Well-posedness of eight problems of multi-modal statistical image-matching

Olivier Faugeras, Gerardo Hermosillo

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Well-posedness of eight problems of multi-modal statistical image-matching

Olivier Faugeras — Gerardo Hermosillo

N° 4235

THÈME 3

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4235--FR+ENG

Well-posedness of eight problems of multi-modal statistical image-matching

Olivier Faugeras , Gerardo Hermosillo

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Robotvis

Rapport de recherche n° 4235 — — 64 pages

Abstract: Multi-Modal Statistical Image-Matching techniques look for a deformation field that minimizes some error criterion between two images. This is achieved by computing a solution of the parabolic system obtained from the Euler-Lagrange equations of the error criterion. We prove the existence and uniqueness of a *classical* solution of this parabolic system in eight cases corresponding to the following alternatives. We consider that the images are realizations of spatial random processes that are either stationary or nonstationary. In each case we measure the similarity between the two images either by their mutual information or by their correlation ratio. In each case we regularize the deformation field either by borrowing from the field of Linear elasticity or by using the Nagel-Enkelmann tensor. Our proof uses the Hille-Yosida theorem and the theory of analytical semi-groups. We then briefly describe our numerical scheme and show some experimental results.

Key-words: *Multi-modal Image Matching, Variational Methods, Registration, Optical Flow, Mutual Information, Correlation Ratio, Euler-Lagrange equations, Initial-value problems, Maximal monotone operators, Strongly continuous semigroups of linear bounded operators, Analytical semi-groups of linear bounded operators.*

This work was partially supported by NSF grant DMS-9972228, EC grant Mapawamo QLG3-CT-2000-30161, INRIA ARCs IRMf and MC2, and Conacyt.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Résultats sur le caractère bien posé de huit problèmes de mise en correspondance multimodale d'images

Résumé : Les méthodes de mise en correspondance multimodale cherchent un champ de déformation qui minimise un critère d'erreur entre deux images. Ceci est accompli en calculant une solution du système d'EDP paraboliques obtenu à partir des équations d'Euler Lagrange du critère d'erreur. Nous démontrons existence et unicité de la solution pour ce système parabolique dans huit cas qui correspondent aux alternatives suivantes. Nous considérons que les images sont des réalisations de processus aléatoires spatiaux qui sont soit stationnaires soit non stationnaires. Dans chaque cas on mesure la similarité entre les deux images soit par l'information mutuelle, soit par le rapport de corrélation. Dans chaque cas nous régularisons le champ de déformation soit par un terme d'élasticité linéarisée, soit par de la diffusion anisotrope en utilisant le tenseur de Nagel-Enkelmann. Notre preuve utilise le théorème de Hille-Yosida. Nous décrivons ensuite brièvement la discrétisation des équations et nous montrons quelques résultats expérimentaux.

Mots-clés : *Mise en Correspondance Multimodale, Méthodes Variationnelles, Flot Optique, Information Mutuelle, Rapport de Corrélation, Équations d'Euler-Lagrange, Problèmes d'évolution, Opérateurs maximaux monotones, Semi-groupes d'opérateurs linéaires bornés fortement continus, Semi-groupes d'opérateurs linéaires bornés analytiques.*

Contents

1 Introduction

Image-matching techniques look for a deformation field that minimizes some error criterion between two images. In the variational framework, the deformation field is modeled as an element $\mathbf{h}$ of some functional space H, and the minimization is achieved by solving the Euler-Lagrange equations of the error criterion, which is a functional of the deformation field. In view of the difficulty of directly solving these equations, a gradient descent strategy is adopted, which in turn may be written as a parabolic system of functional equations. The solution is then taken as the asymptotic state (when t goes to infinity) of this system. In [4], we derived a variational approach to the multimodal matching problem by considering that the images are realizations of spatial random processes and taking as error criterion either their mutual information [12] or their correlation ratio [11]. We considered two classes of linear regularization terms corresponding to a linearised elasticity model [5] and to a contour preserving, geometry driven diffusion term based on the Nagel-Enkelmann tensor [9, 1]. The purpose of this paper is twofold. First, we generalize this approach by considering that the mutual information, or the correlation ratio, are functions of the space variable, and take as error criterion the integral of these functions over the image domain. This approach is well suited for dealing with nonstationarities of the statistical relation between intensity pairs. Second, we prove existence and uniqueness of a *classical* solution of the parabolic system of equations which are obtained in the eight cases which encompass the four statistical criteria and the two classes of linear regularization terms.

2 The framework

We are given two images I_1 and I_2, which we model as functions $I_1 : \mathbb{R}^n \rightarrow [0, \mathcal{A}]$ and $I_2 : \mathbb{R}^n \rightarrow [0, \mathcal{A}]$, where $[0, \mathcal{A}]$ is the closed interval (we limit ourselves to the cases $n = 2, 3$ and note $|\mathbf{x}|$ the Euclidean norm in $\mathbb{R}^n$). We make the weak assumption that these functions are zero outside the "square" $S \equiv [0, 1]^n$ and square-integrable, i.e. they belong to $L^2(\mathbb{R}^n)$. This assumption is completed by another one, i.e. that we observe regularised versions I_1^σ and I_2^σ of I_1 and I_2 by a gaussian with standard deviation equal to σ. This is realistic in terms of modelling (it is the scale-space idea []) and has several mathematical advantages. Let Ω be any open bounded set containing S (in particular we may require that Ω be regular, i.e. that its boundary $\partial\Omega$ be of class C^2): I_1^σ and I_2^σ are in the space $C^\infty(\Omega)$ of the infinitely continuously differentiable real functions on Ω; they are therefore, as all their derivatives, bounded on Ω; they are also, as all their derivatives, Lipschitz continuous on Ω.

The matching problem consists in finding a "regular" mapping $\mathbf{h} : \Omega \rightarrow \mathbb{R}^n$, such that the two images $I_2^\sigma(\mathbf{Id} + \mathbf{h})$ and I_1^σ minimize some criterion which is a functional of the mapping $\mathbf{h}$ ($\mathbf{Id}$ is the identity mapping of $\mathbb{R}^n$).

2.1 Functional spaces

The mapping $\mathbf{h}$ belongs to the functional space $H = \mathbf{L}^2(\Omega)$, where $\mathbf{L}^2(\Omega) = (L^2(\Omega))^n$.

We will consider in most parts of this paper $L^2(\Omega)$ as a real Hilbert space equipped with the Hilbert product

$$(h,k)_{L^2(\Omega)} = \int_\Omega h(\mathbf{x})k(\mathbf{x})\,d\mathbf{x},$$

which induces the norm

$$\|h\|_{L^2(\Omega)} = \left(\int_\Omega h(\mathbf{x})^2\,d\mathbf{x}\right)^{1/2}$$

$H = \mathbf{L}^2(\Omega)$ is also considered most of the time as a real Hilbert space equipped with the Hilbert product

$$(\mathbf{h},\mathbf{k})_H = (\mathbf{h},\mathbf{k})_{\mathbf{L}^2(\Omega)} = \int_\Omega \mathbf{h}(\mathbf{x})^T\mathbf{k}(\mathbf{x})\,d\mathbf{x} = \sum_{i=1}^n \int_\Omega h_i(\mathbf{x})k_i(\mathbf{x})\,d\mathbf{x}$$

which induces the product norm

$$\|\mathbf{h}\|_H = \|\mathbf{h}\|_{\mathbf{L}^2(\Omega)} = \left(\int_\Omega |\mathbf{h}(\mathbf{x})|^2\,d\mathbf{x}\right)^{1/2} = \left(\sum_{i=1}^n \int_\Omega h_i^2(\mathbf{x})\,d\mathbf{x}\right)^{1/2}$$

The criterion that we minimize with respect to $\mathbf{h}$ is the sum of two terms, a data term, noted $\mathcal{J}[\mathbf{h}]$, measuring the dissimilarity of the two images and a regularization term, noted $\mathcal{R}[\mathbf{h}]$, which ensures that the mapping $\mathbf{h}$ is regular. This regularization term is a function φ of the Jacobian $D\mathbf{h}$ ($\mathbf{h}$ therefore belongs to a subspace of $\mathbf{H}^1(\Omega) = (H^1(\Omega))^n$ of H which we choose to be $\mathbf{H}_0^1(\Omega) = (H_0^1(\Omega))^n$, and

$$\mathcal{R}[\mathbf{h}] = \int_\Omega \varphi(D\mathbf{h}(\mathbf{x}))d\mathbf{x}$$

$H^1(\Omega)$ and $H_0^1(\Omega)$ are, in most parts of this paper, considered to be real Hilbert spaces equipped with the Hilbert product

$$(h,k)_{H^1(\Omega)} = \int_\Omega h(\mathbf{x})k(\mathbf{x})\,d\mathbf{x} + \int_\Omega \nabla h(\mathbf{x})^T\nabla k(\mathbf{x})\,d\mathbf{x},$$

$\mathbf{H}^1(\Omega)$ and $\mathbf{H}_0^1(\Omega)$ are also considered to be real Hilbert spaces equipped with the Hilbert product

$$(\mathbf{h},\mathbf{k})_{\mathbf{H}^1(\Omega)} = \int_\Omega \mathbf{h}(\mathbf{x})^T\mathbf{k}(\mathbf{x})\,d\mathbf{x} + \int_\Omega D\mathbf{h}(\mathbf{x}) : D\mathbf{k}(\mathbf{x})\,d\mathbf{x} = \sum_{i=1}^n \left(\int_\Omega h_i(\mathbf{x})k_i(\mathbf{x})\,d\mathbf{x} + \int_\Omega \nabla h_i(\mathbf{x})^T\nabla k_i(\mathbf{x})\,d\mathbf{x}\right),$$

where $D\mathbf{h}(\mathbf{x}) : D\mathbf{k}(\mathbf{x}) = \text{trace}(D\mathbf{h}(\mathbf{x})^T D\mathbf{k}(\mathbf{x}))$. This Hilbert product induces the product norm

$$\|\mathbf{h}\|_{\mathbf{H}^1(\Omega)} = \left(\int_\Omega \mathbf{h}(\mathbf{x})^2\,d\mathbf{x} + \int_\Omega D\mathbf{h}(\mathbf{x}) : D\mathbf{h}(\mathbf{x})\,d\mathbf{x}\right)^{1/2} = \left(\sum_{i=1}^n \int_\Omega h_i^2(\mathbf{x}),d\mathbf{x} + \int_\Omega \nabla h_i(\mathbf{x})^T\nabla h_i(\mathbf{x})\,d\mathbf{x}\right)^{1/2}$$

Despite the fact that the mappings that we consider are from Ω into $\mathbb{R}^n$ it will be convenient in section 7.2 to consider also mappings from Ω into $\mathbb{C}^n$. $\mathbf{L}^2_{\mathbb{C}}(\Omega) = H_{\mathbb{C}}$ is a complex Hilbert space equipped with the Hilbert product

$$(\mathbf{h}, \mathbf{k})_{H_{\mathbb{C}}} = (\mathbf{h}, \mathbf{k})_{\mathbf{L}^2_{\mathbb{C}}(\Omega)} = \int_\Omega \mathbf{h}(\mathbf{x})^T \mathbf{k}^*(\mathbf{x})\, d\mathbf{x} = \sum_{i=1}^n \int_\Omega h_i(\mathbf{x}) k_i^*(\mathbf{x})\, d\mathbf{x}$$

where * indicates the complex conjugate and which induces the product norm

$$\|\mathbf{h}\|_{H_{\mathbb{C}}} = \|\mathbf{h}\|_{\mathbf{L}^2_{\mathbb{C}}(\Omega)} = \left(\int_\Omega |\mathbf{h}(\mathbf{x})|^2\, d\mathbf{x} \right)^{1/2} = \left(\sum_{i=1}^n \int_\Omega |h_i(\mathbf{x})|^2\, d\mathbf{x} \right)^{1/2}$$

Similarly, $\mathbf{H}^1_{\mathbb{C}}(\Omega)$ and $\mathbf{H}^1_{0\mathbb{C}}(\Omega)$ are complex Hilbert spaces equipped with the Hilbert product

$$(\mathbf{h}, \mathbf{k})_{\mathbf{H}^1_{\mathbb{C}}(\Omega)} = \int_\Omega \mathbf{h}(\mathbf{x}) \cdot \mathbf{k}^*(\mathbf{x})\, d\mathbf{x} + \int_\Omega D\mathbf{h}(\mathbf{x}) : D\mathbf{k}^*(\mathbf{x}))\, d\mathbf{x} + \equiv$$
$$(\mathbf{h}, \mathbf{k})_{\mathbf{L}^2_{\mathbb{C}}(\Omega)} + \int_\Omega D\mathbf{h}(\mathbf{x}) : D\mathbf{k}^*(\mathbf{x})\, d\mathbf{x}, \quad (1)$$

which induces the product norm

$$\|\mathbf{h}\|^2_{\mathbf{H}^1_{\mathbb{C}}(\Omega)} =$$
$$\|\mathrm{Re}(\mathbf{h})\|^2_{\mathbf{L}^2(\Omega)} + \|\mathrm{Im}(\mathbf{h})\|^2_{\mathbf{L}^2(\Omega)} + \sum_{i=1}^n \left(\|\mathrm{Re}(\nabla h_i)\|^2_{\mathbf{L}^2(\Omega)} + \|\mathrm{Im}(\nabla h_i)\|^2_{\mathbf{L}^2(\Omega)} \right) =$$
$$\|\mathrm{Re}(\mathbf{h})\|^2_{\mathbf{H}^1(\Omega)} + \|\mathrm{Im}(\mathbf{h})\|^2_{\mathbf{H}^1(\Omega)}$$

For technical reasons, we will also have to consider in section 9.2 the functional space $K = \mathbf{L}^\infty(\Omega) = (L^\infty(\Omega))^n$. $L^\infty(\Omega)$ is a real Banach space equipped with the norm

$$\|h\|_{L^\infty(\Omega)} = \inf\{c; |h(x)| \leq c \quad \text{a.e.} \quad \text{in} \quad \Omega\}.$$

Similarly $\mathbf{L}^\infty(\Omega)$ is a real Banach space equipped with the product norm

$$\|\mathbf{h}\|_{\mathbf{L}^\infty(\Omega)} = sup_{i=1,\cdots,n} \|h_i\|_{L^\infty(\Omega)}$$

We have the useful

Lemma 1 *We have the following continuous imbedding:*

$$\mathbf{L}^\infty(\Omega) \subset \mathbf{L}^2(\Omega),$$

or equivalently

$$K \subset H$$

Proof : If $\mathbf{h} \in \mathbf{L}^\infty(\Omega)$ we have, according to the definitions

$$|h_i(\mathbf{x})| \leq \|\mathbf{h}\|_{\mathbf{L}^\infty(\Omega)} \quad \text{a.e.} \quad i = 1, \cdots, n$$

Therefore

$$\|\mathbf{h}\|_{\mathbf{L}^2(\Omega)} = \left(\sum_{i=1}^n \int_\Omega |h_i(\mathbf{x})|^2 \, d\mathbf{x} \right)^{1/2} \leq |\Omega| \sqrt{n} \|\mathbf{h}\|_{\mathbf{L}^\infty(\Omega)}.$$

□

2.2 The minimization problem and the associated initial value problem

In summary, we look for the minima of the following functional of the mapping $\mathbf{h}$:

$$\mathcal{I}[\mathbf{h}] = \mathcal{J}[\mathbf{h}] + \kappa \int_\Omega \varphi(D\mathbf{h}(x)) \, dx,$$

where κ is a positive weighting parameter, i.e.

$$\min_{\mathbf{h} \in H_0^1(\Omega)} \mathcal{I}[\mathbf{h}] = \min_{\mathbf{h} \in H_0^1(\Omega)} \left(\mathcal{J}[\mathbf{h}] + \kappa \int_\Omega \varphi(D\mathbf{h}(x)) \, dx \right) \tag{2}$$

The case $\kappa = 0$ corresponds to no regularization.

The search for minima is done by computing the first variation (sometimes also called the Euler-Lagrange equations or the infinitesimal gradient) $\nabla_{\mathbf{h}} \mathcal{I}[\mathbf{h}]$ of the functional $\mathcal{I}[\mathbf{h}]$ with respect to $\mathbf{h}$ and equating it to 0 to find the extrema. The set of equations

$$\nabla_{\mathbf{h}} \mathcal{I}[\mathbf{h}] = \mathbf{0}$$

is known as the Euler-Lagrange equations of the error criterion $\mathcal{I}$. They have to be completed with the boundary conditions

$$\mathbf{h} = \mathbf{0} \quad \text{on} \quad \partial\Omega$$

Rather than solving them directly in the unknown $\mathbf{h}$, a task which is usually impossible, one introduces time and a differentiable function, also noted $\mathbf{h}$ from the interval $[0, T]$ into H (we say that $\mathbf{h} \in C^1([0, T]; H)$) such that $\mathbf{h}(0)$ is an initial field $\mathbf{h}_0$ and we solve the following initial value problem:

$$\begin{cases} \dfrac{d\mathbf{h}}{dt} = -\nabla_{\mathbf{h}} \mathcal{I}[\mathbf{h}] \\ \mathbf{h}(0)(\cdot) = \mathbf{h}_0(\cdot), \end{cases} \tag{3}$$

i.e. we start from the initial field $\mathbf{h}_0$ and follow the infinitesimal gradient of the functional $\mathcal{I}$ (the - sign is because we are minimising). The question that we answer in this article is that of the existence and uniqueness of a solution to (3).

In our case $\nabla_{\mathbf{h}} \mathcal{I}$ is the sum of two terms, one corresponding to the dissimilarity functional $\mathcal{J}[\mathbf{h}]$ and one to the regularization term $\mathcal{R}[\mathbf{h}]$. For this term, the computation of $\nabla_{\mathbf{h}} \mathcal{R}[\mathbf{h}]$ is standard

$$\nabla_{\mathbf{h}} \mathcal{R}[\mathbf{h}] = -\mathbf{div}(\varphi_{D\mathbf{h}}),$$

where $\varphi_{D\mathbf{h}}$ is the derivative of φ with respect to $D\mathbf{h}$.

3 The regularization terms: convexity and coerciveness

We introduce two regularization functionals and show that they are coercive and convex. We recall that, for a functional

$$\mathcal{R}[\mathbf{h}] = \int_{\Omega} \varphi(D\mathbf{h}(\mathbf{x}))\, d\mathbf{x},$$

coerciveness means

$$\begin{cases} \text{There exist constants} \quad \beta > 0,\, \gamma \geq 0 \quad \text{such that} \\ \varphi(P) \geq \beta |P|^q - \gamma \\ \text{for all} \quad P \in \mathbb{M}^{n\times n} \end{cases}$$

$|P| = (\text{trace}(P^T P))^{1/2}$ is the usual matrix norm defined on the set $\mathbb{M}^{n\times n}$ of $n \times n$ matrices. In our case we use $q = 2$.

3.1 The linearized elasticity operator

Our first regularization operator arises from elasticity theory and is defined by the Saint-Venant/Kirchoff model [5] for which

$$\varphi_{El}(D\mathbf{h}) = \frac{\lambda}{8}\big(\,\text{trace}\,(D\mathbf{h}^T + D\mathbf{h})\big)^2 + \frac{\mu}{2}\,\text{trace}\,\big((D\mathbf{h}^T + D\mathbf{h})^2\big), \tag{4}$$

$\lambda \geq 0$ and $\mu > 0$ are constants called the Lamé coefficients.

Proposition 1 *The mapping*

$$\begin{array}{llll} \varphi_{El}: & \mathbb{M}^{n\times n} & \mapsto & \mathbb{R}^+ \\ & \mathbf{X} & \mapsto & \lambda\,\big(\mathit{trace}(\mathbf{X} + \mathbf{X}^T)\big)^2 + \mu\;\mathit{trace}\big((\mathbf{X} + \mathbf{X}^T)^2\big) \end{array}$$

is convex.

Proof :

We write φ_{El} as a quadratic form of the components X_k of $\mathbf{X}$,

$$\varphi_{El}(\mathbf{X}) = \sum_{i}^{n^2} \sum_{j}^{n^2} a_{ij}\; X_i\, X_j$$

and notice that the matrix a_{ij} has 3 (resp. 2) eigenvalues equal to $4(\lambda+\mu)$ and 6 (resp. 2) eigenvalues equal to $2\,\mu$ in the case $n = 3$ (resp. $n = 2$). The result follows from the fact that $\mu > 0$ and $\lambda \geq 0$ □

Proposition 2 *The functional*

$$\mathcal{R}_{El}[\mathbf{h}] = \int_{\Omega} \varphi_{El}(D\mathbf{h}(\mathbf{x}))\; d\mathbf{x},$$

is coercitive, i.e. $\exists\ c_1 > 0,\ c_2 \geq 0$ *such that:*

$$\varphi_{El}(D\mathbf{h}(\mathbf{x})) \geq c_1 |D\mathbf{h}|^2 - c_2$$

Proof : Clear from the the previous proposition if we choose c_1 equal to the smallest eigenvalue of φ_{El} and $c_2 = 0$. $\square$

3.2 The Nagel-Enkelmann tensor

Our second regularization operator is defined by functions φ of the form

$$\varphi_N(D\mathbf{h}) = \frac{1}{2}\,\mathrm{trace}\,(D\mathbf{h}\mathbf{T}_{I_1^\sigma} D\mathbf{h}^T), \tag{5}$$

where $\mathbf{T}_{I_1^\sigma}$ is a $n \times n$ symmetric matrix defined at every point of Ω by the following expression:

$$\mathbf{T}_f = \frac{(\lambda + |\nabla f|^2)\mathbf{Id} - \nabla f \nabla f^T}{(n-1)|\nabla f|^2 + n\lambda}, \quad \text{for} \quad f : \mathbb{R}^n \to R$$

This matrix was first proposed by Nagel and Enkelmann, [9] for optical flow computation and used more recently by Alvarez et al. [1].

We consider each of its scalar components since in this case this separation is possible. As pointed out in [1], $\mathbf{T}_{I_1^\sigma}$ has strictly positive eigenvalues.

Proposition 3 *The mapping*

$$\begin{array}{rccc} \varphi_N : & \mathbb{R}^n & \mapsto & \mathbb{R}^+ \\ & \mathbf{X} & \mapsto & \mathbf{X}\mathbf{T}_{I_1^\sigma}\mathbf{X}^T \end{array}$$

is convex.

Proof : Clear, since $\mathbf{T}_{I_1^\sigma}$ has strictly positive eigenvalues. $\square$

Proposition 4 *The functional*

$$\mathcal{R}_N[\mathbf{h}] = \int_\Omega \varphi_N(D\mathbf{h}(\mathbf{x}))\, d\mathbf{x},$$

is coercitive, i.e. $\exists\ c_1 > 0,\ c_2 \geq 0$ *such that:*

$$\varphi_N(D\mathbf{h}(\mathbf{x}) \geq c_1 |D\mathbf{h}|^2 - c_2$$

Proof : We have

$$\nabla u^T \mathbf{T}_{I_1^\sigma} \nabla u \geq \theta |\nabla u|^2 \quad \forall \mathbf{x} \in \Omega$$

Where $\theta > 0$ is the smallest eigenvalue of $\mathbf{T}_{I_1^\sigma}$. $\square$

4 The Dissimilarity terms

We analyse two classes of statistical similarity criteria between two images that we call global and local. Both classes are based upon the use of some estimates of the joint probability of the grey levels in the two images. This joint probability, noted $P_{\mathbf{h}}(i_1, i_2)$, is estimated by the Parzen window method []. It depends upon the mapping $\mathbf{h}$ since we estimate the joint probability distribution between the images $I_2^\sigma(\mathbf{Id} + \mathbf{h})$ and I_1^σ. To be compatible with the scale-space idea and for computational convenience, we choose a Gaussian window with variance $\beta > 0$ for the Parzen window. We use the notation $\mathbf{i} = [i_1, i_2]^T$ and note:

$$G_\beta(\mathbf{i}) = g_\beta(i_1)g_\beta(i_2) = \frac{1}{2\pi\beta}\exp\left(-\frac{|\mathbf{i}|^2}{2\beta}\right) = \frac{1}{\sqrt{2\pi\beta}}\exp(-\frac{i_1^2}{2\beta})\,\frac{1}{\sqrt{2\pi\beta}}\exp(-\frac{i_2^2}{2\beta})$$

Notice that G_β and all its partial derivatives are bounded and Lipschitz. We will need in the sequel the infinite norms $\|g_\beta\|_\infty$ and $\|g'_\beta\|_\infty$.

For conciseness, we also use the following notation when making reference to a pair of grey-level intensities at a point $\mathbf{x}$:

$$\mathbf{I}_{\mathbf{h}}(\mathbf{x}) = [I_1^\sigma(\mathbf{x}), I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x}))]^T$$

4.1 Definition of the global dissimilarity criteria

We note $X_{I_1^\sigma}$ the random variable whose samples are the values $I_1^\sigma(\mathbf{x})$ and $X_{I_2^\sigma,\mathbf{h}}$ the random variable whose samples are the values $I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x}))$.

The joint probability density distribution (pdf) of $X^g_{I_1^\sigma}$ and $X^g_{I_2^\sigma,\mathbf{h}}$ (the upper index g stands for global) is defined by the function $P_{\mathbf{h}} : [0, \mathcal{A}] \times [0, \mathcal{A}] \to [0, 1]$:

$$P_{\mathbf{h}}(\mathbf{i}) = \frac{1}{|\Omega|}\int_\Omega G_\beta(\mathbf{I}_{\mathbf{h}}(\mathbf{x}) - \mathbf{i})\; d\mathbf{x} \tag{6}$$

With the help of the estimate (6) we define the mutual information between the two images I_1^σ and I_2^σ, noted $\mathbf{MI}^g[\mathbf{h}]$, and the correlation ratio, noted $\mathbf{CR}^g[\mathbf{h}]$. In order to do this we need to introduce more random variables besides $X^g_{I_1^\sigma}$ and $X^g_{I_2^\sigma,\mathbf{h}}$. They are summarized in table 1.

We have introduced in this table the conditional law of $X^g_{I_2^\sigma,\mathbf{h}}$ with respect to $X^g_{I_1^\sigma}$, noted $P^g_{\mathbf{h}}(i_2|i_1)$:

$$P^g_{\mathbf{h}}(i_2|i_1) = \frac{P_{\mathbf{h}}(i_1, i_2)}{p^g(i_1)},$$

and the conditional expectation $\mathbf{E}[X^g_{I_2^\sigma,\mathbf{h}}|X^g_{I_1^\sigma}]$ of the intensity in the second image $I_2^\sigma(\mathbf{Id} + \mathbf{h})$ conditionally to the intensity in the first image I_1^σ. For conciseness, we note the value of this random variable $\mu_{2|1}(i_1, \mathbf{h})$, indicating that it depends on the intensity value i_1 and on the field $\mathbf{h}$. Similarly the conditional variance of the intensity in the second image conditionally to the intensity in the first image is noted $\mathbf{Var}[X^g_{I_2^\sigma,\mathbf{h}}|X^g_{I_1^\sigma}]$ and its value is abbreviated $v_{2|1}(i_1, \mathbf{h})$. The mean and variance

Random Variable	Value	PDF
$(X^g_{I_1^\sigma}, X^g_{I_2^\sigma,\mathbf{h}})$	(i_1, i_2)	$P_\mathbf{h}(i_1, i_2)$
$X^g_{I_1^\sigma}$	i_1	$p^g(i_1) = \int_\mathbb{R} P_\mathbf{h}(i_1, i_2)\, di_2$
$X^g_{I_2^\sigma,\mathbf{h}}$	i_2	$p^g_\mathbf{h}(i_2) = \int_\mathbb{R} P_\mathbf{h}(i_1, i_2)\, di_1$
$\mathbf{E}[X^g_{I_2^\sigma,\mathbf{h}} \vert X^g_{I_1^\sigma}]$	$\mu_{2\vert 1}(i_1, \mathbf{h}) \equiv \int_\mathbb{R} i_2 P^g_\mathbf{h}(i_2\vert i_1)\, di_2$	$p^g(i_1)$
$\mathbf{Var}[X^g_{I_2^\sigma,\mathbf{h}} \vert X^g_{I_1^\sigma}]$	$v_{2\vert 1}(i_1, \mathbf{h}) \equiv$ $\int_\mathbb{R} i_2^2 P^g_\mathbf{h}(i_2\vert i_1)\, di_2 - \mu_{2\vert 1}(i_1, \mathbf{h})^2$	$p^g(i_1)$

Table 1: Random variables: global case.

of the second image will also be used. Note that these are not random variables but that they are functions of $\mathbf{h}$:

$$\mu_2(\mathbf{h}) \equiv \int_\mathbb{R} i_2 p^g_\mathbf{h}(i_2)\, di_2 \tag{7}$$

$$v_2(\mathbf{h}) \equiv \int_\mathbb{R} i_2^2 p^g_\mathbf{h}(i_2)\, di_2 - (\mu_2(\mathbf{h}))^2 \tag{8}$$

The similarity measures can be written in terms of the quantities defined in table 1

$$\mathbf{MI}^g[\mathbf{h}] = \int_{\mathbb{R}^2} P_\mathbf{h}(\mathbf{i}) \log \frac{P_\mathbf{h}(\mathbf{i})}{p^g(i_1) p^g_\mathbf{h}(i_2)}\, d\mathbf{i}$$

$$\mathbf{CR}^g[\mathbf{h}] = 1 - \frac{\mathbf{E}[\mathbf{Var}[X^g_{I_2^\sigma,\mathbf{h}} | X^g_{I_1^\sigma}]]}{v_2(\mathbf{h})}.$$

Regarding the correlation ratio, and according to table 1:

$$\mathbf{E}[\mathbf{Var}[X^g_{I_2^\sigma,\mathbf{h}} | X^g_{I_1^\sigma}]] = \int_\mathbb{R} v_{2|1}(i_1, \mathbf{h}) p^g(i_1)\, di_1$$

The mutual information and the correlation ratio are positive and should be maximized with respect to the field $\mathbf{h}$. Therefore we propose the following

Definition 1 *The two global dissimilarity measures based on the mutual information and the correlation ratio are as follows:*

$$\mathcal{J}_{\boldsymbol{MI}^g}[\mathbf{h}] = -\boldsymbol{MI}^g[\mathbf{h}]$$

$$\mathcal{J}_{\boldsymbol{CR}^g}[\mathbf{h}] = -\frac{\mathbf{E}[\mathbf{Var}[X^g_{I_2^\sigma,\mathbf{h}}|X^g_{I_1^\sigma}]]}{v_2(\mathbf{h})} = \boldsymbol{CR}^g[\mathbf{h}] - 1$$

Note that this definition shows that the mappings $\mathbf{h} \rightarrow \mathcal{J}_{\mathbf{MI}^g}[\mathbf{h}]$ and $\mathbf{h} \rightarrow \mathcal{J}_{\mathbf{CR}^g}[\mathbf{h}]$ are not of the form $\mathbf{h} \rightarrow \int_\Omega L(\mathbf{h}(\mathbf{x}))\, d\mathbf{x}$, for some smooth function $L : \mathbb{R}^n \rightarrow \mathbb{R}$. Therefore the Euler-Lagrange equations will be slightly more complicated to compute than in this classical case.

4.2 Definition of the local dissimilarity criteria

An interesting generalisation of the ideas developed in the previous section is to make the estimator (6) local. This allows us to take into account nonstationarities in the distributions of the intensities. We weight our estimate (6) with a spatial Gaussian of variance $\gamma > 0$ centered at $\mathbf{x}_0$. This means that we for each point $\mathbf{x}_0$ in Ω we have two random variables, noted $X^l_{I_1^\sigma,\mathbf{x}_0}$ and $X^l_{I_2^\sigma,\mathbf{x}_0,\mathbf{h}}$ (the upper index l stands for local) whose joint pdf is defined by:

$$P_\mathbf{h}(\mathbf{i},\mathbf{x}_0) = \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} \int_\Omega G_\beta(\mathbf{I_h}(\mathbf{x}) - \mathbf{i}) G_\gamma(\mathbf{x}-\mathbf{x}_0)\, d\mathbf{x}, \tag{9}$$

where

$$G_\gamma(\mathbf{x}-\mathbf{x}_0) = \frac{1}{(\sqrt{2\pi\gamma})^n} \exp(-\frac{|\mathbf{x}-\mathbf{x}_0|^2}{2\gamma}),$$

and

$$\mathcal{G}_\gamma(\mathbf{x}_0) = \int_\Omega G_\gamma(\mathbf{x}-\mathbf{x}_0)\, d\mathbf{x} \leq |\Omega| G_\gamma(\mathbf{0}) \tag{10}$$

With the help of the estimate (9) we define at every point $\mathbf{x}_0$ of Ω the local mutual information between the two images I_1^σ and I_2^σ, noted $\mathbf{MI}^l[\mathbf{h}](\mathbf{x}_0)$, and the local correlation ratio, noted $\mathbf{CR}^l[\mathbf{h}](\mathbf{x}_0)$. In order to do this we need to introduce more random variables besides $X^l_{I_1^\sigma,\mathbf{x}_0}$ and $X^l_{I_2^\sigma,\mathbf{x}_0,\mathbf{h}}$. We summarise our notations and definitions in the table 2.

The similarity measures can be written in terms of the quantities defined in table 2

$$\mathbf{MI}^l[\mathbf{h}](\mathbf{x}_0) = \int_{\mathbb{R}^2} P_\mathbf{h}(\mathbf{i},\mathbf{x}_0) \log \frac{P_\mathbf{h}(\mathbf{i},\mathbf{x}_0)}{p^l(i_1,\mathbf{x}_0) p^l_\mathbf{h}(i_2,\mathbf{x}_0)}\, d\mathbf{i}$$

$$\mathbf{CR}^l[\mathbf{h}](\mathbf{x}_0) = 1 - \frac{\mathbf{E}[\mathbf{Var}[X^l_{I_2^\sigma,\mathbf{x}_0,\mathbf{h}}|X^l_{I_1^\sigma,\mathbf{x}_0}]]}{v_2(\mathbf{x}_0,\mathbf{h})}.$$

As in the global case, the mean and variance of the second image are also used. Note that they are not random variables but that they are functions of $\mathbf{x}_0$ and $\mathbf{h}$:

$$\mu_2(\mathbf{x}_0,\mathbf{h}) \equiv \int_\mathbb{R} i_2 p^l_\mathbf{h}(i_2,\mathbf{x}_0)\, di_2 \tag{11}$$

Random Variable	Value	PDF
$(X^l_{I_1^\sigma,\mathbf{x}_0}, X^l_{I_2^\sigma,\mathbf{x}_0,\mathbf{h}})$	(i_1, i_2)	$P_\mathbf{h}(i_1, i_2, \mathbf{x}_0)$
$X^l_{I_1^\sigma,\mathbf{x}_0}$	i_1	$p^l(i_1,\mathbf{x}_0) = \int_\mathbb{R} P_\mathbf{h}(i_1, i_2, \mathbf{x}_0)\, di_2$
$X^l_{I_2^\sigma,\mathbf{x}_0,\mathbf{h}}$	i_2	$p^l_\mathbf{h}(i_2,\mathbf{x}_0) = \int_\mathbb{R} P_\mathbf{h}(i_1, i_2, \mathbf{x}_0)\, di_1$
$\mathbf{E}[X^l_{I_2^\sigma,\mathbf{x}_0,\mathbf{h}} \vert X^l_{I_1^\sigma,\mathbf{x}_0}]$	$\mu_{2\vert 1}(i_1,\mathbf{x}_0,\mathbf{h}) \equiv \int_\mathbb{R} i_2 P^l_\mathbf{h}(i_2,\mathbf{x}_0 \vert i_1)\, di_2$	$p^l(i_1,\mathbf{x}_0)$
$\mathbf{Var}[X^l_{I_2^\sigma,\mathbf{x}_0,\mathbf{h}} \vert X^l_{I_1^\sigma,\mathbf{x}_0}]$	$v_{2\vert 1}(i_1,\mathbf{x}_0,\mathbf{h}) \equiv \int_\mathbb{R} i_2^2 P^l_\mathbf{h}(i_2,\mathbf{x}_0 \vert i_1)\, di_2 - \mu_{2\vert 1}(i_1,\mathbf{x}_0,\mathbf{h})^2$	$p^l(i_1,\mathbf{x}_0)$

Table 2: Random variables: local case.

$$v_2(\mathbf{x}_0,\mathbf{h}) \equiv \int_\mathbb{R} i_2^2 p^l_\mathbf{h}(i_2,\mathbf{x}_0)\, di_2 - (\mu_2(\mathbf{x}_0,\mathbf{h}))^2 \tag{12}$$

We define our similarity measures by aggregating the local ones:

$$\mathbf{MI}^l[\mathbf{h}] = \int_\Omega \mathbf{MI}^l[\mathbf{h}](\mathbf{x}_0)\, d\mathbf{x}_0 = \int_\Omega \int_{\mathbb{R}^2} P_\mathbf{h}(\mathbf{i},\mathbf{x}_0) \log \frac{P_\mathbf{h}(\mathbf{i},\mathbf{x}_0)}{p^l(i_1,\mathbf{x}_0) p^l_\mathbf{h}(i_2,\mathbf{x}_0)}\, d\mathbf{i}\, d\mathbf{x}_0$$

$$\mathbf{CR}^l[\mathbf{h}] = \int_\Omega \mathbf{CR}^l[\mathbf{h}](\mathbf{x}_0)\, dx_0 = \int_\Omega \left(1 - \frac{\mathbf{E}[\mathbf{Var}[X^l_{I_2^\sigma,\mathbf{x}_0,\mathbf{h}} \vert X^l_{I_1^\sigma,\mathbf{x}_0}]]}{v_2(\mathbf{x}_0,\mathbf{h})}\right) d\mathbf{x}_0$$

The mutual information and the correlation ratio are positive and should be maximized with respect to the field $\mathbf{h}$. Therefore we propose the following

Definition 2 *The two local dissimilarity measures based on the mutual information and the correlation ratio are as follows:*

$$\mathcal{J}_{\boldsymbol{MI}^l}[\mathbf{h}] = -\boldsymbol{MI}^l[\mathbf{h}]$$

$$\mathcal{J}_{\boldsymbol{CR}^l}[\mathbf{h}] = -\int_\Omega \frac{\mathbf{E}[\mathbf{Var}[X^l_{I_2^\sigma,\mathbf{x}_0,\mathbf{h}} \vert X^l_{I_1^\sigma,\mathbf{x}_0}]]}{v_2(\mathbf{x}_0,\mathbf{h})}\, d\mathbf{x}_0 = \boldsymbol{CR}^l[\mathbf{h}] - |\Omega|$$

Note that, just as in the global case, this definition shows that the mappings $\mathbf{h} \rightarrow \mathcal{J}_{\mathbf{MI}^l}[\mathbf{h}]$ and $\mathbf{h} \rightarrow \mathcal{J}_{\mathbf{CR}^l}[\mathbf{h}]$ are not of the form $\mathbf{h} \rightarrow \int_\Omega L(\mathbf{h}(\mathbf{x}))\, d\mathbf{x}$, for some smooth function $L : \mathbb{R}^n \rightarrow \mathbb{R}$. Therefore the Euler-Lagrange equations will be slightly more complicated to compute than in this classical case.

5 Existence of minimizers

In this section, we consider the existence of minimizers for (2). Our error criterion is the sum of a "classical" regularization term $\int_\Omega \varphi(D\mathbf{h}(\mathbf{x}))\, d\mathbf{x}$ where $\varphi : \mathbb{M}^{n\times n} \rightarrow \mathbb{R}$ is smooth and of a dissimilarity term which is only a function of $\mathbf{h}$ but cannot be written as $\int_\Omega L(\mathbf{h}(\mathbf{x}), \mathbf{x})\, d\mathbf{x}$. Since the proof of theorem 5 in section 8.2 of chapter 8 of [7] assumes that this is the case, we have to adapt it. Examining this proof we see we need to prove that the dissimilarity term $\mathcal{J}[\mathbf{h}]$ is continuous in H. This is proved in theorems 8 and 12 for the correlation ratio in the global and local cases, respectively. Therefore we have the

Theorem 1 *There exists at least one function $\mathbf{k} \in \mathbf{H}_0^1(\Omega)$ satisfying*

$$\mathcal{I}[\mathbf{k}] = \min_{\mathbf{h} \in \mathbf{H}_0^1(\Omega)} (\mathcal{J}[\mathbf{h}] + \kappa \mathcal{R}[\mathbf{h}]) \quad \kappa > 0$$

for $\mathcal{J}[\mathbf{h}] = \mathcal{J}_{\mathbf{CR}^g}[\mathbf{h}]$ (definition 1) or $\mathcal{J}[\mathbf{h}] = \mathcal{J}_{\mathbf{CR}^l}[\mathbf{h}]$ (definition 2) and $\mathcal{R}[\mathbf{h}] = \int_\Omega \varphi_{El}(D\mathbf{h}(\mathbf{x}))\, d\mathbf{x}$ (equation (4)) or $\mathcal{R}[\mathbf{h}] = \int_\Omega \varphi_N(D\mathbf{h}(\mathbf{x}))\, d\mathbf{x}$ (equation (5)).

Proof : This is a consequence of propositions 2 and 4, theorems 8 and 12 and theorem 5 in section 8.2 of chapter 8 of [7]. $\square$

In order to prove the same theorem in the case of the mutual information, we need to prove continuity. We have the following

Proposition 5 *Let $\mathbf{h}_n, n = 1, \cdots, \infty$ be a sequence of functions of H such that $\mathbf{h}_n \rightarrow \mathbf{h}$ almost everywhere in Ω then $\boldsymbol{MI}^g[\mathbf{h}_n] \rightarrow \boldsymbol{MI}^g[\mathbf{h}]$.*

Proof : Because I_2^g and g_β are continuous, $G_\beta(\mathbf{I}_{\mathbf{h}_n}(\mathbf{x}) - \mathbf{i}) \rightarrow G_\beta(\mathbf{I}_{\mathbf{h}}(\mathbf{x}) - \mathbf{i})$ a.e. in $\Omega \times \mathbb{R}^2$. Since $G_\beta(\mathbf{I}_{\mathbf{h}_n}(\mathbf{x}) - \mathbf{i}) \leq g_\beta(0)^2$ the dominated convergence theorem implies that $P_{\mathbf{h}_n}(\mathbf{i}) \rightarrow P_{\mathbf{h}}(\mathbf{i})$ for all $\mathbf{i} \in \mathbb{R}^2$. A similar reasoning shows that $p^g_{\mathbf{h}_n}(i_2) \rightarrow p^g_{\mathbf{h}}(i_2)$ for all $i_2 \in \mathbb{R}$. Hence, the logarithm being continuous

$$P_{\mathbf{h}_n}(\mathbf{i}) \log \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1) p^g_{\mathbf{h}_n}(i_2)} \rightarrow P_{\mathbf{h}}(\mathbf{i}) \log \frac{P_{\mathbf{h}}(\mathbf{i})}{p^g(i_1) p^g_{\mathbf{h}}(i_2)} \quad \forall \mathbf{i} \in \mathbb{R}^2$$

We next consider three cases to find an upper bound for $P_{\mathbf{h}_n}(\mathbf{i}) \log \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1) p^g_{\mathbf{h}}(i_2)}$:

$i_2 \leq 0$

This is the case where

$$0 \leq |i_2| \leq |i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}_n(\mathbf{x}))| \qquad \leq |i_2 - \mathcal{A}| \quad n \geq 1$$

Hence

$$g_\beta(i_2 - \mathcal{A}) \leq g_\beta(i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}_n(\mathbf{x}))) \qquad \leq g_\beta(i_2) \quad n \geq 1$$

This yields

$$\frac{g_\beta(i_2 - \mathcal{A})}{g_\beta(i_2)} \leq \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1)p^g_{\mathbf{h}}(i_2)} \leq \frac{g_\beta(i_2)}{g_\beta(i_2 - \mathcal{A})}$$

and

$$\left|\log \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1)p^g_{\mathbf{h}}(i_2)}\right| \leq \log \frac{g_\beta(i_2)}{g_\beta(i_2 - \mathcal{A})},$$

and therefore

$$P_{\mathbf{h}_n}(\mathbf{i}) \left|\log \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1)p^g_{\mathbf{h}}(i_2)}\right| \leq g_\beta(i_2)p^g(i_1) \log \frac{g_\beta(i_2)}{g_\beta(i_2 - \mathcal{A})}$$

The function on the righthand side is continuous and integrable in $\mathbb{R}\times] - \infty, \mathcal{A}]$.

$0 \leq i_2 \leq \mathcal{A}$

We have

$$0 \leq |i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}_n(\mathbf{x}))| \qquad \leq \mathcal{A} \quad n \geq 1$$

Hence

$$g_\beta(\mathcal{A}) \leq g_\beta(i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}_j(\mathbf{x}))) \qquad \leq g_\beta(0) \quad n \geq 1$$

This yields

$$\frac{g_\beta(\mathcal{A})}{g_\beta(0)} \leq \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1)p^g_{\mathbf{h}}(i_2)} \leq \frac{g_\beta(0)}{g_\beta(\mathcal{A})},$$

and

$$\left|\log \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1)p^g_{\mathbf{h}_n}(i_2)}\right| \leq \log \frac{g_\beta(0)}{g_\beta(\mathcal{A})},$$

and therefore

$$P_{\mathbf{h}_n}(\mathbf{i}) \left|\log \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1)p^g_{\mathbf{h}}(i_2)}\right| \leq g_\beta(0)p^g(i_1) \log \frac{g_\beta(0)}{g_\beta(\mathcal{A})}$$

The function on the righthand side is continuous and integrable in $\mathbb{R} \times [0, \mathcal{A}]$.

$i_2 \geq \mathcal{A}$

This is the case where

$$0 \leq i_2 - \mathcal{A} \leq i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}_n(\mathbf{x})) \qquad \leq i_2 \quad n \geq 1$$

Hence

$$g_\beta(i_2) \leq \quad g_\beta(i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}_n(\mathbf{x}))) \quad \leq g_\beta(i_2 - \mathcal{A}) \quad n \geq 1$$

This yields

$$\frac{g_\beta(i_2)}{g_\beta(i_2 - \mathcal{A})} \leq \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1)p^g_{\mathbf{h}}(i_2)} \leq \frac{g_\beta(i_2 - \mathcal{A})}{g_\beta(i_2)},$$

and

$$\left|\log \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1)p^g_{\mathbf{h}}(i_2)}\right| \leq \log \frac{g_\beta(i_2 - \mathcal{A})}{g_\beta(i_2)},$$

and therefore

$$P_{\mathbf{h}_n}(\mathbf{i}) \left|\log \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1)p^g_{\mathbf{h}}(i_2)}\right| \leq g_\beta(i_2 - \mathcal{A}) p^g(i_1) \log \frac{g_\beta(i_2 - \mathcal{A})}{g_\beta(i_2)}$$

The function on the righthand side is continuous and integrable in $\mathbb{R}\times]\mathcal{A}, +\infty]$.

The dominated convergence theorem implies that

$$\mathbf{MI}^g[\mathbf{h}_n] = \int_{\mathbb{R}^2} P_{\mathbf{h}_n}(\mathbf{i}) \log \frac{P_{\mathbf{h}_n}(\mathbf{i})}{p^g(i_1)p^g_{\mathbf{h}_n}(i_2)}\, d\mathbf{i} \rightarrow$$

$$\mathbf{MI}^g[\mathbf{h}] = \int_{\mathbb{R}^2} P_{\mathbf{h}}(\mathbf{i}) \log \frac{P_{\mathbf{h}}(\mathbf{i})}{p^g(i_1)p^g_{\mathbf{h}}(i_2)}\, d\mathbf{i}$$

□

We also have the

Proposition 6 *Let* $\mathbf{h}_n, n = 1, \cdots, \infty$ *be a sequence of functions of* H *such that* $\mathbf{h}_n \rightarrow \mathbf{h}$ *almost everywhere in* Ω *then* $\boldsymbol{MI}^l[\mathbf{h}_n] \rightarrow \boldsymbol{MI}^l[\mathbf{h}]$.

Proof : The proof is similar to that of proposition 5 □

Theorem 2 *There exists at least one function* $\mathbf{k} \in \mathbf{H}^1_0(\Omega)$ *satisfying*

$$\mathcal{I}[\mathbf{k}] = \min_{\mathbf{h} \in \mathbf{H}^1_0(\Omega)} (\mathcal{J}[\mathbf{h}] + \kappa \mathcal{R}[\mathbf{h}]) \quad \kappa > 0$$

for $\mathcal{J}[\mathbf{h}] = \mathcal{J}_{\boldsymbol{MI}^g}[\mathbf{h}]$ *(definition 1) or* $\mathcal{J}[\mathbf{h}] = \mathcal{J}_{\boldsymbol{MI}^l}[\mathbf{h}]$ *(definition 2) and* $\mathcal{R}[\mathbf{h}] = \int_\Omega \varphi_{El}(D\mathbf{h}(\mathbf{x}))\, d\mathbf{x}$ *(equation (4)) or* $\mathcal{R}[\mathbf{h}] = \int_\Omega \varphi_N(D\mathbf{h}(\mathbf{x}))\, d\mathbf{x}$ *(equation (5)).*

Proof : This is a consequence of propositions 2 and 4, propositions 5 and 6 and theorem 5 in section 8.2 of chapter 8 of [7]. □

6 The Euler-Lagrange equations

In this section we compute the Euler-Lagrange equations of the error criterion (2). As pointed out earlier, this is classical for the regularization terms and slightly more involved for the dissimilarity terms.

6.1 The first variation of the regularization terms

It is straightforward to verify that

$$\mathbf{div}(\varphi_{El_{D\mathbf{h}}}(D\mathbf{h})) = \mu\Delta\mathbf{h} + (\lambda+\mu)\nabla(\nabla\cdot\mathbf{h}),$$

We will find it useful to rewrite the righthand side in divergence form:

$$\mu\Delta\mathbf{h} + (\lambda+\mu)\nabla(\nabla\cdot\mathbf{h}) = \mathbf{div}(\lambda\,\mathrm{trace}\,(e(\mathbf{h}))\mathbf{Id} + 2\mu e(\mathbf{h})),$$

where we have noted

$$e(\mathbf{h}) = \frac{1}{2}(D\mathbf{h}^T + D\mathbf{h}), \tag{13}$$

and $\mathbf{div}(X) = [\sum_{j=1}^n \partial_j X_{1j}, \cdots, \sum_{j=1}^n \partial_j X_{nj}]^T$ which is equal to $[\mathrm{div}(X_1), \cdots, \mathrm{div}(X_n)]^T$, where X_i, $i = 1, \cdots, n$ is the ith row vector of the matrix $X : \Omega \rightarrow \mathbb{M}^{n\times n}$. We define the linear operator $A_1 : \mathcal{D}(A_1) \rightarrow H$ by

$$A_1\mathbf{h} = \mathbf{div}(\lambda\,\mathrm{trace}\,(e(\mathbf{h}))\mathbf{Id} + 2\mu e(\mathbf{h})) \tag{14}$$

The domain $\mathcal{D}(A_1)$ of the operator A_1 is a subspace of $\mathbf{L}^2(\Omega)$ defined in section 7.

In the case of the Nagel-Enkelman tensor, it is also straightforward to verify that

$$\mathbf{div}(\varphi_{N_{D\mathbf{h}}}(D\mathbf{h})) = \begin{pmatrix} \mathrm{div}(\mathbf{T}_{I_1^\sigma}\nabla h_1) \\ \vdots \\ \mathrm{div}(\mathbf{T}_{I_1^\sigma}\nabla h_n) \end{pmatrix},$$

and we define the linear operator $A_2 : \mathcal{D}(A_2) \rightarrow H$ by

$$A_2\mathbf{h} = \begin{pmatrix} \mathrm{div}(\mathbf{T}_{I_1^\sigma}\nabla h_1) \\ \vdots \\ \mathrm{div}(\mathbf{T}_{I_1^\sigma}\nabla h_n) \end{pmatrix} \tag{15}$$

The domain $\mathcal{D}(A_2)$ of the operator A_2 is a subspace of $\mathbf{L}^2(\Omega)$ defined in section 7.

6.2 The first variation of the dissimilarity terms

In [4], we have given the expressions of the infinitesimal gradients $\nabla_{\mathbf{h}}\mathcal{I}[\mathbf{h}]$ of the global criteria. We reproduce this result here without proof, the proof being found in [8], since we will need it later.

Theorem 3 *In the global case, the infinitesimal gradient is given by*

$$F^g(\mathbf{h})(\mathbf{x}) = \nabla_{\mathbf{h}}\mathcal{J}[\mathbf{h}](\mathbf{x}) = (G_\beta \star L^g_{\mathbf{h}})(I_1^\sigma(\mathbf{x}), I_2^\sigma(\mathbf{x}+\mathbf{h}(x)))\nabla I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x})),$$

where the function $L^g_{\mathbf{h}}(\mathbf{i})$ *is equal to*

$$L^g_{MI,\mathbf{h}}(\mathbf{i}) = \frac{1}{|\Omega|}\left(\frac{\partial_2 P_{\mathbf{h}}(\mathbf{i})}{P_{\mathbf{h}}(\mathbf{i})} - \frac{p'_{\mathbf{h}}(i_2)}{p_{\mathbf{h}}(i_2)}\right), \tag{16}$$

in the case of the mutual information and to

$$L^g_{CR,\mathbf{h}}(\mathbf{i}) = -\frac{2}{|\Omega|\, v_2(\mathbf{h})}\left(i_2\mathbf{CR}^g[\mathbf{h}] - \mu_{2|1}(i_1,\mathbf{h}) - (\mathbf{CR}^g[\mathbf{h}]-1)\mu_2(\mathbf{h})\right), \tag{17}$$

in the case of the correlation ratio.

This defines two functions $H \to H$*:*

$$F^g_{MI}(\mathbf{h}) = (G_\beta \star L^g_{MI,\mathbf{h}})(I_1^\sigma, I_2^\sigma(\mathbf{Id}+\mathbf{h}))\nabla I_2^\sigma(\mathbf{Id}+\mathbf{h})$$

and

$$F^g_{CR}(\mathbf{h}) = (G_\beta \star L^g_{CR,\mathbf{h}})(I_1^\sigma, I_2^\sigma(\mathbf{Id}+\mathbf{h}))\nabla I_2^\sigma(\mathbf{Id}+\mathbf{h}).$$

Proof : The only point that is not contained in [8] is the fact that $F^g_{\mathrm{MI}}(\mathbf{h}) \in H$ (respectively that $F^g_{\mathrm{CR}}(\mathbf{h}) \in H$). This is a consequence of theorems 7 and 10. □

A point to be noticed is that the function $L^g_{\mathbf{h}} : \mathbb{R}^2 \to \mathbb{R}$ is convolved with the 2D gaussian G_β and the result evaluated at the intensity pair $(I_1^\sigma(\mathbf{x}), I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x})))$. The direction of descent at the point $\mathbf{x}$ is then obtained by multiplying this value by the gradient of the second image at the point $\mathbf{x}+\mathbf{h}(\mathbf{x})$.

The case of the local criteria is very similar. We have the following

Theorem 4 *In the local case, the infinitesimal gradient is given by*

$$F^l(\mathbf{h})(\mathbf{x}) = \nabla_{\mathbf{h}}\mathcal{I}[\mathbf{h}](\mathbf{x}) = (G_\gamma \star (G_\beta \star L^l_{\mathbf{h}}))(I_1^\sigma(\mathbf{x}), I_2^\sigma(\mathbf{x}+\mathbf{h}(x)), \mathbf{x})\nabla I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x})),$$

where the function $L^l_{\mathbf{h}}(i_1, i_2, \mathbf{x})$ *is equal to*

$$L^l_{MI,\mathbf{h}}(\mathbf{i},\mathbf{x}) = \frac{1}{\mathcal{G}_\gamma(\mathbf{x})}\left(\frac{\partial_2 P_{\mathbf{h}}(\mathbf{i},\mathbf{x})}{P_{\mathbf{h}}(\mathbf{i},\mathbf{x})} - \frac{p'_{\mathbf{h}}(i_2,\mathbf{x})}{p_{\mathbf{h}}(i_2,\mathbf{x})}\right),$$

in the case of the mutual information and to

$$L^l_{CR,\mathbf{h}}(\mathbf{i},\mathbf{x}) = -\frac{2}{\mathcal{G}_\gamma(\mathbf{x})\, v_2(\mathbf{x},\mathbf{h})}\left(i_2\mathbf{CR}^l[\mathbf{h}](\mathbf{x}) - \mu_{2|1}(i_1,\mathbf{x},\mathbf{h}) - (\mathbf{CR}^l[\mathbf{h}](\mathbf{x})-1)\mu_2(\mathbf{x},\mathbf{h})\right), \tag{18}$$

in the case of the correlation ratio. $\mathcal{G}_\gamma(\mathbf{x})$ *is defined by (10).*

This defines two functions $H \to H$*:*

$$F^l_{MI}(\mathbf{h}) = (G_\gamma \star G_\beta \star L^l_{MI,\mathbf{h}})(I_1^\sigma, I_2^\sigma(\mathbf{Id}+\mathbf{h}), \mathbf{Id})\nabla I_2^\sigma(\mathbf{Id}+\mathbf{h})$$

and

$$F^l_{CR}(\mathbf{h}) = (G_\gamma \star G_\beta \star L^l_{CR,\mathbf{h}})(I_1^\sigma, I_2^\sigma(\mathbf{Id}+\mathbf{h}), Id)\nabla I_2^\sigma(\mathbf{Id}+\mathbf{h}).$$

Proof : The fact that $F^l_{\mathrm{MI}}(\mathbf{h}) \in H$ (respectively that $F^l_{\mathrm{CR}}(\mathbf{h}) \in H$) is a consequence of theorems 11 and 16.

We give the remaining of the proof only in the case of the mutual information, the proof in the case of the correlation ratio can be obtained from this proof and [8]. We compute the Gâteaux derivative $\mathcal{J}_{\mathbf{MI},\mathbf{k}}[\mathbf{h}]$ of $\mathcal{J}_{\mathbf{MI}}[\mathbf{h}]$ in the direction of $\mathbf{k} \in H$:

$$\mathcal{J}_{\mathbf{MI},\mathbf{k}}[\mathbf{h}] = \left.\frac{\partial \mathcal{J}_{\mathbf{MI}}[\mathbf{h}+\epsilon\mathbf{k}]}{\partial \epsilon}\right|_{\epsilon=0}$$

We have

$$\begin{aligned}\frac{\partial \mathcal{J}_{\mathbf{MI}}[\mathbf{h}+\epsilon\mathbf{k}]}{\partial \epsilon} &= -\int_\Omega \int_{\mathbb{R}^2} \frac{\partial}{\partial \epsilon}\left[P_{\mathbf{h}+\epsilon\mathbf{k}}(\mathbf{i},\mathbf{x}_0)\log\frac{P_{\mathbf{h}+\epsilon\mathbf{k}}(\mathbf{i},\mathbf{x}_0)}{p(i_1,\mathbf{x}_0)p^g_{\mathbf{h}+\epsilon\mathbf{k}}(i_2)}\right] d\mathbf{i}d\mathbf{x}_0\\ &= -\int_\Omega \int_{\mathbb{R}^2} \underbrace{\left(1+\log\frac{P_{\mathbf{h}+\epsilon\mathbf{k}}(\mathbf{i},\mathbf{x}_0)}{p(i_1,\mathbf{x}_0)p_{\mathbf{h}+\epsilon\mathbf{k}}(i_2,\mathbf{x}_0)}\right)}_{L^{\mathbf{h}+\epsilon\mathbf{k}}(\mathbf{i},\mathbf{x}_0)} \frac{\partial P_{\mathbf{h}+\epsilon\mathbf{k}}(\mathbf{i},\mathbf{x}_0)}{\partial \epsilon} d\mathbf{i}d\mathbf{x}_0\\ &\quad \underbrace{-\int_\Omega \int_{\mathbb{R}^2} \frac{P_{\mathbf{h}+\epsilon\mathbf{k}}(\mathbf{i},\mathbf{x}_0)}{p_{\mathbf{h}+\epsilon\mathbf{k}}(i_2,\mathbf{x}_0)} \frac{\partial p_{\mathbf{h}+\epsilon\mathbf{k}}(i_2,\mathbf{x}_0)}{\partial \epsilon} d\mathbf{i}d\mathbf{x}_0}_{P}\end{aligned}$$

We first notice that

$$P = \int_{\mathbb{R}} \frac{\partial p_{\mathbf{h}+\epsilon\mathbf{k}}(i_2,\mathbf{x}_0)}{\partial \epsilon} \frac{1}{p_{\mathbf{h}+\epsilon\mathbf{k}}(i_2,\mathbf{x}_0)} \underbrace{\left(\int_{\mathbb{R}} P_{\mathbf{h}+\epsilon\mathbf{k}}(\mathbf{i},\mathbf{x}_0)di_1\right)}_{p_{\mathbf{h}+\epsilon\mathbf{k}}(i_2,\mathbf{x}_0)} di_2 = \frac{\partial}{\partial \epsilon}\underbrace{\left[\int_{\mathbb{R}} p_{\mathbf{h}+\epsilon\mathbf{k}}(i_2,\mathbf{x}_0)di_2\right]}_{1} = 0$$

The law $P_{\mathbf{h}+\epsilon\mathbf{k}}(\mathbf{i},\mathbf{x}_0)$ is given by equation (9):

$$P_{\mathbf{h}+\epsilon\mathbf{k}}(\mathbf{i},\mathbf{x}_0) = \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)}\int_\Omega G_\beta(\mathbf{I}_{\mathbf{h}+\epsilon\mathbf{k}}(\mathbf{x})-\mathbf{i})G_\gamma(\mathbf{x}-\mathbf{x}_0)\,d\mathbf{x}.$$

Therefore

$$\frac{\partial P_{\mathbf{h}+\epsilon\mathbf{k}}(\mathbf{i},\mathbf{x}_0)}{\partial\epsilon} = \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} \int_\Omega G_\gamma(\mathbf{x}-\mathbf{x}_0)\partial_2 G_\beta(I_1^\sigma(\mathbf{x})-i_1, I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x})+\epsilon\mathbf{k}(\mathbf{x}))-i_2) \nabla I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x})+\epsilon\mathbf{k}(\mathbf{x}))\cdot\mathbf{k}(\mathbf{x})\,d\mathbf{x}$$

We now let $\epsilon = 0$,

$$\mathcal{J}_{\mathbf{MI},\mathbf{k}}[\mathbf{h}] = -\int_\Omega\int_{\mathbb{R}^2} L^{\mathbf{h}}(\mathbf{i},\mathbf{x}_0) \int_\Omega \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} G_\gamma(\mathbf{x}-\mathbf{x}_0)\partial_2 G_\beta(I_1^\sigma(\mathbf{x})-i_1, I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))-i_2) \nabla I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))\cdot\mathbf{k}(\mathbf{x})\,d\mathbf{x}\,d\mathbf{i}\,d\mathbf{x}_0$$

This expression can be rewritten

$$\mathcal{J}_{\mathbf{MI},\mathbf{k}}[\mathbf{h}] = -\int_\Omega\Big(\int_{\mathbb{R}^2}\big(\int_\Omega G_\gamma(\mathbf{x}-\mathbf{x}_0)\frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)}L^{\mathbf{h}}(\mathbf{i},\mathbf{x}_0)\,d\mathbf{x}_0\big) \partial_2 G_\beta(I_1^\sigma(\mathbf{x})-i_1, I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))-i_2)\,d\mathbf{i}\Big)\nabla I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))\cdot\mathbf{k}(\mathbf{x})\,d\mathbf{x}$$

Two convolutions appear, one with respect to the space variable $\mathbf{x}$ and the other one with respect to the intensity variable $\mathbf{i}$. This last convolution commutes with the partial derivative ∂_2 with respect to the second intensity variable i_2:

$$\mathcal{J}_{\mathbf{MI},\mathbf{k}}[\mathbf{h}] = -\int_\Omega\Big(\int_{\mathbb{R}^2}\big(\int_\Omega G_\gamma(\mathbf{x}-\mathbf{x}_0)\frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)}\partial_2 L^{\mathbf{h}}(\mathbf{i},\mathbf{x}_0)\,d\mathbf{x}_0\big) G_\beta(I_1^\sigma(\mathbf{x})-i_1, I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))-i_2)\,d\mathbf{i}\Big)\nabla I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))\cdot\mathbf{k}(\mathbf{x})\,d\mathbf{x}$$

Since

$$\partial_2 L^{\mathbf{h}}(\mathbf{i},\mathbf{x}_0) = \left(\frac{\partial_2 P_{\mathbf{h}}(\mathbf{i},\mathbf{x}_0)}{P_{\mathbf{h}}(\mathbf{i},\mathbf{x}_0)} - \frac{p'_{\mathbf{h}}(i_2,\mathbf{x}_0)}{p_{\mathbf{h}}(i_2,\mathbf{x}_0)}\right),$$

we find the announced Euler-Lagrange equation. □

The point to be noticed is that the function $L^l_{\mathbf{h}} : \mathbb{R}^2\times\mathbb{R}^n \to \mathbb{R}$ is convolved with the 2D gaussian G_β for the first two variables (intensities) and the nD gaussian G_γ for the remaining n variables (spatial), and the result is evaluated at the point $((I_1^\sigma(\mathbf{x}), I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))), \mathbf{x})$ of $\mathbb{R}^2\times\mathbb{R}^n$. The direction of descent at the point $\mathbf{x}$ is then obtained by multiplying this value by the gradient of the second image at the point $\mathbf{x}+\mathbf{h}(\mathbf{x})$.

In the remaining of this paper we prove that the operators $-A_1$ and $-A_2$ defined in section 6.1 are self-adjoint, maximal monotone, that A_1 and A_2 are the infinitesimal generators of C_0 semigroups of contractions of H, are also the infinitesimal generators of analytical semigroups of contractions of H and that the four functions F^g_{MI}, F^g_{CR}, F^l_{MI}, and F^l_{CR}, defined in section 6.2 are Lipschitz continuous. We then use these results to prove the existence and uniqueness of various types of solutions for the initial value problem (3).

7 Properties of the operators A_1 and A_2

In this section we prove that the operators $-A_1$ and $-A_2$ defined by our two regularization terms are self-adjoint maximal monotone and therefore that A_1 and A_2 generate C_0 semigroups. We then show that they also generate analytical semi-groups of operators.

7.1 $-A_1$ and $-A_2$ are self-adjoint maximal monotone

We recall that a linear operator A from a linear subspace $\mathcal{D}(A)$ of the Hilbert space H is said to be maximal monotone if and only if

- It is monotone:
$$(Av, v)_H \ \geq \ 0 \qquad \forall v \in \mathcal{D}(A),$$
where $(\cdot, \cdot)_H$ indicates the Hilbert product in H.

- It is maximal:
$$\text{Ran}(\mathbf{Id} + A) = H,$$
where Ran$(\mathbf{Id} + A)$ denotes the range of the linear operator $\mathbf{Id} + A$. In other words, an operator is maximal if the equation
$$v + A\, v = f$$
has a solution $\forall f \in H$.

7.1.1 The linearized elasticity operator

We begin with the linear operator $-A_1$ and apply the standard variational approach [7, 5]:

Proposition 7 *The operator $\mathbf{Id} - A_1$ defines a bilinear form B_1 on the space $\mathbf{H}^1_0$ which is continuous and coercive (elliptic).*

Proof : The proof can be found for example in [5]. We reproduce it here schematically. Let $\mathbf{k} \in \mathbf{H}^1_0$, we consider the bilinear form C_1 defined as

$$C_1(\mathbf{h}, \mathbf{k}) = -\int_\Omega \mathbf{k}^T A_1 \mathbf{h}\, d\mathbf{x}$$

We use the definition (14) of A_1, integrate by parts using the Green formula

$$\int_\Omega \mathbf{div} S \cdot \mathbf{h}\, d\mathbf{x} = -\int_\Omega S : D\mathbf{h}\, d\mathbf{x} + \int_{\partial\Omega} S\mathbf{n} \cdot \mathbf{h}\, da,$$

true for each smooth enough field $S : \overline{\Omega} \to \mathbb{S}^n$ ($\mathbb{S}^n$ is the set of symmetric $n \times n$ matrices) and vector field $\mathbf{h} : \overline{\Omega} \to \mathbb{R}^n$. In this formula, $S : D\mathbf{h} = \text{trace}(S^T D\mathbf{h})$, $\mathbf{n}$ is the inside pointing normal vector to $\partial\Omega$ and da is the area element of $\partial\Omega$. Using the fact that $\mathbf{k} \in \mathbf{H}_0^1$, we find

$$C_1(\mathbf{h}, \mathbf{k}) = \int_\Omega \left(\lambda \,\text{trace}\,(e(\mathbf{h}))\,\text{trace}\,(e(\mathbf{k})) + 2\mu e(\mathbf{h}) : e(\mathbf{k})\right)\, d\mathbf{x},$$

where $e(\mathbf{h}) : e(\mathbf{k}) = \text{trace}\,(e(\mathbf{h})^T e(\mathbf{k}))$. Hence $B_1(\mathbf{h}, \mathbf{k}) = C_1(\mathbf{h}, \mathbf{k}) + \int_\Omega \mathbf{h}(\mathbf{x}) \cdot \mathbf{k}(\mathbf{x})\, d\mathbf{x}$.

By applying several times Cauchy-Schwarz, we find that

$$|C_1(\mathbf{h}, \mathbf{k})| \leq c_1 \|\mathbf{h}\|_{\mathbf{H}^1(\Omega)} \|\mathbf{k}\|_{\mathbf{H}^1(\Omega)}, \quad c_1 > 0,$$

and hence, using Cauchy-Schwarz again, $|B_1(\mathbf{h}, \mathbf{k})| \leq b_1 \|\mathbf{h}\|_{\mathbf{H}^1(\Omega)} \|\mathbf{k}\|_{\mathbf{H}^1(\Omega)}, \quad b_1 > 0$, which proves continuity. Next we note that

$$B_1(\mathbf{h}, \mathbf{h}) \geq 2\mu \int_\Omega e(\mathbf{h}) : e(\mathbf{h})\, d\mathbf{x},$$

It is proved in [5], theorem 6.3-4, that if $\mathbf{h} \in \mathbf{H}_0^1(\Omega)$ there exists a constant $c > 0$ such that $\int_\Omega e(\mathbf{h}) : e(\mathbf{h})\, d\mathbf{x} \geq c \|\mathbf{h}\|^2_{\mathbf{H}^1(\Omega)}$, hence

$$B_1(\mathbf{h}, \mathbf{h}) \geq 2\mu c \|\mathbf{h}\|^2_{\mathbf{H}^1(\Omega)},$$

and the coerciveness is proved. □

We therefore have the

Proposition 8 *$-A_1$ is a maximal monotone self-adjoint operator from $\mathcal{D}(A_1) = \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$ into $\mathbf{L}^2(\Omega)$.*

Proof : Monotonicity follows from the coerciveness of B_1 proved in the previous proposition. More precisely, since $(-A_1\mathbf{h}, \mathbf{h})_{\mathbf{L}^2(\Omega)} = C_1(\mathbf{h}, \mathbf{h})$, the proof shows that $(-A_1\mathbf{h}, \mathbf{h})_{\mathbf{L}^2(\Omega)} \geq 2\mu \int_\Omega e(\mathbf{h}) : e(\mathbf{h})\, d\mathbf{x} \geq 0$.

Regarding maximality, proposition 7 shows that the bilinear form B_1 associated to the operator $\mathbf{Id} - A_1$ is continuous and coercive in $\mathbf{H}^1(\Omega)$. We can therefore apply the Lax-Milgram theorem and state the existence and uniqueness of a weak solution in $\mathbf{H}_0^1(\Omega)$ to the equation $\mathbf{h} - A_1\mathbf{h} = \mathbf{f}$ for all $\mathbf{f} \in \mathbf{L}^2(\Omega)$. Since Ω is regular (in particular C^2), the solution is is in $\mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$ and is a strong solution (see [5], theorem 6.3-6).

Therefore we have $\mathcal{D}(A_1) = \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$ and $\text{Ran}(\mathbf{Id} - A_1) = H$. In order to prove that the operator is self-adjoint, it is sufficient, since it is maximal monotone, to prove that is symmetric ([3], proposition VII.6), i.e. that $(-A_1\mathbf{h}, \mathbf{k})_{\mathbf{L}^2(\Omega)} = (\mathbf{h}, -A_1\mathbf{k})_{\mathbf{L}^2(\Omega)}$ and this is obvious from the proof of proposition 7. □

7.1.2 The Nagel-Enkelmann operator

We now treat the case of $-A_2$:

$$-A_2\ \mathbf{h} = -\begin{pmatrix} \mathrm{div}(\mathbf{T}_{I_1^\sigma}\nabla h_1) \\ \vdots \\ \mathrm{div}(\mathbf{T}_{I_1^\sigma}\nabla h_n) \end{pmatrix}$$

We prove the analog of proposition 7.

Proposition 9 *The operator $\mathbf{Id} - A_2$ defines a bilinear form B_2 on the space $\mathbf{H}_0^1(\Omega)$ which is continuous and coercive (elliptic).*

Proof : Because of the form of the operator A_2, it is sufficient to work on one of the coordinates and consider the operator $a_2 : \mathcal{D}(a_2) \to L^2(\Omega)$ defined by

$$a_2\ u = \mathrm{div}(\mathbf{T}_{I_1^\sigma}\nabla u),$$

and to show that the operator $u \to u - a_2\ u$ defines a bilinear form b_2 on the space $H_0^1(\Omega) \times H_0^1(\Omega)$ which is continuous and coercive. Indeed, we define

$$b_2(u,v) = \int_\Omega \left(uv - v\mathrm{div}(\mathbf{T}_{I_1^\sigma}\nabla u)\right)\ d\mathbf{x}$$

We integrate by parts the divergence term, use the fact that $v \in H_0^1(\Omega)$, and obtain

$$b_2(u,v) = \int_\Omega \left(uv + \nabla v^T \mathbf{T}_{I_1^\sigma}\nabla u\right)\ d\mathbf{x}$$

Because the coefficients of $\mathbf{T}_{I_1^\sigma}$ are all bounded, we obtain, by applying Cauchy-Schwartz:

$$|b_2(u,v)| \le c_2\|u\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)},$$

where c_2 is a positive constant, hence continuity.

Because the eigenvalues of the symmetric matrix $\mathbf{T}_{I_1^\sigma}$ are strictly positive, we have $\mathbf{T}_{I_1^\sigma} \ge c_\mathbf{T}\mathbf{Id}$, where $c_\mathbf{T}$ is a positive constant. This implies that

$$\nabla u^T \mathbf{T}_{I_1^\sigma}\nabla u = b_2(u,u) - \|u\|^2_{L^2(\Omega)} \ge c_\mathbf{T}\|\nabla u\|^2_{L^2(\Omega)},$$

from which it follows that

$$b_2(u,u) \ge c_3\|u\|^2_{H^1(\Omega)},$$

for some positive constant $c_3 > 0$ and hence we have coerciveness. □

We can therefore apply the Lax-Milgram theorem and state the existence and uniqueness of a weak solution in $\mathbf{H}_0^1(\Omega)$ to the equation $\mathbf{h} - A_2\mathbf{h} = \mathbf{f}$ for all $\mathbf{f} \in \mathbf{L}^2(\Omega)$. Since Ω is regular (in particular C^2), the coefficients of $\mathbf{T}_{I_1^\sigma}$ in $C^1(\overline{\Omega})$, the solution is in $\mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$ and is a strong solution (see e.g. [7]).

We therefore have the

Proposition 10 *$-A_2$ is a maximal monotone self-adjoint operator from $\mathcal{D}(A_2) = \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$ into $\mathbf{L}^2(\Omega)$.*

Proof : Monotonicity follows from the coerciveness of B_2 proved in the previous proposition. Maximality also follows from the proof of proposition 9. According to the same proposition, we have $\mathcal{D}(A_2) = \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$ and $\mathrm{Ran}(\mathbf{Id} - A_2) = H$ (application of the Lax-Milgram theorem). Finally, $-A_2$ is self-adjoint for the same reasons as those indicated in the proof of proposition 8. □

7.1.3 κA_1 and κA_2 are the infinitesimal generators of C_0 semigroups

We recall the definition of a C_0 semigroup of bounded linear operators on a Banach space H. Consider a one-parameter family $S(t)$, $0 \leq t \leq +\infty$ of bounded linear operators from H to H. This family is said to be a C_0 semigroup of bounded linear operators if

Definition 3 *1. $S(0) = I$, I is the identity operator on H.*

2. $S(t_1 + t_2) = S(t_1)S(t_2)$ for every $t_1, t_2 \geq 0$.

3. $\lim_{t \to 0^+} S(t)\mathbf{h} = \mathbf{h}$ for every $\mathbf{h} \in H$.

The Hille-Yosida theorem says that there is a one-to-one correspondence between C_0 semigroup of contractions ($\|S(t)\|_{\mathcal{L}(H)} \leq 1$ for all $t \geq 0$) and maximal monotone operators in a Hilbert space. A maximal monotone operator $-A$ in a Hilbert space H is said to be the infinitesimal generator of the corresponding C_0 semigroup noted $S_A(t),\ t \geq 0$.

The relation between $-A$ and $S_A(t)$ is the following. Consider the initial value problem

$$\begin{cases} \dfrac{d\mathbf{h}}{dt} - A\mathbf{h}(t) = \mathbf{0} \\ \mathbf{h}(0)(\cdot) = \mathbf{h}_0(\cdot), \end{cases} \tag{19}$$

Because $-A$ is maximal monotone selfadjoint for all $\mathbf{h}_0 \in H$ there exists a unique function $\mathbf{h} \in C([0, +\infty[, H) \cap C^1(]0, +\infty[, H) \cap C(]0, +\infty[, \mathcal{D}(A))$ such that $\|\mathbf{h}(t)\|_H \leq \|\mathbf{h}_0\|_H$. The linear application $S_A(t)$, $\mathcal{D}(A) \to \mathcal{D}(A)$ is defined by $S_A(t)\mathbf{h}_0 = \mathbf{h}(t)$, where $\mathbf{h}(t)$ is the solution of (19) at time t. Since $\|S_A(t)\mathbf{h}_0\|_H \leq \|\mathbf{h}_0\|_H$ it is possible to extend $S_A(t)$ by continuity and density to a linear continuous operator $H \to H$. This family of operators, also noted $S_A(t)$, is the C_0 semigroup of contractions corresponding to $-A$.

Thus we have the

Theorem 5 *The operators κA_1 and κA_2 are the infinitesimal generators of two C_0 semigroups of bounded linear operators on H for all $\kappa > 0$.*

Proof : This is a consequence of the Hille-Yosida theorem and of propositions 8 and 10. □

A property of the C_0 semigroups of bounded operators that we will need later is given in the following

Proposition 11 *For all* $\mathbf{h} \in \mathcal{D}(A)$, $S_A(t)\mathbf{h} \in \mathcal{D}(A)$ *and*

$$\frac{d}{dt} S_A(t)\mathbf{h} = A S_A(t)\mathbf{h} = S_A(t) A\mathbf{h}$$

Proof : The proof can be found for example in theorem 2.4 of chapter 1 of [10]. □

We will also need the following two lemmas.

Lemma 2 *The linear operator* $S_A(t)$ *is bounded by 1, for* $A = \kappa A_1$ *or* $A = \kappa A_2$ *for all* $\kappa > 0$.

Proof : It is a consequence of the Hille-Yosida theorem: because $-A$ is a maximal-monotone operator (propositions 8 and 10) we have $\|S_A(t)\mathbf{h}_0\|_H \leq \|\mathbf{h}_0\|$ for all $\mathbf{h}_0 \in H$ ([3], theorem VII.7). Hence $\|S_A(t)\|_{\mathcal{L}(H)} = 1$. □

We will also need the following

Lemma 3 *The linear operator* $-A$ *is invertible, for* $A = \kappa A_1$ *or* $A = \kappa A_2$ *for all* $\kappa > 0$.

Proof : It is sufficient to show that the equation $-A\mathbf{h} = \mathbf{f}$ has a unique solution for all $\mathbf{f} \in \mathbf{L}^2(\Omega)$. The proofs of propositions 7 and 9 show that the bilinear forms associated to the operators $-\kappa A_1$ and $-\kappa A_2$ are continuous and coercive in $\mathbf{H}^1(\Omega)$, hence the Lax-Milgram theorem tells us that the equation $-A\mathbf{h} = \mathbf{f}$ has a unique weak solution in $\mathbf{H}_0^1(\Omega)$ for $A = \kappa A_1$ and $A = \kappa A_2$ for all $\mathbf{f} \in \mathbf{L}^2(\Omega)$. Since Ω is regular the weak solution is in $\mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega)$ and is a strong solution. □

7.2 κA_1 and κA_2 are the infinitesimal generators of analytical semigroups

It turns out that, because of the special form of the operators A_1 and A_2, the corresponding C_0 semigroups can be extended to analytical semigroups. This is necessary in order to obtain the existence, uniqueness and regularity results for the solution of (31).

We recall the definition of an analytical semigroup of bounded linear operators. For more details, the interested reader is referred to [10].

Definition 4 *Let* $\Delta = \{z \in \mathbb{C} : \varphi_1 < arg\, z < \varphi_2,\, \varphi_1 < 0 < \varphi_2\}$ *and for* $z \in \Delta$, *let* $S(z)$ *be a bounded linear operator. The family* $S(z),\ z \in \Delta$ *is an* analytic *semigroup in* Δ *if*

1. $z \rightarrow S(z)$ *is analytic in* Δ.
2. $S(0) = I$ *and* $\lim_{z \rightarrow 0,\, z \in \Delta} S(z)\mathbf{h} = \mathbf{h}$ *for every* $\mathbf{h} \in H$.
3. $S(z_1 + z_2) = S(z_1)S(z_2)$ *for* $z_1, z_2 \in \Delta$.

A semigroup $S(t)$ will be called *analytic* if it is analytic in some sector Δ containing the nonnegative real axis. Clearly, the restriction of an analytic semigroup to the real axis is a C_0 semigroup. We are interested in the possibility of extending a given C_0 semigroup (i.e. $S_{\kappa A_1}(t)$ and $S_{\kappa A_2}(t)$) to an analytic semigroup in some sector around the nonnegative real axis. The reason for this interest is that if A is the infinitesimal generator of an analytic semigroup, one can define fractional powers $(-A)^\alpha$ of $-A$ which are useful in the study of our semilinear initial value problem (31).

Theorem 6 *The operators κA_1 and κA_2 are the infinitesimal generators of two analytical semi-groups of operators on H.*

Proof : To carry out the proof which is simply an adaptation of the proof of theorem 7.2.7 in [10], we extend the Hilbert product in $\mathbf{L}^2(\Omega)$ to complex valued functions $\mathbf{h}$ as shown in equation (1). Let $\mathbf{h} = \mathbf{h}_1 + i\mathbf{h}_2$, $\mathbf{h}_1 = \text{Re}(\mathbf{h})$ and $\mathbf{h}_2 = \text{Im}(\mathbf{h})$. A simple computation shows that

$$(-A\mathbf{h}, \mathbf{h})_{\mathbf{L}^2_{\mathbb{C}}(\Omega)} = (-A\mathbf{h}_1, \mathbf{h}_1)_{\mathbf{L}^2(\Omega)} + (-A\mathbf{h}_2, \mathbf{h}_2)_{\mathbf{L}^2(\Omega)},$$

since the operators $-\kappa A_1$ and $-\kappa A_2$ are symmetric (propositions 8 and 10). From the proofs of propositions 7 and 9 we deduce

$$(-A\mathbf{h}_1, \mathbf{h}_1)_{\mathbf{L}^2(\Omega)} + (-A\mathbf{h}_2, \mathbf{h}_2)_{\mathbf{L}^2(\Omega)} \geq c(\|\mathbf{h}_1\|^2_{\mathbf{H}^1(\Omega)} + \|\mathbf{h}_2\|^2_{\mathbf{H}^1(\Omega)}) = c\|\mathbf{h}\|^2_{\mathbf{H}^1_{\mathbb{C}}(\Omega)}, \quad c > 0$$

Since $\text{Im}(-A\mathbf{h}, \mathbf{h})_H = 0$, we certainly have also

$$|\text{Im}(-A\mathbf{h}, \mathbf{h})_H| \leq b\|\mathbf{h}\|^2_{\mathbf{H}^1_{\mathbb{C}}(\Omega)} \quad b > 0$$

It follows that the numerical range $S(-A) = \{(-A\mathbf{h}, \mathbf{h})_{\mathbf{L}^2_{\mathbb{C}}(\Omega)},\ \|\mathbf{h}\|_{\mathbf{H}^1_{\mathbb{C}}(\Omega)} = 1\}$ is included in the set $S_{\theta_1} = \{\lambda \in \mathbb{C}, |arg\,\lambda| < \theta_1$, where $\theta_1 = \arctan(b/c) < \pi/2$. Choosing θ such that $\theta_1 < \theta < \pi/2$ and denoting $\Sigma_\theta = \{\lambda \in \mathbb{C}, |\text{arg}\lambda| > \theta\}$ there exists a constant C_θ such that

$$d(\lambda : S(-A)) \geq C_\theta |\lambda| \quad \text{for all} \quad \lambda \in \Sigma_\theta, \tag{20}$$

where $d(\lambda : S)$ denotes the distance between λ and the set $S \in \mathbb{C}$.

Next we note that all reals $\mu, \mu < 0$ are in the resolvent set $\rho(-A)$ of $-A$. This is because the resolvent of $-A$ is the set of μs such that $\mu I + A$ is invertible, i.e. the set of μs such that $-A - \mu I$ is invertible. But since the complex equation $-A\mathbf{h} - \mu\mathbf{h} = \mathbf{f}$ ($\mathbf{h} = \mathbf{h}_1 + i\mathbf{h}_2$ and $\mathbf{f} = \mathbf{f}_1 + i\mathbf{f}_2$) is equivalent to the two real equations $-A\mathbf{h}_1 - \mu\mathbf{h}_1 = \mathbf{f}_1$ and $-A\mathbf{h}_2 - \mu\mathbf{h}_2 = \mathbf{f}_2$, we can apply the results in the proofs of propositions 7 and 9 and the Lax-Milgram theorem to show that each equation has a unique strong solution in $\mathbf{H}^1_0(\Omega) \cap \mathbf{H}^2(\Omega)$ for each $\mathbf{f}_1,\ \mathbf{f}_2 \in \mathbf{L}^2(\Omega)$ and therefore that the equation $-A\mathbf{h} - \mu\mathbf{h} = \mathbf{f}$ has a unique strong solution in $\mathbf{H}^1_{0\mathbb{C}}(\Omega) \cap \mathbf{H}^2_{\mathbb{C}}(\Omega)$ for each $\mathbf{f} \in \mathbf{L}^2_{\mathbb{C}}(\Omega)$.

Since $\mu, \mu < 0$ are in the resolvent set of $-A$, this shows that Σ_θ, which contains the negative real axis, is contained in a component of the complement of the closure $\overline{S(-A)}$ of $S(-A)$ that has a nonempty intersection with $\rho(-A)$. This implies (theorem 1.3.9 of [10]) that $\rho(-A) \supset \Sigma_\theta$ and that for every λ in Σ_θ

$$\|R(\lambda : -A)\| \leq d(\lambda : \overline{S(-A)})^{-1} \leq \frac{1}{C_\theta |\lambda|},$$

where $R(\lambda : -A)$ is the operator $(\lambda I + A)^{-1}$.

Using the fact that $\rho(-A) = -\rho(A)$ and defining $\Sigma_{-\theta} = -\Sigma_\theta = \{\lambda \in \mathbb{C}, |\text{arg}\lambda| < \pi - \theta\}$, we infer that $\rho(A) \supset \Sigma_{-\theta}$. Moreover, for all $\lambda \in \Sigma_{-\theta}$ we have $\|R(\lambda : A)\| = \|R(-\lambda : -A)\|$ and, according to (20),

$$\|R(-\lambda : -A)\| \leq d(-\lambda : \overline{S(-A)})^{-1} \leq \frac{1}{C_\theta |\lambda|}.$$

We can therefore apply theorem 2.5.2 in [10] which allows us to conclude that the C_0 semigroup generated by A can be extended to an analytical semigroup in a sector $\Delta_\delta = \{z \in \mathbb{C},\ |\text{arg}z| < \delta\}$ where $0 < \delta < \pi/2$ is defined by $\pi - \theta = \pi/2 + \delta$. $\square$

8 Properties of the functions F^g and F^l

In this section we prove that the functions F defined by our various criteria (theorems 3 and 4) satisfy some Lipschitz continuity conditions.

8.1 Preliminary results

We remind the reader of the following results on Lipschitz continuous functions. These results will be used several times in the following.

Proposition 12 *Let $\mathcal{H}$ be a Banach space and let us denote its norm by $\|.\|_{\mathcal{H}}$. Let $f_i,\ i = 1, 2 : \mathcal{H} \longrightarrow \mathbb{R}$ be two Lipschitz continuous functions. We have the following:*

1. *$f_1 + f_2$ is Lipschitz continuous.*
2. *If f_1 and f_2 are bounded then the product $f_1 f_2$ is Lipschitz continuous.*
3. *If $f_2 > 0$ and if f_1 and f_2 are bounded, then the ratio $\frac{f_1}{f_2}$ is Lipschitz continuous.*

Proof : We prove only 2 and 3. Let $\mathbf{h}$ and $\mathbf{h}'$ be two vectors of $\mathcal{H}$:

$$
\begin{aligned}
&|f_1(\mathbf{h})f_2(\mathbf{h}) - f_1(\mathbf{h}')f_2(\mathbf{h}')| = \\
&\qquad |(f_1(\mathbf{h}) - f_1(\mathbf{h}'))f_2(\mathbf{h}) + f_1(\mathbf{h}')(f_2(\mathbf{h}) - f_2(\mathbf{h}')| \leq \\
&\qquad\qquad |f_2(\mathbf{h})|\,|f_1(\mathbf{h}) - f_1(\mathbf{h}')| + |f_1(\mathbf{h}')|\,|f_2(\mathbf{h}) - f_2(\mathbf{h}')|,
\end{aligned}
$$

from which point 2 above follows. Similarly

$$
\begin{aligned}
&\left|\frac{f_1(\mathbf{h})}{f_2(\mathbf{h})} - \frac{f_1(\mathbf{h}')}{f_2(\mathbf{h}')}\right| = \frac{|f_1(\mathbf{h})f_2(\mathbf{h}') - f_2(\mathbf{h})f_1(\mathbf{h}')|}{f_2(\mathbf{h})f_2(\mathbf{h}')} \leq \\
&\qquad \frac{|f_1(\mathbf{h}) - f_1(\mathbf{h}')|f_2(\mathbf{h}') + |f_1(\mathbf{h}')|\,|f_2(\mathbf{h}) - f_2(\mathbf{h}')|}{f_2(\mathbf{h})f_2(\mathbf{h}')}
\end{aligned}
$$

If $f_2 > 0$, there exists $a > 0$ such that $f_2 > a$. Hence

$$
\left|\frac{f_1(\mathbf{h})}{f_2(\mathbf{h})} - \frac{f_1(\mathbf{h}')}{f_2(\mathbf{h}')}\right| \leq \frac{1}{a^2}|f_1(\mathbf{h}) - f_1(\mathbf{h}')|f_2(\mathbf{h}') + |f_1(\mathbf{h}')|\,|f_2(\mathbf{h}) - f_2(\mathbf{h}')|
$$

from which point 3 above follows. $\square$

In the following, we will need the following definitions and notations

Definition 5 *We note $\mathcal{H}_1 = [0, \mathcal{A}] \times H$ and $\mathcal{H}_2 = [0, \mathcal{A}]^2 \times H$ the Banach spaces equipped with the norms $\|(z, \mathbf{h})\|_{\mathcal{H}_1} = |z| + \|\mathbf{h}\|_H$ and $\|(z_1, z_2, \mathbf{h})\|_{\mathcal{H}_2} = |z_1| + |z_2| + \|\mathbf{h}\|_H$, respectively.*

We will use several times the following result

Lemma 4 *Let $f : \mathcal{H}_2 \longrightarrow \mathbb{R}$ be such that $(z_1, z_2) \longrightarrow f(z_1, z_2, \mathbf{h})$ is Lipschitz continuous with a Lipschitz constant l_f independent of $\mathbf{h}$ and such that $\mathbf{h} \longrightarrow f(z_1, z_2, \mathbf{h})$ is Lipschitz continuous with a Lipschitz constant L_f independent of (z_1, z_2), then f is Lipschitz continuous.*

Proof : We have

$$
\begin{aligned}
&|f(z_1, z_2, \mathbf{h}) - f(z_1', z_2', \mathbf{h}')| \leq \\
&\quad |f(z_1, z_2, \mathbf{h}) - f(z_1', z_2', \mathbf{h})| + |f(z_1', z_2', \mathbf{h}) - f(z_1', z_2', \mathbf{h}')| \leq \\
&\quad l_f(|z_1 - z_1'| + |z_2 - z_2'|) + L_f\|\mathbf{h} - \mathbf{h}'\|_H \leq \\
&\quad \max(l_f, L_f)(|z_1 - z_1'| + |z_2 - z_2'| + \|\mathbf{h} - \mathbf{h}'\|_H)
\end{aligned}
$$

□

In section 8.3, we will need a slightly more general version of this lemma.

Lemma 5 *Let $f : [0, \mathcal{A}]^2 \times \Omega \times H \longrightarrow \mathbb{R}$ be such that $(z_1, z_2) \longrightarrow f(z_1, z_2, \mathbf{x}, \mathbf{h})$ is Lipschitz continuous with a Lipschitz constant l_f independent of $\mathbf{x}$ and $\mathbf{h}$ and such that $\mathbf{h} \longrightarrow f(z_1, z_2, \mathbf{x}, \mathbf{h})$ is Lipschitz continuous with a Lipschitz constant L_f independent of $(z_1, z_2, \mathbf{x})$, then f is Lipschitz continuous on $[0, \mathcal{A}]^2 \times H$ uniformly on Ω.*

Proof : Indeed,

$$
\begin{aligned}
&|f(z_1, z_2, \mathbf{x}, \mathbf{h}) - f(z_1', z_2', \mathbf{x}, \mathbf{h}')| \leq \\
&\quad |f(z_1, z_2, \mathbf{x}, \mathbf{h}) - f(z_1', z_2', \mathbf{x}, \mathbf{h})| + |f(z_1', z_2', \mathbf{x}, \mathbf{h}) - f(z_1', z_2', \mathbf{h}')| \leq \\
&\quad l_f(|z_1 - z_1'| + |z_2 - z_2'|) + L_f\|\mathbf{h} - \mathbf{h}'\|_H \leq \\
&\quad \max(l_f, L_f)(|z_1 - z_1'| + |z_2 - z_2'| + \|\mathbf{h} - \mathbf{h}'\|_H) \forall \mathbf{x} \in \Omega,
\end{aligned}
$$

and the Lipschitz constant $\max(l_f, L_f)$ is independent of $\mathbf{x} \in \Omega$. □

8.2 Global criteria

We discuss the case of the two global criteria.

8.2.1 Mutual Information

We first prove that in the mutual information case, there is a neat separation in the definition of the function F between its local and global dependency in the field $\mathbf{h}$. More precisely we have the following

Proposition 13 *The function* $q_{\mathbf{h}} : [0, \mathcal{A}] \to \mathbb{R}$ *defined by*

$$q_{\mathbf{h}}(i_2) = \frac{p'_{\mathbf{h}}(i_2)}{p_{\mathbf{h}}(i_2)}$$

satisfies the following equation:

$$q_{\mathbf{h}}(i_2) = a(i_2, \mathbf{h}) - \frac{i_2}{\beta},$$

where the function $0 \le a(i_2, \mathbf{h}) \le \frac{\mathcal{A}}{\beta}$.

Proof : $p_{\mathbf{h}}$ is defined by

$$p_{\mathbf{h}}(i_2) = \int_\Omega g_\beta(I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) - i_2)\, d\mathbf{x},$$

hence

$$p'_{\mathbf{h}}(i_2) = \frac{1}{\beta\,\Omega} \int_\Omega (I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) - i_2) g_\beta(I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) - i_2)\, d\mathbf{x}.$$

The function $a(i_2, \mathbf{h})$ is equal to

$$a(i_2, \mathbf{h}) = \frac{1}{\beta} \frac{\displaystyle\int_\Omega I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) g_\beta(I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) - i_2)\, d\mathbf{x}}{\displaystyle\int_\Omega g_\beta(I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) - i_2)\, d\mathbf{x}}, \tag{21}$$

and the result follows from the fact that $I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) \in [0, \mathcal{A}]$. □

A simple variation of the previous proof shows the truth of the following

Proposition 14 *The function* $Q_{\mathbf{h}} : [0, \mathcal{A}]^2 \to \mathbb{R}$ *defined by*

$$Q_{\mathbf{h}}(\mathbf{i}) = \frac{\partial_2 P_{\mathbf{h}}(\mathbf{i})}{P_{\mathbf{h}}(\mathbf{i})}$$

satisfies the following equation:

$$Q_{\mathbf{h}}(\mathbf{i}) = A(\mathbf{i}, \mathbf{h}) - \frac{i_2}{\beta},$$

where the function $0 \le A(\mathbf{i}, \mathbf{h}) \le \frac{\mathcal{A}}{\beta}$.

In the following, we will use the function $L^g_{\mathrm{MI},\,\mathbf{h}}(\mathbf{i}, \mathbf{h}) : [0, \mathcal{A}]^2 \to \mathbb{R}$ defined as (see theorem 3)

$$L^g_{\mathrm{MI},\,\mathbf{h}}(\mathbf{i}) = \frac{1}{|\Omega|}(Q_{\mathbf{h}}(\mathbf{i}) - q_{\mathbf{h}}(i_2)) = \frac{1}{|\Omega|}(A(\mathbf{i}, \mathbf{h}) - a(i_2, \mathbf{h})) \tag{22}$$

We then consider the result f^g_{MI} of convolving $L^g_{\mathrm{MI},\,\mathbf{h}}$ with G_β, i.e. the two functions $b : [0, \mathcal{A}] \times H \longrightarrow \mathbb{R}$ defined as

$$b(z_2, \mathbf{h}) = (g_\beta \star a)(z_2, \mathbf{h}) = \int_{\mathbb{R}} g_\beta(z_2 - i_2) a(i_2, \mathbf{h})\, di_2, \tag{23}$$

and $B : [0, \mathcal{A}]^2 \times H \longrightarrow \mathbb{R}$ defined as

$$B(z_1, z_2, \mathbf{h}) = (G_\beta \star A)(\mathbf{z}, \mathbf{h}) = \int_{\mathbb{R}^2} G_\beta(\mathbf{z} - \mathbf{i}) A(\mathbf{i}, \mathbf{h}) \, d\mathbf{i} \tag{24}$$

We prove a series of propositions.

Proposition 15 *The function* $[0, \mathcal{A}] \longrightarrow \mathbb{R}^+$ *defined by* $z_2 \longrightarrow b(z_2, \mathbf{h})$ *is Lipschitz continuous with a Lipschitz constant* l_b^g *which is independent of* $\mathbf{h}$. *Moreover, it is bounded by* $\frac{\mathcal{A}}{\beta}$.

Proof : The second part of the proposition follows from the fact that $0 \leq a(i_2, \mathbf{h}) \leq \frac{\mathcal{A}}{\beta} \, \forall i_2 \in \mathbb{R}$ and $\forall \mathbf{h} \in H$ (proposition 13).

In order to prove the first part, we prove that the magnitude of the derivative of the function is bounded independently of $\mathbf{h}$. Indeed

$$|b'(z_2, \mathbf{h})| = \frac{1}{\beta} \left| \int_{-\infty}^{+\infty} (z_2 - i_2) g_\beta(z_2 - i_2) a(i_2, \mathbf{h}) di_2 \right| \leq \frac{\mathcal{A}}{\beta} \int_{-\infty}^{+\infty} |z_2 - i_2| g_\beta(z_2 - i_2) di_2$$

The function on the righthand side of the inequality is independent of $\mathbf{h}$ and continuous on $[0, \mathcal{A}]$, therefore upperbounded. $\Box$

Proposition 16 *The function* $\mathbf{h} \longrightarrow b(z_2, \mathbf{h}) : \mathbf{L}^2(\Omega) \longrightarrow \mathbb{R}$ *is Lipschitz continuous on* $\mathbf{L}^2(\Omega)$ *with Lipschitz constant* L_b^g *which is independent of* $z_2 \in [0, \mathcal{A}]$.

Proof : We consider

$$b(z_2, \mathbf{h}_1) - b(z_2, \mathbf{h}_2) = \int_{\mathbb{R}} g_\beta(z_2 - i_2) \left(a(i_2, \mathbf{h}_1) - a(i_2, \mathbf{h}_2) \right) \, di_2 \tag{25}$$

According to equation (21), $a(i_2, \mathbf{h})$ is the ratio $N(i_2, \mathbf{h})/D(i_2, \mathbf{h})$ of the two functions

$$N(i_2, \mathbf{h}) = \int_\Omega I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) g_\beta(I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) - i_2) \, d\mathbf{x},$$

and

$$D(i_2, \mathbf{h}) = \int_\Omega g_\beta(I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) - i_2) \, d\mathbf{x}.$$

We ignore the factor $1/\beta$ which is irrelevant in the proof. We write

$$|b(z_2, \mathbf{h}_1) - b(z_2, \mathbf{h}_2)| \leq \int_{\mathbb{R}} g_\beta(z_2 - i_2) \frac{|N(i_2, \mathbf{h}_2)| \, |D(i_2, \mathbf{h}_2) - D(i_2, \mathbf{h}_1)|}{D(i_2, \mathbf{h}_1) D(i_2, \mathbf{h}_2)} \, di_2 + \int_{\mathbb{R}} g_\beta(z_2 - i_2) \frac{D(i_2, \mathbf{h}_2) \, |N(i_2, \mathbf{h}_2) - N(i_2, \mathbf{h}_1)|}{D(i_2, \mathbf{h}_1) D(i_2, \mathbf{h}_2)} \, di_2, \tag{26}$$

and consider the first term of the righthand side.

$$D(i_2, \mathbf{h}_2) - D(i_2, \mathbf{h}_1) = \int_\Omega \left(g_\beta(i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}_2(\mathbf{x}))\,) - g_\beta(i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}_1(\mathbf{x}))\,)\right) d\mathbf{x}$$

We use the first order Taylor expansion with integral remainder of the C^1 function g_β. This says that

$$g_\beta(i + t) = g_\beta(i) + t \int_0^1 g'_\beta(i + t\alpha)\, d\alpha,$$

as the reader will easily verify. We can therefore write

$$g_\beta(i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}_2(\mathbf{x}))\,) - g_\beta(i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}_1(\mathbf{x}))\,) = \left(I_2^\sigma(\mathbf{x} + \mathbf{h}_1(\mathbf{x})) - I_2^\sigma(\mathbf{x} + \mathbf{h}_2(\mathbf{x}))\,\right) \int_0^1 g'_\beta\Big(i_2 - \big(\alpha\, I_2^\sigma(\mathbf{x} + \mathbf{h}_2(\mathbf{x})) + (1-\alpha)\, I_2^\sigma(\mathbf{x} + \mathbf{h}_1(\mathbf{x}))\big)\Big)\, d\alpha$$

We use the fact that I_2^σ is Lipschitz continuous and write

$$|D(i_2, \mathbf{h}_2) - D(i_2, \mathbf{h}_1)| \leq Lip(I_2^\sigma) \int_\Omega \Big(|\mathbf{h}_1(\mathbf{x}) - \mathbf{h}_2(\mathbf{x})| \left| \int_0^1 g'_\beta(i_2 - (\alpha\, I_2^\sigma(\mathbf{x} + \mathbf{h}_2(\mathbf{x})) + (1-\alpha)\, I_2^\sigma(\mathbf{x} + \mathbf{h}_1(\mathbf{x}))\,))\, d\alpha \right| \Big)\, d\mathbf{x}$$

Schwarz inequality implies

$$|D(i_2, \mathbf{h}_2) - D(i_2, \mathbf{h}_1)| \leq Lip(I_2^\sigma) \|\mathbf{h}_1 - \mathbf{h}_2\|_H \left(\int_\Omega \left(\int_0^1 g'_\beta(i_2 - (\alpha\, I_2^\sigma(\mathbf{x} + \mathbf{h}_2(\mathbf{x})) + (1-\alpha)\, I_2^\sigma(\mathbf{x} + \mathbf{h}_1(\mathbf{x}))\,))\, d\alpha \right)^2 d\mathbf{x} \right)^{\frac{1}{2}}$$

We introduce the function

$$r(i_2, \mathbf{h}_1, \mathbf{h}_2) = \left(\int_\Omega \left(\int_0^1 \left| i_2 - \Big(\alpha\, I_2^\sigma(\mathbf{x} + \mathbf{h}_2(\mathbf{x})) + (1-\alpha)\, I_2^\sigma(\mathbf{x} + \mathbf{h}_1(\mathbf{x}))\Big) \right| g_\beta\Big(i_2 - \big(\alpha\, I_2^\sigma(\mathbf{x} + \mathbf{h}_2(\mathbf{x})) + (1-\alpha)\, I_2^\sigma(\mathbf{x} + \mathbf{h}_1(\mathbf{x}))\big)\Big)\, d\alpha \right)^2 d\mathbf{x} \right)^{\frac{1}{2}}$$

We notice that

$$\left(\int_\Omega \left(\int_0^1 g'_\beta(i_2 - (\alpha\, I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x}))\; + (1-\alpha)\, I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x})))\,)\, d\alpha\right)^2 d\mathbf{x}\right)^{\frac{1}{2}} \\ \leq \frac{1}{\beta} r(i_2, \mathbf{h}_1, \mathbf{h}_2).$$

So far we have

$$\int_{\mathbb{R}} g_\beta(z_2 - i_2) \frac{|N(i_2,\mathbf{h}_2)|\,|D(i_2,\mathbf{h}_2) - D(i_2,\mathbf{h}_1)|}{D(i_2,\mathbf{h}_1)D(i_2,\mathbf{h}_2)}\, di_2 \leq \\ \frac{Lip(I_2^\sigma)}{\beta}\|\mathbf{h}_1 - \mathbf{h}_2\|_H \left(\int_{\mathbb{R}} g_\beta(z_2 - i_2)\frac{|N(i_2,\mathbf{h}_2)|\; r_H(i_2,\mathbf{h}_1,\mathbf{h}_2)}{D(i_2,\mathbf{h}_1)D(i_2,\mathbf{h}_2)}\, di_2\right) \quad (27)$$

We study the function of z_2 that is on the righthand side of this inequality. First we note that the function is well defined since no problems occur when i_2 goes to infinity because ”there are three gaussians in the numerator and two in the denominator”. We then show that this function is bounded independently of $\mathbf{h}_1$ and $\mathbf{h}_2$ for all $z_2 \in [0, \mathcal{A}]$. We consider three cases:

$i_2 \leq 0$

This is the case where

$$\begin{array}{lll} 0 \leq |i_2| \leq & |i_2 - I_2^\sigma(\mathbf{x}+\mathbf{h}_j(\mathbf{x}))| & \leq |i_2 - \mathcal{A}| \quad j = 1,2 \\ 0 \leq |i_2| \leq & |i_2 - (\alpha I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x})) + & \\ & (1-\alpha) I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x})))| & \leq |i_2 - \mathcal{A}| \quad 0 \leq \alpha \leq 1 \end{array}$$

Hence

$$\begin{array}{ll} g_\beta(i_2 - \mathcal{A}) \leq g_\beta(i_2 - I_2^\sigma(\mathbf{x}+\mathbf{h}_j(\mathbf{x}))) & \leq g_\beta(i_2) \quad j = 1,2 \\ g_\beta(i_2 - \mathcal{A}) \leq g_\beta(i_2 - (\alpha I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x})) + & \\ \qquad (1-\alpha) I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x})))) & \leq g_\beta(i_2) \quad 0 \leq \alpha \leq 1 \end{array}$$

This yields

$$\int_{-\infty}^0 g_\beta(z_2 - i_2)\frac{|N(i_2,\mathbf{h}_2)|\; r(i_2,\mathbf{h}_1,\mathbf{h}_2)}{D(i_2,\mathbf{h}_1)D(i_2,\mathbf{h}_2)}\, di_2 \leq \\ |\Omega|^{1/2}\mathcal{A}\int_{-\infty}^0 g_\beta(z_2 - i_2)|i_2 - \mathcal{A}|\left(\frac{g_\beta(i_2)}{g_\beta(i_2 - \mathcal{A})}\right)^2 di_2,$$

The integral on the righthand side is well-defined and defines a continuous function of z_2.

$0 \le i_2 \le \mathcal{A}$

$$\begin{aligned} 0 \le & \quad |i_2 - I_2^\sigma(\mathbf{x}+\mathbf{h}_j(\mathbf{x}))| & \le \mathcal{A} \quad j=1,2 \\ 0 \le & |i_2 - (\alpha I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x})) + (1-\alpha) I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x})))| & \le \mathcal{A} \quad 0 \le \alpha \le 1 \end{aligned}$$

Hence

$$\begin{aligned} g_\beta(\mathcal{A}) \le & \quad g_\beta(i_2 - I_2^\sigma(\mathbf{x}+\mathbf{h}_j(\mathbf{x}))) & \le g_\beta(0) \quad j=1,2 \\ g_\beta(\mathcal{A}) \le & \quad g_\beta(i_2 - (\alpha I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x})) + & \\ & \quad (1-\alpha) I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x})))) & \le g_\beta(0) \quad 0 \le \alpha \le 1 \end{aligned}$$

This yields

$$\int_0^{\mathcal{A}} g_\beta(z_2 - i_2) \frac{|N(i_2,\mathbf{h}_2)|\, r(i_2,\mathbf{h}_1,\mathbf{h}_2)}{D(i_2,\mathbf{h}_1) D(i_2,\mathbf{h}_2)}\, di_2 \le \\ |\Omega|^{1/2} \left(\frac{\mathcal{A} g_\beta(0)}{g_\beta(\mathcal{A})} \right)^2 \int_0^{\mathcal{A}} g_\beta(z_2 - i_2)\, di_2,$$

The integral on the righthand side is convergent and defines a continuous function of z_2.

$i_2 \ge \mathcal{A}$
This is the case where

$$\begin{aligned} 0 \le & i_2 - \mathcal{A} \le \quad i_2 - I_2^\sigma(\mathbf{x}+\mathbf{h}_j(\mathbf{x})) & \le i_2 \quad j=1,2 \\ 0 \le & i_2 - \mathcal{A} \le \quad i_2 - (\alpha I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x})) + & \\ & \quad (1-\alpha) I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x}))) & \le i_2 \quad 0 \le \alpha \le 1 \end{aligned}$$

Hence

$$\begin{aligned} g_\beta(i_2) \le & \quad g_\beta(i_2 - I_2^\sigma(\mathbf{x}+\mathbf{h}_j(\mathbf{x}))) & \le g_\beta(i_2 - \mathcal{A}) \quad j=1,2 \\ g_\beta(i_2) \le & \quad g_\beta(i_2 - (\alpha I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x})) + & \\ & \quad (1-\alpha) I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x})))) & \le g_\beta(i_2 - \mathcal{A}) \quad 0 \le \alpha \le 1 \end{aligned}$$

This yields

$$\int_{\mathcal{A}}^{+\infty} g_\beta(z_2 - i_2) \frac{|N(i_2,\mathbf{h}_2)|\, r(i_2,\mathbf{h}_1,\mathbf{h}_2)}{D(i_2,\mathbf{h}_1) D(i_2,\mathbf{h}_2)}\, di_2 \le \\ |\Omega|^{1/2} \mathcal{A} \int_{\mathcal{A}}^{+\infty} g_\beta(z_2 - i_2) i_2 \left(\frac{g_\beta(i_2 - \mathcal{A})}{g_\beta(i_2)} \right)^2 di_2,$$

The integral on the righthand side is convergent and defines a continuous function of z_2.

In all three cases, the functions of z_2 appearing on the righthand side are continuous, independent of $\mathbf{h}_1$ and $\mathbf{h}_2$, therefore upperbounded on $[0, \mathcal{A}]$ by a constant independent of $\mathbf{h}_1$ and $\mathbf{h}_2$. Returning to inequality (27), we have proved that there existed a positive constant C independent of z_2 such that

$$\int_{\mathbb{R}} g_\beta(z_2 - i_2) \frac{|N(i_2, \mathbf{h}_2)|\,|D(i_2, \mathbf{h}_2) - D(i_2, \mathbf{h}_1)|}{D(i_2, \mathbf{h}_1) D(i_2, \mathbf{h}_2)}\, di_2 \leq \\ C \|\mathbf{h}_1 - \mathbf{h}_2\|_H \quad \forall z_2 \in [0, \mathcal{A}] \quad \forall \mathbf{h}_1,\, \mathbf{h}_2 \in H$$

A similar proof can be developed for the second term in the righthand side of the inequality (26). In conclusion we have proved that there existed a constant L_b^g, independent of z_2 such that

$$|b(z_2, \mathbf{h}_1) - b(z_2, \mathbf{h}_2)| \leq L_b^g \|\mathbf{h}_1 - \mathbf{h}_2\|_H \quad \forall z_2 \in [0, \mathcal{A}] \quad \forall \mathbf{h}_1,\, \mathbf{h}_2 \in H$$

□

We can state the

Proposition 17 *The function $b : \mathcal{H}_1 \to \mathbb{R}$ is Lipschitz continuous.*

Proof : The proof follows from propositions 15, 16 and lemma 4. □

We now proceed with showing the same kind of properties for the function B. We start with the

Proposition 18 *The function $[0, \mathcal{A}]^2 \longrightarrow \mathbb{R}^+$ defined by $(z_1, z_2) \longrightarrow B(z_1, z_2, \mathbf{h})$ is Lipschitz continuous with a Lipschitz constant l_B^g which is independent of $\mathbf{h}$. Moreover, it is bounded by $\frac{\mathcal{A}}{\beta}$.*

Proof : The second part of the proposition follows from the fact that, $\forall i_1, i_2 \in \mathbb{R}^2$ and $\forall \mathbf{h} \in H$, we have (proposition 14):

$$0 \leq A(i_1, i_2, \mathbf{h}) \leq \frac{\mathcal{A}}{\beta}.$$

The first part follows the same pattern as the proof of proposition 15. □

We also have the

Proposition 19 *The function $\mathbf{h} \longrightarrow B(z_1, z_2, \mathbf{h})$, $H \longrightarrow \mathbb{R}$ is Lipschitz continuous with a Lipschitz constant L_B^g which is independent of $(z_1, z_2) \in [0, \mathcal{A}]^2$.*

Proof : The proof follows the same pattern as the one of proposition 16. □

Therefore we can state the

Proposition 20 *The function $B : \mathcal{H}_2 \to \mathbb{R}$ is Lipschitz continuous.*

Proof : The proof follows of propositions 18, 19 and lemma 4. □

From propositions 17, 20 and 12 we obtain the

Corollary 1 *The function $f_{MI}^g : \mathcal{H}_2 \longrightarrow \mathbb{R}$ defined by $(z_1, z_2, \mathbf{h}) \longrightarrow \frac{1}{|\Omega|}(B(z_1, z_2, \mathbf{h}) - b(z_2, \mathbf{h}))$ is Lipschitz continuous and bounded by $2\mathcal{A}/\beta|\Omega|$. We note $Lip(f_{MI}^g)$ the corresponding Lipschitz constant.*

We can now state the main result of this section:

Theorem 7 *The function $F^g_{MI} : H \longrightarrow H$ defined by*

$$F^g_{MI}(\mathbf{h}) = f^g_{MI}(I^\sigma_1, I^\sigma_2(\mathbf{Id}+\mathbf{h}), \mathbf{h})\nabla I^\sigma_2(\mathbf{Id}+\mathbf{h}) = \frac{1}{|\Omega|}(B(I^\sigma_1, I^\sigma_2(\mathbf{Id}+\mathbf{h}), \mathbf{h}) - b(I^\sigma_2(\mathbf{Id}+\mathbf{h}), \mathbf{h}))\nabla I^\sigma_2(\mathbf{Id}+\mathbf{h})$$

is Lipschitz continuous and bounded.

Proof : Boundedness comes from the fact that b and B are bounded (propositions 15 and 18, respectively) and that $|\nabla I^\sigma_2|$ is bounded. This implies that $F^g_{\text{MI}}(\mathbf{h}) \in H = \mathbf{L}^2(\Omega) \forall \mathbf{h} \in H$.

We consider the ith component $F^{g\,i}_{\text{MI}}$ of F_{MI}:

$$F^{g\,i}_{\text{MI}}(\mathbf{h}_1)(\mathbf{x}) - F^{g\,i}_{\text{MI}}(\mathbf{h}_2)(\mathbf{x}) = \frac{1}{|\Omega|}(S_1T_1 - S_2T_2),$$

with

$$\begin{aligned} S_j &= B(I^\sigma_1(\mathbf{x}), I^\sigma_2(\mathbf{x}+\mathbf{h}_j(\mathbf{x})), \mathbf{h}_j) - b(I^\sigma_2(\mathbf{x}+\mathbf{h}_j(\mathbf{x})), \mathbf{h}_j) \\ T_j &= \partial_i I^\sigma_2(\mathbf{x}+\mathbf{h}_j(\mathbf{x}))\ j = 1, 2, \end{aligned}$$

$j = 1, 2$. We continue with

$$|F^{g\,i}_{\text{MI}}(\mathbf{h}_1)(\mathbf{x}) - F^{g\,i}_{\text{MI}}(\mathbf{h}_2)(\mathbf{x})| \leq \frac{1}{|\Omega|}(|S_1 - S_2||T_1| + |S_2||T_1 - T_2|)$$

Because $\partial_i I^\sigma_2$ is bounded, $|T_j| \leq \|\partial_i I^\sigma_2\|_\infty$. Because of propositions 15 and 18, $|S_2| \leq 2\frac{\mathcal{A}}{\beta}$. Because $\partial_i I^\sigma_2$ is Lipschitz continuous $|T_1 - T_2| \leq Lip(\partial_i I^\sigma_2)|\mathbf{h}_1(\mathbf{x}) - \mathbf{h}_2(\mathbf{x})|$.

Finally, because of corollary 1 and the fact that I^σ_2 is Lipschitz continuous,

$$|S_1 - S_2| \leq Lip(f^g_{\text{MI}})\left(Lip(I^\sigma_2)|\mathbf{h}_1(\mathbf{x}) - \mathbf{h}_2(\mathbf{x})| + \|\mathbf{h}_1 - \mathbf{h}_2\|_H\right).$$

Collecting all terms we obtain

$$|F^{g\,i}_{\text{MI}}(\mathbf{h}_1)(\mathbf{x}) - F^{g\,i}_{\text{MI}}(\mathbf{h}_2)(\mathbf{x})| \leq C^i(|\mathbf{h}_1(\mathbf{x}) - \mathbf{h}_2(\mathbf{x})| + \|\mathbf{h}_1 - \mathbf{h}_2\|_H),$$

for some positive constant C^i, $i = 1, \cdots, n$. The last inequality yields, through the application of Cauchy-Schwarz:

$$\|F^g_{\text{MI}}(\mathbf{h}_1) - F^g_{\text{MI}}(\mathbf{h}_2)\|_H \leq L^g_F\|\mathbf{h}_1 - \mathbf{h}_2\|_H$$

for some positive constant L^g_F and this completes the proof. □

The following proposition will be needed later

Proposition 21 *The function $\Omega \to \mathbb{R}^n$ such that $\mathbf{x} \to F^g_{MI}(\mathbf{h}(\mathbf{x}))$ satisfies*

$$|F^g_{MI}(\mathbf{h}(\mathbf{x})) - F^g_{MI}(\mathbf{h}(\mathbf{y}))| \leq K(|\mathbf{x} - \mathbf{y}| + |\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{y})|),$$

for some constant $K > 0$.

Proof : We write

$$F^g_{MI}(\mathbf{h}(\mathbf{x})) - F^g_{MI}(\mathbf{h}(\mathbf{y})) = f^g_{MI}(\mathbf{h}(\mathbf{x}))\nabla I^\sigma_2(\mathbf{x}+\mathbf{h}(\mathbf{x})) - f^g_{MI}(\mathbf{h}(\mathbf{x}))\nabla I^\sigma_2(\mathbf{y}+\mathbf{h}(\mathbf{y})) + f^g_{MI}(\mathbf{h}(\mathbf{x}))\nabla I^\sigma_2(\mathbf{y}+\mathbf{h}(\mathbf{y})) - f^g_{MI}(\mathbf{h}(\mathbf{y}))\nabla I^\sigma_2(\mathbf{y}+\mathbf{h}(\mathbf{y}))$$

Hence

$$|F^g_{MI}(\mathbf{h}(\mathbf{x})) - F^g_{MI}(\mathbf{h}(\mathbf{y}))| \le |f^g_{MI}(\mathbf{h}(\mathbf{x}))|\,|\nabla I^\sigma_2(\mathbf{x}+\mathbf{h}(\mathbf{x})) - \nabla I^\sigma_2(\mathbf{y}+\mathbf{h}(\mathbf{y}))| + |\nabla I^\sigma_2(\mathbf{y}+\mathbf{h}(\mathbf{y}))|\,|f^g_{MI}(\mathbf{h}(\mathbf{x})) - f^g_{MI}(\mathbf{h}(\mathbf{y}))|$$

Corollary 1 and the fact that the functions I^σ_1, I^σ_2 and its first order derivative, are Lipschitz continuous imply

$$|F^g_{MI}(\mathbf{h}(\mathbf{x})) - F^g_{MI}(\mathbf{h}(\mathbf{y}))| \le \frac{2\mathcal{A}}{\beta|\Omega|} Lip(\nabla I^\sigma_2)(|\mathbf{x}-\mathbf{y}| + |\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{y})|) + \||\nabla I^\sigma_2|\|_\infty Lip(f^g_{\mathrm{MI}})|\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{y})|,$$

hence the result. □

8.2.2 Correlation ratio

We produce a simple expression of $\mathbf{CR}^g[\mathbf{h}]$ in terms of the two images I^σ_1 and I^σ_2 and use it to prove that the correlation ratio is Lipschitz continuous as a function of $\mathbf{h}$. In the sequel, we drop the upper index g. We begin with some estimates.

Lemma 6

$$0 \le \mu_2(\mathbf{h}) = \mathbf{E}[X_{I^\sigma_2,\mathbf{h}}] = \frac{1}{|\Omega|}\int_\Omega I^\sigma_2(\mathbf{x}+\mathbf{h}(\mathbf{x}))\,d\mathbf{x} \le \mathcal{A}$$

$$\beta \le v_2(\mathbf{h}) = \mathbf{Var}[X_{I^\sigma_2,\mathbf{h}}] = \beta + \mathbf{Var}[I^\sigma_2(\mathbf{Id}+\mathbf{h})] \le \beta + \mathcal{A}^2,$$

where

$$\mathbf{Var}[I^\sigma_2(\mathbf{Id}+\mathbf{h})] = \frac{1}{|\Omega|}\int_\Omega I^\sigma_2(\mathbf{x}+\mathbf{h}(\mathbf{x}))^2\,d\mathbf{x} - \left(\frac{1}{|\Omega|}\int_\Omega I^\sigma_2(\mathbf{x}+\mathbf{h}(\mathbf{x}))\,d\mathbf{x}\right)^2$$

Proof : Because of equation (7) we have

$$\mu_2(\mathbf{h}) = \frac{1}{|\Omega|}\int_{\mathbb{R}} i_2\left(\int_\Omega g_\beta(i_2 - I^\sigma_2(\mathbf{x}+\mathbf{h}(\mathbf{x})))\,d\mathbf{x}\right)di_2 = \frac{1}{|\Omega|}\int_\Omega I^\sigma_2(\mathbf{x}+\mathbf{h}(\mathbf{x}))\,d\mathbf{x}$$

This yields the first part of the lemma. For the second part, we use equation (8):

$$v_2(\mathbf{h}) = \mathbf{Var}[X_{I^\sigma_2,\mathbf{h}}] = \int_{\mathbb{R}} i_2^2\left(\frac{1}{|\Omega|}\int_\Omega g_\beta(i_2 - I^\sigma_2(\mathbf{x}+\mathbf{h}(\mathbf{x}))\,d\mathbf{x}\right)di_2 - \mu_2(\mathbf{h})^2,$$

and hence

$$\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}] = \beta + \frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))^2\,d\mathbf{x} - \left(\frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))\,d\mathbf{x}\right)^2,$$

from which the upper and lower bounds of the lemma follow. □

We next take care of $\mathbf{E}[\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}|X_{I_1^\sigma}]]$ with the following

Lemma 7

$$\mathbf{E}[\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}|X_{I_1^\sigma}]] = \beta + \frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))^2\,d\mathbf{x} - M[\mathbf{h}],$$

where

$$M[\mathbf{h}] = \int_{\Omega\times\Omega} f(\mathbf{x},\mathbf{x}')I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))I_2^\sigma(\mathbf{x}'+\mathbf{h}(\mathbf{x}'))\,d\mathbf{x}\,d\mathbf{x}',$$

and

$$f(\mathbf{x},\mathbf{x}') = \frac{1}{|\Omega|^2}\int_{\mathbb{R}} \frac{g_\beta(i_1 - I_1^\sigma(\mathbf{x}))g_\beta(i_1 - I_1^\sigma(\mathbf{x}'))}{p(i_1)}\,di_1,$$

is such that $\int_\Omega f(\mathbf{x},\mathbf{x}')\,d\mathbf{x} = \int_\Omega f(\mathbf{x},\mathbf{x}')\,d\mathbf{x}' = \frac{1}{|\Omega|}$

Proof : According to table 1 we have $v_{2|1}(i_1,\mathbf{h}) = S(i_1) - T^2(i_1)$ and

$$\mathbf{E}[\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}|X_{I_1^\sigma}]] = \int_{\mathbb{R}} v_{2|1}(i_1,\mathbf{h})p(i_1)\,di_1,$$

with

$$S(i_1) = \int_{\mathbb{R}} i_2^2 \frac{P_{\mathbf{h}}(i_1,i_2)}{p(i_1)}\,di_2,$$

and

$$T(i_1) = \int_{\mathbb{R}} i_2 \frac{P_{\mathbf{h}}(i_1,i_2)}{p(i_1)}\,di_2 = \mu_{2|1}(i_1,\mathbf{h}).$$

It is straightforward to show that

$$\int_{\mathbb{R}} S(i_1)p(i_1)\,di_1 = \beta + \frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))^2\,d\mathbf{x}.$$

It is also straightforward to show that

$$T(i_1) = \frac{1}{|\Omega|p(i_1)}\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))\,d\mathbf{x},$$

and hence that

$$\int_{\mathbb{R}} T^2(i_1)p(i_1)\,di_1 = \frac{1}{|\Omega|^2}\int_{\mathbb{R}} \frac{1}{p(i_1)}\left(\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))\,d\mathbf{x}\right)^2 di_1.$$

We next write

$$\left(\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))\,d\mathbf{x}\right)^2 = \int_{\Omega\times\Omega} g_\beta(i_1 - I_1^\sigma(\mathbf{x}))I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))g_\beta(i_1 - I_1^\sigma(\mathbf{x}'))I_2^\sigma(\mathbf{x}'+\mathbf{h}(\mathbf{x}'))\,d\mathbf{x}\,d\mathbf{x}',$$

commute the integration with respect to i_1 with that with respect to $\mathbf{x}$ and $\mathbf{x}'$ to obtain the result. □

We pursue with another

Lemma 8 *The function $M : H \longrightarrow \mathbb{R}$ defined in lemma 7 is bounded and Lipschitz continuous.*

Proof : For the first part, $|M[\mathbf{h}]| \le \mathcal{A}^2 \int_{\Omega\times\Omega} f(\mathbf{x},\mathbf{x}')\,d\mathbf{x}\,d\mathbf{x}' = \mathcal{A}^2$, according to lemma 7. For the second part we compute

$$M[\mathbf{h}_1] - M[\mathbf{h}_2] = \int_{\Omega\times\Omega} f(\mathbf{x},\mathbf{x}')\Big(I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x}))I_2^\sigma(\mathbf{x}'+\mathbf{h}_1(\mathbf{x}')) - I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x}))I_2^\sigma(\mathbf{x}'+\mathbf{h}_2(\mathbf{x}'))\Big)\,d\mathbf{x}\,d\mathbf{x}'$$

$$|M[\mathbf{h}_1] - M[\mathbf{h}_2]| \le \int_{\Omega\times\Omega} f(\mathbf{x},\mathbf{x}')(|I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x})) - I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x}))|I_2^\sigma(\mathbf{x}'+\mathbf{h}_1(\mathbf{x}'))+ |I_2^\sigma(\mathbf{x}'+\mathbf{h}_1(\mathbf{x}')) - I_2^\sigma(\mathbf{x}'+\mathbf{h}_2(\mathbf{x}'))|I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x})))\,d\mathbf{x}\,d\mathbf{x}'$$

Because I_2^σ is Lipschitz continuous and bounded

$$|M[\mathbf{h}_1] - M[\mathbf{h}_2]| \le \|I_2^\sigma\|_\infty Lip(I_2^\sigma)\int_{\Omega\times\Omega} f(\mathbf{x},\mathbf{x}')\left(|\mathbf{h}_1(\mathbf{x})-\mathbf{h}_2(\mathbf{x})| + |\mathbf{h}_1(\mathbf{x}')-\mathbf{h}_2(\mathbf{x}')|\right)\,d\mathbf{x}\,d\mathbf{x}' = \frac{\|I_2^\sigma\|_\infty Lip(I_2^\sigma)}{|\Omega|}\left(\int_\Omega |\mathbf{h}_1(\mathbf{x})-\mathbf{h}_2(\mathbf{x})|\,d\mathbf{x} + \int_\Omega |\mathbf{h}_1(\mathbf{x}')-\mathbf{h}_2(\mathbf{x}')|\,d\mathbf{x}'\right) = \frac{2\|I_2^\sigma\|_\infty Lip(I_2^\sigma)}{|\Omega|}\int_\Omega |\mathbf{h}_1(\mathbf{x})-\mathbf{h}_2(\mathbf{x})|\,d\mathbf{x}$$

The previous to the last equality is obtained from lemma 7. Therefore we have from Cauchy-Schwarz:

$$|M[\mathbf{h}_1] - M[\mathbf{h}_2]| \le \frac{2\|I_2^\sigma\|_\infty Lip(I_2^\sigma)}{|\Omega|^{1/2}}\|\mathbf{h}_1-\mathbf{h}_2\|_H,$$

□

We continue with the following

Lemma 9 *The functions $H \longrightarrow \mathbb{R}^+$ defined by*

$$\mathbf{h} \longrightarrow \frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))\ d\mathbf{x} \quad \textit{and} \quad \mathbf{h} \longrightarrow \frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))^2\ d\mathbf{x}$$

are bounded and Lipschitz continuous.

Proof : Boundedness has been proved in lemma 6 for the first function. For the second we have

$$\frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))^2\ d\mathbf{x} \le \mathcal{A}^2$$

Next, for Lipschitz continuity:

$$\left|\frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x}))\ d\mathbf{x} - \frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x}))\ d\mathbf{x}\right| \le Lip(I_2^\sigma)\frac{1}{|\Omega|}\int_\Omega |\mathbf{h}_1(\mathbf{x})-\mathbf{h}_2(\mathbf{x})|\ d\mathbf{x},$$

because I_2^σ is Lipschitz, and hence (Cauchy-Schwarz)

$$\left|\frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x}))\ d\mathbf{x} - \frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x}))\ d\mathbf{x}\right| \le \frac{Lip(I_2^\sigma)}{|\Omega|^{1/2}}\|\mathbf{h}_1-\mathbf{h}_2\|_H.$$

Similarly (Cauchy-Schwarz)

$$\left|\frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}_1(\mathbf{x}))^2\ d\mathbf{x} - \frac{1}{|\Omega|}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}_2(\mathbf{x}))^2\ d\mathbf{x}\right| \le 2\ \frac{\mathcal{A}Lip(I_2^\sigma)}{|\Omega|^{1/2}}\|\mathbf{h}_1-\mathbf{h}_2\|_H.$$

□

From this lemma, proposition 12 and lemma 6 we deduce the following

Corollary 2 *The function $H \longrightarrow \mathbb{R}$ defined by $\mathbf{h} \longrightarrow \mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}]$ is Lipschitz continuous.*

We also prove the

Lemma 10 *The function $H \longrightarrow \mathbb{R}$ defined by $\mathbf{h} \longrightarrow \mathbf{E}[\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}|X^g_{I_1^\sigma}]]$ is bounded and Lipschitz continuous.*

Proof : This follows from lemmas 7, 8, and 9. □

We can now prove the important intermediary result that the correlation ratio, as a function of the field $\mathbf{h}$, is Lipschitz continuous.

Theorem 8 *The function $H \longrightarrow \mathbb{R}$ defined by $\mathbf{h} \longrightarrow \frac{\mathbf{E}[\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}|X^g_{I_1^\sigma}]]}{\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}]}$ is Lipschitz continuous.*

Proof : This follows from proposition 12 and from lemmas 10, 6 and corollary 2. □

We pursue with another

Lemma 11 *The function $f^g_{CR} = G_\beta \star L^g_{CR} : \mathcal{H}_2 \longrightarrow \mathbb{R}$, where L^g_{CR} is given by equation (17), is equal to the following expression*

$$f^g_{CR}(z_1, z_2, \mathbf{h}) = -\frac{2}{|\Omega|\, v_2(\mathbf{h})} \left(z_2 \mathbf{CR}[\mathbf{h}] - d(z_1, \mathbf{h}) + \mu_2(\mathbf{h})(1 - \mathbf{CR}[\mathbf{h}])\right),$$

where

$$d(z_1, \mathbf{h}) = \int_{\mathbb{R}} g_\beta(z_1 - i_1) \left(\frac{\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x})) I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x}))\, d\mathbf{x}}{\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))\, d\mathbf{x}} \right) di_1.$$

Proof : We use equation (17) and apply the convolution to it. The value of d is obtained from:

$$d(z_1, \mathbf{h}) = \int_{\mathbb{R}} g_\beta(z_1 - i_1) \mu_{2|1}(i_1, \mathbf{h})\, di_1 = \int_{\mathbb{R}} \frac{g_\beta(z_1 - i_1)}{p(i_1)} \left(\int_{\mathbb{R}} i_2 P_{\mathbf{h}}(i_1, i_2)\, di_2 \right) di_1 = \int_{\mathbb{R}} \frac{g_\beta(z_1 - i_1)}{p(i_1)} \left(\int_\Omega I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) g_\beta(i_1 - I_1^\sigma(\mathbf{x}))\, d\mathbf{x} \right) di_1$$

□

We next prove the following

Lemma 12 *The function $d : \mathcal{H}_1 \longrightarrow \mathbb{R}$ is bounded and Lipschitz continuous.*

Proof : The proof of the first part uses exactly the same ideas as those of the second part of proposition 16. For the second part, we first prove that the function $H \longrightarrow \mathbb{R}$, $\mathbf{h} \longrightarrow d(z_1, \mathbf{h})$ is Lipschitz continuous with a Lipschitz constant L_d that is independent of $z_1 \in [0, \mathcal{A}]$ and second prove that the function $[0, \mathcal{A}] \longrightarrow \mathbb{R}$, $z_1 \longrightarrow \frac{\partial d}{\partial z_1}$ is upperbounded independently of $\mathbf{h} \in H$. Indeed,

$$|d(z_1, \mathbf{h}_1) - d(z_1, \mathbf{h}_2)| \leq \int_{\mathbb{R}} g_\beta(z_1 - i_1) \left(\frac{\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x})) |I_2^\sigma(\mathbf{x} + \mathbf{h}_1(\mathbf{x})) - I_2^\sigma(\mathbf{x} + \mathbf{h}_2(\mathbf{x}))|\, d\mathbf{x}}{\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))\, d\mathbf{x}} \right) di_1$$

Because I_2^σ is Lipschitz continuous and of Schwarz inequality, we have

$$|d(z_1, \mathbf{h}_1) - d(z_1, \mathbf{h}_2)| \leq |\Omega| Lip(I_2^\sigma) \|\mathbf{h}_1 - \mathbf{h}_2\|_H \int_{\mathbb{R}} g_\beta(z_1 - i_1) \left(\frac{\left(\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))^2\, d\mathbf{x} \right)^{\frac{1}{2}}}{\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))\, d\mathbf{x}} \right) di_1$$

The function of z_1 that appears on the righthand side of this inequality does not depend on $\mathbf{h}$, is continuous and therefore bounded on $[0, \mathcal{A}]$.

We now notice that

$$\frac{\partial d}{\partial z_1} = \frac{1}{\beta}\int_{\mathbb{R}}(i_1 - z_1)g_\beta(z_1 - i_1)\left(\frac{\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))\,d\mathbf{x}}{\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))\,d\mathbf{x}}\right)\,di_1,$$

and, since

$$\frac{\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))\,d\mathbf{x}}{\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))\,d\mathbf{x}} \le \mathcal{A},$$

$$\left|\frac{\partial d}{\partial z_1}\right| \le \frac{\mathcal{A}}{\beta}\int_{\mathbb{R}}|z_1 - i_1|g_\beta(z_1 - i_1)\,di_1.$$

The righthand side of this inequality is equal to $2\frac{\mathcal{A}}{\beta}\int_0^{+\infty} z g_\beta(z)dz$ from where the conclusion follows. □

This allows us to prove the

Theorem 9 *The function $f_{CR}^g : \mathcal{H}_2 \longrightarrow \mathbb{R}$ is Lipschitz continuous and bounded.*

Proof : The denominator $|\Omega|v_2(\mathbf{h})$ is > 0 and bounded (lemma 6), and Lipschitz continuous on H and K (corollary 2).

The numerator is bounded because $\mathbf{CR}[\mathbf{h}]$ is bounded by 1, d is bounded (lemma 12) and $\mu_2(\mathbf{h})$ is bounded (lemma 6).

The product $\mu_2(\mathbf{h})(1 - \mathbf{CR}[\mathbf{h}])$ is Lipschitz continuous on H as the product of two bounded Lipschitz continuous functions (lemma 12 and theorem 8). Hence we have proved the boundedness of f_{CR}^g.

The function $d : \mathcal{H}_1 \to \mathbb{R}$ is Lipschitz continuous (lemma 12).

The function $r : \mathcal{H}_1 \longrightarrow \mathbb{R}$ defined by $(z_2, \mathbf{h}) \longrightarrow z_2\mathbf{CR}[\mathbf{h}]$ is Lipschitz continuous because of theorem 8 and

$$\begin{aligned}|z_2\mathbf{CR}[\mathbf{h}] - z_2'\mathbf{CR}[\mathbf{h}']| = \\ |z_2(\mathbf{CR}[\mathbf{h}] - \mathbf{CR}[\mathbf{h}']) + \mathbf{CR}[\mathbf{h}'](z_2 - z_2')| \le \\ \mathcal{A}|\mathbf{CR}[\mathbf{h}] - \mathbf{CR}[\mathbf{h}']| + |z_2 - z_2'|.\end{aligned}$$

Hence the numerator is also Lipschitz continuous and, from proposition 12, so is f_{CR}^g. □

We finally obtain the

Theorem 10 *The function $F_{CR}^g : H \longrightarrow H$ defined by*

$$F_{CR}^g(\mathbf{h}) = f_{CR}^g(I_1^\sigma, I_2^\sigma(\mathbf{Id}+\mathbf{h}))\nabla I_2^\sigma(\mathbf{Id}+\mathbf{h})$$

is Lipschitz continuous and bounded.

Proof : The boundedness follows from theorem 9 and the fact that $|\nabla I_2^\sigma|$ is bounded. It implies that $F_{\text{CR}}^g(\mathbf{h}) \in H \forall \mathbf{h} \in H$. The rest of the proof follows exactly the same pattern as the proof of theorem 7 and uses theorem 9. □

We finish this section with the following

Proposition 22 *The function* $\Omega \to \mathbb{R}^n$ *such that* $\mathbf{x} \to F_{CR}^g(\mathbf{h}(\mathbf{x}))$ *satisfies*

$$|F_{CR}^g(\mathbf{h}(\mathbf{x})) - F_{CR}^g(\mathbf{h}(\mathbf{y}))| \leq K(|\mathbf{x}-\mathbf{y}| + |\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{y})|),$$

for some constant $K > 0$.

Proof : The proof is similar to that of proposition 21 and follows from theorem 9 and the fact that the functions I_1^σ, I_2^σ and all its derivatives, are Lipschitz continuous. □

8.3 Local criteria

The analysis of the local criteria follows pretty much directly from the analysis of the global ones and from theorem 4. The main difference with the global case is that we have an extra spatial dependency. In the next lemma we introduce a constant that is needed in the sequel.

Lemma 13 *Let diam*(Ω) *be the diameter of the open bounded set* Ω*:*

$$\textit{diam}(\Omega) = \sup_{\mathbf{x},\mathbf{y} \in \Omega} \|\mathbf{x} - \mathbf{y}\|.$$

We note $G_\gamma(\textit{diam}(\Omega))$ *the value* $\inf_{\mathbf{x},\mathbf{y}\in\Omega} G_\gamma(\mathbf{x}-\mathbf{y})$ *and define*

$$K_\Omega = \frac{G_\gamma(\mathbf{0})}{G_\gamma(\textit{diam}(\Omega))}. \tag{28}$$

We say that

$$\int_\Omega G_\gamma(\mathbf{x}-\mathbf{x}_0)\frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)}\,d\mathbf{x}_0 \leq K_\Omega \quad \forall \mathbf{x} \in \Omega$$

Proof : Since $\mathcal{G}_\gamma(\mathbf{x}_0) = \int_\Omega G_\gamma(\mathbf{y}-\mathbf{x}_0)\,d\mathbf{y}$, we have $\mathcal{G}_\gamma(\mathbf{x}_0) \geq |\Omega| G_\gamma(\text{diam}(\Omega))$. Therefore

$$\int_\Omega G_\gamma(\mathbf{x}-\mathbf{x}_0)\frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)}\,d\mathbf{x}_0 \leq \frac{1}{|\Omega| G_\gamma(\text{diam}(\Omega))}\int_\Omega G_\gamma(\mathbf{x}-\mathbf{x}_0)d\mathbf{x}_0 \leq \frac{1}{|\Omega| G_\gamma(\text{diam}(\Omega))} \times |\Omega| G_\gamma(\mathbf{0}) = K_\Omega$$

□

8.3.1 Mutual information

The functions $q_{\mathbf{h}}$ and $Q_{\mathbf{h}}$ defined in propositions 13 and 14 are now functions of $\mathbf{x}_0$ but the propositions are unchanged. The function $a(i_2, \mathbf{x}_0, \mathbf{h})$ defined in proposition 13 is equal to:

$$a(i_2, \mathbf{x}_0, \mathbf{h}) = \frac{1}{\beta} \frac{\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x})) g_\beta(I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x})) - i_2)\, G_\gamma(\mathbf{x}-\mathbf{x}_0)\, d\mathbf{x}}{\int_\Omega g_\beta(I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x})) - i_2)\, G_\gamma(\mathbf{x}-\mathbf{x}_0)\, d\mathbf{x}}$$

Similarly, the function $L^g_{\text{MI, }\mathbf{h}}(\mathbf{i})$ defined by equation (22) must be modified as follows:

$$L^l_{\text{MI, }\mathbf{h}}(\mathbf{i}, \mathbf{x}_0) = \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} (A(\mathbf{i}, \mathbf{x}_0, \mathbf{h}) - a(i_2, \mathbf{x}_0, \mathbf{h})),$$

as well as the function b of equation (23):

$$b(z_2, \mathbf{x}, \mathbf{h}) = (G_\gamma \star g_\beta \star \frac{a}{\mathcal{G}_\gamma})(z_2, \mathbf{x}, \mathbf{h}) = \int_\Omega \int_{\mathbb{R}} G_\gamma(\mathbf{x}-\mathbf{x}_0) g_\beta(z_2 - i_2) \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} a(i_2, \mathbf{x}_0, \mathbf{h})\, di_2\, d\mathbf{x}_0, \quad (29)$$

and the function B of equation (24):

$$B(\mathbf{z}, \mathbf{x}, \mathbf{h}) = (G_\gamma \star G_\beta \star \frac{A}{\mathcal{G}_\gamma})(\mathbf{z}, \mathbf{x}, \mathbf{h}) = \int_\Omega \int_{\mathbb{R}^2} G_\gamma(\mathbf{x}-\mathbf{x}_0) G_\beta(\mathbf{z}-\mathbf{i}) \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} A(\mathbf{i}, \mathbf{x}_0, \mathbf{h})\, d\mathbf{i}\, d\mathbf{x}_0. \quad (30)$$

This being done, propositions 15 and 16 can be adapted to the present case as follows

Proposition 23 *The function $[0, \mathcal{A}] \longrightarrow \mathbb{R}^+$ defined by $z_2 \longrightarrow b(z_2, \mathbf{x}, \mathbf{h})$ is Lipschitz continuous with a Lipschitz constant l_b^l which is independent of $\mathbf{h}$ and $\mathbf{x}$. Moreover, it is bounded by $\frac{\mathcal{A}}{\beta}$.*

Proof : We give the proof in this particular simple case to give the flavour of the ideas which extend to the more complicated cases that come later.

The second part of the proposition follows from the fact that $0 \le a(i_2, \mathbf{x}, \mathbf{h}) \le \frac{\mathcal{A}}{\beta}\ \forall i_2 \in \mathbb{R}$ and $\forall \mathbf{h} \in H$ (local version of proposition 13) and lemma 13.

In order to prove the first part, we prove that the magnitude of the derivative of the function is bounded independently of $\mathbf{h}$ and $\mathbf{x}$. Indeed

$$\left| \frac{\partial(z_2, \mathbf{x}, \mathbf{h})}{\partial z_2} \right| = \frac{1}{\beta} \left| \int_\Omega \int_{\mathbb{R}} (i_2 - z_2) G_\gamma(\mathbf{x}-\mathbf{x}_0) \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} g_\beta(z_2 - i_2) a(i_2, \mathbf{x}_0, \mathbf{h}) di_2\, d\mathbf{x}_0 \right| \le \frac{\mathcal{A} K_\Omega}{\beta} \int_{\mathbb{R}} |z_2 - i_2| g_\beta(z_2 - i_2)\, di_2$$

In order to derive the last inequality we have used the local version of proposition 13 and lemma 13. The function on the righthand side of the last inequality is independent of $\mathbf{h}$ and $\mathbf{x}$ and continuous on $[0, \mathcal{A}]$, therefore bounded. $\square$

Proposition 24 *The function $H \longrightarrow \mathbb{R}$ defined by $\mathbf{h} \longrightarrow b(z_2, \mathbf{x}, \mathbf{h})$ is Lipschitz continuous with a Lipschitz constant L_b^l which is independent of $z_2 \in [0, \mathcal{A}]$ and $\mathbf{x} \in \Omega$.*

Proof : The proof is similar to that of proposition 16. $\square$

We also have the

Proposition 25 *The function $\Omega \rightarrow \mathbb{R}$ defined by $\mathbf{x} \rightarrow b(z_2, \mathbf{x}, \mathbf{h})$ is Lipschitz continuous uniformly on $\mathcal{H}_1$.*

Proof : Because of equation (29) and proposition 23 we have

$$|b(z_2, \mathbf{x}, \mathbf{h}) - b(z_2, \mathbf{y}, \mathbf{h})| \leq \frac{\mathcal{A}}{\beta} \int_\Omega \int_\mathbb{R} \frac{|G_\gamma(\mathbf{x} - \mathbf{x}_0) - G_\gamma(\mathbf{y} - \mathbf{x}_0)|}{\mathcal{G}_\gamma(\mathbf{x}_0)} g_\beta(z_2 - i_2)\, di_2\, d\mathbf{x}_0$$

The proof of lemma 13 allows us to write

$$|b(z_2, \mathbf{x}, \mathbf{h}) - b(z_2, \mathbf{y}, \mathbf{h})| \leq \frac{\mathcal{A}}{\beta |\Omega| G_\gamma(\text{diam}(\Omega))} \int_\Omega \int_\mathbb{R} |G_\gamma(\mathbf{x} - \mathbf{x}_0) - G_\gamma(\mathbf{y} - \mathbf{x}_0)| g_\beta(z_2 - i_2)\, di_2\, d\mathbf{x}_0,$$

and therefore

$$|b(z_2, \mathbf{x}, \mathbf{h}) - b(z_2, \mathbf{y}, \mathbf{h})| \leq \frac{\mathcal{A} Lip(G_\gamma)}{\beta G_\gamma(\text{diam}(\Omega))} |\mathbf{x} - \mathbf{y}|.$$

$\square$

As a consequence of lemma 5 we can state the following

Proposition 26 *The function $b : \mathcal{H}_1 \times \Omega \rightarrow \mathbb{R}$ is Lipschitz continuous on $\mathcal{H}_1$ uniformly on Ω.*

Proof : The proof follows from lemma 5 and proposition 23 and 24. $\square$

Similarly we have the

Proposition 27 *The function $\Omega \rightarrow \mathbb{R}$ defined by $\mathbf{x} \rightarrow B(z_1, z_2, \mathbf{x}, \mathbf{h})$ is Lipschitz continuous uniformly on $\mathcal{H}_2$.*

Proof : The proof is similar to that of proposition 25. $\square$

The following proposition is also needed.

Proposition 28 *The function $B : \mathcal{H}_2 \times \Omega \rightarrow \mathbb{R}$ is Lipschitz continuous on $\mathcal{H}_2$ uniformly on Ω. It is bounded by $\frac{\mathcal{A}}{\beta}$.*

Proof : The proof is similar to that of proposition 26. □

And therefore

Corollary 3 *The function $f^l_{MI} : \mathcal{H}_2 \times \Omega \longrightarrow \mathbb{R}$ defined by*

$$(z_1, z_2, \mathbf{x}, \mathbf{h}) \longrightarrow B(z_1, z_2, \mathbf{x}, \mathbf{h}) - b(z_2, \mathbf{x}, \mathbf{h})$$

is Lipschitz continuous on $\mathcal{H}_2$ uniformly on Ω and bounded.

From this follows the local version of theorem 7

Theorem 11 *The function $F^l_{MI} : H \longrightarrow H$ defined by*

$$F^l_{MI}(\mathbf{h}) = f^l_{MI}(I^\sigma_1, I^\sigma_2(\mathbf{Id} + \mathbf{h}), \mathbf{Id}, \mathbf{h}) = \\ (B(I^\sigma_1, I^\sigma_2(\mathbf{Id} + \mathbf{h}), \mathbf{Id}, \mathbf{h}) - b(I^\sigma_2(\mathbf{Id} + \mathbf{h}), \mathbf{Id}, \mathbf{h}))\nabla I^\sigma_2(\mathbf{Id} + \mathbf{h})$$

is Lipschitz continuous and bounded.

Proof : Boundedness follows from corollary 3 and the fact that $|\nabla I^\sigma_2|$ is bounded. It implies that $F^l_{\text{MI}}(\mathbf{h}) \in H \forall \mathbf{h} \in H$.

We next consider the ith component $F^{l\,i}_{\text{MI}}$ of F^l_{MI}:

$$F^{l\,i}_{\text{MI}}(\mathbf{h}_1)(\mathbf{x}) - F^{l\,i}_{\text{MI}}(\mathbf{h}_2)(\mathbf{x}) = S_1 T_1 - S_2 T_2,$$

with

$$\begin{aligned} S_j &= B(I^\sigma_1(\mathbf{x}), I^\sigma_2(\mathbf{x} + \mathbf{h}_j(\mathbf{x})), \mathbf{x}, \mathbf{h}_j) - b(I^\sigma_2(\mathbf{x} + \mathbf{h}_j(\mathbf{x})), \mathbf{x}, \mathbf{h}_j) \\ T_j &= \partial_i I^\sigma_2(\mathbf{x} + \mathbf{h}_j(\mathbf{x}))\ j = 1, 2, \end{aligned}$$

$j = 1, 2$. We continue with

$$|F^{l\,i}_{\text{MI}}(\mathbf{h}_1)(\mathbf{x}) - F^{l\,i}_{\text{MI}}(\mathbf{h}_2)(\mathbf{x})| \le |S_1 - S_2||T_1| + |S_2||T_1 - T_2|$$

Because $\partial_i I^\sigma_2$ is bounded, $|T_j| \le \|\partial_i I^\sigma_2\|_\infty$. Because of propositions 23 and 28, $|S_2| \le 2\frac{\mathcal{A}}{\beta}$. Because $\partial_i I^\sigma_2$ is Lipschitz continuous $|T_1 - T_2| \le Lip(\partial_i I^\sigma_2)|\mathbf{h}_1(\mathbf{x}) - \mathbf{h}_2(\mathbf{x})|$. Finally, because of corollary 3 and the fact that I^σ_2 is Lipschitz continuous,

$$|S_1 - S_2| \le Lip(f^l_{\text{MI}})\left(Lip(I^\sigma_2)|\mathbf{h}_1(\mathbf{x}) - \mathbf{h}_2(\mathbf{x})| + \|\mathbf{h}_1 - \mathbf{h}_2\|_H\right).$$

The conclusion of the theorem follows from these inequalities through the same procedures as in the proof of theorem 7. □

We finish this section with the

Proposition 29 *The function $\Omega \to \mathbb{R}^n$ such that $\mathbf{x} \to F^l_{MI}(\mathbf{h}(\mathbf{x}))$ satisfies*

$$|F^l_{MI}(\mathbf{h}(\mathbf{x})) - F^l_{MI}(\mathbf{h}(\mathbf{y}))| \le K(|\mathbf{x} - \mathbf{y}| + |\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{y})|),$$

for some constant $K > 0$.

Proof : The proof is similar to that of proposition 21 and uses propositions 25 and 27. □ ‘

8.3.2 Correlation ratio

In this case also, the proofs follow pretty much the same pattern as those in the global case. In detail, the analog of lemma 6 is the

Lemma 14

$$0 \le \mu_2(\mathbf{x}_0, \mathbf{h}) = \mathbf{E}[X_{I_2^\sigma,\mathbf{h}}](\mathbf{x}_0) = \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} \int_\Omega I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) G_\gamma(\mathbf{x} - \mathbf{x}_0)\, d\mathbf{x} \le \mathcal{A}$$

$$\beta \le v_2(\mathbf{x}_0, \mathbf{h}) = \mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}](\mathbf{x}_0) = \beta + \mathbf{Var}[I_2^\sigma(\mathbf{Id} + \mathbf{h})](\mathbf{x}_0) \le \beta + \mathcal{A}^2,$$

where

$$\mathbf{Var}[I_2^\sigma(\mathbf{Id} + \mathbf{h})](\mathbf{x}_0) = \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} \int_\Omega I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x}))^2 G_\gamma(\mathbf{x} - \mathbf{x}_0)\, d\mathbf{x} - \left(\frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} \int_\Omega I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) G_\gamma(\mathbf{x} - \mathbf{x}_0)\, d\mathbf{x} \right)^2$$

Proof : Because of equation (11) we have

$$\mu_2(\mathbf{x}_0, \mathbf{h}) = \int_\mathbb{R} i_2 \left(\frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} \int_\Omega g_\beta(i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x}))) G_\gamma(\mathbf{x} - \mathbf{x}_0)\, d\mathbf{x} \right) di_2 = \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} \int_\Omega I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) G_\gamma(\mathbf{x} - \mathbf{x}_0)\, d\mathbf{x}$$

This yields the first part of the lemma. For the second part, we use equation (12):

$$\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}](\mathbf{x}_0) = \int_\mathbb{R} i_2^2 \left(\frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} \int_\Omega g_\beta(i_2 - I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) G_\gamma(\mathbf{x} - \mathbf{x}_0)\, d\mathbf{x} \right) di_2 - \mu_2(\mathbf{x}_0, \mathbf{h})^2,$$

and hence

$$\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}](\mathbf{x}_0) = \beta + \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} \int_\Omega I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x}))^2 G_\gamma(\mathbf{x} - \mathbf{x}_0)\, d\mathbf{x} - \left(\frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} \int_\Omega I_2^\sigma(\mathbf{x} + \mathbf{h}(\mathbf{x})) G_\gamma(\mathbf{x} - \mathbf{x}_0)\, d\mathbf{x} \right)^2,$$

from which the upper and lower bounds of the lemma follow. □

We next take care of $\mathbf{E}[\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}|X_{I_1^\sigma}](\mathbf{x}_0)]$ with the analog of lemma 7

Lemma 15

$$\mathbf{E}[\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}|X_{I_1^\sigma}](\mathbf{x}_0)] = \beta + \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))^2 G_\gamma(\mathbf{x}-\mathbf{x}_0)\, d\mathbf{x} - M[\mathbf{x}_0,\mathbf{h}],$$

where

$$M[\mathbf{x}_0,\mathbf{h}] = \int_{\Omega\times\Omega} f(\mathbf{x}_0,\mathbf{x},\mathbf{x}') I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x})) I_2^\sigma(\mathbf{x}'+\mathbf{h}(\mathbf{x}'))\, d\mathbf{x}\, d\mathbf{x}',$$

and

$$f(\mathbf{x}_0,\mathbf{x},\mathbf{x}') = \frac{G_\gamma(\mathbf{x}-\mathbf{x}_0)G_\gamma(\mathbf{x}'-\mathbf{x}_0)}{\mathcal{G}_\gamma(\mathbf{x}_0)^2}\int_{\mathbb{R}} \frac{g_\beta(i_1 - I_1^\sigma(\mathbf{x}))g_\beta(i_1 - I_1^\sigma(\mathbf{x}'))}{p(i_1,\mathbf{x}_0)}\, di_1,$$

is such that

$$\begin{aligned}\int_\Omega f(\mathbf{x}_0,\mathbf{x},\mathbf{x}')\, d\mathbf{x} &= \frac{G_\gamma(\mathbf{x}'-\mathbf{x}_0)}{\mathcal{G}_\gamma(\mathbf{x}_0)}\\ \int_\Omega f(\mathbf{x}_0,\mathbf{x},\mathbf{x}')\, d\mathbf{x}' &= \frac{G_\gamma(\mathbf{x}-\mathbf{x}_0)}{\mathcal{G}_\gamma(\mathbf{x}_0)}\end{aligned}$$

and

$$\int_{\Omega\times\Omega} f(\mathbf{x}_0,\mathbf{x},\mathbf{x}')\, d\mathbf{x}\, d\mathbf{x}' = 1.$$

Proof : According to table 2 we have $v_{2|1}(i_1,\mathbf{x}_0,\mathbf{h}) = S(i_1,\mathbf{x}_0) - T^2(i_1,\mathbf{x}_0)$ and

$$\mathbf{E}[\mathbf{Var}[X_{I_2^\sigma,\mathbf{h}}|X_{I_1^\sigma}](\mathbf{x}_0)] = \int_{\mathbb{R}} v_{2|1}(i_1,\mathbf{x}_0,\mathbf{h}) p(i_1,\mathbf{x}_0)\, di_1,$$

with

$$S(i_1,\mathbf{x}_0) = \int_{\mathbb{R}} i_2^2 \frac{P_{\mathbf{h}}(i_1,i_2,\mathbf{x}_0)}{p(i_1,\mathbf{x}_0)}\, di_2,$$

and

$$T(i_1,\mathbf{x}_0) = \int_{\mathbb{R}} i_2 \frac{P_{\mathbf{h}}(i_1,i_2,\mathbf{x}_0)}{p(i_1,\mathbf{x}_0)}\, di_2 = \mu_{2|1}(i_1,\mathbf{x}_0,\mathbf{h}).$$

It is straightforward to show that

$$\int_{\mathbb{R}} S(i_1,\mathbf{x}_0)p(i_1,\mathbf{x}_0)\, di_1 = \beta + \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)}\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))^2 G_\gamma(\mathbf{x}-\mathbf{x}_0)\, d\mathbf{x}.$$

It is also straightforward to show that

$$T(i_1) = \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)p(i_1,\mathbf{x}_0)}\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x})) I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x})) G_\gamma(\mathbf{x}-\mathbf{x}_0)\, d\mathbf{x},$$

and hence that

$$\int_{\mathbb{R}} T^2(i_1,\mathbf{x}_0)p(i_1,\mathbf{x}_0)\,di_1 = \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)^2}\int_{\mathbb{R}}\frac{1}{p(i_1)}\left(\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))G_\gamma(\mathbf{x}-\mathbf{x}_0)\,d\mathbf{x}\right)^2 di_1.$$

We next write

$$\left(\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))G_\gamma(\mathbf{x}-\mathbf{x}_0)\,d\mathbf{x}\right)^2 = \int_{\Omega\times\Omega} g_\beta(i_1 - I_1^\sigma(\mathbf{x}))I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))G_\gamma(\mathbf{x}-\mathbf{x}_0) \\ g_\beta(i_1 - I_1^\sigma(\mathbf{x}'))I_2^\sigma(\mathbf{x}'+\mathbf{h}(\mathbf{x}'))G_\gamma(\mathbf{x}'-\mathbf{x}_0)\,d\mathbf{x}\,d\mathbf{x}',$$

commute the integration with respect to i_1 with that with respect to $\mathbf{x}$ and $\mathbf{x}'$ to obtain the result. □

We continue with the analog of lemma 11:

Lemma 16 *The function $h^l_{CR} = G_\beta \star L^l_{CR} : \mathcal{H}_2 \times \Omega \longrightarrow \mathbb{R}$, where L^l_{CR} is given by equation (18), is equal to the following expression*

$$h^l_{CR}(z_1,z_2,\mathbf{x}_0,\mathbf{h}) = -\frac{2}{\mathcal{G}_\gamma(\mathbf{x}_0)\,v_2(\mathbf{x}_0,\mathbf{h})} \Big(z_2\mathbf{CR}^l[\mathbf{h}](\mathbf{x}_0) - d(z_1,\mathbf{x}_0,\mathbf{h}) + \mu_2(\mathbf{x}_0,\mathbf{h})(1-\mathbf{CR}^l[\mathbf{h}](\mathbf{x}_0))\Big),$$

where

$$d(z_1,\mathbf{x}_0,\mathbf{h}) = \int_{\mathbb{R}} g_\beta(z_1 - i_1)\left(\frac{\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))G_\gamma(\mathbf{x}-\mathbf{x}_0)\,d\mathbf{x}}{\int_\Omega g_\beta(i_1 - I_1^\sigma(\mathbf{x}))G_\gamma(\mathbf{x}-\mathbf{x}_0)\,d\mathbf{x}}\right) di_1.$$

Proof : We use equation (18) and apply the convolution to it. The value of d is obtained from:

$$d(z_1,\mathbf{x}_0,\mathbf{h}) = \int_{\mathbb{R}} g_\beta(z_1 - i_1)\mu_{2|1}(i_1,\mathbf{x}_0,\mathbf{h})\,di_1 = \int_{\mathbb{R}}\frac{g_\beta(z_1-i_1)}{p(i_1,\mathbf{x}_0)}\left(\int_{\mathbb{R}} i_2 P_{\mathbf{h}}(i_1,i_2,\mathbf{x}_0)\,di_2\right) di_1 = \\ \int_{\mathbb{R}}\frac{g_\beta(z_1-i_1)}{\mathcal{G}_\gamma(\mathbf{x}_0)p(i_1,\mathbf{x}_0)}\left(\int_\Omega I_2^\sigma(\mathbf{x}+\mathbf{h}(\mathbf{x}))g_\beta(i_1 - I_1^\sigma(\mathbf{x}))G_\gamma(\mathbf{x}-\mathbf{x}_0)\,d\mathbf{x}\right) di_1$$

from where the result follows. □

Our goal is to prove that $f^l_{\text{CR}} = G_\gamma \star h^l_{\text{CR}}$, a function from $\mathcal{H}_2 \times \Omega$ in $\mathbb{R}$ is Lipschitz continuous in $\mathcal{H}_2$ uniformly on Ω.

In order to prove this, it is sufficient to prove that the numerator and the denominator of h^l_{CR} are bounded and Lipschitz continuous in $\mathcal{H}_2$ uniformly on Ω, and that the denominator is strictly positive.

Indeed, we have the following

Lemma 17 *Let $N : \mathcal{H}_2 \times \Omega \to \mathbb{R}$ be bounded and Lipschitz continuous in $\mathcal{H}_2$ uniformly on Ω. Let also $D : \mathcal{H}_2 \times \Omega \to \mathbb{R}^+$ be bounded, strictly positive, and Lipschitz continuous in $\mathcal{H}_2$ uniformly on Ω. Then the function $G_\gamma \star \frac{N}{D} : \mathcal{H}_2 \times \Omega \to \mathbb{R}$ is Lipschitz continuous in $\mathcal{H}_2$ uniformly on Ω.*

Proof : We form

$$\left| \int_\Omega G_\gamma(\mathbf{x} - \mathbf{x}_0) \left(\frac{N(\mathbf{x}_0)[\mathbf{z}', \mathbf{h}']}{D(\mathbf{x}_0)[\mathbf{z}', \mathbf{h}']} - \frac{N(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]}{D(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]} \right) d\mathbf{x}_0 \right| \leq \\ G_\gamma(\mathbf{0}) \int_\Omega \frac{|N(\mathbf{x}_0)[\mathbf{z}', \mathbf{h}'] - N(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]|}{D(\mathbf{x}_0)[\mathbf{z}', \mathbf{h}']} d\mathbf{x}_0 + \\ G_\gamma(\mathbf{0}) \int_\Omega \frac{|N(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]| \, |D(\mathbf{x}_0)[\mathbf{z}', \mathbf{h}'] - D(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]|}{D(\mathbf{x}_0)[\mathbf{z}', \mathbf{h}'] D(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]} d\mathbf{x}_0$$

According to the hypotheses, there exists $a > 0$ such that $a \leq D(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}] \forall \mathbf{z}, \mathbf{x}_0, \mathbf{h}$, there exists $K_N \geq 0$ such that $|N(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]| \leq K_N \forall \mathbf{z}, \mathbf{x}_0, \mathbf{h}$, and there exists L_N and L_D such that

$$|N(\mathbf{x}_0)[\mathbf{z}', \mathbf{h}'] - N(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]| \leq L_N(|\mathbf{z} - \mathbf{z}'| + \|\mathbf{h} - \mathbf{h}'\|_H) \quad \text{and} \\ |D(\mathbf{x}_0)[\mathbf{z}', \mathbf{h}'] - D(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]| \leq L_D(|\mathbf{z} - \mathbf{z}'| + \|\mathbf{h} - \mathbf{h}'\|_H) \; \forall \mathbf{z}, \mathbf{z}', \mathbf{x}_0, \mathbf{h}, \mathbf{h}'.$$

We therefore have through Cauchy-Schwarz inequality:

$$\left| \int_\Omega G_\gamma(\mathbf{x} - \mathbf{x}_0) \left(\frac{N(\mathbf{x}_0)[\mathbf{z}', \mathbf{h}']}{D(\mathbf{x}_0)[\mathbf{z}', \mathbf{h}']} - \frac{N(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]}{D(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]} \right) d\mathbf{x}_0 \right| \leq \\ C(|\mathbf{z} - \mathbf{z}'| + \|\mathbf{h} - \mathbf{h}'\|_H)$$

for some positive constant C. □

We prove these properties for the numerator and the denominator of h^l_{CR}.

Lemma 18 *We have*

$$|\Omega| \textit{diam}(\Omega) \beta \leq \mathcal{G}_\gamma(\mathbf{x}_0) \, v_2(\mathbf{x}_0, \mathbf{h}) \leq |\Omega| G_\gamma(\mathbf{0})(\beta + \mathcal{A}^2)$$

Proof : The proof is a direct consequence of the definition of $\mathcal{G}(\mathbf{x}_0)$ and of lemma 14. □

We then prove the following

Lemma 19 *The function $\Omega \times H \to \mathbb{R}^+$ such that $(\mathbf{x}_0, \mathbf{h}) \to \mathcal{G}_\gamma(\mathbf{x}_0) \, v_2(\mathbf{x}_0, \mathbf{h})$ is Lipschitz continuous in H uniformly in Ω.*

Lemma 20 *The function $\mathcal{H}_2 \times \Omega \rightarrow \mathbb{R}$ such that*

$$(\mathbf{z}, \mathbf{x}_0, \mathbf{h}) \rightarrow z_2 \mathbf{CR}^l[\mathbf{h}](\mathbf{x}_0) - d(z_1, \mathbf{x}_0, \mathbf{h}) + \mu_2(\mathbf{x}_0, \mathbf{h})(1 - \mathbf{CR}^l[\mathbf{h}](\mathbf{x}_0)),$$

is bounded by $3\mathcal{A}$.

Proof : This is because $0 \leq \mathbf{CR}[\mathbf{h}](\mathbf{x}_0) \leq 1$, $0 \leq \mu_2(\mathbf{x}_0, \mathbf{h}) \leq \mathcal{A}$, and, according to lemma 16, because $0 \leq d(z_1, \mathbf{x}_0, \mathbf{h}) \leq \mathcal{A}$. □

At this point we can prove the following

Proposition 30 *The function $\Omega \rightarrow \mathbb{R}^n$ such that $\mathbf{x} \rightarrow f^l_{CR}(\mathbf{z}, \mathbf{x}, \mathbf{h})$ is Lipschitz continuous uniformly in $\mathcal{H}_2$.*

Proof : With the notations of lemma 17 we have

$$|G_\gamma \star h^l_{CR}(\mathbf{x}) - G_\gamma \star h^l_{CR}(\mathbf{y})| \leq \int_\Omega |G_\gamma(\mathbf{x} - \mathbf{x}_0) - G_\gamma(\mathbf{y} - \mathbf{x}_0)| \frac{|N(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]|}{D(\mathbf{x}_0)[\mathbf{z}, \mathbf{h}]} \, d\mathbf{x}_0$$

Because of lemmas 18 and 20 we have

$$|G_\gamma \star h^l_{CR}(\mathbf{x}) - G_\gamma \star h^l_{CR}(\mathbf{y})| \leq \frac{3\mathcal{A}}{|\Omega| \mathrm{diam}(\Omega)\beta} \int_\Omega |G_\gamma(\mathbf{x} - \mathbf{x}_0) - G_\gamma(\mathbf{y} - \mathbf{x}_0)| \, d\mathbf{x}_0 \leq \frac{3\mathcal{A}Lip(G_\gamma)}{\mathrm{diam}(\Omega)\beta} |\mathbf{x} - \mathbf{y}|,$$

hence the result. □

We also prove the analog of lemmas 12, 9, and theorem 8.

Lemma 21 *The function $d : \mathcal{H}_1 \times \Omega \longrightarrow \mathbb{R}$ is bounded and Lipschitz continuous in $\mathcal{H}_1$ uniformly in Ω.*

Lemma 22 *The function $\Omega \times H \longrightarrow \mathbb{R}$ defined by*

$$(\mathbf{x}_0, \mathbf{h}) \longrightarrow \frac{1}{\mathcal{G}_\gamma(\mathbf{x}_0)} \int_\Omega I^\sigma_2(\mathbf{x} + \mathbf{h}(\mathbf{x})) G_\gamma(\mathbf{x} - \mathbf{x}_0) \, d\mathbf{x}$$

is bounded and Lipschitz continuous in H uniformly in Ω.

Theorem 12 *The function $\Omega \times H \longrightarrow \mathbb{R}$ defined by*

$$(\mathbf{x}_0, \mathbf{h}) \longrightarrow \frac{\mathbf{E}[\mathbf{Var}[X_{I^\sigma_2, \mathbf{h}} | X^l_{I^\sigma_1}](\mathbf{x}_0)]}{v_2(\mathbf{x}_0, \mathbf{h})}$$

is Lipschitz continuous in H uniformly in Ω.

From which we deduce the

Theorem 13 *The function $f^l_{CR} : \mathcal{H}_2 \times \Omega \rightarrow \mathbb{R}$ such that*

$$(\mathbf{z}, \mathbf{x}_0, \mathbf{h}) \rightarrow z_2 \mathbf{CR}^l[\mathbf{h}](\mathbf{x}_0) - d(z_1, \mathbf{x}_0, \mathbf{h}) + \mu_2(\mathbf{x}_0, \mathbf{h})(1 - \mathbf{CR}^l[\mathbf{h}](\mathbf{x}_0)),$$

is Lipschitz continuous in $\mathcal{H}_2$ uniformly in Ω.

We can now prove the

Theorem 14 *The function $f^l_{CR} : \mathcal{H}_2 \times \Omega \rightarrow \mathbb{R}$ such that $(\mathbf{z}, \mathbf{x}, \mathbf{h}) \rightarrow f^l_{CR}(\mathbf{z}, \mathbf{x}, \mathbf{h})$ is Lipschitz continuous in $\mathcal{H}_2$ uniformly in Ω.*

Proof : The proof is just an application of lemma 17 to f^l_{CR}. □

The combination of proposition 30 and theorem 14 yields the following

Theorem 15 *The function $f^l_{CR} : \mathcal{H}_2 \times \Omega \rightarrow \mathbb{R}$ such that $(\mathbf{z}, \mathbf{x}, \mathbf{h}) \rightarrow f^l_{CR}(\mathbf{z}, \mathbf{x}, \mathbf{h})$ is Lipschitz continuous.*

And we can conclude with the following theorem and proposition.

Theorem 16 *The function $F^l_{CR} : H \longrightarrow H$ defined by*

$$F^l_{CR}(\mathbf{h}) = f^l_{CR}(I^\sigma_1, I^\sigma_2(\mathbf{Id} + \mathbf{h}), \mathbf{Id}) \nabla I^\sigma_2(\mathbf{Id} + \mathbf{h})$$

is Lipschitz continuous and bounded.

Proof : The proof follows exactly the same pattern as the proof of theorem 11 and uses theorem 14. □

Proposition 31 *The function $\Omega \rightarrow \mathbb{R}^n$ such that $\mathbf{x} \rightarrow F^l_{CR}(\mathbf{h}(\mathbf{x}))$ satisfies*

$$|F^l_{CR}(\mathbf{h}(\mathbf{x})) - F^l_{CR}(\mathbf{h}(\mathbf{y}))| \leq K(|\mathbf{x} - \mathbf{y}| + |\mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{y})|),$$

for some constant $K > 0$.

Proof : The proof is similar to that of proposition 22 and follows from theorem 15 and the fact that the functions I^σ_1, I^σ_2 and all its derivatives, are Lipschitz continuous. □

8.4 Summary

At this stage it is wise to stop for a minute and recall what we have achieved in this long section. Starting from theorems 3 and 4 that give explicit expressions for the infinitesimal gradients of the four dissimilarity measures defined in definitions 1 and 2, we have proved that the four corresponding functions $H \rightarrow H$ (or $H \rightarrow K$) that they define and which appear on the righthand side of equation (3), are Lipschitz continuous.

This has been done in theorems 7 and 11 for the mutual information in the global case (F^g_{MI}) and in the local case (F^l_{MI}), respectively.

This has been done in theorems 10 and 16 for the correlation ratio in the global case (F^g_{CR}) and in the local case (F^l_{CR}), respectively.

9 Existence and uniqueness of a solution to the initial value problem (3)

We now consider the initial value problem (3). According to the previous analysis, (3) can be rewritten

$$\begin{cases} \dfrac{d\mathbf{h}}{dt} - A\mathbf{h}(t) = -F(\mathbf{h}(t)) \\ \mathbf{h}(0)(\cdot) = \mathbf{h}_0(\cdot), \end{cases} \tag{31}$$

where the operator A is either κA_1 defined in (14) or κA_2 defined in (15), F is one the functions defined in theorems 3 and 4, and κ is the positive weighting parameter introduced in (2). We intend to show that (31) has a unique solution in a sense to be made more precise later.

The proofs use three ingredients

1. The fact that the operator $-A$ is maximal monotone and therefore that A generates a C_0 semigroup (proved in propositions 8 and 10).
2. The fact that the operator A is the infinitesimal generator of an analytic semigroup of operators on H (proved in theorem 6).
3. The fact that the function $F : H \to H$ is bounded and Lipschitz continuous (proved in theorems 7, 10, 11, and 16).

9.1 Weak and strong solution

We first make use of the fact, proved in propositions 8 and 10, that the operator $-A$ is maximal monotone implies, through the Hille-Yosida theorem, that A is the infinitesimal generator of a C_0 semigroup of bounded linear operators in $H = \mathbf{L}^2(\Omega)$

We begin with a definition

Definition 6 *A function* $\mathbf{h} : [0, T[\to H$ *is a* classical *solution of (31) if* $\mathbf{h} \in C([0, T[; H) \cap C^1(]0, T[; H)$ *and* $\mathbf{h}(t) \in \mathcal{D}(A)$ *for* $0 < t < T$*, and (31) is satisfied on* $[0, T[$.

The initial problem (31) does not necessarily have a solution of any kind. However, if it has a classical solution then the H valued function $\mathbf{k}(s) = S_A(t-s)\mathbf{h}(s)$ is differentiable for $0 < s < t$ and (proposition 11):

$$\frac{d\mathbf{k}}{ds} = -AS_A(t-s)\mathbf{h}(s) + S_A(t-s)\mathbf{h}'(s) = \\ -AS_A(t-s)\mathbf{k}(s) + S_A(t-s)A\mathbf{h}(s) - S_A(t-s)F(\mathbf{h}(s)) = -S_A(t-s)F(\mathbf{h}(s)) \tag{32}$$

If $F \circ \mathbf{h} \in L^1([0, T[; H)$ then $S_A(t-s)F(\mathbf{h}(s))$ is integrable and integrating (32) from 0 to t yields

$$\mathbf{k}(t) - \mathbf{k}(0) = \mathbf{h}(t) - S_A(t)\mathbf{h}_0 = -\int_0^t S_A(t-s)F(\mathbf{h}(s))\, ds$$

hence

$$\mathbf{h}(t) = S_A(t)\mathbf{h}_0 - \int_0^t S_A(t-s)F(\mathbf{h}(s))\,ds \tag{33}$$

The following definition is then natural

Definition 7 *A continuous solution* $\mathbf{h}$ *of the integral equation (33) is called a* mild solution *of the initial value problem (31).*

Given this definition, we have the following,

Theorem 17 *The initial value problem (31) has a unique mild solution* $\mathbf{h} \in C([0,T];H)$ *(given by (33)) for all* $\mathbf{h}_0 \in H$*. Moreover, the mapping* $\mathbf{h}_0 \rightarrow \mathbf{h}$ *is Lipschitz continuous from* H *into* $C([0,T];H)$.

Proof : The proof follows from theorem 6.1.2 in [10]: A is the infinitesimal generator of a C_0 semigroup $S_A(t),\ t \geq 0$ and the function $F: H \rightarrow H$ is Lipschitz continuous (theorems 7, 10, 11, and 16). □

To prove that $\mathbf{h}$ is a classical solution would require that $F: H \rightarrow H$ be continuously differentiable (in the Frechet sense) according to theorem 6.1.5 of [10]. Since we have only proved that F is Lipschitz continuous we can only prove the existence of a somewhat more regular solution of the initial problem (31). We first define a *strong* solution of (31):

Definition 8 *A function* $\mathbf{h}$ *which is differentiable almost everywhere on* $[0,T]$ *such that* $d\mathbf{h}/dt \in L^1(]0,T[;H)$ *is called a* strong solution *of the initial value problem (31) if* $\mathbf{h}(0) = \mathbf{h}_0$ *and* $d\mathbf{h}/dt - A\mathbf{h}(t) = -F(\mathbf{h}(t))$ *almost everywhere on* $[0,T]$.

We can in effect prove that if we take the right initial condition, we obtain a strong solution,

Theorem 18 *The initial value problem (31) has a unique strong solution for all* $\mathbf{h}_0 \in \mathcal{D}(A)$.

Proof : This is a direct consequence of theorem 6.1.6 in [10]: H being a Hilbert space is a reflexive Banach space, $F: H \rightarrow H$ is Lipschitz continuous (theorems 7, 10, 11, and 16), hence if $\mathbf{h}_0 \in \mathcal{D}(A)$ then, the mild solution of (31) is the strong solution. □

Note that despite the results of theorems 17 and 18 we have not very much regularity for our solutions.

9.2 Classical and regular solution

It follows from theorem 6 and from [10], chapter 2, section 2.6, that $(-A)^\alpha$ can be defined for $0 < \alpha \leq 1$ and that $(-A)^\alpha$ is a closed linear invertible operator with domain $\mathcal{D}((-A)^\alpha)$ dense in H. It is invertible because $-A$ is (lemma 3). The closedness of $(-A)^\alpha$ implies that $\mathcal{D}((-A)^\alpha)$ endowed with the graph norm of $(-A)^\alpha$, i.e. the norm $|||\mathbf{h}||| = \|\mathbf{h}\|_H + \|(-A)^\alpha\mathbf{h}\|_H$ is a Banach space. Since $(-A)^\alpha$ is invertible its graph norm is equivalent to the norm $\|\mathbf{h}\|_\alpha = \|(-A)^\alpha\mathbf{h}\|_H$. Thus, $\mathcal{D}((-A)^\alpha)$ equipped with the norm $\|\cdot\|_\alpha$ is a Banach space which we denote by H_α.

We will need the following result about the imbedding of H_α in various functional spaces.

Proposition 32 *Let Ω be a bounded open set of $\mathbb{R}^n$ with smooth boundary $\partial\Omega$. If $0 \le \alpha \le 1$ then*

$$H_\alpha \subset \mathbf{W}^{k,q}(\Omega) \quad \textit{for} \quad k - \frac{n}{q} < 2\alpha - \frac{n}{2},$$

$$H_\alpha \subset \mathbf{C}^\nu(\overline{\Omega}) \quad \textit{for} \quad 0 \le \nu < 2\alpha - \frac{n}{2},$$

and the imbeddings are continuous.

Proof : This is a special case of theorem 8.4.3 in [10]. In particular if $k = 1$, $q = \infty$ and $n = 3$ (resp. 2) we see that as soon as $\alpha > 3/4$ (resp. $\alpha > 1/2$) we have $H_\alpha \subset \mathbf{W}^{0,\infty}(\Omega) = \mathbf{L}^\infty(\Omega)$ and $H^\alpha \subset \mathbf{C}^\nu(\overline{\Omega})$, i.e. is a set of Hölder continuous functions. □

The main result that we will use is the following.

Theorem 19 *If the function F satisfies the following two conditions*

$$\|F(\mathbf{h}_1) - F(\mathbf{h}_2)\|_H \le L\|\mathbf{h}_1 - \mathbf{h}_2\|_\alpha \, \forall \mathbf{h}_1, \, \mathbf{h}_2 \in H_\alpha \tag{34}$$

and

$$\|F(\mathbf{h})\|_H \le K \, \forall \mathbf{h} \in H_\alpha, \tag{35}$$

for $0 \le \alpha_0 < \alpha < 1$ then for every $\mathbf{h}_0 \in H_\alpha$ the initial value problem (31) has a unique classical solution $\mathbf{h} \in C([0, +\infty[; H) \cap C^1(]0, +\infty[; H)$. Moreover, the function $t \to d\mathbf{h}/dt$ from $]0, +\infty[$ into H_α is Hölder continuous.

Proof : The proof is a special case of theorems 6.3.1 and 6.3.3 in [10]. Indeed, A is the infinitesimal generator of an analytic semigroup $S_A(t)$ satisfying $\|S_A(t)\| \le 1$ (lemma 2) and $0 \in \rho(A)$ since $-A$ is invertible (lemma 3). If F satisfies the Lipschitz condition (34), according to theorem 6.3.1 of [10], the initial value problem (31) has a unique local solution $\mathbf{h} \in C([0, t[, H) \cap C^1(]0, t[, H) \cap C(]0, t[, \mathcal{D}(A))$ where $t = t(\mathbf{h}_0)$.

Next, because of theorems 7, 10, 11, and 16, the function F is bounded in H and therefore in H_α: $|F(\mathbf{h}(\mathbf{x}))| \le M_F \, \forall \mathbf{h} \in H$ and $\forall \mathbf{x} \in \Omega$ implies $\|F(\mathbf{h})\|_H \le M_F \, |\Omega|^{1/2} = K$. In particular, F is well-defined on H_α. We can therefore adapt theorem 6.3.3 of [10] to this case and try to continue the solution of (31) by showing that $\|\mathbf{h}(t)\|_\alpha$ stays bounded. It is sufficient to show that if $\mathbf{h}$ exists on $[0, T[$ then $\|\mathbf{h}(t)\|_\alpha$ is bounded as $t \to T$. We have, operating on both sides of (33) with $(-A)^\alpha$:

$$(-A)^\alpha \mathbf{h}(t) = (-A)^\alpha S_A(t)\mathbf{h}_0 - \int_0^t (-A)^\alpha S_A(t-s) F(\mathbf{h}(s)) \, ds,$$

We next use the facts (proved in theorem 2.6.13 of [10]) that $S_A(t) : H \to \mathcal{D}((-A)^\alpha)$ for every $t > 0$ and $\alpha \ge 0$, that for every $\mathbf{h} \in \mathcal{D}((-A)^\alpha))$ we have $S_A(t)(-A)^\alpha \mathbf{h} = (-A)^\alpha S_A(t)\mathbf{h}$, that $\|(-A)^\alpha S_A(t)\|_{\mathcal{L}(H)} \le C_\alpha t^{-\alpha}$, that $S_A(t)$ is bounded by one (lemma 2) and that $\|F(\mathbf{h})\|_H \le K$ for all $\mathbf{h} \in H_\alpha$ to obtain

$$\|(-A)^\alpha \mathbf{h}(t)\|_H = \|\mathbf{h}(t)\|_\alpha \le \|(-A)^\alpha \mathbf{h}_0\|_H + C_\alpha K \int_0^t (t-s)^{-\alpha} \, ds =$$

$$\|(-A)^\alpha \mathbf{h}_0\|_H + C_\alpha K \frac{t^{1-\alpha}}{1-\alpha},$$

and $\|\mathbf{h}(t)\|_\alpha$ stays bounded when $t \to T$.

The Hölder continuity follows from corollary 6.3.2 in [10] which also shows that the Hölder exponent β verifies $0 < \beta < 1 - \alpha$. $\square$

We therefore prove that our function F satisfies (34):

Proposition 33 *The function F is such that*

$$\|F(\mathbf{h}_1) - F(\mathbf{h}_2)\|_H \leq L_F \|\mathbf{h}_1 - \mathbf{h}_2\|_\alpha$$

Proof : We have seen in the proof of theorem 19 that F is well defined on H_α. From theorems 7, 10, 11, and 16, we deduce that

$$\|F(\mathbf{h}_1) - F(\mathbf{h}_2)\|_H \leq L_F \|\mathbf{h}_1 - \mathbf{h}_2\|_H,$$

and the conclusion follows from the fact that $H_\alpha \subset \mathbf{L}^\infty(\Omega) \subset \mathbf{L}^2(\Omega)$ with continuous imbedding (proposition 32 and lemma 1) for $\alpha > 3/4$ ($n = 3$) and $\alpha > 1/2$ ($n = 2$). $\square$

It follows from theorem 19 that the solution (mild, strong, classical) $\mathbf{h}$ of the initial problem (31) satisfies

$$\mathbf{h} \in C([0, +\infty[, \mathbf{L}^2(\Omega)) \cap C(]0, +\infty[, \mathbf{H}^2(\Omega) \cap \mathbf{H}^1_0(\Omega)) \cap C^1(]0, +\infty[, \mathbf{L}^2(\Omega))$$

We next show that it is in fact a regular solution of this initial value problem for $t > 0$. We have the proposition

Proposition 34 *The functions $(t, \mathbf{x}) \to \mathbf{h}(t, \mathbf{x})$ and $(t, \mathbf{x}) \to (\partial/\partial t)\mathbf{h}(t, \mathbf{x})$ are continuous on $]0, +\infty[\times\overline{\Omega}$ and for each $t > 0$ the function $\mathbf{x} \to \mathbf{h}(t, \mathbf{x})$ is in $\mathbf{C}^2(\Omega)$.*

Proof : It follows from a theorem due to Sobolev that $\mathbf{H}^2(\Omega) \subset \mathbf{C}(\overline{\Omega})$ for $n = 2, 3$, hence $\mathcal{D}(A) \subset \mathbf{H}^2(\Omega) \subset \mathbf{C}(\overline{\Omega})$ and since $\mathbf{h}(t) \in \mathcal{D}(A)$ for $t > 0$ this proves the continuity of $\mathbf{h}$ on $]0, +\infty[\times\overline{\Omega}$. Next, because of theorem 19, $t \to d\mathbf{h}/dt \in H_\alpha$ is Hölder continuous for $t > 0$. Moreover, if $\alpha > 3/4$ ($n = 3$) and $\alpha > 1/2$ ($n = 2$) H_α is a set of Hölder continuous functions (proposition 32) and we have the continuity of $d\mathbf{h}/dt$ on $]0, +\infty[\times\overline{\Omega}$.

It remains to show that $\mathbf{h}(t, \cdot) \in \mathbf{C}^2(\Omega)$. First the function $\mathbf{x} \to F(\mathbf{h}(\mathbf{x}))$ from $\Omega \to \mathbb{R}^n$ is Hölder continuous if $\mathbf{h}$ is (propositions 21, 22, 29 and 31). Since $\mathbf{h}(t) \in \mathbf{H}^2(\Omega)$ for $t > 0$ and (Sobolev) $\mathbf{H}^2(\Omega) \subset \mathbf{C}^\gamma(\overline{\Omega})$ ($0 \leq \gamma < 0.5$ for $n = 3$ and $0 \leq \gamma < 1$ for $n = 2$) $\mathbf{x} \to \mathbf{h}(t, \mathbf{x})$ is Hölder continuous for $t > 0$. Finally, since $(\partial/\partial t)\mathbf{h}(t, .)$ is Hölder continuous in Ω it follows that $A\mathbf{h} = F(\mathbf{h}) - d\mathbf{h}/dt$ is Hölder continuous in Ω and by a classical regularity theorem for elliptic equations it follows that $\mathbf{h}(t, \cdot) \in \mathbf{C}^{2+\delta}(\Omega)$ for some $\delta > 0$, i.e. has second order Hölder continuous derivatives in the space variable and is thus a regular solution. $\square$

10 Comments

Let us comment on what we have achieved. Our main result is that, given an initial field $\mathbf{h}_0$ in the domain $\mathcal{D}(A)$ of the operator A, the initial value problem (31) has a unique *classical* solution $\mathbf{h}(t)$

in $\mathcal{D}(A)$. This solution and its time derivative are continuous on $]0,+\infty[\times\overline{\Omega}$ and for each $t > 0$, $\mathbf{x} \to \mathbf{h}(t,\mathbf{x})$ is in $\mathbf{C}^2(\Omega)$. $d\mathbf{h}/dt$ is by construction equal to $-\nabla\mathcal{I}_{\mathbf{h}(t)}[\mathbf{h}(t)]$ (equation (3)). How does this relate to the original minimization problem (2)?

Consider the function $r : [0,+\infty[\to \mathbb{R}^+$ defined by $r(t) = \mathcal{I}[\mathbf{h}(t)]$. Because of the smoothness of $\mathbf{h}$ the function r is differentiable in $]0,+\infty[$ and at each point where it is differentiable its derivative is equal to $-\|\nabla_{\mathbf{h}(t)}[\mathbf{h}(t)]\|_H^2$, hence ≤ 0. Therefore, by solving (31) and by following the differential "curve" $t \to \mathbf{h}(t)$ in H we guarantee that we decrease the original value $\mathcal{I}[\mathbf{h}_0]$ of our criterion as long as $\|\nabla_{\mathbf{h}(t)}[\mathbf{h}(t)]\|_H \neq 0$ or that we find an extremum of our criterion if $\|\nabla_{\mathbf{h}(t_0)}[\mathbf{h}(0)]\|_H = 0$ for some $t_0 > 0$.

11 Implementation Issues

Several numerical schemes are possible for the discretization of the gradient flows. The computation of the warped image requires interpolation of the image values. Our implementation combines the so called tri-linear interpolation (it is not actually linear) with centered differences schemes for the spatial derivatives. This choice has given good results in practice. As usual for this kind of problem, a multi-resolution approach is used. The gradient descent is applied to a set of smoothed and subsampled images. The corresponding functionals are non convex and this coarse to fine strategy helps avoiding irrelevant extrema. A larger class of deformations can be recovered and the computational cost of the algorithms is decreased. In practice, we use for this purpose a convolution by a gaussian kernel followed by subsampling. We discuss the density estimation step in some length in the next section.

11.1 Density estimation

For the global algorithms, Parzen-Rozenblatt estimates are obtained by convolution of the joint histogram with a 2D gaussian kernel, implemented as a recursive filter [6]. For the local case, this approximation is done by explicit discretization of the integral in equation (9). The parameter β of the gaussian kernel is determined automatically. A very large amount of literature has been published on the problem of determining an adequate value for β (we refer to [2] for a recent comprehensive study on nonparametric density estimation, containing many references to the forementioned literature). We adopt a cross-validation method technique based on an empirical maximum likelihood method. We note $\{\mathbf{i}_k\}$ a set of m intensity pair samples ($k = 1 \ldots m$) and take the value of β which maximizes the empirical likelihood:

$$L(\beta) = \prod_{k=1}^{m} \hat{P}_{\beta,k}(\mathbf{i}_k)$$

where

$$\hat{P}_{\beta,k}(\mathbf{i}_k) = \frac{1}{m - n_k} \sum_{\{s:\ \mathbf{i}_s \neq \mathbf{i}_k\}} G_\beta(\mathbf{i}_k - \mathbf{i}_s)$$

and n_k is the number of data samples for which $\mathbf{i}_s = \mathbf{i}_k$.

11.2 Experimental results

In this section we present some experimental results obtained with the four matching algorithms. The global algorithms are relatively fast, specially in the 2D case. Typical running times are 5 to 10 minutes for images of size 256x256. The local algorithms are more computationally expensive, since the Parzen-Rozenblatt estimations must be carried out in a large neighbourhood around each pixel. For running the 3D experiments, they have been parallelized and run on a cluster of 20 processors (1 GHz clock speed each). Their running time in these conditions was approximately 30 minutes for a volume of size 80^3. Despite this disadvantage, the results obtained can replace several hours of tedious, manual alignment.

Experiment and Figure 1

This experiment shows the result of the local algorithms on synthetic data. This allows a quantitative an objective validation of the results. The reference and target image where both taken from the same 2D plane in a MRI data volume. The reference image J was then transformed in the following way ($|\Omega|_x$ is the size of the domain in the x direction):

$$J'(x,y) = \sin\left(2\,\pi\; J(x,y)\right) - \cos\left(\frac{2\,\pi}{|\Omega|}\;\left(x\,+\,y\,|\Omega|_x\,\right)\right)$$

and then linearly renormalised in $[0, \mathcal{A}]$. Notice that the effect of this manipulation produces a bias in the intensities of the reference image which resembles a commonly observed bias in MRI data. A non-rigid smooth deformation was then applied to the target image. As expected, the global algorithms failed to align these two images, due to the severe non-stationarity in the intensity distributions. The local algorithms (both **MI** and **CR**) succeeded with a good (sub-pixel) precision (0.53 and 0.59 pixels of mean error respectively).

Experiment and Figure 2

Our next experiment shows the result of the local algorithms with real 3D MR data. The reference image is a T1 weighted anatomical MRI of a human brain. The target image is an MRI from the same patient which is acquired using a special magnetic field gradient as part of the process of obtaining an image of the water diffusion tensor at each point. Notice that the intensities in this modality are qualitatively close to our simulated experiment. The estimated deformation field has a dominant y component, a property which is physically coherent with the applied gradient. Both **MI** and **CR** yielded similar results in this case.

Experiment and Figure 3

This example shows an experiment with real MR data of the brain of a macaque monkey. The reference image is a T1-weighted anatomical volume and the target image is a functional, mion contrast MRI (fMRI). The contrast in this modality is related to blood oxygenation level. The figure

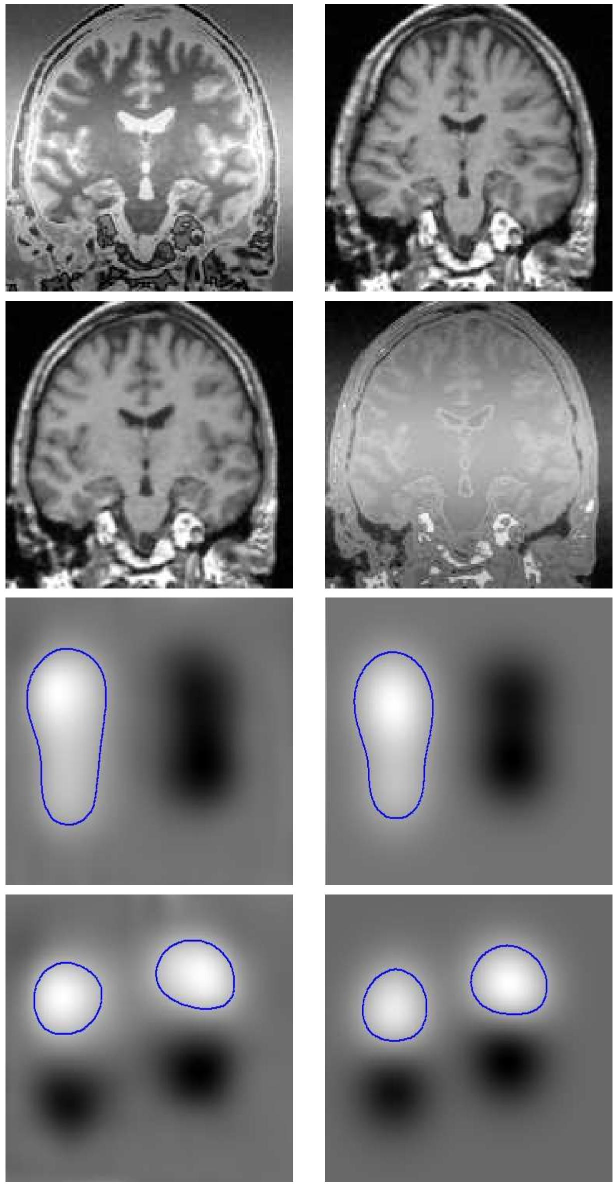

Figure 1: Local algorithm (**MI**) with synthetic data. First row: reference and target images. Second row: corrected target image (left) and its superposition with the reference image (right). Third and fourth rows: horizontal and vertical components of the estimated (left) and true (right) deformations fields (isolevel 3.4 is outlined).

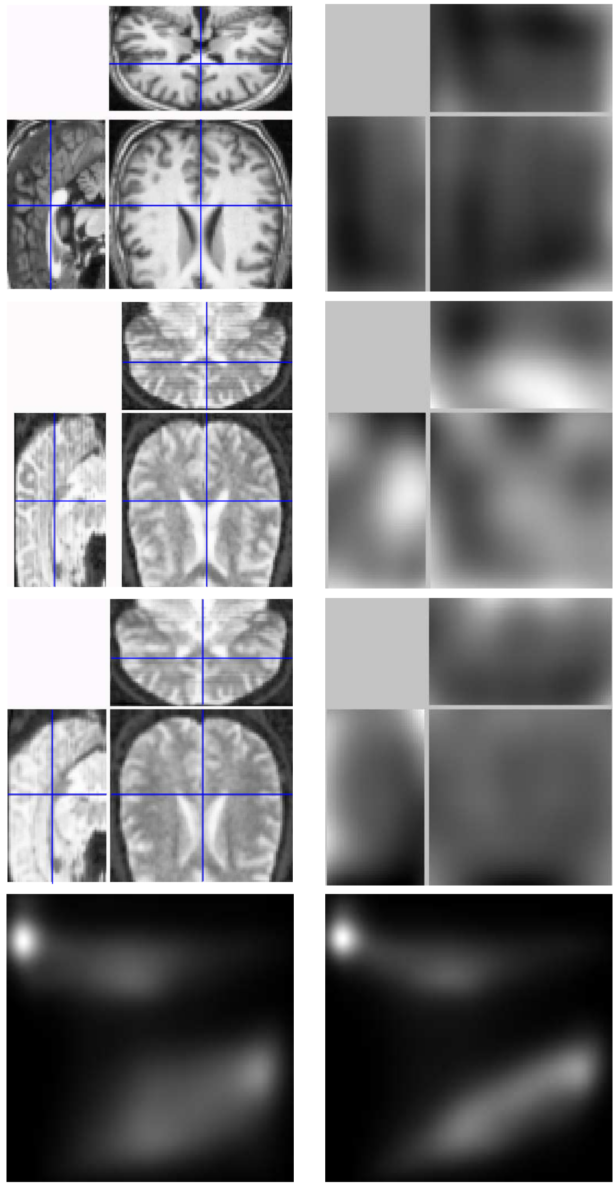

Figure 2: Local algorithm with real data. First three rows from top to bottom : left column: reference, target and corrected image; right column: x, y and z components of the found deformation field. Bottom row: estimated joint density for the intensity pairs before (left) and after (right) correction.

shows the result of the global algorithms. Notice that the alignment of main axis of the volume has been corrected.

Experiment and Figure 4

This last experiment shows the use of the global algorithms to perform template matching of human faces. In this case the illuminating conditions are the same in both photographs. If different, the local algorithms should be used. The different albedos of the two skins create a “multimodal” situation and the transformation is truly non rigid due to the different shapes of the noses and mouths. Notice the excellent matching of the different features. This result was obtained completely automatically with the same sets of parameters as the rest of the experiments, using global mutual information. The running time was approximately five minutes on a PC at 755 MHz. With the correspondences, one can interpolate the displacement field and the texture to perform fully automatic morphing (see video).

12 Summary and conclusion

In this paper, we have proved existence and uniqueness of a *classical* solution of a parabolic system of equations derived from multimodal, dense registration variational problems. We treated eight cases corresponding to the following alternatives. We considered that the images are realizations of spatial random processes that are either stationary or nonstationary. In each case we measured the similarity between the two images either by their mutual information or by their correlation ratio. In each case we regularized the deformation field either by borrowing from the field of linearized elasticity or by using geometry driven anisotropic diffusion.

References

[1] L. Alvarez, J. Weickert, and J. Sánchez. Reliable Estimation of Dense Optical Flow Fields with Large Displacements. Technical report, Cuadernos del Instituto Universitario de Ciencias y Tecnologías Cibernéticas, 2000. A revised version has appeared at IJCV 39(1):41-56,2000.

[2] D. Bosq. *Nonparametric Statistics for Stochastic Processes*, volume 110 of *Lecture Notes in Statistics*. Springer–Verlag, 2nd edition, 1998.

[3] H. Brezis. *Analyse fonctionnelle. Théorie et applications*. Masson, 1983.

[4] C. Chefd’hotel, G. Hermosillo, and O. Faugeras. A variational approach to multi-modal image matching. In *1st IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pages 21–28, University of British Columbia, Vancouver, Canada, July 13, 2001, July 2001. IEEE Computer Society.

[5] P. Ciarlet. *Mathematical Elasticity*, volume 1. North Holland, 1988.

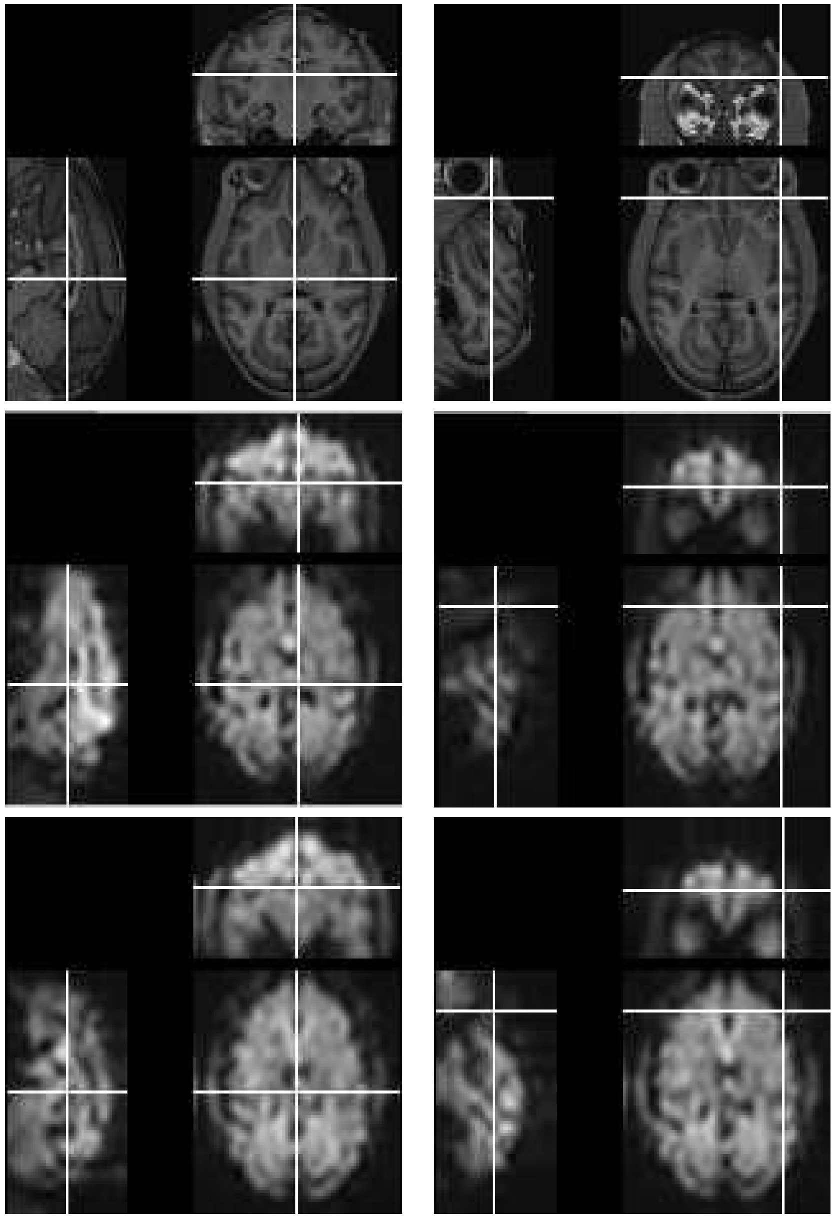

Figure 3: Global algorithm with fMRI data. Top row: reference anatomical MRI. Middle row: initial fMRI volume. Bottom row: final (corrected) fMRI volume. The two columns show two different points in the volumes.

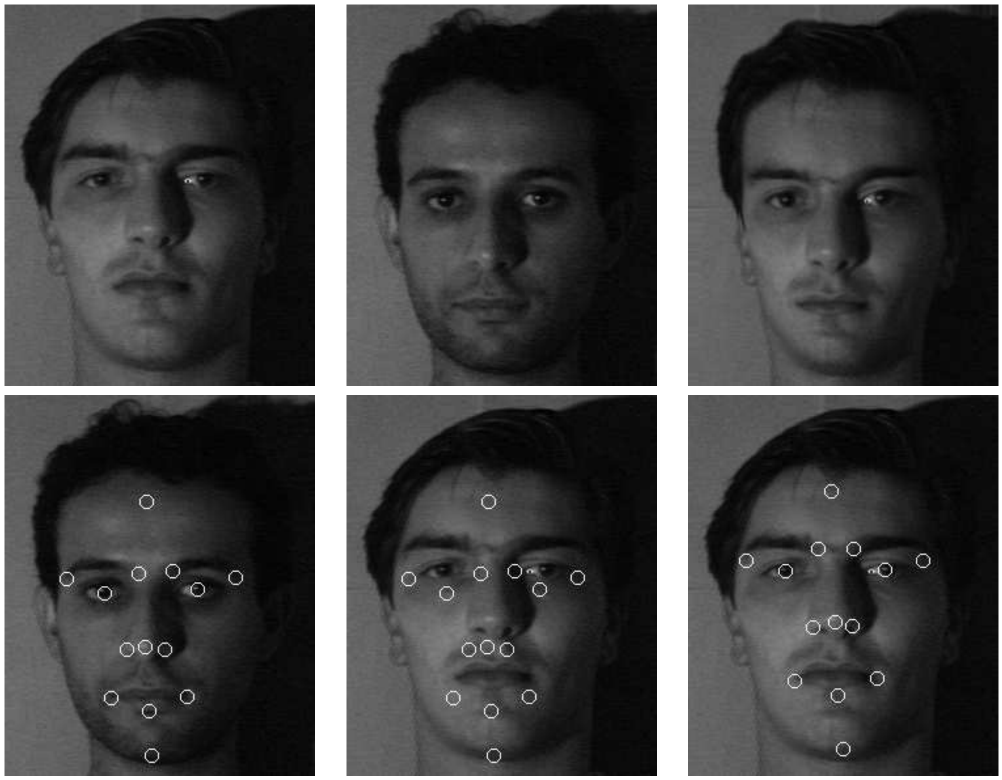

Figure 4: Global mutual information algorithm applied to human face template matching. In the top row from left to right: target, reference and warped image. The bottom row points out some interesting points in the reference image (left) and their respective correspondences before (center) and after warping (right).

[6] R. Deriche. Fast algorithms for low-level vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1(12):78–88, Jan. 1990.

[7] L. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. Proceedings of the American Mathematical Society, 1998.

[8] G. Hermosillo, C. Chefd'Hotel, and O. Faugeras. A variational approach to multi-modal image matching. RR 4117, INRIA, 01. Also appeared in the Proceedings of the 1st IEEE Workshop on Variational and Level Set Methods in Computer Vision, University of British Columbia, Vancouver, Canada, July 13, 2001, published by the IEEE Computer Society, pages 21-28.

[9] H. Nagel and W. Enkelmann. An investigation of smoothness constraint for the estimation of displacement vector fiels from images sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:565–593, 1986.

[10] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer–Verlag, 1983.

[11] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as new similarity metric for multimodal image registration. In W. Wells, A. Colchester, and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention-MICCAI'98*, number 1496 in Lecture Notes in Computer Science, pages 1115–1124, Cambridge, MA, USA, Oct. 1998. Springer.

[12] P. Viola and W. M. Wells III. Alignement by maximization of mutual information. *The International Journal of Computer Vision*, 24(2):137–154, 1997.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifi que
615, rue du Jardin Botanique - BP 101 - 54602 Villers-l`es-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399