. Since the distance as several independent components, it is advisable to first clusterize traffics using centroids translations, scalings and $\mathbf{SO}(2)$ action. In fact, is not even needed to use exact expressions: taking only the endpoints on the trajectory and letting the centroid be the middle point of the segment, the scaling the inverse of the length and the rotation the one between the normalized segment and the x -axis is enough to segregate efficiently. Using this procedure, a coarse clustering may be produced in less than 1s even on a set of thousands of trajectories. Once obtained, the shape part will be used only within an upper level cluster in order to refine it. The time needed to fully clusterize a complete day of traffic over France (8000 flights when keeping only airliners) is less than 1 minute. A work is conducted so as to use the capabilities of modern graphics hardware in order to be able to get it with a latency compatibly with human interaction (less than 1s). Completing this program will provide stakeholders with the first tool allowing trajectory design in real time.

3. Metrics on spaces of curves

As in the case of shapes manifolds, the studies in the geometry of curves were initiated for dealing with questions in computer vision. Basically, one wants to obtain a sound mean of saying that two objects are similar and a possible answer is to compare the curves bounding them. It turns out that the convenient settings is the space of planar closed curves, that are assumed to be smooth, even if some sharp edges may be encountered in applications: it enough to approximate them smoothly without spoiling the overall aspect of the shape. For aircraft trajectories analysis, flight path are non closed curves, so an adaptation of the general framework has to be done.

3.1. Spaces of planar curves

For shape recognition applications, smooth simple planar closed curves are adequately modeled as embeddings from the unit circle to \mathbb{R}^2 , the set of which will be referred to as $\mathbf{Emb}(\mathbb{S}^1, \mathbb{R}^2)$. It is sometimes convenient to allow self-intersections in the considered curves, and this becomes mandatory for applications in air traffic: some aircraft waiting for availability of

runways will be bound to follow closed round paths known as racetracks. and sketched in figure 1. Furthermore, while uncommon, some trajectories

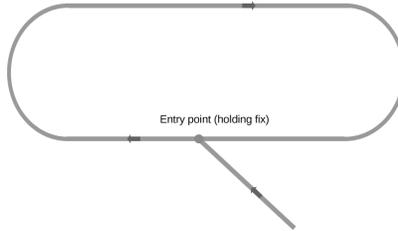


Figure 1. — A standard racetrack holding pattern

generated during radio nav aids calibration or special flights, like sport event coverage, may exhibit intricate patterns with many self-crossings. The set of immersions $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$ will thus be considered also. Going back to the original problem of shape classification or recognition, it is clear that changing the way a curve is parametrized will not influence its appearance. Translating this into mathematical language amounts of quotienting out the space $\mathbf{Emb}(\mathbb{S}^1, \mathbb{R}^2)$ (resp. $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$) by the group of smooth diffeomorphisms of \mathbb{S}^1 to itself, so as to get:

$$\begin{aligned} \mathbf{Emb}(\mathbb{S}^1, \mathbb{R}^2) &\rightarrow \mathbf{Emb}(\mathbb{S}^1, \mathbb{R}^2) / \mathbf{Diff}(\mathbb{S}^1) \simeq B_e \\ \mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2) &\rightarrow \mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2) / \mathbf{Diff}(\mathbb{S}^1) \simeq B_i \end{aligned} \quad (3.1)$$

The notation B_e, B_i used for the quotient space is borrowed from [13] and will be used in the sequel. A riemannian metric will be introduced on either $\mathbf{Emb}(\mathbb{S}^1, \mathbb{R}^2)$ or $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$ in such a way that the quotient map turns into a Riemannian submersion [14]. In the general case, given a mapping:

$$\pi: M \rightarrow B.$$

such that the tangent map $d\pi$ to π at any point is onto, it will be a Riemannian submersion if it preserves length of horizontal vectors. In more detail, if the tangent space to M is split into a sub-bundle tangent to the fibers $\pi^{-1}(b), b \in B$ and an orthogonal one (with respect to the given metric G on TM), then there is an induced metric G_B on TB such that for any two tangent vectors u, v in TB , $G(d\pi^{-1}(u), d\pi^{-1}(v))$ does not depend on the particular choice of the representatives $d\pi^{-1}(u), d\pi^{-1}(v)$ and $G_B(u, v) = G(d\pi^{-1}(u), d\pi^{-1}(v))$. The situations depicted in the diagrams 3.1 give rise to Riemannian submersion, so that the metric in the top space can be propagated to the quotient.

It turns out to be quite easy to obtain a metric on $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$ that is invariant to re-parametrization of the curves. Starting with $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$

as the top manifold, a tangent vector at the curve γ is simply a smooth mapping from $\gamma(\mathbb{S}^1)$ to \mathbb{R}^2 , represented by an element of $C^\infty(\mathbb{S}^1, \mathbb{R}^2)$. Vertical tangent vectors are elements sent by the derivative of the quotient map to 0 and are of the form $t \in \mathbb{S}^1 \mapsto \alpha(t)\gamma'(t)$ with α a nowhere vanishing real valued smooth mapping. A metric G on $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$ will give rise to a Riemannian submersion if given any base curve γ , and any couple (h, k) of tangent vectors with k vertical, $G_\gamma(h, k) = 0$ or equivalently $G_\gamma(h, \gamma') = 0$ for any tangent vector h .

A quite general form for suitable metrics is:

$$G_\gamma^L(h, k) = \int_{\mathbb{S}^1} \langle L_\gamma h(t), k(t) \rangle \|\gamma'(t)\| dt.$$

where γ is the base point. L is a field of a positive definite pseudo-differential operators on $C^\infty(\mathbb{S}^1, \mathbb{R}^2)$ in the following sense:

$$L: T\mathbf{Imm}(S^1, \mathbb{R}^2) \rightarrow T\mathbf{Imm}(S^1, \mathbb{R}^2).$$

is a smooth bundle isomorphism and for any base point γ , L_γ is a pseudo-differential operator, symmetric and positive for the L^2 metric. Invariance by the action of $\mathbf{Diff}(\mathbb{S}^1)$ is due to the $\|\gamma'(t)\|$ factor for if $\phi \in \mathbf{Diff}(\mathbb{S}^1)$:

$$\|\gamma \circ \phi'(t)\| = |\phi'(t)| \|\gamma'(\phi(t))\|.$$

so that:

$$\begin{aligned} G_{\gamma \circ \phi}(h, k) &= \int_{\mathbb{S}^1} \langle Lh(t), k(t) \rangle |\phi'(t)| \|\gamma'(\phi(t))\| dt \\ &= \int_{\mathbb{S}^1} \langle Lh(t), k(t) \rangle \|\gamma'(t)\| dt \\ &= G_\gamma(h, k) \end{aligned}$$

The simplest of member of this family is obtained with $L = \mathbf{Id}$:

$$G_\gamma^0(h, k) = \int_{\mathbb{S}^1} \langle h(t), k(t) \rangle \|\gamma'(t)\| dt.$$

The metric induced in the quotient space is expressed easily by noting that any tangent vector at base point γ orthogonal to the vertical space can be written as a mapping:

$$t \in \mathbb{S}^1 \mapsto \alpha(t)N(t).$$

with α a real valued smooth mapping and $N(t)$ the unit normal vector to γ at t . Given any couple (α, β) of smooth real mappings on \mathbb{S}^1 , it comes:

$$G_\gamma^{B_i}(\alpha N, \beta N) = \int_{\mathbb{S}} \alpha(t)\beta(t) \|\gamma'(t)\| dt.$$

Unfortunately, it turns out to be of no interest as geodesic distance vanishes everywhere.

An other classical metric is obtained using Sobolev inner products:

$$G_\gamma^N(h, k) = \int_{\mathbb{S}^1} \sum_{j=0}^N \langle D^j h(t), D^j k(t) \rangle \|\gamma'(t)\| dt.$$

with D the parametrization insensitive operator acting on a tangent vector h at base point γ as:

$$Dh(t) = \|\gamma'(t)\|^{-1} h'(t).$$

which does not induces degenerate geodesic distance. As mentioned in [13], the Sobolev metric of order N can be simplified to:

$$G_\gamma^N(h, k) = \int_{\mathbb{S}^1} \langle h(t), k(t) \rangle + \lambda \langle D^N h(t), D^N k(t) \rangle \|\gamma'(t)\| dt.$$

where $\lambda > 0$ acts as a weight between the two terms. Once again, no degeneracy occurs, but it turns out that if the metric is expressed as:

$$G_\gamma^N(h, k) = \int_{\mathbb{S}^1} \langle Lh(t), k(t) \rangle dt.$$

with:

$$L = \mathbf{Id} + (-1)^N \lambda D^{2N}.$$

the inverse operator L^{-1} admits a simple closed form expression as an integral operator. When using Sobolev metrics, the splitting into vertical and horizontal spaces has to be changed accordingly. In fact, the horizontal tangent vectors at γ for the general metric G^L is the set :

$$\{h \in T\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2) \mid L_\gamma h = \alpha N\}.$$

where N is the unit normal vector to γ and α is a mapping in $C^\infty(\mathbb{S}^2, \mathbb{R}^2)$. The description of the horizontal space is thus implicit unless L^{-1} is known, which makes the simplified Sobolev metric easier to manipulate than the full one.

Finally, it is worth to mention the curvature based metric introduced in [11]:

$$G_\gamma^c(h, k) = \int_{\mathbb{S}^1} \langle h(t), k(t) \rangle (1 + \lambda \kappa^2(t)) \|\gamma'(t)\| dt.$$

where κ is the curvature of γ . Since it is a weighted version of the G_γ^0 , the splitting into horizontal and vertical space is the same as in the simple case. However, no degeneracy occurs for the geodesic distance. Here again the tuning parameter $\lambda > 0$ controls the influence of the curvature term.

3.2. Geodesic computation

For the simplest metric G^0 , the geodesic equation can be obtained readily, and gives some insights for more complex cases. To begin with, a smooth path in $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$ is a smooth mapping:

$$\phi: [a, b] \times \mathbb{S}^1 \rightarrow \mathbb{R}^2.$$

such that for any $t \in [a, b]$, the mapping $\phi(t, \bullet)$ belongs to $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$. The energy of a path is thus:

$$E(\phi) = \int_a^b \int_{\mathbb{S}^1} \langle \partial_t \phi(t, s), \partial_t \phi(t, s) \rangle \|\partial_s \phi(t, s)\| ds dt.$$

Given any two curves γ_1, γ_2 , a smooth homotopy ϕ between them in $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$, ϕ will be critical for the energy if giving an admissible variation ϵ the first order term in ϵ in the expression of $E(\phi + \epsilon)$ vanishes. By direct computation this term is:

$$\int_0^1 \int_{\mathbb{S}^1} 2 \langle \partial_t \phi, \partial_t \epsilon \rangle \|\partial_s \phi\| ds dt + \int_0^1 \int_{\mathbb{S}^1} 2 \langle \partial_t \phi, \partial_t \phi \rangle \langle \partial_s \phi, \partial_s \epsilon \rangle \|\partial_t \phi\|^2 ds dt = \tag{3.1}$$

$$- \int_0^1 \int_{\mathbb{S}^1} 2 \langle \partial_t (\partial_t \phi \|\partial_s \phi\|), \epsilon \rangle ds dt - \int_0^1 \int_{\mathbb{S}^1} \left\langle \partial_s \left(\|\partial_t \phi\|^2 \frac{\partial_s \phi}{\|\partial_s \phi\|} \right), \epsilon \right\rangle ds dt \tag{3.2}$$

yielding the geodesic equation in $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$:

$$\partial_s \left(\|\partial_t \phi\|^2 \frac{\partial_s \phi}{\|\partial_s \phi\|} \right) = -2 \partial_t (\partial_t \phi \|\partial_s \phi\|).$$

Passing to the quotient amounts to write $\partial_t \phi$ as a normal component only, $\partial_t \phi = \alpha N$ with α a smooth real valued mapping and N the unit normal vector for the curve $\phi(t, \bullet)$. Keeping only the normal contribution yields the quotient (or horizontal) geodesic equation as:

$$\alpha^2 \kappa = 2 \partial_t \alpha.$$

where an arclength parametrization is assumed. As mentioned before, geodesic distance for the metric G^0 is vanishing using the following theorem:

THEOREM 3.1. — *Let γ_1 and γ_2 be elements of $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$. For any $\epsilon > 0$ it exists a smooth homotopy ϕ in B_i between $\pi \gamma_1$ and $\pi \gamma_2$ such that $E(\phi) < \epsilon$.*

Proof. — The main idea behind the proof is to select a family of paths with steep slopes. Since the normal component is inversely proportional to the length of the path, and is squared in the computation of the energy, its product with the length can be made as close to 0 as wanted. For the sake of simplicity, the re-parametrization ψ that will be used for the proof is only piecewise continuous, being piecewise linear, but can be approximated arbitrary well by smooth ones. The domain of ψ is selected to be $[0, 1] \times [0, 1]$. Let N be a fixed integer. ψ is defined to be:

$$\psi(t, s) \mapsto \begin{cases} 4t(sN - k), & \text{if } s \in \left[\frac{k}{N}, \frac{k+1/2}{N} \right] \\ 4t(sN - k), & \text{if } s \in \left[\frac{k+1/2}{N}, \frac{k+1}{N} \right] \end{cases} .$$

when $t < 1/2$ and for $t \geq 1/2$:

$$\psi(t, s) \mapsto \begin{cases} 2t - 1 + 4(1-t)(sN - k), & \text{if } s \in \left[\frac{k}{N}, \frac{k+1/2}{N} \right] \\ 2t - 1 + 4(1-t)(sN - k), & \text{if } s \in \left[\frac{k+1/2}{N}, \frac{k+1}{N} \right] \end{cases} .$$

Let ϕ be an homotopy between γ_1, γ_2 . Since $\psi(0, \bullet) = 0, \psi(1, \bullet) = 1$, the mapping:

$$\tilde{\psi}: (t, s) \mapsto \phi(\psi(t, s), s).$$

is still an homotopy between the same curves. The partial derivatives with respect to t, s are given by:

$$\partial_t \tilde{\psi} = \partial_t \phi \partial_t \psi, \quad \partial_s \tilde{\psi} = \partial_t \phi \partial_s \psi + \partial_s \phi.$$

Since only the horizontal part of the vector is of interest, it may be assumed that $\partial_t \phi$ and $\partial_s \phi$ are orthogonal, so that:

$$\|\partial_s \tilde{\phi}\|^2 = \|\partial_s \phi\|^2 + \partial_s \psi^2 \|\partial_t \phi\|^2.$$

To compute the normal component of $\partial_t \tilde{\phi}$, using again the orthogonality of $\partial_t \phi$ and $\partial_s \phi$, a normal vector to $\partial_s \tilde{\phi}$ is obtained easily as:

$$\|\partial \phi_s\|^2 \partial_t \phi - \partial_s \psi \|\partial_t \phi\|^2 \partial_s \phi.$$

that can be normalized to:

$$\frac{\frac{\|\partial_s \phi\|}{\|\partial_t \phi\|} \partial_t \phi - \frac{\|\partial_t \phi\|}{\|\partial_s \phi\|} \partial_s \phi}{\sqrt{\|\partial_s \phi\|^2 + (\partial_s \psi)^2 \|\partial_t \phi\|^2}}.$$

The horizontal part of $\partial_t \tilde{\phi}$ is then:

$$\frac{\partial_t \psi \|\partial_t \phi\| \|\partial_s \phi\|}{\|\partial_s \tilde{\phi}\|}.$$

The length of the path $\tilde{\phi}$ is:

$$\int_0^1 \left(\int_0^1 \frac{(\partial_t \psi)^2 \|\partial_t \phi\|^2 \|\partial_s \phi\|^2}{\|\partial_s \tilde{\phi}\|} ds \right)^{1/2} dt = \tag{3.3}$$

$$\int_0^1 \left(\int_0^1 \frac{(\partial_t \psi)^2 \|\partial_t \phi\|^2 \|\partial_s \phi\|}{\sqrt{1 + (\partial_s \psi)^2 \frac{\|\partial_t \phi\|^2}{\|\partial_s \phi\|^2}}} ds \right)^{1/2} dt \tag{3.4}$$

without going into complete computation that can be found in [12], it is enough to note first that being continuous mappings on the the compact $[0, 1]^2$, $\|\partial_s \phi\|$, $\|\partial_t \phi\|$ are bounded, the fist one being also bounded away from 0. Assuming large N , the inner integral will behave roughly as N^{-1} and thus goes to 0 as $N \rightarrow \infty$. \square

The shape of the mapping ψ along with its effect on a linear homotopy on two circles is given on figures 2, 3.

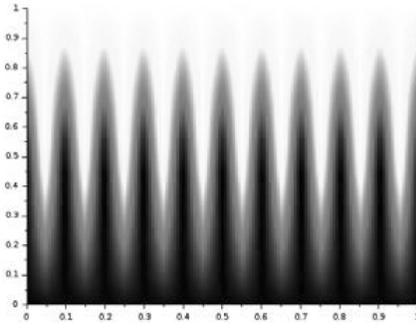


Figure 2. — The ϕ mapping with $N = 10$. White is 1, and 0 is black

The degeneracy of G^0 precludes its use for curve distance computation. Even when using numerical approximation, geodesics will tend to pinch and will not yield interesting results. Looking at the above proof, the degeneracy can be avoided by introducing second order terms. As invariance with respect to diffeomorphic reparametrization is needed to have a metric yielding a Riemannian submersion when passing to the quotient, it is natural to introduce curvature. The energy of an homotopy ϕ between two immersions is defined to be:

$$E(\phi) = \int_0^1 \int_0^1 (1 + \lambda \kappa^2) \langle \partial_t \phi, \partial_t \phi \rangle \|\partial_s \phi\| ds dt.$$

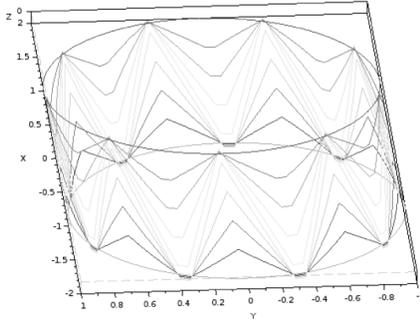


Figure 3. — Vanishing length path between circles

where $\lambda > 0$ is a tuning parameter and $\kappa(t, s)$ is, for fixed t , the curvature of the path:

$$s \mapsto \phi(t, s).$$

As previously mentioned, passing to the quotient B_i simply amounts to restrict to horizontal components. The method used for G^0 can be reproduced verbatim to obtain (after much more computations !) the geodesic equation in $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)$:

$$\partial_t \left((1 + \lambda \kappa^2) \|\partial_s \phi\| \partial_t \phi \right) = \partial_s \left(\frac{-1 + \lambda \kappa^2}{2} \frac{\|\partial_t \phi\|^2}{\|\partial_s \phi\|} + \lambda \frac{\partial_s (\kappa \|\partial_t \phi\|^2)}{\|\partial_s \phi\|} N \right).$$

with N being, for any fixed t , the unit normal vector to the curve $s\phi(t, s)$. The expression is significantly more intricate than with the metric G^0 and in turn yields to non straightforward numerical implementations.

For application to aircraft trajectories, the original energy has to be modified so as to reflect the fact that they are not closed curves. An extra term:

$$\int_0^1 \|\partial_t \phi(t, 0)\|^2 + \|\partial_t \phi(t, 1)\|^2 dt.$$

is added to the energy in order to include boundary effects. Solving the geodesic equation is a computationally intensive process that makes difficult the integration with a clustering algorithm. The choice was made not to solve the partial differential equation giving ϕ but to directly minimize the energy after discretization. Partial derivatives are approximated using a standard centered finite difference scheme, except at the boundaries of the domain, and the curvature is obtained from:

$$\kappa = \frac{|\det(\partial_s \phi, \partial_{ss} \phi)|}{\|\partial_s \phi\|^3}.$$

with again finite difference approximations of the partial derivatives. Please note that no special effort has been made on the numerical implementation itself, that may be clearly improved a lot. A limited memory BFGS algorithm was used in final to get the optimal values for the discretization points.

Since curvature is non-linear and somewhat hard to obtain from a numerical point of view, it is advisable to switch to a Sobolev metric. Unfortunately, as mentioned above, the splitting of the tangent bundle is dependant on the differential operator used. A clever approach presented first in [16] for a special case and generalized later in [1] is used to split the Sobolev metric into a tangential and a normal part. For a base curve γ and tangent vectors h, k it is expressed as:

$$G^{srvt}(h, k) = \int_{\mathbb{S}^1} \lambda \langle \partial_t h(t), N(t) \rangle \langle \partial_t k(t), N(t) \rangle + \mu \langle \partial_t h(t), T(t) \rangle \langle \partial_t k(t), T(t) \rangle dt.$$

where λ, μ are positive real numbers and $N(t), T(t)$ are respectively the unit normal and tangent vector to γ at t . Since h, k contribute only with their first derivative, the metric is translation invariant and must be used on the quotient space $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)/\mathcal{T}$ with \mathcal{T} the group of translations. The choice $\lambda = 1, \mu = 1/\sqrt{2}$ made in [16] allows to use a trick referred to as the Square Root Velocity Transform (SRVT). A element γ in $\mathbf{Imm}(\mathbb{S}^1, \mathbb{R}^2)/\mathcal{T}$ is mapped to a smooth curve by:

$$\gamma \mapsto \sqrt{\|\partial_t \gamma\|} T.$$

with $T(t)$ the unit tangent vector to γ at t . The transform is invertible:

$$\phi \mapsto \left(t \mapsto \int_0^t \|\phi(s)\| \phi(s) ds \right).$$

furthermore, it turns out that the metric G^{srvt} is the pullback of the L^2 metric by the SRVT transform. This fact allows an easy and efficient computation of geodesics, but also of Karcher means. Here again, underlying objects are closed curves, and it is not possible to plug the method as-is in the context of aircraft trajectories. However, a simple workaround, similar to the geometric-Sobolev metric of [10], may be applied. The main idea is to write a path $\gamma: [0, 1] \rightarrow \mathbb{R}^2$ as a curve $\tilde{\gamma}: [0, 1] \rightarrow \mathbb{R}^2$ with $\tilde{\gamma}(0) = \tilde{\gamma}(1) = 0$ that is submitted to a centroid translation, scaling and rotation as in the case of shape spaces based on landmarks.

4. Application to major flows identification

A classical problem in air traffic management is to extract from trajectories databases the most probable paths that a flight is prone to follow. Most of the work done on the subject amounts to compute aircraft density maps and goes back to the eighties [5]. An example of such a map, computed using modern graphics hardware, is given in figure 4. One approach to

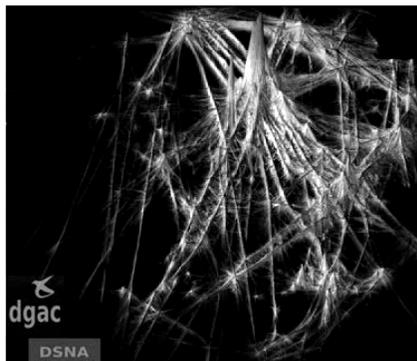


Figure 4. — A density map produced by the visualization software “FromDady”

major flows identification is to start with pixels as seeds and to grow path by connecting neighboring points with highest density. However, it is quite easy to construct counter-examples where the tracks produced are in fact followed by no aircraft. Using the geometric-Sobolev metric and a simple k -means algorithm, no such oddities occur. The result presented on figure 5 was obtained on a set of 1700 landing trajectories at Blagnac airport with a pre-selected number of clusters of 4. Mean trajectory in each cluster is plotted in black for the highest density ones and in red for the remaining two. It has be noted that all mean trajectories are sound from an operational point of view and are related to standard procedures. Some atypic paths can be seen on the picture, with no apparent effect on the mean trajectories. Among them, one can identify clearly outliers and perhaps aberrant measures: when using density maps, the net effect of such curves is to distort the computed mean track, sometimes moving it away from the real areas of interest. One the picture 5, no such method was able to separate the two clusters to the right as their differ only by the final turn. Without careful tuning, even the two clusters to the left were merged at the end. On the complete french airspace, the recorded flight path over one day looks like the picture 6. Different behaviours are observed depending on the altitude. Lower airspace flights are mostly general aviation using visual flying rules (VFR) and are color-coded blue while the upper airspace is dedicated to

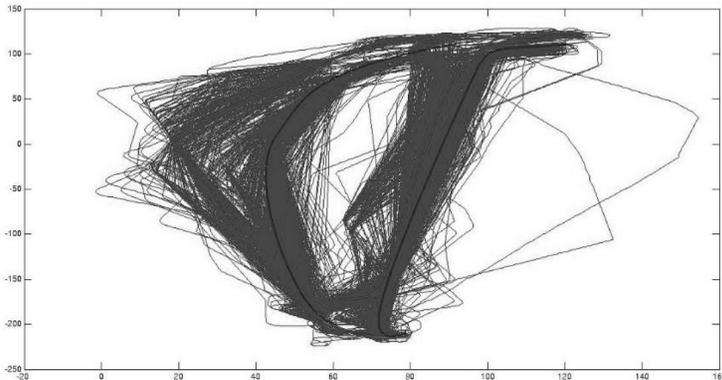


Figure 5. — Landing trajectories at Blagnac Airport

commercial airliners appearing in green on the figure. Due to the high number of displayed trajectories (18000), a special visual trick was used: the luminance of a pixel is computed according to the number of tracks above it, so that areas with low density fade. Applying a classification procedure similar to the one used for landing trajectories yields the picture 7. A visualization technique referred to as “bundling” [6] was adapted and used in the display in order to get some thickness around the cluster mean trajectories. In this special view, trajectories are gathered around the cluster center, but endpoints are kept fixed. Furthermore, in order to have a usable view, all tracks have a transparency factor close to the maximum so that only paths with a sufficient number of neighboring trajectories are apparent.

The algorithm not only allows to pick up the major flows, but also extracts some uncommon trajectories (that must be flown however by a significant number of aircraft): some transversal routes between national airports are made apparent, that were masked on the original picture.

Day to day variations can be spotted on the pictures. Looking at the two images in figures 8, 9 shows globally similar traffic patterns, but noticeable differences from point to point. It turns out that the two days used in the study have differences that can be explained: the first one is a Monday, at the the beginning of the All Saints Holidays while the second is located at the middle of the week: destinations are mainly leisure in the first case and business in the second.

Pictures are produced in less than 1 minute with a i7 quad core unit, equipped with 32Go of RAM. As for the landmarks based approach, the use of modern graphics hardware computational capabilities is under investiga-

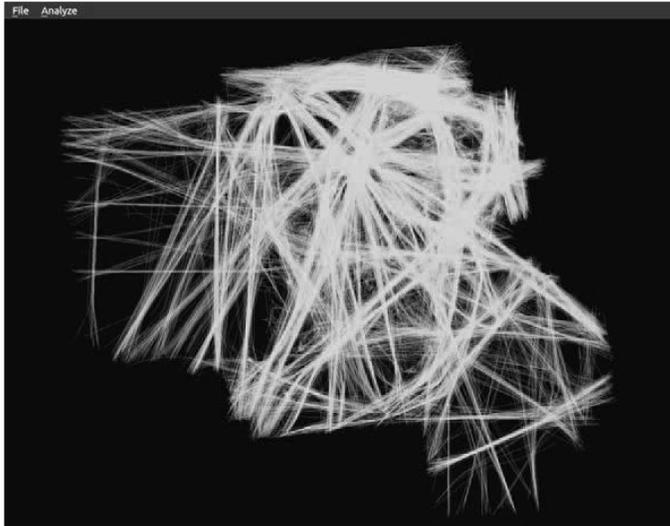


Figure 6. — One day of traffic over France.

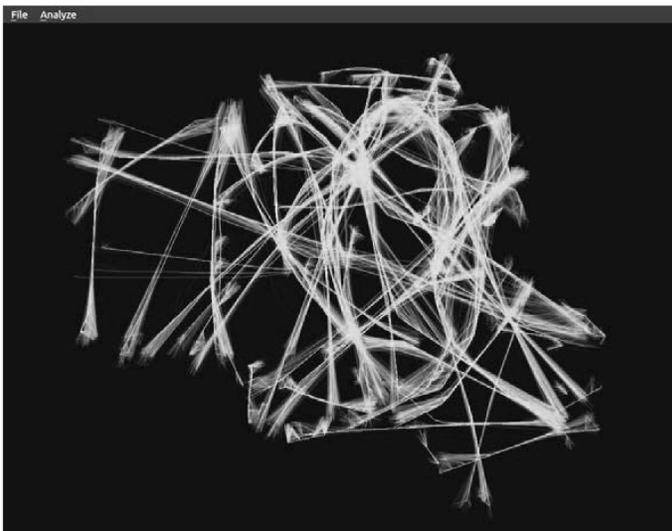


Figure 7. — One day of traffic after bundling procedure.

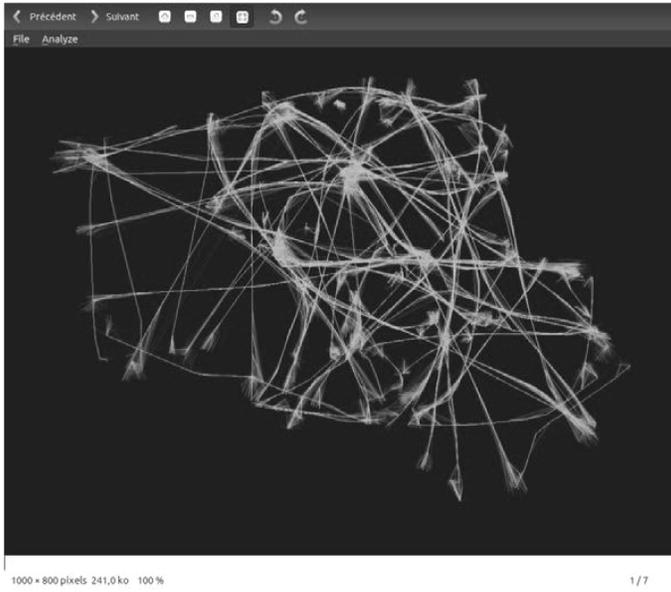


Figure 8. — Bundled traffic on the 21-10-2013

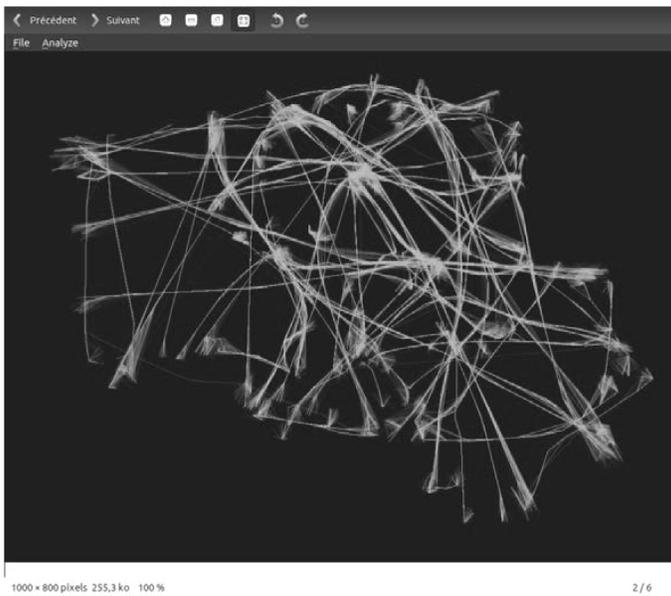


Figure 9. — Bundled traffic on the 24-10-2013

tion and it is expected that an integrated environment will be made available under open source licence, freely downloadable from the Ecole Nationale de l'Aviation Civile (ENAC) website.

5. Conclusion and future work

Finding metrics in geometric spaces of curves is an active area of research, mainly with applications to shape recognition and computer vision in mind. Within the frame of air traffic management, a similar problem arises, that is the analysis of aircraft trajectories. As ATM will gradually shift from airspace to trajectory based operations, it will become a central question in the context of future systems. Only a few studies have been conducted on the subject, as a suitable framework has yet to be introduced. However, the potential applications are numerous and have all a very high impact on the performance of the system. Using geometric measures similar to those designed for computer vision allows a new approach to be taken in order to overcome the limitations of sample based ones. It has been successfully applied to the problem of flight path classification, yielding the first usable analyzing system.

Future works on the topic will make use of descriptions of trajectories as currents, for which a notion of geodesic distance can be defined. The main idea will be to get rid of the curve parametrization, in the spirit of landmark based approaches.

Finally, developments are under progress in order to make use of modern hardware computational abilities so as to leverage the computation power of desktop computers. It is expected that a new tool will be soon released that can be applied to days of traffic in an interactive fashion, allowing the operator to generate and tune clusterings on the fly.

Bibliography

- [1] BAUER (M.), M. BRUVERIS (M.), S. MARSLAND (S.), and MICHOR (P. W.). Constructing reparametrization invariant metrics on spaces of plane curves. ArXiv e-prints, July 2012.
- [2] CARNE (T.K.). — The geometry of shape space. Proc. of the London Math. Society, p. 407-432 (1989).
- [3] ENRIQUEZ (M.). — Identifying temporally persistent ows in the terminal airspace via spectral clustering. In Tenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2013), (2013).

- [4] EUROCONTROL/NMD/STATFOR. Eurocontrol seven-year forecast. <https://www.eurocontrol.int/sites/default/files/content/documents/official-documents/forecasts/seven-year-flights-service-units-forecast-2014-2020-sep2014.pdf> (2014).
- [5] HARMAN (W.H.). — Air traffic density and distribution measurements. Lincoln Laboratory, MIT, Report ATC-80, May 1979.
- [6] HURTER (C.), ERSOY (O.), and TELEA (A.). — Smooth bundling of large streaming and sequence graphs. In Visualization Symposium (Pacific Vis), 2013 IEEE Pacific, p. 41-48, Feb 2013.
- [7] KENDALL (D. G.). — Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2), p. 81-121 (1984).
- [8] KRAKOWSKI (K.), HÜPER (K.), and MANTON (J.). — On the computation of the karcher mean on spheres and special orthogonal groups. *Proc. Workshop Robot. Math. (RoboMat07)* (2007).
- [9] LE (H.). — On geodesics in euclidean shape spaces. *Journal of the London Mathematical Society*, s2-44(2), p. 360-372 (1991).
- [10] MENNUCCI (A. C. G.), SOATTO (S.), SUNDARAMOORTHY (G.), and YEZZI (A.). — A new geometric metric in the space of curves, and applications to tracking deforming objects by prediction and filtering. *SIAM Journal on Imaging Sciences*, 4(1) (2011).
- [11] MICHOR (P. W.) and DAVID MUMFORD (D.). — Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Documenta Math*, p. 217-245 (2005).
- [12] MICHOR (P. W.) and MUMFORD (D.). — Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)*, 8, p. 1-48 (2006).
- [13] MICHOR (P. W.) and MUMFORD (D.). — An overview of the riemannian metrics on spaces of curves using the hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23(1), p. 74-113 (2007).
- [14] O'NEILL (B.). — The fundamental equations of a submersion. *The Michigan Mathematical Journal*, 13(4), p. 459-469, 12 (1966).
- [15] SIBSON (R.). — Studies in the robustness of multidimensional scaling: Perturbational analysis of classical scaling. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(2), p. 217-229 (1979).
- [16] SRIVASTAVA (A.), ERIC KLASSEN (E.), SHANTANU H. JOSHI (S. H.), and JERMYN (I. H.). — Shape analysis of elastic curves in euclidean spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(7) p. 1415-1428 (2011).