Digital Signal Processing 83 (2018) 45–58

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Motion and acceleration from image assimilation with evolution models

Dominique Béréziat [a,*], Isabelle Herlin [b]

[a] Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6, France
[b] Inria, France

ARTICLE INFO

Article history:
Available online 23 August 2018

Keywords:
Acceleration
Data assimilation
Fluid flow
Motion
Shallow water

ABSTRACT

Image sequences allow visualizing dynamic systems and understanding their intrinsic characteristics. One first component of this dynamics is retrieved from the estimation of the velocity displayed on the sequence. Motion estimation has been extensively studied in the literature of image processing and computer vision. In this paper, we step beyond the traditional optical flow methods and address the problem of recovering the acceleration from the whole temporal sequence. This issue has been poorly investigated, even if this is of major importance for major data types, such as fluid flow images. Acceleration is defined as the space–time function resulting from the forces applied to the studied system. To estimate its value, we propose a variational approach where an energy function is designed to model both the motion and the acceleration fields. The contributions of the paper are twofold: first, we introduce a unified variational formulation of motion and acceleration under space–time constraints; second, we describe the minimization scheme, which allows retrieving the estimations, and provide the full information on the discretization schemes. Last, experiments illustrate the potentiality of the method on synthetic and real image sequences, visualizing fluid-like flows, where direct and precise calculation of acceleration is of primary importance.

1. Introduction

Estimating motion from images is of major importance for a large range of environmental applications. Analyzing satellite acquisitions of Sea Surface Temperature allows, for instance, to detect precursors of extreme events and better mitigate their risks. Processing fish-eye sky images acquired on a solar plant allows forecasting the solar irradiance and accurately estimating the photovoltaic production.

Motion estimation on fluid flows has been extensively discussed in, for instance, Heitz et al. [1]. The underlying problems of the fluid flows context are quite different of those that are usually occurring in most computer vision applications. One major difference is that these fluid flows data require a dense in space and dense in time description and can not be summarized by local features, which are tracked in time. This is the spirit of the research work described in the following and it explains most of the technical choices that were done when implementing the approach.

Acceleration estimation seems having been seldom investigated in the state-of-the-art, whatever the type of image acquisitions. In Arnspang [2], the assumption of Lagrangian constancy of both brightness and velocity is used to estimate motion and acceleration on three consecutive frames. However, these fields are supposed to be locally constant. Hu and Ahuja [3] propose an approach, which is not based on the image data but on a set of characteristics points tracked on consecutive acquisitions, and compute the affine and projective parameters of motion and a constant, in space and time, acceleration. Staković et al. [4] determine the motion and acceleration fields using Fourier-based techniques, but motion is restricted to be locally translational and acceleration is also locally constant.

Multi-frame motion has been the focus of a large number of research works. Tomasi and Kanade [5] and Irani [6] extend the pioneer work of Lucas and Kanade [7] in a multi-frame context. Under the assumption of locally stationary motion, they prove that motion fields are included in a low rank subspace. The main advantage of their methods is the robustness to noise and the capability to solve the aperture problem without applying any regularization process. In the same spirit, Garg et al. [8] also compute non rigid motion fields from their projection on a low rank subspace. Ricco and Tomasi [9] define a multi-frame method to assess long-range

* Corresponding author.
E-mail addresses: Dominique.Bereziat@lip6.fr (D. Béréziat), Isabelle.Herlin@inria.fr (I. Herlin).

https://doi.org/10.1016/j.dsp.2018.08.008

motion and detect the occlusions by computing the Lagrangian trajectories of points on the low rank subspace. Garg et al. [10] apply the same concept for image registration purpose. The major limitation of this whole set of these methods, compared with our approach, is the poor representation of the dynamics, with a small size basis, which is precise enough for their application domains but fails for the complex motion fields of fluid flows. This remark originates our motivation for implementing a dynamic model $\mathbb{M}$ describing both the temporal evolution of the motion field and the transport of the image data.

Accuracy of motion estimation, both in direction and intensity is also a key component for forecasting images at short temporal horizon and mitigating future events. This concern is particularly crucial for environmental issues as, for instance, the short-term forecast of heavy rain or clear sky, which are required for mitigating flashfloods or estimating and regulating the production of photovoltaic energy. However, the pertinence of the forecasted images relies on the full knowledge of the space–time dynamics, and not only on the motion field component. In the operational application which is considered for the discussion given in the paper, such forecast is applied with a temporal sliding-window setting: a set of images is first processed for estimating motion and acceleration, which are further used for forecasting the image data at a given temporal horizon, then the temporal window is iteratively incremented in time. In such context, the only knowledge of motion snapshots is not sufficient for a correct forecast of the future motion fields and a dense in time model of motion and acceleration is strongly required.

For a large range of environmental systems observed with image data, mathematical models of the physical processes are available. This is well-known for meteorology and oceanography for instance, which are both based on the Navier–Stokes equations. These physical laws should then be used when processing the image acquisitions, in order to allow a full and reliable estimation of the dynamics. The paper describes the design of an image model that includes evolution equations for all studied quantities such as image brightness, velocity and acceleration. Then, we discuss the estimation of the full dynamics, motion and acceleration, with a data assimilation approach, which originates in the meteorology forecasting community and is currently used in meteorological institutes all over the world. These data assimilation techniques appeared in the last decade in the image processing and computer vision community for estimating motion from image sequences, as for instance in Papadakis et al. [11,12], Titaud et al. [13] and Béréziat et al. [14]. One primary output of these approaches is an elegant solution of the well-known aperture ambiguity by an explicit model of motion. But the paper makes a strong improvement, compared to these state-of-the-art methods, as it allows the estimation of forces applied to the system, or equivalently the simultaneous estimation of motion and acceleration. For an accurate comparison with the literature, we highlight that Papadakis et al. [11, 12] or Heas et al. [15] also include an additional quantity, which could be viewed as an acceleration term. However, their mathematical formulation constrains this quantity to be small or sparse, consequently suppressing any physical interpretation. In this paper, we focus on a class of approaches named 4D-Var in the data assimilation literature that are relying on an adjoint formulation. However ensemble-based approaches are also possible [16]. Our approach solves the following inverse problem: given N^O images I_l^O, $l = 1, \cdots, N^O$, the motion and acceleration fields are estimated under the constraint of the given dynamic model, expressed by partial differential equations, and space–time regularity properties. Compared to the previously mentioned motion estimation methods that rely on data assimilation, an additional equation is added to the model, which corresponds to the description of the acceleration. If this equation includes a parametric formulation of the acceleration [17], the problem reduces to the estimation of the parameters values without any strong difficulty. But, in the general case of fluid-flows images, the parametric assumption is not valid on the data and a variational data assimilation technique is applied for estimating a dense acceleration field [18]. A specific energy is then designed whose control variables are the values of all variables at the beginning of the studied temporal interval and the acceleration field at each space–time value. The optimization is conducted by computing iteratively the values of the energy and of its gradient, which are the input of the optimization solver, named BFGS [19]. The outputs are the motion and acceleration fields on a continuous temporal interval.

Paper Organization: Section 2 discusses the problem faced in the paper, provides the basic notations and describes the mathematical content. The variational data assimilation technique is then shortly discussed in Section 3, which provides the main mathematical and technical components for understanding the approach. Our method is extensively described in Section 4 for allowing interested Readers to reproduce the full implementation and the experiments. Results are thoroughly discussed in Section 5, and Section 6 concludes the paper and gives indications on future work.

2. Mathematical setting

In order to improve the understanding of our approach, all symbols included in the paper are the same for continuous and discrete descriptions, even if not always fully correct from the mathematical point of view.

Ω is the image domain and $[0, \mathrm{T}]$ is the temporal interval on which images are acquired and processed. The set $A = \Omega \times [0, \mathrm{T}]$ is the studied space–time domain. If a function f is defined on A, $f(\mathbf{x}, t)$ denotes the value at point $\mathbf{x}$ and time t and $f(t)$ describes the spatial field at time t.

The motion vector at point $\mathbf{x}$ and time t is written $\mathbf{w}(\mathbf{x}, t) = \begin{pmatrix} u(\mathbf{x}, t) & v(\mathbf{x}, t) \end{pmatrix}^T$ with $.^T$ being the transpose operator and u and v the horizontal and vertical components. The acceleration is written $\mathbf{a}(\mathbf{x}, t) = \begin{pmatrix} a_u(\mathbf{x}, t) & a_v(\mathbf{x}, t) \end{pmatrix}^T$.

$\langle f, g \rangle$ denotes the scalar product of functions f and g in the continuous domain (or in the discrete domain) and verifies:

$$\langle f, g \rangle = \int_\Omega f(x) g(x) dx \tag{1}$$

A discrete sequence of images I_l^O, $l = 1, \cdots, N^O$, is available and processed for estimating motion and acceleration. I_l^O is acquired at time t_l and is a snapshot of the continuous function I^O, defined on A, with values $I^O(\mathbf{x}, t_l)$.

As pointed out in the introduction, our approach estimates motion and acceleration from images and is based on a dynamic model $\mathbb{M}$. A state vector $\mathbf{X}$ is first defined on A: $\mathbf{X}(\mathbf{x}, t) = \begin{pmatrix} u(\mathbf{x}, t) & v(\mathbf{x}, t) & I(\mathbf{x}, t) \end{pmatrix}^T$, which includes the two components u and v of the velocity $\mathbf{w}$ and a synthetic image I. The function I satisfies the same physical and mathematical properties than the real image acquisitions. It is initialized with the first image I_1^O, at the beginning of the studied temporal interval, and transported by the motion field $\mathbf{w}(\mathbf{x}, t)$. If this transport is correctly performed by the estimated motion field, the brightness values $I(\mathbf{x}, t)$ should be almost identical to the image values $I_l^O(\mathbf{x})$ at each acquisition time t_l. Consequently, our method estimates motion and acceleration by forcing the image function I to be almost identical to the observed images I_l^O.

The model $\mathbb{M}$, expressing the evolution of the state vector, is defined by partial differential equations regulating the time evolution of $\mathbf{w}$, $\mathbf{a}$ and I. The motion and acceleration functions are mathematically linked by: