Space-time Reconstruction of Oceanic Sea States via Variational Stereo Methods

Guillermo Gallego [1], *Anthony Yezzi* [2], *Francesco Fedele* [2,3], *Alvise Benetazzo* [4]
1. Grupo de Tratamiento de Imagenes, Universidad Politecnica de Madrid, Madrid, Spain.
2. School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, USA.
3. School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, USA.
4. CNR-ISMAR, Venice, Italy.

ABSTRACT

We present a remote sensing observational method for the measurement of the spatio-temporal dynamics of ocean waves. Variational techniques are used to recover a coherent space-time reconstruction of oceanic sea states given stereo video imagery. The stereoscopic reconstruction problem is expressed in a variational optimization framework. There, we design an energy functional whose minimizer is the desired temporal sequence of wave heights. The functional combines photometric observations as well as spatial and temporal regularizers. A nested iterative scheme is devised to numerically solve, via 3-D multigrid methods, the system of partial differential equations resulting from the optimality condition of the energy functional. The output of our method is the coherent, simultaneous estimation of the wave surface height and radiance at multiple snapshots. We demonstrate our algorithm on real data collected off-shore. Statistical and spectral analysis are performed. Comparison with respect to an existing sequential method is analyzed.

KEY WORDS

Marine technology, remote sensing, hydrodynamics, stereo vision, image processing, variational methods, multigrid methods

INTRODUCTION

The study of the dynamics of oceanographic phenomena using vision systems has a long tradition that dates back to the first half of the previous century (Schumacher, 1939; Shemdin et al., 1988; Holland et al., 1997), etc. This topic has gained popularity in previous years due to the economical and practical benefits of these systems (Santel et al., 2004; Benetazzo, 2006; Wanek & Wu, 2006; MacHutchon & Liu, 2007; Fedele et al., 2008, 2011; Hsiao & Huang, 2009), and more recently in (de Vries et al., 2011; Bechle & Wu, 2011; Astruc et al., 2012).

In this paper, we build upon the graph variational method presented in (Gallego et al., 2011a,b) for the stereoscopic reconstruction of oceanic sea states. The variational method for still images was extended to process stereo video on a sequential, snapshot-by-snapshot basis. Besides this, the variational framework also allows for more ways to incorporate temporal coherence on the reconstructed surface. The ultimate goal of such an observational technique is to include the spatial and temporal physics of the waves in the reconstruction step (Gallego, 2011). Because ocean waves are governed by the wave equation, it would be desirable to include such law in the estimation process. This is however, a challenging problem. Before considering such approach, a natural way to enforce temporal coherence of the reconstructed wave height, besides the purely sequential processing above mentioned, is to include a temporal regularizer in our energy model and solve the resulting variational problem. This is the object of this work and it implies a *simultaneous* estimation of the unknowns (wave height, radiance, etc.) for *all* snapshots in a sequence. The resulting reconstruction is called a *manifold* reconstruction (MR) because the developed method estimates a manifold of graphs (wave heights or elevation maps).

THEORETICAL MODEL

In a multi-camera setup recording synchronized videos of the ocean surface, let the video signal acquired by the i-th camera be denoted by $I_i(\mathbf{x}_i, \tau)$, where $\mathbf{x}_i = (x, y)^\top$ and τ denote the (continuous) spatial and temporal variables, respectively. In reality, a digital video signal only specifies the intensity values at discrete locations of the space-time axes (pixels and time instants), however, it is more beneficial to model the problem in the continuum. Following the physical model of the scene presented in (Gallego et al., 2011a), consider the surface shape (wave height) and surface radiance as functions of space *and time*, i.e., as 3-D functions: $Z(u, v, \tau)$ and $f(u, v, \tau)$, with $\mathbf{u} = (u, v) \in U$ and $\tau \in [0, T]$. Recall that the surface radiance is a function that lives on top of the surface and represents the "color" of the wave that is seen from the viewpoints of the cameras. Let us denote the domain of Z and f by $U_T = U \times [0, T]$. Let the symbol ∇Z denote the gradient of Z with respect to all variables, not only the spatial ones (u, v),

i.e.

$$\nabla Z = (Z_u, Z_v, Z_\tau)^\top,$$

and similarly for ∇f. With this notation, the three variables u, v and τ are treated uniformly.

Next, we define an energy functional to measure the goodness of fit of candidate space-time functions Z and f to the observed data (i.e., the stereo videos):

$$E(Z, f) = E_{\text{data}}(Z, f) + \alpha E_{\text{geom}}(Z) + \beta E_{\text{rad}}(f), \tag{0.1}$$

with weights $\alpha, \beta \in \mathbb{R}^+$. Let the data fidelity term, which measures the photo-consistency throughout the video for a candidate wave height function, be

$$E_{\text{data}} = \sum_{i=1}^{N_c} E_i,$$

with

$$E_i \doteq \int_0^T \int_{\Omega_i} \phi_i \, d\Omega_i d\tau \tag{0.2}$$

and photometric criterion

$$\phi_i = \tfrac{1}{2}\big(I_i(\mathbf{x}_i, \tau) - f(\mathbf{x}_i, \tau)\big)^2. \tag{0.3}$$

The definition of the data fidelity term as an integral over the image domain Ω_i (rather than over U) has two advantages: the data term is independent of the choice of domain for the graph and the resulting optimality conditions for the minimization of (0.1) lack image derivatives. This desirable property is inherited from the modeling and mathematical principles that we follow from (Mumford & Shah, 1988). The resulting method is less sensitive to image noise than other variational approaches for stereo 3-D reconstruction.

Also, let the spatio-temporal regularizers on the smoothness of the surface shape and radiance be:

$$E_{\text{geom}} = \int_{U_T} \tfrac{1}{2}\|\nabla Z\|^2 \, d\mathbf{u}_T, \tag{0.4}$$

$$E_{\text{rad}} = \int_{U_T} \tfrac{1}{2}\|\nabla f\|^2 \, d\mathbf{u}_T. \tag{0.5}$$

where $d\mathbf{u}_T = d\mathbf{u}\, d\tau$. The new terms with respect to the still-image case (Gallego et al., 2011a), Z_τ^2 and f_τ^2, are the source of temporal coherence in the model.

The radiance function is an auxiliary variable with a twofold purpose: it simplifies pairwise image-to-image comparisons with fewer image-to-model comparisons and it also extends the influence of the point-wise photometric criterion by means of neighborhood effects due to the regularizer E_{rad}. The latter endows the energy functional with a global influence and improves the robustness of the energy functional against image noise.

The data fidelity term (0.2) corresponding to the i-th camera can also be expressed as

$$E_i = \int_{U_T} \phi_i \mathtt{J}_i \, d\mathbf{u}_T,$$

where the Jacobian $\mathtt{J}_i$ (the geometric conversion factor between integrals) has the same expression as in the still-image case despite the fact that now the Jacobian is also a function of time τ (since $\mathtt{J}_i$ depends on the surface shape, which is a function of time). Therefore, the energy (0.1) can be expressed in the common and fixed domain of integration U_T as the integrand of the so-called Lagrangian $L(Z, \nabla Z, f, \nabla f, u, v, \tau)$:

$$E = \int_{U_T} L \, d\mathbf{u}_T.$$

To minimize the proposed energy, let us compute the necessary optimality condition, which follows from the first variation (directional derivative) of the energy with respect to the unknowns:

$$D_{(h,w)}E(Z, f) = \frac{d}{d\epsilon} E(Z + \epsilon h, f + \epsilon w)\bigg|_{\epsilon=0}.$$

Setting to zero the directional derivative for all admissible perturbations (h, w) we arrive at a coupled system of partial differential equations (PDE), the so-called Euler-Lagrange (EL) equations:

$$\begin{aligned} L_Z - (L_{Z_u})_u - (L_{Z_v})_v - (L_{Z_\tau})_\tau &= 0 \quad \text{in } U_T, \\ L_{Z_u}\nu^u + L_{Z_v}\nu^v + L_{Z_\tau}\nu^\tau &= 0 \quad \text{on } \partial U_T, \\ L_f - (L_{f_u})_u - (L_{f_v})_v - (L_{f_\tau})_\tau &= 0 \quad \text{in } U_T, \\ L_{f_u}\nu^u + L_{f_v}\nu^v + L_{f_\tau}\nu^\tau &= 0 \quad \text{on } \partial U_T. \end{aligned}$$

After calculations, similar expressions to those of the still-image optimality conditions are derived. That is why this approach is considered a natural extension of the aforementioned still-image case. Now, however, variables Z, f, and consequently $\mathtt{J}_i$ and I_i depend on the new temporal variable τ:

$$g(Z, f) - \alpha \Delta Z = 0 \quad \text{in } U_T, \tag{0.6}$$

$$b(Z, f) + \alpha \frac{\partial Z}{\partial \boldsymbol{\nu}} = 0 \quad \text{on } \partial U_T, \tag{0.7}$$

$$-\sum_{i=1}^{N_c} (I_i - f)\mathtt{J}_i(Z) - \beta \Delta f = 0 \quad \text{in } U_T, \tag{0.8}$$

$$\beta \frac{\partial f}{\partial \boldsymbol{\nu}} = 0 \quad \text{on } \partial U_T, \tag{0.9}$$

where the non-linear terms due to the data fidelity energy have the same expression as those of the still-image energy since the data fidelity energy does not depend on the new derivatives Z_τ, f_τ:

$$g(Z, f) = \nabla_{\mathbf{u}} f \cdot \sum_{i=1}^{N_c} |\mathtt{M}^i| \tilde{Z}_i^{-3} (I_i - f)(u - C_i^1, v - C_i^2),$$

$$b(Z, f) = \sum_{i=1}^{N_c} \phi_i |\mathtt{M}^i| \tilde{Z}_i^{-3} \big((u - C_i^1)\nu^u + (v - C_i^2)\nu^v\big).$$

Recall that, for the i-th camera, $\mathbf{C}_i = (C_i^1, C_i^2, C_i^3)^\top$ is the optical center, $|\mathtt{M}^i|$ is the product of the focal lengths and $\tilde{Z}_i$ is the depth of the surface point with respect to the camera. The 3-D Laplacians $\Delta Z = Z_{uu} + Z_{vv} + Z_{\tau\tau}$ and, similarly, Δf arise from the regularizing terms (0.4) and (0.5), respectively. $\partial * / \partial \boldsymbol{\nu} = (\nabla *) \cdot \boldsymbol{\nu}$ is the usual notation for the directional derivative along $\boldsymbol{\nu} = (\nu^u, \nu^v, \nu^\tau)^\top$, the normal to the integration domain U_T in the parameter space. $\nabla_{\mathbf{u}} f$ is the gradient of f with respect to the original spatial variables only. To simplify the model, we replace (0.7) by homogeneous Neumann boundary conditions, as in the still-image case.

Having introduced the manifold reconstruction model, let us make it more flexible by including an extra parameter $\rho^2 \geq 0$ to control the amount of temporal regularization of the solution relative to the amount of spatial regularization. The effect of this

parameter is the substitution of the 3-D Laplacian $\Delta Z = Z_{uu} + Z_{vv} + Z_{\tau\tau}$ by the weighted sum $Z_{uu} + Z_{vv} + \rho^2 Z_{\tau\tau}$, and similarly for Δf. There are two possible interpretations to this anisotropic diffusion operator. The simplest one is that the operator arises by replacing the integrand in (0.4) by $\frac{1}{2}(Z_u^2 + Z_v^2 + \rho^2 Z_\tau^2)$, thus using a weighted norm instead of the Euclidean norm (and similarly for (0.5)). The second interpretation is that the anisotropic diffusion operator arises by using the Euclidean norm (0.4) in a deformed space where variable τ is scaled by ρ with respect to variables (u, v) to yield the desired non-uniform scaling of the gradients and the Laplacians, $\nabla Z, \nabla f, \Delta Z$ and Δf.

NUMERICAL SOLUTION

An iterative method is used to find the minimum of energy (0.1) via the solution of the coupled system of equations that arise from the the necessary optimality condition of the energy. Observe that, for a fixed height, equation (0.8) is a linear PDE in the radiance, which is easier to solve than the non-linear PDE in the height (0.6) for a fixed radiance. By exploiting this asymmetry one may devise a minimization strategy consisting of a nested iterative scheme: an outer loop performing a gradient descent in the height, and an inner loop implementing a direct optimization for the radiance. The scheme is initialized by an approximate solution, usually consisting of the zero height function and the corresponding optimal radiance without regularizer ($\beta = 0$). The linear PDE in the radiance is faster to solve using classical stationary iteration methods such as Jacobi or Gauss-Seidel rather than setting up a gradient descent equation and using time-stepping[1] solvers. The PDE in the height is more complicated and it is solved via time-stepping methods.

Numerical discretization of the PDEs is carried out using Finite Difference Methods (FDM). Therefore, the integration domain U_T is discretized by means of a 3-D grid. Forward differences in fictitious time and central differences in (u, v, τ) approximate the derivatives in the PDEs. The quantity $\tilde{\rho} = \rho h / \Delta\tau$ can be interpreted as the step ratio that states the relationship between the grid steps $h = \Delta u = \Delta v$ and $\Delta\tau$ in the anisotropic space previously mentioned.

Both updating schemes (stationary method for f and time-stepping method for Z) are used as relaxation procedures inside a 3-D multigrid method (Briggs et al., 2000; Trottenberg, 2000) that approximately solves the EL equations. Multigrid methods are the most efficient numerical tools for solving elliptic boundary value problems. Each relaxation iteration implies the update of the values of the approximate solution (either Z or f) at *all* the grid points, that is, all physical time slices of the 3-D grid are updated in the same iteration. This simultaneous processing property is the main difference with respect to the sequential scheme (Gallego et al., 2011a).

Due to the extra dimension added to the problem, memory becomes a precious resource. Storage is required not only for the unknowns Z and f at the different levels of multigrid, but also for the coefficients of the linear/linearized PDEs and other temporary variables. For the same amount of memory, in-place updating schemes allow longer sequences of images to be processed. Full Multigrid (Briggs et al., 2000) with zero initial condition is used to initialize the 3-D multigrid solver. It is also possible to initialize the iterator with the solution from a fast sequential 2-D multigrid, possibly computed using a coarser version of the problem.

[1] artificial time, not the physical time τ.

Comparison to sequential 2-D multigrid

Although the sequential and manifold reconstruction methods (SR and MR, respectively) share a common base theoretical modeling, their major difference is that SR relies on 2-D multigrid (snapshot-wise), whereas MR is based on 3-D multigrid. A comparison of the speed performance between both methods is not the focus of this work, but we discuss this topic of practical relevance. In general, we may conclude in a first analysis that SR is simpler and faster but more sensitive to noise than MR.

The two main components of any multigrid method are the intergrid transfer operators and the smoothers to update the approximate solution. Let us compare both qualitatively for SR and MR. On the one hand, 3-D restriction is faster than sequential 2-D restriction because each coarser level of 3-D multigrid has roughly half the number of snapshots of the previous level, whereas in sequential 2-D multigrid the number of snapshots remains constant and maximal for all levels. On the other hand, assuming full weighting restriction, each point in the coarse grid requires a weighted sum of 9 points in case of 2-D restriction and 27 points in case of 3-D restriction. Thus, each point-wise 3-D restriction involves more neighbors and calculations than each 2-D restriction. Similar comments apply to the prolongation operators: trilinear interpolation is applied to a smaller number of grid points than sequential bilinear interpolation, but the former is more expensive point-wise. For details on 2-D and 3-D inter-grid transfer operators, see (Trottenberg, 2000).

The update of each grid point in MR is slower than the update of each point in SR because the former depends on more neighbors (temporal derivatives and 3-D Laplacian). In addition, SR only performs full multigrid for the reconstruction of the first snapshot (and forward propagation of the solution to initialize the remaining snapshots), whereas MR carries out full multigrid for all snapshots. Moreover, the artificial time step to evolve the non-linear gradient descent PDE in the height is smaller in MR than in SR because the former takes the minimum step over all snapshot-wise steps of the latter. Hence, MR requires more iterations than SR to reach the same final, artificial time. On the other hand, MR is more robust than SR because it is less prone to be trapped in a local minimum caused by the poor reconstruction of a snapshot. Errors are distributed better among all snapshots in MR than in SR.

EXPERIMENTS

Our variational method has been tested on stereo video acquired at an off-shore platform near the southern seashore of the Crimean peninsula, in the Black Sea. Two cameras mounted 12 m above the mean sea level and 2.5 m apart provide images of 1624×1236 pixels, acquired at a frame rate of 10 Hz.

Current implementation of the MR method requires all relevant information (input images, processing grids, etc.) to be stored in the computer's main memory (RAM). Thus, the limit in the amount of available main memory in the computer performing the reconstruction imposes a constraint on the size of the achievable reconstruction. For a fixed amount of memory, there is a trade-off between the spatial and temporal sizes of the grid within the MR method: one can either reconstruct a few snapshots with high spatial resolution or a larger number of snapshots at a coarser spatial resolution.

For example, in a computer with 2 GBytes of RAM, a 6-level multigrid solver can roughly handle a reconstruction on a spatial

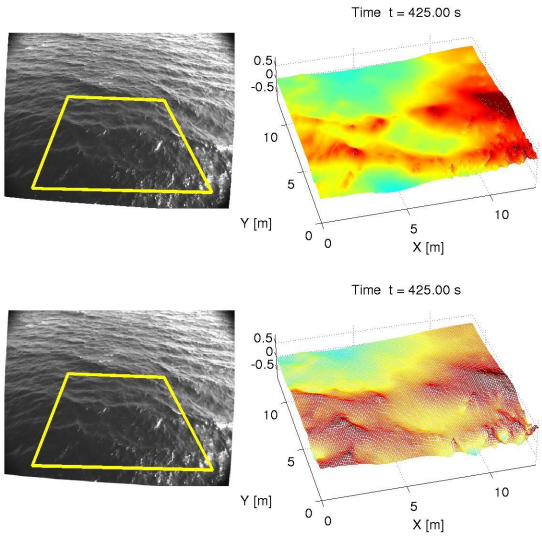

Figure 0.1: Right: reconstruction of a snapshot at spatial resolutions $h = 2.5$ cm (top) and $h = 10$ cm (bottom). Wave height is pseudo-colored, from blue (low) to red (high). Left: one of the two input images at the corresponding resolution (top: 1624×1236 pixels, bottom: 406×309 pixels), with highlighted reconstructed region of interest.

grid of 513×513 points (resolution $h = 2.5$ cm) on 65 snapshots of the original dataset or a reconstruction on a grid of 129×129 points ($h = 10$ cm) on 1025 snapshots of a $\times 4$ coarser version of the images. Fig. 0.1 shows an example of the reconstructions obtained at both spatial resolutions previously mentioned. This figure shows that the reconstruction at the low resolution captures the gist (i.e., low spatial frequency components) of the wave heights, thus resembling the reconstruction at the high spatial resolution.

The MR algorithm acts on the entire sequence of images by reconstructing pieces (sub-sequences) of consecutive snapshots. Other coarse-to-fine strategies are also possible to ensure a smooth transition near the temporal boundaries of the sub-sequences.

Snapshots can also be decimated in time. The linear dispersion relation that relates spatial and temporal frequencies of waves in deep water, $k = \omega^2/g$ (where $k = 2\pi/\lambda$ is the wave number, $\omega = 2\pi f$ is the angular frequency and g is the gravity acceleration), serves as a physical criterion to choose a reasonable frame rate for a given spatial resolution, and vice versa. For instance, in the example with grid resolution $h = 10$ cm, assuming the minimum spatial wavenumber that the algorithm reconstructs reliably is $\lambda = 4h$, the corresponding frequency of the wave is

$$f = \frac{\omega}{2\pi} = \frac{1}{2\pi}\sqrt{g\frac{2\pi}{\lambda}} = \sqrt{\frac{g}{2\pi\, 4h}} \approx 2 \text{ Hz}.$$

If this is the maximum temporal frequency that the algorithm reconstructs reliably, the corresponding Nyquist rate (minimum sampling rate required to avoid aliasing) is twice as much, 4 Hz, meaning that snapshots should be at most $\Delta\tau \leq 1/4 = 0.25$ seconds apart. In the experiments, we used a smaller resolution $\Delta\tau = 0.1$ s since the acquisition rate (10 Hz) allowed so, but we could have used $\Delta\tau = 0.2$ s (i.e., temporal snapshot decimation by a factor of 2) to achieve results with similar validity and interpretation, albeit expanding twice the physical time interval.

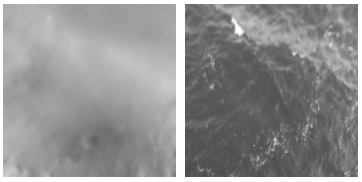

Figure 0.4: Experiment I. A slice at constant $\tau = \tau_0$. Size: 129×129 grid points. Left: surface height $Z(u, v, \tau_0)$ (grayscale encoded, from dark (low) to white (high)). Right: surface radiance $f(u, v, \tau_0)$.

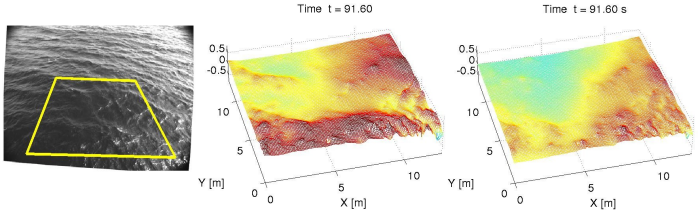

Figure 0.5: Experiment I. Manifold reconstruction at one of the snapshots. Left: a coarse input image of the stereo pair. Center: reconstructed wave height with $\tilde{\rho} = 0.1$. Right: reconstructed wave height with $\tilde{\rho} = 1.0$.

Experiment I

The MR method was tested on a sequence of 4100 consecutive snapshots, split in sub-sequences of 1025 snapshots. A 6-level full multigrid method with 1000 iterations per level and 1500 iterations at the finest level was performed, with 2 V-cycles per iteration and one pre- and post-relaxation sweeps per level. The weights $\alpha = 0.1$ and $\beta = 0.025$ were empirically determined. The initial manifold surface consisted of the zero height surface (for all snapshots). Different values of the parameter ρ were tested: $\tilde{\rho} = \rho h/\Delta\tau = \{0.1, 0.2, 0.5, 1\}$. Observe that the temporal coherence of the MR decreases as $\rho \to 0$. In the limit, $\rho = 0$ is equivalent to the reconstruction of each snapshot independently, using full multigrid on each of them, but with a common artificial time step.

Figs. 0.2 through 0.4 show slices of the surface height and radiance functions obtained by the MR method. The computational grid has $129 \times 129 \times 1025$ points, with spatial and temporal resolutions $h = 10$ and $\Delta\tau = 0.1$ s, respectively. In particular, the results correspond to the case $\tilde{\rho} = 0.1$. Observe the oscillating patterns of the ocean waves in both the height and radiance functions of the u and v slices (Figs. 0.2 and 0.3). The radiance function also captures the location and motion of sea foam caused by breaking waves. Fig 0.4 gives the reconstructed surface height and radiance for a particular snapshot of the sequence.

By visual inspection of the reconstruction (see Fig. 0.5), one concludes that the values $\tilde{\rho} = \{0.5, 1\}$ are too large: temporal derivatives are penalized too much with respect to spatial derivatives, yielding a reconstructed surface shape that is very smooth in time and does not capture the wave patterns present in the stereo video data. Drawing an analogy with linear signal processing, the anisotropic diffusion carried out by the (weighted) Laplacian operator has a low-pass filtering effect: it limits the temporal bandwidth of the output signal, thus reducing noise but also destroying the desired wave signal.

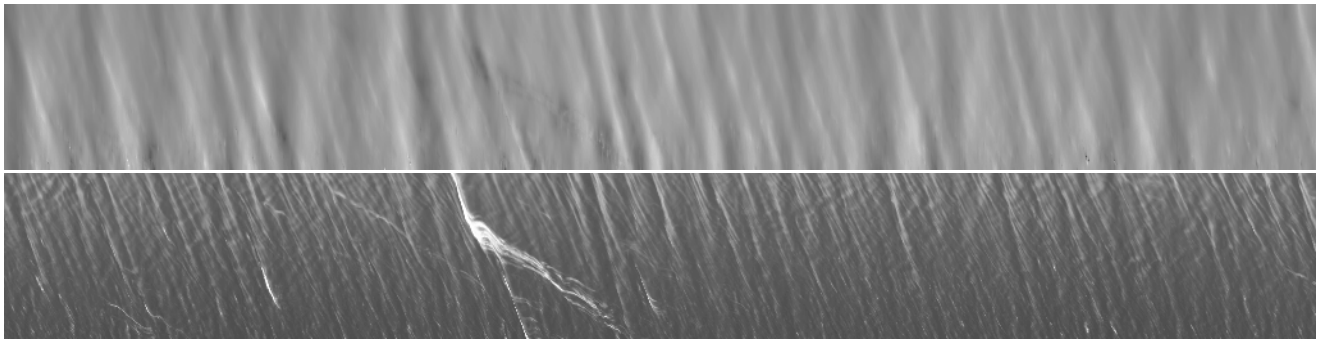

Figure 0.2: Experiment I. A slice at constant $u = u_0$. Top: surface height $Z(u_0, v, \tau)$ (grayscale encoded, from dark (low) to white (high)). Bottom: surface radiance $f(u_0, v, \tau)$. Horizontal axis is time τ.

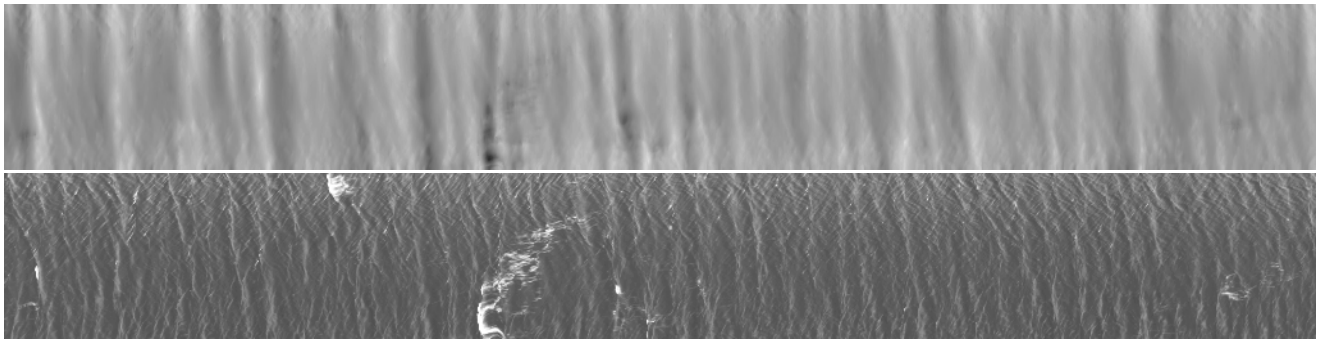

Figure 0.3: Experiment I. A slice at constant $v = v_0$. Top: surface height $Z(u, v_0, \tau)$ (grayscale encoded, from dark (low) to white (high)). Bottom: surface radiance $f(u, v_0, \tau)$. Horizontal axis is time τ.

Using the following formulas to measure the photometric quality of the modeled image given by the reconstructed surface at a single snapshot,

$$\tilde{E}_{\text{data}} = \frac{E_{\text{data}}}{N} \approx \frac{h^2}{N} \sum_{i=1}^{N_c} \sum_{j=1}^{N} \phi_i(\mathbf{u}_j) \mathtt{J}_i(\mathbf{u}_j), \tag{0.10}$$

$$\text{RMS}\left[\frac{E_i}{\text{Area}(\Omega_i)}\right] \approx \left(\frac{\tilde{E}_{\text{data}}}{2} \frac{N}{\text{Area}(\Omega_i)}\right)^{\frac{1}{2}}, \tag{0.11}$$

it is possible to measure the photometric quality of the modeled images given by the reconstructed space-time surface, i.e., for all snapshots. Table 1 compares the photometric error of the reconstructed sequence of wave heights obtained by means of the SR and MR methods with different values of the parameter ρ. Observe that the magnitude of the photometric error is small and very similar for all values of ρ, however the differences in reconstructed surface shape are more noticeable by means of other methods such as the statistics of the wave heights. The case $\tilde{\rho} = 0.1$ is further analyzed.

A piece of the reconstructed space-time volume of wave heights, $V = Z(x, y, t)$, (i.e., $x \equiv u, y \equiv v$ and $t \equiv \tau$) is shown in Fig. 0.6, where the oscillating color pattern encodes the oscillating pattern of the waves. Next, spectral and statistical analysis are applied to validate the reconstruction obtained with the MR method, yielding similar results as those in (Gallego et al., 2011a).

Applying Fourier analysis to the wave space-time volume $Z(x, y, t)$, yields the 3-D spectrum $\mathcal{Z}(k_x, k_y, \omega)$, where frequencies k_x, k_y and ω are the Fourier transformed variables of x, y and t, respectively. Here, $\omega = 2\pi f$ is the angular frequency and $f \equiv f_t$ is the linear frequency, measured in Hertzs. Similarly, $f_x = k_x/(2\pi)$ and $f_y = k_y/(2\pi)$ are the wave numbers in units of cycles/m. The space-time grid sampling yields the Nyquist frequencies $[f_x, f_y, f]_{\max} = [h^{-1}, h^{-1}, \Delta t^{-1}] = [5\,\text{cycles}/m, 5\,\text{cycles}/m, 5\text{Hz}]$. The 3-D spectrum of the reconstructed $129 \times 129 \times 4100$ wave height grid is computed by averaging the 3-D spectra of overlapping pieces of $N_t = 512$ snapshots selected by a Gaussian 3-D window. A 3-D FFT with $N = 512$ points in each dimension is applied to each piece, thus oversampling in space. The spectral spacings are $\Delta f_x = \Delta f_y = 1/(Nh) \approx 0.02\,\text{cycles}/m$ and $\Delta f = 1/(N\Delta t) \approx 0.02\,\text{Hz}$. The 3-D wave spectrum and two of its slices (the frequency-wave number spectra ω-k_x and ω-k_y), are shown in Figs. 0.7 and 0.8, respectively. The white curve in Fig. 0.8 corresponds to planar projections of the linear dispersion manifold in deep water, namely $\sqrt{k_x^2 + k_y^2} = \omega^2/g$. At the Crimean platform, the water depth of 30 m is regarded as being infinite for our range of wave numbers. The shift of the en-

Table 1: Comparison of photometric error (0.11) for several methods used to reconstruct the same coarse stereo image sequence: the variational graph sequential method (SR) (Gallego et al., 2011a) and the variational graph manifold method (MR).

Method	mean	standard deviation
Sequential	3.613	0.398
Manifold, $\tilde{\rho} = 0.1$	3.615	0.393
Manifold, $\tilde{\rho} = 0.2$	3.621	0.391
Manifold, $\tilde{\rho} = 0.5$	3.657	0.392

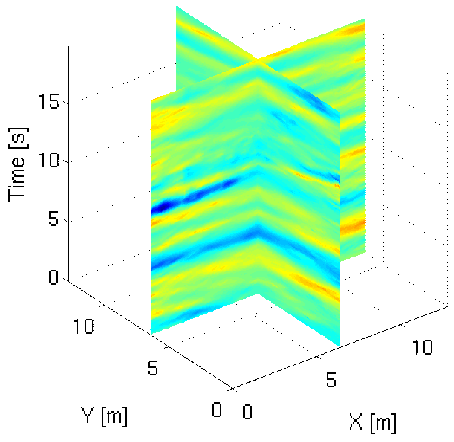

Figure 0.6: Experiment I. Vertical slicing of the wave space-time volume $Z(x, y, t)$ showing the reconstruction for fixed x-t or y-t planes. The wave heights Z have been pseudo-colored from blue (low) to red (high).

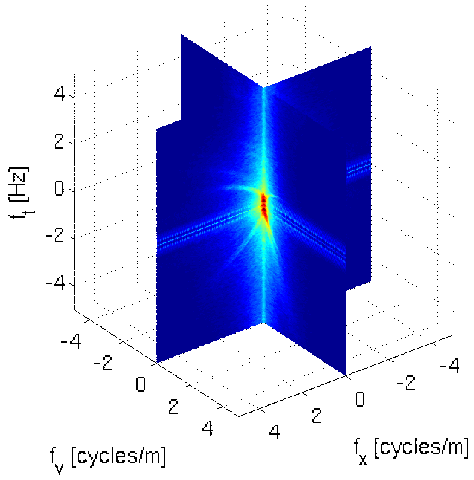

Figure 0.7: Experiment I. Two orthogonal slices of the 3-D wave spectrum $\mathcal{Z}(k_x, k_y, \omega)$ through the frequency origin. Slices correspond to the values of $\mathcal{Z}(k_x, k_y, \omega)$ at planes $k_x = 0$ and $k_y = 0$. Axes f_x, f_y and f_t stand for $k_x/(2\pi)$, $k_y/(2\pi)$ and $\omega/(2\pi)$, respectively.

ergy spectrum from the theoretical dispersion curve (Fig. 0.8) is explained by the presence of a wave-current with velocity vector $\mathbf{v} \approx (-0.17, -0.45)$ m/s (Gallego et al., 2011a). The wave-current dispersion relation is $k = (\omega - \mathbf{k} \cdot \mathbf{v})^2/g$, where $k = |\mathbf{k}| = \sqrt{k_x^2 + k_y^2}$.

Time series of wave heights $Z_i(t) = Z(x_i, x_i, t)$ were extracted from the reconstructed space-time wave volume by selecting virtual probe points (x_i, x_i) in the study area U. The Fourier analysis of the time series leads to the observed Power Spectral Density reported in Fig. 0.9. The tail of the frequency spectrum decays as f^{-4}, which agrees with the tail of the wave-number spectrum decaying as $k^{-2.5}$ (Zakharov, 1999). Finally, the MR method shows an improvement in the estimation of the crest-trough asymmetry of wave heights (see Fig. 0.10).

Experiment II

We tested the MR method on the same image sequence, but with a different grid, $33 \times 33 \times 4097$, maintaining the same spatial and temporal spacings ($h = 10$ cm and $\Delta\tau = 0.1$ s) so that more snapshots are processed simultaneously at the expense of reducing the study area on the surface of the sea. The new region of interest is approximately 16 times smaller (10.89 m^2) than the region in Experiment I (166.41 m^2) and it is located in the center of the highlighted region in Fig. 0.1. The weights α, β, the amount of temporal coherence $\tilde{\rho}$ and the parameters for the full multigrid method are the same as those in Experiment I, except for the number of levels in multigrid (5 instead of 6 due to the size of the new grid).

Fig. 0.11 shows the Power Spectral Density (computed from time series of wave heights of the reconstructed surface), whose tail decays as f^{-4}, as in Fig. 0.9. The larger the observed time interval (410 s vs. 102.5 s), the better waves are resolved in the Fourier domain, as a consequence of the uncertainty principle (localization in time vs. localization in frequency).

Because the study area is small compared to the observed wavelengths, all wave heights corresponding to the same snapshot move in unison, hence there is a high degree of correlation among the time series. This illustrates the fact that, in the limit, as the spatial grid size decreases and the temporal grid size increases, the MR method acts as a remote sensing technology to extract coherent wave heights at a *single* probe point. Note, however, the flexibility in the selection of the size of the 3-D grid in the MR method to handle different situations, exchanging spatial and temporal resolutions, depending on whether we are interested in resolving better spatial or temporal characteristics of the waves.

CONCLUSION AND FUTURE WORK

A variational graph manifold method for the space-time coherent reconstruction of ocean waves has been discussed and developed. Due to the convenient representation of the ocean surface as a height function, the incorporation of a temporal dimension is straightforward. This benefit of the graph representation makes the reconstruction problem significantly simpler than what it would be if the level set framework was used instead. The manifold method has been tested on stereo video data from real ocean waves at an off-shore platform in the Black Sea. The qualitative and photometric performance of the method has been demonstrated on a coarser version of the dataset due to memory limitations of the implementation with respect to the size of the original dataset. By design, the manifold reconstruction method is more robust than the sequential method, at the expense of speed. In addition, it captures better some physical properties of wave heights than the sequential method, such as the crest-trough asymmetry of wave heights. The variational graph manifold method can be used to study both large and small areas of the surface of the ocean, exchanging spatial and temporal localization goals while maintaining coherence of the estimated wave heights.

Evidence shows that the incorporation of the physics and coherence in our variational method produces tangible improvements and encourages us to continue with this line of research to achieve better results. These insights justify the research on the assimilation of the wave equation in the reconstruction process. Finally, the manifold view developed in this paper can also be applied to the variational disparity method in (Alvarez et al., 2002). This

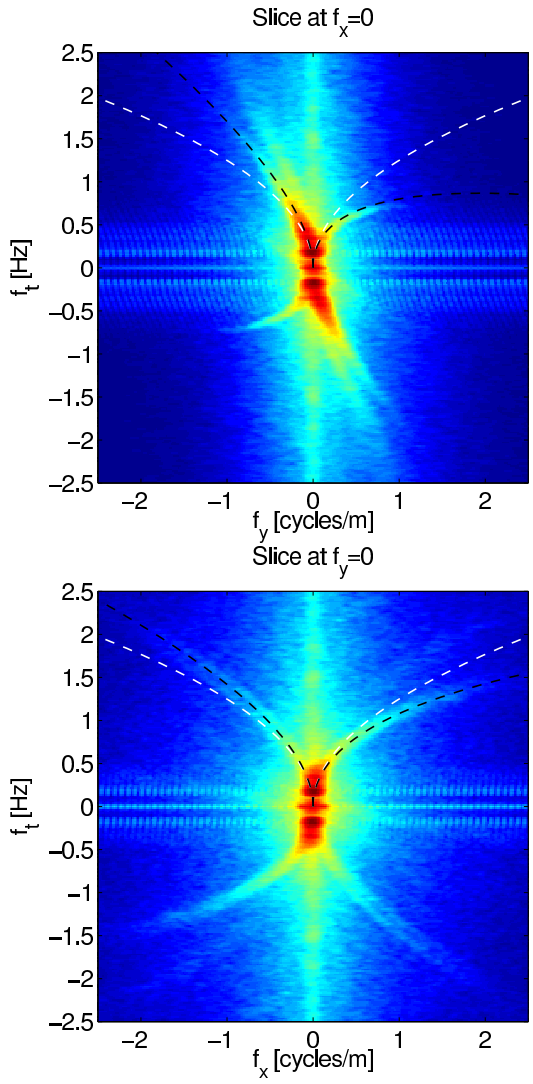

Figure 0.8: Experiment I. Slices of the 3-D wave spectrum at $k_x = 0$ (top) and $k_y = 0$ (bottom). Superimposed on top half of both plots: (white curve) vertical slice of the linear wave dispersion manifold $|\mathbf{k}| = \omega^2/g$, with $\omega = 2\pi f_t$, and (black curve) vertical slice of the wave-current dispersion manifold $|\mathbf{k}| = (\omega - \mathbf{k}\cdot\mathbf{v})^2/g$, with $\mathbf{v} \approx (-0.17, -0.45)$ m/s. Axes f_x, f_y and f_t stand for $k_x/(2\pi)$, $k_y/(2\pi)$ and $\omega/(2\pi)$, respectively.

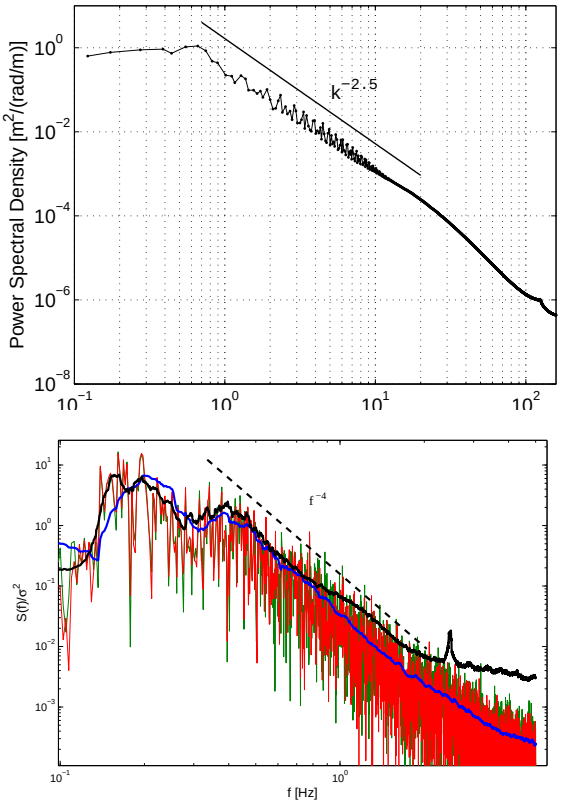

Figure 0.9: Experiment I. Top: Omni-directional wave number spectrum. Bottom: Normalized frequency spectrum (σ^2 is the variance of the wave surface) averaged over several virtual probes in time (blue line) and compared to the spectrum by the epipolar method (black line) (Benetazzo, 2006).

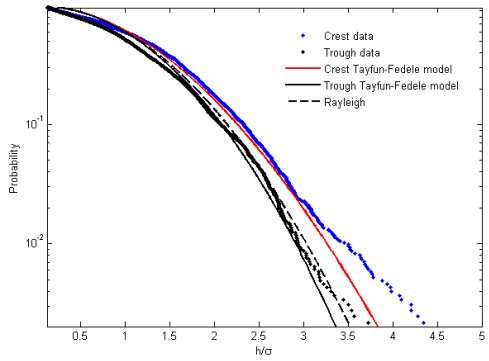

Figure 0.10: Empirical exceedance probabilities of wave crests and troughs from time series extracted at virtual probes. Rayleigh and Tayfun-Fedele theoretical models are also shown for comparison.

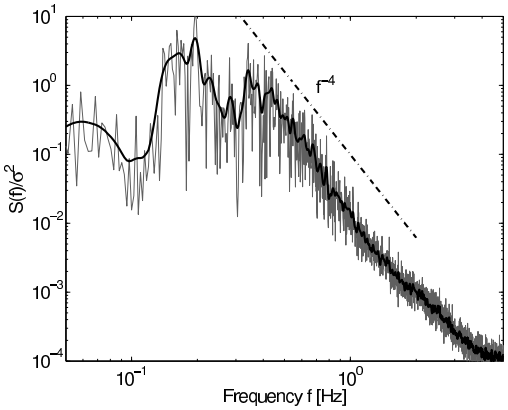

Figure 0.11: Experiment II. Normalized frequency spectrum averaged over several virtual probes in time. The tail of the frequency spectrum decays as f^{-4}.

topic may be further investigated in the future.

ACKNOWLEDGMENTS

Research supported by ONR grant BAA 09-012: "Ocean Wave Dissipation and energy Balance (WAVE-DB): toward reliable spectra and first breaking statistics". G. Gallego is supported by the Marie-Curie COFUND Programme of the EU.

REFERENCES

Alvarez, L., Deriche, R., Sánchez, J., & Weickert, J. (2002). Dense disparity map estimation respecting image discontinuities : A pde and scale-space based approach. *Journal of Visual Communication and Image Representation*, 13, 3–21.

Astruc, D. et al. (2012). A stereoscopic method for rapid monitoring of the spatio-temporal evolution of the sand-bed elevation in the swash zone. *Coastal Engineering*, 60, 11–20.

Bechle, A. J. & Wu, C. H. (2011). Virtual wave gauges based upon stereo imaging for measuring surface wave characteristics. *Coastal Engineering*, 58(4), 305–316.

Benetazzo, A. (2006). Measurements of short water waves using stereo matched image sequences. *Coastal Engineering*, 53(12), 1013–1032.

Briggs, W. L., Henson, V. E., & McCormick, S. F. (2000). *A Multigrid Tutorial, Second Edition*. SIAM.

de Vries, S. et al. (2011). Remote sensing of surf zone waves using stereo imaging. *Coastal Engineering*, 58(3), 239–250.

Fedele, F., Benetazzo, A., & Forristall, G. (2011). Space-time waves and spectra in the northern adriatic sea via a wave acquisition system. In *ASME 30th Int. Conf. OMAE*. Rotterdam, The Netherlands.

Fedele, F., Gallego, G., Yezzi, A., & Benetazzo, A. (2008). Understanding extreme waves via a variational wave acquisition stereo system. In *Workshop ROGUE WAVES*. Oct 13-15, Brest, France.

Gallego, G. (2011). *Variational Image Processing Algorithms for the Stereoscopic Space-Time Reconstruction of Water Waves*. PhD thesis, Georgia Institute of Technology, Atlanta, GA, USA. Directors: Yezzi, A. and Fedele, F.

Gallego, G., Yezzi, A., Fedele, F., & Benetazzo, A. (2011a). A variational stereo method for the three-dimensional reconstruction of ocean waves. *Geoscience and Remote Sensing, IEEE Trans.*, 49(11), 4445–4457.

Gallego, G., Yezzi, A., Fedele, F., & Benetazzo, A. (2011b). A variational wave acquisition stereo system for the 3-d reconstruction of oceanic sea states. In *ASME 30th Int. Conf. OMAE*.

Holland, K., Holman, R., Lippmann, T., Stanley, J., & Plant, N. (1997). Practical use of video imagery in nearshore oceanographic field studies. *Oceanic Engineering, IEEE Journal of*, 22(1), 81–92.

Hsiao, Y.-H. & Huang, M.-C. (2009). Application of active contour model in tracking sequential nearshore waves. *China Ocean Engineering*, 23(12), 251–266.

MacHutchon, K. & Liu, P. (2007). Measurement and analysis of ocean wave fields in four dimensions. In *ASME 26th Int. Conf. OMAE*, volume 1 (pp. 923–927).

Mumford, D. & Shah, J. (1988). Optimal approximations by piecewise smooth functions and variational problems. *Comm. Pure Appl. Math*, XLII(5), 577–685.

Santel, F., Linder, W., & Heipke, C. (2004). Stereoscopic 3d-image sequence analysis of sea surfaces. *Proc. of the ISPRS Commission V Symposium*, 35(5), 708–712.

Schumacher, A. (1939). Stereophotogrammetrische wellenaufnahmen. wiss ergeb. dtsch. atlant. exped. forschungs vermessung. meteor 1925-1927. *Ozeanographische Sonderuntersuchungen, Erste Lieferung*.

Shemdin, O., Tran, H., & Wu, S. (1988). Directional measurements of short ocean waves with stereography. *J. Geophys. Res.*, 93, 13891–13901.

Trottenberg, U. (2000). *Multigrid*. Academic Press.

Wanek, J. M. & Wu, C. H. (2006). Automated trinocular stereo imaging system for three-dimensional surface wave measurements. *Ocean Engineering*, 33(5-6), 723–747.

Zakharov, V. E. (1999). Statistical theory of gravity and capillary waves on the surface of a finite-depth fluid. *Eur. J. Mech. B - Fluids*, 18(3), 327–344.