Regularization terms for motion estimation. Links with spatial correlations

Yann Lepoittevin, Isabelle Herlin

Regularization terms for motion estimation
Links with spatial correlations

Yann Lepoittevin
PhD student, Inria, Institut National de Recherche en Informatique et Automatique, Le Chesnay, France
Isabelle Herlin
Inria, Institut National de Recherche en Informatique et Automatique, Le Chesnay, France
Lepoittevin.Yann@gmail.com,Isabelle.Herlin@inria.fr

Keywords: Covariance Matrix, Image Assimilation, Motion, Tikhonov regularization.

Abstract: Motion estimation from image data has been widely studied in the literature. Due to the aperture problem, one equation with two unknowns, a Tikhonov regularization is usually applied, which constrains the estimated motion field. The paper demonstrates that the use of regularization functions is equivalent to the definition of correlations between pixels and the formulation of the corresponding correlation matrices is given. This equivalence allows to better understand the impact of the regularization with a display of the correlation values as images. Such equivalence is of major interest in the context of image assimilation as these methods are based on the minimization of errors that are correlated on the space-time domain. It also allows to characterize the role of the errors during the assimilation process.

1 INTRODUCTION

As well known and extensively discussed in the literature of image processing, motion estimation from image data is an ill-posed problem, according to the Hadamard definition (Hadamard, 1923). This comes from the fact that only one equation is available, the optical flow equation (Horn and Schunk, 1981), for estimating two unknown variables, the horizontal and vertical components, u and v, of the motion vector $\mathbf{w}$.

Smoothing of the motion field, according to the design of Tikhonov regularization terms (Tikhonov, 1963), is often used in the literature in order to get a unique solution, as seen for instance in the papers (Nagel and Enkelmann, 1986), (Nielsen et al., 1994) or more recently in (Werlberger et al., 2010). A huge literature is available on the subject. Survey papers on optical flow have been published, as for instance (Sun et al., 2010) and (Fortun et al., 2015).

An alternative to the Tikhonov regularization comes from the use of image assimilation methods, which include, in the estimation process, the available heuristics on the temporal evolution of the observed system. The reader can refer, for example, to the methods presented in (Papadakis et al., 2010), (Ridal et al., 2011) or in (Béréziat and Herlin, 2011). In the last few years, a number of such techniques were defined for various contexts of motion estimation from image sequences.

The data assimilation approach used in the paper is a 4D-Var method, based on the control theory. The foundational paper of Le Dimet and Talagrand (Le Dimet and Talagrand, 1986) describes the computation of the solution of a 4D-Var data assimilation algorithm, thanks to the adjoint method.

The 4D-Var image assimilation, which is applied in the paper, works as follows. Starting from a background value, a simulation model is integrated in time, producing a state vector value at each time step of the studied temporal interval. At each acquisition date, the state vector is compared to characteristics calculated on the image observations. For minimizing their difference on the whole temporal interval, the data assimilation method computes an optimal initial value, named the analysis vector. The whole temporal trajectory is then obtained by integrating the model from that analysis value. Section 2 describes the main mathematical components of the 4D-Var framework.

In order to estimate motion, the 4D-Var approach defines a cost function J, which is minimized for computing the result. This cost function is depending on the discrepancy between the state vector and the image data, or image characteristics, at acquisition dates. Regularization terms are often added to that cost function, as described by Béréziat et al. in (Béréziat and Herlin, 2011), in order to determine the vectorial sub-

space on which motion is estimated. These regularization terms ensure that, during the minimization process, the motion field keeps the chosen regularity properties, which are based on the available knowledge of the observed system.

Three types of regularization are analyzed in the paper and given in Section 2, simultaneously with a discussion of their impact on the estimated motion field. They concern the gradient of motion, its divergence and its norm.

The core of the paper concerns an extensive discussion on the interpretation of these three regularization terms as correlations between pixels of the image domain. Some related work has been done, for instance, by Dean S. Oliver in (Oliver, 1998), where the regularization terms are associated to the inverse of a covariance matrix. Section 3 demonstrates, in the context of 4D-Var image assimilation, that the estimation result obtained with the regularization terms and no correlation between pixels is exactly the same than the one obtained with specific correlations. The correlation matrices corresponding to the three types of regularization, on the gradient of motion, on its divergence and on its norm, will be given and discussed. The equivalence between regularization and correlation allows to visualize their joint impact on the estimation and to get insights on the choice of the parameters values, which weight these regularization terms. The displays of correlation matrices are also given in Section 3.

1.1 Notations

In the remaining of this introduction, the main mathematical notations used in the paper are given.

- Let Ω be an open subset of $\mathbb{R}^2$. Ω is the image domain on which motion is estimated.
- Let $[0,T]$ be a closed subset of $\mathbb{R}$, corresponding to the time interval on which image acquisitions are available.
- $\Omega_T = \Omega \times [0,T]$ defines the studied spatio-temporal interval, on which image assimilation is applied.
- A point of the image domain Ω is denoted by:

$$\mathbf{x} = \begin{pmatrix} x & y \end{pmatrix}^T \tag{1}$$

  with x and y corresponding respectively to the abscissa and the ordinate, in a Cartesian system defined on Ω.
- Let $\mathbf{w}$ denote the motion function, defined on Ω_T, such that:

$$\mathbf{w}(\mathbf{x},t) = \begin{pmatrix} u(\mathbf{x},t) & v(\mathbf{x},t) \end{pmatrix}^T \tag{2}$$

  with u and v quantifying respectively the values of motion along the abscissa and the ordinate.
- An image function I is defined on Ω_T, with the same physical properties as the image acquisitions. I is supposed to be transported by the motion function $\mathbf{w}$. Consequently, this image function corresponds to a passive tracer of the motion function.
- Let introduce the notation $\mathbf{X}$, denoting the state vector of the observed system, depending on $\mathbf{x}$ and t and defined on Ω_T by:

$$\mathbf{X}(\mathbf{x},t) = \begin{pmatrix} \mathbf{w}(\mathbf{x},t) & I(\mathbf{x},t) \end{pmatrix}^T \tag{3}$$

- The image function I and both components, u and v, of the motion field $\mathbf{w}$ are defined on Ω_T. For sake of simplicity, we denote u the space-time function, $u(t)$ the field at date t and $u(\mathbf{x},t)$ the value at pixel $\mathbf{x}$ and date t of the image domain Ω. The same rule is applied for all functions defined on Ω_T.
- Data assimilation methods are functioning by comparing a model output with observed values of the studied system. The observation vector $\mathbf{Y}$ is defined on Ω_T. Its value at date t and point $\mathbf{x}$ is $\mathbf{Y}(\mathbf{x},t)$. Its components correspond to image acquisitions or to image features computed on these acquisitions. They are denoted by using the superscript $\cdot^O$. For instance, the image acquisition is denoted I^O. $I^O(\mathbf{x},t)$ is the value at date t and point $\mathbf{x}$.
- When describing a data assimilation method, projection operators are needed that are denoted $\mathbb{P}$. For instance, $\mathbb{P}_{\mathbf{w}}$ is the projection from the space of the state vector on the space of the motion fields.
- When defining the formulation of the optimal estimation, error terms, denoted ε, are needed. These error terms will be considered as Gaussian and zero-mean. They are therefore described by a covariance function. The covariance function of the error term denoted ε_B is B.
- For describing the implementation, the image domain Ω is discretized but is still denoted with the same symbol, for sake of simplicity. In the same spirit, $\mathbf{x}$ denotes either the point of the continuous domain or the pixel of the discrete domain, with indexes i and j. The same rule is applied for all quantities, $\mathbf{X}$, $\mathbf{Y}$, u, v, I, ...
  The image domain is composed of N_Ω pixels.
  The state vector $\mathbf{X}$ has $N_{\mathbf{X}} = 3N_\Omega$ components, as it includes the value of motion and image for each pixel.

The vector u has N_Ω components, which are the values of u at all pixels. The same goes for v and I.
The discrete observation vector $\mathbf{Y}$ has $N_\mathbf{Y}$ components.
The notation t of the continuous time variable is also kept for the discrete time index.

2 MOTION ESTIMATION AND REGULARIZATION TERMS

This section summarizes the issue of motion estimation, based on data assimilation methods, as described for instance in (Béréziat and Herlin, 2011).

The first element to be defined is the Image Model, expressing the heuristics on the observed system and on the image acquisitions. The design of this Image Model depends on the duration of the studied temporal interval. On a short term, the motion field is usually considered as stationary, which is mathematically written as:

$$\frac{\partial \mathbf{w}}{\partial t} = 0 \tag{4}$$

Such simple evolution law has a great potential for operational applications, as no temporal integration of the motion field is required: $\mathbf{w}(t) = \mathbf{w}(0)$ for each value of t. On a longer duration, this assumption is no more valid and has to be released. In this paper, motion is considered as advected by itself. This is written as:

$$\frac{\partial \mathbf{w}}{\partial t} + (\mathbf{w} \cdot \nabla)\mathbf{w} = 0 \tag{5}$$

It corresponds to the Lagrangian conservation of motion on the whole trajectory:

$$\frac{d\mathbf{w}}{dt}(\mathbf{x},t) = 0 \tag{6}$$

Expressing the motion field $\mathbf{w}$ with its two components u and v,

$$\mathbf{w} = \begin{pmatrix} u & v \end{pmatrix}^T, \tag{7}$$

allows to decompose Equation (5) with two partial differential equations:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = 0 \tag{8}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = 0 \tag{9}$$

Considering the hypothesis that the image brightness is a physical property, which is preserved over time accordingly to the displacement of objects on the image domain, leads to:

$$I(\mathbf{x},t) = I(\mathbf{x}+\delta\mathbf{x}, t+\delta t) \tag{10}$$

Assuming that the displacement $\delta\mathbf{x}$ and the time interval δt are small, Equation (10) is developed, accordingly to Taylor series, into:

$$I(\mathbf{x}+\delta\mathbf{x}, t+\delta t) = I(\mathbf{x},t) + \delta\mathbf{x}\frac{\partial I}{\partial \mathbf{x}} + \delta t\frac{\partial I}{\partial t} + \dots \tag{11}$$

From Equations (10) and (11), it comes:

$$\frac{\partial I}{\partial t} \approx -\frac{\delta\mathbf{x}}{\delta t}\frac{\partial I}{\partial \mathbf{x}} \tag{12}$$

Therefore, the image brightness is considered transported by the motion field, which conducts to the optical flow equation:

$$\frac{\partial I}{\partial t} + \mathbf{w} \cdot \nabla I = 0 \tag{13}$$

The image assimilation approach, which estimates $\mathbf{X}$ with a 4D-Var algorithm, is then based on the following system of three equations:

$$\frac{\partial \mathbf{X}}{\partial t}(\mathbf{x},t) + \mathbb{M}(\mathbf{X})(\mathbf{x},t) = 0 \tag{14}$$

$$\mathbf{X}(\mathbf{x},0) = \mathbf{X}^{(b)}(\mathbf{x}) + \varepsilon_B(\mathbf{x}) \tag{15}$$

$$\mathbb{H}(\mathbf{X},\mathbf{Y})(\mathbf{x},t) = \varepsilon_R(\mathbf{x},t) \tag{16}$$

Equation (14) is the partial differential equation ruling the temporal evolution of $\mathbf{X}(\mathbf{x},t)$. This equation comes either from Equations (4, 13) or (5, 13). The value $\mathbf{X}(\mathbf{x},t)$ is determined, for any date t, from the initial value $\mathbf{X}(\mathbf{x},0)$ and the temporal integration of the model $\mathbb{M}$.

Equation (15) expresses the a priori knowledge, named the background value and denoted $\mathbf{X}^{(b)}(\mathbf{x})$, that is available on the state vector at initial date 0. An error term, $\varepsilon_B(\mathbf{x})$, is added in order to express the uncertainty on this a priori knowledge. This error term is supposed to be Gaussian and zero-mean, with the covariance function denoted B. The choice of the background value is depending on the experiment that is conducted and is described together with the studied images. However, as the objective is to estimate the motion field from the image data, no constraint will be applied for ensuring that the result stay close to the background value of motion. This background motion field is only used as a starting point for the iterative minimization process. The background of the image function I is generally taken as the first acquisition of the sequence. Equation (15) is then equivalent to:

$$I(\mathbf{x},0) - I^{(b)}(\mathbf{x}) = \varepsilon_B(\mathbf{x}) \tag{17}$$

where the symbol $\varepsilon_B(\mathbf{x})$ is now used to indicate the zero-mean Gaussian error on the image component, associated to its covariance function B. We assume that the image error is uncorrelated in space. Therefore B is a diagonal matrix, whose diagonal values are denoted σ^2.

Equation (16) is the observation equation that links the values of the image acquisitions to the state vector, at each date of the studied interval. In the paper, the observation vector $\mathbf{Y}(\mathbf{x},t)$ is equal to the observed image $I^O(\mathbf{x},t)$ and $\mathbb{H}$ allows to compare the image function $I(\mathbf{x},t)$ to the image acquisitions $I^O(\mathbf{x},t)$. The operator $\mathbb{H}$ is then defined by:

$$\mathbb{H}(\mathbf{X},\mathbf{Y})(\mathbf{x},t) = I(\mathbf{x},t) - I^O(\mathbf{x},t) \tag{18}$$

The observation equation, Equation (16), then rewrites:

$$I(\mathbf{x},t) - I^O(\mathbf{x},t) = \varepsilon_R(\mathbf{x},t) \tag{19}$$

The discrepancy between the image function $I(\mathbf{x},t)$ and the acquisition $I^O(\mathbf{x},t)$ is described by the error $\varepsilon_R(\mathbf{x},t)$ on the image component. $\varepsilon_R(\mathbf{x},t)$ represents both the acquisition and representativity errors. This error term is also supposed Gaussian, zero-mean and uncorrelated with ε_B. In the paper, the covariance function R, associated to $\varepsilon_R(\mathbf{x},t)$, considers no covariance between two locations. Therefore R is a diagonal matrix, whose diagonal values are also taken equal to σ^2.

Solving System (14, 17, 19) is equivalent with the minimization of the error terms ε_R and ε_B. This is obtained by designing a cost function J, depending on the control variable $\mathbf{X}(0)$, with the following formulation (where the space variable $\mathbf{x}$ is suppressed for sake of clarity):

$$J(\mathbf{X}(0)) = \int_\Omega \left(I(0) - I^{(b)}\right) B(\mathbf{x})^{-1} \left(I(0) - I^{(b)}\right) + \int_{\Omega_T} \left(I(t) - I^O\right) R^{-1} \left(I(t) - I^O\right) \tag{20}$$

Three regularization terms $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$ are added to the cost function of Equation (20). This new cost function, still denoted J, is minimized during the data assimilation process, which estimates $\mathbf{X}(0)$ and its motion component.

The first regularization term, named $\mathcal{R}_1$, acts on the norm of the gradient of the motion field. It is designed as follows:

$$\mathcal{R}_1(\mathbf{X}(0)) = \alpha \int_\Omega ||\nabla(\mathbb{P}_\mathbf{w}(\mathbf{X}(\mathbf{x},0)))||^2 d\mathbf{x} \tag{21}$$

or equivalently:

$$\mathcal{R}_1(\mathbf{X}(0)) = \alpha \int_\Omega ||\nabla \mathbf{w}(\mathbf{x},0)||^2 d\mathbf{x} \tag{22}$$

$\mathcal{R}_1$ ensures the spatial smoothness of the estimation. It is weighted by the parameter α.

When working on the issue of sea surface circulation, the estimated motion field should be divergence free, due to the incompressibility property. In other applications, even if the divergence is non null, its value should be small as aliasing effects could appear, during the temporal integration of the image model, if the divergence is high. A second regularization term $\mathcal{R}_2$, acting on the divergence, is then added to the cost function J:

$$\mathcal{R}_2(\mathbf{X}(0)) = \beta \int_\Omega [\mathrm{div}(\mathbb{P}_\mathbf{w}(\mathbf{X}(\mathbf{x},0)))]^2 d\mathbf{x} \tag{23}$$

or equivalently:

$$\mathcal{R}_2(\mathbf{X}(0)) = \beta \int_\Omega [\mathrm{div}(\mathbf{w}(\mathbf{x},0))]^2 d\mathbf{x} \tag{24}$$

where:

$$\mathrm{div}(\mathbf{w}(\mathbf{x},0)) = \frac{\partial u}{\partial x}(\mathbf{x},0) + \frac{\partial v}{\partial y}(\mathbf{x},0) \tag{25}$$

A regularization term acting on the norm of the motion field is also included in the function J, in order to avoid having spurious high values of $\mathbf{w}$. This term $\mathcal{R}_3$ is defined by:

$$\mathcal{R}_3(\mathbf{X}(0)) = \gamma \int_\Omega ||\mathbb{P}_\mathbf{w}(\mathbf{X}(\mathbf{x},0))||^2 d\mathbf{x} \tag{26}$$

or equivalently:

$$\mathcal{R}_3(\mathbf{X}(0)) = \gamma \int_\Omega ||\mathbf{w}(\mathbf{x},0)||^2 d\mathbf{x} \tag{27}$$

Let sum up these three regularization into a global term $\mathcal{R}$, defined as:

$$\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2 + \mathcal{R}_3 \tag{28}$$

and depending on the initial value $\mathbf{X}(0)$ through its motion component $\mathbf{w}(0)$.

Having defined the regularization terms, the next section will discuss and illustrate their significance and action during the estimation process.

3 SIGNIFICANCE OF THE REGULARIZATION

The regularization terms, which are included in the cost function are given, with a variational formulation, in Equations (22, 24, 27). Keeping in mind that $\mathbf{w} = \begin{pmatrix} u \\ v \end{pmatrix}$, it is possible to rewrite the formulation of $\mathcal{R}_1$ Equation (22), as:

$$\mathcal{R}_1(\mathbf{X}(0)) = \alpha \int_\Omega \frac{\partial u}{\partial x}(\mathbf{x},0)^2 + \frac{\partial u}{\partial y}(\mathbf{x},0)^2 + \frac{\partial v}{\partial x}(\mathbf{x},0)^2 + \frac{\partial v}{\partial y}(\mathbf{x},0)^2 \tag{29}$$

The formulation of $\mathcal{R}_2$ in Equation (24) is rewritten as:

$$\mathcal{R}_2(\mathbf{X}(0)) = \beta \int_\Omega \left(\frac{\partial u}{\partial x}(\mathbf{x},0) + \frac{\partial v}{\partial y}(\mathbf{x},0) \right)^2 \quad (30)$$

Last, $\mathcal{R}_3$ of Equation (27) is equal to:

$$\mathcal{R}_3(\mathbf{X}(0)) = \gamma \int_\Omega u(\mathbf{x},0)^2 + v(\mathbf{x},0)^2 \quad (31)$$

When implementing the method on a discrete the image domain Ω, the derivatives along x and y are computed by filters $\mathcal{D}_x$ and $\mathcal{D}_y$, whose values depend on the chosen discretization schemes. If the derivatives are, for instance, approximated with a forward scheme, the filter $\mathcal{D}_x$ is defined by:

$$\mathcal{D}_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{-1}{dx} & \frac{1}{dx} \\ 0 & 0 & 0 \end{bmatrix} \quad (32)$$

and the filter $\mathcal{D}_y$ by:

$$\mathcal{D}_y = \begin{bmatrix} 0 & \frac{1}{dy} & 0 \\ 0 & \frac{-1}{dy} & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (33)$$

The derivative filters being applied on the whole domain, let introduce the matrices D_x and D_y, which compute the discrete derivatives at every pixel, respectively along the directions x and y. By definition, D_x and D_y are Toeplitz matrices and their coefficients along descending diagonals are constant. For instance, D_x has the value $\frac{-1}{dx}$ on its main diagonal and the value $\frac{1}{dx}$ on the first above diagonal:

$$D_x = \frac{1}{dx} \begin{bmatrix} -1 & 1 & & & & & \\ & -1 & 1 & & & & \\ & & -1 & 1 & & & \\ & & & \ddots & \ddots & & \\ & & & & -1 & 1 & \\ & & & & & -1 \end{bmatrix} \quad (34)$$

It is then possible to rewrite the discrete formulation of each regularization term from these notations.
The discrete version of Equation (29) is:

$$\begin{aligned} \mathcal{R}_1(\mathbf{X}(0)) = \alpha\big(& \langle D_x u \ , \ D_x u \rangle + \langle D_y u \ , \ D_y u \rangle \\ & + \langle D_x v \ , \ D_x v \rangle + \langle D_y v \ , \ D_y v \rangle \big) \end{aligned} \quad (35)$$

where $\langle f_1 \ , \ f_2 \rangle$ denotes the scalar product of the vectors v_1 and v_2. Equation (30) leads to:

$$\mathcal{R}_2(\mathbf{X}(0)) = \beta \langle D_x u + D_y v \ , \ D_x u + D_y v \rangle \quad (36)$$

Equation (31) is discretized by:

$$\mathcal{R}_3(\mathbf{X}(0)) = \gamma\big(\langle u \ , \ u \rangle + \langle v \ , \ v \rangle \big) \quad (37)$$

Let introduce the vector $\begin{pmatrix} u \\ v \end{pmatrix}$ of size $2\mathrm{N}_\Omega$ in the previous scalar products. Let also use the fact that D_x and D_y being matrices with real coefficients, their adjoint is equal to their transpose. Let furthermore use the bilinearity of the scalar product. These three points lead to rewrite the discrete formulation of $\mathcal{R}_1$ in Equation (35) as:

$$\left\langle \begin{pmatrix} u \\ v \end{pmatrix} \ , \ \alpha \begin{pmatrix} K & 0 \\ 0 & K \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \right\rangle \quad (38)$$

with K being defined by:

$$K = D_x^T D_x + D_y^T D_y \quad (39)$$

The discrete formulation of $\mathcal{R}_2$, in Equation (36), leads to:

$$\left\langle \begin{pmatrix} u \\ v \end{pmatrix} \ , \ \beta \begin{pmatrix} D_x^T D_x & D_x^T D_y \\ D_y^T D_x & D_y^T D_y \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \right\rangle \quad (40)$$

Last, the formulation of $\mathcal{R}_3$, in Equation (37), becomes:

$$\left\langle \begin{pmatrix} u \\ v \end{pmatrix} \ , \ \gamma \begin{pmatrix} \mathbb{I} & 0 \\ 0 & \mathbb{I} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \right\rangle \quad (41)$$

where $\mathbb{I}$ is the identity matrix.

Let denote C_1 the matrix involved in the computation of $\mathcal{R}_1$, Equation (38):

$$C_1 = \alpha \begin{pmatrix} K & 0 \\ 0 & K \end{pmatrix} \quad (42)$$

Let denote C_2 the matrix involved in Equation (40):

$$C_2 = \beta \begin{pmatrix} D_x^T D_x & D_x^T D_y \\ D_y^T D_x & D_y^T D_y \end{pmatrix} \quad (43)$$

Let denote C_3 the matrix obtained in Equation (41):

$$C_3 = \gamma \begin{pmatrix} \mathbb{I} & 0 \\ 0 & \mathbb{I} \end{pmatrix} \quad (44)$$

Last, let define the matrix C:

$$\begin{aligned} C &= C_1 + C_2 + C_3 \\ &= \begin{pmatrix} \alpha K + \beta D_x^T D_x + \gamma \mathbb{I} & \beta D_x^T D_y \\ \beta D_y^T D_x & \alpha K + \beta D_y^T D_y + \gamma \mathbb{I} \end{pmatrix} \end{aligned} \quad (45)$$

The definition of C leads to the following equality for the regularization term involved in the cost function:

$$\mathcal{R}(X(0)) = \left\langle \begin{pmatrix} u \\ v \end{pmatrix} \ , \ C \begin{pmatrix} u \\ v \end{pmatrix} \right\rangle \quad (46)$$

As $\mathcal{R}_1$ (in Equation (22)), $\mathcal{R}_2$ (in Equation (24)) and $\mathcal{R}_3$ (in qEquation (27)) are positive or null, as long as α, β and γ are positive, the regularization value expressed in Equation (28) is also positive or null. Moreover, $\mathcal{R}(X(0))$ is null if and only if $\mathbf{w}(0)$ is null. As both formulation of the regularization,

Equations (28) and (46) are equivalent, the matrix C is symmetric definite positive and can be considered as the inverse of a covariance matrix $B_{\mathcal{R}}$.

It comes that the two following formulations of the discrete cost function (where the space and time indexes are suppressed for sake of clarity) are equivalent:

$$\left\langle I(0)-I^{(b)}\ ,\ B^{-1}\left(I(0)-I^{(b)}\right)\right\rangle + \sum_{[0,T]}\left\langle I-I^{O}\ ,\ R^{-1}\left(I-I^{O}\right)\right\rangle+\mathcal{R}(\mathbf{X}(0)) \quad (47)$$

and:

$$\left\langle \mathbf{X}(0)-\mathbf{X}^{(b)}\ ,\ B_{\mathcal{R}}^{-1}\left(\mathbf{X}(0)-\mathbf{X}^{(b)}\right)\right\rangle + \sum_{[0,T]}\left\langle I-I^{O}\ ,\ R^{-1}\left(I-I^{O}\right)\right\rangle \quad (48)$$

where the new covariance matrix $B_{\mathcal{R}}$ verifies:

$$B_{\mathcal{R}}^{-1}=\begin{pmatrix} C & 0 \\ 0 & B^{-1}\end{pmatrix} \quad (49)$$

It should be noted that, in Equation (47), only the image component $I^{(b)}$ of $\mathbf{X}^{(b)}$ is involved and is chosen equal to the first image acquisition $I^{O}(0)$. On another hand, in Equation (48), the whole background vector is involved and defined by:

$$\mathbf{X}^{(b)} \quad = \quad \begin{pmatrix} 0 \\ 0 \\ I(0)\end{pmatrix} \quad (50)$$

where the image component is the same and the motion component is given a null value, which provides the heuristic of smoothness for the motion field. The error covariance matrix B of the image background keeps the same value from Equation (47) to Equation (49).

This concludes the demonstration that the use of regularization terms is equivalent to the use of a non diagonal covariance matrix $B_{\mathcal{R}}$ in the cost function minimized for estimating motion.

When implementing the image assimilation method, the state vector $\mathbf{X}(0)$ is composed of the three components $u(0)$, $v(0)$, and $I(0)$. Each of these components is defined on the discrete image domain Ω, composed of N_Ω pixels. Therefore, $\mathbf{X}(0)$ has $3N_\Omega$ components. The size of the covariance matrix B is equal to the square of the size of the state vector. This would lead to unaffordable memory costs if one wants to store the whole matrix. For instance, for a 100×100 pixels image, this leads to a 54 gigabytes matrix. However, the inverse matrix designed in Equation (49) is sparse and contains a high number of zero values. A sparse storage of this covariance matrix is feasible, but would lead to high computational costs when performing the matrix inversion or the product of the matrix by a vector, for instance in Equation (48). Therefore, the solution of minimizing Equation (48) by designing the covariance matrix $B_{\mathcal{R}}$ and inverting it is not considered for the operational use of the image assimilation method.

Let however remark that $B_{\mathcal{R}}$ is not required for computing the cost function with Equation (48) but only its inverse $B_{\mathcal{R}}^{-1}$ is. As the blocks included in $B_{\mathcal{R}}^{-1}$ are Toeplitz matrices, the best way to compute the value of the cost function J with Equation (48) is to consider each block of $B_{\mathcal{R}}^{-1}$ as a discrete filter. Let first remark that the filter associated to B^{-1} is defined by:

$$\mathcal{B}^{-1}=\begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\sigma^2} & 0 \\ 0 & 0 & 0\end{bmatrix} \quad (51)$$

For further illustrating the discussion, let consider that the derivatives are computed with forward schemes, which are determined by the following convolution filters:

$$\mathcal{D}x=\begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{-1}{dx} & \frac{1}{dx} \\ 0 & 0 & 0\end{bmatrix},\ \mathcal{D}y=\begin{bmatrix} 0 & \frac{1}{dy} & 0 \\ 0 & \frac{-1}{dy} & 0 \\ 0 & 0 & 0\end{bmatrix} \quad (52)$$

Let denote $B_{\mathcal{R}_{i,j}}^{-1}$ the bloc on the i^{th} line and j^{th} column of $B_{\mathcal{R}}^{-1}$, as it is written in Equation (49), considering the definition of C given by Equations (45) and (39). Let denote $\mathcal{B}_{\mathcal{R}_{i,j}}^{-1}$ the corresponding convolution filter. Using the mathematical rules for addition and composition of filters, it comes:

$$\mathcal{B}_{\mathcal{R}_{1,1}}^{-1} \quad =\begin{bmatrix} 0 & -\alpha & 0 \\ -(\alpha+\beta) & L_1 & -(\alpha+\beta) \\ 0 & -\alpha & 0\end{bmatrix} \quad (53)$$

where:

$$L_1=2\left(\frac{\alpha+\beta}{dx^2}+\frac{\beta}{dy^2}\right)+\gamma \quad (54)$$

$$\mathcal{B}_{\mathcal{R}_{2,2}}^{-1} \quad =\begin{bmatrix} 0 & -(\alpha+\beta) & 0 \\ -\alpha & L_2 & -\alpha \\ 0 & -(\alpha+\beta) & 0\end{bmatrix} \quad (55)$$

where:

$$L_2=2\left(\frac{\beta}{dx^2}+\frac{\alpha+\beta}{dy^2}\right)+\gamma \quad (56)$$

$$\mathcal{B}^{-1}_{\mathcal{R}_{1,2}} = \begin{bmatrix} \frac{\beta}{dxdy} & \frac{-\beta}{dxdy} & 0 \\ \frac{-\beta}{dxdy} & \frac{\beta}{dxdy} & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (57)$$

$$\mathcal{B}^{-1}_{\mathcal{R}_{2,1}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{\beta}{dxdy} & \frac{-\beta}{dxdy} \\ 0 & \frac{-\beta}{dxdy} & \frac{\beta}{dxdy} \end{bmatrix} \quad (58)$$

The use of $B^{-1}_{\mathcal{R}}$ during the computation of J with Equation (48) is replaced by the use of the four previous filters. The design of this non diagonal matrix $B_{\mathcal{R}}$ is equivalent, as demonstrated above, to apply the regularization $\mathcal{R}$ to the state vector. However, the covariance method has the advantage, compared to the regularization method, that the derivatives of the regularization functions defined by Equations (22, 24, 27) are no more required during the minimization. Moreover, the filters included in the matrix $B^{-1}_{\mathcal{R}}$, Equations (49, 45, 39), are applied both in the forward integration, computing the cost function J of Equation (48), and in the backward integration, which computes the gradient $\frac{dJ}{d\mathbf{X}(0)}$:

$$\frac{dJ}{d\mathbf{X}(0)} = 2B^{-1}_{\mathcal{R}}\left(\mathbf{X}(0) - \mathbf{X}^{(b)}\right) + \lambda(0) \quad (59)$$

Studying the values of the covariance matrix $B_{\mathcal{R}}$, corresponding to the values of the coefficients α, β and γ is a tool for better understanding the impact of the regularization $\mathcal{R}$ on the estimation. For doing this, it is first required to invert the matrix $B^{-1}_{\mathcal{R}}$, defined in Equations (49, 45, 39), in order to obtain the covariance matrix $B_{\mathcal{R}}$. This can not be done in operational use, due to the large size of the involved state vectors (3 times the size of the image domain). Moreover, it has no interest apart having a complete knowledge of the links imposed between variables of the state vector and between pixels of the spatial domain. However, when designing an operational use of motion estimation, this allows visualizing and understanding how the regularization terms act on the estimation results. This can be applied, during a learning phase for calibrating the operational use, on small sub-windows on the whole image domain as explained in the following.

For being able to easily compute the inverse of the matrix $B^{-1}_{\mathcal{R}}$, we consider a small size sub-image of 35×35 pixels. One can extract the $\mathbf{x}^{th}$ line of the covariance matrix. It corresponds to the covariance values of that pixel $\mathbf{x}$ with all other pixels of the domain. In the following, we focus on the visualization of the covariances in $B_{\mathcal{R}_{11}}$, as they involve the three regularization terms and the three parameters α, β and γ as visible in Equations (53) and (54). $B_{\mathcal{R}_{22}}$ is a rotated version of $B_{\mathcal{R}_{11}}$ and would lead to a redundant visualization. $B_{\mathcal{R}_{12}}$ and $B_{\mathcal{R}_{21}}$ are only depending on the term $\mathcal{R}_2$ and on the parameter β (see Equation (55) and (56)). Their visualization would not allow to improve the understanding of the joint effect of the three regularization terms.

The term $\mathcal{R}_3$, regularizing the norm of the motion field with the parameter γ, acts on the individual variance and does not add any correlation between pixels. Varying the two terms $\mathcal{R}_1$, regularizing the gradient norm of $\mathbf{w}$ with the parameter α, and $\mathcal{R}_2$, regularizing the divergence of $\mathbf{w}$ with the parameter β, allows to display the covariance between a reference point and the rest of the domain. The state vector is composed of the three fields corresponding to the values of u, v and I at all locations. The covariance matrix associated to each field may be displayed as an image.

Figure 1 gives the covariance of the component u of pixel $(17, 17)$ with the rest of the sub-image. On the left, the coefficient of $\mathcal{R}_1$ is preponderant. In the middle, $\mathcal{R}_1$ and $\mathcal{R}_2$ have the same importance in the computation. On the right, $\mathcal{R}_2$ is preponderant. It can be

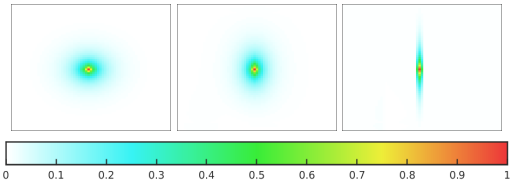

Figure 1: Covariance values associated to the central point *(red pixel)*; when $\mathcal{R}_1$ is preponderant *(on the left)*; when $\mathcal{R}_1$ and $\mathcal{R}_2$ are of same weight *(in the middle)*; and when $\mathcal{R}_2$ is preponderant *(on the right)*.

seen that $\mathcal{R}_1$ mimics an homogeneous diffusion process. On another hand, $\mathcal{R}_2$ favors specific directions for creating vortices and limiting the divergence of the motion field.

The range of the covariance values is parametrized by the values of α and β. This is, first, illustrated on Figure 2, which displays the covariance values associated to the regularization term $\mathcal{R}_1$, according to a small α, on the left, and a higher one, on the right. Similarly, Figure 3 shows the covariance values associated to the regularization term $\mathcal{R}_2$. On the left image, a small value of β is used, whereas the right image shows the covariance values for a higher β.

It can be seen, by analyzing Figure 2 and Figure 3, that the region of high covariance increases with the value of the regularization parameters α and β. Displaying a number of such images should help, for a given, application, to define the parameters values ac-

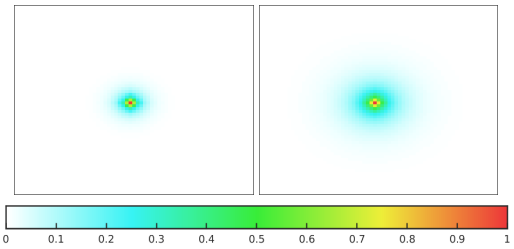

Figure 2: Covariance values associated to the central point *(red pixel)* for $\mathcal{R}_1$ only; with a small α *(on the left)* and a higher one *(on the right)*.

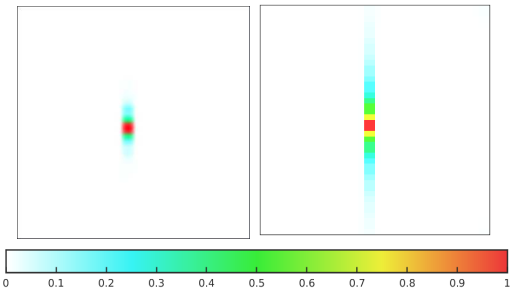

Figure 3: Covariance values associated to the central point *(red pixel)* for $\mathcal{R}_2$ only; with a small β *(on the left)* and a higher one *(on the right)*.

cording to the size of the structures to be found on the images.

For visualizing the impact of the parameters on the estimation, several motion results are given. On Figure 4, the estimation is computed on traffic data from the database KOGS/IAKS of the Karlsruhe University (Nagel, 1995). Motion is displayed with the color code of the Middlebury data base described by Baker et al. (Baker et al., 2011). The top image displays the result obtained with only $\mathcal{R}_3$. It can be seen that the motion field is irregular even if its whole shape is overall well recovered. The second image displays the estimation result when adding $\mathcal{R}_1$, with a small value of α (this corresponds to the left image of Figure 2). The estimation is smoother than the one on the top image, but irregularities remain. The bottom image is obtained with a large α value (this corresponds to the right image of Figure 2). The estimation is smooth without any irregularities.

A sequence of Sea Surface Temperature (SST) is processed as another illustration of the impact on the estimation of different parametrizations. Some images of the sequence are displayed on Figure 5. Result are shown on Figure 6, where the estimation is either obtained with $\mathcal{R}_1$ and $\mathcal{R}_3$ or $\mathcal{R}_2$ and $\mathcal{R}_3$. It can bee seen, from the left image of Figure 6, that $\mathcal{R}_1$ tends to favor smooth and homogene motion fields. In the contrary, as seen from the right image of Figure 6, $\mathcal{R}_2$

Figure 4: From top to bottom: Motion result with $\mathcal{R}_3$ only; $\mathcal{R}_3$ and $\mathcal{R}_1$ and a small value of α; $\mathcal{R}_3$ and $\mathcal{R}_1$ and a high value of α

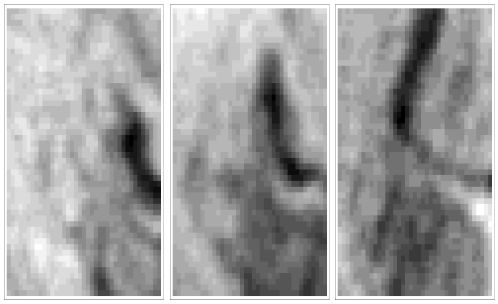

Figure 5: Sequence of Sea Surface Temperature images.

favors gyral structures to explain the temporal evolution of the gray level values.

4 CONCLUSIONS

The paper discusses the mathematical links between the Tikhonov regularization terms and the spatial covariances applied between pixels. The application concerns the issue of motion estimation, which is

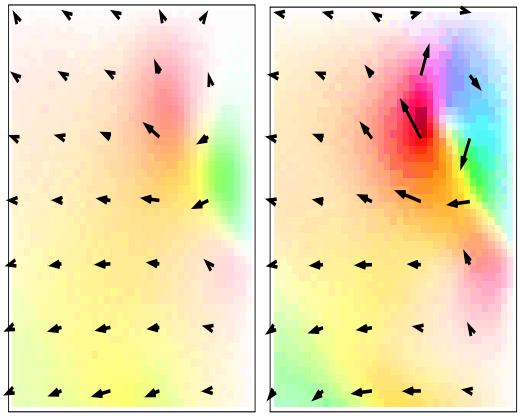

Figure 6: Estimation results obtained on the SST sequence. Left: Motion result with $\mathcal{R}_1$ and $\mathcal{R}_3$. Right: Motion result with $\mathcal{R}_2$ and $\mathcal{R}_3$.

an ill-posed problem that is often solved by adding regularization terms in a cost function. In the paper, the framework of motion estimation relies on image assimilation, which also implies to model the covariances between pixels, variables and dates. The major result of that research comes from the display of the regularization terms as images of correlation values. Analyzing these display regarding the parametrization of the regularization, enables to visualize the region of high covariance of the regularization and allows to objectively determine the values of the weighting coefficients according to image properties. The perspectives concern the design and interpretation of regularization terms, which are suitable to model the structures displayed on image sequences.

ACKNOWLEDGEMENTS

This research has been partially funded by the DGA.

REFERENCES

Baker, S., Scharstein, D., Lewis, J. P., Roth, S., Black, M. J., and Szeliski, R. (2011). A database and evaluation methodology for optical flow. *International Journal on Computer Vision*, 92(1):1–31.

Bérézíat, D. and Herlin, I. (2011). Solving ill-posed image processing problems using data assimilation. *Numerical Algorithms*, 56(2):219–252.

Fortun, D., Bouthemy, P., and Kervrann, C. (2015). Optical flow modeling and computation: A survey. *Computer Vision and Image Understanding*, 134:1 – 21.

Hadamard, J. (1923). *Lecture on Cauchy's Problem in Linear Partial Differential Equations*. Yale University Press, New Haven.

Horn, B. and Schunk, B. (1981). Determining optical flow. *Artificial Intelligence*, 17:185–203.

Le Dimet, F. and Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus Series A : Dynamic meteorology and oceanography*, 38(2):97–110.

Nagel, H.-H. (1995). Nibelungen-platz. *www.ira.uka.de*.

Nagel, H. H. and Enkelmann, W. (1986). An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *Pattern Analysis and Machine Intelligence*, PAMI-8(5):565–593.

Nielsen, M., Florack, L., and Deriche, R. (1994). Regularisation and scale space. Technical Report RR 2352, INRIA.

Oliver, D. (1998). Calculation of the inverse of the covariance. *Mathematical Geology*, 30(7):911–933.

Papadakis, N., Mémin, E., Cuzol, A., and Gengembre, N. (2010). Data assimilation with the weighted ensemble Kalman filter. *Tellus Series A : Dynamic meteorology and oceanography*, 62(5):673–697.

Ridal, M., Lindskog, M., Gustafsson, N., and Haase, G. (2011). Optimized advection of radar reflectivities. *Atmospheric Research*, 100(2–3):213–225.

Sun, D., Roth, S., and Black, M. (2010). Secrets of optical flow estimation and their principles. In *European Conference on Computer Vision*, pages 2432–2439.

Tikhonov, A. N. (1963). Regularization of incorrectly posed problems. *Soviet Mathematics - Doklady*, 4:1624–1627.

Werlberger, M., Pock, T., and Bischof, H. (2010). Motion estimation with non-local total variation regularization. In *Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, USA.