Supporting Information for "Directional Breaking Kinematics Observations from 3D Stereo Reconstruction of Ocean Waves"

Bernard Akaawase[1], Leonel Romero[1], Alvise Benetazzo[2]

[1]Department of Marine Science, University of Connecticut, Groton, CT, USA

[2]Institute of Marine Science, Italian National Research Council (ISMAR-CNR), Italy

1. Introduction

In this material, we have provided more information about the Wave Acquisition Stereo System (WASS) processing, wave-breaking detection, and Particle Image Velocimetry (PIV). Additionally, 3 salient points presented in the manuscript are further expatiated on: 1) We show a one-day time series of wind conditions preceding this experiment, 2) we demonstrate the adaptive tapering procedure engaged during spectral analysis on the trapezoidal-shaped wave measurement area. The trapezoidal shape is a result of projecting from the image coordinate system to the local cartesian frame. 3) Show the impact of camera motion correction on the mean sea level displacement between consecutive images.

2. Wave Acquisition Stereo System (WASS)

The WASS pipeline is comprised of four C++ executables namely, WASS-prepare, feature-match, auto-calibrate, and WASS-stereo. The WASS-prepare fetches the corresponding image pairs from the stereo-storage alongside the intrinsic parameters and the configuration information into sequential folders for processing. It also removes the radial and tangential distortion effect based on the provided intrinsic parameters from the images. WASS-match and auto-calibrate are used to determine the camera pose (extrinsic parameters) and the matrix relating to the corresponding points detected in the stereo images. Auto-calibrate is only engaged if the user does not provide the extrinsic parameters. The camera pose is described by both the translation and rotation matrices which are recovered through correspondence based on epipole constraints.

The final step of WASS processing is the WASS-stereo. During this step, the algorithm uses the intrinsic and extrinsic parameters to align the two-camera coordinate system to an ideal stereo plane, such that the lines of epipole in corresponding images become parallel and the complex 3D search-problem is reduced to 2D. Such that the corresponding points within the image pairs lie within the same scan lines. Next, the disparity map is calculated between the two perspectives, and the triangulation of matched pixels is executed to obtain the 3D point clouds. The sea plane coefficients are estimated by combining Random Sample Consensus (RANSAC) and least square fitting. The sea surface anomalies are obtained by removing the mean sea plane from the 3D reconstructed points. Following Bergamasco et al. (2017), the reconstructed surface elevation in the world reference frame $\eta(x, y)$ was mapped to the camera sensor coordinates $\eta(i, j)$. Where i and j are the

horizontal and vertical image pixels indices, respectively. Figure 1a shows 3D point clouds of the stereo triangulation superimposed to the left camera image of the stereo imaging system.

As with every stereographic processing, uncertainties are expected. These include calibration errors which is related to the inaccuracy of getting the exact internal and external calibration matrices. Also, features matching error which originates from the difficulty of matching the corresponding pixel. Several checks have been incorporated into WASS to minimize these errors. The most concerning is the resolution or quantization error, which is the error associated with recovering the 3D coordinates of the reconstructed cloud points due to the increasing area within the footprint of a pixel or sub-pixel. The expected error increases with distance from the cameras along the (y) axis. For our setup, the maximum expected error is ±3cm. The errors are much smaller for locations near the cameras and tend to increase away from the camera.

3. Breakers Detection in Visible Imagery

3.1. Image background removal

The 30-minute stereo dataset analyzed in this study was acquired under a relatively steady temporal illumination. However, there was a slight spatial gradient across frames. The spatial non-uniformity in the images is corrected with equation (1) using the pixel-wise division approach presented in Kleiss (2009) for mean background intensity S(x,y) removal.

$$I'(x,y) = \frac{I(x,y)\overline{S}}{S(x,y)}, \tag{1}$$

where $I'(x, y)$ is the corrected image, $\overline{S}$ is a scalar corresponding to the mean spatial intensity of the image, and $I(x, y)$ is the undistorted image. The background of each image was obtained from a morphological opening operation utilizing a diamond-shaped kernel of 300×300 pixels, this is different from Kleiss 2009, who did temporal averaging. The image correction helps prevent false identification of whitecaps by enhancing the contrast of the dim sections of the image. It is worth mentioning that we excluded the sky and the Acqua Alta structure from the spatial mean intensity calculation.

3.2. Brightness Thresholding

To compute kinematics, we first identify wave-breaking candidates via brightness thresholding in images and track them across frames. The brightness threshold was determined analogous to Kleiss and Melville (2011), see Figure S3. The brightness probability distribution function $p(I)$ was computed from all corrected images, mapping relative intensity I from 0.0 (darkest) to 1.0 (brightest). The portions of images above a threshold intensity (I_t) are determined using the complementary cumulative distribution function $W(I_t)$ given by

$$W(I_t) = 1 - \int_0^{I_t} p(I)\, dI. \tag{2}$$

$p(I)$ was discretized with a resolution $\Delta I = 0.01$. The natural logarithm of equation (2), here defined as $L(I_t) = ln[W(I_t)]$ was differentiated twice to obtain a curvature

$$L''(I_t) = \frac{d^2 L}{dI^2}. \tag{3}$$

The resulting curvature $L''(I_t)$ was used to estimate the brightness threshold ($I_t = 0.55$) as the end of the positive curvature, which is defined as the point that first falls below 20%

of the curvature peak value (see Figure S3). Note, that the 20% is not arbitrary. Similar to Kleiss and Melville (2011) using the curvature's peak value included false breakers.

The potential breakers are obtained by using the determined brightness threshold to contour the corrected images (lens distortion and background lighting removal) of the left camera. The brightness threshold is used to contour the corrected images of the stereo left camera. Before contouring, areas brighter than the threshold (I_t) are smoothed using a Gaussian blur filter with a 3×3 kernel, similar to Sutherland and Melville (2013).

4. PIV Processing

We calculated the wave-breaking kinematics between consecutive image pairs via a MATLAB Particle Image Velocimetry (PIV) toolbox (Thielicke and Sonntag, 2021) which track visible foam or bubbles in combination with the 3D topography information. The tracking of features is performed through normalized cross-correlation. In this work, we configured PIVlab to run four (4) passes of window sizes 256, 128, 64 and 32 pixels with an overlap of 50% for each window. This window selection is capable of resolving speeds up to 30ms^{-1}.

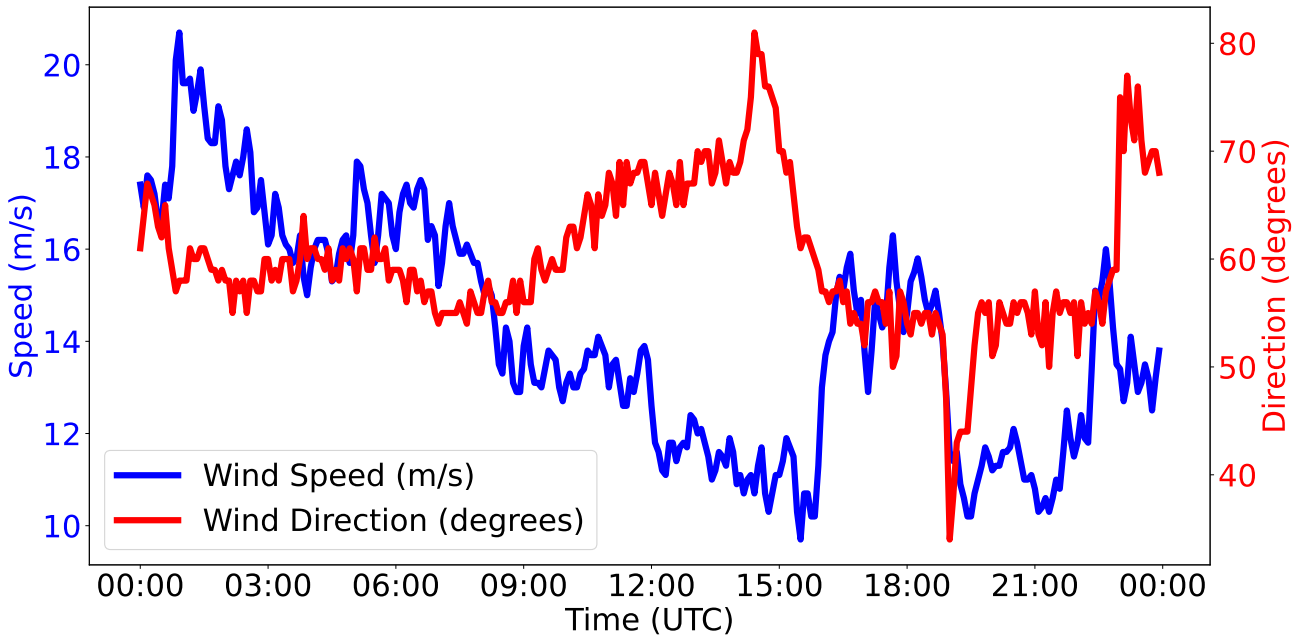

Figure S1. A one-day time series of wind measurement U_{10} obtained at the Acqua Alta oceanographic research platform showing the before and during data acquisition wind information. The data acquisition started at 10:35:00 UTC and lasted for 30 minutes. The direction is based on atmospheric convention, i.e 0 radian corresponds to true North with angles increasing due East. This is converted to standard cartesian angles where 0 radian is due East with angles increasing due North, one may convert using $\theta_{cart} = \frac{\pi}{2} - \theta_{atm}$.

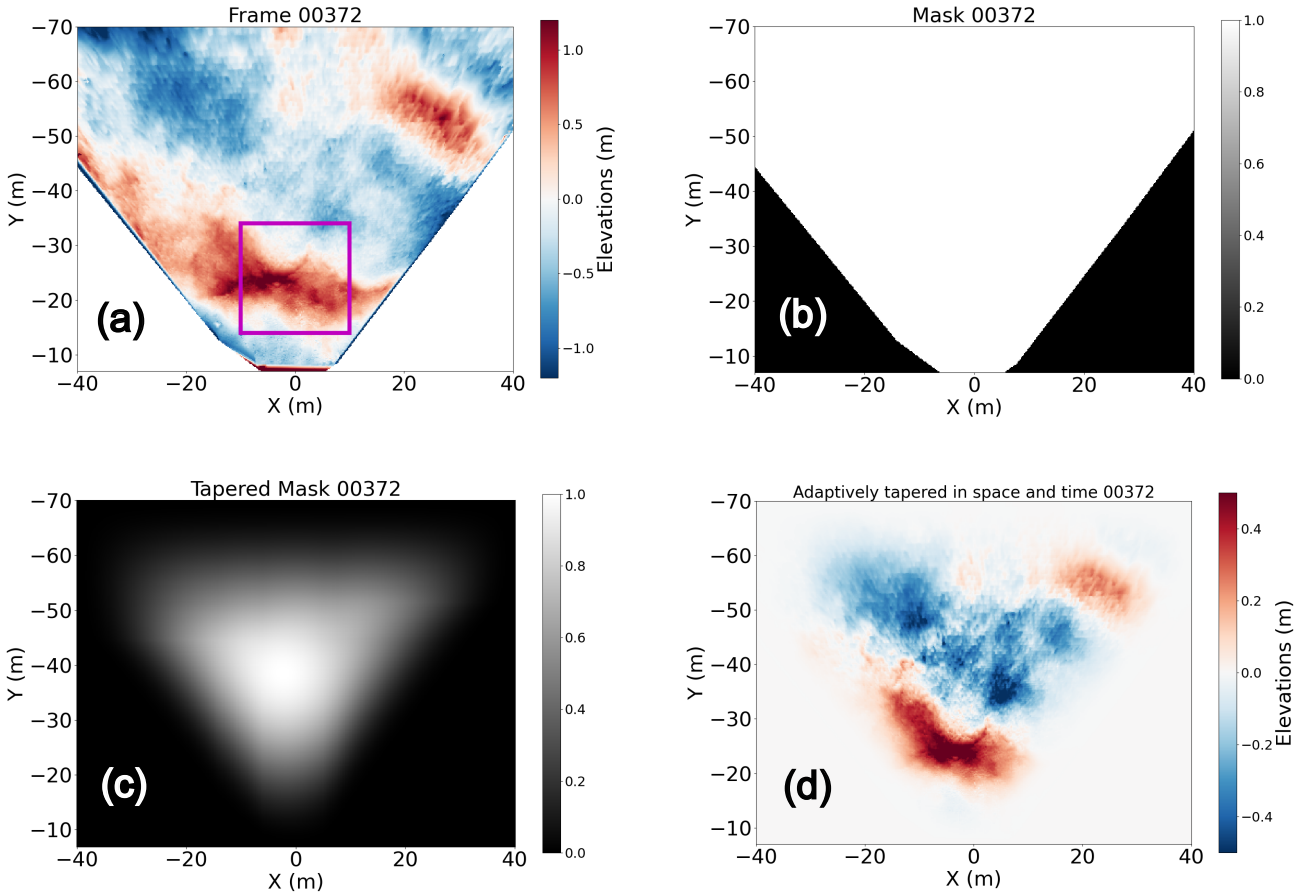

Figure S2. The adaptive tapering procedure. a) An example snapshot of visible stereo reconstructed sea surface elevation. The magenta box corresponds to the 19 m × 19 m region used to resolve high frequencies during the wave spectrum computation. b) The mask derived from the wave elevation map presented in panel (a). c) An instantaneous output of space-time tapering of the mask. d) Tapered mask of panel (c) applied on the wave data shown in panel (a).

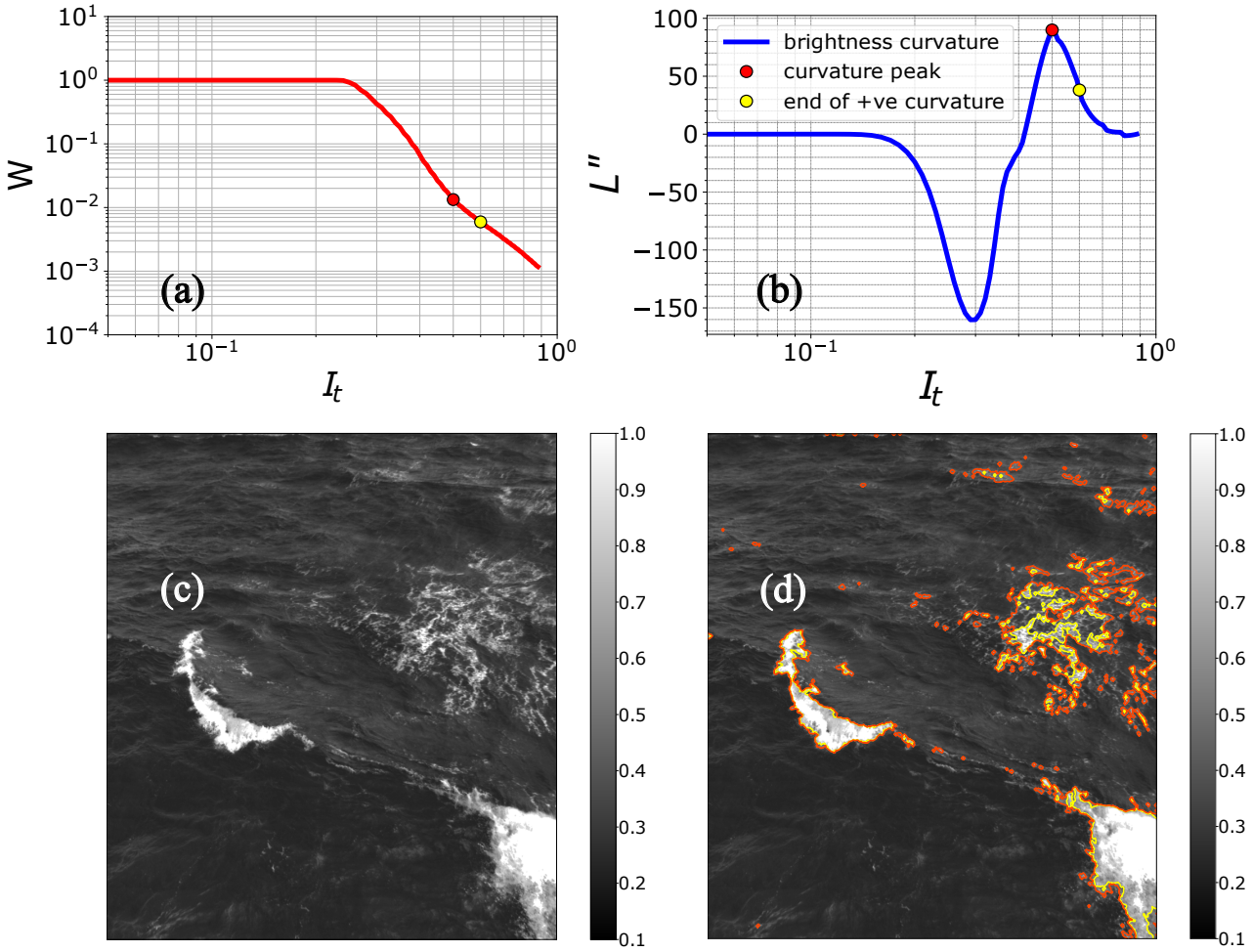

Figure S3. Brightness thresholding from probability density functions. a) Inverse cumulative distribution W as a function of I_t. b) The second derivative of the natural Log of $W(i_t)$. We smoothed before each differentiation with a triangular kernel of size 10 to reduce the noise. c) A randomly selected frame containing both active and passive foam. d) Contours drawn on the chosen frame with orange color corresponding to the peak curvature and yellow contours for the end of positive curvature. Yellow corresponds to $I_t = 0.55$, which is the threshold value used in our study.

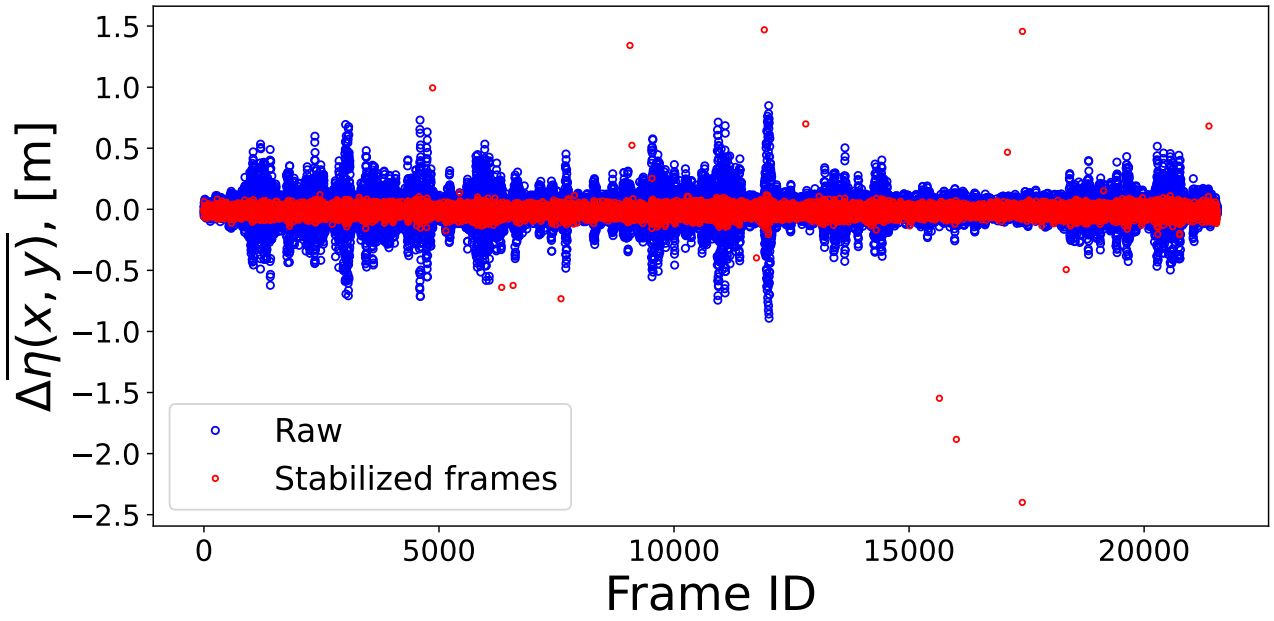

Figure S4. Correction for the camera vibratory-motion using mean elevation. The relatively stable frames are the near zero centered red open-circles comprising of about 95% of the entire dataset. Whereas the open-blue circles correspond to the mean difference in elevation computed from the raw data of the Wave Acquisition Stereo System (WASS) prior to camera motion correction.