Available online at www.sciencedirect.com
ScienceDirect
C. R. Acad. Sci. Paris, Ser. I 347 (2009) 27–32
COMPTES RENDUS MATHEMATIQUE
http://france.elsevier.com/direct/CRASS1/

Mathematical Analysis

Extension to BV functions of the large deformation diffeomorphisms matching approach

François-Xavier Vialard [a,b,*], Filippo Santambrogio [a,b]

[a] *CMLA, ENS Cachan, CNRS, UniverSud, 61, avenue President Wilson, 94230 Cachan cedex, France*
[b] *CEREMADE, Université Paris-Dauphine, place de Lattre de Tassigny, 75775 Paris cedex 16, France*

Received 13 October 2008; accepted after revision 24 November 2008

Available online 18 December 2008

Presented by Yves Meyer

Abstract

The image matching within the framework of large deformations via diffeomorphisms is extended to the space of bounded variation functions. Thanks to a semi-differentiation lemma, which is the central new result of this article, we derive the geodesic equations for a general penalty term and we describe the associated momentum. ***To cite this article: F.-X. Vialard, F. Santambrogio, C. R. Acad. Sci. Paris, Ser. I 347 (2009).***

Résumé

Extension de l'appariement d'images par difféomorphismes aux fonctions à variation bornée. Le cadre de l'appariement d'images via des groupes de difféomorphismes est étendu à l'ensemble des fonctions à variation bornée. Le résultat principal de l'article est la différenciation du terme d'attache aux données de la fonctionnelle à minimiser. On établit ainsi les équations géodésiques associées à ce problème variationnel pour un terme d'attache général et on décrit le moment relatif à cette géodésique. ***Pour citer cet article : F.-X. Vialard, F. Santambrogio, C. R. Acad. Sci. Paris, Ser. I 347 (2009).***

Version française abrégée

L'appariement d'images par difféomorphismes a été récemment développé dans le but d'applications à l'anatomie computationnelle et l'imagerie médicale. Mathématiquement, on minimise une fonctionnelle avec un terme d'énergie associé à la déformation de l'espace ambiant et un terme d'attache aux données qui est le carré de la différence L^2 entre l'image initiale déformée et l'image cible :

$$\mathcal{J} = \frac{1}{2}\int_0^1 |v_t|_V^2 \,\mathrm{d}t + \frac{1}{\sigma^2}\left\| I_0 \circ \phi_1^{-1} - I_{\mathrm{targ}} \right\|_{L^2}^2, \tag{1}$$

* Corresponding author.
E-mail addresses: fxvialard@normalesup.org (F.-X. Vialard), filippo@ceremade.dauphine.fr (F. Santambrogio).

1631-073X/$ – see front matter © 2008 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.crma.2008.11.019

avec ϕ_1 le flot au temps 1 du champ de vecteur dépendant du temps $v_t \in L^2([0,1], V)$. On supposera que I_0 et I_{targ} sont des fonctions bornées à variation bornée. Le premier terme de la fonctionnelle $\mathcal{J}$ une fois minimisée est une distance (premier terme dans (2)) sur un groupe de difféomorphismses définie dans le paragraphe 2. L'espace des champs de vecteur est un espace de Hilbert V, tel qu'il existe une injection continue dans les champs de vecteurs C^1, ce qui assure l'existence du flot en tout temps. L'existence d'un minimum pour cette fonctionnelle s'obtient par un argument classique de semi-continuité inférieure. Pour obtenir plus d'information sur la géodésique, on différencie par rapport à une variation du champs de vecteur minimisant la fonctionnelle. Le résultat principal (Théorème 3.1) est la dérivation du terme d'attache aux données, qui est classiquement la différence L^2, mais qui peut être plus général (par exemple L^p pour $p > 1$). Lorque I_0 est suffisamment régulière ($I_0 \in H^1$), la dérivation est évidente. Un premier pas pour comprendre le cas de fonctions discontinues a été fait dans [3] pour le cas de fonctions caractéristiques d'un domaine du plan délimité par une courbe C^1 par morceaux. La preuve dans le cas des fonctions à variation bornée repose sur plusieurs réductions successives. Essentiellement, on démontre le résultat en dimension 1 en utilisant la décomposition d'une fonction BV en somme d'une fonction SBV et d'une fonction continue. On montre le résultat pour les fonctions Lipschitz par morceaux et on conclut par densité des fonctions Lipschitz par morceaux dans l'ensemble des fonctions SBV. Le passage en dimension quelconque se fait par l'utilisation du théorème du redressement du flot d'un champ de vecteur C^1 et de la décomposition de la distribution dérivée d'une fonction BV selon ses restrictions en dimension 1.

On explicite donc au paragraphe 4 la structure du moment associé au problème initial, qui est transporté par l'action du flot (voir l'Éq. (6)). Au cas d'une image $I_0 \in SBV$, le plus important pour les applications, est associé un moment décomposé en une partie dense (homogène à ∇I_0) et une partie singulière supportée par l'ensemble de saut de la fonction I_0. Cette décomposition permettra l'élaboration d'algorithmes numériques efficients prenant en compte les discontinuités des images médicales.

1. Introduction

In this Note, our aim is to give an overview with short proofs of the state of the art of the large deformation diffeomorphic approach for the case of images. This field has been widely studied since the first preliminary works by Grenander et al., and especially by A. Trouvé, L. Younes and M. Miller [6,4]. One of the most important and now widely developed application can be found in computational anatomy and medical imaging. The idea of this field is to match objects which can be deformed by the action of a diffeomorphism group of $\mathbb{R}^n$. These objects can be images, measures or group of points. A functional is minimized over the group of diffeomorphisms, the sum of a cost term and a penalty term, which is, in the case of image matching, the square of L^2 distance,

$$\mathcal{J} = \frac{1}{2} D(\mathrm{Id}, \phi)^2 + \frac{1}{\sigma^2} \left\| I_0 \circ \phi^{-1} - I_{\text{targ}} \right\|_{L^2}^2. \tag{2}$$

In order to get a proper distance on the diffeomorphism group, we need to work with sufficiently smooth vector fields, and the distance will be chosen right invariant.

If the template image I_0 is smooth, namely H^1 then the differentiation of $\mathcal{J}$ with respect to a small variation of the vector field can be performed and therefore we obtain the equation of a minimizing path which is a geodesic on the group of diffeomorphisms. What happens if I_0 is not sufficiently smooth? An attempt to answer this question was presented in [3] for the case of piecewise C^1 closed curves. We will present the derivation of the geodesic equation in the case of BV functions (functions of bounded variation). We will also extend our work to other penalty terms like for example the L^p norm for $p > 1$. And we will detail the structure of the momentum associated to these geodesic. (The momentum is to be understood as in the Hamiltonian formulation of the geodesic equation on a Riemannian manifold.)

The Note will be divided as follows. In the first part, we will briefly present the framework of large deformation via diffeomorphisms, the space of images and the space of vector fields. In the second part, we will establish the existence of a minimizer for the functional $\mathcal{J}$ and we will differentiate the functional. In the last part, we will derive the geodesic equation and enlight our result within the framework of large deformation via diffeomorphisms.

2. Basic framework: images and diffeomorphisms

Let U be a Lipschitz domain in $\mathbb{R}^n$, $BV(U)$ is the space of BV functions on U. In the following, we will work with $\mathrm{Im}(U) := BV(U) \cap L^\infty(U)$. We will extensively use in the next section the one-dimensional restriction of BV

functions. To fix the notations, let $f \in \mathrm{Im}(U)$, we denote by (f^+, f^-, ν) the precise representative of f and J_f is the jump set of f. As a BV function, we write the distributional derivative of f, $Df = \nabla f + D^c f + j(f)(x)\mathcal{H}^{n-1} \llcorner J_f$. The gradient ∇f is the absolutely continuous part of the distributional derivative with respect to the Lebesgue measure and $D^c f$ is the Cantor part of the derivative. The jump part is written as $j(f)(x) = (f^+(x) - f^-(x))\nu_f(x)$.

Now, we turn to define the value of a function in $\mathrm{Im}(U)$ with respect to a vector field. Although not a standard one, this definition is straightforward and useful for the central result stated in the paper. In the definition below, a vector field is an application from U to $\mathbb{R}^n$ without any further assumption but to be measurable:

Definition 2.1. If X is a (measurable) vector field on U and $g \in \mathrm{Im}(U)$, we define g_X by $g_X(x) = g(x)$ if $x \notin J_g$. On J_g, we define $\mathcal{H}^{n-1}$ a.e.

- if $\langle \nu(x), X(x)\rangle > 0$, $g_X(x) = g^+(x)$,
- if $\langle \nu(x), X(x)\rangle < 0$, $g_X(x) = g^-(x)$,
- else $\langle \nu(x), X(x)\rangle = 0$, $g_X(x) = \frac{g^-(x)+g^+(x)}{2}$.

Hence, g_X lies in $\mathrm{Im}(U) \times L^\infty(J_g; \mathcal{H}^{n-1})$.

Remark 1. In order to make use of change of variables formulas, the action by a diffeomorphism ψ is given by

$$(g \circ \psi)_X \circ \psi^{-1} = g_{\mathrm{d}\psi(X \circ \psi^{-1})}.$$

We will describe the group of diffeomorphisms as usual with a reproducing kernel Hilbert space of vector fields. We suppose that V is a Hilbert space of C^1 diffeomorphisms which can be continuously embedded in $C^1(U, \mathbb{R}^n)$, i.e. there exists $c > 0$ such that for any vector field $v \in V$, we have

$$\|v\|_{1,\infty}^2 \leqslant c^2 \langle v, v\rangle_V.$$

Consider $H = L^2([0,1], V)$ the Hilbert space of square integrable time vector fields, then using Gronwall lemma, it has been proved in [6] that the flow of such a vector field exists for all time and that under weak convergence in H, $v^n \rightharpoonup v$ then there is a uniform convergence of the flow on every compact sets: $\phi_1^n \mapsto \phi_1^n$ uniformly on every compact sets in U. The dual operator for the scalar product will be denoted by L, i.e. for the duality relation we have, $(Lv, v) = \langle v, v\rangle_V$.

Last, we define precisely the functional we aim to differentiate:

$$\mathcal{J} = \frac{1}{2} D(\mathrm{Id}, \phi_1)^2 + \int_U F\big(I_0 \circ \phi_1^{-1}(x), I_{\mathrm{targ}}\big)\,\mathrm{d}x \tag{3}$$

with $(I_0, I_{\mathrm{targ}}) \in \mathrm{Im}(U)^2$. The distance D is a right invariant metric defined as follows,

$$D(\phi, \psi)^2 = D\big(\mathrm{Id}, \phi \circ \psi^{-1}\big)^2 = \inf\left\{ |v|_{H^2}^2 = \int_0^1 |v_t|^2\,\mathrm{d}t \;\middle|\; \Phi_1 = \phi \circ \psi^{-1} \right\}, \tag{4}$$

with $v \in H$ and Φ_1 the flow at time 1 of the time dependent vector field v. The penalty term involves the function $F : \mathbb{R}^2 \mapsto \mathbb{R}$ which will be locally Lipschitz and C^1 in the first variable and such that $F(0,0) = 0$ (hypothesis which is required in our derivation result).

3. Differentiation of the functional

The existence of a minimizer for the functional (3) is well known and use the compactness of bounded balls in H for the weak topology. The proof relies on the lower semi-continuity of the functional with respect to $v \in H$. Thanks to a convexity argument, the first term is weakly lower semi-continuous. For the penalty term, thanks to the uniform convergence of the flow (presented in Section 2) and by means of a smooth approximation, we also get the lower semi-continuity. Remark that we do not need any regularity assumption on both functions I_0, I_{targ}. A sufficient condition would be that I_0 and I_{targ} are both in L^∞. What was not known until now is the differentiation of the functional with discontinuities in both images. The first step was done for the case of piecewise C^1 closed curves in the plane in [3]. That work contains an important lemma (see Appendix A, Lemma A.1) which enables the statement of our result.

Returning to the differentiation, the first term gives, in dual notation, $\int_0^1 (Lv_t, u_t)\,\mathrm{d}t$. The penalty term is much more difficult to differentiate and in fact, it is only semi-differentiable. We state hereunder the main theorem. Note that the flow at time 1 can be differentiated with respect to the vector field (see Lemma A.2 in Appendix A).

Theorem 3.1. *Let F be a locally Lipschitz function $F:\mathbb{R}^2 \mapsto \mathbb{R}$ and C^1 in the first variable such that $F(0,0)=0$, $(f,g)\in \mathrm{Im}(U)^2$, X a Lipschitz time dependent vector field C^1 in space and ϕ_t its associated flow. Defining, $J_t(f,g)=\int_{\mathbb{R}^n} F(f\circ\phi_t^{-1}(x), g(x))\,\mathrm{d}x$, then we have*

$$\partial_{t=0^+} J_t = \int \bigl\langle \partial_1 F\bigl(f(x), g_{X_0}(x)\bigr), -X_0\bigr\rangle \,\mathrm{d}x,$$

where $\partial_1 F(f, g_{X_0})$ is a part of the BV derivative of $F(f,g)$, defined by $\partial_1 F(f(x),l) = \nabla_1 F(f(x),l)(\nabla f(x) + D^c f(x)) + j_F(x)\mathcal{H}^{n-1}\llcorner J_f$, $j_F(x) = (F(f^+(x),l) - F(f^-(x),l))\nu_f(x)$.

Remark 1. Had we assumed $I_0 \in H^1$ in the functional (3), we could have dealt with $I_{\mathrm{targ}} \in L^2$. The weaker the regularity assumptions on I_0, the stronger are the assumptions on I_{targ}. Hence, we need I_{targ} to be smooth enough in order to give a sense to the dual pairing with the distributional derivative of $I_0 \in BV$. It turns out that functions in BV gives a natural pairing (not to be understood here as dual) with a distributional derivative in BV.

Sketch of the proof. The proof will follow three reductions:

• *First reduction*

It is sufficient to prove the result for autonomous vector fields (which do not depend on the time variable). Comparing the penalty term for the constant vector field X_0 and the initial one, we get the result with the estimation in $o(t)$ of the distance between the flows.

• *Second reduction*

It is sufficient to prove the result for $F(x,y)=xy$. As $\mathrm{Im}(U)$ is an algebra since the functions are bounded, the formula is easily true for polynomial functions. Then, approximating $\nabla_1 F$ on the first coordinate by a polynomial function, we get the result with the following control if f is C^1

$$\bigl|J_t(f,g) - J_0(f,g)\bigr| \leqslant \mathrm{Lip}_1(F)\int_U \bigl|f\circ\phi_t^{-1} - f\bigr|\,\mathrm{d}x \leqslant \mathrm{Lip}_1(F)\|X\|_\infty \int_0^t\!\!\int_U \bigl|\nabla\bigl(f\circ\phi_s^{-1}\bigr)\bigr|\,\mathrm{d}x\,\mathrm{d}s.$$

Here, $\mathrm{Lip}_1(F)$ is the Lipschitz constant of F on $\{(f(x),g(y)) \mid (x,y)\in U^2\}$. The inequality is also valid for $f\in BV$ by approximation. By a change of variable, we get the result. Remark the sup norm of the vector field on a small time interval $[0,t_0]$ in the bound, so that the contribution of the equilibrium points $(X^{-1}(\{0\}))$ to the differentiation result is null. In the case of the product xy, we have the following estimation: for $t\leqslant t_0$, if $g\in \mathrm{Im}(U)$ and $f\in BV(U)$ (also true if $g\in L^\infty(U)$), we have

$$\bigl|J_t(f,g) - J_0(f,g)\bigr| \leqslant Ct\|g\|_\infty\|f\|_{BV}.$$

We emphasize the continuity of this result with respect to the sup norm for g and the BV norm for f. By a change of variable, we have the same result switching the role f and g.

• *Third reduction*

Remark first that if the result is true for $F(x,y)=xy$ then it is true if g is an uniform limit in $\mathrm{Im}(U)$. We claim that it is sufficient to prove the result for the one-dimensional case.

As we pointed it out above, we only need to focus on points x such that $X(x)\neq 0$. By the flow-box theorem, we obtain trough a change of variables, $J_t = \int_{\mathbb{R}^n} f\circ\psi(x-tv) g\circ\psi\,\mathrm{Jac}(\psi)\,\mathrm{d}x$. Remark that we have to deal with the Jacobian of ψ, we need to assume that it is continuous in order to apply the result we prove in one dimension. This is allowed thanks to our assumptions on V. Hence $g\circ\psi\,\mathrm{Jac}(\psi)$ lies in the closure of $\mathrm{Im}(U)$ under the uniform norm. We then use the theorem (3.108 in [1]) which exhaustingly explains the behaviour of the one-dimensional restrictions of a BV function and the dominated convergence theorem to conclude.

Now, the conclusion in one dimension is easier. However it appeared that we are not able to prove it without the use of a density argument: this led in [5] to a more general proof which is interesting in itself, due to its higher dimensional taste and to weaker assumptions on the vector fields. In dimension 1, integrating the derivatives, a BV function is the sum of an SBV function and a continuous BV function. It is not difficult to prove that piecewise Lipschitz functions are dense in SBV. Actually, we can prove an equivalent statement in any dimension. Our result is continuous with respect to the BV norm of f. Hence we first treat the case of two piecewise C^1 functions which is straightforward. If f is in BV and g is C^1, the very definition of BV derivative gives the result. As the result is continuous for the sup norm of g, we get the result for $f \in SBV$ and $g \in BV$. To complete the proof, we have to deal with the case $f \in BV \cap C^0$ and $g \in SBV$. Switching the role of the two functions through a change of variable and then integrating by part, we obtain the result. $\square$

4. Geodesic equations and conclusion

Let us define the adjoint for ϕ a diffeomorphism of $\mathbb{R}^n$ and v a vector field on $\mathbb{R}^n$ by $\mathrm{Ad}_\phi v = (\mathrm{d}\phi v) \circ \phi^{-1}$. The conclusion of the previous section is that the differentiation of the functional gives, for a perturbation u_t of the minimizer v_t and $\delta u = -\int_0^1 \mathrm{Ad}_{\phi_{t,1}}(u)\,\mathrm{d}t$ (recall that the flow ϕ is associated to the minimizer v_t),

$$\int_0^1 \left[\langle v_t, u_t \rangle_V - \int \left\langle \mathrm{Ad}^*_{\phi_{t,1}} \partial_1 F\left(I_1, (I_{\mathrm{targ}})_{-\delta u}\right), u_t \right\rangle \right] \mathrm{d}t = 0. \tag{5}$$

In this equation one should notice that the distributional derivative in the second term belongs to the dual of V thanks to the smoothness of vector fields.

Now we could conclude that there exists $Z_t \in \overline{\mathrm{Conv}(\{\mathrm{Ad}^*_{\phi_{t,1}} \partial_1 F(I_1, (I_{\mathrm{targ}})_U) \mid u_t \in H\})}$ such that $\int_0^1 (Lv_t, u_t) - (Z_t, u_t)\,\mathrm{d}t = 0$, applying the first lemma in Appendix A. Yet to be more precise on the structure of Z_t, we need to apply the lemma on a different space in order to control Z_t with a stronger norm. If $\mu = \mathrm{d}x \otimes |D^c f| \otimes |j(x)| \mathcal{H}^{n-1} \llcorner J_f$, we consider Eq. (5) as a scalar product on $H \times L^2([0,1], L^2(\mu))$ and we apply the lemma. Since F is continuous, by the intermediate value theorem there exists $\tilde{I}_{\mathrm{targ}}$ which is a modification of the precise representative of I_{targ} only on the jump set J_f such that $\tilde{I}_{\mathrm{targ}}(x) \in [I^-_{\mathrm{targ}}, I^+_{\mathrm{targ}}]$ for which we have a.e. for $t \in [0,1]$:

$$Lv_t = \int \mathrm{Ad}^*_{\phi_{t,1}} \partial_1 F(I_1, \tilde{I}_{\mathrm{targ}}). \tag{6}$$

This equation shows the transportation of the momentum by the action of the flow which is similar in its expression to the smooth case. The structure of the momentum is given by the structure of the distributional derivative of f, which has an absolutely continuous part, a singular part on the jump set of I_0, and a Cantor part. The absolutely continuous part and the Cantor part of the momentum are just the restriction of a BV function to the sets involved. The Cantor part of the momentum behaves like the absolutely continuous one, and this is essentially due to the chain rule on BV functions.

In this work, we have detailed the structure of the momentum in the geodesic equations of the matching via diffeomorphisms for images as BV functions. Future work may essentially be devoted to develop numerical scheme to take into account this structure in the simpler case of SBV functions, which is sufficient for practical applications.

Acknowledgements

We warmly thank Alain Trouvé and Laurent Younes.

Appendix A. Lemmas

Lemma A.1. *Let E a Hilbert space, $H \subset E$ a vector space and B a non-empty bounded subset of E. Assume that for any $a \in H$, there exists $b_a \in B$ such that $\langle b_a, a \rangle \geqslant 0$. Then, there exists $b \in \overline{\mathrm{Conv}(B)}$ such that $\langle b, a \rangle = 0$, $\forall a \in H$.*

For a proof, refer to [7] or [3].

With the notation $\phi_{t,t'} = \phi_{t'} \circ \phi_t^{-1}$, we have the differentiation lemma (refer to [7] for a short proof, or [2] for more details)

Lemma A.2. *Let $(u_t, v_t) \in H$ be two time dependent vector fields on $\mathbb{R}^n$, and denote by $\phi_{0,t}^{\varepsilon}$ the flow generated by the vector field $u_t + \varepsilon v_t$, then we have*:

$$\partial_\varepsilon \phi_{0,1}(x) = \int_0^1 [\mathrm{d}\phi_{t,1}]_{\phi_{0,t}(x)} v\bigl(\phi_{0,t}(x)\bigr)\,\mathrm{d}t = \mathrm{d}\phi_1\left(\int_0^1 \mathrm{Ad}_{\phi_{t,0}}(v_t)\,\mathrm{d}t\right).$$

We can rewrite the formula to derive the expression used for the geodesic equations

$$\partial_\varepsilon\bigl[\phi_1 \circ \bigl(\phi_1^\varepsilon\bigr)^{-1}\bigr] = -\int_0^1 \mathrm{Ad}_{\phi_{t,1}}(v)\,\mathrm{d}t.$$

References

[1] L. Ambrosio, N. Fusco, D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, Oxford Mathematical Monographs, 2000.
[2] J. Glaunès, Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de forme et l'anatomie numérique, PhD Thesis, 2005.
[3] J. Glaunès, A. Trouvé, L. Younes, Modelling planar shape variation via Hamiltonian flows of curves, in: Statistics and Analysis of Shapes, Springer-Verlag, 2006.
[4] M.I. Miller, A. Trouvé, L. Younes, Geodesic shooting for computational anatomy, J. Math. Imaging Vis. (2006).
[5] F. Santambrogio, F.X. Vialard, A semi-derivation lemma on BV functions, oai:hal.archives-ouvertes.fr:hal-00317751_v1, 2008.
[6] A. Trouvé, Infinite dimensional group action and pattern recognition, 1995.
[7] F.X. Vialard, Hamiltonian approach to geodesic image matching, arXiv:0801.2095v1 [math.DC], 2008.