Coastal Engineering 58 (2011) 239–250

Contents lists available at ScienceDirect

Coastal Engineering

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

Remote sensing of surf zone waves using stereo imaging

S. de Vries [a,c,*,1], D.F. Hill [b], M.A. de Schipper [a,c,1], M.J.F. Stive [a]

[a] *Section of Hydraulic Engineering, Delft University of Technology, Delft, 2628 CN, The Netherlands*
[b] *School of Civil and Construction Engineering, Oregon State University, 207 Owen Hall, Corvallis, OR 97330, USA*
[c] *Building with Nature, EcoShape, Burgraadt Building, Dordrecht, 3311 JG, The Netherlands*

ARTICLE INFO

Article history:
Received 3 December 2009
Received in revised form 24 September 2010
Accepted 18 October 2010
Available online 3 December 2010

Keywords:
Stereo imaging
Remote sensing
Surf zone
Wave measurement

ABSTRACT

The measurement of water wave characteristics, such as wavelength and wave height, in the surf zone is important for monitoring, prediction of erosion, and numerical model calibration. Traditional methods of measuring wave heights have either been limited to a small number of points or have required contact with the water. An experimental study of the remote sensing of water wave elevations, through the application of stereo photogrammetry, is presented. This method uses two spatially offset cameras, with overlapping fields of view, to determine water surface elevation. This remote sensing approach provides data with excellent spatial coverage and spatial and temporal resolution. Additionally, the hardware needs are minimal and the system is quickly deployed, calibrated, and operational.

In the present study, a phased approach was taken, with medium scale (domain ~50 m^2) laboratory experiments being followed by a large scale (domain ~10^3–10^4 m^2) field test of the method. In the laboratory, reconstructed surface elevations were validated using a pressure sensor and demonstrated excellent agreement. In the field, measured wave heights and periods were found to agree well with available buoy data.

1. Introduction

Wave measurements in the nearshore region play a valuable role in coastal management and forecasting. The most commonly measured parameters are wave period and wave height and these data are of use to the shipping industry, to coastal developers, and to public safety officials, among others. Point measurements of wave conditions are routinely provided by wave buoys. The National Data Buoy Center,[2] operated by the United States' National Oceanographic and Atmospheric Administration, provides a useful portal to real-time and archival buoy data around the globe. Higher resolution information on wave conditions is often obtained by using offshore buoy data to initialize nowcast numerical models of nearshore waves (e.g., the Coastal Data Information Program[3]).

In recent decades, remote sensing methods have been applied to coastal and offshore measurements with increasing frequency. Two of the key advantages of remote sensing are that (i) capital and maintenance expenses are relatively low and (ii) the exposure of instrumentation and personnel to potentially hazardous field conditions is minimized. There are many different types of remote sensing methods, reflective of the numerous possibilities for sensing wavelength. For example, Jessup et al. (1997) reported on an experiment where infrared imaging was used to identify deep-water wave breaking events. Irish and White (1998) review the SHOALS (Scanning Hydrographic Operational Airborne Lidar Survey) system and its applications to coastal engineering problems, highlighting its ability to determine nearshore bathymetry. Hyperspectral imaging (e.g., Sandidge and Holyer, 1998) has also been applied to the problem of determining nearshore bathymetry. Finally, numerous studies have used radar methods to identify and quantify waves using satellite-based (see Krogstad and Barstow, 1999 for a review) and shore-based (Haller and Lyzenga, 2003) platforms.

The use of the visible range (video or photographic imagery) as a remote sensing technique is well established. Early studies (Monahan, 1971) were necessarily limited to the manual analysis of a limited set of photographs. Advances in hardware and image processing software have led to increases in the spatial and temporal resolution capabilities of visible range imaging and have expanded the range of data products that can be derived from the images. For example, Holland et al. (2001) applied particle image velocimetry (PIV) methods to the swash zone. This technique uses two images of the same field of view, but taken at different times, in order to determine the spatial displacements (and therefore velocities) of "tracers" in the images. These tracers provide visual texture in the images and, in the coastal environment, take the form of bubbles, foam, ripples, and other features.

As another example, the Argus program is used to monitor morphologic behaviour of coastal zones (Holman and Stanley, 2007).

* Corresponding author. Building with Nature, EcoShape, Burgraadt Building, Dordrecht, 3311 JG, The Netherlands.
E-mail addresses: Sierd.deVries@TUdelft.nl (S. de Vries), david.hill@oregonstate.edu (D.F. Hill), M.A.deSchipper@TUdelft.nl (M.A. de Schipper), M.J.F.Stive@TUdelft.nl (M.J.F. Stive).

[1] http://www.ecoshape.nl.
[2] http://www.ndbc.noaa.gov.
[3] http://cdip.ucsd.edu.

0378-3839/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.coastaleng.2010.10.004

The initial focus of the Argus program was on time exposure photographs where the image intensity at each pixel could be taken as a proxy for wave dissipation. Then, and as reviewed by Lippmann and Holman (1989), a wave model could be used to link the dissipation to bathymetry. In essence, therefore, the Argus data provided inexpensive non-contact bathymetry information. The automated nature of the Argus instrumentation meant that individual coastal sites could be sampled repeatedly over periods of days, months, and years. Davidson et al. (2007) described the practical application of these video-based measurements to coastal management. Special emphasis was placed upon identifying and prioritizing derived products that are of use to managers and policy makers.

The visible range imaging methods described above are unable to directly measure wave amplitude since time-of-flight is not recorded by standard video. It is worth pointing out that there are circumstances where a single sensor could measure wave amplitude. For example, a downward-looking visible-range laser would be able to determine the water surface elevation. Recently, there have been some efforts using a stereoscopic (two camera) setup to measure the elevations of a dynamic water surface. A successful stereo imaging setup would yield the same data (wave height and period) as a wave buoy, but would also be able to directly measure wavelength and would have the advantage of providing data over a finite two-dimensional area as opposed to a single point.

Holthuijsen (1983) provided an initial description of the method and of an operational system using two airborne cameras. The lack of digital cameras and automated analysis methods at the time of the study led to prohibitive post-processing times. Piepmeier and Waters (2004) more recently discussed a laboratory implementation of the stereo imaging of waves. In their study, they considered only monochromatic waves and found it necessary to artificially “roughen” the water surface in order to provide adequate visual texture. Wanek and Wu (2005) investigated the use of trinocular imaging for the purposes of imaging waves in a field environment. However, their study was limited to a very small field of view ($O(1\ m^2)$) and it relied upon a delicate laboratory-based calibration method that may not be suitable for large scale applications.

Moving up in spatial scale, Benetazzo (2006) presented results of the stereo imaging of waves from two field campaigns, one 4-m^2 in area, the other 400 m^2 in area. Santel et al. (2004) described a true field test of stereo imaging which covers a domain approximately 40,000 m^2 in area. While promising, the accuracy of their camera setup was limited mainly due to the high distance between the cameras and the area of interest.

The goal of the present study is to improve the ability of stereo imaging methods to measure waves in the surf zone. Therefore the results of both a laboratory setup and a field setup are assessed. The ultimate goal is a system that will be able to measure waves over a large area with good spatial and temporal resolution, that will be able to be rapidly installed and calibrated, and that will have reasonable image processing times.

2. Methods

The general procedure of stereo imaging is to first take stereo images using calibrated cameras. These images are then correlated pixel by pixel for matching features creating pixel pairs. Finally, the pixel pairs are triangulated towards real world XYZ-coordinates. Two experimental configurations are discussed.

2.1. Experimental configuration

2.1.1. Laboratory experiment

The laboratory stereo imaging measurements were carried out as part of a larger project studying the effects of wave attack on dune erosion (van Thiel de Vries et al., 2008). The experiments were in the

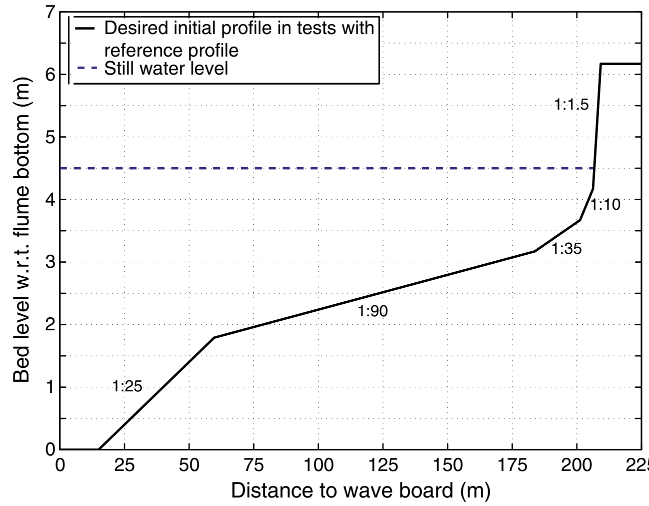

Fig. 1. Elevation view of the initial bed profile for the laboratory flume experiments.

Delta Flume operated by Deltares (formerly WL-Delft Hydraulics), located in the Netherlands. The flume is 225 m long, 5 m wide, and 7 m deep. The wave generator is equipped with active reflection compensation and 2nd order wave steering. The experimental trial described here had a water depth of 4.5 m near the wave generator and used a single-peaked Pierson–Moskowitz wave spectrum to generate the irregular wave field. The wave spectrum's characteristic wave height (H_{m0}) was 1.5 m and peak period (T_p) was 4.9 s at the wave board. Details on the experimental trial can be found in van Thiel de Vries et al. (2008) from which we use run T01E.

The flume was set up with an initial bed profile as shown in Fig. 1 and was equipped with a wide variety of sensing equipment, including wave gauges, acoustic Doppler velocimeters, pressure sensors, and optical backscatter sensors, among others. A total of 12 pressure sensors (type Kulite HKM-375M-1) were affixed to the flume wall at various distances from the wave generator. Table 1 provides the horizontal and vertical locations of the four sensors in close proximity to the swash zone. The pressure readings were used to derive water surface elevation by translating the pressures to water elevations with linear wave theory.

In addition to this instrumentation, the inner surf and swash zones were imaged with several cameras. Fig. 2 illustrates that two cameras were mounted at an elevation of 8.3 m above the top of the flume sidewalls. Both of these cameras, called C1 and C2, were aimed in the direction of wave propagation and imaged areas approximately 12 m (along the flume axis) by 6 m (perpendicular to the flume axis). The two cameras were separated by a streamwise baseline of approximately 6.4 m. During the experimental trials, the cameras obtained images at a rate of 2 Hz. Sample raw images from the two cameras are provided in Fig. 3.

2.1.2. Field experiment

The field experiments were carried out at the Scheveningen beach, (Fig. 4) on the southern coast of the Netherlands, on July 15, 2008. A tourist pier (~400 m long) present served as the staging area and

Table 1
Positions of selected pressure sensors located on flume wall.

Pressure sensor	Distance from Wave generator (m)	Distance from Flume bottom (m)
PS07	190	3.95
PS08	200	4.15
PS09	205	4.30
PS10	209	4.25

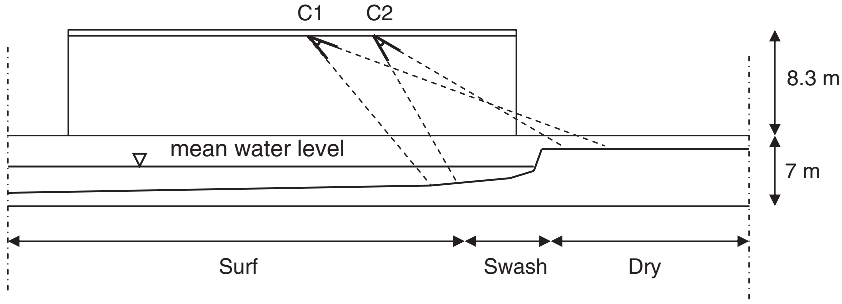

Fig. 2. Schematic detail of the surf/swash zone, showing the location of the stereo cameras.

platform for the measurements. Two identical cameras (Scorpion IEEE-1394) were mounted on the railing of a viewing balcony on the southern side of the pier with a horizontal baseline of 9 m. The terms "left" and "right" will hereafter be used to refer to the shoreward and seaward cameras. The elevation of the railing above the water surface was O(10 m).

An individual experimental trial consisted of the acquisition of 5 min of images at a temporal rate of 8 Hz. A sample raw image pair is provided in Fig. 5. The experimental trial occurred on an overcast day during a rising tide from 10.08 to 10.13 a.m. Coordinated Universal Time (UTC). The preceding low of −0.7 m NAP[4] occurred at 7.33 a.m. and the following high of 0.8 m NAP occurred at 12.20 p.m. Wave parameters were recorded by the IJmuiden station, approximately 50 km to the northeast along the Dutch coast. The reported significant wave height offshore was around 1 m with a mean period (T_m) on the order of 4 s throughout the experiment. The combination of the foam from the spilling breakers and the high contrast from the diffuse lighting resulted in very good visual texture in the images.

As illustrated by Fig. 4, a local coordinate system was defined with its origin at the foot of the pier and X pointing seaward along the pier and Y pointing south along the coast. The vertical reference plane is NAP. Fig. 6 shows the projection of a typical image from the left camera onto this coordinate system, together with the right camera projection outline. The overlap close to the cameras is limited and increases with distance from the cameras. The relative coverage per pixel similarly increases with increasing Y.

2.2. Camera calibration

The general procedure in stereo imaging is to first collect overlapping stereoscopic images of a field of view of interest. Each of these images is therefore a two-dimensional (2d) representation of the actual three-dimensional (3d) world. Defining the relation between the 2d representation and 3d world is called camera calibration.

2.2.1. Camera model

Following Hartley and Zisserman (2000, Chap. 6), and referring to Fig. 7, the 2d image coordinates x_i and 3d world coordinates X_i are related through a camera projection matrix P according to

$$x_i = PX_i, \qquad x_i = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}, \qquad X_i = \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}. \tag{1}$$

With the camera projection matrix known, for all 3d coordinates corresponding 2d coordinates can be calculated. The projection matrix, which describes the intrinsic and extrinsic camera parameters, is subdivided by

$$P = KR\left[I \,|\, -\tilde{C}\right], \tag{2}$$

where I is the 3×3 identity matrix and

$$K = \begin{bmatrix} \alpha_u & s & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

accounts for the intrinsic camera parameters. In this matrix, $\alpha_u = f_l m_u$ and $\alpha_v = f_l m_v$, with f_l being the focal length, and m_u and m_v the scale factors (pixels length^{-1}) in the u and v directions. The parameter α_v is also known as the relative focal length (pixels) f_p and the pixel ratio λ_u is given by α_u/α_v. The parameter s accounts for the skew in the CCD array of the camera. Finally, u_0 and v_0 are the principal point coordinates in pixels.

Next, the extrinsic camera parameters are accounted for by $\tilde{C} = (X_c, Y_c, Z_c)^T$, which represents the camera position in world coordinates, and R, which is a 3×3 rotation matrix describing the orientation of the camera coordinate system with respect to the world coordinate system. Using the tilt-swing-azimuth (Wolf, 1983, Chap. 11) system, also, known as the pitch-roll-azimuth system. Unknowns in the rotation matrix are the pitch (τ), roll (σ) and azimuth (ϕ) angles indicated in Fig. 7.

2.2.2. Distortion correction

First, note that (1) is a projection that assumes no distortion. In reality, however, camera lens distortions result in variable magnification across the image plane. These distortions are often modeled as

$$\Delta r = k_1 r^3 + k_2 r, \tag{4}$$

where r is the radial distance from (u_0, v_0), Δr is the distortion pixel displacement, and k_1 and k_2 are coefficients. Using (4) the distortions can be corrected for.

2.2.3. Practical application and calibration results

Generally speaking, camera calibration is carried out with a two step approach, as reviewed by Holland et al. (1997). First the distortion coefficients, and the other intrinsic camera parameters, are determined in a laboratory setting. In this step, images are taken of a 'target,' which is typically a regularly-spaced grid of high-contrast features, such as dots or other markers. The appearance of this grid on the image plane is used to derive the distortion coefficients (k_1 and k_2 and intrinsic parameters (u_0, v_0 and λ_u)).

With the intrinsic parameters determined, the second step consists of determining the extrinsic parameters in-situ. This is done by measuring the camera positions (X_c, Y_c and Z_c) and imaging a number of Ground Control Points (GCPs). These ground control points are points with exact known world and image coordinates. Each GCP pair yields 2 equations and therefore a minimum of 2 ground control points is needed to solve for the remaining 4 parameters (σ, τ, ϕ and f_l). When

[4] NAP is the Dutch reference level close to mean sea level.

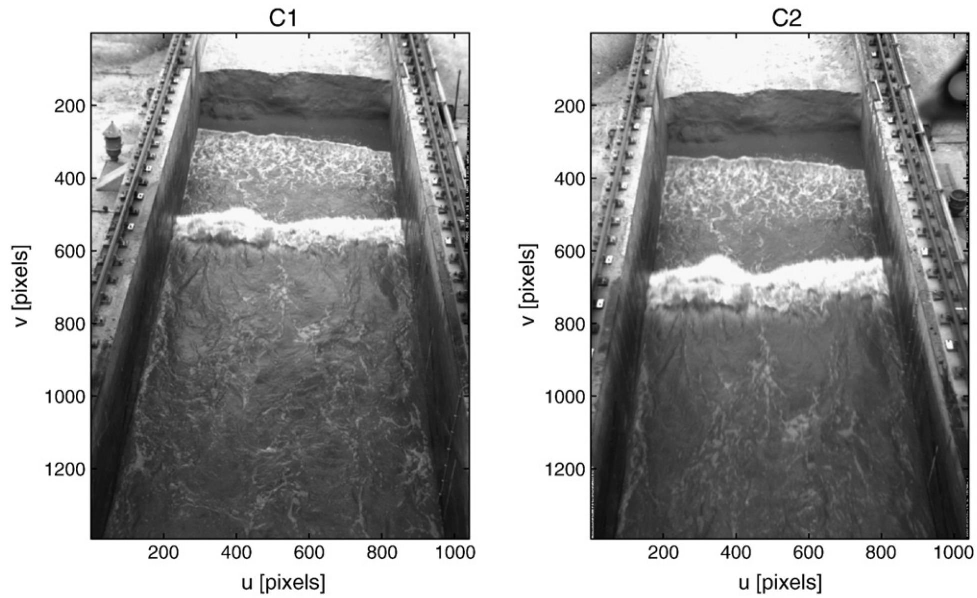

Fig. 3. Sample of synchronized images from the C1 (left) and C2 (right) cameras. Waves are propagating from the bottom to the top of the images.

acquiring more than 2 ground control points an overdetermined system can be solved using a least squares method for better accuracy.

In the laboratory experiment presented here, GCPs were readily available, due to the existence of numerous static features of precisely known location. A total of 17 GCPs, all bolts on the flume rail, were used for the calibration. For the field experiment presented here, where the cameras were imaging the sea surface, with no distinguishing features (buildings, other structures, etc.) of known location, the in-situ calibration was more challenging.

The positions of GCPs in field experiments are commonly determined through the use of a Global Positioning System (GPS). For the present field experiment, a Septentrio Polarx2e GPS receiver was affixed to a raft which was then navigated through the field of view. The GPS system was set up in Real Time Kinetic (RTK) mode with correction values from a nearby base station in Rijswijk (provided by LnR Globalcom B.V.) leading to an accuracy of 2–5 cm in all directions. Unlike static environments, with multiple fixed objects in the field of view, it was not possible to determine multiple GCPs in a single camera image. Therefore, images were acquired continuously, defining the 2d images locations of the GPS over time. When these images were time-synchronized with the 3d GPS measurements, the GCPs were fully determined. A total of 47 GPS points were obtained of which 24 are used as GCP for calibration and 23 to asses the validity of the calibration. GCPs are depicted in Fig. 6.

Appendix A provides the results of the calibrations for the laboratory and field experiments for completeness.

2.3. Feature matching

Image correlation essentially locates or tracks identifiable features between the images making up a stereo image pair. The refinement of correlation methods remains an active area of research since these methods provide the foundation for fields as diverse as experimental fluid mechanics measurements and computer vision. In PIV measurements (Adrian, 1991), the correlation is done in the frequency domain using Fast Fourier Transforms (FFTs). In early PIV studies, image subwindows were $2^n \times 2^n$ ($3 \le n \le 7$, typically) pixels in dimension and were static in location between the two images. This necessarily limited tracer displacements to a small (1/4, ideally) fraction of the window size. Later refinements allowed for the displacement and

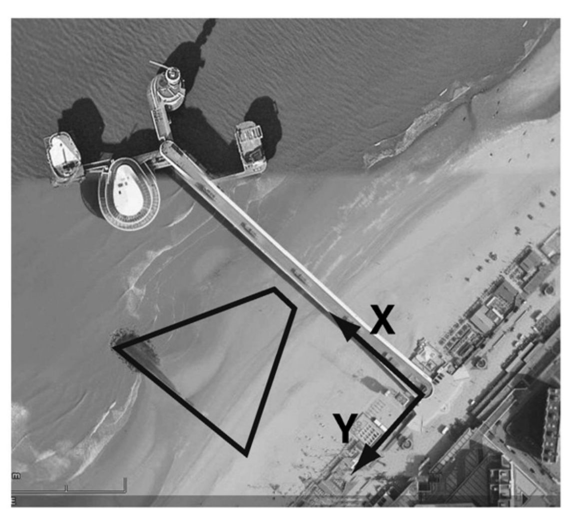

Fig. 4. Overview (left) of the Netherlands showing (circle) the location of Scheveningen. A satellite photo of the field site, with the approximate field of view indicated by the trapezoidal area, is given on the right.

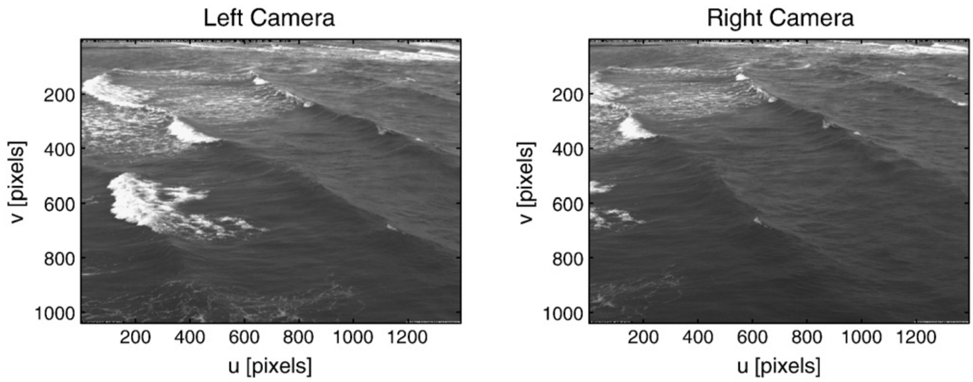

Fig. 5. Sample raw image pair during the experimental trial.

deformation of subwindows between images, greatly improving the accuracy of the correlation estimates.

In computer vision applications (Scharstein and Szeliski, 2002), the correlation is usually done in the spatial domain. The actual correlation process itself is highly variable, depending upon the specific application. For example, relatively low texture regions may require relatively large window sizes in order to find a valid match while high texture regions can be interrogated with smaller windows. Additionally, some strategies use an iterative, multiple-pass approach. In this case, images are initially correlated with large window sizes and the results from this coarse first approximation are used to constrain the search region for successive, higher-resolution passes.

2.3.1. Image rectification

Both photos of a stereo photo pair have separate image planes. With knowledge of the camera locations and orientations, it is possible to rectify the images. Rectifying the images fits both images in the same (rectified) image plane. The rectified image plane is defined as parallel with the cameras' baseline. Either the U or the V coordinate of the rectified image plane will be directed parallel with the baseline. In the case of the laboratory experiments, the V coordinate is parallel to the baseline and in the case of the field experiments, the U coordinate is parallel to the baseline. This rectification process creates epipolar lines that help to constrain the correlation process, thereby reducing computational effort. For example, in the rectified laboratory images (Fig. 8), a unique feature appearing in both images will have the same U coordinate, but will be separated by a disparity in the V direction. In the rectified field images (Fig. 9), a unique feature will have the same V coordinate but will be separated by a disparity in the U direction.

Using Eq. (1) it is possible to define a conversion matrix (H) which converts the uv coordinates of the original image plane to UV coordinates in the new (rectified) image plane:

$$x_{UV} = Hx_{uv} \tag{5}$$

where

$$H = P_{rect}P'. \tag{6}$$

The conversion matrix H is called a homography (Hartley and Zisserman, 2000, p.327). The P_{rect} matrices are chosen to define an image plane parallel to the baseline.

Due to experimental uncertainty, the epipolar lines of the rectified image pairs (Figs. 8 and 9) were not perfectly lined up in both the laboratory and the field cases. Therefore, it was necessary to perform a two dimensional correlation in order to track features from both

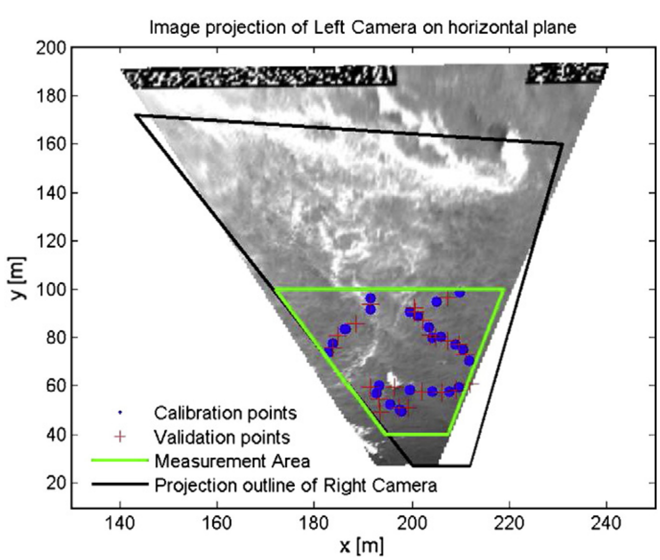

Fig. 6. Projection of a typical image from the left camera onto the local axis system. Also shown is the projection outline of the area imaged by the right camera, the measurement area and the Ground Control Points (GCPs) used to calibrate and validate the camera setup.

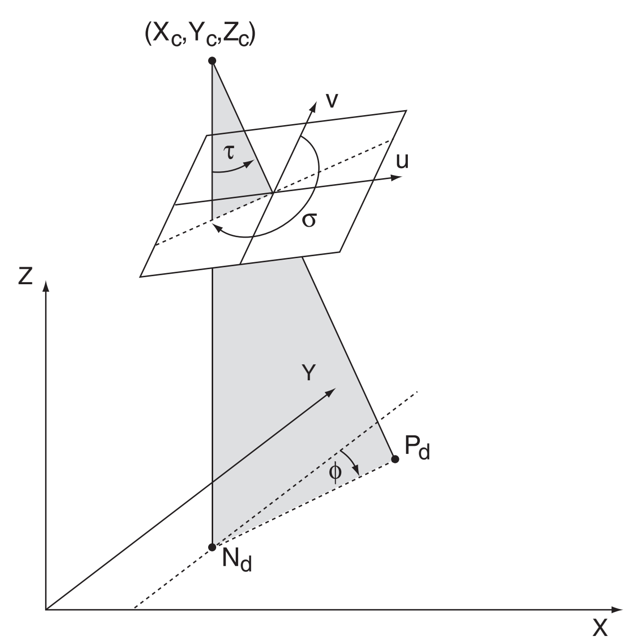

Fig. 7. Sketch illustrating camera parameters and relationship between world coordinates and image plane coordinates.

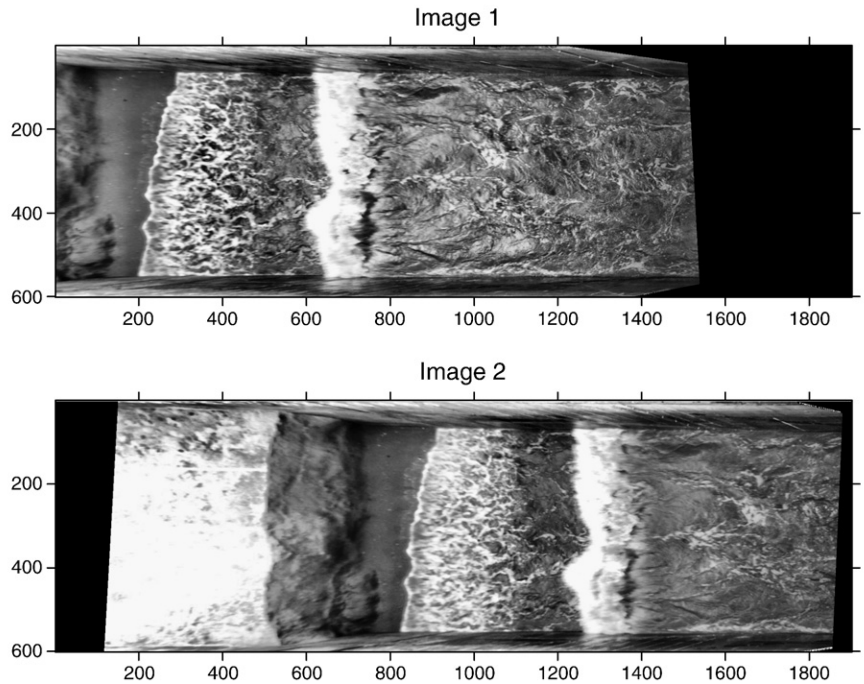

Fig. 8. Sample of rectified images from the C1 and C2 cameras for the laboratory experiment.

cameras. However, this misalignment was only slight, typically 1–10 pixels, which meant that the search was highly constrained in the direction perpendicular to the baseline. Therefore, the rectification process still led to a substantial computational savings.

2.3.2. Normalized cross correlation

After rectifying the images a normalized cross-correlation method was used to identify matching features. The correlation coefficient C is formally defined as

$$C(u,v) = \frac{\sum_{x,y}\left[f(x,y)-\bar{f}_{u,v}\right]\left[t(x-u,y-v)-\bar{t}\right]}{\left(\sum_{x,y}\left[f(x,y)-\bar{f}_{u,v}\right]^2\sum_{x,y}\left[t(x-u,y-v)-\bar{t}\right]^2\right)^{1/2}}. \quad (7)$$

In this expression, and as illustrated in Fig. 10, f represents an image and t a template, or feature, which overlies the image and is centered at (u,v). In Fig. 10, the location of the template in the image has been identified. The mean of the template is denoted by $\bar{t}$ and the mean of the image, in the region under the template, is given by $\bar{f}_{u,v}$. The summations are carried out over the portion of the image that underlies the template. The result of the correlation is also shown in Fig. 10, with the grayscale being mapped to the range of −1 (black - uncorrelated) to 1 (white—perfect correlation). The correlation result clearly identifies the correct location (whitest pixel) of the template in the image.

As applied to the current experiments, the template is simply a sub-image from one camera (C1 or the left camera) and the image is part or all of the image from the other camera (C2 or the right camera). While, in principle, the entirety of the C2/right camera image can be used, it is preferable, in terms of reducing false matches and speeding up processing, to constrain the search space. In the present study, this was done by constraining the water surface elevations to be within a certain range. An illustration of the result of the correlation procedure to sample images from the laboratory is given in Fig. 11. A spatial distribution of horizontal disparity values is given together with the distribution of the magnitude of the correlation coefficient (C).

2.3.3. Confidence filtering

We use two filter methods to discard false correlations. First any correlation coefficient smaller than a certain threshold value gets discarded. Fig. 11 shows the spatial distribution of the correlation coefficient over the rectified image of the laboratory experiment. Overall the correlation is very high except in the area where low texture is present. In areas of low texture, false matches are easily identified. For instance, strong discontinuities in disparity values are

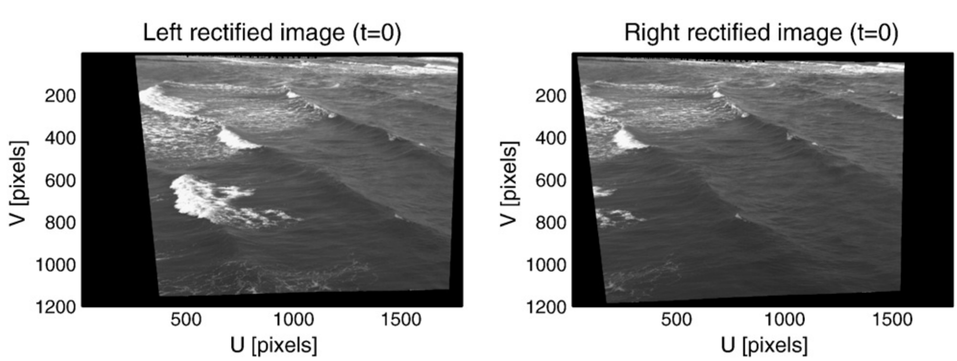

Fig. 9. Sample of rectified images from left and right cameras for the field experiment.

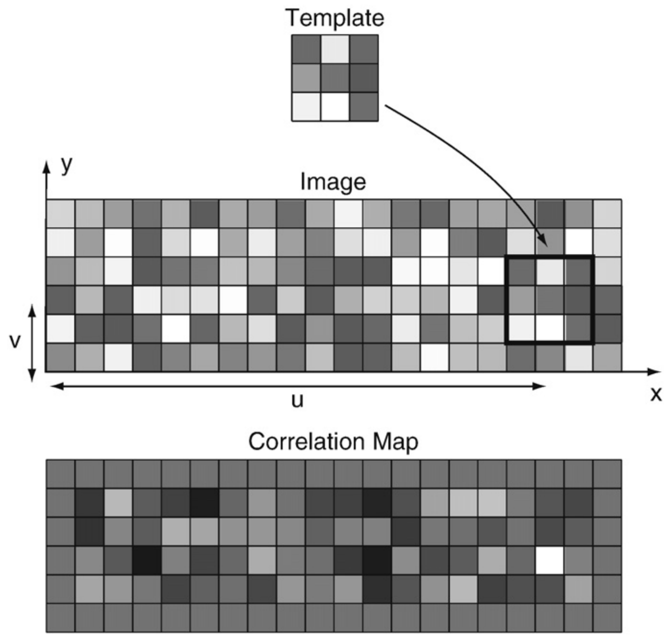

Fig. 10. Conceptual sketch of the image correlation procedure.

observed in Fig. 11. These false matches are found because there was no strong correlation peak due to low texture and get discarded.

A consecutive two-dimensional median filter discards any disparity value which deviates a certain factor from the disparity values of neighboring pixels resulting in the removal of disparity peaks (spikes). Discarded disparity values get re-interpolated using standard procedures.

2.4. Triangulation and reconstruction

Using the image pixel pairs, which are the two camera coordinates of individual features, and available P matrices, each pixel pair can then be back projected. In principle, the intersection of such projection lines determines the real world location corresponding to the pixel pair. In reality, an exact intersection is impossible due to the discrete nature of the pixels and inaccuracies along the way. Therefore this intersection must be estimated. Combining (1) for both cameras, a system of four equations with $X_i = (X, Y, Z, 1)^T$ coordinates as unknowns is derived. This system is then solved for every pixel pair using Single Value Decomposition. Further details on this triangulation procedure are given in Hartley and Zisserman (2000) and van Thiel de Vries (2009).

3. Results

3.1. Laboratory Experiments

Reconstructions of the laboratory tests span a measurement area of roughly 5×10 meters. Sample results are shown in Figs. 12 and 13. The former figure shows color contours of water surface elevation and the latter maps the original camera image onto the reconstructed water surface. Qualitatively, the stereo imaging appears to successfully capture the key features of the water surface. The spilling breaker crest at $x \sim 208$ m is clearly evident, as is the trough just shoreward of this location.

To provide a more quantitative analysis of the success of the stereo imaging, the results were benchmarked against water surface elevations calculated using the pressure sensor located at $(x,y) = (205, -2)$ m. Recall that the images were acquired at a rate of 2 Hz. A time series of elevation was derived from the sequential stereo images by extracting elevation estimates at this horizontal location. A comparison of water surface elevations, as determined from the two methods, is provided in Fig. 14. Overall, the agreement is excellent, with the stereo imaging being able to pick up both high- and low-frequency waves as well as small and large amplitude waves. Comparing every data point of the full time series of the available 4136 consecutive photo pairs (34.5 min) and pressure data, the root mean square (RMS) of the difference between both estimates (at location $(x,y) = (205, -2)$ m) is 0.034 m.

Analyzing the measured variance density spectra (Fig. 15) it is shown that for all frequencies the stereo photo data slightly under predicts the variance with respect to the pressure sensor data. When converting measured total variance (m_0) of both our signals' wind wave frequencies (0.05–0.3 Hz) towards H_s (using $H_s = 4\sqrt{m_0}$) deviations are found to be less than 10%. Moreover peak periods derived from the spectra ($T_p = \sqrt{m_0 / m_2}$) agree well. See Table 2 for an overview.

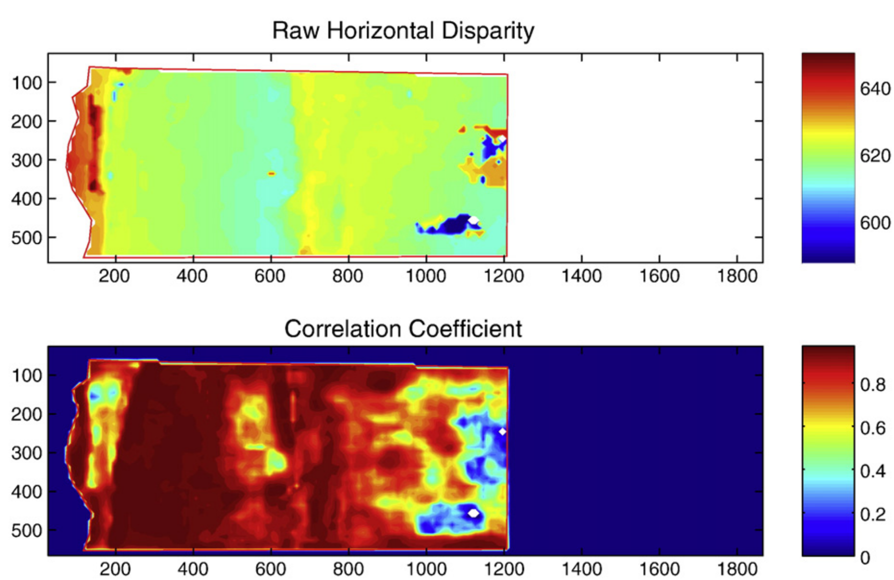

Fig. 11. Top image shows calculated horizontal disparity for photo pair shown in Fig. 8. Relevant area is bounded by the overlap area of the two photos, the waterline and the flume walls. Bottom image shows the accompanying correlation coefficient.

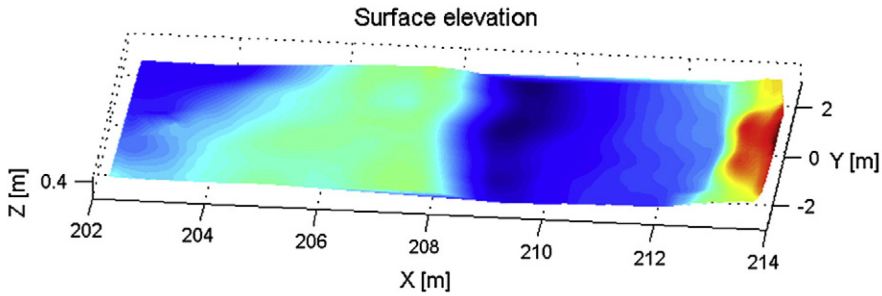

Fig. 12. Perspective view of the sample reconstruction of water surface elevation. Color contours denote elevation in meters.

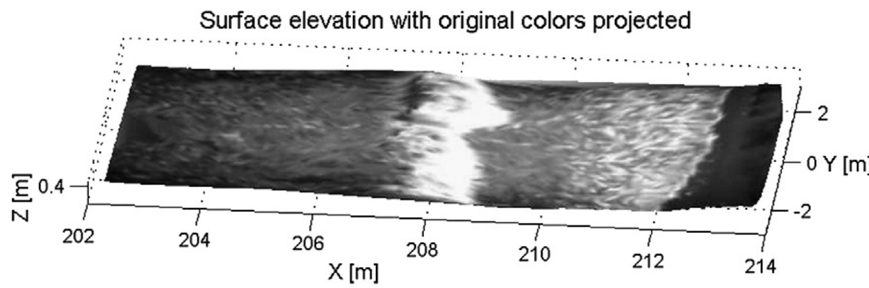

Fig. 13. Perspective view of original camera image mapped onto the three-dimensional water surface elevation.

This good if imperfect agreement between the stereo imaging results and the pressure sensor data seems to be consistent with the findings of Guza and Thornton (1980). They investigated the validity of using pressure sensors and linear wave theory to estimate water surface elevation in and outside of the surf zone. Their results also showed deviations of significant wave height H_s in the order of 10%.

3.2. Field experiments

Reconstructions of the field experiments span a measurement area of around 1800 m^2. A sequence of images (sub-sampled from the 8 Hz raw data to 2 Hz) is shown in Fig. 16. In this figure, the column on the right contains successive images from the right camera. The polygon superimposed on the images roughly delineates the region that was correlated with the left camera images. The left column of Fig. 16 shows the reconstructed water surface elevations in meters. Upon inspection of Fig. 16, it is observed that, as with the laboratory results, surface elevations are successfully reconstructed over the area of interrogation.

The acquired dataset, with its high temporal and spatial resolutions provides a very complete picture of the time history of the water surface over a large area. The dataset contains over 1500 (X,Y) points, which can be seen as closely spaced virtual wave gauges, and can be analyzed as such. Moreover, while a 5-min 2-Hz dataset is available a spatial distribution of statistical parameters can be assessed.

No ground truth data are available in terms of wave time series within the measurement area. However, there are a number of quantitative measures that can be considered, to help assess the performance of the stereo imaging process. For example, the nearest wave station (IJmuiden) provides spectral information and the Scheveningen tide gauge provides water level information. Additional GPS measurements have also been taken.

Using the IJmuiden wave station data the wave period is considered only. Wave height and energy density cannot be compared due to the large distance between the wave station and the measurement area. Using the time series of the stereo reconstructions, the mean period at every spatial location can be derived using spectral moments m_0 and m_2 considering the wind wave band (frequencies between 0.03 and 0.5 Hz). Averaged over the measurement area the mean period, T_m, is derived to be 3.9 s. This measured mean period relates very well to the measured offshore mean period which ranges between 3.5 and 4.1 s.[5]

Additional RTK-GPS measurements, acquired during the calibration process, of 23 locations within the measurement area not used for calibration are available. The 2D locations of these points are manually identified on the photo pairs and reprojected towards their 3D real world coordinates. Fig. 6 shows the locations of the used "validation" points. Re-projecting the image coordinates towards 3d coordinates, deviations are found when compared to the measured GPS locations. The RMS values of these deviations are 0.071 m, 0.073 m and 0.041 m corresponding to X, Y and Z direction respectively. Note that these deviations are of similar order as the uncertainty related to the GPS measurements.

The mean water level is calculated using the total spatial mean of the water level for every time step. The mean water level of all time steps combined is −0.058 m NAP with a standard deviation of 0.014 m over the measurement domain. This value is coherent with the −0.06 m. measured at Scheveningen tidal station[6] located 2.5 km south of the measurement area. Calculating the mean surface elevation in time for every point individually a spatial pattern relative to the reference plane is found, see Fig. 17. This variation shows the characteristics of a plane with offset angle of 0.08 degrees in Y direction and −0.08 degrees in Z direction relative to the reference NAP plane. This deviation from the reference plane is attributed to inaccuracies in the calibration and coordinate registration method. Spectral parameters analyzed hereafter are based on the detrended time series at every measured location, therefore the relatively small deviation of the mean is not of relevance.

4. Discussion

4.1. Potential Products Derived from Measurements

The stereo imagery on medium and large scales (domain ~10^4 m^2) as presented in the present paper, provides large opportunities in the research of nearshore hydrodynamics and morphology. Time series of the surface elevation as collected during this research show detailed

[5] Measured at IJmuiden at 15-7-2008 between 8.00 and 12.00 UTC.
[6] Measurement at Scheveningen tidal station at 15-7-2008 at 10.10 UTC.

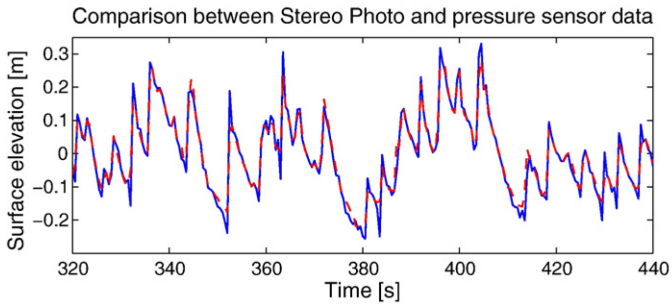

Fig. 14. A 2-min (out of 34.5 available minutes) comparison of the stereo photo and pressure sensor data. The blue solid line indicates pressure sensor data and the red dashed line represents the stereo photo data extracted at the location of the pressure sensor.

characteristics of the nearshore wave field, such as wave asymmetry (sharp wave crests and flatter troughs), wave skewness (steep wavefronts and mild back slopes) and time-varying wave amplitude due to the presence of wave groups. These aspects, as well as more trivial characteristics such as wave height and period can be evaluated at every point in the domain without an extensive array of in situ instruments. Especially at locations with large spatial variations, such as nearshore bar rip morphology, this can provide valuable information.

A cross shore transect of the wave heights over the domain is shown in Fig. 18, to illustrate the differences that can be observed over a domain of this size. Significant wave height H_s, about 0.42 m at the offshore edge of the domain, increases until $X \approx 200$ m. From $X \approx 200$ m towards shore the significant waveheight H_s decreases. The variations in wave height can be explained by wave breaking, visible as white regions on the photos. The photos indeed show that wave breaking is initiated near $X = 208$ m, and the majority of the waves break around $X = 200$ m. This visually determined breakpoint in the middle of the domain explains the cross shore variations in waveheight as seen in Fig. 18.

A comparison between surface elevation spectra on both sides of the breaker line is shown in Fig. 19. We observe a decrease in energy at the

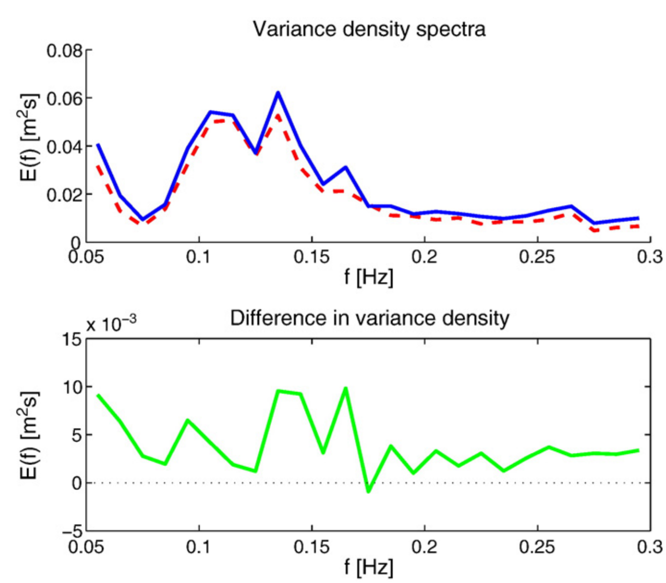

Fig. 15. Top panel shows derived spectra from a single run. The blue solid line indicates the pressure sensor data and the red dashed line represents the stereo photo data extracted at the location of the pressure sensor. Bottom panel shows a positive difference between the variance densities, indicating that the stereo photo technique generally underpredicts the variance for all frequencies relative to the pressure results.

Table 2
Comparison of time series results between the stereo imaging and in-situ pressure sensor.

Parameter	Stereo photo	Pressure sensor
T_p [s]	6.4	6.3
H_s [m]	0.27	0.29
Time series RMS difference [m]	0.034	

peak frequency as well as an increase in energy in the super- and sub-harmonics. This is in accordance with Guza and Thornton (1980) who observed similar spectral behavior analyzing nearshore waves.

Besides the waveshape and height information, the closely spaced high resolution data also provide information on wave propagation in the domain and the interaction of the incoming waves with bottom topography. One promising application of stereo photogrammetry is the potential contribution to refine the determination of in field bathymetry. In-situ collection of bathymetry data can be difficult and expensive and remote sensing methods are an attractive alternative. As an example, Stockdon and Holman (2000) describe measurements of the surf zone using individual cameras. Their estimates of wavenumber and wave period, as derived from the camera images, were combined with linear wave theory in order to estimate depth. While their results showed good agreement with ground truth data in the case of small wave amplitudes, the depth estimates were found to be less satisfactory in regions of known wave nonlinearity.

Catalan and Haller (2008) state that composite models, which retain both wave dispersion and nonlinearity effects, are superior to linear methods in the surf zone. Their laboratory experiments investigate the effect of wave nonlinearity by using individual cameras for video imagery to measure the phase speed and wave gauges to measure the wave height. They estimated water depth using a variety of phase speed models and compared the estimates against the measured water depth. For a correct application of these composite models, it is paramount to know the wave height in order to estimate the nonlinearity. Especially in very shallow waters, this formulation cuts the RMS error between the estimated and actual water depth to one third of the error associated with the linear theory estimates, as shown by Catalan and Haller (2008).

4.2. Accuracy of measurements

The inaccuracies and errors in the camera model, calibration, correlation and reconstruction procedures determine the uncertainties of the method as a whole. One important systematic uncertainty, as reviewed by (Benetazzo, 2006) is the quantization, or resolution, uncertainty. To estimate this quantity, consider the schematic drawing of the stereo camera arrangement in Fig. 20. In this figure, the $\hat{X}$ axis is the camera baseline and the $\hat{Z}$ axis is perpendicular to the baseline and in the plane defined by the (non-parallel) lines of sight of the two cameras. The $\hat{Y}$ axis is defined in the usual right-hand sense. The camera view angle is given by 2β and the included angle between the cameras' lines of sight is 2α. (Benetazzo, 2006) then gives the maximum quantization errors as

$$
\begin{aligned}
er\,\hat{Z} &= \frac{\hat{Z}^2}{2TN}\frac{sin(2\beta)}{cos(\beta+\alpha)^2} \\
er\,\hat{X} &= \frac{\hat{Z}}{2N}\frac{sin(2\beta)}{cos(\beta+\alpha)^2} \\
er\,\hat{Y} &= \frac{\hat{Z}}{2N}\frac{sin(2\beta)}{cos(\beta)^2}.
\end{aligned}
$$

In these formulae, the camera baseline is given by T and the CCD array resolution (number of pixels along one side) is given by N. For the field cameras $\beta = 28°$ and α was limited to just a few degrees. Therefore, the above equations give maximum quantization errors (corresponding to the distant, in this case, southern, end of the field of view) of roughly 1 m, 5 cm, and 5 cm in the $\hat{Z}$, $\hat{X}$, and $\hat{Y}$ directions. Not that in Benetazzo's

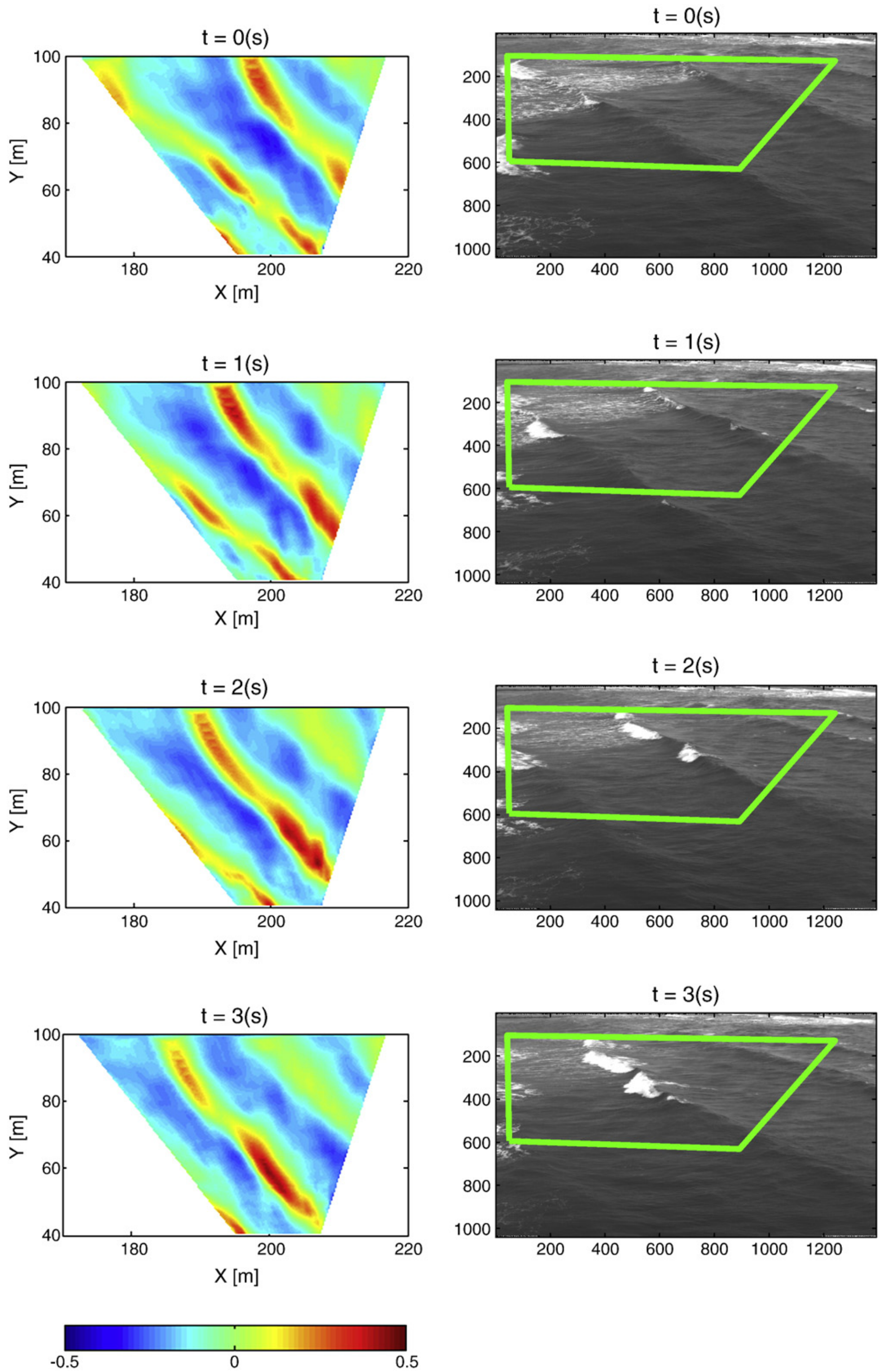

Fig. 16. Sequence of reconstructed images (left) and one of the corresponding original images (right). The colorbar indicates the surface elevation in meters relative to NAP. The green box on the original image represents the actual correlated surface.

schematisation the $\hat{Z}$ axis is in the direction of the line of sight of the camera.

These theoretical quantization errors discussed above were used to design the stereo setup. However, in field situations the camera axes are non perpendicular to the real world axis and the correlation algorithm is able to give sub-pixel solutions for the disparity. Moreover, the accuracies of the determined angles regarding the stereo setup (in Benetazzo's case α and β) are dependent on each other and difficult to derive. Therefore Benetazzo's formulas give only indications at best.

In this paper synoptic indications of accuracies are given based on parallel measurements using a pressure sensor or aggregated parameters combined with GPS measured GCPs for laboratory and field situation

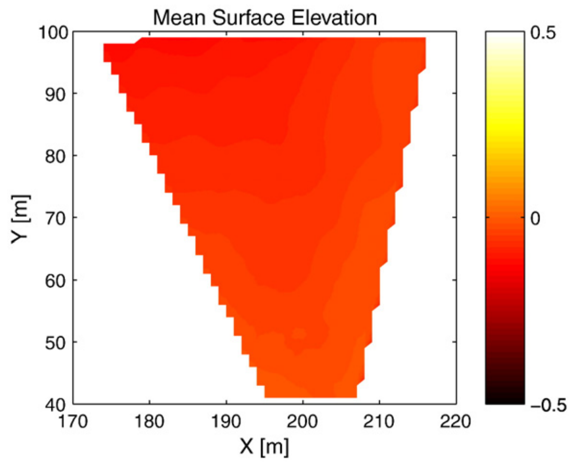

Fig. 17. Mean surface elevation calculated over 5 min.

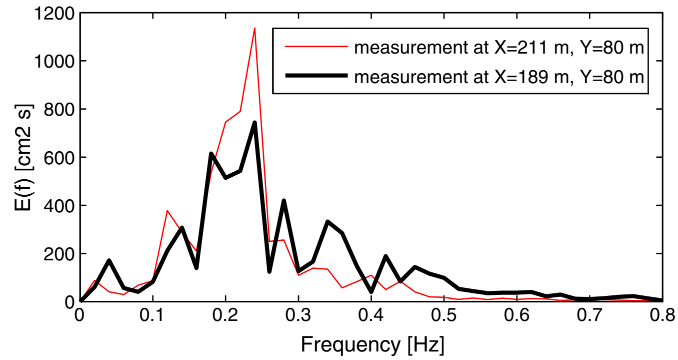

Fig. 19. Surface elevation spectra derived from the stereo imaging near the offshore and onshore edge of the domain (shown in red and thick black line respectively).

respectively. Based on this synoptic analysis the accuracy is conservatively assumed to be (O< 0.1 m) in the field and (O< 0.05 m) for the laboratory situation.

5. Conclusions and future work

In conclusion, it has been demonstrated that stereo imaging of the surf zone is a viable and competitive method for obtaining wave data. The method provides for highly resolved (both in time and space) data on the water surface elevation over large areas. The present study is particularly of note for the execution of these experiments over a much greater area than has been reported previously. However, it is worth recognizing the inherent limits of this method that are best addressed through future testing and development. First of all, being based upon visible-range imagery, the method requires adequate illumination and visual texture. The field data reported here were obtained on an overcast day with diffuse lighting. These conditions provided for excellent visual texture and successful correlation of the stereo image pairs. As described, the optical system is a passive one and cannot be used at night time or during very low light periods. Additionally, depending upon the angle of the sun, it is anticipated that measurements made on clear days could yield results biased by specular reflections from the water surface. Future field campaigns should seek to obtain and validate measurements over a wide range of environmental (lighting, wave characteristics, etc.) conditions. Also, another item for consideration in the future is whether or not stereo visualizations of the water surface obtained from other sensing techniques (infrared, radar, etc.) could provide for satisfactory estimates of wave heights.

A second area for technique improvement lies in increasing the computational efficiency of the image analysis. Presently the time required to process (rectify, correlate, triangulate) an image pair is O (1 min). There are numerous strategies for decreasing this processing time including (among others) a multi-pass approach and the use of successive images to highly constrain the correspondence problem. In the first case, an initial coarse pass is used to obtain a rough estimate of the disparity field and this is then used to constrain a second higher-resolution pass. In the second case, which is appropriate for image sequences obtained at a high frequency, the disparity field from one image pair is used to highly constrain the matching process in the next image pair, since it is known that the waves will move only slightly during the time interval between successive images. Also current hardware developments where more and more parallel processes are assigned to dedicated (mostly graphic) parallel processors are of interest in reducing computational time. Since the time consuming correlation process is parallel in its essence it might quite easily be implemented on a parallel system. In addition to reducing image processing time, data acquisition can be optimized by using a ‘burst’ strategy, whereby computational tasking alternates between acquiring a sequence of images and then processing them. Combining these improvements will advance stereo imaging of coastal waves towards a real time observation system.

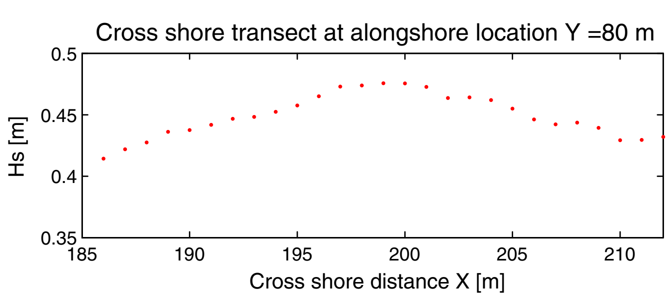

Fig. 18. Cross shore transect of significant wave height Hs.

Acknowledgements

Sierd de Vries and Matthieu de Schipper were funded by the innovation program Building with Nature. The Building with Nature program is funded from several sources, including the Subsidieregeling Innovatieketen Water (SIW, Staatscourant nrs 953 and 17009) sponsored by the Dutch Ministry of Transport, Public Works and Water Management and partner contributions of the participants to the Foundation EcoShape. The program receives co-funding from the

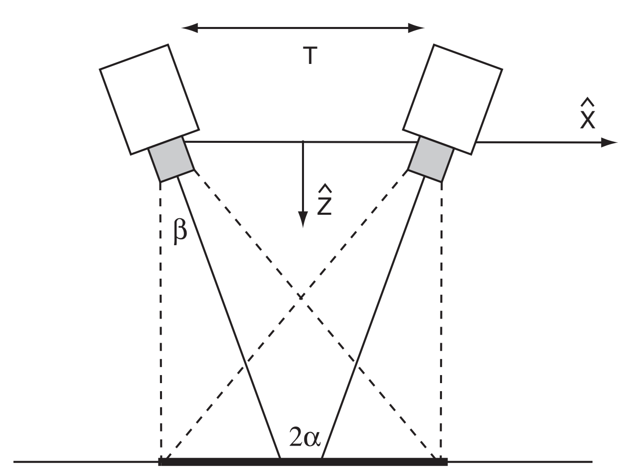

Fig. 20. Schematic of experimental configuration for the purposes of quantization error estimation. Figure adapted from (Benetazzo, 2006).

European Fund for Regional Development EFRO and the Municipality of Dordrecht. D. Hill thanks the Technical University of Delft for the financial support provided by the TU Grants Program. Jaap van Thiel de Vries is acknowledged for providing the Delta Flume laboratory dataset. Irv Elshoff and Deltares are acknowledged for their supporting role in the field work and camera calibration.

Appendix A. Calibration parameters

Camera parameters for the laboratory and field experiments.

		Lab		Field	
		C1	C2	Left	Right
Intrinsic parameters					
Principal point (pixels)	u_0	696.0	696.0	669.8	679.6
Principal point (pixels)	v_0	520.0	520.0	532.7	528.1
Focal length (mm)	f_l			12	12
Skewness	s	0	0	0	0
Pixel ratio	λ_u	1.004	1.004	1.006	1.006
Distortion parameters					
	k_1	$-3.050e^{-7}$	0	$-6.171e^{-8}$	$-6.055e^{-8}$
	k_2	0.0223	0	0.0219	0.0212
Extrinsic parameters					
Roll (°)	σ	−87.62	−90.58	1.49	0.24
Pitch (°)	τ	55.06	45.98	74.39	74.15
Azimuth (°)	ϕ	89.49	91.40	−2.94	−7.42
Location (m)	X_c	190.409	196.822	200.00	208.99
Location (m)	Y_c	0.000	0.049	5.00	5.00
Location (m)	Z_c	8.511	8.453	11.17	11.38

References

Adrian, R., 1991. Particle-imaging techniques for experimental fluid mechanics. Annual Review of Fluid Mechanics 23, 261–304.

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

Catalan, P., Haller, M., 2008. Remote sensing of breaking wave phase speeds with application to non-linear depth inversions. Coastal Engineering 55, 93–111.

Davidson, M., Koningsveld, M.V., de Kruif, A., Rawson, J., Holman, R., Lamberti, A., Medina, R., Kroon, A., Aarninkhof, S., 2007. The CoastView project: developing video-derived Coastal Sate Indicators in support of coastal zone management. Coastal Engineering 54, 463–475.

Guza, R., Thornton, E., 1980. Local and shoaled comparisons of sea-surface elevations, pressures and velocities. Journal of Geophysical Research-Oceans and Atmospheres 85 (NC3), 1524–1530.

Haller, M., Lyzenga, D., 2003. Comparison of radar and video observations of shallow water breaking waves. IEEE Transactions on Geoscience and Remote Sensing 41 (4), 832–843.

Hartley, R., Zisserman, A., 2000. Multiple view geometry in computer vision. Cambridge University Press.

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

Holland, K., Puleo, J., Kooney, T., 2001. Quantification of swash flows using video-based particle image velocimetry. Coastal Engineering 44, 65–77.

Holman, R., Stanley, J., 2007. The history and technical capabilities of Argus. Coastal Engineering 54, 477–491.

Holthuijsen, L., 1983. Stereophotography of ocean waves. Applied Ocean Research 5 (4), 204–209.

Irish, J.L., White, T.E., 1998. Coastal engineering applications of high-resolution lidar bathymetry. Coastal Engineering 35 (1–2), 47–71.

Jessup, A., Zappa, C., Loewen, M., Hesany, V., 1997. Infrared remote sensing of breaking waves. Nature 385, 52–55.

Krogstad, H.E., Barstow, S.F., 1999. Satellite wave measurements for coastal engineering applications. Coastal Engineering 37 (3–4), 283–307.

Lippmann, T., Holman, R., 1989. Quantification of sand-bar morphology—a video technique based on wave dissipation. Journal of Geophysical Research-Oceans 94 (C1), 995–1011 JAN 15.

Monahan, E., 1971. Oceanic whitecaps. Journal of Physical Oceanography 1 (2), 139–144.

Piepmeier, J., Waters, J., 2004. Analysis of stereo vision based measurements of laboratory water waves. IEEE International Geoscience and Remote Sensing Symposium Proceedings: Science for Society: Exploring and Managing a Changing Planet, pp. 358–3591.

Sandidge, J.C., Holyer, R.J., 1998. Coastal bathymetry from hyperspectral observations of water radiance. Remote Sensing of Environment 65 (3), 341–352.

Santel, F., Linder, W., Heipke, C., 2004. Stereoscopic 3d-image analysis of sea surface. Proceedings of the ISPRS Commission V Symposium. Istanbul, Turkey, pp. 708–712. July.

Scharstein, D., Szeliski, R., 2002. A taxonomy and evaluation of dense two-fram stereo correspondence algorithms. International Journal of Computer Vision 47 (1/2/3), 7–42.

Stockdon, H., Holman, R., 2000. Estimation of wave phase speed and nearshore bathymetry from video imagery. Journal of Geophysical Research 105 (C9), 22015–22033.

van Thiel de Vries, J., 2009. Dune erosion during storm surges. Ph.D. thesis, Delft University of Technology.

van Thiel de Vries, J., van Gent, M., Walstra, D., Reniers, A., 2008. Analysis of dune erosion processes in large-scale flume experiments. Coastal Engineering 55 (12), 1028–1040.

Wanek, J., Wu, C., 2005. Automated trinocular stereo imaging system for three-dimensional surface wave measurements. Ocean Engineering 33, 723–747.

Wolf, P., 1983. Elements of Photogrammetry. McGraw-Hill Book Company.