Coastal Engineering 109 (2016) 114–127

Contents lists available at ScienceDirect

Coastal Engineering

journal homepage: www.elsevier.com/locate/coastaleng

Stereo wave imaging from moving vessels: Practical use and applications

Alvise Benetazzo [a,*], Francesco Barbariol [a], Filippo Bergamasco [b], Andrea Torsello [b], Sandro Carniel [a], Mauro Sclavo [a]

[a] Institute of Marine Sciences, Italian National Research Council (ISMAR-CNR), Venice 30122, Italy
[b] DAIS - Università Ca' Foscari, Venice, Italy

ARTICLE INFO

Article history:
Received 11 June 2015
Received in revised form 27 November 2015
Accepted 10 December 2015
Available online 12 January 2016

Keywords:
Space–time sea wave data
Stereo wave imaging
Moving vessel
Wave spectra

ABSTRACT

Stereo wave imaging of the sea surface elevation has become an effective instrumentation to gather small- and medium-range 3-D wind wave data. Indeed, fruitful applications of stereo techniques have provided new insights into directional wave spectra, space–time distributions of wave maxima, and small-scale wave statistics. So far, however, stereo systems have been deployed mainly on fixed structures (e.g. oceanographic platforms or lighthouses) in order to simplify the installation and maintenance procedures. Nonetheless, advances in stereo calibration and processing suggest that stereo deployments are also feasible onboard moving vessels, thus broadening the impact of these observations on the study of wind waves. In this context, this study aims at discussing how the stereo processing designed to gather reliable wave data from fixed structures should be managed to operate on a moving structure. In particular, estimate of stereo cameras orientation and position with respect to the mean sea plane is of utmost importance. We discuss this aspect by using a synthetic sea state and stereo data collected during an oceanographic campaign onboard a research vessel. Results suggest that, without complementary data sources for ship motion compensation, the sea surface elevation field should include at least about sixteen spatial (2-D) waves to gather a robust estimate of the mean sea plane and consequently realistic wave parameters (e.g. the significant wave height). In this respect, our results provide also insight into the uncertainty of estimates in case of a limited number of 2-D waves is collected by the stereo system. Finally, applications of stereo wave imaging on a moving structure are discussed, with particular emphasis on the collection of space–time wave fields for assessment of numerical models and operational wave observation onboard vessels.

1. Introduction

Statistical and spectral properties of wind waves are typically inferred from time records of sea surface elevations retrieved from instruments (like buoys or wave gauges) installed at fixed locations of the ocean. These observatories have provided unique datasets that have been extensively used over the years by generations of oceanographers and engineers. However, the information content of a single time series does not accurately predict the complete wave dynamics, which must be assumed as developing over the 2-D space as well as time (Boccotti, 2000). In this respect, in the recent past new classes of instrumentations (e.g. radars or lidars) have started to provide sufficient resolution and accuracy for measuring waves at different spatial scales, usually larger than some meters (Hwang et al., 2000; Nieto Borge, 2004; Romero and Melville, 2010). At smaller scales, however, where most of the air–sea exchanges occur, the optical systems (e.g. Jähne and Riemer, 1990; Zappa et al., 2008) have proved to gather sea elevations spatial data with higher accuracy. In this context, stereovision systems have started to gain credit as a tool to collect accurate 3-D fields of sea surface elevations. Starting from the pioneering studies of Schumacher (1939) and subsequent applications (e.g., Banner et al., 1989), thanks to a noticeable merging of image analysis techniques and available computational resources, stereo analysis has become only in the recent years a well explored technique for measuring sea waves remotely (Benetazzo et al., 2012; de Vries et al., 2011; Gallego et al., 2011; Kosnik and Dulov, 2011; Liu, 2013; Mironov et al., 2012). As mentioned, the added value of the stereo systems is the possibility to gather 3-D wave fields as they evolve in time, thus providing inputs to deepen the scientist knowledge on how the surface waves actually behave when they are treated as space–time fields (e.g. Banner et al., 2014; Benetazzo et al., 2015).

So far, however, stereo systems have been mostly installed on fixed platforms or piers over the sea. These conditions greatly ease the calibration and mounting procedure, as well as the entire processing necessary to get accurate sea waves information. Two stages have always been considered as critical when using stereo cameras at sea. Firstly, the computation of the so-called external parameters (Ma et al., 2004) that provide, for a given stereo setup, rotation and displacements

* Corresponding author at: Institute of Marine Sciences (ISMAR), Italian National Research Council (CNR), Arsenale Tesa 104, Castello 2737/F, 30122 Venice, Italy.
E-mail address: alvise.benetazzo@ve.ismar.cnr.it (A. Benetazzo).

http://dx.doi.org/10.1016/j.coastaleng.2015.12.008
0378-3839/

between the cameras (generally two) is crucial for field applications in which it may be unfeasible to take apart or even physically access the device. Then, the pose (i.e. the orientation and position in 3-D space) of the stereo-camera system with respect to the mean sea surface must be accounted for to attain a proper space–time sea surface representation. For deployment on fixed structures, it has been verified (Benetazzo, 2006) that the mean sea plane can be accurately determined by a time-averaging procedure of the planes best-fitting the 3-D wave field mapped in the camera reference frame. When installing the stereo system on a moving structure, however, the 3-D mapping onto the sea reference is much more complicated, since this averaging is not feasible in a straightforward manner. A different strategy, therefore, must be adopted to transform the stereo 3-D data onto a reference system consistent with the horizontal mean sea plane. This topic was partially resolved, for instance, by Brandt et al. (2010) using the horizon visible in each frame to track the stereo cameras rotational motion, and by Schwendeman and Thomson (2015) who developed a horizon-tracking method for shipboard video stabilization and rectification.

In this paper we study how robust is the mean sea plane estimation when the plane orientation is determined using the 3-D data only (i.e. without complementary data sources for motion compensation), a condition that would greatly simplify the stereo processing. For the analysis reported in the study we have taken advantage of a Wave Acquisition Stereo System (WASS; Benetazzo, 2006) that was deployed onboard a vessel during a research cruise (Section 3.2). Preliminarily (Section 3.1), sea elevation data from a synthetic sea state have been used to asses to what extent the wave parameters (as the significant wave height) are well determined when a limited portion of the sea surface is retrieved by the stereo system. With reference to the wave spectral moments, this problem has already been investigated on time series and spatially distributed data by Krogstad et al. (1999). Section 2 of the paper reports the recent improvements of the WASS pipeline with respect to the layout described in Benetazzo et al. (2012). The developments proposed in this study allow getting more accurate and resolved 3-D data of the sea surface elevation. The study is completed (Section 4) with a detailed analysis of the possible uses of stereo systems mounted onboard research vessels and ships of opportunity.

2. The WASS observatory

WASS is an optical-based system used to collect space–time data of sea surface elevations. It relies upon a pair of high-resolution digital cameras, which, once synchronized, allow the sea waves to be observed from two distinct points of view. With respect to the layout described in Benetazzo et al. (2012), WASS has been improved to ease the installation and calibration phases, which are critical to get accurate measurements. Such developments are described in the following sections.

2.1. Calibration: intrinsic parameters and recovery of the stereo camera pose

Intrinsic parameters of each camera composing the stereo system are calibrated using a hand crafted known target (i.e., a chessboard). Since we expect such target being generally affected by some imperfections (i.e. printing misalignments, small bumps or glitches) we implemented the method described in Albarelli et al. (2010) that suggest a bundle adjustment step to optimize both camera parameters and target geometry. Each camera is therefore calibrated independently by acquiring ~50 snapshots of the target with different orientations and distances from the camera, spanning a space about 3 m depth and 5 m wide in front of it. All the parameters are estimated by imposing zero skewness, square pixels, and a five coefficient polynomial radial distortion model.

The estimation of the intrinsic parameters is not enough to perform the stereo reconstruction from a pair of images. In fact, the reciprocal position of the two cameras (the so-called extrinsic parameters) must be provided to recover the full geometry of the scene through triangulation. The extrinsic parameters define the displacement $\boldsymbol{\tau}$ and the rotation R between the left and right camera frames according to the Euclidean transformation $g = (R, \boldsymbol{\tau})$. In previous WASS deployments, the rigid motion g was estimated by exposing an ad-hoc calibration target to both cameras, and by relating the known 3-D geometry of the target with its re-projection onto the image planes. However, even if this is the standard de-facto way to calibrate a stereo rig in laboratory conditions, this approach manifests several drawbacks when applied to stereo systems with large baseline.

At first, since for field applications we usually require a baseline $\boldsymbol{\tau}$ between cameras larger than 2 m, the target size has to be wider than $1 \times 1\ m^2$, and placed at a distance greater than about 5 m from the cameras. Due to the target size, the manufacturing process may lead to some coarse imperfections and allowing the protrusion of such target meters away from the vessel hull can be time consuming or even dangerous. Moreover, the calibration procedure is time intensive and requires taking apart the device from its working position. As such, it is very difficult to modify the system geometry on-the-fly to accommodate different acquisition requirements. For instance, it may be reasonable to take the device closer to the sea when the waves are slight, so a small but highly resolved sea surface region can be acquired. On the other hand, large waves demand a broader area, requiring the device to be repositioned farther from the surface. Finally, the "calibrate once and for all" strategy is not reliable since vibrations of the support and environmental factors, as wind, can modify the relative angle between cameras and jeopardize the reconstruction accuracy.

To overcome these limitations, we developed a calibration procedure that relies on the photometric consistency of the sea surface itself, and thus can be carried out during the acquisition without the need of a calibration target. Specifically, it is well known that it is possible to estimate the relative pose of two cameras, up to scale, from a set of corresponding points between the two images (Fig. 1).

Therefore, taking some points (in homogeneous coordinates) $p_1 \ldots p_n$ extracted from a frame captured by the first camera (say left), and the corresponding set $p'_1 \ldots p'_n$ by the second (say right), the epipolar constraint can be exploited to estimate the essential matrix M such that (Ma et al., 2004)

$$p_i M\, p'_i = 0, \ \forall i = 1 \ldots n\,. \tag{1}$$

Moreover, the essential matrix can be decomposed through singular value decomposition to recover the rigid motion g up to a scale factor for $\boldsymbol{\tau}$ (Hartley and Zisserman, 2004). While conceptually simple, determining such corresponding points can be a difficult task particularly when dealing with un-textured areas or repetitive patterns. Not surprisingly, sea surface is not rich of distinctive features so special care has to be taken to let this process be as reliable as possible. In the stereo pipeline, the extrinsic calibration process starts with the extraction of a set of Speeded Up Robust Features (SURF; Bay et al., 2008) from each image. The algorithm is set to use 3 octaves, 8 intervals per octave and a blob response threshold of 10^{-4}. To obtain a more uniform spatial distribution of interest points in a frame, the image is divided in 16 blocks and points with lower hessian response are iteratively removed from each block to finally collect a set of 2600 features for each image. From these feature points, orientation, scale and a 64-component descriptor are computed.

Since most of these points are located on high textured areas (waves crest and white capped areas), the descriptor itself is not sufficient to establish a reliable set of matches between left and right camera features. Indeed, the local information around each point is not distinctive when dealing with a surface that shows uniform color, smooth shading, and clear but repetitive white areas. To guarantee a good set of point-to-point correspondences, we implemented the state-of-the-art method proposed in Albarelli et al. (2012). The key idea is that, for small motions, the transformation between stereo images that affects a group