

A Horizon-Tracking Method for Shipboard Video Stabilization and Rectification*

MICHAEL SCHWENDEMAN AND JIM THOMSON

Applied Physics Laboratory, University of Washington, Seattle, Washington

(Manuscript received 19 March 2014, in final form 27 August 2014)

ABSTRACT

An algorithm is presented for the stabilization and rectification of digital video from floating platforms. The method relies on a horizon-tracking technique that was tested under a variety of lighting and sea-state conditions for 48 h of video data over 12 days during a research cruise in the North Pacific Ocean. In this dataset, the horizon was correctly labeled in 92% of the frames in which it was present. The idealized camera model assumes pure pitch-and-roll motion, a flat sea surface, and an unobstructed horizon line. Pitch and roll are defined along the camera look direction rather than in traditional ship coordinates, such that the method can be used for any heading relative to the ship. The uncertainty in pitch and roll is estimated from the uncertainties of the horizon-finding method. These errors are found to be of the order 0.6° in roll and 0.3° in pitch. Errors in rectification are shown to be dominated by the uncertainty in camera height, which may change with the heave motion of a floating platform. The propagation of these errors is demonstrated for the breaking-wave distribution $\Lambda(c)$. A toolbox for implementation of this method in MATLAB is shared via the MATLAB File Exchange.

1. Introduction

The use of digital video as a measurement tool has grown tremendously in recent years, as the cost of making and storing high-quality digital videos has decreased. Standard electrooptical (EO) video, which records visible light in the wavelengths spanning 400–700 nm, is part of a larger category of passive remote sensors that also includes infrared (IR) cameras and hyperspectral imaging systems (Holman and Haller 2013; Dickey et al. 2006). These passive systems complement the active remote sensors such as radar and lidar. Oceanographic applications of remote sensing include direct measurements of whitecap coverage (Callaghan and White 2009) and sea ice concentration (Weissling et al. 2009). In addition, underlying properties such as wave dissipation, bathymetry, and currents can sometimes be inferred from remote measurements (Sutherland and Melville 2013; Stockdon

and Holman 2000; Chickadel et al. 2003). Alternatively, hyperspectral measurements of optical properties of the upper water column (i.e., ocean color) may be used to determine concentrations of certain biological quantities (IOCCG 2000).

Across these modes and applications, quantitative use of camera imagery benefits greatly from knowledge of the camera location and orientation, or “pose.” Camera pose information is needed specifically for two common tasks: stabilization and rectification. Stabilization is the removal of camera motion from a set of images, resulting in images appearing as though recorded from a fixed camera. Rectification is the projection of one or more images into a common coordinate system. When images are projected onto a map of the earth’s surface, as from an airplane or satellite, it is commonly termed *orthorectification*.

For a camera mounted to a fixed platform, the pose may be measured once and applied to all images. Such configurations are not often available at sea, but they may be approximated in exceptionally stable ships [such as R/P *Floating Instrument Platform (FLIP)*; see Gemmrich et al. 2008; Zappa et al. 2012; Sutherland and Melville 2013] or in relatively calm conditions (e.g., Thomson et al. 2009; Thomson and Jessup 2009). Reliance on specific ships or conditions is not ideal, however, as it limits the availability of these measurements. Alternatively, the pose may be recorded

* Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JTECH-D-14-00047.s1>.

Corresponding author address: Michael Schwendeman, Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Box 355640, Seattle, WA 98105-6698.
E-mail: mss28@u.washington.edu

continuously from an external system, usually an inertial motion unit (IMU). IMUs measure linear acceleration and angular velocity with a combination of accelerometers and gyroscopes. The position of the IMU is then estimated from dead reckoning. In Kleiss and Melville (2010, 2011) and Romero et al. (2012), an IMU was used to gather the positional data needed to rectify images taken from an airborne system. In Schwendeman et al. (2014), a shipboard camera was actively stabilized with an inertial pan-tilt system. There are two disadvantages of this strategy. First, inertial measurements are sensitive to noise and drift, such that accurate systems can be prohibitively expensive. Second, such systems require precise synchronization with the video frames, otherwise matching errors will occur.

A more attractive option is to use the video frames themselves to estimate the camera orientation. Problems of this type usually require ground control points (GCPs) with known x , y , and z coordinates (Szeliski 2010). GCPs may be available in coastal applications when land makes up a significant portion of the camera field of view, as in Holland et al. (1997). In the open ocean, however, there are no GCPs.

In the following paper, a method for determining the pitch and roll of a shipboard camera is described from the location of the horizon line in the image. Similar strategies have been used for stabilization of cameras for marine surveillance (Cao and Zhang 2007; Morris et al. 2007; Fefilatyeu et al. 2012) and unmanned aerial vehicle (UAV) flight control (Bao et al. 2005; Thurrowgood et al. 2009; Zhang et al. 2011). The focus here is an image rectification to sea surface coordinates for oceanic measurement tasks. One such application is the calculation of the wave-breaking distribution $\Lambda(c)$ (Phillips 1985). Calculating this distribution requires the measurement of the length L and speed c of many breaking-wave events.

In section 2, the horizon-finding method is described using an example image from a research cruise in the North Pacific. In section 3, this algorithm is applied to the full video dataset and its performance is evaluated. The uncertainty and sensitivity in the method are examined in section 4, using $\Lambda(c)$ as a case study. Finally, section 5 concludes and points the user to a publicly available MATLAB toolbox for implementation of these algorithms.

2. Methods

The general problem of interest is outlined in Fig. 1. A camera looks obliquely outward over a flat sea surface from a known height H at a constant heading (or azimuth) γ . The horizon line is unobstructed and visible in

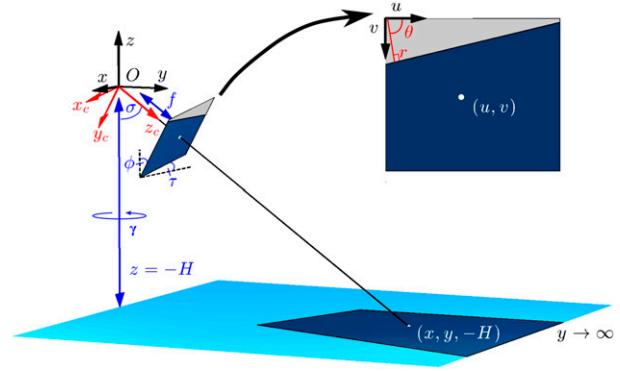


FIG. 1. Problem schematic, defining the camera parameters and coordinate systems. The sea surface is shown in blue, with dark blue corresponding to the camera field of view, while gray areas denote sky. The small rotated square represents the camera sensor (not to scale), and the top-right square shows the resulting image.

the image. The camera orientation is fixed except for two angles: pitch ϕ and roll τ , which are unknown and are defined in the camera reference frame rather than the ship reference frame. Throughout the remainder of this paper, the incidence angle, $\sigma = \pi/2 - \phi$, is used rather than pitch ϕ for notational convenience in the calculations.

The ultimate goal of the following procedure is to determine σ and τ , such that it becomes possible to project a point located at pixel (u, v) to sea surface coordinates $(x, y, -H)$, and vice versa. This goal will be achieved by relating the equation of the horizon line to σ and τ , as detailed below.

a. Data collection

The images shown here were collected during a research cruise on board the R/V *New Horizon* in the North Pacific in 2012. The ship sailed from San Diego, California, to Ocean Station Papa (50°N , 145°W) and back over the course of 3 weeks. On 12 of these days, the ship held station for several hours to make a variety of wind and wave measurements. These measurements include approximately 48 h of video, taken from a 3.6-mm “bullet” camera with $1/3$ -in. high-resolution Sony ExView black-and-white (B/W) charge-coupled device (CCD). The data were sampled at 30 Hz and 8-bit grayscale, and compressed to MPEG-2 video. For the first half of the experiment, the camera was mounted on the top rail ($H = 10.7$ m) facing off the starboard side. On 7 October, it was moved to the ship’s centerline (still on the top rail) to be able to look off either port or starboard.

The camera was attached to a pan-tilt stabilization system (Directed Perception PTU-D100), which actively corrected for rotation in the incidence ($= 90^\circ - \text{pitch}$) and azimuth angles. This was the same stabilization

package used in Schwendeman et al. (2014), in which a time-consuming manual horizon-finding method was used to compensate for the drift in incidence, while roll was left uncorrected. By using active stabilization, the camera orientation is decoupled from the ship's pitch, roll, and heading angles. Therefore, we must make the distinction that in the following sections, all angles are of the camera relative to the sea surface, which are unrelated to the ship's orientation. Although helpful in maintaining a given field of view, the active stabilization has been shown to be insufficient for quantitative processing of moving signals, such as breaking waves. This motivates the postprocessing method that follows.

b. Horizon-finding algorithm

This image-based stabilization depends on an automated method for fast, robust, and accurate detection of the horizon line. Possible complications are poor lighting conditions, partial horizon views, and noise from features in the sea surface (e.g., whitecaps) and sky (e.g., clouds). As noted in section 1, horizon detection has proven useful for a number of applications, and several strategies have been described in the literature. Recently, Fefilat'ev et al. (2012) developed a "separation criterion" approach for finding the optimal horizon line from a rapidly moving, buoy-mounted camera. While their problem requires a very robust detection algorithm, here we find that a simpler, faster approach is acceptable for our data and application.

Figure 2a shows an example video frame with the horizon clearly visible in the image. In Fig. 2b, the image has been filtered with the Canny edge detector, which isolates and thresholds local maxima in the image gradient along the direction perpendicular to the gradient direction (Canny 1986). Because the Canny detection method uses the gradient direction information, it has a number of advantages over other edge filters. First, the user can specify to return only the maximum pixel for a given edge (called *nonmaximal suppression*), which leads to sharper edges. Additionally, the Canny detector can find connected edges with a lower threshold through hysteresis and edge linking (see Szeliski 2010 for more details). Figure 2b shows the result of MATLAB's Canny detector with an automated threshold that identifies 5% of pixels as edges.

Next, the binary edge image is passed to the Hough transform, a widely used method for identifying lines in images (Duda and Hart 1972; Illingworth and Kittler 1988). In the Hough transform, candidate lines are parameterized in polar coordinates by r , the distance of the line to the image origin, and θ , the direction of the line, as shown in Fig. 1. Then, the number of edge pixels in the

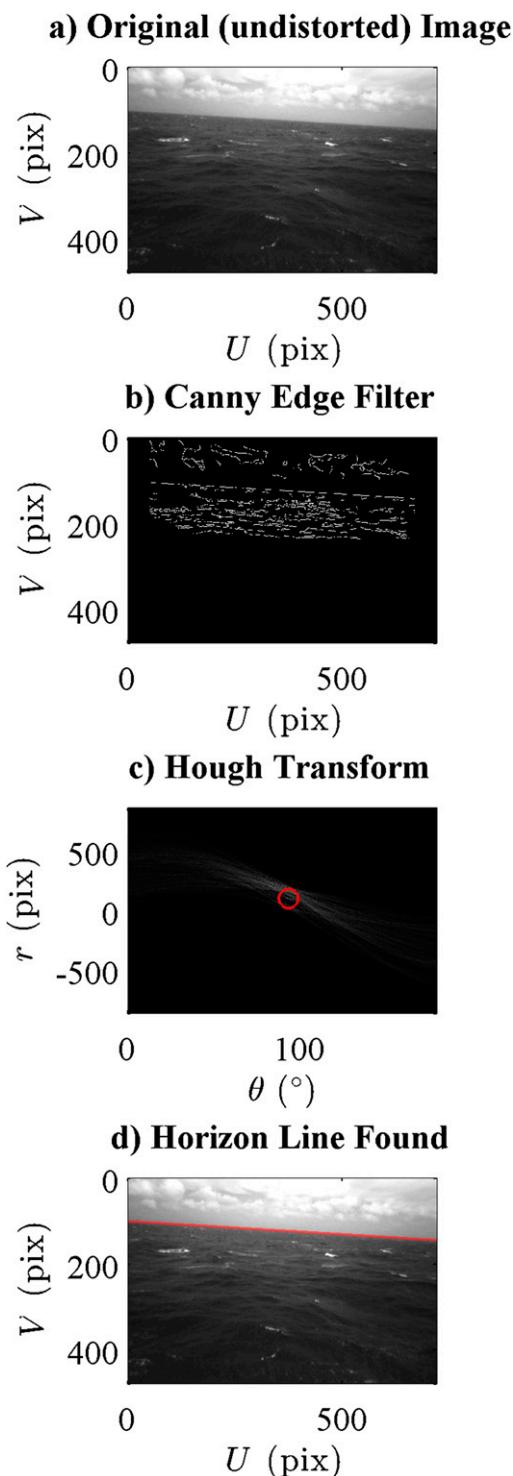


FIG. 2. An example of the horizon-finding procedure for a frame from 1958 UTC 8 Oct. (a) The original image after removal of distortion. (b) The binary image resulting from filtering with the Canny edge detector. (c) The Hough transform accumulator array, with maximum value circled in red. (d) The original image with computed horizon line overlaid. For this image, the calculated camera angles are $\tau = 3.9^\circ$ and $\sigma = 78.7^\circ$.

binary edge image is tallied along each line candidate. Thus, an *accumulator array* is assembled, whose elements each correspond to a line in (r, θ) space, and whose values equal the number of edge pixels located along this line. Since the horizon line is usually the longest or most prominent edge in the image, its r and θ parameterization corresponds to the maximum value of the accumulator array. Figure 2c shows the resulting accumulator array with the maximum value circled. Finally, Fig. 2d shows the resulting line drawn on the original frame, exactly aligned with the horizon.

c. *Projection equations*

With the horizon line found, the camera incidence and roll can be calculated from the camera projection equation (also called *perspective transform* or *homography*). The geometric basis for the projection equation is the pinhole camera model, in which all light rays pass through a principal point O before hitting the sensor located one focal length f from O. Although the sensor physically lies behind the principal point, it can be easier to visualize the problem with it in front, as in Fig. 1.

A necessary first step is to determine the camera’s individual distortion coefficients and intrinsic parameters. This is done using a geometric calibration technique, based on the algorithms of Zhang (2000) and Heikkila and Silven (1997), that is available in MATLAB’s Computer Vision System Toolbox (or similar open source software packages). Distortion can be thought of as the deviation of the camera from an ideal pinhole camera and is often noticeable for wide-angle lenses. After correcting for distortion, the horizon is seen as a straight line in the image, as in Fig. 2. Here, a distortion correction is used with two radial coefficients, no skew, and no tangential distortion, resulting in average errors

TABLE 1. Distortion coefficients and intrinsic parameters for the 3.6-mm “bullet” camera used in this experiment.

Parameter name	Value
Number of columns n_u	640
Number of rows n_v	480
First radial distortion coefficient κ_1	-0.601
Second radial distortion coefficient κ_2	0.364
Focal length per pixel width f_u	673.4
Focal length per pixel height f_v	610.0
Image column center c_u	339.2
Image row center c_v	244.5

of less than one pixel from a pinhole camera. With distortion removed, the camera is fully characterized by four intrinsic parameters. The first of these are f_u and f_v , which are the camera focal length expressed in terms of the pixel width and pixel height, respectively. The other two are c_u and c_v , which correspond to the (u, v) coordinates, respectively, of the pixel located directly behind O, otherwise known as the image origin. The distortion coefficients and intrinsic parameters for the camera used in this experiment are listed in Table 1.

The invertible perspective equation for projecting real-world (x, y, z) coordinates to (u, v) pixel coordinates is

$$s[u, v, 1, d]^T = [\mathbf{P}][x, y, z, 1]^T, \tag{1}$$

where \mathbf{P} is the 4×4 “camera matrix,” d is the inverse depth or “disparity,” and s is an arbitrary scale factor. Equation (1) can be inverted to find (x, y) from (u, v) by assuming $s = d = 1$, and then normalizing the resulting (x_0, y_0, z_0) coordinates by $-z_0/H$, leading to coordinates at $(x, y, -H)$, that is, the flat sea surface. For a pure rotation in σ and τ , the form of \mathbf{P} is

$$[\mathbf{P}] = \begin{bmatrix} f_u & 0 & c_u & 0 \\ 0 & f_v & c_v & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\tau & \sin\tau \cos\sigma & \sin\tau \sin\sigma & 0 \\ \sin\tau & -\cos\tau \cos\sigma & -\cos\tau \sin\sigma & 0 \\ 0 & \sin\sigma & -\cos\sigma & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{2}$$

where $f_u, f_v, c_u,$ and c_v are the intrinsic parameters described above.

Angles σ and τ are uniquely determined by the location of the horizon line. In the limit of $y \rightarrow \infty$, Eq. (1) leads to u and v , which lie along a straight line in the image corresponding to the horizon line, and defined by r and θ as found in the previous section. The equations relating these parameters are

$$\tau = \tan^{-1} \left[\frac{-f_u}{f_v \tan(\theta)} \right] \tag{3}$$

$$\sigma = \tan^{-1} \left[\frac{f_u \sin(\tau) \cos(\theta) - f_v \cos(\tau) \sin(\theta)}{r - c_u \cos(\theta) - c_v \sin(\theta)} \right]. \tag{4}$$

Note that the above equations do not include the third camera angle, azimuth, given by γ in Fig. 1. The azimuth

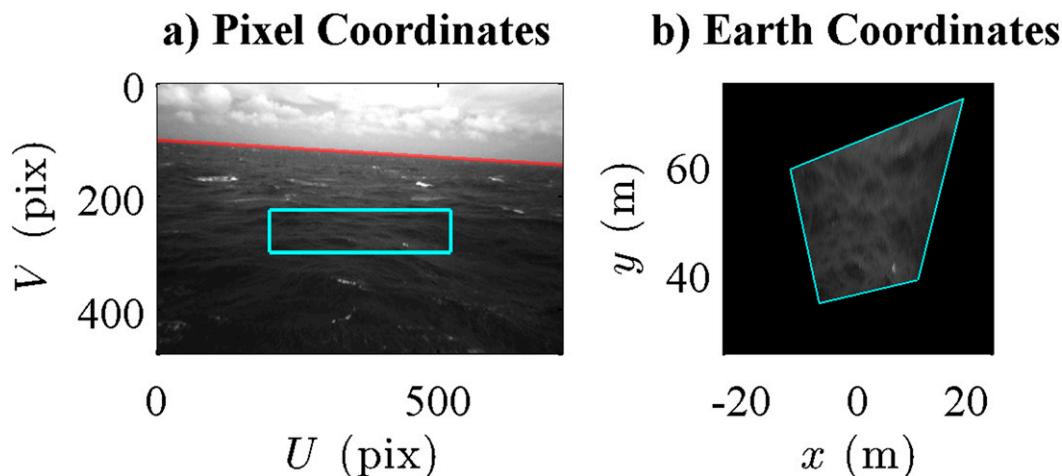


FIG. 3. Example rectification using the image from Fig. 2. (a) The image in pixel coordinates, with horizon detected correctly (red) and a sample rectangle (blue). (b) The projection of the pixels within the blue rectangle to Earth coordinates.

angle may be needed for some applications if it varies significantly across images and directional information is required. Like camera height, however, azimuth cannot be determined with the horizon-finding procedure. Therefore, it must be measured or estimated from some other data source, such as the ship's heading. With the pan-tilt stabilization system used here, the camera azimuth is kept relatively constant over time scales of several minutes, but this makes it independent of the ship's heading. We therefore define the line of sight of the camera to be at azimuth = 0° unless otherwise noted.

Figure 3 demonstrates the rectification from image coordinates to world coordinates, for the same example as Fig. 2. The red line again shows the accurately detected horizon line, which is used to measure roll and incidence. For the example shown, the camera was found to be oriented at 3.9° roll and 78.7° incidence. The blue rectangle in the original image maps to an oblique near-trapezoidal shape in Earth coordinates.

3. Results

a. Horizon-finding statistics

Next, we apply the method to the full dataset of 48 h over 12 days. Images were subsampled at 30-s intervals for a total of 5761 frames. Figure 4 shows 12 sample frames from this set, one for each day of the experiment. This figure demonstrates the robustness of the horizon-finding procedure over the large variety of lighting conditions seen during the experiment. Then all the frames are projected as though taken from a stationary camera positioned at 0° roll and 75° incidence. Example videos showing this stabilization for several minutes at a time are also available as online supplemental material.

All 5761 processed images were manually reviewed and classified into one of three categories: horizon visible and correctly identified, horizon visible but incorrect, and horizon not visible. The results of this analysis are shown in Table 2, separated by date. The rightmost column of the table shows the percentage of frames in which a visible horizon has been correctly detected. The results totaled over the entire experiment are shown in the bottom row. Overall, 92.3% of visible horizon lines were correctly identified with the Hough transform method.

The accuracy is quite variable over the experiment, reaching a minimum of 70.8% on 5 October, compared with a perfect 100% on 11 October. This accuracy is primarily a function of lighting, as noted in the Conditions column. The most unfavorable conditions are on clear or partly cloudy days when the sun is low in the sky, resulting in high glare and saturation of the image near the horizon. The worst days for these conditions are 5 and 6 October. This motivated moving the camera to the ship's centerline on 7 October, allowing the camera to be directed either port or starboard to avoid glare, thereby dramatically improving the horizon detection. Secondary effects include low light in the early morning or evening and rain drops on the window of the camera housing.

b. Quality control

It is crucial to be able to identify when the horizon-finding algorithm fails, without relying on the manual inspection described above. Given a series of sequential images, it is possible to identify an incorrect horizon as an outlier from the expected smooth variation in σ and τ . Such a method is unsuitable, however, in cases where the

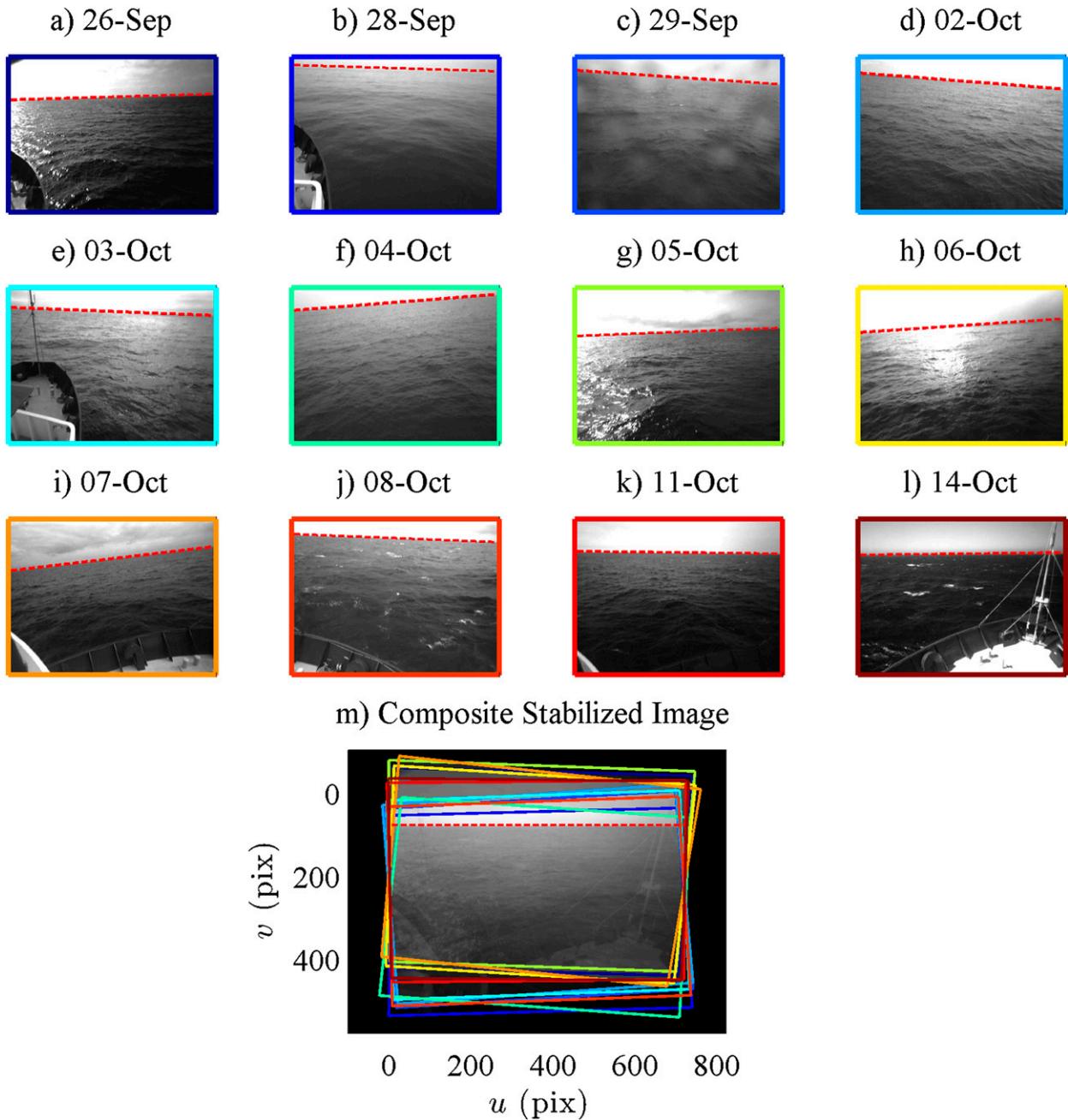


FIG. 4. (a)–(l) Twelve example video frames showing successful horizon detection (red dashed line), each from a different day of the experiment. (m) Composite image formed by transforming each image as though taken from a camera at 75° incidence and 0° roll, and averaging over all images. The red dashed line shows the expected horizon for this orientation, and the colored rectangles mark the outline of each transformed image.

horizon is rapidly moving or when nonsequential images are examined.

Instead, a predictive variable is defined to identify spurious horizons. Since the horizon is chosen as the maximum of the Hough transform, the prominence of this maximum is related to the strength of the horizon line. This first peak is compared with the second

highest peak, and the ratio m of these peaks is used as an inverse quality metric. To avoid values near a prominent horizon, the second maximum is found a distance of $r > 10$ pixels and $\theta > 2.5^\circ$ from the first peak. A ratio $m = 1$ indicates that both maxima are identical, whereas $m = 0$ corresponds to an infinitely large primary maximum.

TABLE 2. Horizon-finding results, separated by day, and totaled over the entire dataset. Images were output every 30 s and manually classified. Percent correct column indicates the percentage of visible horizon lines that are correctly identified in the Hough transform.

Date (2012)	Conditions	Horizon visible, correct	Horizon visible, incorrect	Horizon not visible	Percent correct (%)
26 Sep	Partly cloudy	208	17	29	92.4
28 Sep	Low light early	189	28	37	87.1
29 Sep	Occasional rain on glass	145	31	26	82.4
2 Oct	Partly cloudy, occasional glare	470	8	30	98.3
3 Oct	Partly cloudy, late glare	296	58	13	83.6
4 Oct	Mostly cloudy	377	1	3	99.7
5 Oct	Clear, frequent glare	318	131	59	70.8
6 Oct	Partly cloudy, frequent glare	299	74	8	80.2
7 Oct	Early glare, late darkness	959	37	40	96.3
8 Oct	Partly cloudy	898	1	5	99.9
11 Oct	Mostly cloudy	330	0	1	100
14 Oct	Early glare, clear later	598	38	0	94.0
Total		5086	424	251	92.3

The horizon-finding method can fail for one of two reasons: either the horizon is not visible in the image or the horizon is not the strongest line in the image. In both cases, a large value of m is expected. Using the full dataset of 5761 images, the ability of m to diagnose false horizons is examined. For a given threshold value m_t images with $m < m_t$ are predicted to be successful horizon identifications (positives), while images with $m > m_t$ are predicted failures (negatives). By comparing with the manual classification, images are characterized as true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). Positive predictive value (PPV), or “precision,” is the ratio of true positives to all predicted successful horizon identifications,

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (5)$$

Conversely, true positive rate (TPR), or “recall,” is the ratio of true positives to all actual successful horizon identifications,

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (6)$$

Figure 5 shows the precision and recall curves made by varying m_t . Daily curves are shown, as well as for the full dataset. As with Table 2 the results are encouraging overall, but they depend greatly on the individual day. For $m_t = 0.8$, the total PPV is 99.6% (four false positives per 1000 predicted positives) and the total TPR is 96.2% (3.8% unnecessarily discarded images). For the same threshold, the minimum daily PPV and TPR values are 98.9% and 81.8%. Thus, if anything, this cutoff has the potential to be overly conservative when predicting failure in the horizon-finding algorithm.

4. Discussion

a. Uncertainty analysis

Three uncertainties can be quantified in this method: the uncertainty in the horizon line (r, θ), the uncertainty in camera angles (σ, τ), and the uncertainty in the projected real-world coordinates (x, y, z).

The uncertainty in (r, θ) can be due to the resolution of the camera, the chosen quantization in the Hough transform, and uncertainty as a result of ambiguity in the horizon location. Camera intrinsic calibration indicates that the pinhole model is only accurate to within 1 pixel. Similarly, the Hough transform is discretized by 1 pixel in r . The resolution in θ is specified by the user and was

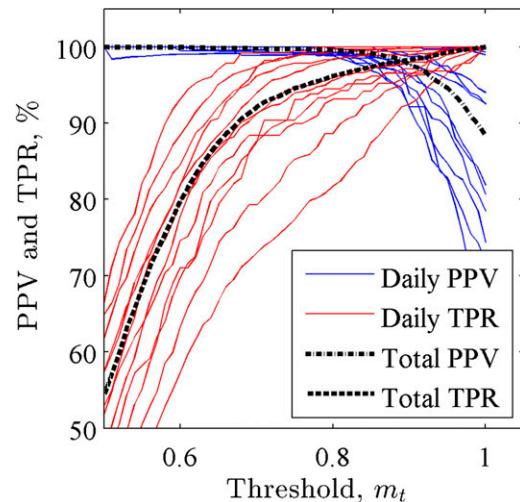


FIG. 5. Plots of PPV and TPR are plotted as a function of the threshold of the Hough transform peak ratio m_t . Thin curves in blue (PPV) and red (TPR) are daily values (12 lines each), while the dashed (PPV) and dashed-dotted (TPR) curves are for the full dataset.

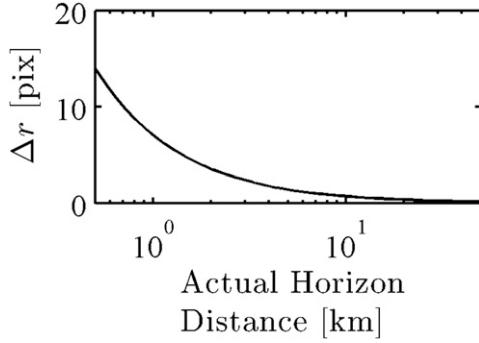


FIG. 6. Uncertainty in image horizon location (Δr , pixels) for finite horizon distances. The line shown is for $\sigma = 75^\circ$, $\tau = 0^\circ$, and $H = 10.7$ m.

chosen here to be 0.5° to reduce computation time and to ensure robust statistics in the maximization.

Additionally, there is some uncertainty in the actual horizon location. The assumption made in deriving Eqs. (3) and (4) was of a horizon located infinitely far from the camera. For a number of reasons, including the curvature of the earth and atmospheric scattering, the distance to the “apparent horizon” is actually finite (see French 1982; Bohren and Fraser 1986). Ignoring atmospheric effects and assuming an observer height much less than the earth’s radius, the apparent horizon distance is

$$D \approx \sqrt{2RH}, \quad (7)$$

where R is the radius of the earth and H is the height of the observer. So, for our camera located at $H = 10.7$ m, the distance to the apparent horizon is roughly 11.7 km. This distance can be shortened as a result of atmospheric scattering, which decreases the contrast between sea and sky (Bohren and Fraser 1986). Alternatively, the distance can be lengthened as much as 10% as a result of atmospheric refraction (French 1982).

Figure 6 shows the errors in r from assuming an infinite horizon, for $\sigma = 75^\circ$, $\tau = 0^\circ$, and $H = 10.7$ m. This camera pose is representative of a typical view during this experiment and will be used repeatedly in the following analysis. At the estimated horizon distance of 11.7 km, the error in r is only 0.6 pixels, which is less than the uncertainty noted above. As the apparent horizon nears the camera, the horizon appears lower in the image, and the errors in r increase. This error is not sensitive to changes in σ , but scales proportional to H . Since D scales with $H^{1/2}$, these errors will become slightly more pronounced for higher cameras. This plot is also useful in estimating the errors stemming from using the horizon information under nonideal conditions. For

example, if a coastline obscures the horizon at a distance of 1 km, then an error of 6 pixels can be expected in r .

The uncertainty in σ and τ is directly related to the uncertainty in r and θ through Eqs. (3) and (4). Around the reference orientation ($\tau = 0^\circ$, $\sigma = 75^\circ$), an error of $\pm 0.5^\circ$ in θ leads to errors of $\pm 0.55^\circ$ in τ and $\pm 0.26^\circ$ in σ . Around the same value, errors of ± 1 pixel in r have no effect on τ but correspond to errors of 0.09° in σ . Therefore, under ideal lighting, the horizon method as described here is accurate to within 0.6° in τ and 0.3° in σ .

In section 2, it was shown that the projected coordinates (x, y, z) are functions of the pixel coordinates (u, v) , the intrinsic parameters (f_u, f_v, c_u, c_v) , the camera angles (σ, τ, γ) , and the camera height H . Since z of the water surface is taken to be $-H$, the uncertainty in z is simply equal to the uncertainty in H . However, the errors in x and y are functions of the uncertainties of all the parameters,

$$\begin{aligned} \Delta x = & x(u, v; H + \Delta H, \sigma + \Delta\sigma, \tau + \Delta\tau, \gamma + \Delta\gamma, f_u, f_v, c_u, c_v) \\ & - x(u, v; H, \sigma, \tau, \gamma, f_u, f_v, c_u, c_v), \end{aligned} \quad (8)$$

$$\begin{aligned} \Delta y = & y(u, v; H + \Delta H, \sigma + \Delta\sigma, \tau + \Delta\tau, \gamma + \Delta\gamma, f_u, f_v, c_u, c_v) \\ & - y(u, v; H, \sigma, \tau, \gamma, f_u, f_v, c_u, c_v). \end{aligned} \quad (9)$$

Potential errors in $f_u, f_v, c_u,$ and c_v are not addressed here based on the positive results of the calibration procedure, which showed average errors of less than one pixel in the calibration images. In section 3, it was shown that after stabilization, the error in incidence and roll angles were on the order of $\Delta\sigma \sim 0.3^\circ$ and $\Delta\tau \sim 0.6^\circ$, respectively. The horizon technique cannot be used to estimate H and γ , so these uncertainties must be estimated from other sources. Term ΔH is a function of the sea surface elevation variability, or wave height. During this experiment, the average wave height was roughly 2 m, which will be used as a representative value of ΔH . The specifications for the pan-tilt stabilization unit give azimuth errors of $\Delta\gamma < 0.25^\circ$ for underlying sine-wave motions at 1 Hz, although the system tended to drift in azimuth on longer time scales.

Figures 7 and 8 show the errors in x and y resulting from these sources, around the reference state of $\sigma = 75^\circ$, $\tau = 0^\circ$, $\gamma = 0^\circ$, and $H = 10.7$ m. Figure 7 shows the camera field of view in pixel coordinates, with the area above the horizon colored gray. Pixels are colored by error components Δx and Δy for the errors in $\tau, \sigma, \gamma,$ and H discussed above. In Fig. 8, these errors are shown as a function of y , for vertical slices in the left, middle, and right of the image. The errors grow as y increases

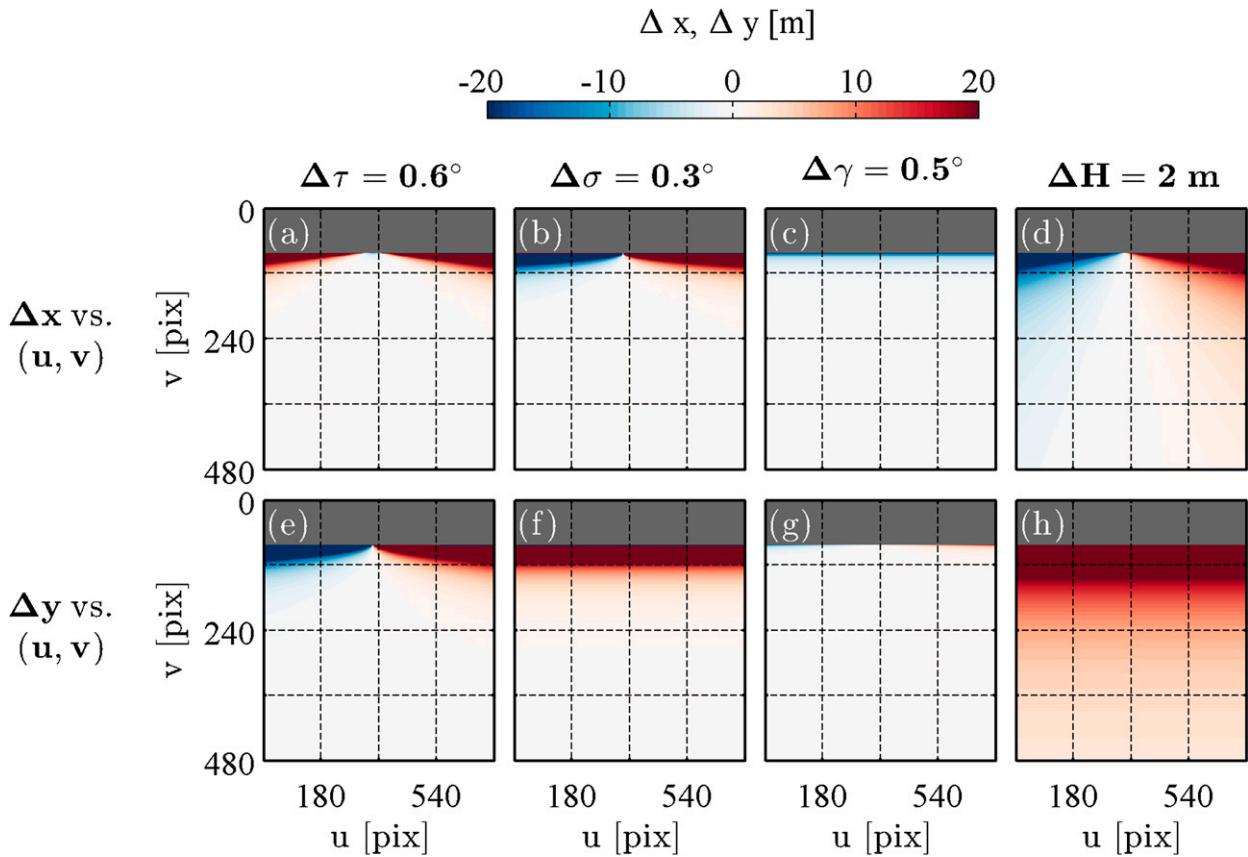


FIG. 7. Uncertainty in rectification around a reference camera orientation of $\sigma = 75^\circ$, $\tau = 0^\circ$, $\gamma = 0^\circ$, and $H = 10.7$ m. The errors are shown in (a)–(d) x and (e)–(h) y at each pixel (u, v) for corresponding errors of (a),(e) $\Delta\tau = 0.6^\circ$; (b),(f) $\Delta\sigma = 0.3^\circ$; (c),(g) $\Delta\gamma = 0.5^\circ$; and (d),(h) $\Delta H = 2$ m. The gray region represents the pixels above the horizon in the image.

for both components and all uncertainty sources. For all sources except γ , uncertainty in y is larger than in x . The most severe source of error is in Δy because of ΔH . For uncertainty in H , the relative errors, $\Delta y/y$ and $\Delta x/x$, are equal to $\Delta H/H$, or 18.6% in this case. So, at a distance of 100 m, Δy is nearly 20 m for $\Delta H = 2$ m, whereas all other errors are less than 5 m at this distance. Although this error analysis was performed around one camera pose, with single values for $\Delta\tau$, $\Delta\sigma$, $\Delta\gamma$, and ΔH , it would be easy to adapt this procedure to different camera views and error magnitudes.

b. Uncertainty in $\Lambda(c)$

Figure 7 demonstrates that errors in the assumed camera height can lead to large uncertainties during rectification, even after stabilization in roll and incidence. The manifestation of these uncertainties will vary depending on the application. As a case study, we demonstrate the result of such errors on estimates of $\Lambda(c)$, a common remotely sensed wave measurement (see, e.g., Gemmrich et al. 2008; Thomson et al. 2009; Kleiss and Melville 2010; Schwendeman et al. 2014).

Term $\Lambda(c)$ is defined as the sum of all breaking-wave crest lengths per unit area and speed. It has become commonly used in estimating wave energy dissipation as a result of breaking (see Phillips 1985). Calculation of $\Lambda(c)$ requires the estimation of wave crest length L_i and speed c_i for many breaking events. Then the calculation is carried out by binning the events by c_i , summing over L_i , and applying the proper normalization. The final formula for $\Lambda(c)$ is

$$\Lambda(c) = \frac{\sum_i L_i}{AN dc}, \quad (10)$$

where N is the total observed events, A is the spatial domain, and dc is the speed bandwidth, such that the summation is over all L_i with speeds $(c - dc/2) < c_i < (c + dc/2)$. If the breaking waves are observed from a moving camera, it is clear that estimates of the crest lengths and speeds will be affected by the camera motion, but it is not apparent how these errors will propagate through the $\Lambda(c)$ computation.

The sensitivity of $\Lambda(c)$ to errors of this kind can be examined through the use of synthetic data. These

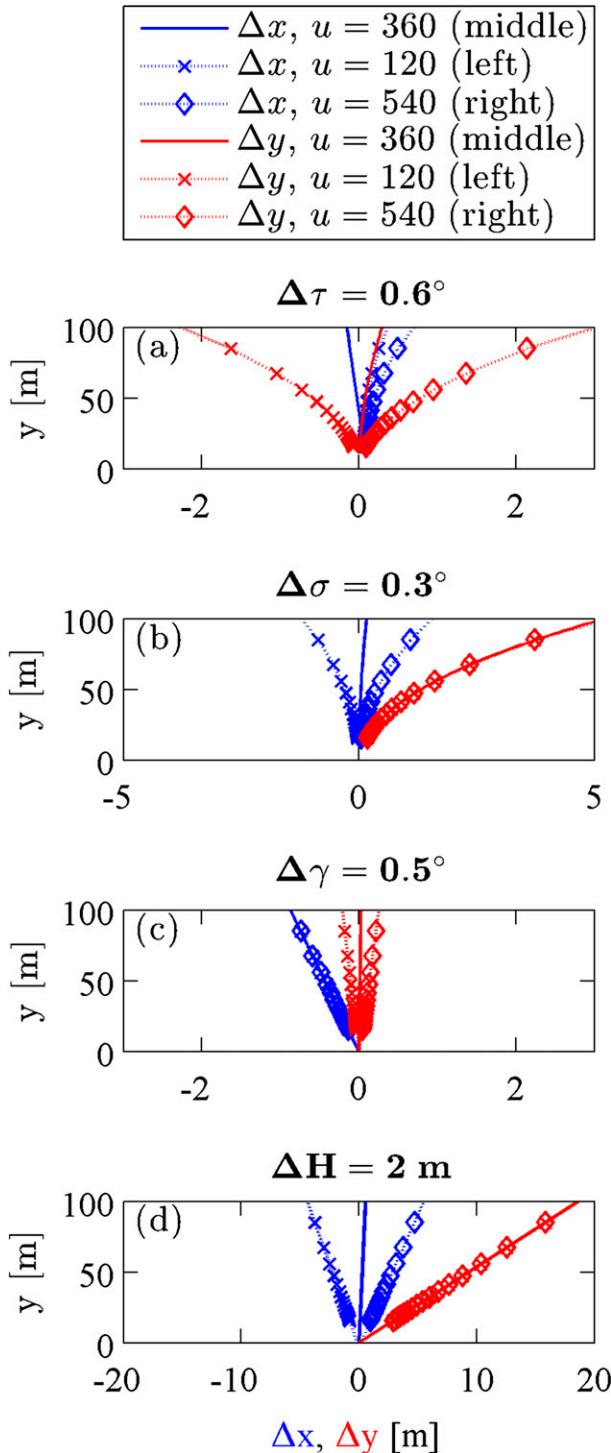


FIG. 8. Uncertainty in rectification around a reference camera orientation of $\sigma = 75^\circ$, $\tau = 0^\circ$, $\gamma = 0^\circ$, and $H = 10.7$ m. The Δx (blue) and Δy (red) errors are shown as a function of distance from the ship (y) for corresponding errors of (a) $\Delta\tau = 0.6^\circ$, (b) $\Delta\sigma = 0.3^\circ$, (c) $\Delta\gamma = 0.5^\circ$, and (d) $\Delta H = 2$ m. Curves are plotted for three constant pixel columns, corresponding to lines in the middle ($u = 360$, solid line), left ($u = 120$, dotted line with crosses), and right ($u = 540$, dotted line with diamonds) of the image.

synthetic data have a known distribution of L and c , making errors easy to determine. Moreover, in practice there are many subtle variations for determining L_i and c_i that are mostly avoided with the synthetic data. The creation of the synthetic data proceeds as follows: Line segments with constant length 2 m are randomly given a starting position within a domain of $30 \text{ m} < y < 50 \text{ m}$ and $-10 \text{ m} < x < 10 \text{ m}$. They are assigned a speed from a Gaussian distribution centered around 3 m s^{-1} with a standard deviation of 1 m s^{-1} . They are propagated forward with that speed, over a duration of 2 s, with their position recorded every $dt = 1/(15 \text{ Hz})$. Next, these crests are “observed” from a camera at the familiar pose of $\sigma = 75^\circ$, $\tau = 0^\circ$, $\gamma = 0^\circ$, and $H = 10.7$ m. The camera is subject to a sinusoidal oscillation in height, with $\Delta H = 2$ m (i.e., amplitude of 1 m), at a period of $T = 4$ s, to simulate a shipboard camera in a rough sea state. Each wave crest is randomly given a start time, which determines the camera phase during the event. The observer, however, is not aware that the camera height is changing and therefore the observer attributes the camera motion to the crests themselves. Examples of true and apparent crest motions are shown in Figs. 9a,b for two scenarios: crests moving directly across the image, in the $+x$ direction, and directly away from the camera, in the $+y$ direction.

Figures 9c,d show the results of calculating $\Lambda(c)$ as in Eq. (10) for the true and apparent crest motion. Over 10000 crests were simulated in each case to ensure robust statistics. In making Fig. 9c, L_i is calculated as the Δy between the two endpoints of the crest, and c_i is $\Delta x/dt$ (vice versa for Fig. 9d). The true $\Lambda(c)$ distribution is shown in black. Five cases are tested for each scenario, the first being the base case from Figs. 9a,b ($\Delta H = 2$ m, $T = 4$ s), in blue. The red line shows the results for halving ΔH ($\Delta H = 1$ m, $T = 4$ s) and the green line for doubling T ($\Delta H = 2$ m, $T = 8$ s). The cyan curve is a low-motion case ($\Delta H = 0.25$ m, $T = 4$ s). The magenta distribution is for a similar scenario, with the exception that the camera is looking down from a nadir position ($\sigma = 0^\circ$).

The results from the synthetic data lead to a number of conclusions. First, in all cases, sinusoidal errors in camera height tended to broaden the observed $\Lambda(c)$ from the true distribution, while the total sum of breaking crests [equal to the integral of $\Lambda(c)$] stayed within 1%. Thus, the measurement of crest length was unbiased by the camera motion, but the crests speeds were biased toward more extreme values. This is especially true for the scenario with crests moving away from the camera, where for the large camera motions, the peak is obscured completely. This orientation appears to be a worse-case scenario for observing $\Lambda(c)$, although the errors may be manageable for lower sea states, as the

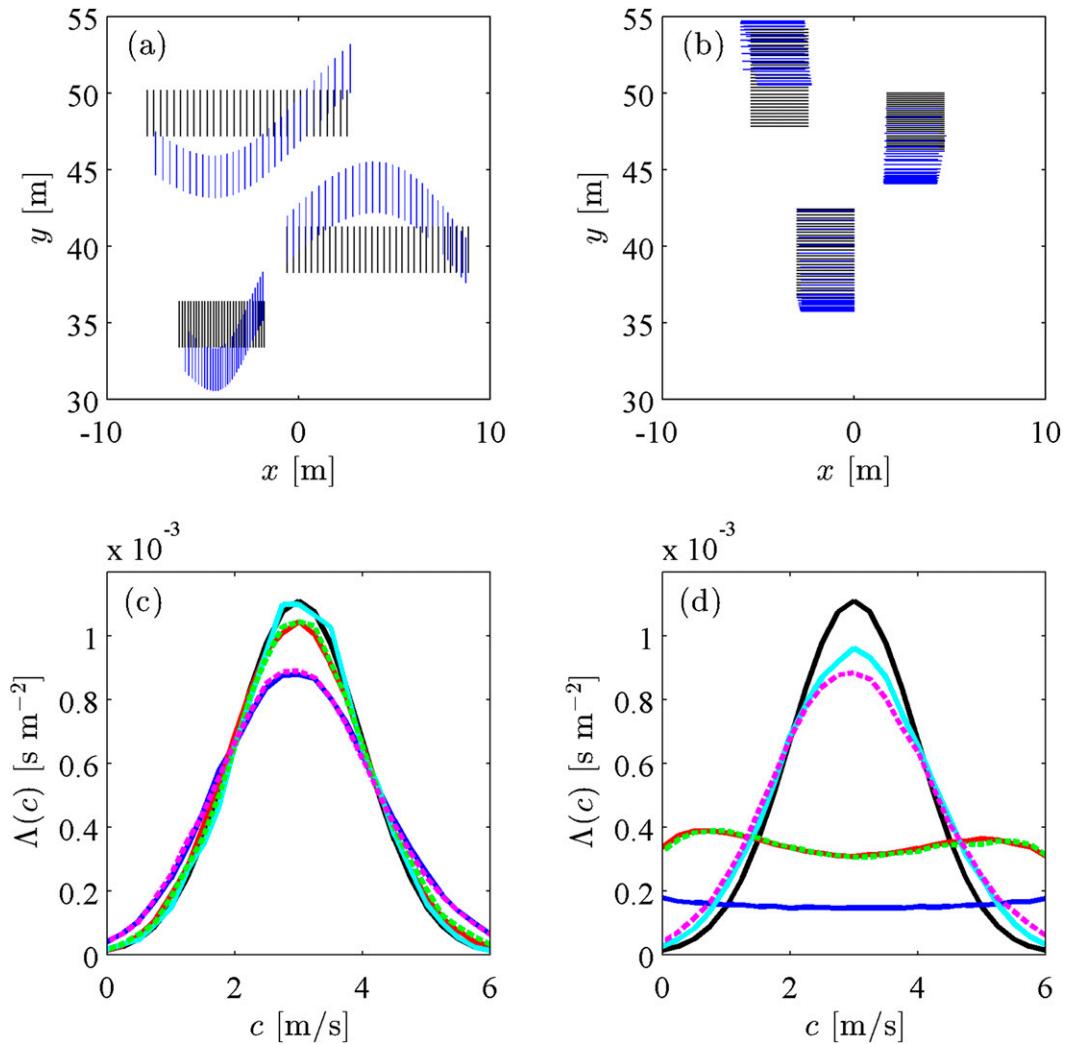


FIG. 9. Error in $\Lambda(c)$ as a result of varying camera height with synthetic breaking waves for (a),(c) breaking waves propagating in the $+x$ direction and (b),(d) breaking waves propagating in the $+y$ direction. (a),(b) Three examples of the true breaker propagation (black) and the apparent propagation (blue) for sinusoidal camera motion of $\Delta H = 2 \text{ m}$ at $T = 4 \text{ s}$. (c),(d) The resulting $\Lambda(c)$ for Gaussian input centered at 3 m s^{-1} (black) for $\Delta H = 2 \text{ m}$, $T = 4 \text{ s}$ (blue), $\Delta H = 1 \text{ m}$, $T = 4 \text{ s}$ (red), $\Delta H = 2 \text{ m}$, $T = 8 \text{ s}$ (green, dashed), $\Delta H = 0.25 \text{ m}$, $T = 4 \text{ s}$ (cyan), and $\Delta H = 2 \text{ m}$, $T = 4 \text{ s}$, at $\sigma = 0^\circ$ (i.e., nadir) (magenta, dashed).

$\Delta H = 0.25 \text{ m}$ case shows. For both geometries, the curves are nearly identical for halving wave height (red) as for the doubling wave period (green), suggesting that the errors scale with the vertical velocity of the camera motion. Finally, these plots show that the errors are not restricted to oblique views, and in fact are as large for nadir camera orientations (magenta) as for x -propagating crests under similar sea states (left, blue).

It should be noted that the above-mentioned case study is an idealized model for shipboard camera motion. The relationship between sea state and ship motion will depend greatly on the ship, and in most cases the vertical motion of the camera will be less than the

significant wave height. Additionally, the frequency of camera motion will depend on the size of the ship relative to components in the incident wave spectrum. However, even for a camera without any underlying motion, there are uncertainties in this rectification as a result of projecting a wavy sea surface onto a flat plane. Acceptable uncertainties will vary with application; these simulations of $\Lambda(c)$ suggest that camera height variations less than 1 m are necessary for high-quality results.

c. Limitations

Here we summarize the limitations of this horizon stabilization method, many of which have already been

mentioned. The most fundamental requirement is that the horizon be visible in the camera field of view. This is not guaranteed if the platform is subject to large motions or the camera lens has a narrow field of view. The accuracy assessment reveals that the method is especially sensitive to the lighting conditions (see Table 2). Possible complications include glare from direct or reflected sunlight, and lack of contrast as a result of low light, heavy clouds, or rain. Furthermore, the Hough transform detection method requires the horizon to be the most well-defined line in the image, meaning obstacles or other instrumentation in the field of view may impede its detection. In that case, one of the several other horizon detection algorithms (e.g., Bao et al. 2005; Morris et al. 2007; Zhang et al. 2011; Thurrowgood et al. 2009; Moore et al. 2011a; Fefilatyeu et al. 2012) may be more suitable. Finally, if a coastline is used in place of the horizon, at distances less than roughly 10 km, the camera angle calculation will be negatively impacted, as shown Fig. 6.

The horizon only provides information on the camera incidence and roll angles and cannot be used to calculate the heading angle and x , y , and z positions. On large ships and in low sea states, as in this dataset, the assumption of small motions in these other degrees of freedom may be reasonable. The two most problematic of these unknowns are the heading and camera height, because small errors in these variables lead to rectification errors that become large as the distance from the camera increases (as shown in Figs. 7, 8). One option is to use some combination of GPS, IMU, and magnetic compass data to estimate these other motions. Here, the overlapping pitch-and-roll estimates from the video and GPS/IMU may be useful simply to synchronize the data streams. This unfortunately does not solve the issue of noise (at short time scales) and drift (at longer time scales) from which these systems can suffer. Ideally, the camera frames themselves could also be used to estimate these other variables. Toward that end, Moore et al. (2011b) and Thurrowgood et al. (2014) showed that by using two cameras with fish-eye lenses oriented skyward, they could create a “visual compass”—a nearly complete sky panorama that they could compare with an individual view to calculate the relative heading of their UAV. Unfortunately, our limited view of the sky in this dataset makes this approach challenging, but future experiments may take advantage of this method. Similarly, the height of the camera could be estimated visually using a stereo camera system and fitting a plane to the three-dimensional points on the water surface, as in Benetazzo (2006). This assumes that the camera field of view covers a sufficiently large footprint that the fitted plane approximates the mean water level. Note also that

this provides its own estimate of the camera pitch and roll, potentially making the horizon method redundant. However, stereo imaging brings with it a host of significant processing challenges, and thus researchers may still be drawn to this relatively simple method if the conditions are suitable.

5. Conclusions

The horizon line, if it is visible and unobstructed, can be used to calculate the incidence (or pitch) and roll angles of shipboard cameras in the absence of stationary ground control points or external IMU systems. The Hough transform is shown to be a robust and computationally efficient means for horizon detection and a derived quality control metric is effective for isolating spurious horizons. Upon detecting the horizon, the conversion to the incidence and roll of the camera is exact and given in Eqs. (3) and (4). The performance of the horizon tracking technique is largely dependent on the lighting conditions.

For images shown here, the uncertainties in incidence and roll were determined to be on the order of 0.3° and 0.6° , respectively. The effect of horizon distance is negligible in this case, though larger uncertainties are possible if the apparent horizon is within a few kilometers of the camera as a result of poor atmospheric conditions or distant land. The largest source of uncertainty in rectification is due to camera height variations from waves. The percent errors in x and y are similar to the ratio of $\Delta H/H$. These errors can in turn negatively affect estimates of such quantities as $\Lambda(c)$, as a result of the biasing of wave crest speeds.

The implementation of our method in MATLAB is available as an open source toolbox via the MATLAB File Exchange under the name Horizon Stabilization. The package includes all codes for horizon detection, image stabilization, rectification, and error analysis. These codes require either the MATLAB Computer Vision System Toolbox or the Image Processing Toolbox for implementation of the Hough transform.

Acknowledgments. We thank the field engineers of the Applied Physics Laboratory—Joe Talbert, Alex De Klerk, Stephanie Downey, and Mike Ohmart—as well as the crew of the R/V *New Horizon* for all their efforts in support of this work. Thanks also to the National Science Foundation (NSF), the University-National Oceanographic Laboratory System (UNOLS), and the ARCS Foundation. Finally, thanks to Professor Ali Farhadi of the University of Washington for his brilliant introduction to the subject of computer vision.

REFERENCES

- Bao, G.-Q., S.-S. Xiong, and Z.-Z. Zhou, 2005: Vision-based horizon extraction for micro air vehicle flight control. *IEEE Trans. Instrum. Meas.*, **54**, 1067–1072, doi:10.1109/TIM.2005.847234.
- Benetazzo, A., 2006: Measurements of short water waves using stereo matched image sequences. *Coastal Eng.*, **53**, 1013–1032, doi:10.1016/j.coastaleng.2006.06.012.
- Bohren, C. F., and A. B. Fraser, 1986: At what altitude does the horizon cease to be visible? *Amer. J. Phys.*, **54**, 222–227, doi:10.1119/1.14659.
- Callaghan, A. H., and M. White, 2009: Automated processing of sea surface images for the determination of whitecap coverage. *J. Atmos. Oceanic Technol.*, **26**, 383–394, doi:10.1175/2008JTECH0634.1.
- Canny, J., 1986: A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-8**, 679–698, doi:10.1109/TPAMI.1986.4767851.
- Cao, H., and J. Zhang, 2007: Video stabilizing and tracking by horizontal line for maritime cruise ship. *Proceedings of the IEEE International Conference on Control and Automation, 2007*, IEEE, 1202–1206.
- Chickadel, C. C., R. A. Holman, and M. H. Freilich, 2003: An optical technique for the measurement of longshore currents. *J. Geophys. Res.*, **108**, 3364, doi:10.1029/2003JC001774.
- Dickey, T., M. Lewis, and G. Chang, 2006: Optical oceanography: Recent advances and future directions using global remote sensing and in situ observations. *Rev. Geophys.*, **44**, RG1001, doi:10.1029/2003RG000148.
- Duda, R. O., and P. E. Hart, 1972: Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM*, **15**, 11–15, doi:10.1145/361237.361242.
- Fefilat'ev, S., D. Goldgof, M. Shreve, and C. Lembke, 2012: Detection and tracking of ships in open sea with rapidly moving buoy-mounted camera system. *Ocean Eng.*, **54**, 1–12, doi:10.1016/j.oceaneng.2012.06.028.
- French, A. P., 1982: How far away is the horizon? *Amer. J. Phys.*, **50**, 795–799, doi:10.1119/1.13099.
- Gemmrich, J. R., M. L. Banner, and C. Garrett, 2008: Spectrally resolved energy dissipation rate and momentum flux of breaking waves. *J. Phys. Oceanogr.*, **38**, 1296–1312, doi:10.1175/2007JPO3762.1.
- Heikkila, J., and O. Silven, 1997: A four-step camera calibration procedure with implicit image correction. *Proceedings: 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, D. Plummer and I. Torwick, Eds., IEEE, 1106–1112.
- Holland, K. T., R. A. Holman, T. C. Lippmann, J. Stanley, and N. Plant, 1997: Practical use of video imagery in nearshore oceanographic field studies. *IEEE J. Oceanic Eng.*, **22**, 81–92, doi:10.1109/48.557542.
- Holman, R., and M. C. Haller, 2013: Remote sensing of the nearshore. *Annu. Rev. Mar. Sci.*, **5**, 95–113, doi:10.1146/annurev-marine-121211-172408.
- Illingworth, J., and J. Kittler, 1988: A survey of the Hough transform. *Comput. Vision Graphics Image Process.*, **44**, 87–116, doi:10.1016/S0734-189X(88)80033-1.
- IOCCG, 2000: Remote sensing of ocean colour in coastal, and other optically-complex, waters. S. Sathyendranath, Ed., IOCCG Rep. 3, 140 pp.
- Kleiss, J. M., and W. K. Melville, 2010: Observations of wave breaking kinematics in fetch-limited seas. *J. Phys. Oceanogr.*, **40**, 2575–2604, doi:10.1175/2010JPO4383.1.
- , and —, 2011: The analysis of sea surface imagery for whitecap kinematics. *J. Atmos. Oceanic Technol.*, **28**, 219–243, doi:10.1175/2010JTECH0744.1.
- Moore, R. J. D., S. Thurrowgood, D. Bland, D. Soccol, and M. Srinivasan, 2011a: A fast and adaptive method for estimating UAV attitude from the visual horizon. *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, N. M. Amato et al., Eds., IEEE, 4935–4940.
- , —, D. Soccol, D. Bland, and M. V. Srinivasan, 2011b: A method for the visual estimation and control of 3-DOF attitude for UAVs. *Australasian Conference on Robotics and Automation 2011 (ACRA 11)*, Australian Robotics and Automation Association, 267–275.
- Morris, D. D., B. R. Colonna, and F. D. Snyder, 2007: Image-based motion stabilization for maritime surveillance. *Image Processing: Algorithms and Systems V*, J. T. Astola, K. O. Egiazarian, and E. R. Dougherty, Eds., International Society for Optical Engineering (SPIE Proceedings, Vol. 6497), 64970F, doi:10.1117/12.706099.
- Phillips, O. M., 1985: Spectral and statistical properties of the equilibrium range in wind-generated gravity waves. *J. Fluid Mech.*, **156**, 505–531, doi:10.1017/S0022112085002221.
- Romero, L., W. K. Melville, and J. M. Kleiss, 2012: Spectral energy dissipation due to surface-wave breaking. *J. Phys. Oceanogr.*, **42**, 1421–1444, doi:10.1175/JPO-D-11-072.1.
- Schwendeman, M., J. Thomson, and J. R. Gemmrich, 2014: Wave breaking dissipation in a fetch-limited sea. *J. Phys. Oceanogr.*, **44**, 104–127, doi:10.1175/JPO-D-12-0237.1.
- Stockdon, H. F., and R. A. Holman, 2000: Estimation of wave phase speed and nearshore bathymetry from video imagery. *J. Geophys. Res.*, **105**, 22 015–22 033, doi:10.1029/1999JC000124.
- Sutherland, P., and W. K. Melville, 2013: Field measurements and scaling of ocean surface wave-breaking statistics. *Geophys. Res. Lett.*, **40**, 3074–3079, doi:10.1002/grl.50584.
- Szeliski, R., 2010: *Computer Vision: Algorithms and Applications*. Texts in Computer Science, Springer, 812 pp.
- Thomson, J., and A. T. Jessup, 2009: A Fourier-based method for the distribution of breaking crests from video observations. *J. Atmos. Oceanic Technol.*, **26**, 1663–1671, doi:10.1175/2009JTECH0622.1.
- , J. R. Gemmrich, and A. T. Jessup, 2009: Energy dissipation and the spectral distribution of whitecaps. *Geophys. Res. Lett.*, **36**, L11601, doi:10.1029/2009GL038201.
- Thurrowgood, S., D. Soccol, R. Moore, D. Bland, and M. V. Srinivasan, 2009: A vision based system for attitude estimation of UAVs. *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2009)*, IEEE, 5725–5730.
- , R. J. D. Moore, D. Soccol, M. Knight, and M. V. Srinivasan, 2014: A biologically inspired, vision-based guidance system for automatic landing of a fixed-wing aircraft. *J. Field Rob.*, **31**, 699–727, doi:10.1002/rob.21527.
- Weissling, B., S. Ackley, P. Wagner, and H. Xie, 2009: Eiscam—Digital image acquisition and processing for sea ice parameters from ships. *Cold Reg. Sci. Technol.*, **57**, 49–60, doi:10.1016/j.coldregions.2009.01.001.
- Zappa, C. J., M. L. Banner, H. Schultz, J. R. Gemmrich, R. P. Morison, D. A. LeBel, and T. Dickey, 2012: An overview of sea state conditions and air-sea fluxes during RaDyO. *J. Geophys. Res.*, **117**, C00H19, doi:10.1029/2011JC007336.
- Zhang, H., P. Yin, X. Zhang, and X. Shen, 2011: A robust adaptive horizon recognizing algorithm based on projection. *Trans. Inst. Meas. Control*, **33**, 734–751, doi:10.1177/0142331209342201.
- Zhang, Z., 2000: A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22**, 1330–1334, doi:10.1109/34.888718.