ASSIMILATION OF SST SATELLITE IMAGES FOR ESTIMATION OF OCEAN CIRCULATION VELOCITY

E. Huot\textsuperscript{1,2,3} \quad I. Herlin\textsuperscript{1,2} \quad G. Korotaev\textsuperscript{4}

\textsuperscript{1} Institut National de Recherche en Informatique et Automatique – France
\textsuperscript{2} Université Paris-Est, CEREA, Joint Laboratory ENPC/EDF R&D Paris-EST – France
\textsuperscript{3} Université de Versailles - Saint-Quentin-en-Yvelines – France
\textsuperscript{4} Marine Hydrophysical Institute – National Academy of Science – Ukraine

ABSTRACT

The objective of this study is to compute the surface circulation velocity from oceanographic image sequences. Data assimilation provides a mathematical solution to combine optimally observations and models. We define a dedicated Image Model consistent with the physical knowledge of ocean dynamic. Satellite images are then assimilated into this model using a variational data assimilation scheme. This technique relies on the adjoint model definition to compute the minimum of a cost function measuring the discrepancy between model’s outputs and observations. The method takes into account the physical knowledge of the ocean circulation, which is relevant for explaining images evolution and, in case of missing data due to clouds occlusion or lack of acquisitions, the estimation is still consistent with the physical evolution laws.

Index Terms— image sequence analysis, partial differential equations, optimal control, sea surface

1. INTRODUCTION

The problem of apparent motion estimation from a sequence of images has long been addressed by the image processing community. It can be tackled by solving a PDEs system composed of a conservation equation and a regularity hypothesis. On the other hand, oceanographers use circulation models and data assimilation to estimate and forecast ocean state.

We recently defined \cite{1} a new method for velocity estimation using the data assimilation framework: the available satellite images constitute observations of Sea Surface Temperature (SST) to be assimilated within a dedicated Image Model (IM). The major drawback of this approach was the lack of physical meaning of the evolution equations used in the IM. To take into account the physics of underlying processes, it is possible to consider these estimated velocities as pseudo-observations which are further assimilated in an oceanographic circulation model \cite{2}.

The herein proposed method is slightly different: it relies on the definition of a new image model directly taking into account some physical knowledge of ocean processes which is relevant to describe images evolution. This so-called Extended Image Model (EIM) is based on both image evolution properties and some physical equations expressing the evolution of the ocean state. Variational data assimilation is used to assimilate SST observations into the EIM in order to estimate ocean surface velocity. The paper is organized as follow: section 2 summarizes the state of the art concerning fluid motion estimation in the context of image assimilation; section 3 presents the model used to describe image evolution taking into account physical knowledge of ocean dynamic; and section 4 explains how variational data assimilation principles are used to set up image assimilation and to estimate the sea surface velocity.

2. STATE OF THE ART

Ocean surface velocity can be approached by image processing techniques applied to sequences of satellite images. Classical methods for dense apparent motion computation rely on a conservation equation \cite{3}, which is not sufficient to obtain the two components of the velocity vector. An additional constraint has to be set up, generally relying on a regularity hypothesis of the result. To apply this approach on fluid motion, several authors proposed either a dedicated conservation equation and/or dedicated regularity constraints. For instance, in \cite{4, 5} authors propose to replace the usual luminance conservation by a mass conservation; in \cite{5, 6} authors use a regularity constraint relying on a solenoidal and irrotational description of the motion field. These approaches have two major drawbacks. First, they strongly rely on the computation of spatial and temporal derivatives, which is very difficult in case of missing data. This point becomes crucial with satellite images that can be significantly occluded by clouds. Second, the equations used for the estimation do not have any physical meaning. Recently, the image processing community has been interested in data assimilation tools \cite{7}, which provide an optimal way to solve the problem of missing data. According to the same idea, we introduced in \cite{1} an Image Model that expresses the link between image observations (SST) and ap-
parent motion (surface velocity):

\[
\begin{aligned}
&\frac{\partial T}{\partial t} + \nabla \cdot \mathbf{v} = K_T \Delta T \\
&\frac{\partial u}{\partial t} = 0 \\
&\frac{\partial v}{\partial t} = 0,
\end{aligned}
\]

(1)

where the first line is a simplification of the advection-diffusion equation governing the transport of temperature of an uncompressible fluid and \( K_T \) stands for the temperature diffusivity parameter. The two other lines constrain the evolution of the two components of velocity \( \mathbf{v} = (u,v) \). SST satellite data are assimilated into the IM in order to estimate surface velocity.

The major drawback of this approach is that the equations of velocity evolution are quiet restrictive and do not have any physical origin.

3. EXTENDED IMAGE MODEL

3.1. Definition

To define an Extended Image Model taking into account physical knowledge of the observed flow, it is necessary to address ocean models conceived by oceanographers and applied mathematicians. In a first stage, we are interested in models based on the Saint-Venant approximation [8]. They correspond to Navier-Stokes equations in a 2D formulation leading to shallow water models. These models can be used for atmospheric simulation, hydrology and operational oceanography [9]. They link the 2D velocity of a merged layer of the fluid to the thickness \( h \) of this layer.

The EIM definition includes the same transport equation of temperature as the IM but the velocity evolution equation comes from the shallow water model. Moreover, the EIM includes an equation on thickness \( h \) evolution:

\[
\begin{aligned}
&\frac{\partial T}{\partial t} + \nabla \cdot \mathbf{v} = K_T \Delta T \\
&\frac{\partial u}{\partial t} = f_v = g \frac{\partial h}{\partial x} + K_v \Delta u \\
&\frac{\partial v}{\partial t} = f_u = g \frac{\partial h}{\partial y} + K_v \Delta v \\
&\frac{\partial h}{\partial t} + \frac{\partial (uh)}{\partial x} + \frac{\partial (vh)}{\partial y} = 0,
\end{aligned}
\]

(2)

where \( f \) is the Coriolis parameter; \( K_v \) the horizontal diffusivity of the mixed layer; \( g' = g(\rho_0 - \rho_1)/\rho_0 \) the reduced gravity, with \( \rho_0 \) corresponding to the reference density and \( \rho_1 \) to the average density of the mixed layer. The first line of system (2) corresponds to the evolution of image properties and the latter lines correspond to oceanographic parameters evolution.

3.2. Discretization

In the context of data assimilation it is important to work with the discretized version of the model, especially when the adjoint model is needed. We can rewrite the EIM system as:

\[
\begin{aligned}
&\frac{\partial T}{\partial t} = -\frac{\partial T}{\partial x} u - \frac{\partial T}{\partial y} v + K_T \left( \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) \\
&\frac{\partial u}{\partial t} = -\frac{\partial u}{\partial x} u - \frac{\partial u}{\partial y} v + f_v + g' \frac{\partial h}{\partial x} + K_v \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \\
&\frac{\partial v}{\partial t} = -\frac{\partial v}{\partial x} u - \frac{\partial v}{\partial y} v - f_u + g' \frac{\partial h}{\partial y} + K_v \left( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) \\
&\frac{\partial h}{\partial t} = -\frac{\partial u}{\partial x} h - u \frac{\partial h}{\partial x} - \frac{\partial v}{\partial y} h - v \frac{\partial h}{\partial y}.
\end{aligned}
\]

(3)

The system (3) corresponds to:

\[
\frac{\partial X}{\partial t} = -F(X,t),
\]

(4)

where \( X = (T, u, v, h) \) is the state variable of the model and the operator \( F \) stands for the differential dynamic model itself, ie the right part of system (3) equations.

The equation (4) is discretized in time using an explicit Euler scheme:

\[
X^{t+\Delta t} = X^t - F(X^t, t) \Delta t.
\]

Spatially, the partial derivatives are estimated using finite differences computed on Arakawa shifted C-grids [10]. Unfortunately the two terms: \( \frac{\partial u}{\partial x} u \) and \( \frac{\partial v}{\partial y} v \) of the second and third equations lead to numerical unstabilities. In order to avoid this problem it is possible to use the Friedriech scheme [11]. For instance, we introduce an intermediate variable \( f_u = \frac{1}{2} u^2 \) and replace \( \frac{\partial u}{\partial x} u \) by \( \frac{\partial f_u}{\partial x} \). Finally, the spatial derivative of \( f_u \) is estimated using a centered finite difference.

4. VARIATIONAL DATA ASSIMILATION

4.1. Introduction

Data assimilation is a mathematical tool related to optimal control theory. It allows to estimate the actual state of a system of interest according to an initial state, a dynamic model and partial noisy observations. In practice, the evolution of the state variable \( X \) is assumed to be described by a differential dynamic model \( F \) and the initial state \( X_0 \) (which is the control parameter):

\[
\begin{aligned}
&\frac{\partial X}{\partial t} + F(X) = 0 \\
&X(t = 0) = X_0 + \varepsilon_0.
\end{aligned}
\]

(5)
Partial noisy observations \( Y \) are available and linked to \( X \) through an observation operator \( H \):

\[
Y = H(X) + \varepsilon_o.
\] (6)

A least square estimation of the control variable regarding the whole set of available measurements (within a considered time range \([0, \tau]\)) is obtained by minimizing, with respect to the control variable \( X_0 \), a cost function:

\[
J(X_0) = \frac{1}{2} \int_{\Omega,t} (X - X_b)^T B^{-1} (X - X_b) + \frac{1}{2} \int_{\Omega,t} (H(X) - Y)^T R^{-1} (H(X) - Y)
\] (7)

where \( B \) and \( R \) are the covariance error matrices of respectively the background state \( X_b \) and the observations \( Y \). In our case, we can notice that we haven’t any information on \( X_b \), so the first term of equation (7) disappears and the cost function becomes:

\[
J(X_0) = \frac{1}{2} \int_{\Omega,t} (H(X) - Y)^T R^{-1} (H(X) - Y)
\] (8)

Most minimization methods rely on the functional gradient computation. As the state dimension is about \( 10^5 \) to \( 10^6 \) (for instance \( X \) has 240000 for the \( EIM \) with 6 image observations of size \( 100 \times 100 \) pixels) it becomes impossible to compute \( \nabla J(X_0) \) by finite differences. However, adjoint methods [12] authorize the computation of the gradient functional in only one integration of the adjoint model.

### 4.2. Adjoint model

To obtain the adjoint model, the gradient of the functional \( J \) in the direction \( x_0 \), is computed as:

\[
\left\langle \frac{\partial J}{\partial X_0}, x_0 \right\rangle = \int_{\Omega,t} \langle (H(X) - Y), H\left( \frac{\partial X}{\partial X_0} \right) \rangle R
\]

\[
= \int_{\Omega,t} \langle H^T R (H(X) - Y), dX \rangle
\] (9)

The system (5) is differentiated as:

\[
\begin{cases}
\frac{\partial (dX)}{\partial t} + \frac{\partial F(X)}{\partial X} dX = 0 \\
dX(t_0) = DX_0.
\end{cases}
\] (10)

The first equation of (10) is multiplied by a new variable \( p \), called adjoint variable, and integrated in time:

\[
\int_{\tau} \langle \frac{\partial (dX)}{\partial t} + \frac{\partial F(X)}{\partial X} dX, p \rangle = 0.
\] (11)

After an integration by parts, we obtain:

\[
\int_{\tau} \left\langle \frac{\partial p}{\partial t} + \frac{\partial F'(p)}{\partial X}, dX \right\rangle = \langle p(0), dX(0) \rangle - \langle p(\tau), dX(\tau) \rangle
\] (12)

where \( F' \) is the adjoint operator defined by:

\[
\forall X_1, X_2: \langle X_1, \frac{\partial F'(X_2)}{\partial X} \rangle = \langle \frac{\partial F(X_1)}{\partial X}, X_2 \rangle
\]

Let assume that \( p(\tau) = 0 \) we obtain the adjoint system of (5):

\[
\begin{cases}
\frac{\partial p}{\partial t} + \frac{\partial F'(p)}{\partial X} = H^T R (Y - H(X)) \\
p(\tau) = 0.
\end{cases}
\] (13)

By combining (9), (12) and (13) we obtain:

\[
\frac{\partial J}{\partial X_0} = p(0),
\] (14)

and the gradient of \( J \) is obtain by a backward integration of the adjoint model. This technique is called the 4DVAR assimilation scheme.

### 4.3. Application

In this paper, SST observations are assimilated into the \( EIM \) in order to estimate \( v_0 = (v_{0x}, v_{0y})^T \). Considering we haven’t any information on the initial state, the assimilation system can be rewritten as:

\[
\begin{cases}
\frac{\partial X}{\partial t} + F(X) = 0 \\
X(0) = X_0 \\
Y = H(X) + \varepsilon_o
\end{cases}
\] (15)

where the first line corresponds to the evolution model \( EIM \) defined by system (3); the second line is the initial background condition; and the third line directly links \( X \) to the observations \( Y = T \) through the projection operator \( H \). In a first stage, we assume that observation errors are uncorrelated in space and time: \( \varepsilon_o \) gives infinite values on position occluded by clouds (missing data), and a constant value everywhere else (this constant should be determined by the sensor performance).

In practice the adjoint of system (15) is computed by using an automatic differentiation tool [13] and the minimization is performed by the quasi-Newton algorithm M1QN3 [14] commonly used in operational forecast systems.

In order to validate this method, we have used simulation of SST images computed by a 3D ocean simulation model [15]. The figure 1 displays simulation of three SST images given by the reference model, the \( IM \) and the \( EIM \). In order to perform correct estimations the \( IM \) and \( EIM \) should provide the same kind of simulation than the reference model. It is clear on figure 1 that the simulation computed with the \( EIM \) are closer to the reference model images than the one computed with the \( IM \). The figure 2 displays simulation of surface velocities simulated by the reference model and the \( EIM \); note that \( IM \) velocities are not depicted here: they are supposed to be constant. The figure 3 shows an example of velocity estimation result.
5. CONCLUSION

This paper presents a velocity estimation method applied to oceanographic image sequences. It is based on variational assimilation of image information into a dedicated image model. This approach has two major advantages: it can deal with missing data due to cloud occlusion, and it takes into account some physical knowledge of the ocean circulation.

6. REFERENCES


Fig. 1. Comparison of SST simulations coming from a 3D model (top), the first IM (middle), and the EIM (bottom).

Fig. 2. Comparison of surface velocity simulations coming from a 3D model (top) and the EIM (bottom).

Fig. 3. Result of velocity estimation.


