Computers & Geosciences 95 (2016) 105–117

Contents lists available at ScienceDirect

Computers & Geosciences

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

Research paper

# Multi-view horizon-driven sea plane estimation for stereo wave imaging on moving vessels

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

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

ARTICLE INFO

*Article history:*
Received 31 January 2016
Received in revised form
21 July 2016
Accepted 22 July 2016
Available online 25 July 2016

*Keywords:*
Stereo vision
Horizon estimation
Sea surface reconstruction
3D wave field

ABSTRACT

In the last few years we faced an increased popularity of stereo imaging as an effective tool to investigate wind sea waves at short and medium scales. Given the advances of computer vision techniques, the recovery of a scattered point-cloud from a sea surface area is nowadays a well consolidated technique producing excellent results both in terms of wave data resolution and accuracy. Nevertheless, almost all the subsequent analyses tasks, from the recovery of directional wave spectra to the estimation of significant wave height, are bound to two limiting conditions. First, wave data are required to be aligned to the mean sea plane. Second, a uniform distribution of 3D point samples is assumed. Since the stereo-camera rig is placed tilted with respect to the sea surface, perspective distortion do not allow these conditions to be met. Errors due to this problem are even more challenging if the optical instrumentation is mounted on a moving vessel, so that the mean sea plane cannot be simply obtained by averaging data from multiple subsequent frames. We address the first problem with two main contributions. First, we propose a novel horizon estimation technique to recover the attitude of a moving stereo rig with respect to the sea plane. Second, an effective weighting scheme is described to account for the non-uniform sampling of the scattered data in the estimation of the sea-plane distance. The interplay of the two allows us to provide a precise point cloud alignment without any external positioning sensor or rig viewpoint pre-calibration. The advantages of the proposed technique are evaluated throughout an experimental section spanning both synthetic and real-world scenarios.



## 1. Introduction

In the recent years there has been a growing interest in remote and proximal observation of sea surface waves. New classes of instruments (e.g. radars, laser scanners, optical) have been placed aboard satellites, airplanes, or ships at sea, facing the need for new processing tools for specific tasks. In this context, observations of sea surface fields by means of optical-based stereo systems, after the pioneering work of Schumacher (1939), Holthuijsen (1979), and Banner et al. (1989), are nowadays becoming accessible to a large number of scientists given the advances in the processing of digital images (Kosnik and Dulov, 2011; Gallego et al., 2011; Benetazzo et al., 2012, 2015). The efforts made to develop both hardware and software for stereo processing has allowed stereo wave imaging to become a well consolidated and accurate tool for wave observations, that indeed is providing new insights into the field of sea surface waves (Banner et al., 2014; Leckler et al., 2015; Yurovskaya et al., 2013). So far, snapshots of the sea surface have been mostly taken from fixed platforms at sea, as this condition greatly eases the cameras deployment and the image processing. Indeed, the time-constant rigid-motion that places all the acquired surfaces into a common geo-referenced frame can be estimated a priori exploiting the statistical nature of the waves (Benetazzo et al., 2016). However, this design for stereo systems is starting to become too restrictive for oceanographers, who need to collect wave data in as many different conditions and locations as possible, eventually where the use of a fixed platform is unfeasible. Thus, first applications of stereo systems mounted on oceanographic vessels is taking place, facing the nontrivial task of accounting for the camera motion during the stereo processing (Brandt et al., 2010).

In this study we present an analysis of spatio-temporal 3D wave fields that take advantage of a multi-view horizon estimate approach for compensating ship motion. The paper is organized as follows. We start with an overview of the related work in the field of horizon estimation, highlighting the novelties of our proposed

* Corresponding author.
*E-mail address:* filippo.bergamasco@unive.it (F. Bergamasco).

http://dx.doi.org/10.1016/j.cageo.2016.07.012
0098-3004/

approach with respect to the literature. Then, in Section 2 we describe the stereo system used to reconstruct the sea surface. In Section 3 we present our multi-view horizon estimation approach to recover the sea surface plane normal. Then, in Section 4 we describe how to use the recovered normal to estimate a homography relating to the camera and sea plane. Such homography is used to define a spatial density estimator to weight the 3D point samples and account for perspective distortion. This allows a robust recovery of plane distance and, consequently, a precise alignment of the scattered point data. An extensive experimental part (Section 5) demonstrates the strengths of our method comparing with both synthetic and real-world data acquired in a controlled environment (i.e. a fixed oceanographic platform). Furthermore, a use-case is presented to show the behaviour of the approach "in the wild" while investigating the wave statistics on a moving ship during a cruise. Finally, in Section 6 we draw some concluding considerations.

### 1.1. Related work

Being able to recover the attitude and position of a moving vehicle subject to 6 degrees of freedom has been a topic of pivotal importance for both aerial and marine operations. With the former, the recent availability of inexpensive unmanned aerial vehicles to be used in dangerous or remote situations raised the need to improve the automatic control and stabilization beyond the accuracy obtainable with purely inertial sensors. In the latter, the precise estimation of ship motion is an aid for navigation safety (Liu et al., 2008), vessel detection (Fefilatyev et al., 2007) or environmental sciences (Williams and Howard, 2011).

In this plethora of different applications, vision based approaches have gained an increased popularity due to their potential ability to hinge the estimation on well-localizable visual features occurring on the landscape. Among all, the horizon line is usually the weapon of choice as is expected being always visible if no other assumptions can be made to the application scenario. Even if seeing the horizon is not a sufficient condition to recover the full 6 degrees of freedom of a vessel (the estimation is recovered up to any rotation around an arbitrary axis orthogonal to the ground plane), position can be usually obtained with enough accuracy through GPS sensors and is less critical with respect to attitude recovery. In fact, by constraining to a physically plausible kinematics of a ship, statistical filtering methods can improve the accuracy of common marine GPS to tens of centimeters. Given the importance of the horizon as a visual feature, sophisticated techniques have been proposed to provide a precise estimation even when the acquired images are affected by cluttering objects or optical artifacts (sun-glares, fog, low-contrast, etc.).

Following the taxonomy introduced by Shabayek et al. (2012), horizon line estimation methods are essentially divided into two categories. In the former, a Sky/Ground segmentation is performed by means of statistical models over the feature space being either color (Gallagher et al., 2004), textures (Cherian et al., 2009; Schwendeman and Thomson, 2015), polarization (Shabayek et al., 2012) or a combination of the previous (Todorovic et al., 2003). In the latter, the horizon line is directly estimated by extracting significant image edges (usually after a sequence of multiple filtering stages) and a statistically robust aggregation of contour segments composing the horizon or localizing vanishing points. For instance, Wang et al. (2009) propose an horizon extraction method from ocean observations by weighting the color channels to enhance sky and sea signals while reducing the effect of noise and highlights. Thereafter, contours are extracted with Canny edge detection and the predominant line clustered with a classical Hough transform. Similarly, Williams and Howard (2011) describe a rather sophisticated sequence of adaptive histogram thresholding, region growing and Gaussian mixture model to preprocess the images before binarization and contour extraction.

When horizon estimation is used to control an UAV, temporal coherency can be exploited to aid the extraction considering multiple video frames (Bao et al., 2003). Moreover, optical flow can be extracted between subsequent images as an aid to cluster horizon edges from spurious clutter arising from natural or human-made features on the landscape (Dusha et al., 2007). When an inertial measurement unit (IMU) is available, its data can be used as an initial prior to limit the image analysis on a reduced area of the image (Hugues et al., 2014). Furthermore, Angelino et al. (2013) suggest the usage of an Unscented Kalman Filter (UKF) to describe the dynamic system non-linear equations to effectively fuse attitude information coming from GPS, IMU and cameras.

If stereo vision is available, the complete pose can be obtained as a natural consequence of multi-view triangulation. For instance, Wang et al. (2005) use stereo vision to recover the full position and attitude of a UAV flying near the ground whereas Eynard et al. (2010) use an hybrid stereo system composed of a central and a fish-eye camera to measure the altitude without any external sensor. However, stereo reconstruction can only be used for pose estimation if we assume a non-moving set of landmarks being present in a scene. This is obviously not the case when observing the continuously changing sea surface.

### 1.2. Contributions

In this paper we propose a novel horizon estimation technique to recover the stereo rig attitude with respect to the sea surface. Once the relative orientation of the rig and the sea plane is assessed (up to a rotation around the plane normal, that cannot be inferred by the horizon alone), its distance is estimated in a robust way that compensates the non-uniform distribution of all the reconstructed point samples.

With respect to the aforementioned existing approaches, our method introduces several novelties. To our knowledge, it presents the first intrinsically multi-view approach being able to fuse the horizon line estimates coming from different images. In our case, just the two images of the stereo rig are used but the method itself is flexible enough to comprise many different (pre-calibrated) image sources. Second, we use a prior given by the 3D reconstruction as a first guess to refine the estimated horizon line. In this sense, our method is more like an horizon refinement approach rather than a full estimator being able to locate the horizon anywhere in the image. Nevertheless, this prior allows us to pose the problem as an energy optimization spanning the entire image domain and avoiding voting schemes such as the commonly used Hough Transform. This both improves the accuracy of the method and reduces the computational effort. Third, our method is not based on a preliminary edge extraction step. Instead, it directly operates on image pixels in a way similar to active contours (Kass et al., 1988). This has the double benefit of getting rid of multiple edge extraction parameters and allowing a better line localization since no information is loss due to binarization.

The horizon line allows to recover the parameters defining the normal of the sea plane. This is of crucial importance for the alignment of the reconstructed data, that can be directly rotated so that the sea plane normal coincides with the $z$-axis. Also, we can exploit the relative angles to correct for perspective distortion in an optimal way to account for the non-uniform sampling of the reconstructed points with respect to the sea plane. This non-uniformity is typical of all the stereo methods that operate by matching corresponding pixels directly on the image space. These methods are the only feasible on a moving scenario (like the one investigated in this work) for which the pose of the cameras from the observed scene cannot be reliably provided a priori. On the