NON-LINEAR OBSERVATION EQUATION FOR MOTION ESTIMATION

D. Béréziat

LIP6,
Université Pierre et Marie Curie,
4 place Jussieu 75005 Paris, France

I. Herlin

INRIA, Domaine de Voluceau, Rocquencourt
BP 105, 78153 Le Chesnay Cedex, France
CEREA, joint laboratory ENPC - EDF R&D,
Université Paris-Est,
77455 Marne la Vallée Cedex 2, France

ABSTRACT

The paper addresses the estimation of motion on an image sequence by data assimilation methods. The core of the study concerns the definition of the data term, or observation equation, that links images to the underlying motion field. In the image processing literature, the optical flow equation is usually chosen to characterize these links. It expresses the Lagrangian constancy of grey level values in time. However, this optical flow equation is obtained by linearization and is no more valid in case of large displacements. The paper discusses the improvement obtained with the original non-linear transport equation of the image brightness by the velocity field. A 4D-Var data assimilation method is applied that solves the evolution equation of motion and the observation equation in its non-linear and linear forms. The comparison of results obtained with both observation equations is quantified on synthetic data and discussed on oceanographic Sea Surface Temperature images.

Index Terms— Image Assimilation, Optical flow, SST images, Variational Data Assimilation.

1. INTRODUCTION

Motion estimation is one major task of Image Processing. In the literature, motion fields are often inferred from the Optical Flow Constraint [1]. This equation is only valid for small displacements as it results from a first order Taylor development of the brightness conservation equation. However, large displacements are visualized on discrete image sequences if the time period of the acquisition process is high compared to the observed dynamics. There are two main approaches in this case. The first one solves the Optical Flow Constraint equation with an incremental algorithm [2, 3], each iteration remaining a linear problem. The second one directly solves the non-linear equation modeling the image brightness transport. Such method is named a warping method [4]. Nevertheless, the underlying dynamics is sometimes partially known. This is the case if the physical processes, visualized on images, are approximately modeled by mathematical equations. This knowledge on dynamics is used in the computation of the solution by Data Assimilation methods as it had been discussed in [5]. As a positive side effect, it allows to retrieve a motion field continuous in time, and not only at the acquisition dates.

The issue of motion estimation is discussed in the following context: the displacements between two consecutive images may be too large, due to the temporal sampling period, so that the linear transport of brightness becomes inadequate.The paper focuses on the observation equation, or data term, and discusses improvements obtained by going back to the non-linear equation compared to its linear version. In Section 2, the issue of motion estimation is introduced. Equations ruling the non-linear and linear evolution of brightness and the dynamics are discussed. In Section 3, the variational data assimilation framework is briefly summarized. References are given to allow the Reader having a full knowledge on the method. In Section 4, results, with the linear observation model and the non-linear one, are quantified on twin experiments and displayed on SST data acquired over Black Sea. Concluding remarks and perspectives are given in Section 5.

2. MATHEMATICAL SETTINGS

Apparent motion measures the displacement of pixels on a sequence of images. This is one signature of the physical processes occurring during the acquisition. Equations ruling it therefore depend on the studied context. The commonly used assumption considers that motion $\mathbf{V}$ transports the image brightness I:

$$I(\mathbf{x} + \triangle t\mathbf{V}(\mathbf{x}; t); t + \triangle t) = I(\mathbf{x}; t) \tag{1}$$

$\mathbf{x}$ and t being the space-time coordinate, and $\triangle t$ the time interval between two frames. As this is non-linear, it is often approximated by a first-order Taylor expansion, commonly referred as Optical Flow Constraint equation:

$$\frac{\partial I}{\partial t} + \nabla I^T \mathbf{V} = 0 \tag{2}$$

Eqs. (1) and (2) are named *observation equations* as they link the image observations to the quantity to be computed, named the state vector, which is the velocity field $\mathbf{V}$.
If the time interval between frames, $\triangle t$, is too large compared to the dynamics time scale, Eq. (2) is no more valid, and Eq. (1) has to be considered in order to obtain correct results, as it will be proven in Section 4.

In the absence of additional information on the dynamics, we consider the heuristics of Lagrangian constancy of velocity. It corresponds to the transport of motion by itself. As this is only an approximation, an error term $\mathcal{E}_m(\mathbf{x}, t)$ is added to the equation:

$$\frac{d\mathbf{V}}{dt} = \frac{\partial \mathbf{V}}{\partial t} + (\mathbf{V}^T \nabla)\mathbf{V} = \frac{\partial \mathbf{V}}{\partial t} + \mathbb{M}(\mathbf{V}) = \mathcal{E}_m(\mathbf{x}, t) \quad (3)$$

The operator $\mathbb{M}$ is named *evolution model*.

In the paper, the observation equation described in Eq. (1) is compared with that of Eq. (2) in the case of large displacements on the discrete image sequence. Let us assume that N images are available. For each image acquisition i, $i = 1, \cdots, N-1$, the observation equation is rewritten as:

- in the non-linear case:

$$I(\mathbf{x} + \triangle t_i \mathbf{V}(x, t_i); t_{i+1}) - I(\mathbf{x}; t_i) = \mathcal{E}_o(\mathbf{x}, t_i) \quad (4)$$

  with $\triangle t_i = t_{i+1} - t_i$ the time interval between acquisition dates t_i and t_{i+1},

- in the linear case:

$$\frac{\partial I}{\partial t}(\mathbf{x}, t_i) + \nabla I^T(\mathbf{x}, t_i)\mathbf{V}(\mathbf{x}, t_i) = \mathcal{E}_o(\mathbf{x}, t_i) \quad (5)$$

$\mathcal{E}_o(\mathbf{x}, t_i)$ is the observation error. It is required as, first, observations are contaminated by noise and, second, their real evolution is only approximated by the equations. Eqs. (4) and (5) are summarized, in a generic way, by:

$$\mathbb{H}(\mathbf{V}, I)(\mathbf{x}, t) = \mathcal{E}_o(\mathbf{x}, t) \quad (6)$$

only valid at acquisition dates t_i.

3. DATA ASSIMILATION

Given the evolution (Eq. (3)) and observation (Eq. (6)) equations on a bounded space-time domain $A = \Omega \times [t_0, t_1]$, the section summarizes the data assimilation process used to retrieve the solution $\mathbf{V}(\mathbf{x}, t)$, from the observed image data $I(\mathbf{x}, t)$.

Errors $\mathcal{E}_m$ and $\mathcal{E}_o$ are supposed unbiased, Gaussian, and characterized by their covariance matrices Q and R, whose definition and properties have been described in [5]. In order to minimize these two errors, the following functional $E(\mathbf{V})$ has to be minimized:

$$\int_A \left[\left(\frac{\partial \mathbf{V}}{\partial t} + \mathbb{M}(\mathbf{V}) \right)^T (\mathbf{x}, t) Q^{-1}(\mathbf{x}, t) \left(\frac{\partial \mathbf{V}}{\partial t} + \mathbb{M}(\mathbf{V}) \right)(\mathbf{x}, t) + \mathbb{H}(\mathbf{V}, I)^T(\mathbf{x}, t) R^{-1}(\mathbf{x}, t) \mathbb{H}(\mathbf{V}, I)(\mathbf{x}, t) \right] d\mathbf{x} dt$$

A weak 4D-Var method is applied, whose complete description is available in [5, 6]. The minimization is performed iteratively. At each iteration, a forward integration of $\mathbf{V}$ and a backward integration of the adjoint variable λ are performed. The forward integration requires the evolution model to be discretized in space. The numerical scheme described in [5] are used. The backward problem requires the definition of the adjoint of the evolution and observation operators. Adjoint of $\mathbb{M}$, formally defined as the dual operator of $\frac{\partial \mathbb{M}}{\partial \mathbf{V}}$, is obtained from the discrete operator $\mathbb{M}$ by an automatic differentiation software [7]. Adjoint of $\mathbb{H}$ is explicitly determined. In the non-linear case, it comes:

$$\left(\frac{\partial \mathbb{H}}{\partial \mathbf{V}} \right)^* (\mathbf{V})(\mathbf{x}, t_i) = \triangle t_i \nabla I \left(\mathbf{x} + \triangle t_i \mathbf{V}(\mathbf{x}, t_i), t_{i+1} \right)$$

and in the linear one:

$$\left(\frac{\partial \mathbb{H}}{\partial \mathbf{V}} \right)^* (\mathbf{V})(\mathbf{x}, t_i) = \nabla I \left(\mathbf{x}, t_i \right)$$

4. RESULTS

4.1. Twin experiments

The improvements, obtained thanks to the non-linear observation equation, are first quantified on twin experiments. One rectangle is moving horizontally, from the left to the right, with a velocity of 50 m/s. On the image data, this corresponds to several displacement values, ranging from 8 to 12 pixels, according to the temporal acquisition period. The discrete sequence, obtained for a displacement of 8 pixels, is visualized on Fig. 1.

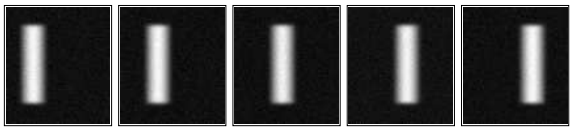

Fig. 1. Image observations.

Data assimilation is applied with the non-linear and linear observation equations. Statistics are computed on the motion result, at the observation dates, and given in Table 1. This table provides the mean of the motion norm ($\bar{N}_{\text{NL}}$ for the non-linear equation and $\bar{N}_{\text{L}}$ for the linear one) and the mean ($\bar{\theta}_{\text{NL}}$ and $\bar{\theta}_{\text{L}}$ respectively) of orientation (in degree). Values should then be 50 m/s for the norm and 0 degree for the orientation. The first conclusion is that the results obtained with the linear

observation equation, for displacements of 10 and 12 pixels, are random and consequently replaced by NR (Not Relevant) in the table. Even with a displacement of 8 pixels, the non-linear observation equation outperforms the linear one, both in norm and orientation, as the linear equation underestimates the norm value.

Displacement	$\bar{N}_{\mathrm{NL}}$	$\bar{\theta}_{\mathrm{NL}}$	$\bar{N}_{\mathrm{L}}$	$\bar{\theta}_{\mathrm{L}}$
8 pixels	43.07	0.89	21.26	-2.57
10 pixels	45.36	1.36	NR	NR
12 pixels	46.60	1.72	NR	NR

Table 1. Statistics on results.

4.2. Satellite data

Data assimilation is also applied on satellite data to estimate motion with the two observation equations. A sequence of four Sea Surface Temperature (SST) images, acquired by NOAA-AVHRR sensor in July 2005 over the Black Sea, is displayed on Figure 2. Two Regions Of Interest (ROI) are

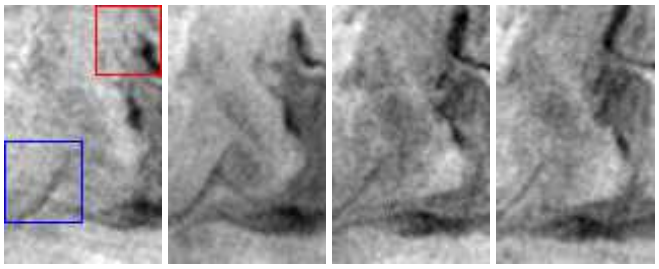

Fig. 2. Four observations. ROI#1 (red) and ROI#2 (blue).

chosen and visualized on the first observation of Fig. 2.

ROI#1 is located at the top-right (red) and displays a structure with a displacement of maximum 12 pixels over the temporal sequence. The non-linear and linear methods compute a displacement field with a maximal norm of respectively 12.3 and 7.8 pixels (see Table 2). Results are displayed on Fig. 3, with the satellite data as background: the linear method clearly under estimates the structure's displacement.

The maximal displacement on ROI#2 (blue) is of 6 pixels. The maximal norms obtained with the non-linear and linear methods are computed (see Table 2). The conclusion is similar: the linear observation equation under estimates the displacements. Results are displayed on Fig. 4 with the image data as background.

ROI	Max visual displ.	$\max N_{\mathrm{NL}}$	$\max N_{\mathrm{L}}$
#1 (red)	12 pixels	12.3	7.8
#2 (blue)	6 pixels	7.0	4.5

Table 2. Evaluation on SST images.

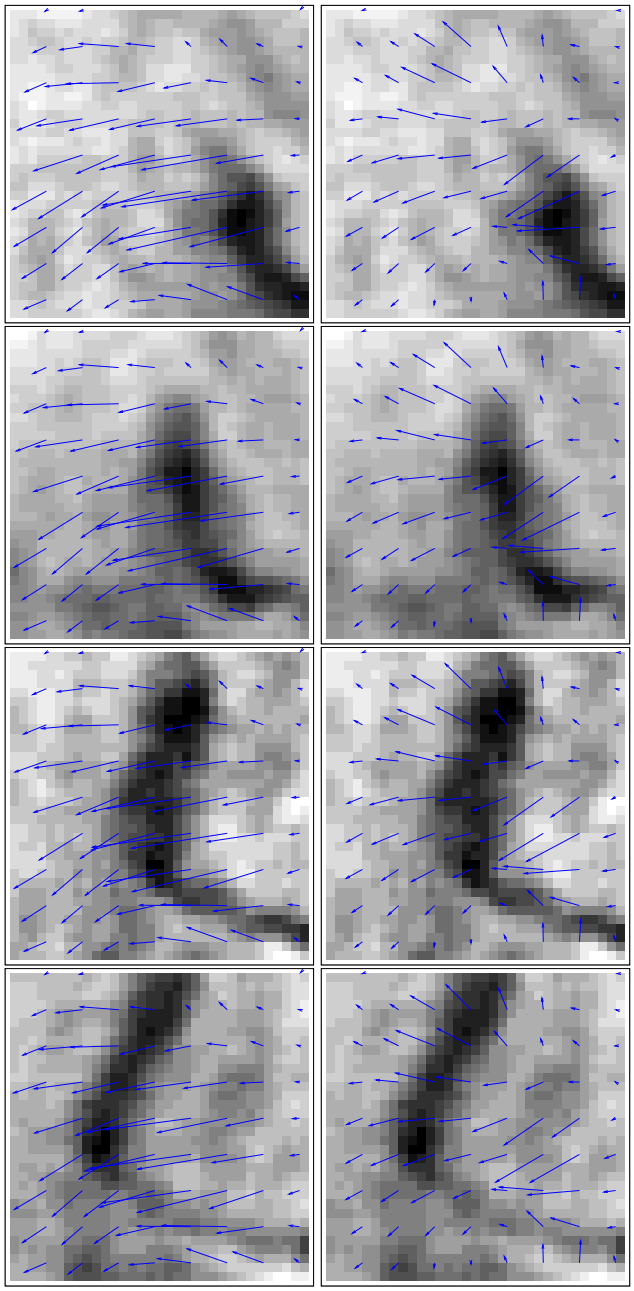

Fig. 3. Results on ROI#1. Left: non-linear observation equation. Right: linear. Arrows correspond to displacements. A 4-pixel sub-sampling is applied on velocity field.

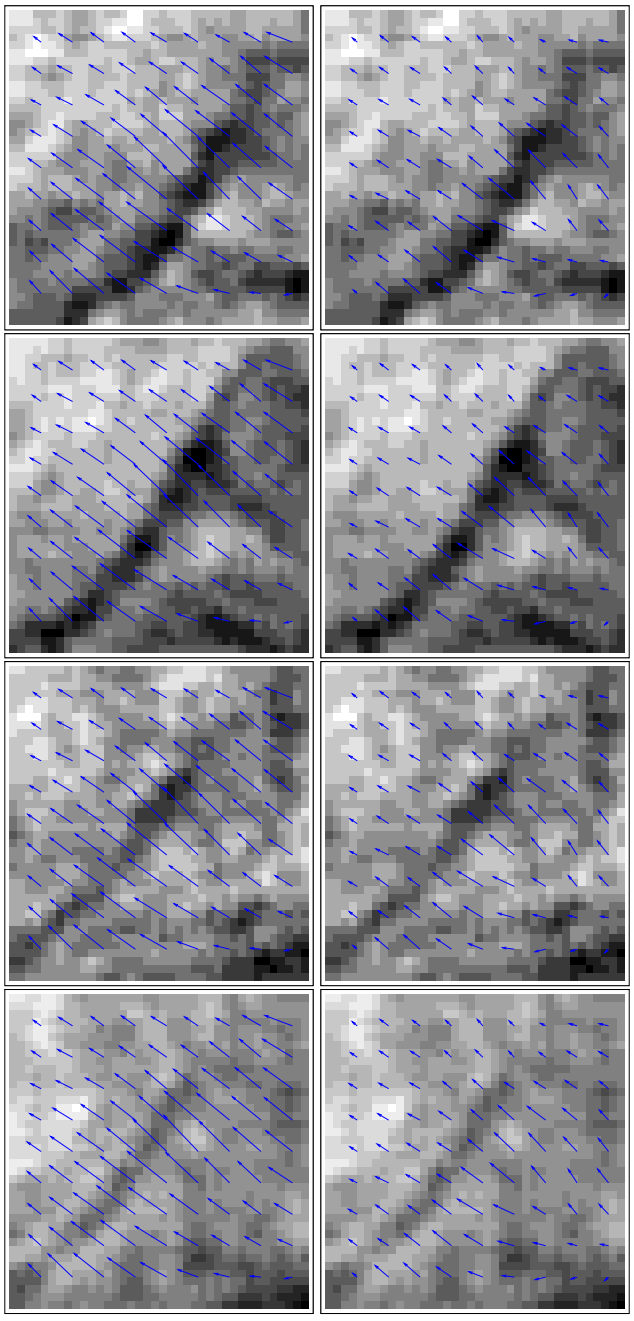

Fig. 4. Result on ROI#2. Left: non-linear observation equation. Right: linear.

5. CONCLUSION

In this paper, a non-linear observation model is compared to its linearized version in an image assimilation process. The Lagrangian constancy of velocity is considered to approximately describe the motion dynamics. A weak formulation of 4D-Var data assimilation is implemented. The estimations, provided by the two observation models, are quantified with twin experiments and tested on satellite SST data, acquired by NOAA-AVHRR sensors over the Black Sea. The results, obtained both on synthetic and real data, confirm the performance of the non-linear equation, when the temporal acquisition period is high compared to the observed underlying dynamics.

The main perspective of this research is to achieve the determination of motion if the object displacement is greater than its size. In this case, coarse-to-fine incremental methods, as well as the non-linear method, fail to correctly recover motion. We then propose to add, in the state vector, a variable describing the trajectory of pixels. The observation operator will consequently measure the effective displacement of pixels, according to their trajectories, and allow a better estimation of motion values.

ACKNOWLEDGEMENTS

Satellite data have been provided by E. Plotnikov and G. Korotaev from the MHI, Sevastopol, Ukraine. Research has been parly funded by the ANR project Geo-FLUIDS (ANR 09 SYSC 005 02).

6. REFERENCES

[1] B.K.P. Horn and B.G. Schunk, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.

[2] J.-M. Odobez and P. Bouthemy, "Direct incremental model-based image motion segmentation for video analysis.," *Signal Processing*, vol. 66, no. 2, pp. 143–155, 1998.

[3] M. Proesmans, L. Van Gool, E. Pauwels, and A. Oosterlinck, "Determination of optical flow and its discontinuities using non-linear diffusion," in *ECCV*, 1994, vol. 2, pp. 295–304.

[4] N. Papenberg, A. Bruhn, T. Brox, S. Didas, and J. Weickert, "Highly accurate optic flow computation with theoretically justified warping," *International Journal on Computer Vision*, vol. 67, no. 2, pp. 141–158, 2006.

[5] D. Béréziat and I. Herlin, "Solving ill-posed image processing problems using data assimilation," *Numerical Algorithms*, vol. 56, no. 2, pp. 219–252, 2011.

[6] E. Valur Hólm, "Lectures notes on assimilation algorithms," European Centre for Medium-Range Weather Forecasts Reading, Apr. 2008.

[7] L. Hascoët and V. Pascual, "Tapenade 2.1 user's guide," Technical Report 0300, INRIA, 2004.