Appariement difféomorphique robuste de faisceaux neuronaux

Pierre ROUSSILLON, Jean-Marc THIERY, Isabelle BLOCH, Pietro GORI

LTCI, Télécom ParisTech, Université Paris Saclay, Paris
pierre.roussillon@telecom-paristech.fr

Résumé – Dans cet article, nous nous intéressons à l'appariement difféomorphique de faisceaux du cerveau. Ce problème est complexe car deux faisceaux n'ont presque jamais le même nombre de fibres. Ces variations topologiques peuvent détériorer les déformations obtenues et fausser les interprétations cliniques. Nous proposons ici une méthode robuste aux différences topologiques, grâce à la combinaison de normes L^p et de métriques à noyaux adaptées à l'espace des fibres. Les résultats sur des exemples jouets et sur des données réelles sont prometteurs.

Abstract – In this article, we study the diffeomorphic matching of brain fibers bundles. This problem is complex because there is no topological correspondence between two bundles, namely two bundles do not have the same number of fibers. This strongly deteriorates the deformations obtained and can lead to a wrong clinical intepretation. We present a method that is robust to topological differences, by combining L^p norms and kernel-based metrics adapted to the space of fibers. We test the effectivenes of our method on both toy examples and brain fiber bundles.

Introduction

Contexte. Les méthodes récentes de tractographie à partir d'imagerie IRM de diffusion permettent de modéliser la cartographie des connexions entre les différentes zones du cerveau. Le tractogramme obtenu est un ensemble de polylignes 3D, appelées fibres, qui estiment la trajectoire de nombreux chemins neuronaux. Ces données sont d'un grand intérêt pour la caractérisation de certaines pathologies ou des études pré-opératoires. L'analyse de telles données et leur comparaison entre plusieurs individus nécessite d'apparier ces faisceaux de fibres. Le problème générique d'appariement consiste à trouver une déformation « optimale » qui rend les faisceaux de fibres sources Q proches des faisceaux cibles Q'. Ce problème est souvent formulé de manière variationelle sur les déformations φ [10] :

$$\min_{\varphi} J(\varphi) := E(\varphi) + A(\varphi(Q), Q'), \tag{1}$$

qui exprime le compromis entre la régularité de la déformation (énergie $E(\varphi)$), et la qualité de l'appariemment (distance résiduelle entre la cible Q' et la source déformée $\varphi(Q)$, mesurée par A). En imagerie médicale, les difféomorphismes sont très utilisés comme modèle de déformation, car ils ne détruisent pas l'organisation anatomique. En revanche, lorsqu'il y a un changement topologique (comme c'est le cas sur les données de fibres car il n'y a presque jamais le même nombre de fibres entre deux faisceaux), ce modèle peut créer des déformations irrégulières et peu plausibles anatomiquement. Cela peut fausser l'interprétation de l'appariement et les conclusions médicales. C'est notamment le cas lors de la présence de fibres aberrantes. Par fibres aberrantes, nous entendons ici des fibres qui ne trouvent pas d'appariement évident au cours de la déformation. Elles peuvent être le résultat d'un artefact lors de la génération des données, ou bien être le signe d'une variation anatomique due à une pathologie ou une opération chirurgicale.

Une piste de recherche pour que ces fibres abberantes ne perturbent pas trop le recalage serait alors d'enrichir le modèle de déformation pour prendre en compte un éventuel changement topologique, par exemple grâce à des modèles de *métamorphoses* [9] . Une autre possibilité est de construire des termes d'attache aux données *robustes*, qui pénalisent les fibres aberrantes, dans l'esprit de [5], ou des statistiques robustes [4]. C'est ce que nous proposons par la suite.

Contributions. La prise en compte des fibres aberrantes lors de l'appariement de faisceaux de fibres (voir figure 1) vient du coût quadratique utilisé pour le terme d'attache aux données, que ce soit une simple norme L^2 sur des fibres ré-échantillonnées pour avoir le même nombre de points, ou des normes à noyaux hilbertiens. Ce coût quadratique accroît l'importance des fibres aberrantes lors de la déformation. Nous proposons dans cet article un terme d'attache aux données robuste, de sorte que les fibres aberrantes ne perturbent pas la déformation obtenue. Pour cela, nous combinons des normes L^p, $p \in]0, 1[$ ($p = 0.1$ en pratique), avec des métriques à noyaux bien adaptées à l'espace des fibres. Nous obtenons alors une méthode mathématiquement bien posée, qui permet de réaliser des appariements robustes de fibres sans ajout de paramètres supplémentaires. La déformation optimale obtenue est plus régulière, et nous montrerons l'amélioration des résultats d'appariement sur des exemples jouets et des données réelles.

Méthodes

Modèle de déformations Nous proposons dans cet article le cadre des déformations difféomorphiques en utilisant le modèle appelé *Large Deformation Diffeomorphic Metric Mapping* (LDDMM) [10]. Dans ce modèle, les déformations sont générées par le flot de champs de vecteurs, variant avec le temps :

$t \in [0,1] \mapsto v_t$. Chaque v_t appartient à un espace de Hilbert de champs de vecteurs lisses V, et la norme $\|v_t\|_V$ représente le coût infinitésimal de déformation par v_t. Le problème (1) revient alors à résoudre sur l'ensemble de fonctions de carré intégrable $L^2([0,1],V)$ le problème :

$$\min_{v \in L^2([0,1],V)} \int_0^1 \|v_t\|_V^2 \, dt + A(\varphi_1^v(Q), Q') \qquad (2)$$

où φ_1^v est solution de l'équation de flot $\partial_t \varphi_t^v = v_t \circ \varphi_t^v$, $\varphi_0^v = \mathrm{Id}$ et où $\int_0^1 \|v_t\|_V^2 \, dt$ représente l'énergie totale de la déformation. *In fine*, grâce à une formulation hamiltonnienne et dans le cas discret, il est possible de montrer que cette énergie (et par suite, la déformation) ne dépend que d'impulsions initiales localisées aux sommets de la forme source. La minisation engendre alors un difféomorphisme de l'espace ambiant, qui réalise un compromis entre la régularité de la déformation et la qualité du recalage définie par A.

Distance *weighted varifold* entre les fibres Dans cette section, nous définissons un terme d'attache aux données A entre deux ensembles de fibres. Soit $Q = \{q_i, i = 1...n\}$, $Q' = \{q'_j, j = 1...m\}$ deux ensembles de fibres, où les q_i et q'_j sont des polylignes de segments (n'ayant pas forcément le même nombre de points pour chaque fibre). Parmi les distances les plus utilisées sur les faisceaux de fibres, beaucoup reposent sur une simple distance L^2 [6]. Au-delà du ré-échantillonnage nécessaire, ces distances ne possèdent aucun paramètre d'échelle. Or, dans certaines applications, il est intéressant de pouvoir décrire à quelle échelle caractéristique le problème est considéré, c'est-à-dire à quelle distance nous considérons que deux fibres sont loin l'une de l'autre. C'est pourquoi dans cet article nous nous intéressons plutôt aux métriques à noyaux et plus particulièrement au modèle des *varifolds non orientés* [3], qui est invariant par reparamétrisation et réorientation.

Nous rappelons dans cette partie le strict nécessaire pour la construction d'un produit scalaire sur les courbes au travers des varifolds. Nous renvoyons à [3] et les références qui y sont citées pour plus de détails. La motivation derrière la construction d'un tel produit scalaire est d'avoir une distance entre courbes qui prenne en compte à la fois la position spatiale relative de chaque courbe, mais aussi leur tangente (non orientée). L'idée des varifolds est donc de représenter les polylignes dans un espace produit (position $\times$ tangente). De plus, nous renforçons le poids des extrémités en utilisant les *weighted varifolds*, inspiré du modèle des *weighted currents* [2]. Deux courbes sont considérées comme similaires si elles ont des trajectoires *et* des extrémités proches.

Si l'on considère deux polylignes, $q = \cup_{i=1}^N C_i$ et $q' = \cup_{j=1}^M S_j$, comme des réunions de segments où chaque segment C_i (resp. S_j) est de longueur l_i (resp. l'_j), de milieu x_i (resp. y_j), de vecteur tangent unitaire $\vec{t_i}$ (resp. $\vec{t'_j}$), et d'etrémités respectives (x_a, x_b), (y_a, y_b), on définit le produit scalaire entre q et q' comme :

$$\langle q, q' \rangle_{wvar} = k_e(x_a, y_a) k_e(x_b, y_b) \sum_{i=1}^N \sum_{j=1}^M l_i l'_j k_p(x_i, y_j) \left\langle \vec{t_i}, \vec{t'_j} \right\rangle^2 \qquad (3)$$

où k_e et $k_p : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ sont des noyaux gaussiens : $k_p(x,y) = e^{-\frac{\|x-y\|^2}{\sigma^2}}$ et σ est un paramètre qui représente la distance caractéristique au-delà de laquelle deux points n'interagissent plus dans la métrique. Dans la suite, nous proposons d'utiliser comme terme d'attache aux données A :

$$A(Q, Q') = \sum_{i=1}^n \min_{j=1...n} \|q_i - q_j\|_{wvar}^2 \qquad (4)$$

où $\|\cdot\|_{wvar}$ est la norme sur les weighted varifolds, issue du produit scalaire défini par l'équation 3.

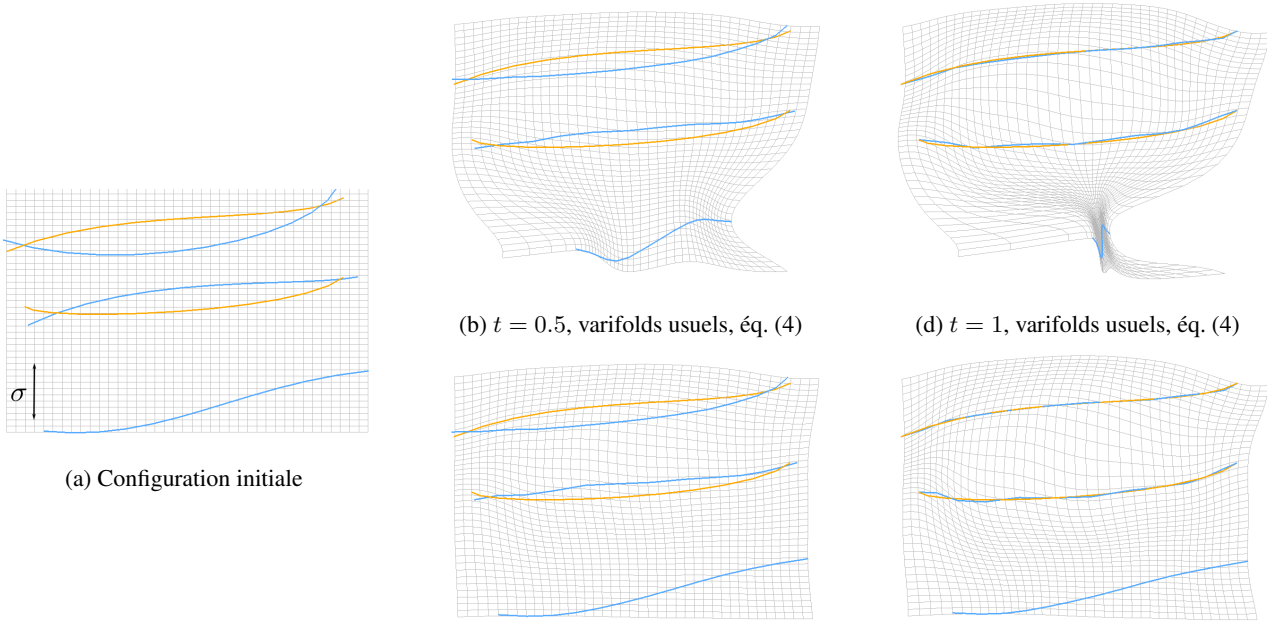
Limites La figure 1 présente un exemple synthétique simple de recalage de fibres dans le cadre décrit précédemment. Le problème d'appariement initial est mal posé : il y a trois fibres sources ($n = 3$), pour deux fibres cibles ($m = 2$). Dans cet exemple, le résultat obtenu est particulièrement mauvais : la fibre tout à gauche, notée q_1, qui peut être considérée comme aberrante dans ces données, se recroqueville complètement durant l'appariement. Pour expliquer cela, il faut revenir à la modélisation de la distance varifold. Sur la figure 1, la taille du noyau choisi est représentée et on vérifie que pour le produit scalaire, on a $\langle q_1, q'_j \rangle = 0$. En effet, q_1 est loin de toutes les autres fibres relativement à la taille du noyau. Ainsi, comme au cours de la déformation le modèle cherche à minimiser entre autres $\|q_1 - q'_j\|^2 = \|q_1\|^2 + \|q'_j\|^2 - 2\underbrace{\langle q_1, q'_j \rangle}_{=0}$, q'_j étant fixe, la seule possibilité est donc de minimiser $\|q_1\|^2$, ce qui se fait au travers du recroquevillement. Au-delà du problème visuel, ce phénomène fausse l'estimation de la déformation optimale, et cela empêche toute étude statistique ultérieure et donc toute analyse clinique. C'est pourquoi il est intéressant de développer un modèle qui seuille de façon lisse : si une fibre est aberrante, alors il vaut mieux ne pas la prendre en compte dans le problème. C'est l'objectif de la section suivante.

Distance robuste sur l'espace des fibres L'idée des méthodes robustes consiste à éviter la trop grande importance donnée aux valeurs aberrantes dans les termes quadratiques en remplaçant ces termes par des estimateurs robustes. Nous proposons ainsi de travailler avec des distances L^p, pour $p \in]0,1[$, comme cela a été proposé dans [5], ou [1] pour l'appariement rigide de nuages de points. Appliquée à notre espace de fibre, cette approche donne un nouveau terme A :

$$A(Q, Q') = \sum_{i=1}^n \min_{j=1,...,m} \left(\|q_i - q'_j\|_{var}^2 \right)^{p/2}. \qquad (5)$$

Théorème 1. *Il existe un minimiseur pour le problème d'appariement inexact (équation 2) avec A défini dans l'équation 5.*

La preuve est similaire à celle du thm. 37, [8] dans le cas des cycles normaux. Elle repose sur la continuité faible de $v \mapsto \sum_{i,j} \left(\|\varphi_1^v(q_i) - q'_j\|_{var}^2 \right)^{p/2}$. Cette dernière est bien faiblement continue car la $v \mapsto \|\varphi_1^v(q_i) - q'_j\|_{var}^2$ l'est, et $x \mapsto x^p$ est continue.



(a) Configuration initiale

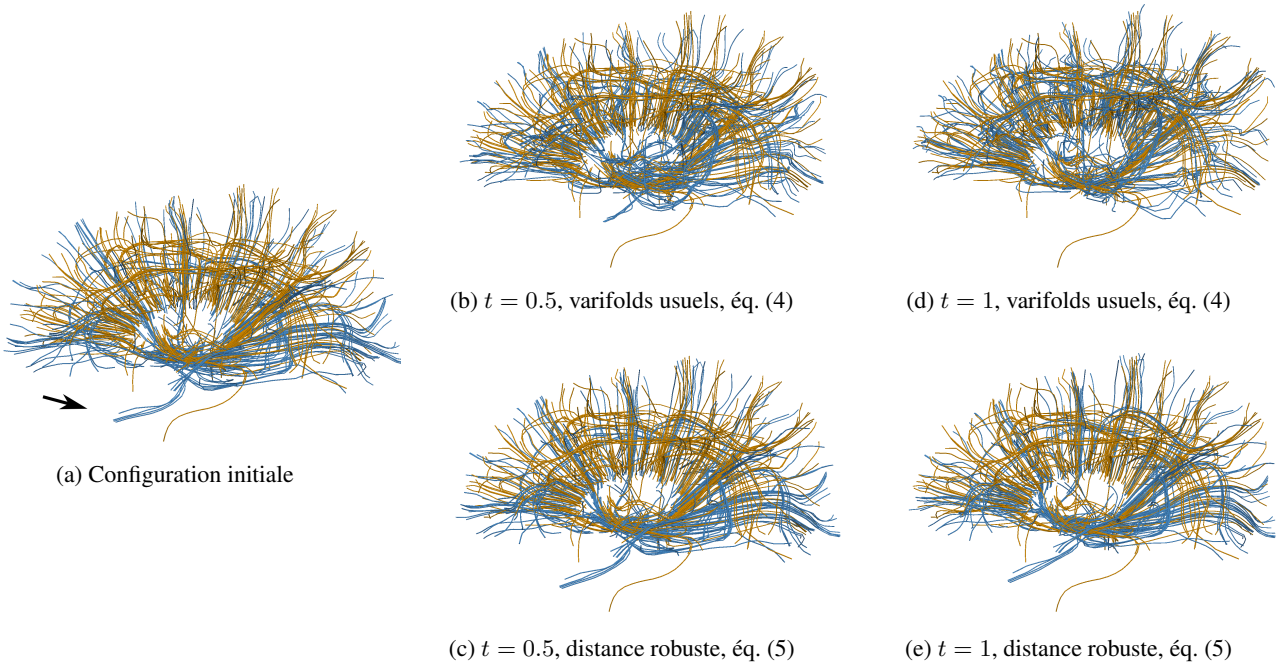
(b) $t = 0.5$, varifolds usuels, éq. (4)

(d) $t = 1$, varifolds usuels, éq. (4)

(c) $t = 0.5$, distance robuste, éq. (5)

(e) $t = 1$, distance robuste, éq. (5)

FIGURE 1 – Recalage de trois fibres (en bleu) sur deux fibres (en orange) avec un terme d'attache aux données varifolds usuel (équation 4) (b, d) et avec une distance varifolds robuste (équation 5) (c, e). En haut, la grille montre que la déformation est fortement perturbée par la fibre aberrante. En bas, la fibre aberrante n'est pas prise en compte durant le recalage, sans aucun pré-traitement ni nouveau paramètre. La taille du noyau σ est représentée en (a)



(a) Configuration initiale

(b) $t = 0.5$, varifolds usuels, éq. (4)

(d) $t = 1$, varifolds usuels, éq. (4)

(c) $t = 0.5$, distance robuste, éq. (5)

(e) $t = 1$, distance robuste, éq. (5)

FIGURE 2 – Recalage d'un faisceau neuronal source (bleu) vers un faisceau cible (orange). La ligne supérieure est obtenue avec un terme d'attache aux données varifolds. La ligne inférieur avec le terme robuste eq. (5), pour $p = 0.1$. Pour les deux recalages, la déformation est générée à partir d'un noyau gaussien, de paramètre $\sigma_V = 10mm$ et la taille du noyau k_p est $\sigma = 10mm$. La flèche noire indique l'ensemble de fibres qui sont recroquevillée avec la métrique varifold usuelle (ligne supérieure). Avec notre terme robuste, elles ne sont pratiquement pas prises en compte (ligne inférieure). La déformation finale est donc plus régulière.

Il est à noter que pour $p \in]0,1[$, nous aurions pu choisir aussi $A(Q,Q') = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(\left\|q_i - q'_j\right\|^2_{var}\right)^{p/2}$. En effet, minimiser A revient alors à minimiser pour chaque fibre source i la quantité $\sum_{j=1}^{m}\left(\left\|q_i - q'_j\right\|^2_{var}\right)^{p/2}$. Pour $p = 2$, minimiser la somme des écarts quadratiques revient à estimer la moyenne. Tel quel, ce terme forcerait donc l'appariement de toutes les fibres sources sur une fibre « moyenne » de la cible et c'est pourquoi pour $p = 2$ nous avons choisi de prendre le minimum sur j plutôt que la somme (équation 4). Pour $p \in]0,1[$ cependant, le modèle va favoriser la mise à zéro d'un des $\left\|q_i - q'_j\right\|_{wvar}$ (voir discussion dans [5], section 4). Intégré dans la fonctionnelle d'appariement, ce terme va donc favoriser le recalage de chaque fibre source vers la fibre cible la plus proche. Il semble donc identique de choisir le minimum sur j ou la somme sur j dans ce cas de figure. De plus, pour $p \in]0,1[$, la concavité de $x \mapsto x^p$ fait que les fibres sources aberrantes auront une contribution moindre dans le terme A, ainsi que dans son gradient. Cela amène à une prise en compte douce des fibres aberrantes lors de la procédure d'appariement.

Mise en œuvre. Pour résoudre numériquement le problème de recalage, la minimisation de la fonctionnelle est réalisée en PyTorch, avec A comme dans l'équation 5. Le gradient de la fonctionnelle se calcule sans peine en utilisant des librairies d'auto-différentiation (voir une explication plus détaillée dans [7], section 6). Il est alors possible de brancher n'importe quelle méthode de descente sur ce calcul de gradient. Nous utilisons une méthode L-BFGS pour ce problème non convexe.

Résultats

Nous présentons ici des recalages obtenus avec la minimisation de l'équation 2 avec A construit comme dans la section précédente. Pour l'exemple jouet qui avait motivé l'introduction de distances plus robustes, la figure 2, où $p = 0.1$, montre que cette nouvelle méthode est effectivement robuste aux fibres aberrantes : la fibre source de gauche ne trouvant pas d'appariement convaincant, elle est ignorée au cours de la minimisation et ne perturbe pas la déformation finale. C'est exactement ce genre de comportement que nous voulons lorsque le problème d'appariement est mal posé. En pratique, nous avons étudié les différents $p \in]0,1[$ possibles. La sensibilité à ce choix est assez faible, et $p = 0.1$ donne de bons résultats dans tous les cas de figure. C'est cette valeur qui est utilisée dans la littérature [1]. Pour réaliser la figure 2, deux sujets ont été choisis au hasard parmi la base de données de l'HCP [1]. Les tractogrammes des cerveaux entiers, sont estimés avec MRTrix [2], en utilisant un algorithme probabiliste (iFOD2) et un modèle de diffusion CSD. Les faisceaux neuronaux ensuite segmentés relient l'hémisphère gauche de la surface corticale et le putamen gauche. Comme nous l'expliquons dans la figure 2, et contrairement à la métrique varifold usuelle, notre méthode est robuste aux fibres aberrantes qui sont indiquées par une flèche noire.

1. https ://db.humanconnectome.org
2. http ://www.mrtrix.org

Conclusion

Après avoir expliqué le phénomène de recroquevillement de fibres aberrantes lors de recalages difféomorphiques, nous avons proposé une méthode inspirée des statistiques robustes afin de résoudre ce problème. Cette méthode est simple à implémenter et s'insère facilement dans de nombreux termes d'attache aux données. De plus, avec les bibliothèques de programmes d'auto-différentiation, cela ne rajoute pas de coût supplémentaire pour effectuer la minimisation de la fonctionnelle associée. Les premiers résultats (sur des exemples jouets et des données réelles) sont très convaincants. Pour aller plus loin, une étude plus approfondie est aussi nécessaire pour comprendre les garanties théoriques (test de robustesse, comportement du gradient en fonction du nombre de fibres aberrantes, comparaison systématique avec les normes L^2, ...). À plus long terme, il nous semble pertinent d'adapter ce travail pour pouvoir détecter automatiquement des changements topologiques, ce qui peut être très intéressant dans un contexte clinique.

Références

[1] S. BOUAZIZ, A. TAGLIASACCHI, AND M. PAULY, *Sparse iterative closest point*, SGP, 2013, pp. 113–123.

[2] P. GORI, O. COLLIOT, L. MARRAKCHI-KACEM, AND Y. WORBE ET AL., *A Prototype Representation to Approximate White Matter Bundles with Weighted Currents*, in MICCAI, LNCS, 2014, pp. 289–296.

[3] I. KALTENMARK, B. CHARLIER, AND N. CHARON, *A general framework for curve and surface comparison and registration with oriented varifolds*, in CVPR, 2017.

[4] R. A. MARONNA, R. D. MARTIN, V. J. YOHAI, AND M. SALIBIÁN-BARRERA, *Robust statistics : theory and methods (with R)*, Wiley, 2018.

[5] M. NIKOLOVA, *Minimizers of cost-functions involving nonsmooth data-fidelity terms. application to the processing of outliers*, J Num Ana, 40 (2002), pp. 965–994.

[6] E. OLIVETTI, G. BERTO, P. GORI, N. SHARMIN, AND P. AVESANI, *Comparison of distances for supervised segmentation of white matter tractography*, in IEEE PRNI, 2017, pp. 1–4.

[7] P. ROUSSILLON AND J. GLAUNÈS, *Representation of Surfaces with Normal Cycles. Application to Surface Registration*. working paper or preprint, Jan. 2019.

[8] P. ROUSSILLON AND J. GLAUNÈS, *Kernel Metrics on Normal Cycles and Application to Curve Matching*, SIAM J Imag Sciences, 9 (2016), pp. 1991–2038.

[9] A. TROUVÉ AND L. YOUNES, *Metamorphoses Through Lie Group Action*, Foundations of Computational Mathematics, 5 (2005), pp. 173–198.

[10] L. YOUNÈS, *Shapes and Diffeomorphisms*, Springer, 2010.