

Séparation de la forme et du temps dans les séries temporelles - application à l'authentification de signatures

Pierre-François Marteau*

*IRISA, Université de Bretagne Sud, Campus de Tohannic, 56000 Vannes
prénoms.nom@univ-ubs.fr,
<http://people.irisa.fr/Pierre-Francois.Marteau/>

Résumé. Nous nous intéressons au problème de la séparation des composantes de forme et de temps dans les séries temporelles. En statistique ce problème concerne la séparation des variations d'amplitude et de phase dans les données fonctionnelles. La forme recherchée est *neutre* temporellement dans le sens où celle-ci peut éventuellement exister en dehors de toute spécification temporelle, comme c'est le cas pour une forme géométrique. Nous proposons d'exploiter et d'adapter un algorithme d'alignement temporel probabiliste, initialement conçu pour estimer le centroïde d'un ensemble de séries temporelles, pour apporter des éléments de solution à ce problème de séparation. Nous montrons sur des données de synthèse que cet algorithme répond empiriquement au besoin, puis l'évaluons sur des données réelles dans le cadre de tâches de détection de signatures falsifiées produites en-ligne. Sur les *benchmarks* exploités, l'approche proposée se positionne légèrement au dessus de l'état de l'art, illustrant ainsi l'intérêt applicatif de l'approche.

1 Introduction

La notion de forme telle qu'appréhendée par la communauté de la reconnaissance des formes (Andreopoulos et Tsotsos, 2013; Niennattrakul et al., 2012) peut-être assimilée à la notion de *motif* associé en général à un support statistique, susceptible d'être extrait de données très diverses, telles que des images, des séries temporelles, des séquences symboliques, etc. Nous adoptons dans cet article une définition plus restrictive, proche d'une caractérisation géométrique selon laquelle une forme peut exister en dehors de toute considération temporelle. C'est le cas par exemple lorsque l'on définit une ellipse à partir de l'équation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. L'ensemble des solutions de cette équation définit complètement et de manière *atemporelle* une ellipse. Cette forme peut cependant être définie via des équations paramétriques (Eq. (1)) qui réintroduisent un paramètre, t , que l'on convient d'appeler en général variable temporelle.

$$\begin{cases} x(t) = a.\cos(\omega t) \\ y(t) = b.\sin(\omega t) \end{cases} \quad (1)$$

Dans cette seconde formulation, les séries temporelles $x(t)$ et $y(t)$ définissent également complètement la forme d'ellipse. Lorsque ce paramètre t est échantillonné, celui-ci est assimilable à un index permettant d'énumérer des couples $(x(t_k), y(t_k))$ solutions de l'équation non

Séparation de la forme et du temps

paramétrique et de les regrouper cette fois-ci dans un multi-ensemble $\mathcal{E} = \{(x(t_k), y(t_k)); k = 0, 1, \dots\}$, avec $t_k = k/f_e$ et f_e la fréquence d'échantillonnage. On peut alors identifier la forme associée à l'ellipse au multi-ensemble \mathcal{E} .

Les formes caractérisables de cette manière sont relativement diverses et potentiellement complexes. Elles émergent en particulier dans les comportements asymptotiques de certains systèmes dynamiques non linéaires, stationnaires ou non, stochastiques ou déterministes, décrits par des équations différentielles ou aux différences, sous la forme de cycles limites ou d'attracteurs étranges comme les attracteurs fractals de Rosler ou de Lorentz.

Nous nous intéressons ici au problème de la séparation des composantes de forme (supposée *atemporelle*) et de temps dans un sous-ensemble de séries temporelles produites par un processus stochastique générateur. Ce problème a été abordé sous l'angle de la caractérisation des variabilités d'amplitude et de phase, en particulier dans le cas des réalisations multiples de processus ponctuels aléatoires (Panaretos et Zemel, 2016). Pour le clustering de données fonctionnelles non alignées, des solutions d'optimisation sous contrainte ont également été proposées dans (Sangalli et al., 2010) en considérant des fonctions paramétriques (linéaires) d'alignement temporel.

En nous inspirant de la formalisation proposée dans (Panaretos et Zemel, 2016) pour les données fonctionnelles, nous étudions cette séparation pour des séries temporelles discrètes déphasées de manière quelconque à travers une interprétation probabiliste des noyaux d'alignement élastiques (Marteau, 2019) et détaillons l'approche heuristique qui en découle. La troisième partie de cet article est consacrée à une expérimentation qui permet d'évaluer dans un premier temps la solution algorithmique développée sur des données synthétiques que nous contrôlons. Dans un deuxième temps nous testons notre algorithme sur des données réelles dans le cadre de tâches d'authentification de signatures en ligne. Dans la dernière partie conclusive, nous discutons les résultats obtenus et les perspectives que ceux-ci suggèrent.

2 Formalisation du problème et de la solution proposée

Nous nous inspirons ici de la formalisation proposée dans (Panaretos et Zemel, 2016).

2.1 Contexte

Pour les données fonctionnelles, la variabilité dite d'amplitude est caractérisée en considérant N réalisations d'un élément de (L^2, D) (ensemble des fonctions définies sur $D \subset \mathbb{R}$ de carré sommable), $\mathcal{O} = \{o_i(t) : t \in D; i = 1, \dots, N\}$ admettant une fonction moyenne $\mu(t)$ dans (L^2, D) telle que :

$$o_i(t) = \mu(t) + z_i(t), \quad i = 1, \dots, N \quad (2)$$

où z_i un processus *iid* de moyenne nulle et de noyau de covariance $K(u, v)$ supposé continu.

On considère par ailleurs que la variabilité de phase découle de la présence de variations (non linéaires) sur l'axe t qui engendrent des courbes \tilde{o}_i distordues temporellement :

$$\tilde{o}_i(t) = o_i(h_i^{-1}(t)) = \mu(h_i^{-1}(t)) + z_i(h_i^{-1}(t)), \quad i = 1, \dots, N \quad (3)$$

Les fonctions de distorsion temporelle $h_i^{-1} : D \rightarrow D$ sont supposées aléatoires et monotones croissantes, indépendantes des z_i et telles que $\mathbb{E}[h_i(t)] = t$.

Généralement, les données se présentent sous la forme de mesures discrètes disposées sur une grille $\{t_j\}_j^m \subset D$ et soumises à un bruit additif de variance σ^2 ,

$$\tilde{o}_{ij} = \tilde{o}_i(t_j) + \epsilon_{ij}, \quad i = 1, \dots, N; j = 1, \dots, m \quad (4)$$

Le problème de séparation de la variabilité d'amplitude et de phase consiste alors à retrouver les fonctions $\{h_i\}$ et $\{o_i\}$ à partir des données discrètes $\{\tilde{o}_i\}$.

Résoudre ce problème de séparation nécessite d'introduire des hypothèses supplémentaires sur les fonction h_i , au delà de la monotonie et de l'identité en espérance. Ces hypothèses prennent la forme, par exemple, de points remarquables (*extrema*, plateaux, etc.) sur les courbes à aligner (Gervini et Gasser, 2004), de patrons temporels paramétriques (Ramsay et Li, 1998) ou encore de moments d'inerties à aligner (Tang et Müller, 2008), etc.

Nous abordons ce problème de séparation en nous limitant à l'extraction des fonctions $\{h_i(t)\}$, ce qui permet en principe de remonter à $\mu(t)$ puisque $\tilde{o}_i(h_i^{-1} \circ h_i(t)) = o_i(t)$. En pratique l'approche développée ci-après construit simultanément les estimations des fonctions $\{h_i(t)\}$ et $\mu(t)$.

2.2 Éléments de solution heuristique au problème de séparabilité

Le problème général de séparabilité tel que formalisé précédemment est sous-contraint (ou mal posé) et à ce titre n'admet probablement pas de solution dans le cas général, comme semble l'indiquer l'état de l'art. En revanche, nous montrons que l'algorithme proposé dans (Marteau, 2019) pour estimer une série temporelle moyenne, ou centroïde, d'un ensemble de données fonctionnelles, répond empiriquement au problème posé. Nous présentons ci-dessous les concepts généraux de cet algorithme basé sur une formalisation probabiliste du mécanisme d'alignement temporel (pour les détails, cf. (Marteau, 2019)) et montrons en quoi il apporte des éléments de réponse au problème de séparabilité posé.

Contrairement à l'algorithme de comparaison dynamique (DTW) (Sakoe et Chiba, 1971) qui considère uniquement le meilleur chemin pour appairer les paires d'échantillons de deux séries temporelles, les noyaux temporellement élastiques (Cuturi et al., 2007; Marteau et Gibet, 2014) considèrent la somme des scores d'alignement de tous les chemins d'appariement possibles. Le score attribué à un chemin est le résultat du produit des noyaux locaux d'appariement évalués le long du chemin.

Soient $o_1^n = o_1 o_2 \dots o_n$ et $o_1^{n'} = o_1' o_2' \dots o_{n'}'$ deux séries temporelles (observations) de longueurs respectives n et n' . Le noyau local quantifiant la proximité de deux échantillons o_i et o_j' s'exprime en général sous la forme d'un noyau exponentiel : $\kappa(o_i, o_j') = e^{-\nu \|o_i - o_j'\|^2}$. En interprétant le noyau local d'appariement comme une fonction de probabilité (Marteau, 2019), il est possible de concevoir l'alignement de deux séries temporelles comme l'espérance mathématique du résultat produit par un automate d'états finis stochastique.

L'automate stochastique considéré est un modèle de Markov caché défini par un ensemble fini d'états $\mathcal{S} = \{S_{1,1}, S_{1,2}, \dots, S_{n,n'}, S_{n+1,n'+1}\}$. Chaque état $S_{i,j}$ caractérise l'alignement entre deux échantillons o_i et o_j' . $S_{1,1}$ est assimilé à l'état initial et $S_{n,n'}$ à l'état final.

La variable d'état cachée (à l'étape τ) z_τ prend valeur dans \mathcal{S} . L'index d'étape τ suit un chemin d'alignement admissible et prend valeur dans $\{(1,1), \dots, (n,n')\}$. Pour tenir compte des contraintes d'alignement (de type DTW), on impose : $\tau = (i,j) \rightarrow \tau + 1 \in \{(i,j) +$

Séparation de la forme et du temps

1), $(i + 1, j), (i + 1, j + 1)\}$ et $\tau - 1 \in \{(i, j - 1), (i - 1, j), (i - 1, j - 1)\}$.

La probabilité *a posteriori* pour que le processus soit dans l'état $S_{i,j}$ à l'étape τ , étant données les séquences d'observations o_1^n et $o_1^{n'}$, est $P(z_\tau = S_{i,j} | o_1^n, o_1^{n'}, \theta)$, où θ regroupe les paramètres du modèle. Cette probabilité *a posteriori* est estimée grâce aux récursions dites avant (*forward*) et arrière (*backward*) dérivables pour les modèles de Markov cachés.

L'appariement local de deux échantillons particuliers o_i et o'_j peut être alors interprété comme la probabilité de présence de l'automate sur la cellule $S_{i,j}$.

La probabilité marginale pour que le processus soit à l'étape τ , dans l'un des états du sous ensemble $S_{t,\bullet} = \{S_{t,1}, S_{t,2}, \dots, S_{t,n'}\}$ étant données les observations o_1^n et $o_1^{n'}$, (autrement dit, la probabilité pour que le processus visite à l'étape τ l'un des états de $S_{t,\bullet}$, ou encore que l'échantillon o_t soit aligné avec l'un des échantillons de la série $o_1^{n'}$) s'écrit :

$$P(z_\tau \in S_{t,\bullet} | o_1^n, o_1^{n'}, \theta) = \sum_{t'} P(z_\tau = S_{t,t'} | o_1^n, o_1^{n'}, \theta) \quad (5)$$

Pour tout t et t' , la probabilité conditionnelle de visiter l'état $S_{t,t'}$ connaissant les deux séries d'observation, θ et le sous-ensemble $S_{t,\bullet}$, (autrement dit la probabilité que o_t et $o'_{t'}$ soient alignés sachant que o_t est aligné avec l'un des échantillons de la série $o_1^{n'}$) s'écrit :

$$P(z_\tau = S_{t,t'} | o_1^n, o_1^{n'}, z_\tau \in S_{t,\bullet}, \theta) = \frac{P(z_\tau = S_{t,t'} | o_1^n, o_1^{n'}, \theta)}{P(z_\tau \in S_{t,\bullet} | o_1^n, o_1^{n'}, \theta)} \quad (6)$$

L'espérance mathématique des alignements des échantillons de la série $o_1^{n'}$ avec l'échantillon o_t (sachant que o_t est aligné par l'automate) et de manière similaire l'espérance mathématique des temps d'occurrence associés aux échantillons de la série $o_1^{n'}$ alignés avec o_t sont alors définis par :

$$\begin{aligned} \mathbb{E}(o' | o_t) &\propto \sum_{t'=1}^{n'} o'_{t'} P(z_\tau = S_{t,t'} | o_1^n, o_1^{n'}, z_\tau \in S_{t,\bullet}, \theta) \\ \mathbb{E}(t' | o_t) &\propto \sum_{t'=1}^{n'} t' P(z_\tau = S_{t,t'} | o_1^n, o_1^{n'}, z_\tau \in S_{t,\bullet}, \theta) \end{aligned} \quad (7)$$

Les espérances précédentes (Eq. (7)) sont à la base de la procédure permettant de calculer une moyenne *temporellement élastique* pour un ensemble de séries temporelles.

Soit $\mathcal{O} = \{o_1^{n_i}\}_{i=1 \dots N}$ un sous-ensemble de séries temporelles, r_1^n une série de référence (r_1^n peut être initialisé par le médoïde de \mathcal{O}). L'estimation du centroïde de \mathcal{O} est défini par le couple $\hat{\mu} = (\hat{a}_\mu, \hat{h}_\mu)$ où \hat{a}_μ représente les amplitudes du centroïde de longueur n et \hat{h}_μ la séquence des occurrences temporelles associées aux amplitudes de \hat{a}_μ .

$$\begin{aligned}
\hat{a}_{\mu,t} &= \frac{1}{N} \sum_{i=1}^N \mathbb{E}(i|r_t) \\
&\propto \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{n_i} i o_{it} P(z_\tau = S_{t,i} | r_1^n, i o_1^{n_i}, z_\tau \in S_{t,\bullet}, \theta) \\
\hat{h}_{\mu,t} &= \frac{1}{N} \sum_{i=1}^N \mathbb{E}(t|r_t) \\
&\propto \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{n_i} t P(z_\tau = S_{t,i} | r_1^n, i o_1^{n_i}, z_\tau \in S_{t,\bullet}, \theta)
\end{aligned} \tag{8}$$

$(\hat{a}_\mu, \hat{h}_\mu)$ est une série temporelle non uniformément échantillonnée : $\hat{h}_{\mu,t}$ est l'espérance mathématique du temps d'occurrence pour le $t^{\text{ième}}$ échantillon $\hat{a}_{\mu,t}$.

Par itérations successives, en remplaçant la série temporelle de référence r_1^n par l'estimation $(\hat{a}_\mu, \hat{h}_\mu)$ obtenue à l'itération précédente, jusqu'à ne plus obtenir d'amélioration (i.e. de diminution de l'inertie), on obtient ainsi une estimation finale approchée du centroïde de \mathcal{O} .

Il est à noter ici que le problème des alignements multiples sous-jacent étant NP-complet (Wang et Jiang, 1994) et de complexité exponentielle (Just et Just, 1999) avec la taille de \mathcal{O} , seules des heuristiques sous optimales permettent d'approcher cette notion de centroïde temporellement élastique pour des ensembles \mathcal{O} de plus d'une dizaine de séries temporelles.

Nous conjecturons que, à la fin du processus itératif, $(\hat{a}_\mu, \hat{h}_\mu)$, tel que défini par les équations (8), est un estimateur empirique pour la forme moyenne μ associée à l'ensemble \mathcal{O} . D'autre-part, les équations (7) fournissent également l'accès à des estimateurs pour les fonctions temporelles h_i associées aux éléments de \mathcal{O} . Ainsi, pour $o_i = o_{i1}^{n_i} \in \mathcal{O}$, étant donnée l'estimation de la moyenne $\hat{\mu}$ obtenue sur \mathcal{O} , l'estimation proposée pour la fonction temporelle \hat{h}_i associée à o_i est :

$$\hat{h}_i(t) = \mathbb{E}(t'|\hat{a}_{\mu,t}(t)) \propto \sum_{t'=1}^{n_i} t' P(z_\tau = S_{t,t'} | \hat{a}_{\mu,1}^n, o_{i1}^{n_i}, z_\tau \in S_{t,\bullet}, \theta) \tag{9}$$

Algorithm 1 Estimation de μ et $\{h_i\}$

Require: $\{o_i, i = 1 \dots N\}$, un ensemble de séries temporelles.

$r \leftarrow \text{Medoid}(\{o_i\})$

$h_{r,t} \leftarrow t, t = 1..|r|$

$h_{i,t} \leftarrow t, t = 1..|o_i|$

$\text{Inertie} \leftarrow +\infty$

do

$\text{Inertie}_0 \leftarrow \text{Inertie}; r_0 \leftarrow r; h_{0r,t} \leftarrow h_{r,t}; h_{0i,t} \leftarrow h_{i,t}, t = 1..|r|, i = 1 \dots N$

$r, h_{r,t} \leftarrow \text{Eq. (8)}, t = 1..|r|$

$h_{i,t} \leftarrow t \text{ Eq. (9)}, t = 1..|r|$ (réalignement de h_i sur h_r), $i = 1 \dots N$

$\text{Inertie} \leftarrow \sum_{i=1}^N \sum_{t=1}^{|r|} ||o_{i,[h_{i,t}]} - r_t||^2$

while $\text{Inertie} < \text{Inertie}_0$

$\hat{\mu} \leftarrow (r_0, h_{0r,t}); \hat{h}_{i,t} \leftarrow h_{0i,t}, t = 1..|r|, i = 1 \dots N$

return $\hat{\mu}, \{\hat{h}_i\}$

Autrement dit, chaque $\hat{h}_{i,t}$ est assimilé à l'espérance mathématique des temps d'alignement des échantillons de la série o_i avec l'échantillon $\hat{a}_{\mu,t}$ de la forme moyenne.

Séparation de la forme et du temps

L'algorithme (1) permet d'estimer μ et $\{h_i, i = 1 \dots N\}$. Cet algorithme est en complexité $O(N \cdot |\hat{a}_\mu|^2)$, c'est à dire quadratique vis-à-vis de la taille moyenne des séries et donc de la taille $|\hat{a}_\mu|$ de la forme moyenne, et linéaire vis-à-vis du cardinal de \mathcal{O} .

3 Expérimentations

Nous détaillons ci-après deux expérimentations pour justifier empiriquement notre conjecture et illustrer le fonctionnement de l'algorithme. Le méta paramètre ν du noyau local $\kappa(o_i, o'_j)$ est optimisé automatiquement sur les données d'apprentissage, de manière à éviter les convergences vers zéro des probabilités calculées.

3.1 Données de synthèse

A titre d'exemple, nous considérons le processus aléatoire (de type ellipse) défini par :

1. la matrice

$$G_{a,b}(t) = \begin{bmatrix} a \cdot \cos(h_\omega(t) + \varphi(t)) & 0 \\ 0 & b \cdot \sin(h_\omega(t) + \varphi(t)) \end{bmatrix}$$

2. les fonctions temporelles $h_\omega(t) = \omega \cdot t$,
3. les contraintes sur les paramètres, $a = a_0 \cdot (1 + U_{[-.5,.5]})$, $b = b_0 \cdot (1 + U_{[-.5,.5]})$, $\omega = \omega_0 \cdot (1 + U_{[-.1,.1]})$, $\varphi = \varphi_0 + U_{[-.1,.1]}$ où $U_{[u,v]}$ est la loi uniforme sur l'intervalle $[u, v]$, et
4. le bruit Gaussien additif de moyenne nulle et de variance σ^2 , $\epsilon_\sigma(t)$.

La forme moyenne pour ce processus est $\mu(t) = G_{a_0, b_0}(\omega_0 t + \varphi_0)$.

Les paramètres sont instanciés de la manière suivante : $\omega_0 = 2\pi f_0$ avec $f_0 = 1Hz$ et $f_e = 400Hz$ la fréquence d'échantillonnage, $\varphi_0 = 0$, $a_0 = 2$, $b_0 = 1$, les ellipses étant toutes centrées en $(1, .5)$, et ϵ est un bruit Gaussien d'écart type $\sigma = 1.5$.

A partir de ce processus, un ensemble $\mathcal{O} = \{\tilde{o}_i\}$ de 20 séries temporelles est produit. La forme moyenne $\hat{\mu}$ et les composantes temporelles $\{\hat{h}_i\}$ associées sont estimées à partir de \mathcal{O} . Les résultats obtenus sont présentés en Fig. 1.

Pour illustrer la variabilité d'amplitude (non bruitée), la sous-figure en haut à gauche présente les projections des 20 ellipses produites aléatoirement. En haut au centre, la sous-figure présente la variabilité temporelle en projetant les 20 fonctions temporelles h_i associées aux formes précédentes. En bas à gauche, sont présentées : les trajectoires bruitées des 20 séries à partir desquelles la séparation est testée (courbes de couleur cyan), la forme moyenne a_μ attendue (courbe noire en pointillé) et la forme moyenne estimée \hat{a}_μ (courbe rouge). La sous-figure en bas au centre présente l'estimation des 20 fonctions temporelles \hat{h}_i , la courbe rouge correspondant à la courbe moyenne associée à la forme $\mu(t)$.

Enfin, les sous-figures de la colonne de droite présentent respectivement les courbes $\{h_i\}$ réelles et leurs estimations produites par le processus en violant les contraintes temporelles fixées précédemment. La courbe rouge correspond à la courbe moyenne h_μ associée à la forme moyenne réelle a_μ ($f_0 = 1hz$). La courbe de couleur cyan correspond à l'estimation d'une

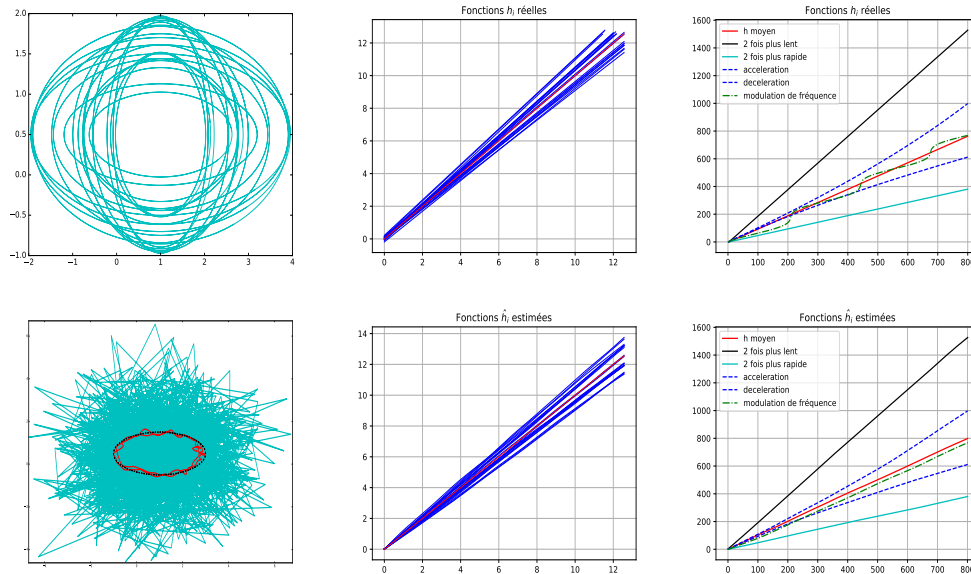


FIG. 1 – En haut à gauche, la variabilité d’amplitude des 20 formes d’ellipse non bruitées. En haut au centre, la variabilité temporelle des 20 fonctions temporelles h_i . En bas à gauche, les 20 séries bruitées de références, la forme moyenne exacte a_μ (courbe noire et tirets), et la forme moyenne estimée \hat{a}_μ (courbe rouge). En bas au centre, les 20 fonctions temporelles \hat{h}_i estimées. En haut à droite, 5 nouvelles fonctions temporelles associées à des ellipses bruitées n’ayant pas servies à l’estimation de la forme moyenne μ , avec en rouge, la fonction h_μ , et en bas à droite, les estimation de ces nouvelles fonctions temporelles et h_μ en rouge.

fonction h générée avec une fréquence double de f_0 ($f = 2hz$). La forme est donc parcourue deux fois plus vite et la pente attendue pour la courbe cyan est 2 fois plus faible que celle de la courbe rouge, ce qui est approximativement le cas sur la figure. La courbe noire correspond à l’estimation d’une fonction h générée avec une fréquence deux fois plus petite que f_0 ($f = 0,5hz$). La forme est donc parcourue deux fois moins vite et la pente attendue pour la courbe noire est 2 fois plus grande que celle de la courbe rouge, ce qui est approximativement le cas sur la figure. Les courbes bleues en pointillé correspondent à des parcours avec accélération (au dessous de la courbe rouge) et décélération (au dessus de la courbe rouge). Les formes de ces deux courbes ne sont plus linéaires, ce qui est également attendu. Enfin, la courbe verte en pointillé correspond à une courbe h modulée en fréquence, qui présente donc des accélérations et des décélérations. Les fluctuations ne sont pas très bien reproduites dans l’estimation de cette courbe, mais celles-ci sont malgré tout visibles dans l’estimation extraite.

Cet exemple synthétique, qui présente des variabilités temporelle et d’amplitude relativement importantes et couplées à un bruit additif couvrant largement la forme moyenne μ , montre que le principe de séparation des composantes d’amplitude et de phase tel que nous l’avons envisagé est possible. L’estimation du centroïde temporellement élastique permet, sur cet exemple, d’approcher le concept de forme moyenne et d’en déduire de manière assez pré-

Séparation de la forme et du temps

cise les fonctions de parcours temporel malgré le haut niveau de bruit présent dans les réalisations aléatoires du processus générateur.

3.2 Vérification de signatures en-ligne

Nous nous intéressons ici à quelques expériences sur des données réelles qui concernent l'identification de signatures dites "en-ligne", c'est à dire capturées via des dispositifs de capture tels que des tables à digitaliser, des stylets numériques ou encore des écrans tactiles. Les données se présentent sous la forme de trajectoires dans \mathbb{R}^d , avec $d \geq 2$ (positions x et y , pression, sur le stylet, orientations du stylet, etc.).

Les tâches que nous considérons sont des tâches de classification semi-supervisées qui peuvent s'apparenter à des tâches de détection d'anomalies : seules quelques signatures légitimes sont disponibles pour l'apprentissage, les signatures falsifiées pouvant être considérées comme des anomalies vis à vis de l'ensemble des signatures légitimes.

Algorithm 2 Vérification de signatures

Require: $\{o_i\}$, un ensemble de signatures légitimes, $\{o'_j\}$ un ensemble de signatures à tester.

$\hat{\mu}, \{\hat{h}_i\} \leftarrow \text{Algorithme1}(\{o_i\})$

$\{\hat{h}'_j\} \leftarrow \text{Eq. (9) appliquée sur } \hat{\mu} \text{ et } \{o'_j\}$

$\{\delta(o'_j, \hat{\mu})\} \leftarrow \text{Eq. (10), (11), 12 appliquées sur } \hat{\mu}, \{o'_j\} \text{ et } \{\hat{h}'_j\}$

L'approche développée ici consiste à appliquer le principe de séparation conformément à l'algorithme 2 sur l'ensemble d'apprentissage $\{o_i\}$ constitué de signatures légitimes, pour en extraire une forme moyenne $\hat{\mu}$ et les composantes temporelles \hat{h}_i associées. Il suffit ensuite d'aligner les signatures à tester $\{o'_j\}$ sur la forme moyenne $\hat{\mu}$ pour obtenir les composantes temporelles $\{\hat{h}'_j\}$. Enfin, nous évaluons un score global qui intègre une mesure de dissimilarité de forme et une mesure de dissimilarité temporelle.

La dissimilarité de forme entre la série o'_j et la forme moyenne $\hat{\mu}$ est estimée en évaluant la moyenne quadratique des écarts entre les échantillons $\{\hat{\mu}(t)\}$ de la forme moyenne et les espérances mathématiques des échantillons de la série o' alignés sur chacun des échantillons $\hat{a}_\mu(t)$, comme formalisé par l'équation Eq. (10).

$$\delta_f(o'_j, \hat{\mu}) = \frac{1}{|\hat{a}_\mu|} \sum_{t=1}^{|\hat{a}_\mu|} (\mathbb{E}(o'_j | \hat{a}_\mu(t)) - \hat{a}_\mu(t))^2 \quad (10)$$

La dissimilarité temporelle est composée du produit de deux termes, comme spécifié dans l'équation Eq. (11). Le premier terme évalue l'écart en absolu à l'unité de la pente de la droite de régression calculée pour la fonction temporelle \hat{h}'_j , associée à la série o'_j , normalisée par la pente moyenne \overline{Slope} des droites de régression des fonctions temporelles $\{\hat{h}_i\}$ évaluées sur les séries d'apprentissage $\{o_i\}$ qui correspondent donc à des signatures légitimes.

Le deuxième terme évalue l'écart en absolu à l'unité entre l'aire sous la courbe (AUC) de la fonction temporelle \hat{h}'_j associée à la série o'_j , normalisée par la moyenne des AUC calculées sur les fonctions temporelles $\{\hat{h}_i\}$ qui correspondent aux séries d'apprentissage, notée \overline{AUC} .

$$\delta_t(o'_j, \hat{\mu}) = \left(\frac{\text{Slope}(\hat{h}'_j)}{\text{Slope}} - 1\right)^2 \cdot \left(\frac{\text{AUC}(\hat{h}'_j)}{\text{AUC}} - 1\right)^2 \quad (11)$$

Le score global de dissimilarité fusionne les deux scores précédents conformément à l'équation (12). Le méta paramètre α permet d'ajuster le poids d'importance entre les dissimilarités temporelle et de forme, la fonction \log permettant d'aplatir le produit de ces dissimilarités. Cette méthode de séparation des composantes de forme et de temps, puis de fusion des scores de dissimilarité, est appelée SFT (séparation forme temps) dans la suite de l'article.

$$\delta(o'_j, \hat{\mu}) = \alpha \cdot \log(1 + \delta_f(o'_j, \hat{\mu})) + (1 - \alpha) \cdot \log(\delta_t(o'_j, \hat{\mu}) + 1) \quad (12)$$

A titre d'illustration, nous présentons en Fig. 2, un cas issu du jeu de données MOBISIG. Sur la sous-figure de gauche, la forme moyenne estimée $\hat{\mu}$ est représentée par la courbe 3D en bleu foncé, superposée aux courbes qui correspondent aux signatures légitimes utilisées pour l'apprentissage. Sur la sous-figure du centre, les composantes temporelles extraites des signatures légitimes sont présentées en couleur cyan tandis que les composantes temporelles extraites des signatures falsifiées sont présentées en couleur rouge. Sur cet exemple, la plupart des falsifications présentent une composante temporelle non linéaire et, comparativement aux signatures légitimes, celles-ci sont réalisées plus lentement. Enfin, sur la sous-figure de droite, les distributions 2D des scores de dissimilarité sont présentés, avec en abscisse les scores de dissimilarité temporelle et en ordonnée les scores de dissimilarité de forme. On peut observer que pour les signatures légitimes (nuage bleu cyan), la distribution des scores est bien plus compacte que la distribution des scores associés aux signatures falsifiées (nuage rouge). On observe également sur cet exemple que les deux types de scores permettent d'identifier une frontière séparatrice entre les deux classes meilleure que si les scores de dissimilarité étaient exploités séparément.

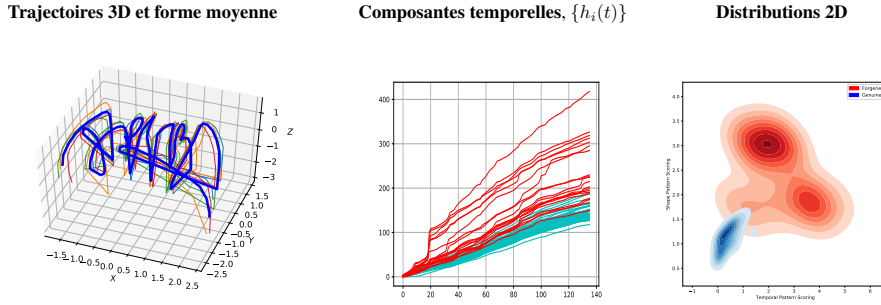


FIG. 2 – Exemple issu du jeu de données MOBISIG. A gauche : forme moyenne estimée en bleu gras superposée aux traces 3D légitimes dont elle est issue. Au centre : composantes temporelles extraites pour les signatures légitimes (courbes de couleur cyan) et falsifiées (courbes de couleur rouge). A droite : Distributions en 2D des scores de dissimilarité.

Pour les *benchmarks* considérés, les mesures d'évaluation utilisées sont les suivantes :

1. EER : (*equal error rate*) donne le taux d'erreur obtenu lorsque le taux de faux positifs est égal au taux de faux négatifs. Dans cette configuration, les taux sont évalués globalement, i.e. lorsque tous les sujets sont traités ensemble, sans distinction.

Séparation de la forme et du temps

2. aEER : (*average equal error rate*) donne le taux d'erreur moyen obtenu lorsque le taux de faux positifs est égal au taux de faux négatifs. Dans cette configuration, les taux sont évalués sujet par sujet puis moyennés.
3. FAR et FRR : taux de faux positif (False Acceptance Rate) et de faux négatifs (False Rejection Rate) réels obtenus au point de fonctionnement le plus proche du point EER.

Le jeu de données **SVC2004** (Yeung et al., 2004) constitue le premier *benchmark* développé spécifiquement pour évaluer les systèmes de vérification de signature en ligne. SVC2004 comporte deux tâches distinctes : la première exploite uniquement la trace 2D (x, y) produite sur une table à digitaliser Wacom Intuos (100hz) ainsi que l'information binaire identifiant le contact du stylet avec la table (*pen-up/pen-down*). La deuxième exploite des informations supplémentaires concernant l'orientation du stylo en particulier. Nous avons utilisé la première tâche pour déterminer une valeur acceptable pour le méta paramètre α (cf. Eq. (12)). Nous avons obtenu $\alpha = .85$ par validation croisée. Cette valeur est utilisée pour traiter la deuxième tâche ainsi que tous les autres *benchmark* décrits ci-dessous.

Le jeu de données **ICDAR2011** (Liwicki et al., 2011) est constitué de traces 3D (x, y, z), produites par 10 sujets et capturées via une table à digitaliser Wacom Intuos3 (200hz) et un stylet. Ici z représente la pression exercée sur le stylet.

Le jeu de données **MOBISIG** (Antal et al., 2018); ce jeu de données est constitué de pseudo-signatures produites par 83 sujets utilisant leur doigt et capturées avec un appareil mobile à écran tactile capacitif à une fréquence de $60hz$. Les trajectoires contiennent les positions x et y , les vitesses vx, vy , les accélérations ax, ay la pression exercée sur l'écran z , et la taille du contact entre le doigt et l'écran tc . Nous avons appliqué SFT sur les trajectoires contenant les descripteurs précédents exceptés les vitesses et les accélérations.

TAB. 1 – Résultats obtenus par jeux de données, mesures exprimées en %

<i>Benchmarks</i>	Méthodes	aEER	EER	FAR	FRR
SVC2004	EdIA, (Sharma et Sundaram, 2016)	2.71	-	-	-
	SFT	2.41	7.44	7.43	7.45
ICDAR2011	EdIA Dutch (Liwicki et al., 2011)	-	-	3.44	3.86
	SFT Dutch	0.77	3.18	3.09	3.27
	EdIA Chinese (Liwicki et al., 2011)	-	-	6.94	6.40
	SFT Chinese	1.12	3.44	3.69	3.20
MOBISIG (2017)	EdIA meilleur EER (Antal et al., 2018)	9.35	14.31	-	-
	EdIA meilleur aEER (Antal et al., 2018)	5.81	29.76	-	-
	SFT	4.19	5.86	5.86	5.86

Le tableau Tab. 1 présente de manière synthétique les résultats obtenus sur les trois *benchmark* exploités. Sur ces jeux de données qui couvrent quatre tâches distinctes, la méthode SFT se positionne légèrement au dessus de l'état de l'art (EdIA). Ces résultats sont d'autant plus encourageants que la nature des dispositifs de capture et l'origine culturelle des sujets (européens et asiatiques) sont relativement diversifiées. L'état de l'art est constitué de méthodes dérivées

de DTW pour SVC2004 et MOBISIG, et d'approches statistiques et heuristiques non détaillées pour ICDAR2011. Les approches neuronales, en particulier celles qui relèvent d'un apprentissage profond ou d'architectures récurrentes, se classent moins bien sur ces tâches du fait de la taille limitée des ensembles d'apprentissage et de la difficulté à les enrichir. Comme pressenti sur l'exemple de la figure 2, la complémentarité des similarités de forme et de temps permet à SFT de supplanter les approches *temporellement invariante* comme DTW par exemple. Enfin, comme l'indique la colonne aEER du tableau Tab. 1, des gains significatifs de performance sont atteints lorsque l'on adapte le seuil de décision à chaque sujet (cf. mesure aEER).

4 Conclusion

Nous avons abordé dans cet article le problème de la séparation des composantes de forme et de temps dans les séries temporelles. Si, dans sa formulation générale, ce problème est sous contraint, nous avons montré empiriquement, sur un cas synthétique constitué d'ellipses bruitées, qu'une heuristique construite autour de la notion de centroïde temporellement élastique permet d'apporter des éléments de solution à ce problème, proches des attendus théoriques¹. Nous avons enfin évalué l'intérêt de cette séparation dans le cadre d'une tâche d'identification de signatures en-ligne falsifiées. Sur trois *benchmarks* mettant en œuvre des dispositifs de capture de nature et de précision variables, l'approche proposée, basée sur une agrégation de scores de dissimilarité temporelle et de forme, améliore légèrement les résultats obtenus par les méthodes issues de l'état de l'art, ce qui est encourageant compte tenu de la simplicité du modèle et de son interprétabilité. Les perspectives de ce travail concernent l'étude et la modélisation des variabilités de forme et de temps dans les séries temporelles à des fins de reconnaissance ou de synthèse des formes. Les finalités applicatives sont diverses, on peut mentionner à titre d'exemple la génération de mouvements expressifs. Par ailleurs, l'algorithme présenté étant basé sur un alignement global, il n'est pas adapté au traitement de séries chaotiques (du fait de la dépendance exponentielle aux conditions initiales). Il conviendrait de proposer pour de telles séries une solution à base d'alignements localisés temporellement et spatialement.

Références

- Andreopoulos, A. et J. K. Tsotsos (2013). 50 years of object recognition : Directions forward. *Computer Vision and Image Understanding* 117(8), 827 – 891.
- Antal, M., L. Z. Szabó, et T. Tordai (2018). Online signature verification on mobisig finger-drawn signature corpus. *Mobile Information Systems* v, 1–15.
- Cuturi, M., J.-P. Vert, O. Birkenes, et T. Matsui (2007). A kernel for time series based on global alignments. In *IEEE ICASSP 2007*, Volume 2, pp. II–413–II–416.
- Gervini, D. et T. Gasser (2004). Self-modelling warping functions. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)* 66(4), 959–971.
- Just, W. et W. Just (1999). Computational complexity of multiple sequence alignment with sp-score. *Journal of Computational Biology* 8, 615–623.

1. Un code C++ et son API "wrapper" Python ainsi que le code Python de l'exemple synthétique sont accessibles sur : <https://github.com/pfmarteau/ShapeTimeSeparation/>.

- Liwicki, M., M. I. Malik, C. E. v. d. Heuvel, X. Chen, C. Berger, R. Stoel, M. Blumenstein, et B. Found (2011). Signature verification competition for online and offline skilled forgeries (sigcomp2011). In *Intern. Conf. on Document Analysis and Recognition*, pp. 1480–1484.
- Marteau, P.-F. (2019). Times series averaging and denoising from a probabilistic perspective on time-elastic kernels. *International Journal of Applied Mathematics and Computer Science* 29(2), 375–392.
- Marteau, P.-F. et S. Gibet (2014). On Recursive Edit Distance Kernels with Application to Time Series Classification. *IEEE Trans. on Neural Networks and Learning Systems*, 1–14.
- Niennattrakul, V., D. Srisai, et C. A. Ratanamahatana (2012). Shape-based template matching for time series data. *Knowledge-Based Systems* 26, 1 – 8.
- Panaretos, V. M. et Y. Zemel (2016). Amplitude and phase variation of point processes. *The Annals of Statistics* 44(2), 771–782.
- Ramsay, J. O. et X. Li (1998). Curve registration. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)* 60(2), 351–363.
- Sakoe, H. et S. Chiba (1971). A dynamic programming approach to continuous speech recognition. In *Proceedings of the Seventh International Congress on Acoustics, Budapest*, Volume 3, Budapest, pp. 65–69. Akadémiai Kiadó.
- Sangalli, L. M., P. Secchi, S. Vantini, et V. Vitelli (2010). k-mean alignment for curve clustering. *Computational Statistics and Data Analysis* 54(5), 1219 – 1233.
- Sharma, A. et S. Sundaram (2016). An enhanced contextual dtw based system for online signature verification using vector quantization. *Pattern Recogn. Lett.* 84(C), 22–28.
- Tang, R. et H.-G. Müller (2008). Pairwise curve synchronization for functional data. *Biometrika* 95(4), 875–889.
- Wang, L. et T. Jiang (1994). On the complexity of multiple sequence alignment. *Jour. of Comp. Biology* 1(4), 337–348.
- Yeung, D. Y., H. Chang, Y. Xiong, S. George, R. Kashi, T. Matsumoto, et G. Rigoll (2004). Svc2004 : First international signature verification competition. In *Proceedings of the Intern. Conf. on Biometric Authentication (ICBA), Hong Kong*, pp. 16–22. Springer.

Summary

In this article we address the problem of separation of shape and time components in time series. In statistics, this problem relates to the separation of amplitude and phase variations. The concept of shape that we tackle is termed temporally *neutral* to consider that it may possibly exist outside of any temporal specification, as it is the case for a geometric form. We propose to exploit and adapt a probabilistic temporal alignment algorithm, initially designed to estimate the centroid of a set of time series, to build some heuristic elements of solution to this separation problem. We show on some controlled synthetic data that this algorithm meets empirically the initial objectives. We finally evaluate it on real data, in the context of some on-line handwritten signature authentication tasks. On the evaluated benchmarks, our approach is positioned slightly above the current state of the art showing the applicative benefit of this separation problem.