Image Registration, Optical Flow and Local Rigidity

Martin Lefébure
was with Poseidon Technologies
mlefebure@compaqnet.fr

Laurent D. COHEN
CEREMADE, UMR 7534
Université Paris 9-Dauphine
75775 Paris cedex 16, France
cohen@ceremade.dauphine.fr

January 23, 2001

To appear in Journal of Mathematical Imaging and Vision, 2001.

Abstract

We address the theoretical problems of optical flow estimation and image registration in a multi-scale framework in any dimension. Much work has been done based on the minimization of a distance between a first image and a second image after applying deformation or motion field. Usually no justification is given about convergence of the algorithm used. We start by showing, in the translation case, that convergence to the global minimum is made easier by applying a low pass filter to the images hence making the energy "convex enough". In order to keep convergence to the global minimum in the general case, we introduce a local rigidity hypothesis on the unknown deformation. We then deduce a new natural motion constraint equation (MCE) at each scale using the Dirichlet low pass operator. This transforms the problem to solving the energy minimization in a finite dimensional subspace of approximation obtained through Fourier Decomposition. This allows us to derive sufficient conditions for convergence of a new multi-scale and iterative motion estimation/registration scheme towards a global minimum of the usual nonlinear energy instead of a local minimum as did all previous methods. Although some of the sufficient conditions cannot always be fulfilled because of the absence of the necessary a priori knowledge on the motion, we use an implicit approach. We illustrate our method by showing results on synthetic and real examples in dimension 1 (signal matching, Stereo) and 2 (Motion, Registration, Morphing), including large deformation experiments.

Keywords: motion estimation, registration, optical flow, multi-scale, motion constraint equation, global minimization, stereo matching

1 Introduction

Registration and motion estimation are one of the most challenging problems in computer vision, having uncountable applications in various domains [17, 18, 7, 1, 13, 28]. These problems occur in many applications like medical image analysis, recognition, visual servoing,

Figure 1: Finding the motion in a two-dimensional images sequence

stereoscopic vision, satellite imagery or indexing. Hence they have constantly been addressed in the literature throughout the development of image processing techniques. As a first example (Figure 1) consider the problem of finding the motion in a two-dimensional images sequence. We then look for a displacement $(h_1(x_1, x_2), h_2(x_1, x_2))$ that minimizes an energy functional:

$$\int\int |I_1(x, y) - I_2(x + h_1(x, y), y + h_2(x, y))|^2 dxdy.$$

Next consider the problem of finding a rigid or non rigid deformation $(f_1(x_1, x_2), f_2(x_1, x_2))$ between two images (Figure 2), minimizing an energy functional:

$$\int\int |I_1(x, y) - I_2(f_1(x, y), f_2(x, y))|^2 dxdy.$$

At last consider the stereoscopic matching problem: given a stereo pair (Figure 3), the epipolar constraint allows to split the two-dimensional matching problem into a series of line by line one-dimensional matching problems. One has just to find, for every line, the disparity $h(x)$ minimizing:

$$\int |I_1(x) - I_2(x + h(x))|^2 dx.$$

Although most papers deal only with motion estimation or matching depending on the application in view, both problems can be formulated the same way and be solved with the same algorithm. Thus the work we present can be applied both to registration for a pair of images to match (stereo, medical or morphing) or motion field / optical flow for a sequence of images. In this paper we will focus our attention on these problems assuming grey level conservation between both signals or images to be matched. Let us denote by $I_1(x)$ and $I_2(x)$ respectively

the study and target signals or images to be matched, where $x \in D = [-M, M]^d \subset I\!R^d$, and $d \geq 1$. In the following I_1 and I_2 are supposed to belong to the space $C_0^1(D)$ of continuously differentiable functions vanishing on the domain boundary ∂D. We will then assume there exists a homeomorphism f^* of D which represents the deformation such that:

$$I_1(x) = I_2 \circ f^*(x), \forall x \in D.$$

In the context of optical flow estimation, let us denote by h^* its associated motion field defined by $h^* = f^* - Id$ on D. We thus have:

$$I_1(x) = I_2(x + h^*(x)). \tag{1}$$

h^* is obviously a global minimum of the nonlinear functional

$$E_{NL}(h) = \frac{1}{2} \int_D |I_1(x) - I_2(x + h(x))|^2 dx. \tag{2}$$

We can deduce from (1) the well known Motion Constraint Equation (also called Optical Flow Constraint):

$$I_1(x) - I_2(x) \simeq < \nabla I_2(x), h^*(x) > , \forall x \in D. \tag{3}$$

E_{NL} is classically replaced in the literature by its quadratic version substituting the integrand with the squared difference between both left and right terms of the MCE, yielding the classical energy for the optical flow problem:

$$E_L(h) = \frac{1}{2} \int_D |I_1(x) - I_2(x) - < \nabla I_2(x), h(x) > |^2 dx.$$

Here ∇ denotes the gradient operator. Since the work of Horn and Schunk [17], MCE (3) has been widely used as a first order differential model in motion estimation and registration algorithms. In order to overcome the too low spatio-temporal sampling problem which causes numerical algorithms to converge to the closest local minimum of the energy E_{NL} instead of a global one, Terzopoulos et al. [23, 28] and Adelson and Bergen [6, 27] proposed to consider it at different scales. This led to the popular coarse-to-fine minimizing technique [18, 10, 13, 24, 14]. It is based on the remark that MCE (3) is a first order expansion which is generally no longer valid with h^* searched for. The idea is then to consider signals or images at a coarse resolution and to refine iteratively the estimation process. Since then many authors pointed out convergence properties of such algorithms towards a dominant motion in the case of motion estimation [12, 10, 11, 20, 5, 15], or an acceptable deformation in the case of registration [13, 24, 25], even if the initial motion were large. Let us mention that many authors assume that

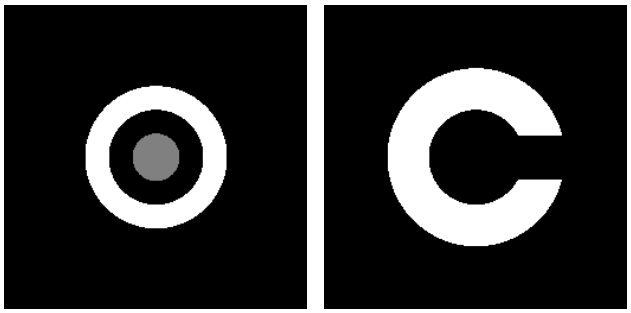


Figure 2: Finding a non rigid deformation between two images

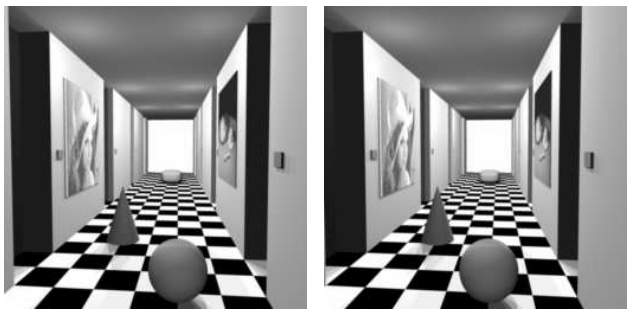


Figure 3: Finding Correspondence in a Stereo pair with epipolar constraint

deformation fields have some continuity or regularity properties, leading to the addition of some particular regularizing terms to the quadratic functional [17, 9, 28, 2, 4]. This very short state-of-the-art is far from being exhaustive but it allows to raise four common features shared by all most effective differential techniques:

1. a motion constraint equation,
2. a regularity hypothesis on the deformation,
3. a multi-scale approach,
4. an iterative scheme.

However, most of the multi-scale approaches assume that the MCE is more "valid" at lower resolutions. But to our knowledge and despite the huge literature, no theoretical analysis can confirm this. It may come from the fact that flattened signals or images are always "more similar". Choosing a particular low pass filter Π_σ (here $\sigma \geq 0$ is proportional to the number of considered harmonics in the Fourier decomposition) and a deformation $f = Id + h$ satisfying some local rigidity hypothesis with respect to a signal or image I_1, we shall find a linear operator $P_\sigma^{I_1}$

depending on I_1 such that:

$$\Pi_\sigma\Big(I_1 - I_2\Big) \simeq P_\sigma^{I_1}(h), \tag{4}$$

the sharpness of this approximation being decreasing with respect to both h norm and resolution parameter σ. We are faced with the following motion size/structure hypothesis trade-off: for some fixed estimation reliability, the larger the motion, the poorer its structure. This transforms the problem to solving the energy minimization in a finite dimensional subspace of approximation obtained through Fourier Decomposition. In this context we are led to consider the new energy to be minimized:

$$E_L(h) = \frac{1}{2}\int_D |\Pi_\sigma\Big(I_1 - I_1 \circ (Id + h)^{-1}\Big) - P_\sigma^{I_1}(h)|^2 dx.$$

Considering general linear parametric motion models for h^*, we give sufficient conditions for asymptotic convergence of the sequence of combined motion estimations towards h^* together with the numerical convergence of the sequence of deformed templates towards the target I_2. Roughly speaking, the shape of the theorem will be the following:
Theorem: If

1. at each step the residual deformation is "locally rigid", and the associated motion can be linearly decomposed onto an "acceptable" set of functions the cardinal of which is not too large with respect to the scale,
2. the initial motion norm is not too large, and the systems conditionings do not decrease "too rapidly" when iterating,
3. the estimated deformations $Id + \hat{h}_i$ are invertible and "locally rigid",

Then the scheme "converges" towards a global minimum of the energy E_{NL}.
The outline of the paper is as follows. In Section 2 we show the energy convexifying properties of multi-scale approaches together with fast convergence of iterative algorithms for the estimation of purely translational motion in any dimension. In Section 3 we turn to the general motion case and introduce a new local rigidity hypothesis and a low pass filter in order to derive a new MCE of the type of equation (4). In Section 4 we design an iterative motion estimation/registration scheme based on the MCE introduced in Section 3 and prove a convergence theorem. In order to avoid the a priori motion representation problem, we try successively two different approaches for the numerical resolution. We first use a level sets approach in Section 5, that

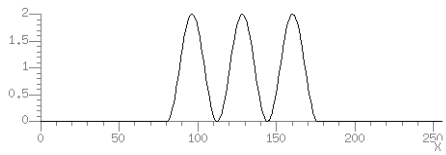


Figure 4: Test Signal. The second signal is the same shifted by 200.

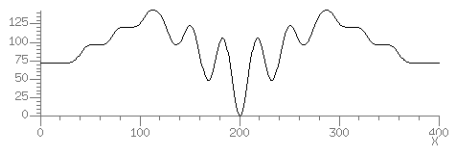


Figure 5: Energy as a function of shift parameter. There are numerous local minima around the global minimum at $x = 200$ at scale 7.

does not prove tractable nor robust. In Section 6 we adopt an implicit approach and constrain each estimated deformation $Id + \hat{h}_i$ to be at least invertible. We show numerical results for some signals and the stereo problem in dimension 1, and for large deformations problems in dimension 2. Section 7 gives a general conclusion to the paper.

2 Purely translational motion estimation

In this section we assume the motion to be found is only translational. This simple case will allow us to show the energy convexifying properties of multi- scale approaches together with fast convergence of iterative algorithms.

2.1 Synthetic 1D energy convexifying example

Consider a test signal (Figure 4) and its purely translated copies. The energy given by the mean quadratic error between shifted test signals and considered as a function of the translational parameter can be convexified using signals at a poorer resolution. Indeed we show the energy as a function of the translation parameter calculated with original test signals (Figure 5) and with same signal at a poorer resolution (Figure 6), namely signals reconstructed with only 5 and 3 first harmonics of the Fourier base. This readily yields more and more convexified energies as the resolution is lower. Based on this convexifying property, a generic algorithm for estimating the translational parameter is as follows:

1. Find the finest resolution j for which the energy is convex enough.
2. Minimize the energy with signals at resolution j.

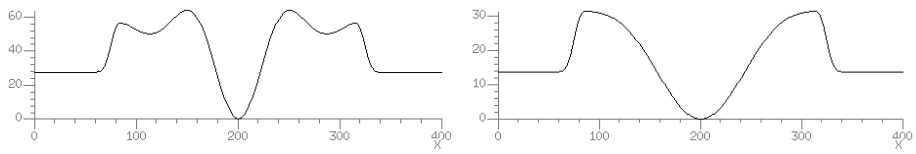


Figure 6: Same energy with signals reconstructed with only 5 harmonics (left) and 3 harmonics (right) using the multiresolution pyramid spanned by the first elements of the Fourier base.

3. Refine the result by increasing the resolution and minimizing the new energy.

2.2 The one dimensional case

Let us introduce some useful notations and technical hypothesis:

- I_1 and I_2 belong to $C_0^1(D)$,
- $D = [-M, M]$,
- h^* satisfies $|h^*| \leq dist(Supp(I_2), \partial D)$, where $dist$ denotes the Hausdorff distance between two sets of points, $Supp(I_2)$ denotes the set of points where I_2 is different from zero, and ∂D denotes the boundary of D,
- $I_1(x) = I_2(x + h^*)$, for all $x \in D$.

The problem we are faced with writes:

$$(P) : \text{Find } \hat{h} = arg\min_h \|I_1(x) - I_2(x+h)\|_{L^2}^2,$$

where L^2 denotes the space of summable squares functions over D. We now define the multiresolution pyramid considering the sequence of spaces

$$V_j = span\{e_k(x) = \tfrac{1}{\sqrt{2M}} e^{-i\pi k x/M}, -2^j \leq k \leq 2^j\}.$$

Let us denote Π_j the projection operator of L^2 onto V_j. The linearized problem in V_j writes:

$$(PL_j)\text{: Find } \hat{h^j} = arg\min_h E_j(h) = \|I_1(x) - I_2(x) - I_2'(x)h\|_{V_j}^2.$$

Our first result will be the

Lemma 1 *If* $\|I_2'\|_{V_j} \neq 0$, *then*

$$\hat{h^j} = \frac{< \Pi_j(I_2'), \Pi_j(I_1 - I_2) >_{V_j}}{\|I_2'\|_{V_j}^2}, \tag{5}$$

and if $|h^*| \leq \frac{M}{2^{j+1}}$ *then we have:* $|\hat{h^j} - h^*| \leq \frac{|h^*|}{2}$.

Proof. See Appendix. ■

To iterate the estimation process we introduce some notations: let $\hat{h_j^0} = 0$, $I_{2,0}(x) = I_2(x)$, and, for each $L > 0$, $\hat{h_j^L} = arg\min_h \|I_1(x) - I_{2,L-1}(x) - I'_{2,L-1}(x)h\|^2_{V_j}$, and $I_{2,L}(x) = I_2(x + \sum_{l=0}^{L} \hat{h_j^l})$. As a result we have the

Theorem 1 *If $\|I'_2\|_{V_j} \neq 0$ and $|h^*| \leq \frac{M}{2^{j+1}}$, the algorithm converges in the sense that, when $L \to \infty$,*

$$\sum_{l=0}^{L} \hat{h^l} \longrightarrow h^*, \text{ and } I_{2,L} \longrightarrow I_1 \text{ uniformly.}$$

Proof. See Appendix ■

Remark. Calculating the translation parameter h^* is not considered to be a difficult task using the classical phase method. This only illustrates theoretical performance of a multiresolution algorithm in a simple case.

2.3 Generalization to dimension $d > 1$

Notations in this context are to be understood as follows:

- D now becomes $[-M, M]^d$ in $I\!R^d$.
- $I_{1,p}$, $I_{2,p}$, I_2, are functions from $I\!R^d$ to $I\!R$.
- h, $\tilde{h^*}$ are vectors in $I\!R^d$.
- $< ., . >$ denotes the scalar product in $I\!R^d$.
- $[., .]$ denotes the scalar product in L^2.

Once again and for technical reasons we assume that I_1 and I_2 belong to C_0^1, and $I_1(x) = I_2(x + h^*)$, $x \in D$, $h^* \in I\!R^d$, and that $dist(Supp(I_2), \partial D) \geq |h^*|$, where ∂D denotes the border of D. Let also consider extended versions of I_1 and I_2 by continuity to the whole space $I\!R^d$, in order that the expression bounding I_1 to I_2 be meaningful if $x + h^* \notin D$. The problem now writes:

$$(P)\ \hat{h} = arg\min_h \|I_1(x) - I_2(x + h)\|^2_{L^2}$$

Consider a set of approximation spaces for the problem, given by the following definition:

Definition 1 *Let i be the chosen component index. We denote by V_i^j the sequence of vector sub-spaces of L^2 defined by:*

$$V_i^j = vect\{e_k(x) = \frac{1}{(2M)^{d/2}} e^{-i\pi k x_i/M}, k = -2^j, .., 0, .., 2^j\}.$$

For each space V_i^j, we denote by Π_i^j the operator from L^2 into L^2 mapping each function f to its reconstruction $\Pi_i^j f$ with its Fourier coefficients $c_k(f)$, $|k| \leq 2^j$:

$$\Pi_i^j f(x) = \sum_{|k| \leq 2^j} c_k(f) e_k(x).$$

Practically, we will have $2M$ samples in each direction, and we can therefore limit the problem study to its approximation in V_i^j spaces, where j is implicitly bounded by inequality $1 + 2^{j+1} \leq 2M$.

Let us call (PL) the problem associated to MCE: Find

$$\hat{h} = arg\min_{h \in I\!R^d} \|I_1(x) - I_2(x) - < \nabla I_2(x), h > \|_{L^2}^2,$$

and (PL_i^j) the problem embedded in V_i^j, $j \geq 1$:

$$(PL_i^j)\text{: Find } \hat{h_i^j} = arg\min_h \|I_1(x) - I_2(x) - \partial_i I_2(x) h\|_{V_i^j}^2,$$

where $\partial_i I_2$ denotes the partial derivative of I_2 w.r.t. component index i. A straightforward result similar to the previous one is given in

Lemma 2 *If $\|\partial_i I_2\|_{V_i^j} > 0$, then*

$$\hat{h_i^j} = \frac{[\Pi_i^j(\partial_i I_2), \Pi_i^j(I_1 - I_2)]}{\|\partial_i I_2\|_{V_i^j}^2}, \tag{6}$$

and if $|h_i^*| \leq \frac{M}{2^{j+1}}$, then we have:

$$|\hat{h_i^j} - h_i^*| \leq \frac{|h_i^*|}{2}. \tag{7}$$

Proof. It is exactly the same as in the one dimensional case. ■

Remark. As a first consequence, if $\|\partial_i I_2\|_2 = 0$, then it has no sense to estimate the translation parameter in this direction (aperture problem). In that particular case we will assume that it is null, and a zero value will then be given to its estimator.

We can see that if we replace I_2 by $I_{2,1}(x) = I_2(x + \hat{h^j})$, then the hypotheses of last Lemma will be satisfied again for $\|\partial_i I_{2,1}\|_{V_i^j} = \|\partial_i I_2\|_{V_i^j} > 0$ and:

$$|\hat{h_i^j} - h_i^*| \leq \frac{|h_i^*|}{2} \leq \frac{M}{2^{j+2}} \leq \frac{M}{2^{j+1}}.$$

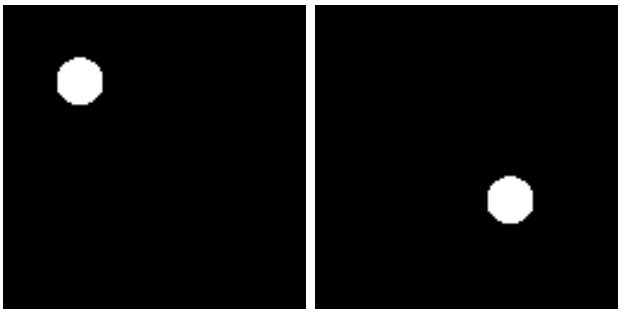


Figure 7: Test image on the left. The second image is the same translated with parameter (100,100). Translation parameters are found exactly without the need for scales greater than 1.

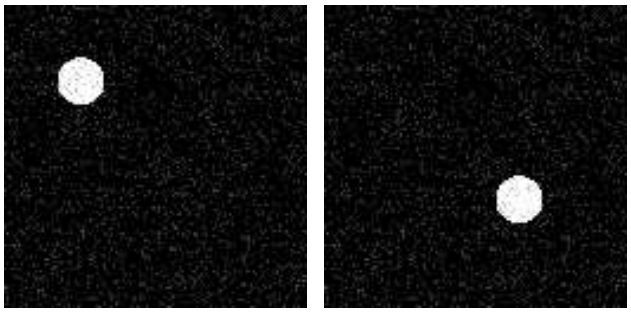


Figure 8: On the left, test image with 50% pixels corrupted by a Gaussian additive noise. The second image is the same before adding noise, then translated by a (100,100) shift, and finally also corrupted by the same type of noise.

We can therefore find

$$h_i^{\hat{j},1} = arg \min_h \|I_1(x) - I_{2,1}(x) - \partial_i I_{2,1}(x) h\|^2_{V_i^j},$$

and the Lemma allows to show that the sequence of functions $I_{2,L}(x) = I_2(x + \sum_{l=0}^{L} h^{\hat{j},l})$ converges uniformly towards I_1 (See Appendix).

We show in figures 7 to 9 some numerical results of two-dimensional purely translational motion estimation and registration. In Figure 8, yielded translation parameter is (99.03,99.07). Surprisingly we note that during the iterative process, the estimated translation parameter was best estimated before reaching the finest resolution, and then became less precise.

3 General motion multiresolution estimation

In Section 2 we have considered only purely translational motion estimation and registration. Our purpose here is to take over the general

case for the motion. We will first try to take some distance with what was done in the past concerning differential models and establish the need and the means of a constructive approach. Our approach is based on the fact that the motion is hidden in the difference between both functions to be matched. This will lead us to analyze this difference at some particular resolution. Making some assumptions on the structure and local behavior of the motion and the type of scale-space, we will find a new MCE and show that we can control the sharpness of it, which has not been taken care of previously.

3.1 Controlling the residuals when mixing differential and scale-space techniques

Using a regularizing kernel G_σ at scale σ, Terzopoulos et al. [23, 28] and Adelson and Bergen [6] were led to consider the following modified MCE:

$$G_\sigma * (I_1 - I_2)(x) \simeq < G_\sigma * \nabla I_2(x), h^*(x) >$$

Remark. One could also consider regularizing both left and right terms of the original MCE, yielding the following alternative:

$$G_\sigma * (I_1 - I_2)(x) \simeq G_\sigma * (< \nabla I_2, h^* >)(x)$$

At finest scales it can be shown that these two propositions are equivalent.
To our knowledge and despite the huge literature on these approaches, no theoretical error analysis can be found when such approximations are done. However it has been reported from numerical experiments that the modified MCE was not performing well at very coarse scales, thus betraying its progressive lack of sharpness. Assuming a local rigidity hypothesis and adopting the Dirichlet operator Π_σ, we will find a different right hand side featuring a "natural" and unique linear operator $P_\sigma^{I_1}$ in the sense that:

$$\Pi_\sigma(I_1 - I_2)(x) \simeq P_\sigma^{I_1}(h^*)(x), \tag{8}$$

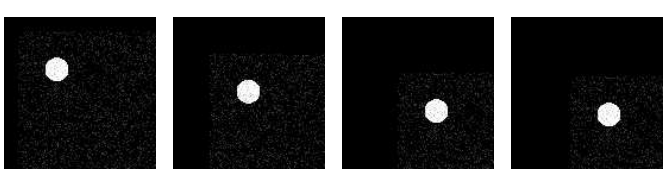


Figure 9: Registration with Gaussian additive noise. From left to right the four first iterations of the process at scale 1. We come up with a translation parameter of (99.66,100.07).

with remainder of the order of $\|h^*\|^2$ for some particular norm and vanishing as the scale is coarser.

3.2 Local rigidity property

In this paragraph we introduce our local rigidity property of deformations.

Definition 2 *$f \in Hom(D)$ is ξ-rigid for $I_1 \in C^1(D)$ iff:*

$$Jac(f)^t.\nabla I_1 = det(Jac(f))\nabla I_1, \tag{9}$$

where $Jac(f)$ denotes the Jacobian matrix of f and $det(A)$ the determinant of matrix A, and $Hom(D)$ the space of continuously differentiable and invertible functions from D to D (homeomorphisms).

All ξ-rigid deformations have the following properties (see [19] for the proofs). Assume f^* is ξ-rigid for $I_1 \in C_0^1(D)$ and $I_1 = I_2 \circ f^*$. Then,

1. equation (9) is always true if dimension d is 1;
2. for all $d \geq 1$,
   (a) $\|\nabla I_1\|_{L^1} = \|\nabla I_2\|_{L^1}$, where L^1 denotes the space of integrable functions over D;
   (b) $\nabla I_1 \ // \ \nabla I_2 \circ f^*$.
   (c) relation $\sim$ defined by
   $$[I_1 \sim I_2] \Longleftrightarrow [\exists f \ \xi\text{-rigid for } I_1 \text{ s.t. } I_1 = I_2 \circ f]$$
   is an equivalence relation on $C_0^1(D)$;
3. suppose $d = 2$: then,
   (a) if $Jac(f^*)$ is symmetric, then (9) means that if $|\nabla I_1| \neq 0$,
   - direction $\eta = \frac{\nabla I_1}{|\nabla I_1|}$ is eigenvector ($\lambda = det(Jac(f)$ is an eigenvalue);
   - direction $\xi = \frac{\nabla I_1^\perp}{|\nabla I_1|}$ is "rigid" ($\lambda = 1$ is an eigenvalue);

   This property can be seen as a non-sliding motion property. We illustrated this interesting property in Figure 10, where we show a level set of I_1, and a motion $h = f - Id$ of a ξ-rigid deformation f for image I_1. h can vary only along the direction of ∇I_1.

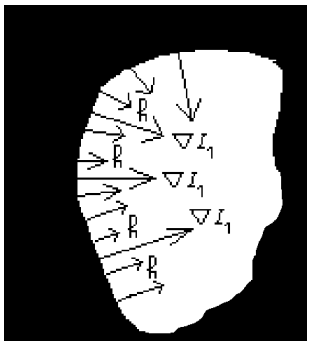


Figure 10: An example of motion $h = f - Id$ of a ξ-rigid deformation f for image I_1. We show a level set of image I_1, and the fields ∇I_1 and h along its boundary. h varies only along the direction of ∇I_1.

(b) $\kappa(I_1) = [Tr(Jac(f^*)) - det(Jac(f^*))].\kappa(I_2) \circ f^*$, where $\kappa(I)(x)$ stands for the curvature of the level line of I passing through x and $Tr(A)$ denotes the trace of matrix A;

4. if $d = 1$ or 2, and
   - h^* is known at
     - 1 point ($d = 1$).
     - each isolated critical point of I_1 and at one interior point of each connected constant set of I_1 ($d = 2$).
   - $h = h^*$ at this(ese) point(s), and

$$I_1 = I_2 \circ (Id + h) \text{ on } D,$$

then for all $x \in D$ where I_1 is not locally constant we have $h(x) = h^*(x)$.

Remark. It is an important issue to know whether such h^* is unique. In case $d \in \{1, 2\}$, property 4 leads to uniqueness if h^* is known at some isolated points. Though it is not proved in the general case, we will assume uniqueness hereafter for simplicity.
As a consequence we can show that ξ-rigid deformations of signals or images can be transferred to test functions. Indeed, we have the following

Lemma 3 *Suppose that*

1. *I_1 and $I_2 \in C_0^1(D)$ are such that: $I_1 = I_2 \circ f$*
2. *f is ξ-rigid for I_1*

3. *$\phi \in C^\infty(D;\mathbb{R})$, and $\Phi \in C^\infty(D;\mathbb{R}^d)$ s.t. $div\Phi = \phi$, where $C^\infty(D;\mathbb{R})$ denotes the space of indefinitely differentiable function from D to $\mathbb{R}$.*

Then, $\int_D (I_1 - I_2)\phi dx = \int_D < \nabla I_1, \Phi \circ f - \Phi > dx$.

Proof. See Appendix ■

3.3 The Dirichlet operator

One choice for the set of test functions in Lemma 3 is the Fourier basis, the simplest projection onto which is the Dirichlet projection operator. Let $D = [-M, M]^d$; $S_\sigma = \{k \in Z^d, \forall i \in [1,d], |k_i| \leq M\sigma^2\}$; $c_k(I)$ denotes the Fourier coefficient of I defined by:

$$c_k(I) = \frac{1}{(2M)^{\frac{d}{2}}} \int_D I(x) e^{-\frac{i\pi < k,x >}{M}} dx.$$

Then the Dirichlet operator Π_σ is the linear mapping associating to each function $I \in C_0^1(D)$ the function $\Pi_\sigma(I) = G_\sigma * I$, where the convolution kernel G_σ is defined by its Fourier coefficients as follows:

$$c_k(G_\sigma) = \begin{cases} 1 \text{ if } k \in S_\sigma \\ 0 \text{ elsewhere} \end{cases}$$

3.4 New MCE by Linearization for the Dirichlet projection

Now that we have introduced our rigidity property of deformations and the Dirichlet projection, let us choose the test functions of Lemma 3 in the Fourier basis. We obtain the

Lemma 4 *If $f^* = Id + h^*$ is ξ-rigid for $I_1 = I_2 \circ f^*$, with I_1, I_2 both in C_0^1, then*

$$c_0(I_1 - I_2) = \frac{1}{d} c_0(< \nabla I_1, h^* >) \tag{10}$$

$$\text{and, for } k \neq 0, \quad c_k(I_1 - I_2) = \frac{iM}{\pi |k|^2} c_k\Big(< \nabla I_1, k > (e^{-i\pi < k, h^* > / M} - 1)\Big). \tag{11}$$

Proof. If $k = 0$, we take $\Phi(x) = \frac{1}{d}x$, which yields the expected $\phi(x) = div(\Phi(x)) = 1$. If $k \neq 0$, then we must find Φ such that $div(\Phi(x)) =$

$\phi(x) = e^{-i\pi<k,x>/M}$. Fortunately in this case we have an explicit solution which is given by $\Phi(x) = \frac{iM}{\pi}\frac{k}{|k|^2}e^{-i\pi<k,x>/M}$. ■

Now taking the linear part of the jet of $e^{-i\pi<k,h^*>/M} - 1$ with respect to h^*, and setting:

$$c_k(P_\sigma^{I_1}(h^*)) = \begin{cases} \frac{1}{d}c_0(<\nabla I_1, h^*>) \text{ if } k = 0 \\ c_k(\frac{<\nabla I_1,k><k,h^*>}{|k|^2}) \text{ if } k \in S_\sigma/\{0\} \\ 0 \text{ if } k \notin S_\sigma \end{cases}$$

we obtain the

Theorem 2 *If $f^* = Id + h^*$ is ξ-rigid for $I_1 = I_2 \circ f^* \in C_0^1(D)$, then we have:*

$$\|\Pi_\sigma(I_1 - I_2) - P_\sigma^{I_1}(h^*)\|_{L^2} \leq \frac{\pi}{2}\sigma^{d+2}\|h^*|\nabla I_1|^{\frac{1}{2}}\|_{L^2}^2.$$

This inequality is nothing but the sharpness of MCE (8):

$$\Pi_\sigma(I_1 - I_2)(x) \simeq P_\sigma^{I_1}(h^*)(x), \tag{12}$$

at scale σ. It clearly expresses the fact that measuring the motion (e.g perceiving the optical flow) h^* is not relevant outside of the support of $|\nabla I_1|$.

Proof. See Appendix ■

4 Theoretical iterative scheme and convergence theorem

In section 3 we found a new MCE and showed that we can control the sharpness of it. In this section we will make a rather general assumption on the motion in the sense that it should belong to some linear parametric motion model without being more specific on the model basis functions. Though it is somewhat restrictive to have motion fields in a finite dimensional functional space, this structural hypothesis will be a key to bounding the residual motion norm after registration in order to iterate the process. This makes it possible to consider a constraint on motion when there is a priori knowledge (like for rigid motion) or consider multi-scale decomposition of motion for an iterative scheme.

4.1 Linear parametric motion models and least square estimation

Let us assume the motion h^* has to be in a finite dimensional space of deformation generated by basis functions $\Psi(x) = (\psi_i(x))_{i=1..n}$. Thus h^*

can be decomposed in the basis: $\exists\ \Theta^* = (\theta_i^*)_{i=1..n}$ unique, such that:

$$h^*(x) =< \Psi(x), \Theta^* >=\sum_{i=1..n} \theta_i^* \psi_i(x),\ \forall x \in Supp(|\nabla I_1|).$$

MCE (8) viewed as a linear model writes:

$$\Pi_\sigma(I_1 - I_2) =< P_\sigma^{I_1}(\Psi), \Theta^* > .$$

Now set, for σ s.t. the $P_\sigma^{I_1}(\psi_i)$ be mutually linearly independent in L^2:

$$M_\sigma = P_\sigma^{I_1}(\Psi) \otimes P_\sigma^{I_1}(\Psi),\ \ Y_\sigma = \Pi_\sigma(I_1 - I_2),$$

where $\otimes$ stands for the tensorial product in L^2. Then applying basic results from the classical theory of linear models yields: $\hat{h} =< \Psi, \hat{\Theta} >=< \Psi, M_\sigma^{-1} B_\sigma >$, where column B_σ's components are defined by $(B_\sigma)_i =< P_\sigma^{I_1}(\psi_i), Y_\sigma >$.

4.2 Estimation error and residual motion

Given the least square estimation of the motion of last paragraph, we have

Lemma 5 *In this framework the motion estimation error is bounded by inequality*

$$\|(\hat{h} - h^*)|\nabla I_1|^{\frac{1}{2}}\|_{L^2} \leq \frac{\pi}{2}\sigma^{d+2}\Big(Tr(M_\sigma^{-1}))\Big)^{\frac{1}{2}}\|h^*|\nabla I_1|^{\frac{1}{2}}\|_{L^2}^2.$$

Proof. See Appendix ■

If $Id + \hat{h}$ is invertible, we can define:

$$I_{1,1} = I_1 \circ \Big(Id + \hat{h}\Big)^{-1}. \tag{13}$$

Letting r_1 denote the residual motion such that $I_{1,1} = I_2 \circ (Id + r_1)$, if $Id + \hat{h}$ is ξ-rigid for I_1 then a variable change yields equality

$$\|(\hat{h} - h^*)|\nabla I_1|^{\frac{1}{2}}\|_{L^2} = \|r_1|\nabla I_{1,1}|^{\frac{1}{2}}\|_{L^2},$$

thus giving by Lemma 5 the following bound on the residual motion norm:

$$\|r_1|\nabla I_{1,1}|^{\frac{1}{2}}\|_{L^2} \leq \frac{\pi}{2}\sigma^{d+2}\Big(Tr(M_\sigma^{-1}))\Big)^{\frac{1}{2}}\|h^*|\nabla I_1|^{\frac{1}{2}}\|_{L^2}^2. \tag{14}$$

In view of equality (13) and inequality (14), iterating the motion estimation/registration process looks completely natural and allows for pointing out sufficient conditions for convergence of such a process. Indeed, provided the same assumptions are made at each step, relations (13) and (14) can be seen as recurrence ones, yielding both r_p and $I_{1,p}$ sequences.

4.3 Theoretical iterative scheme

Having control on the residual motion after one registration step, we deduce the following theoretical iterative motion estimation / registration scheme:

1. Initialization: Enter accuracy $\epsilon > 0$ and the maximal number of iterations N. Set $p = 0$, and $I_{1,0} = I_1$.
2. Iterate while $(\|I_{1,p} - I_2\| \geq \epsilon \ \& \ p \leq N)$
   (a) Enter the set of basis functions $\Psi_p = (\psi_{p,i})_{i=1..n_p}$ that linearly and uniquely decompose r_p on the support of $|\nabla I_{1,p}|$.
   (b) Enter scale σ_p and compute: $\hat{h}_p =< \Psi_p, M_{p,\sigma_p}^{-1} B_{\sigma_p} >$.
   (c) Set $I_{1,p+1} = I_{1,p} \circ (Id + \hat{h}_p)^{-1}$.

4.4 Convergence theorem

Now that we have designed an iterative motion estimation / registration scheme, let us infer sufficient conditions for the residual motion to vanish. This leads us to state our following main result:

Theorem 3 *If:*

1. *For all $p \geq 0$, $I_{1,p} \sim I_2$ (as defined in Section 3.2), and the residual motion r_p can be linearly and uniquely decomposed on a set of basis functions $\{\psi_{p,i}, i = 1..n_p\}$;*
2. *For all $p \geq 0$, there exists a scale $\sigma_p > 0$ such that the set of functions $\{P_{\sigma_p}^{I_{1,p}}(\psi_{p,i}), i = 1..n_p\}$ be free in L^2 and, for $p = 0$, we assume that :*
$$\|h^* |\nabla I_1|^{\frac{1}{2}}\|_{L^2} < \Big(\frac{\pi}{2}\sigma_0^{d+2} Tr(M_{0,\sigma_0})^{\frac{1}{2}}\Big)^{-1};$$
   Set $C_0 = \Big(\frac{\pi}{2}\sigma_0^{d+2} Tr(M_{0,\sigma_0})^{\frac{1}{2}} \|h^* |\nabla I_1|^{\frac{1}{2}}\|_{L^2}\Big)^{-1}$;
3. *The sequence of conditioning ratios satisfy criteria:*
   $\forall p \geq 0, \ \frac{\sigma_{p+1}^{d+2} Tr(M_{p+1,\sigma_{p+1}})^{\frac{1}{2}}}{\sigma_p^{d+2} Tr(M_{p,\sigma_p})^{\frac{1}{2}}} \leq C_0$;
4. *For all $p \geq 0$, the estimated deformations $Id + \hat{h}_p \in Hom(D)$ and are ξ-rigid for $I_{1,p}$;*

Then, $\lim_{p\to\infty} \|r_p |\nabla I_{1,p}|^{1/2}\|_{L^2} = 0$.

Proof. See Appendix ■

4.5 Numerical algorithm requirements

Firstly, due to the fact that h^* is unknown we have to make an arbitrary choice for the scale at each step. Secondly we at least have to ensure that $Id + \hat{h}$ be invertible at each step. Finally we are faced with the motion basis functions choice.

4.5.1 Multi-scale strategy

The scale choice expresses both a priori knowledge on the motion range and its structure complexity. Here we assume that $(\sigma_p)_p$ is an increasing sequence, starting from $\sigma_0 > 0$ such that:

$$\#S_{\sigma_0} \geq \#\{\text{expected independent motions}\}. \tag{15}$$

Then let $\alpha \in]0,1[$. In order to justify the minimization problem at new scale $\sigma_{p+1} > \sigma_p$, we will choose it such that:

$$\|(\Pi_{\sigma_{p+1}} - \Pi_{\sigma_p})(I_{1,p+1} - I_2)\|_{L^2} > \alpha \|I_{1,p+1} - I_2\|_{L^2}, \tag{16}$$

4.5.2 Invertibility of $Id + \hat{h}_p$

Let $\beta > 0$. We will apply to $I_{1,p}$ the inverse of the maximal invertible linear part of the computed deformation e.g. $\left(Id + t^*.\hat{h}_p\right)^{-1}$, where

$$t^* = \sup_{t \in [0,1]} \{t \ / \ det(Jac(Id + t.\hat{h}_p)) \geq \beta\}. \tag{17}$$

Remark **Recursive version of the algorithm**

Set $f^*(I_1, I_2)$ the solution to the correspondence problem between I_1 and I_2. Then, $f^*(I_{1,p}, I_2) = f^*(I_{1,p+1}, I_2) \circ (Id + \hat{h}_p)$. We thus deduce the following alternate recursive motion estimation / registration function $f^*(I_1, I_2)$ defined by:

$$\begin{cases} \text{If } \|I_1 - I_2\| > \epsilon, \\ \text{Then} \begin{cases} \text{Calculate } \hat{h}(I_1, I_2) \\ \text{Deform: } I_{1,1} = I_1 \circ (Id + \hat{h}(I_1, I_2))^{-1} \\ \text{Call } f = f^*(I_{1,1}, I_2) \\ \text{Return } f \circ (Id + \hat{h}(I_1, I_2)) \end{cases} \\ \text{Else return } Id \end{cases}$$

4.5.3 Choosing the set of basis functions

A major difficulty arising in the theoretical scheme comes from the lack of a priori knowledge on the finite set of basis functions to be entered at each step. To alleviate this problem we propose two different approaches. In Section 5 we will consider splitting both signals or images into a collection of pairs of level sets to be matched, whose basis functions are simple Dirac measures in dimension 1, and vector curves in dimension 2. In Section 6 we will use an implicit approach via the optimal step gradient algorithm when minimizing the quadratic energy associated to MCE (8).

5 Level sets approach of basis functions

In Section 4 we derived a theoretical iterative scheme and established a convergence theorem. In order to implement the proposed algorithm we at each step have to choose a finite dimensional motion model. In this section we consider splitting both signals or images into a collection of pairs of level sets to be matched, whose basis functions are simple Dirac measures in dimension 1, and vector curves in dimension 2. Indeed, using the level set decomposition of signals and images, we show that the energy minimization is equivalent to a series of independent distance minimizations between characteristic functions of the level sets of both signals or images. We design a procedure that achieves motion estimation and registration grey level by grey level. At each grey level the motion of the borders of the level sets is estimated recursively.

The initial matching problem can be split into a collection of independent sub- problems: for each $\lambda \in \{I_1(x)/x \in D\} \subset I\!R$, solve:

$$\min_h \|\chi_{\{I_1 \geq \lambda\}}(x) - \chi_{\{I_2 \geq \lambda\}}(x + h(x))\|^2_{L^2},$$

where χ denotes the characteristic function.

Indeed, if $I_1 = I_2 \circ f$, then the level sets of I_1 and $I_2 \circ f$ can be superposed. Conversely, we have:

$$\|I_1 - I_2 \circ f\|^2_{L^2} = \int_D |I_1(x) - I_2 \circ f(x)|^2 dx,$$
$$= \int_D \Big(\int_{-\infty}^{+\infty} |\chi_{\{I_1 \geq \lambda\}}(x) - \chi_{\{I_2 \geq \lambda\}} \circ f(x)| d\lambda \Big)^2 dx.$$

Now set:

$$a = min(\inf_D I_1, \inf_D I_2) \text{ and } b = max(\sup_D I_1, \sup_D I_2).$$

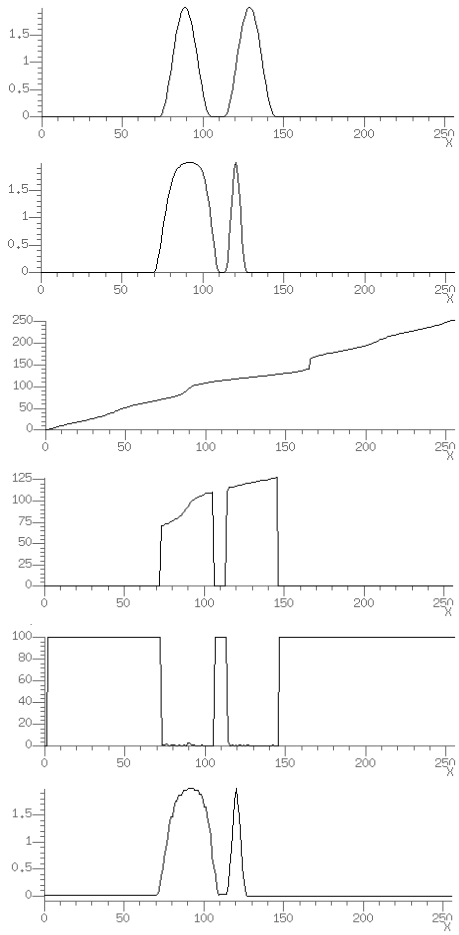


Figure 11: From top to bottom, we show I_2, then $I_1 = I_2 \circ f^*$, f^*, $\hat{f}$, then the error percentage between $\hat{f}$ and f^*, and finally $I_2 \circ \hat{f}$.

We can deduce (see Appendix) that:

$$\|I_1 - I_2 \circ f\|_{L^2}^2 \leq (b-a) \int_{-\infty}^{+\infty} \|\chi_{\{I_1 \geq \lambda\}} - \chi_{\{I_2 \geq \lambda\}} \circ f\|_{L^2}^2 d\lambda.$$

Consequently if the level sets can be superposed for almost each λ, then functions are almost everywhere equal.

As characteristic functions vary only at the borders of connected components, we shall estimate the motion only at the borders of the connected components of the grey level λ sets $\{I_1 \geq \lambda\}$ of I_1.

5.1 The one-dimensional case

In that case we only have to estimate the motion at the borders of each of the chosen level sets, so the unknown is a set of reals that are the amplitudes of the borders motions. Left and right borders of each component are asked to remain in the same order.

We illustrate our algorithm on a pair of 1D synthetic signals (Figure 11).

5.2 The two dimensional case

Here we suggest to proceed to the registration of characteristic functions of the level sets. In this context, we have the following

Proposition 1 *The operator kernel* $KerP_{\sigma}^{\chi_A} \neq \{0\}$ *for some of the open sets* $A \subset \mathbb{R}^2$ *satisfying the necessary and sufficient condition of the Conformal Representation Theorem of Riemann whose conformal application does not imply any local nor global rotation (See [19] for the proofs).*

This shows that the choice of the base functions remains a hard issue in this approach. However, we have tested our algorithm using translational and normal base functions.

Numerical results

We first show two synthetic examples of motion estimation and registration of shapes. Each figure shows the motion of a grey shape which is deformed iteratively to match the second shape represented by its contour. In the first example (Figure 12) we show a Chinese symbol that is translated towards the final contour. In the second example (Figure 13) we show a disk being translated and enlarged to match the final contour.

As a conclusion to this section, we suggest using it only in case the motion is rather uniform or can be modelled with translational and normal basis functions. Finally let us emphasize on its lack of robustness in the presence of noise.

6 Implicit approach of basis functions

In Section 5 we have shown the limitations of the level sets approach to alleviate the motion model choice for the motion estimation / registration scheme of Section 4. Here we suggest to use the optimal step gradient algorithm for the minimization of the quadratic functional associated to MCE (8). There are at least two good reasons for doing this:

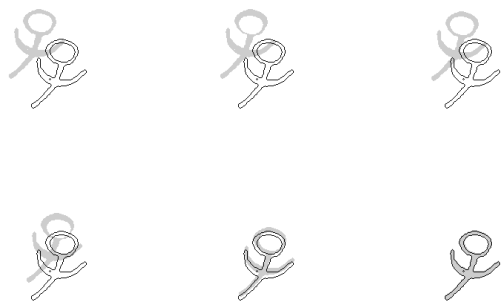


Figure 12: Iterations 1, 7, 13, 19, 25 and 30 of the algorithm for the registration of a Chinese symbol.

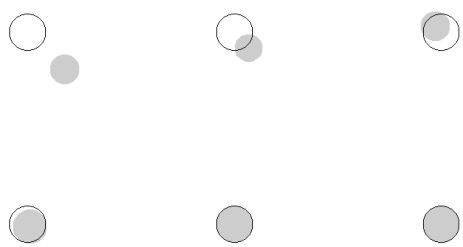


Figure 13: Iterations 1, 2, 3, 4, 5 and 6 of the algorithm for the registration of a translated and reduced disk.

- the choice of base functions is implicit: it depends on the signals or images I_1 and I_2, and the scale space.
- we can control and stop the quadratic minimization if the associated operator is no longer positive definite.

The general algorithm does not guaranty that the resulting matrix M_{p,σ_p} be invertible. Hence we suggest to systematically use a stopping criteria to control the quadratic minimization, based on the descent speed or simply a maximum number of iterations N_G.

In that case our final algorithm writes:

1. Initialization: Enter accuracy $\epsilon > 0$ and the maximal number of iterations N. Set $p = 0$, $I_{1,0} = I_1$, and choose first scale σ_0 according to (15).
2. Iterate while $(\|I_{1,p} - I_2\| \geq \epsilon \ \& \ p \leq N \ \& \ \sigma_p \leq 1)$

(a) Choose σ_p satisfying (16).

(b) Apply N_G iterations of the optimal step gradient algorithm for the minimization of

$$E_p(h) = \|\Pi_{\sigma_p}(I_{1,p} - I_2) - P_{\sigma_p}^{I_{1,p}}(h)\|_{L^2}^2.$$

(c) Compute $I_{1,p+1} = I_{1,p} \circ (Id + t^*.\hat{h}_p)^{-1}$ with t^* defined by (17) and increment p.

In the following experiments we have fixed parameters to $\alpha = 2.5\%$, $N_G = 5$, $\beta = 0.1$.

6.1 Running the 1 dimensional algorithm

In the following we show results on one-dimensional synthetic and real signals, and then with all intensity lines of a stereo pair of synthetic images with some progressively added Gaussian noise.

- **1D signal Matching**: We show 1D synthetic and real examples in Figure 14. Recall that ξ-rigidity is not a constraint when $d = 1$ and thus $\hat{h}_\infty$ is relevant only when $|I_1'(x)| \neq 0$.
- **Stereo Correspondence**: Since we use a synthetic pair of rectified images, epipolar lines are the lines of the images. Image matching is then solved as a sequence of 1D line matchings. In this case, ground truth disparity h^* is available. We see the results in Figures 15 and 16. The Mean Weighted Quadratic Error (indicated at top of each graph as EQPM) is defined as $MWQE = \||\partial_x I_1|^{1/2}|\hat{h} - h^*|\|_{L^2}^2 / \|\partial_x I_1\|_1$.

6.2 Running the 2 dimensional algorithm

We illustrate the algorithm on pairs of images with large deformation for registration applications and movies for motion estimation applications.

- **Registration problems involving large deformation**: In figures 17 and 18 we show the different steps of the algorithm performing the registration between the first and last images. In Figures 19 to 21, we show the study and target images, and the deformed study image after applying the estimated motion. This was applied for two examples of faces and a turbulence image featuring a vortex at two different states.

- **Optical Flow estimation examples**: in Figure 22 we show the sequence of the registered images of the original Cronkite sequence onto first image using the sequence of computed backward motions. The result is expected to be motionless. On top of Figure 23, we show the complete movie obtained by deforming iteratively only the first image of Cronkite movie. For that we use the sequence of computed motions between each pair of consecutive images of the original movie. In Figure 23 on the bottom, we see the error images.

7 Conclusion

We have addressed the theoretical problems of motion estimation and registration of signals or images in any dimension. We have used the main features of previous works on the subject to formalize them in a framework allowing a rigorous mathematical analysis. More specifically we wrote a new ridigity hypothesis that we used to infer a unique Motion Constraint Equation with small remainder at coarse scales. We then showed that upon hypotheses on the motion norm and structure/scale tradeoff, an iterative motion estimation/registration scheme could converge towards the expected solution of the problem e.g. the global minimum of the nonlinear least square problem energy. Since each step of the theoretical scheme needs a set of motion basis functions which are not known, we have first implemented a level sets approach, that prove not tractable nor robust. We then designed an implicit algorithm and illustrated the method in dimension one and two, including large deformation examples.

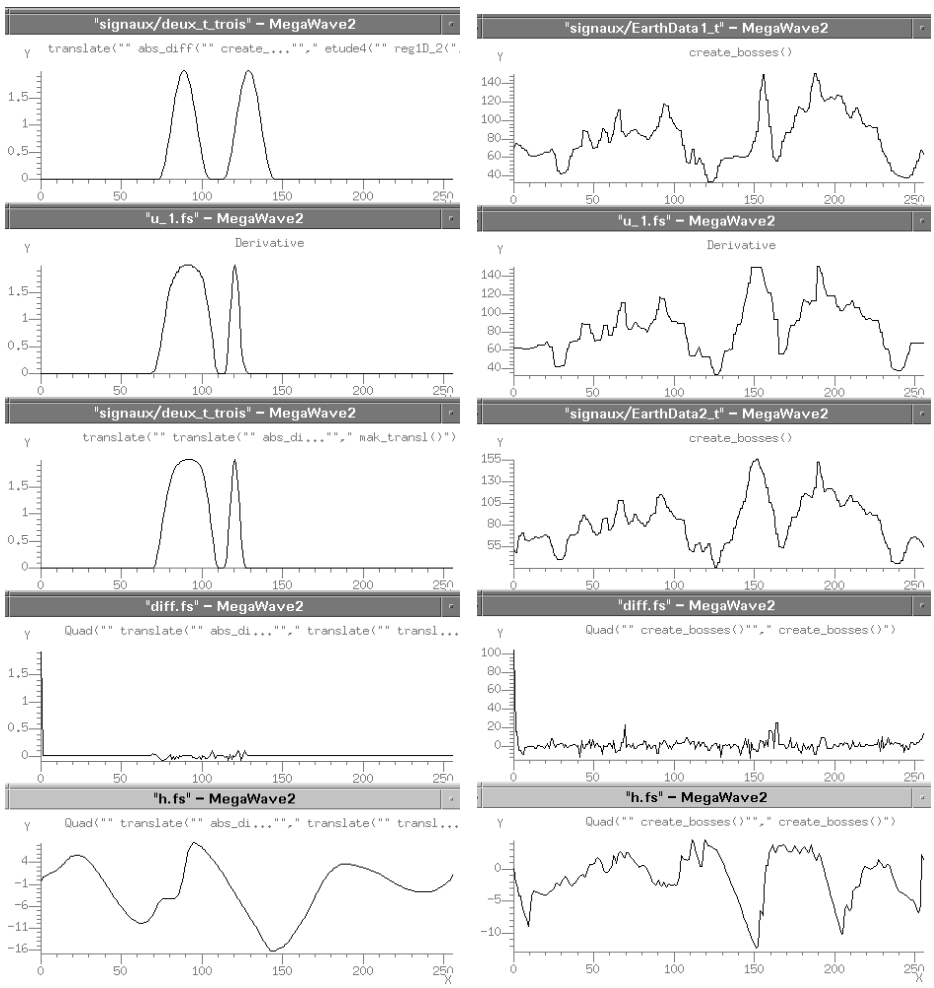


Figure 14: 1D synthetic (left) and 1D real (right) examples: from top to bottom we show signals I_1, $I_{1,\infty}$ (final deformation of I_1), I_2, $I_{1,\infty} - I_2$ and $\hat{h}_\infty$ (final deformation function).

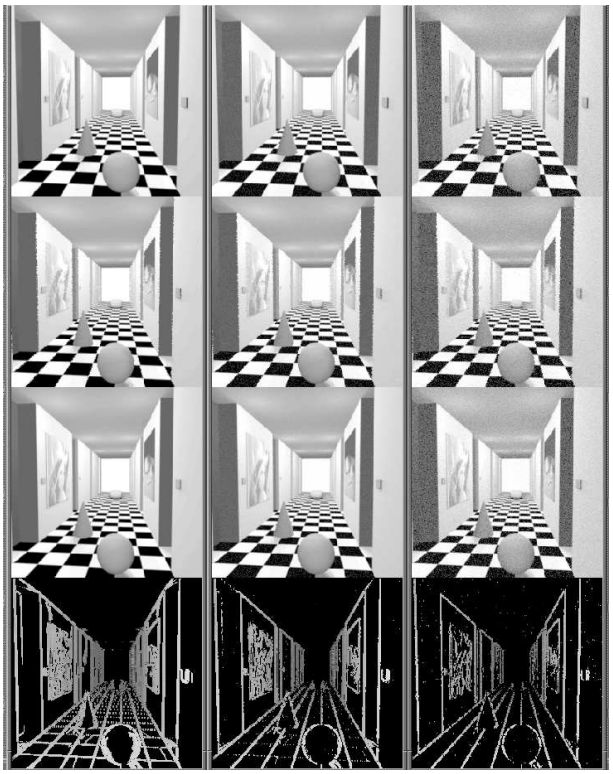


Figure 15: Stereo example: From left to right: noise free data experiments, and then with data corrupted by a Gaussian additive noise with variance 10 and 100. From top to bottom we show images I_1, $I_{1,\infty}$ that was processed line by line, I_2 and $\hat{h}_\infty$ (shown only if $|\partial_x I_1| > 2$, 5 and 12 resp.). We see that this disparity image has lower values (grey level) for points that are at the end of corridor.

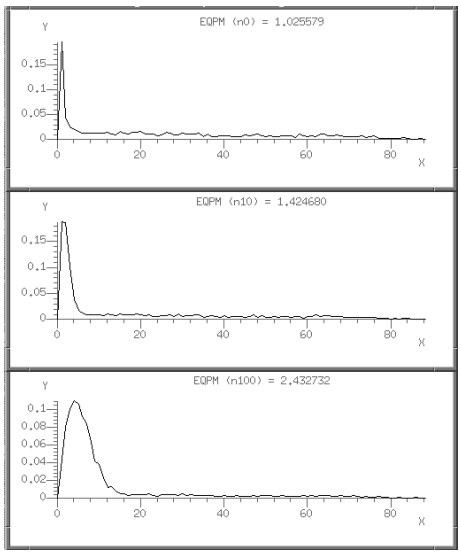


Figure 16: Comparison with ground truth disparity for the stereo example of Figure 15: Mean Weighted Quadratic Error (indicated at top of each graph) distributions as a function of the horizontal spatial gradient of I_1 for the stereo example with noise level 0 (top), 10 (middle) and 100 (bottom). Motion estimation errors are concentrated at low horizontal gradients of I_1, and diffuse to broader values as noise increases.

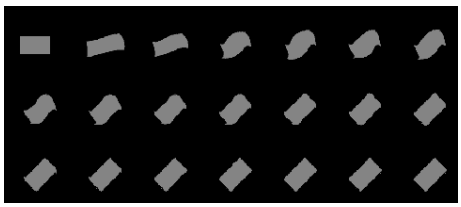


Figure 17: Registration movie of a rotated rectangle: from left to right and from top to bottom we show the different steps of the algorithm performing the registration.

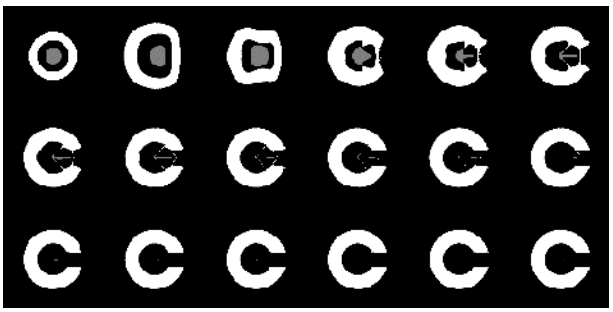


Figure 18: Registration movie of a target to a 'C' letter. Again, each image corresponds to a step in the iterative scheme.

Figure 19: Scene registration example: Study image (left), deformed Study image onto Target image (center), and Target image (right).

Figure 20: Registration of a face with two different expressions: Study image (left), deformed Study image onto Target image (center), and Target image (right).

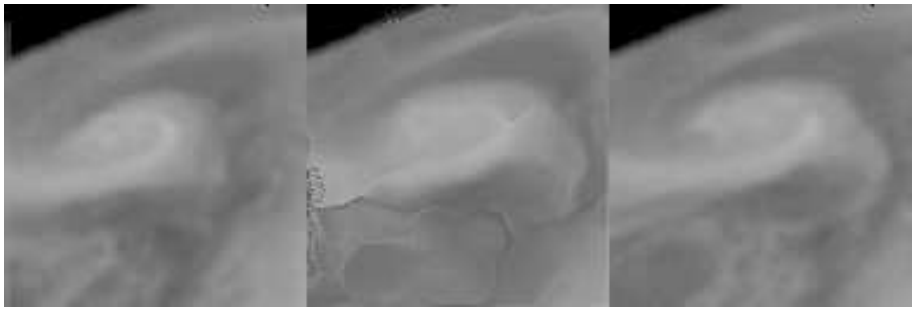


Figure 21: Registration of a vortex at two different states: Study image (left), deformed Study image onto Target image (center), and Target image (right).

Figure 22: Registered sequence of the original sequence onto first image using the computed backward motions.

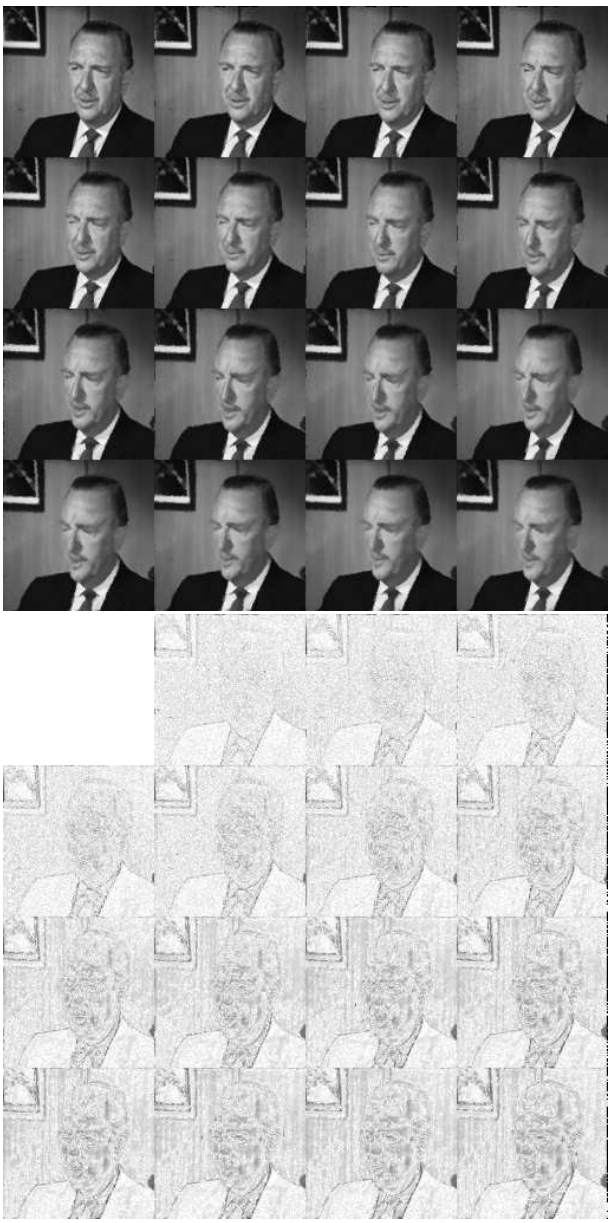


Figure 23: On top, movie obtained by deforming only the first image of Cronkite movie using the sequence of computed motions. On the bottom, enhanced (applying $I' = 255.(1-\sqrt{I/255})$) absolute difference between original and artificially deformed Cronkite sequences.

References

[1] N. Ayache. Medical computer vision, virtual reality and robotics. *Image and Vision Computing*, IVC (13), No 4, May 1995, pp 295-313.

[2] G. Aubert, R. Deriche and P. Kornprobst. Computing optical flow via variational techniques. *SIAM J. Appl. Math.*, Vol. 60 (1), pp. 156–182, 1999.

[3] L. Alvarez, J. Esclarin, M. Lefébure, J. Sanchez. A PDE model for computing the optical flow. *Proc. XVI Congresso de Ecuaciones Diferenciales y Aplicaciones*, Las Palmas, pp. 1349–1356, 1999.

[4] L. Alvarez, J. Weickert, J. Sanchez, Reliable estimation of dense optical flow fields with large displacements. TR 2, Instituto Universitario de Ciencias y Tecnologias Ciberneticas, Universidad de Las Palmas de Gran Canaria, November 1999.

[5] C. Bernard. Discrete wavelet analysis: a new framework for fast optic flow computation. *ECCV*, 1998.

[6] J.R. Bergen and E.H. Adelson. Hierarchical, computationally efficient motion estimation algorithm. *J. of the Optical Society Am.*, 4(35), 1987.

[7] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow. *International Journal of Computer Vision*, 12(1):43–77, 1994.

[8] M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. *Proc. Visualization in Biomedical Computing (VBC'96)* Springer LNCS 1131, Hamburg, Germany, pp 267–276, Sept. 1996.

[9] Ruzena Bajcsy and Stane Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics, and Image Processing*, (46), No 1, April:1–21, 1989.

[10] P. Bouthemy and J.M. Odobez. Robust multiresolution estimation of parametric motion models. *J. of Vis. Comm. and Image Repres.*, 6(4):348–365, 1995.

[11] M. Black and A. Rangajaran. On the unification of line processes, outlier rejection and robust statistics with applications in early vision. *International Journal of Computer Vision*, Vol. 19, 1996.

[12] M. Ben-Ezra, B. Rousso, and S. Peleg. Motion segmentation using convergence properties. In *ARPA Im. Unders. Workshop*, pp II 1233-1235, 1994.

[13] G. Christensen, R.D. Rabbitt, and M.I. Miller. 3D brain mapping using a deformable neuroanatomy. *Physics in Med and Biol*, (39), March :609–618, 1994.

[14] D. Fleet, M. Black, Y. Yacoob and A. Jepson. Design and use of linear models for image motion analysis. *IJCV*, 36(3), 2000.

[15] P.R. Giaccone, D. Greenhill, G.A. Jones. Recovering very large visual motion fields. In *Scandinavian Conference on Image Analysis SCIA97*, pp 917-922, 1997.

[16] B. Galvin, B. McCane, K. Novins, D. Mason and S. Mills. Recovering Motion fields: An analysis of eight optical flow algorithms, *Proc. 1998 British Machine Vision Conference, BMVC'98*, Sept. 1998.

[17] B.K.P. Horn and Brian Schunck. Determining optical flow. *Artificial Intelligence*, (17) (1-3) :185–204, 1981.

[18] M. Irani, B. Rousso, and S. Peleg. Detecting and tracking multiple moving objects using temporal integration. In *ECCV92*, pp 282–287, 1992.

[19] M. Lefébure Estimation de Mouvement et Recalage de Signaux et d'Images: Formalisation et Analyse. PhD Thesis, Université Paris-Dauphine, 1998.

[20] E. Mémin and P. Pérez. Dense estimation and object-based segmentation of the optical flow with robust techniques. IEEE Trans. IP, 1998.

[21] S. Srinivasan, R. Chellappa. Optical flow using overlapped basis functions for solving global motions problems. In *ECCV98*, pp 288–304, 1998.

[22] C. Stiller and J. Konrad. Estimating motion in image sequences. In *IEEE Signal Processing Magazine*, Vol. 16, July, pp 70–91, 1999.

[23] D. Terzopoulos. Multiresolution algorithms in computational vision. *Image Understanding*. S. Ullman, W. Richards, 1986.

[24] J.P. Thirion. Fast non-rigid matching of 3D medical images. Technical Report 2547, INRIA, 1995.

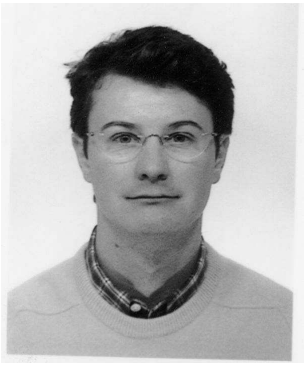
[25] A. Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3), 1998.

[26] J. Weickert. On discontinuity-preserving optic flow. *Proc. Computer Vision and Mobile Robotics Workshop*, Santorini, pp. 115–122, Sept. 1998.

[27] G. Whitten. A framework for adaptive scale space tracking solutions to problems in computational vision. In *Proc. IEEE International Conference on Computer Vision, Osaka*, pp 210–220, Dec 1990.

[28] A. Witkin, D. Terzopoulos, and M. Kass. Signal matching through scale space. *International Journal of Computer Vision*, 1(2):133–144, 1987.

Technical Biography of the Authors



Martin Lefébure was born in Paris, France, in 1969. He received his Master Thesis in mathematics from Paris University in 1993, where he was assistant professor in mathematics from 1994 to 1997. After receiving his PhD in Applied Mathematics at Paris in 1998, he joined Poseidon Technologies research team that built the first computer vision technology based drowning detection system. His research interests include optical flow estimation, image registration, stereoscopy, scale space methods and real time image processing.

Laurent D. Cohen was born in 1962. He was student at the **Ecole Normale Supérieure**, rue d'Ulm in Paris, France from 1981

to 1985. He received the Master's and Ph.D. degrees in Applied Mathematics from University of Paris 6, France, in 1983 and 1986, respectively. From 1985 to 1987, he was member at the Computer Graphics and Image Processing group at Schlumberger Palo Alto Research, Palo Alto, California and Schlumberger Montrouge Research, Montrouge, France and remained consultant for a few years with Schlumberger afterwards. He began working with INRIA, France in 1988, mainly with the medical image understanding group Epidaure. Since 1990, he is Research Scholar with the French National Center for Scientific Research (CNRS) in the Applied Mathematics and Image Processing group at CEREMADE, University Paris-Dauphine, Paris, France. His research interests and teaching at the university are applications of variational methods and Partial Differential Equations to Image Processing and Computer Vision, like deformable models, minimal paths, surface reconstruction, Image registration, Image segmentation and restoration.

Proofs of Theorems

Proof of Lemma 1

Zero crossing the derivative of E_j with respect to h, we necessarily obtain:

$$\hat{h^j} = \frac{< \Pi_j(I_2'), \Pi_j(I_1 - I_2) >_{V_j}}{\|I_2'\|^2_{V_j}}.$$

After simplification, we find:

$$\hat{h^j} = \frac{M}{\pi} \frac{\sum_{|k| \leq 2^j} k |c_k(I_2)|^2 sin(\pi k h^*/M)}{\sum_{|k| \leq 2^j} k^2 |c_k(I_2)|^2}.$$

Writing the difference $|\hat{h^j} - h^*|$, the mean inequality yields:

$$|\hat{h^j} - h^*| \leq |h^*| \frac{\sum_{|k| \leq 2^j} k^2 |c_k(I_2)|^2 |\frac{sin(\pi k h^*/M)}{\pi k h^*/M} - 1|}{\sum_{|k| \leq 2^j} k^2 |c_k(I_2)|^2}.$$

But, for all x, we have $1 - \frac{x^2}{6} \leq \frac{sin(x)}{x} \leq 1$.
Finally, $|h^*| \leq \frac{M}{2^{j+1}}$ and $|k| \leq 2^j$ imply that

$$|\frac{sin(\pi k h^*/M)}{\pi k h^*/M} - 1| \leq \frac{1}{6} (\frac{\pi k h^*}{M})^2 \leq \frac{\pi^2}{24} \leq \frac{1}{2}.$$

Proof of Theorem 1

After L iterations, we have:

$$|\sum_{l=0}^{L} \hat{h_j^l} - h^*| \leq \frac{M}{2^{j+L+2}}.$$

The series $\hat{h_j^l}$ then converges towards h^*. Eventually, as

$$\begin{aligned} &\|I_1(x) - I_2(x + \sum_{l=0}^{L} \hat{h_j^l})\|_\infty \\ &= \|I_2(x + h^*) - I_2(x + \sum_{l=0}^{L} \hat{h_j^l})\|_\infty \\ &\leq \|I_2'\|_\infty |h^* - \sum_{l=0}^{L} \hat{h_j^l}| \leq \frac{M \|I_2'\|_\infty}{2^{j+L+2}}, \end{aligned}$$

we conclude that the functions sequence $I_{2,L}(x) = I_2(x + \sum_{l=0}^{L} \hat{h_j^l})$ uniformly converges towards u.

Proof after Lemma 2 in subsection 2.3

garanties that, for all i,

$$|\hat{h_i^{j,1}} + \hat{h_i^j} - h_i^*| \leq \frac{|\hat{h_i^j} - h_i^*|}{2} \leq \frac{|h_i^*|}{4} \leq \frac{M}{2^{j+1+2}}.$$

Let $\hat{h^{j,0}} = \hat{h^j}$. We iterate, and, at step L, we have:

$$|\sum_{l=0}^{L} \hat{h_i^{j,l}} - h_i^*| \leq \frac{M}{2^{j+L+2}}.$$

Thus the sequence of general term $\hat{h_i^{j,l}}$ converges towards h_i^*. Finally, as

$$\begin{array}{rc} & \|I_1(x) - I_1(x + \sum_{l=0}^{L} \hat{h^{j,l}})\|_\infty \\ = & \|I_1(x + h^*) - I_2(x + \sum_{l=0}^{L} \hat{h^{j,l}})\|_\infty \\ \leq & \|\nabla I_2\|_\infty |h^* - \sum_{l=0}^{L} \hat{h^{j,l}}| \\ \leq & \frac{M\sqrt{d}\|\nabla I_2\|_\infty}{2^{j+L+2}}, \end{array}$$

we can conclude that the sequence of functions $I_{2,L}(x) = I_2(x + \sum_{l=0}^{L} \hat{h^{j,l}})$ converges uniformly towards I_1.

Proof of Lemma 3

Applying Green formula yields:

$$\int_D I_1 \phi dx = \int_{\partial D} I_1 < \Phi, \vec{n} > dl - \int_D < \nabla I_1, \Phi > dx,$$

in which the boundary term disappears because we chose $I_1 \in C_0^1(D)$. The same holds for I_2 and yields $\int_D I_2 \phi dx = -\int_D < \nabla I_2, \Phi > dx$. Now changing the integration variable x by y with $x = f^*(y)$, we have $\int_D < \nabla I_2, \Phi > dx = \int_D < \nabla I_2 \circ f^*, \Phi \circ f^* > det(Jac(f^*))dy$. Using the rigidity hypothesis we remark that:

$$\begin{array}{rcl} \nabla I_2 \circ f^* & = & Jac^{-1}(f^*)^t.Jac(f^*)^t.\nabla I_2 \circ f^*, \\ & = & Jac^{-1}(f^*)^t.\nabla(I_2 \circ f^*), \\ & = & det(Jac(f^*))^{-1}\nabla I_1. \end{array}$$

We thus deduce that

$$\int_D I_2 \phi dx = -\int_D < \nabla I_1, \Phi \circ f^* > dx.$$

Proof of Theorem 2

For Π_σ and $P_\sigma^{I_1}$ project onto a truncated Fourier space with k indices belonging to S_σ, we have to bound the corresponding Fourier coefficients of $\Pi_\sigma(I_1 - I_2) - P_\sigma^{I_1}(h^*)$ and then sum over k indices in S_σ. The

case $k = 0$ is straightforward by Lemma 10:

$$c_0(I_1 - I_2 - \frac{1}{d} < \nabla I_1, h^* >) = 0.$$

Now take $k \in S_\sigma/\{0\}$. By Lemma 10 we have

$$\begin{aligned} & c_k(I_1 - I_2 - \frac{1}{|k|^2} < \nabla I_1, k >< k, h^* >) \\ = & c_k(< \nabla I_1, k > (e^{-i\pi<k,h^*>/M} - 1 + \frac{i\pi}{M} < k, h^* >)). \end{aligned}$$

Using Taylor-Lagrange inequality we have

$$\begin{aligned} & |e^{-i\pi<k,h^*>/M} - 1 + \frac{i\pi}{M} < k, h^* > | \\ & \leq \frac{\pi^2}{2M^2} | < k, h^* > |^2, \end{aligned}$$

and by the mean theorem we deduce that

$$\begin{aligned} & | \int_D < \nabla I_1, k > (e^{-i\pi<k,h^*>/M} - 1 + \frac{i\pi}{M} < k, h^* >) dx| \\ & \leq \frac{\pi^2}{2M^2} \int_D | < \nabla I_1, k > | < k, h^* >^2 dx, \end{aligned}$$

yielding the following bound:

$$\begin{aligned} & |c_k(I_1 - I_2 - \frac{1}{|k|^2} < \nabla I_1, k >< k, h^* >)| \\ & \leq \frac{\pi}{(2M)^{1+d/2}} \frac{1}{|k|} \int_D |\nabla I_1| < k, h^* >^2 dx. \end{aligned}$$

Summing the squared moduli of the coefficients over $k \in S_\sigma/\{0\}$ indices using the following separation:

$$\begin{aligned} & \sum_{k \in S_\sigma\{0\}} \frac{1}{|k|^2} (\int_D |\nabla I_1| < k, h^* >^2 dx)^2 \\ \leq & \sup_{k \in S_\sigma/\{0\}} \{\int_D |\nabla I_1|.|k|^2.|h^*|^2 dx\}. \\ & \sum_{k \in S_\sigma/\{0\}} \frac{1}{|k|^2} \int_D |\nabla I_1| < k.k^t h^*, h^* > dx, \\ \leq & \Big(\frac{\sigma^{d+2}}{2}\Big)^2 (2M)^{d+2} \Big(\int_D |\nabla I_1|.|h^*|^2 dx\Big)^2, \end{aligned}$$

leads to the final result. Though this bound is not optimal because of the choice of the separation and could be reduced by a factor depending essentially on the dimension, it is more tractable and clearly shows the scale dependence in any dimension.

Proof of Lemma 5

Without loss of generality we can assume that the set of basis functions is such that the $(|\nabla I_1|^{\frac{1}{2}}\psi_i)_{i=1..n}$ form an orthonormal basis of the spanned $(L^2)^d$ sub-space, thus yielding $\|(\hat{h}-h^*)|\nabla I_1|^{\frac{1}{2}}\|_{L^2} = \|\hat{\Theta}-\Theta^*\|_2$. Observing that M_σ is symmetric positive definite there exists a diagonal matrix D_σ and a unitary matrix U such that $M_\sigma = U^{-1}D_\sigma U$. Set $\Psi^1 = U\Psi$ and remark that $\|\hat{\Theta}-\Theta^*\|_2 = \|D_\sigma^{-1}D_\sigma U(\hat{\Theta}-\Theta^*)\|_2$. We see that:

$$\begin{aligned} D_\sigma U(\hat{\Theta}-\Theta^*) &= P_\sigma^{I_1}(\Psi^1)\otimes P_\sigma^{I_1}(\Psi^1))U(\hat{\Theta}-\Theta^*), \\ &= \tilde{B_\sigma}, \end{aligned}$$

where column $\tilde{B_\sigma}$'s components are defined by $(\tilde{B_\sigma})_i =< P_\sigma^{I_1}(\Psi_i^1), \Pi_\sigma(I_1 - I_2) - P_\sigma^{I_1}(h^*) >$. Applying Cauchy-Schwartz inequality and Theorem 2 we obtain:

$$\begin{aligned} &|\int_D P_\sigma^{I_1}(\Psi_i^1).\Big(\Pi_\sigma(I_1-I_2) - P_\sigma^{I_1}(h^*)\Big)dx| \\ &\leq \|P_\sigma^{I_1}(\Psi_i^1)\|_{L^2}.\|\Pi_\sigma(I_1-I_2) - P_\sigma^{I_1}(h^*)\|_{L^2}, \\ &\leq \|P_\sigma^{I_1}(\Psi_i^1)\|_{L^2}\frac{\pi}{2}\sigma^{d+2}\|h^*|\nabla I_1|^{\frac{1}{2}}\|_{L^2}^2. \end{aligned}$$

Observing that $(D_\sigma)_{(i,i)} = \|P_\sigma^{I_1}(\Psi_i^1)\|_{L^2}^2$ and using orthogonality between $P_\sigma^{I_1}(\Psi_i^1)$, we deduce that

$$\begin{aligned} &\|(\hat{h}-h^*)|\nabla I_1|^{\frac{1}{2}}\|_{L^2} \\ &= \|D_\sigma^{-1}D_\sigma U(\hat{\Theta}-\Theta^*)\|_2, \\ &\leq \Big(\sum_{i=1}^n \frac{\|P_\sigma^{I_1}(\Psi_i^1)\|_{L^2}^2}{(D_\sigma)_{(i,i)}^2}\Big)^{\frac{1}{2}}\frac{\pi}{2}\sigma^{d+2}\|h^*|\nabla I_1|^{\frac{1}{2}}\|_{L^2}^2, \\ &\leq \frac{\pi}{2}\sigma^{d+2}\Big(Tr(M_\sigma^{-1})\Big)^{\frac{1}{2}}\|h^*|\nabla I_1|^{\frac{1}{2}}\|_{L^2}^2. \end{aligned}$$

Proof of Theorem 3

Let us first remark that by the second hypothesis, we necessarily have $C_0 > 1$. It is thus sufficient to show that, for all $p \geq 1$,

$$\|r_p|\nabla I_{1,p}|^{\frac{1}{2}}\|_{L^2} \leq \frac{1}{C_0^p}\|h^*|\nabla I_{1,0}|^{\frac{1}{2}}\|_{L^2}.$$

If $p = 1$, we obtain by Lemma 5 and the second hypothesis:

$$\begin{aligned}\|r_1|\nabla I_{1,1}|^{\frac{1}{2}}\|_{L^2} &\leq \frac{\pi}{2}\sigma_0^{d+2}Tr(M_{0,\sigma_0})^{\frac{1}{2}}\|h^*|\nabla I_{1,0}|^{\frac{1}{2}}\|_{L^2}^2,\\ &\leq \frac{1}{C_0}\|h^*|\nabla I_{1,0}|^{\frac{1}{2}}\|_{L^2}.\end{aligned}$$

Assume by recurrence hypothesis that we have, for some $p > 1$,

$$\|r_p|\nabla I_{1,p}|^{\frac{1}{2}}\|_{L^2} \leq \frac{1}{C_0^p}\|h^*|\nabla I_{1,0}|^{\frac{1}{2}}\|_{L^2}.$$

Then by Lemma 5 we have again:

$$\|r_{p+1}|\nabla I_{1,p+1}|^{\frac{1}{2}}\|_{L^2} \leq \frac{\pi}{2}\sigma_p^{d+2}Tr(M_{p,\sigma_p})^{\frac{1}{2}}\|r_p|\nabla I_{1,p}|^{\frac{1}{2}}\|_{L^2}^2.$$

Now using the third hypothesis we obtain:

$$\begin{aligned}\frac{\pi}{2}\sigma_p^{d+2}Tr(M_{p,\sigma_p})^{\frac{1}{2}} &\leq C_0\frac{\pi}{2}\sigma_{p-1}^{d+2}Tr(M_{p-1,\sigma_{p-1}})^{\frac{1}{2}},\\ &\leq ...\\ &\leq C_0^p\frac{\pi}{2}\sigma_0^{d+2}Tr(M_{0,\sigma_0})^{\frac{1}{2}}.\end{aligned}$$

By definition of constant C_0 we have

$$C_0^p\frac{\pi}{2}\sigma_0^{d+2}Tr(M_{0,\sigma_0})^{\frac{1}{2}} \leq \frac{C_0^{p-1}}{\|h^*|\nabla I_{1,0}|^{\frac{1}{2}}\|_{L^2}},$$

and we deduce that:

$$\begin{aligned}\|r_{p+1}|\nabla I_{1,p+1}|^{\frac{1}{2}}\|_{L^2} &\leq \frac{C_0^{p-1}}{\|h^*|\nabla I_{1,0}|^{\frac{1}{2}}\|_{L^2}}\frac{\|h^*|\nabla I_{1,0}|^{\frac{1}{2}}\|_{L^2}^2}{C_0^{2p}},\\ &\leq \frac{1}{C_0^{p+1}}\|h^*|\nabla I_{1,0}|^{\frac{1}{2}}\|_{L^2}.\end{aligned}$$

Proof of section 5

Remarking that, for all $(x, \lambda) \in D \times I\!R$,

$$|\chi_{\{I_1 \geq \lambda\}}(x) - \chi_{\{I_2 \geq \lambda\}} \circ f(x)| \leq 1,$$

with

$$\lambda \notin [a, b] \Rightarrow |\chi_{\{I_1 \geq \lambda\}}(x) - \chi_{\{I_2 \geq \lambda\}} \circ f(x)| = 0,$$

we see that:

$$\int_{-\infty}^{+\infty} |\chi_{\{I_1 \geq \lambda\}}(x) - \chi_{\{I_2 \geq \lambda\}} \circ f(x)| d\lambda \leq (b-a).$$

We can deduce that:

$$\|I_1 - I_2 \circ f\|_{L^2}^2 \leq (b-a) \int_{-\infty}^{+\infty} \|\chi_{\{I_1 \geq \lambda\}} - \chi_{\{I_2 \geq \lambda\}} \circ f\|_1 d\lambda,$$

but in this case the L^1 norm is the squared L^2 norm,

$$\|I_1 - I_2 \circ f\|_{L^2}^2 \leq (b-a) \int_{-\infty}^{+\infty} \|\chi_{\{I_1 \geq \lambda\}} - \chi_{\{I_2 \geq \lambda\}} \circ f\|_{L^2}^2 d\lambda.$$