# Practical Use of Video Imagery in Nearshore Oceanographic Field Studies

K. Todd Holland, Robert A. Holman, Thomas C. Lippmann, John Stanley, *Associate Member, IEEE,* and Nathaniel Plant

***Abstract*—An approach was developed for using video imagery to quantify, in terms of both spatial and temporal dimensions, a number of naturally occurring (nearshore) physical processes. The complete method is presented, including the derivation of the geometrical relationships relating image and ground coordinates, principles to be considered when working with video imagery and the two-step strategy for calibration of the camera model. The techniques are founded on the principles of photogrammetry, account for difficulties inherent in the use of video signals, and have been adapted to allow for flexibility of use in field studies. Examples from field experiments indicate that this approach is both accurate and applicable under the conditions typically experienced when sampling in coastal regions. Several applications of the camera model are discussed, including the measurement of nearshore fluid processes, sand bar length scales, foreshore topography, and drifter motions. Although we have applied this method to the measurement of nearshore processes and morphologic features, these same techniques are transferable to studies in other geophysical settings.**

***Index Terms*— Camera calibration, nearshore, morphology, video techniques.**

## I. Introduction

THIS PAPER addresses the problem of extracting quantitative information describing physical processes from video images. The remote nature of the acquisition technology involved is advantageous, allowing measurements over a wide range of spatial scales (centimeters to kilometers) and time periods (seconds to years). Video techniques are particularly appealing in the documentation of nearshore oceanographic processes since the subaerial location of the instrument (distant from the ocean surface) alleviates some difficulties associated with *in situ* instrumentation, namely flow disturbance, biofouling, and sensor deterioration under adverse wave conditions. Moreover, postdesign of sampling arrays can occur following an experiment to allow for the investigation of revised hypotheses. The logistics and cost of sampling nearshore processes using video are generally less than traditional solutions involving the deployment of large arrays of instrumentation at a discrete number of positions. In particular, we will show that nearly any nearshore phenomena that can be discerned visually can be quantified using video image processing techniques.

Manuscript received February 15, 1996; revised October 10, 1996.

K. T. Holland is with the Naval Research Lab, Code 7442, Stennis Space Center, MS 39529 USA.

R. A. Holman, J. Stanley, and N. Plant are with the College of Oceanography, Oregon State University, Corvallis, OR 97331 USA.

T. C. Lippmann is with the Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92037 USA.

Publisher Item Identifier S 0364-9059(97)01267-3.

The concept of remotely documenting oceanographic processes is by no means original given the long history of aerial photogrammetry and remote sensing [1], [2]. Similarly, the use of video equipment to measure physical quantities has gained considerable acceptance in the fields of computer vision and robotics [3]. The basic elements of our techniques are identical to those described in the photogrammetry literature. In addition, our understanding of the intricacies involving the use of off-the-shelf closed-circuit-television (CCTV) lenses, charge-coupled device (CCD) cameras, and digital image processing hardware is considered common knowledge in the science of computer vision. The methods described in the following text, however, differ from the traditional photogrammetric analysis of a single photograph because the use of video allows a near continuous sequence of images to be digitally sampled. Additionally, most previous techniques involving video and computer vision have been applied under controlled conditions in the laboratory, whereas our interest is in processes that occur in a field environment. The distinct differences between these objectives have required appreciable modification to prior methods to allow for greater flexibility in the field applications described here.

Our current technique was several years in development and has resulted in a camera calibration model that is far superior to our previous method, which failed to account for camera roll, lens distortion, image centers, and scale factors [4]. This redesigned approach is to apply a mathematical model that describes the geometric orientation of the camera view relative to a reference coordinate system and account for lens distortion and sampling imprecisions resulting from the digitization of an analog video signal. In practice, the coefficients describing the systematic deviations from this video camera model (irrespective of the camera orientation) are determined in the laboratory prior to the field experiment. When combined with the optional constraint of a fixed camera position, knowledge of these coefficients allows the orientation parameters to be easily estimated on location using as few as two surveyed control points, an operationally realistic condition. In this manner, our implementation is simple, efficient, and minimizes the difficulties introduced in field applications.

In the following section, we review the classical methods of photogrammetry. In Section III, the camera model equations governing the relationship between two-dimensional (2-D) image and three-dimensional (3-D) world coordinate systems are derived. Section IV describes a two-step algorithm for determining the intrinsic and extrinsic camera model parame-

ters and outlines the important aspects to be considered when using video imagery. In Section V, we present results from the application of our algorithm in nearshore field experiments and discuss the algorithm's performance. The final section closes with examples of several applications of this technique to nearshore processes given a time sequence of video images.

## II. Background

One necessary requirement for quantifying the information contained in a video image is knowledge of the photogrammetric transformation between 3-D world and 2-D image coordinates. This transformation is a function of two sets of parameters. The first set comprises the geometrical description of the camera position and orientation relative to the reference coordinate system. These variables are known as the extrinsic parameters. The other set consists of certain intrinsic camera parameters reflecting the physical characteristics of the lens, camera, and image acquisition hardware. The determination of the transformation is called camera calibration. Although a variety of camera calibration methods have been proposed, they usually require as inputs both the world coordinates (in 3-D space) and the corresponding image coordinates (on the 2-D image plane) of visually identifiable ground control points (GCP's).

Most existing techniques for camera calibration can be categorized as either explicit or implicit methods. Explicit calibration methods estimate all the camera parameters in terms of dimensional physical units (millimeters, radians, etc.). The majority of the classical calibration techniques used in photogrammetry belong in this category [5]–[7]. Typically, the parameter estimation technique involves an iterative algorithm that minimizes a set of nonlinear or in some cases linearized equations. The advantage of explicit calibration is that a sophisticated camera model incorporating complicated forms of lens distortion and other systematic errors can be used to obtain a high degree of accuracy. The disadvantages are that the use of nonlinear equations can be somewhat cumbersome and that initial approximations for the unknown parameters are necessary. In addition, since the unknown parameters are not independent, the solution can diverge or be incorrect if the initial guess was poor.

In contrast, implicit schemes compute so-called "intermediate" parameters directly based on a closed-form solution to a linear set of equations. Most of these methods are similar to the direct linear transformation (DLT) method suggested by Abdel-Aziz and Karara [8]. No physical meaning is attached to these intermediate parameters; they merely represent combinations of the intrinsic and extrinsic parameters that easily allow the calculation (projection) of known world coordinates to image coordinates. Implicit calibration algorithms are usually fast since iteration and initial estimates are not needed. Unfortunately, nonlinear effects (like radial distortion) cannot be easily accounted for and the presence of errors in the surveyed location of ground control points may lead to a poor final solution unless a large number of GCP's are used. However, implicit methods have proved particularly appealing in controlled situations where GCP's can be positioned to a high degree of accuracy [9].

One possible alternative to both explicit and implicit techniques is a two-step method. Two-step methods involve a closed form solution for some of the calibration parameters (typically the external parameters and the camera focal length) and an iterative algorithm to compute the remaining parameters. An efficient and commonly used two-step calibration procedure is that suggested by Tsai [3]. Tsai's method incorporates only radial distortions, assumes that the position of the optical center is known, and requires *a priori* knowledge of the dimensions of the camera CCD sensor and the computer frame buffer array. The camera calibration technique that we use is also a two-step method, although the order of the steps is reversed to that of Tsai and the control points used for each of the two steps are different. These changes were necessary given the difficulties encountered when working in a field environment and will be explained fully in Section IV.

## III. Mathematical Derivation of the Camera Model

In this section, we develop a distortion-free, pinhole camera model that will serve as the basis for our camera calibration procedure. We also summarize the direct linear transformation equations suggested by Abdel-Aziz and Karara [8] that we will use in the estimation of distortion coefficients based on deviations of observations from this closed-form solution.

Let $(x, y, z)$ represent the 3-D spatial coordinates of a visible point relative to the Cartesian world coordinate system and $(u, v)$ represent the 2-D coordinates of the same point in a digitized image sampled from a video signal. This idealized image plane is at a distance, $f$ (the effective focal length), from the optical center $(x_c, y_c, z_c)$ of the camera and is assumed to be parallel to, but not necessarily dimensionally equivalent with, the CCD sensing array. The image is assumed to be rectangular with a ratio between horizontal and vertical dimensions of 4:3. Since there are 480 active vertical lines in standard video signals, the number of horizontal picture elements (pixels) is typically 640, but can vary depending upon the capabilities of the image processing system. Image coordinates are referenced relative to the right-hand upper corner of the image plane. The image center is given by $(u_0, v_0)$.

Using the parameters defined in Fig. 1, the coordinate transformation between image and world coordinates can be derived in terms of the following collinearity equations under the condition that the camera center, the image point, and the object point all lie on a straight line:

$$\begin{aligned} u - u_0 &= -C_u \left[ \frac{m_{11}(x - x_c) + m_{12}(y - y_c) + m_{13}(z - z_c)}{m_{31}(x - x_c) + m_{32}(y - y_c) + m_{33}(z - z_c)} \right] \\ v - v_0 &= -C_v \left[ \frac{m_{21}(x - x_c) + m_{22}(y - y_c) + m_{23}(z - z_c)}{m_{31}(x - x_c) + m_{32}(y - y_c) + m_{33}(z - z_c)} \right]. \end{aligned} \tag{1}$$

In this equation, $C_u = f/\lambda_u$ and $C_v = f/\lambda_v$ are coefficients relating horizontal and vertical scale factors, $\lambda_u$ and $\lambda_v$, to the effective focal length. The value of $f$ usually differs from the focal length given by the lens manufacturer, although the two distances are theoretically equivalent if the camera is focused at infinity. The individual elements, $m_{ij}$, of the 3

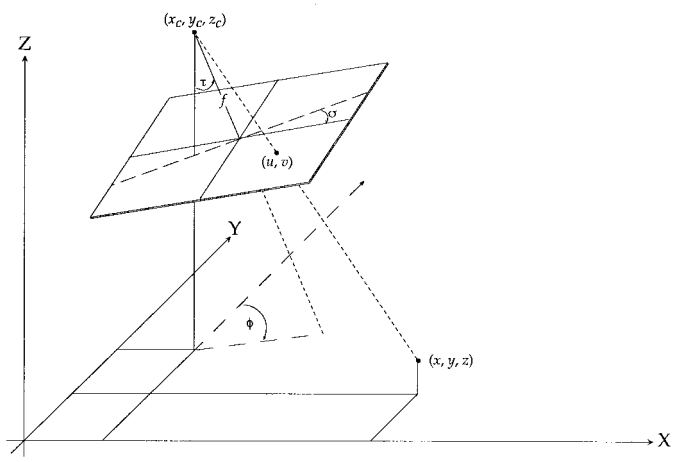



Fig. 1. Collinearity relationship between camera $(x_c, y_c, z_c)$, image $(u, v)$, and world $(x, y, z)$ coordinates and rotation angles $(\phi, \tau, \sigma)$ used in the orientation definition.

$\times$ 3 orthogonal rotation matrix are known as direction cosines and can be derived in terms of three successive rotations about the angles $\phi$ (azimuth), $\tau$ (tilt), and $\sigma$ (roll):

$$\begin{aligned}
m_{11} &= \cos\phi\cos\sigma + \sin\phi\cos\tau\sin\sigma \\
m_{12} &= -\sin\phi\cos\sigma + \cos\phi\cos\tau\sin\sigma \\
m_{13} &= \sin\tau\sin\sigma \\
m_{21} &= -\cos\phi\sin\sigma + \sin\phi\cos\tau\cos\sigma \\
m_{22} &= \sin\phi\sin\sigma + \cos\phi\cos\tau\cos\sigma \\
m_{23} &= \sin\tau\cos\sigma \\
m_{31} &= \sin\phi\sin\tau \\
m_{32} &= \cos\phi\sin\tau \\
m_{33} &= -\cos\tau.
\end{aligned}$$

The angles in the orthonormal rotation matrix may be defined differently by other authors in terms of pan, pitch, and swing; however, the numerical values of the individual elements are the same.

The collinearity equations (1) will serve as our distortion-free camera model and be solved to estimate the parameters relating image and world coordinates. Note that these equations as stated are nonlinear and contain eleven unknowns: the three rotation angles $\phi, \tau$, and $\sigma$; the camera center world coordinates, $x_c, y_c$, and $z_c$; the image center coordinates, $u_0$ and $v_0$; the effective focal length, $f$; and the scale factors, $\lambda_u$ and $\lambda_v$. Typically, surveyed (real-world) GCP's with known image coordinates are needed to solve for the unknowns.

Abdel-Aziz and Karara recognized that the parameters of the collinearity equations can be combined to yield a linear relationship between image and world coordinates:

$$\begin{aligned}
u &= \frac{L_1 x + L_2 y + L_3 z + L_4}{L_9 x + L_{10} y + L_{11} z + 1} \\
v &= \frac{L_5 x + L_6 y + L_7 z + L_8}{L_9 x + L_{10} y + L_{11} z + 1}
\end{aligned} \tag{2}$$

where the coefficients $L_j$, for $j = 1, \cdots, 11$ (known as the DLT coefficients) are listed in Appendix A. The advantage of this rearrangement is that, given the DLT coefficients, the image coordinates of measured world objects can be simply and directly estimated from (2). Note also that the DLT coefficients could, in turn, be estimated given the image and world coordinates of at least six noncoplanar control points. However, given image coordinates from a single camera and the corresponding DLT coefficients, the determination of world coordinates is not possible because the inverse equations

$$\begin{bmatrix} [L_1 - L_9 u] & [L_2 - L_{10} u] & [L_3 - L_{11} u] \\ [L_5 - L_9 v] & [L_6 - L_{10} v] & [L_7 - L_{11} v] \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} [u - L_4] \\ [v - L_8] \end{bmatrix} \tag{3}$$

are underdetermined. If multiple views of the same object space are available (again with predetermined coefficients) or the value of one of the world coordinates is constrained, the solution to (3) may be obtained using the principles of least squares.

## IV. Camera Model Calibration

The formulation of the equations in the above section is considered somewhat standard in the fields of photogrammetry and computer vision. Camera calibration using either the collinearity equations (1) or the direct linear transformation, (2)–(3), is well accepted in situations where control point positioning can be accurately measured. Unfortunately, for many field applications, such control is not always possible. The measurement accuracies often quoted in close-range photogrammetry literature (where the distance separating camera and object is commonly on the order of centimeters) are not applicable in typical nearshore oceanographic study areas, given the magnitude of the regions of interest (often km$^2$). Cameras are often positioned on top of high buildings, towers, or radio antenna sections that are difficult to survey and may change orientation (slightly) over time or abruptly during intense winds. Designing targets that can be accurately surveyed with respect to a reference coordinate system and still be visually identifiable in video images is not trivial. Even after proper design, it is difficult to maintain a stable target in such a hostile environment. Perhaps the biggest limitation for nearshore studies is that only a small portion of the field of view may contain GCP's given that the majority of the view is ocean.

To overcome these limitations, we have developed a new calibration method that solves for the various extrinsic and intrinsic parameters in two steps. Before positioning cameras in the field, we experimentally determine the following intrinsic parameters in the laboratory: the horizontal and vertical image scaling factors, the coefficients necessary to remove lens distortions, and the coordinates of the idealized image plane center. Estimation of these parameters requires a large number of accurately measured control point observations, which precludes their determination in the field. In addition, it is our experience that these parameters are generally invariant for a specific lens-camera-image processing system trio and, therefore, only need to be determined once. We determine the extrinsic parameters and the remaining intrinsic parameter, the effective focal length, during the field calibration step

that is carried out at the onset of each field experiment for each camera view. Although there are more accurate (and considerably more elaborate) methods for calibrating the camera model, we find that this procedure is quite sufficient for our applications, which require only moderate accuracies (i.e., accuracies to within a single picture element).

### A. Laboratory Calibrations

Let us briefly describe the parameters to be determined during the laboratory calibration step. Modern video image acquisition technology requires the digitization of analog video signals into a memory buffer known as a frame grab. Once this "snapshot" is captured, subsequent processing occurs. This processing involves the determination of image scale factors relating the digital frame buffer and the discrete array camera sensor dimensions. These coefficients are integral to the camera model, given our assumption of an image area with a width to height ratio of 4:3, and essentially determine the "squareness" of pixels. The vertical scale factor, $\lambda_v$, is typically set to a value of one since the rows in the computer frame buffer correspond exactly to the scanlines in the video image. The horizontal scale factor, $\lambda_u$, is theoretically related to the ratio of the number of sensor elements in the scan line direction of the CCD array to the number of pixels in a row of the computer image frame buffer, but may vary significantly due to differences in sampling frequencies between the camera and the image acquisition hardware . Note that a 5% error in the horizontal scale factor over the half width of a 640 $\times$ 480 frame buffer translates to an image error of 16 pixels.

Lens distortions can also introduce significant errors for many commonly used CCTV-type lenses. Radial symmetric lens distortion (distortions along radial lines from the center of an image) has been shown to be the largest source of distortion error (compared to other types of distortion) and is typically the only type of distortion accounted for in computer video image processing [3], [12]. The image center coordinates, $u_0$ and $v_0$, are commonly chosen as the center of the image frame buffer; however, the image center for CCD cameras can deviate from the geometric center of the frame buffer by as much as 40 pixels [11], [13]. Interestingly, deviations of this magnitude do not always have an obvious effect on a single camera calibration model because the error introduced by using the wrong $(u_0, v_0)$ is effectively compensated for by translation of the camera position and alteration of the rotation angles. In contrast, choice of the proper center for calibrations involving multiple cameras is critical [3], [11].

Several methods for calculating the horizontal scale factor, the image center, and the distortion coefficients have been proposed [11], [14], [15]. For the most part, we found these suggestions (such as using a spherical target to calculate radial distortion coefficients or electronically measuring camera sampling frequencies to directly determine the horizontal scale factor) enlightening, but somewhat difficult to implement. Instead, we chose to develop a simple method to estimate these parameters based on a severely constrained adaptation of the DLT equations (2). The first constraint we impose is that all of the control point targets lie in a common plane that is parallel to the camera focal plane (our experience suggests that the target plane need be only approximately parallel to achieve meaningful results). We also define a reference world coordinate system such that the elevation value $(z)$ of all the targets in the plane is zero. These constraints allow (2) to be re-written in a simplified form:

$$\begin{aligned} L_1x + L_2y + L_4 - uL_9x - uL_{10}y &= u \\ L_5x + L_6y + L_8 - vL_9x - uL_{10}y &= v. \end{aligned} \tag{4}$$

The final constraint is to further simplify the laboratory setup by requiring the control points to consist of a uniformly spaced array of circular targets (Fig. 2). By postponing the determination of the effective focal length (the remaining intrinsic parameter) until the field calibration step, the distance between the camera and the target plane need not be known, and the units of the world coordinate system are unimportant. This constraint allows definition of the target array at locations $(x^{ij}, y^{ij}) = (\delta i, \delta j)$ where $\delta$ is an equal spacing function and $i$ and $j$ increment from 1 to the desired number of rows $(m)$ and columns $(n)$, respectively. Therefore, the $x$ and $y$ coordinates in (4) represent integer values specifying the column and row of each target in the control point array. The corresponding, distorted image coordinates $(u_d^{ij}, v_d^{ij})$ are determined by calculating the center of mass of each target in the image space in which the gray-scale image has been thresholded to binary values to isolate the near-white targets from the low-intensity background. Having measurements of the image and object space coordinates of the $n \times m$ array of control points allows us to construct a system of equations in terms of the unknown coefficients:

$$\begin{bmatrix} 1 & 1 & 1 & -1u_d^{11} & -1u_d^{11} & 0 & 0 & 0 \\ 2 & 1 & 1 & -2u_d^{12} & -2u_d^{12} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ m & n & 1 & -mu_d^{nm} & -nu_d^{nm} & 0 & 0 & 0 \\ 0 & 0 & 0 & -1v_d^{11} & -1v_d^{11} & 1 & 1 & 1 \\ 0 & 0 & 0 & -2v_d^{12} & -2v_d^{12} & 2 & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & -mv_d^{nm} & -nv_d^{nm} & m & n & 1 \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ L_4 \\ L_9 \\ L_{10} \\ L_5 \\ L_6 \\ L_8 \end{bmatrix} = \begin{bmatrix} u_d^{11} \\ u_d^{12} \\ \vdots \\ u_d^{nm} \\ v_d^{11} \\ v_d^{12} \\ \vdots \\ v_d^{nm} \end{bmatrix}. \tag{5}$$

A least-squares solution for the coefficients $L_1, L_2, L_4, L_5, L_6, L_8, L_9$, and $L_{10}$ can be derived for any target array with more than four elements (we typically sample a 15 $\times$ 20 array). Having determined the calibration coefficients, the predicted (undistorted) image coordinates, $(u_p, v_p)$, can be estimated as

$$\begin{aligned} u_p^{ij} &= \frac{L_1 j + L_2 i + L_4}{L_9 j + L_{10} i + 1} \\ v_p^{ij} &= \frac{L_5 j + L_6 i + L_8}{L_9 j + L_{10} i + 1}. \end{aligned} \tag{6}$$

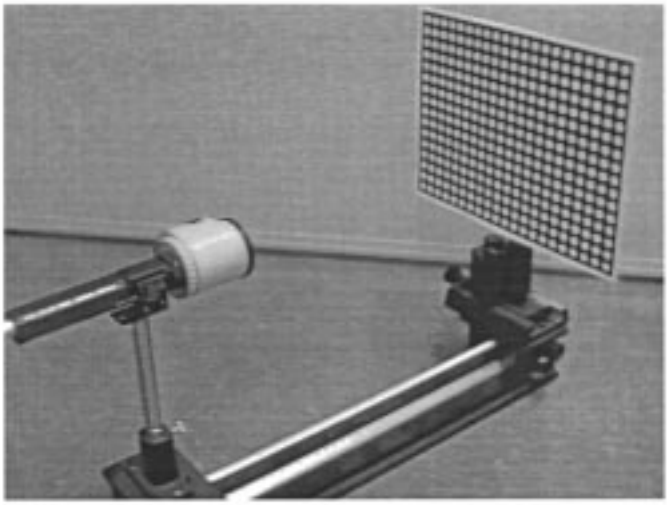


Fig. 2. Laboratory calibration system used in the determination of the intrinsic camera model parameters. Control point testfield consists of regularly spaced white circles on a black background positioned parallel to the camera focal plane. The $z$ axis points toward the camera.

To determine distortion coefficients, $k_1$ and $k_2$, we calculate the deviations between observed and predicted image coordinates corresponding to the control point array given the distortion free model represented by the constrained DLT solution. We choose to model radial distortion in terms of a two-coefficient odd-order polynomial:

$$\Delta r = k_1 r^3 + k_2 r \tag{7}$$

where

$$r = \sqrt{(u_d - u_0)^2 + (v_d - v_0)^2}$$

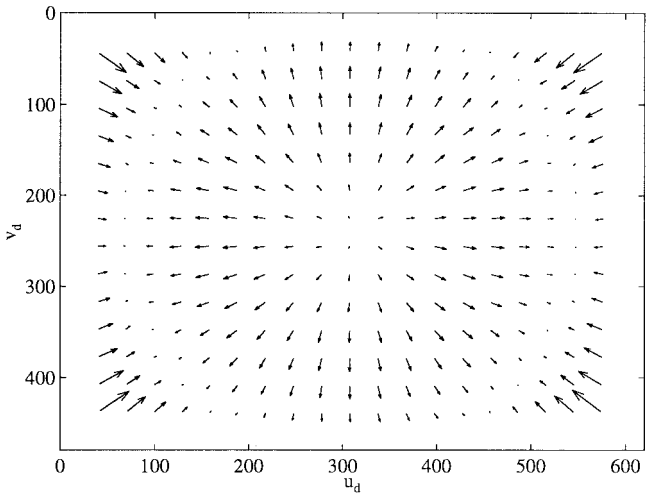
is the distance from the image center at $(u_d, v_d)$ and

$$\Delta r = \sqrt{(u_d - u_0)^2 + (v_d - v_0)^2} - \sqrt{(u_p - u_0)^2 + (v_p - v_0)^2}$$

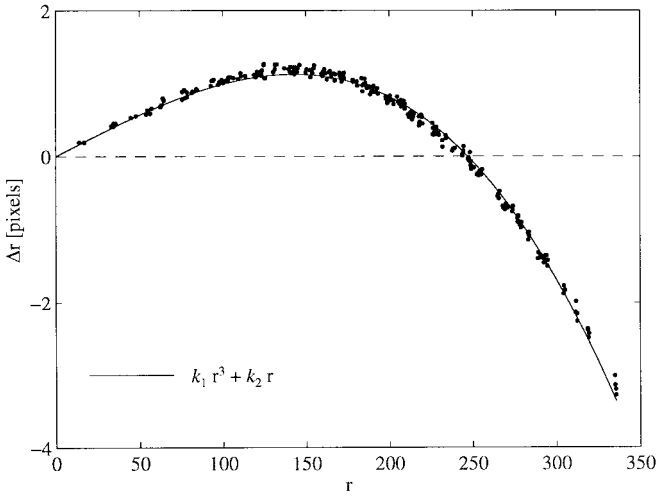
represents the pixel displacement due to distortion. The two distortion coefficients, $k_1$ and $k_2$, are determined as the best fit solution of the polynomial (7) to the $r$ and $\Delta r$ observations. An example solution is shown graphically in Fig. 3 and results for various lenses are listed in Table I.

Although the image center coordinates can be set to the frame buffer center, if greater accuracy is desired, a more elaborate, iterative technique can be used to find the center that minimizes the average error of the distortion polynomial fit (see the local search method proposed by [14]). Willson and Shafer [15] have shown the image center determined using this type of minimization closely approximates the center estimated using the very accurate and robust autocollimated laser approach described by Lenz and Tsai [11], at least for lenses with significant amounts of distortion.

Given the constraint that the original image consisted of a regularly spaced array, $\lambda_u$ is calculated as the ratio of the mean vertical distance (in pixels) between adjacent target points in the same row of targets to the mean horizontal distance




(a)




(b)

Fig. 3. Examples of (a) the radial distortion experienced by an 8-mm lens and (b) the corresponding best fit polynomial used in the distortion correction. The arrows in the top figure represent the difference vector from $(u_d, v_d)$ to $(u_p, v_p)$. The length of the vectors has been increased to emphasize the trends.

TABLE I
EXAMPLE INTRINSIC CAMERA CALIBRATION PARAMETERS DETERMINED DURING THE LABORATORY CALIBRATION STEP FOR THREE DIFFERENT LENSES OF FOCAL LENGTHS GIVEN IN THE TABLE

| | 8 mm | 25 mm | 35 mm |
|---|---|---|---|
| $\lambda_u$ | 1.021 | 1.023 | 1.022 |
| $k_1$ | $-1.94e$-07 | $-4.56e$-08 | $-3.05e$-08 |
| $k_2$ | 0.0119 | 0.0029 | 0.0018 |
| $u_0$ | 314 | 336.5 | 324.5 |
| $v_0$ | 239.5 | 224 | 225 |

The same camera and image processor were used in each case.

between adjacent points in the same column:

$$\lambda_u = \frac{\sum_{i=2,j=1}^{m,n} (v_p^{i,j} - v_p^{i-1,j})}{\sum_{i=1,j=2}^{m,n} (u_p^{i,j} - u_p^{i,j-1})}.$$

For a given camera type and image processing system, the horizontal scale factor is essentially constant for lenses of various focal lengths even though the images from lenses with wider fields of view have larger amounts of distortion (Table I). This consistency suggests that hardware-specific values for $\lambda_u$ can be defined as the average of several observations.

### B. Field Calibrations

Knowledge of the intrinsic parameter values ($\lambda_u,\ \lambda_v,\ u_0,\ v_0,\ k_1$, and $k_2$) reduces the number of unknown parameters in the collinearity equations to seven $(\tau, \phi, \sigma, f, x_c, y_c, z_c)$. Given object coordinates $(x, y, z)$ and undistorted, scale-corrected, image coordinates $(u^*, v^*)$, the unknown parameters are determined using a standard iterative minimization technique that utilizes linearized versions of the collinearity equations. Two equations are constructed for each pair of GCP coordinates by rewriting (1) as

$$F(u^*, \phi, \tau, \sigma, f, x_c, y_c, z_c) = qu^* + of \equiv 0$$
$$G(v^*, \phi, \tau, \sigma, f, x_c, y_c, z_c) = qv^* + pf \equiv 0$$

where

$$\begin{aligned} q &= m_{31}\Delta x + m_{32}\Delta y + m_{33}\Delta z \\ o &= m_{11}\Delta x + m_{12}\Delta y + m_{13}\Delta z \\ p &= m_{21}\Delta x + m_{22}\Delta y + m_{23}\Delta z \\ \Delta x &= x - x_c \\ \Delta y &= y - y_c \\ \Delta z &= z - z_c \\ u^* &= (u - u_0)\lambda_u \\ v^* &= (v - v_0)\lambda_v. \end{aligned}$$

Expanding the functions $F$ and $G$ in terms of a Taylor series in which only the first order terms are retained yields

$$\begin{aligned} 0 = F_0 &+ \frac{\partial F}{\partial u^*}\,du^* + \frac{\partial F}{\partial \phi}\,d\phi + \frac{\partial F}{\partial \tau}\,d\tau + \frac{\partial F}{\partial \sigma}\,d\sigma \\ &+ \frac{\partial F}{\partial f}\,df + \frac{\partial F}{\partial x_c}\,dx_c + \frac{\partial F}{\partial y_c}\,dy_c + \frac{\partial F}{\partial z_c}\,dz_c \\ 0 = G_0 &+ \frac{\partial G}{\partial v^*}\,dv^* + \frac{\partial G}{\partial \phi}\,d\phi + \frac{\partial G}{\partial \tau}\,d\tau + \frac{\partial G}{\partial \sigma}\,d\sigma \\ &+ \frac{\partial G}{\partial f}\,df + \frac{\partial F}{\partial x_c}\,dx_c + \frac{\partial G}{\partial y_c}\,dy_c + \frac{\partial G}{\partial z_c}\,dz_c \end{aligned}$$

where $F_0$ and $G_0$ are initial approximations and the corrections to the unknowns are given as $d\phi, d\tau, d\sigma$, etc. If we note that both partial derivatives $\partial F/\partial u^*$ and $\partial G/\partial v^*$ are equal to $q$, then the finite approximation of the residual errors of the measured image coordinates, $du^* = \Delta u$ and $dv^* = \Delta v$, is

$$\begin{aligned} \Delta u - \frac{F_0}{q} &= b_{11}\,d\phi + b_{12}\,d\tau + b_{13}\,d\sigma + b_{14}\,df \\ &\quad - b_{15}\,dx_c - b_{16}\,dy_c - b_{17}\,dz_c \\ \Delta v - \frac{G_0}{q} &= b_{21}\,d\phi + b_{22}\,d\tau + b_{23}\,d\sigma + b_{24}\,df \\ &\quad - b_{25}\,dx_c - b_{26}\,dy_c - b_{27}\,dz_c \end{aligned} \tag{8}$$

where the $b_{ij}$ coefficients represent partial derivatives of the functions $F$ and $G$ (scaled by $q$) with respect to the unknowns. The coefficient values are listed in Appendix B.

Given initial approximations for all parameters to the functions $F$ and $G$ based on the observations, corrections to the initial values are obtained via a least-squares solution to (8). After the first solution, the computed corrections are added to the initial approximations to obtain revised approximations. This process is repeated (iterated) until the magnitudes of the corrections become negligible and the estimates to the unknowns are determined.

Since each known control point pair yields two equations, all seven of the remaining camera calibration parameters can be determined provided at least four GCP's are visible. By constraining the camera coordinates $(x_c, y_c, z_c)$ to predetermined values, the number of unknowns is reduced to four and the system of equations can be uniquely determined given two GCP's. Having more control points leads to an overdetermined system with redundant information that can be solved in a least-squares sense. Various other combinations of known and unknown field parameters can be used to reduce the number of required GCP's or to increase the level of redundancy. For example, measurement of the horizon line (from which tilt and roll angles can be determined) and knowledge of the camera position allows a model solution for the focal length and azimuth angle with only one surveyed GCP. Of course, any of the field parameters can be easily constrained to predetermined values, regardless of the number of inputs. Once all necessary collinearity parameter values have been calculated, we typically translate these estimates to the equivalent DLT calibration coefficients (Appendix A). This translation simplifies the application of the camera calibration results because the linear equations (2) and (3) can be used.

## V. Technical Performance in Field Studies

In this section, we describe the performance of the calibration technique in a field implementation. Since it is difficult to ground truth the extrinsic parameter values directly, the accuracy of the calibration is measured in two ways. The first method is to determine the discrepancy between manually digitized image points and image points modeled using the measured world coordinates of the corresponding ground control points. In the second approach, the accuracy of the projection of measured image coordinates is compared to the surveyed location of GCP's in the world reference frame. In the later case, example calculations are presented for both single- and multiple-camera scenarios.

### A. Experimental Setup

Results are presented from two distinct field deployments designed for different applications. System #1 consisted of a Sony XC-75 video camera and an 8-mm Fujinon lens mounted atop a 43-m permanent tower at the U.S. Army Corps Field Research Facility in Duck, NC. Video signals from the camera were digitized into 8-bit gray-scale images using a Dipix frame grabber housed in a personal computer. The view of the camera was oriented looking north to document changes in bathymetric features over large, $O(10^2$–$10^3$ m) lengthscales (Fig. 4). Circular targets of 1–2-m diameter were stationed at three positions landward of the dune crest. The

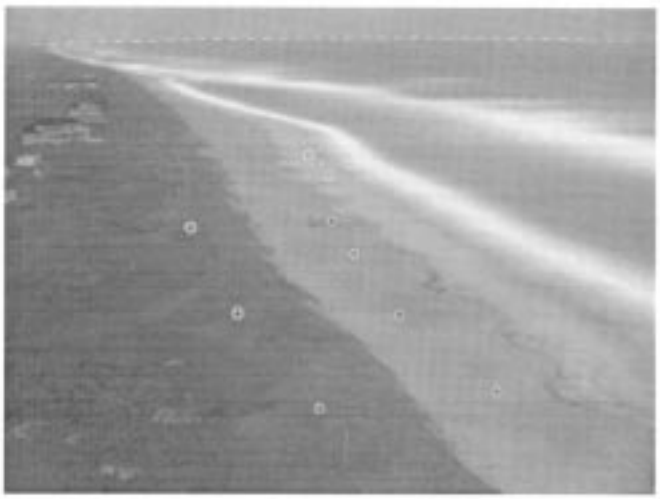


Fig. 4. Example of field implementation of camera calibration approach, October 14, 1994, Duck, NC. The pluses signify the image coordinates of the center of mass of the ground control point targets (the light colored disks of 1–2-m diameter) permanently positioned in the back beach. The circles represent the predicted image coordinates of surveyed, temporary benchmarks on the foreshore. The dashed line represents the predicted horizon line and shows a good correlation with the actual horizon. Given this $640 \times 480$ image as viewed from a camera with a 40° wide angle lens, the vertical pixel footprint of a target at a distance of 225 m is approximately 18 cm.

surveyed centers of these targets served as ground control coordinates for this image. The world coordinates of the camera and several other identifiable targets in the nearshore region were also measured with differential global positioning system (GPS) surveying equipment to centimeter accuracy.

System #2 consisted of three Sony XC-999 cameras with various lenses mounted on temporary towers overlooking a $10 \times 10$ m study region at Gleneden Beach, OR. In this case, the tower elevations were less than 10 m above the beach surface and the distance from the focal point of the cameras to the center of the study region was approximately 20 m. This relatively "close-range" scenario was designed to monitor short term changes in foreshore topography with high resolution. The video imagery was recorded on S-VHS video and subsequently digitized using Imaging Technology image processing hardware. GCP's used in the camera calibration were identified as the location of a dark plastic ball with a diameter of 10 cm mounted on a survey rod. The positions of the rod at various target stations and the camera centers were surveyed using an OMNI total station to within 1 cm.

One important consideration related to the precision of the calibration is the size of GCP targets. The optimal size of a target is dependent upon the size of a pixel back-projected through the camera model to the desired target location. This fundamental limit on image resolution relative to the world reference system is known as the pixel footprint. In order to determine the image position of the control points with high accuracy, the target diameter should be at least an order of magnitude larger than the footprint. For a camera configuration looking in the alongshore, nearly horizontally down the beach, the vertical and horizontal spatial resolutions of a square pixel

TABLE II
ACCURACIES OF FIELD IMPLEMENTATIONS OF CAMERA MODEL

| | $N_{\mathrm{GCP}}$ | $N$ | $E_i$ [pixels] | $E_o$ [cm] | NCE | $D$ [m] |
|---|---|---|---|---|---|---|
| System #1/ Camera #1 | 2 | 7 | 0.63 | 12.1 | 1.60 | 422.0 |
| System #2/ Camera #1 | 4 | 31 | 1.07 | 2.6 | 2.63 | 16.5 |
| System #2/ Camera #2 | 4 | 28 | 1.07 | 2.7 | 2.51 | 18.9 |
| System #2/ Camera #3 | 4 | 27 | 0.60 | 1.4 | 1.59 | 16.8 |

$N_{\mathrm{GCP}}$ represents the number of GCP's used in the camera calibration, $N$ represents the number of GCP's used in the accuracy determination, $E_i$, $E_o$, and NCE are error statistics, and $D$ represents the distance to the target closest to the center of the field of view.

$(\lambda_u = 1)$ in the center of the image are given, respectively, by:

$$\delta_z \approx D \tan\left(\frac{3}{4} * \frac{2\pi\gamma}{360} * \frac{1}{N_u}\right) \quad \text{and} \quad \delta_y \approx \frac{\delta_z}{\tan(\tau)}$$

where $D$ is the horizontal distance of the camera from the ground location, $\gamma$ is the horizontal field of view of the lens in degrees, $N_u$ is the number of horizontal elements in the frame buffer, $\tau$ is the camera tilt, and $\delta_z \ll \delta_y$.

### *B. Calibration Results*

The image coordinates of the visible targets were either manually digitized (to within half a pixel plus operator error) or, in the case of the circular GCP's at Duck, determined to subpixel accuracy using an intensity threshold then center of mass calculation. The camera model parameters were then estimated from the measured object space positions and the corresponding centered, scaled, and undistorted image space coordinates using collinearity equations (8) and the iterative minimization technique. The iteration typically required fewer than five cycles and was completed in less than a second on a SUN Sparc 2 computer. Although the world coordinates of the camera center $(x_c, y_c, z_c)$ could be predicted as part of the minimization, in practice, the most accurate results were obtained by constraining the camera position to the surveyed value.

The accuracies of the model calibrations for each of the four cameras are listed in Table II. Results were calculated in terms of three error estimates. The image space error, $E_i$, represents the average Eucliean distance between the observed (distorted) image coordinates of targets not used in the model solution and the predicted (distorted) coordinates determined using the estimated model parameters and (2). The object space error, $E_0$, defines the average magnitude of closest approach between the measured world coordinate of the targets $(x_i, y_i, z_i)$ and the ray formed by projecting the corresponding image coordinate out through the camera model. The third statistic is the normalized calibration error (NCE) proposed by Weng *et al.* [16] which describes the object space error formed by back projecting the measured image coordinates to the plane formed by the surveyed $z$ component $(\hat{x}_i, \hat{y}_i, z_i)$. This quantity is essentially normalized by the expected resolution

given the separation distance between the camera and target as

$$\mathrm{NCE} = \left[ \frac{1}{N} \sum_{i}^{N} \frac{(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2}{z_i^2 (f_u^{-2} + f_v^{-2})/12} \right]^{1/2}$$

where $N$ represents the number of targets, $f_u = \lambda_u f$, and $f_v = \lambda_v f$. An NCE value $<1$ corresponds to a calibration error that is lower, on average, than the digitization noise of a pixel representation of a target at that separation distance. Values $\approx 1$ indicate a calibration with measured errors comparable to the amount of digitization noise and values $\gg 1$ suggest a poor calibration.

Calibration errors were also determined for the multiple-(three) camera setup. For this scenario, absolute deviations of measured 3-D world coordinates from predicted world coordinates were calculated via stereometric intersection (described in Section VI) using the three sets of camera model parameters and the image coordinates of targets visible in all three views. The average and maximum Eulerian deviations between observations and predictions for the set of 12 targets were 1.7 and 2.8 cm, respectively.

### C. Discussion

In general, the calibration results compared favorably with theoretical expectations. Image space errors for all four cameras were approximately one, and NCE values of order one suggest that the observed errors are consistent with noise inherent in the digitization. The object space errors $(E_o)$ were small relative to the target size and varied depending upon the system setup, with the furthest displaced camera showing the largest error. $E_o$ values for the multiple-camera system was approximately equal to the survey accuracy (1 cm). Finally, the stereometric triangulation error also approximated the survey accuracy and was of smaller scale than most nearshore fluid or bathymetric features of interest.

## VI. Applications

In this paper, we have presented a simple and flexible technique for accurately calibrating video imagery sampled under field conditions. However, the practical utility of any of these models lies not so much in the equation derivations as in the application of the model results.

Our interests usually involve the analysis of 2-D video imagery (as opposed to single frame photogrammetry) to study a wide variety of (potentially 3-D) nearshore phenomena that would have been difficult to sample using conventional methods. Many of these applications have been outlined in previous papers and several are summarized by [17]. In any case, the possible applications of our video methods can generally be divided into two categories: situations where the world-coordinates at the location of the process of interest are known, and situations where the world coordinates are unknown, both of which are described in this section.

### A. World-to-Image Coordinate Transformations

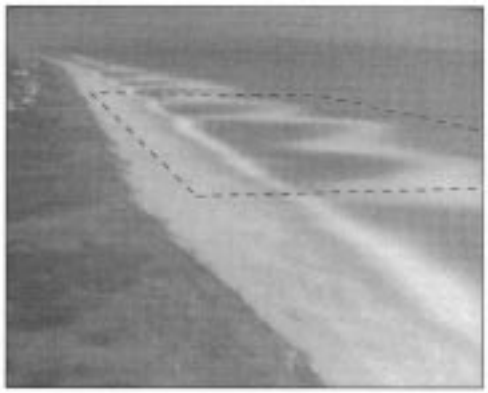
In the first situation, where all of the 3-D world coordinates $(x, y,$ and $z)$ are known, determination of the corresponding image coordinates is straightforward using the relationship given by (2). Typically, the image coordinates corresponding to one or more surveyed locations are determined and the temporal variability in intensity at those pixels is sampled to yield a time series description of the geophysical process of interest. For example, Lippmann and Holman [18] sampled gray-scale intensities at the image location corresponding to a submerged pressure sensor in the breaking wave region. Due to the similarity between the temporal variability in sea–surface elevation and the variability in the brightness of incident wave crests at the same location, the intensity time series served as a proxy for the temporal phase of the co-located pressure measurements and thereby allowed the calculation of wave period, phase speed, and incidence angle at a large number of sensor locations within the breaking-wave region.

In another example, several authors have used 3-D to 2-D transformations to measure the sea surface elevation time series at the ocean edge (runup) [19], [20]. To do so, beach surface locations $(x, y, z)$ along a continuous transect visible in the image are mapped to $(u, v)$ coordinates. The intensity changes along the corresponding image transect are used to define the position of the runup edge over time, and lastly, the pixel coordinates defining the edge position are transformed back to vertical elevations using the same mapping between the image and the measured profile. Holland *et al.* [21] compared nearbed runup measurements sampled using video instrumentation with observations measured at higher elevations using a resistance-wire runup meter. The relationships between the two types of measurements were sensible and resulted in a detailed depiction of runup kinematics.

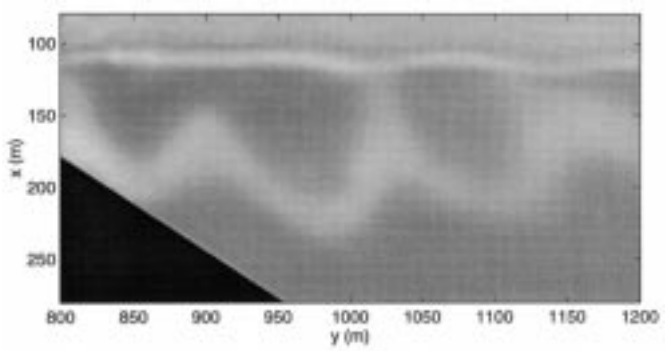
### B. Image-to-World Coordinate Transformations

A more challenging application, at least from an operational point of view, is transformation of image coordinates to world coordinates. For these types of applications, the 3-D coordinates of a 2-D pixel location cannot be estimated because the system of equations is underdetermined (two equations with three unknowns). Two alternatives exist: either constrain at least one spatial aspect of the world coordinate system or sample additional images of the same object space as viewed from different cameras.

*1) Spatial Constraints:* Practical constraints on world coordinates when using video imagery may not be intuitively obvious, but are, in many cases, commonly available. One possibility is that the imagery represents phenomena where one spatial variable can be specified directly, thereby adding a third equation to (3) and yielding a unique solution for the image-to-ground transformation. For example, Lippmann and Holman [4] constrain $z$ to a fixed value (mean sea level as given by *in situ* instrumentation), when making video time exposures (essentially intensity images resulting from averaging multiple video frames over time) of bar morphology. Given a known elevation coordinate allows the mapping of all of the pixel intensities within a region of the oblique time-exposure image to a rectified, planar image in world coordinates. For an alongshore-oriented camera, the horizontal positioning error (absolute difference from a benchmark location) introduced in the rectification by using an



(a)




(b)

Fig. 5. Ten-minute time exposure images of wave-breaking patterns at Duck, NC. (a) The oblique view as seen from the camera. The intensity patterns within the dashed region were rectified using the elevation constraint to construct (b) the horizontal view. The cuspate patterns indicated by the lighter regions represent submerged sand-bar morphology with a rhythmic alongshore length scale of approximately 100 m. Since the original image was sampled at discrete pixel locations, some interpolation of image intensities is required to produce a smooth rectified image.

incorrect elevation scales in terms of the elevation error times the tangent of the camera tilt. However, the relative spacing error (difference between projected measured length scales) is largely insensitive to the error in the elevation estimate and is commonly smaller than the pixel footprint. An example of determining morphological length scales from rectified nearshore imagery is given in Fig. 5. Rectified time-exposure images have shown great utility for long-term (months to decades) studies of morphologic variability given the logistical simplicity and large spatial areas sampled [17].

A similar application, proposed by Holman *et al.* [22] for the video estimation of subaerial beach profiles, is to limit samples to a vertical plane. This plane intersects the ground surface as a visible line which demarcates the world coordinates of the beach profile (unknown). Such a visible line may be artificially fabricated (with a sheet laser for example) or may exist naturally as a recurrent linear pattern (e.g., the shadow cast from a vertical post). If either spatial variable ($x$ or $y$) can be defined as either a constant or in terms of the other (which is equivalent to a rotation of the world coordinate system), the pixel locations corresponding to the line in the image can be used as in the horizontal plane case described above to solve for the world coordinates of the topographic profile.

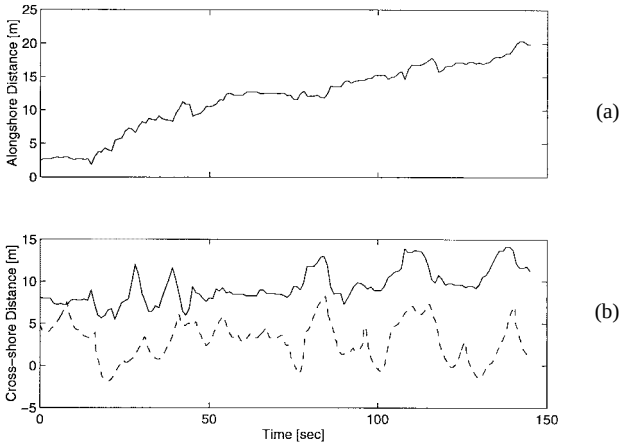



Fig. 6. (a) Alongshore and (b) cross-shore drifter motions as detected using video instrumentation. A slow propagation to the south at a speed of approximately 0.1 m/s was observed. Cross-shore motions were well correlated with large swash excursions (bottom dashed), which were also measured using a video technique (see [19], [21]).

Another alternative for determining world coordinates using imagery from one camera is to constrain the set of world coordinates to a prescribed surface. An example of this type of application involving unknown world coordinates is tracking near-bottom drifters (in this case military-type, anti-invasion mines) in the inner surf and swash zones. This situation is similar to the morphology measurement techniques mentioned previously, except that the constrained spatial dimension need not represent a vertical or horizontal plane (or a plane at all for that matter). Instead, the ground coordinates corresponding to measured pixel locations are derived by defining the ground coordinate surface mathematically. To demonstrate this capability, measured bathymetry from Camp Pendleton, CA, was approximated by a best fit planar surface, $ax + by + cz = d$. Manually determined image coordinates corresponding to the drifter position were transformed to world coordinates using the camera calibration coefficients and the following system of equations derived from (3):

$$\underbrace{\begin{bmatrix} [L_1 - L_9 u] & [L_2 - L_{10} u] & [L_3 - L_{11} u] \\ [L_5 - L_9 v] & [L_6 - L_{10} v] & [L_7 - L_{11} v] \end{bmatrix}}_{\begin{matrix} a & b & c \end{matrix}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} [u - L_4] \\ [v - L_8] \end{bmatrix}_{d}.$$

With one camera, this system of equations is uniquely determined and was solved to yield temporal estimates of drifter position on the nearshore surface. The mines were found to be extremely mobile with typical cross-shore excursions (linked to large swash motions) on the order of 5 m every few seconds (Fig. 6).

*2) Stereometric Methods:* If a spatial constraint is not feasible, the presence of two or more cameras overlooking the same object space will also allow the estimation of ground coordinates from known image coordinates. This determination, termed stereometric intersection, can be accomplished by calibrating the cameras individually and then performing a least-squares solution for the world coordinates, given the

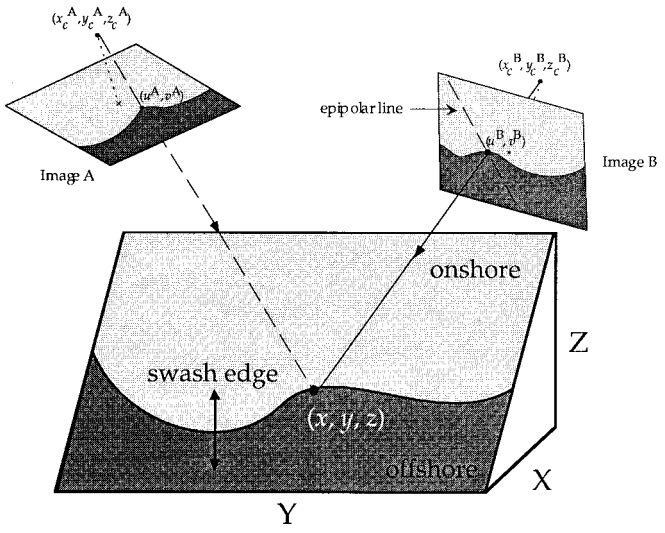



Fig. 7. An illustration of the stereometric method of measuring foreshore topography using multiple cameras. The pixel coordinates in image A define the corresponding coordinates in image B as the intersection of the epipolar line (dashed) and the swash edge. This pair of matching coordinates is used with the camera calibration parameters to determine the world coordinates $(x, y, z)$ of the foreshore surface at the selected point. To map the entire region, the process is repeated for each pixel defined in the swash edge of both images and for successive edges as the swash moves landward and seaward (double-sided arrow).

image coordinate inputs:

$$\begin{bmatrix} [L_1^A - L_9^A u^A] & [L_2^A - L_{10}^A u^A] & [L_3^A - L_{11}^A u^A] \\ [L_5^A - L_9^A v^A] & [L_6^A - L_{10}^A v^A] & [L_7^A - L_{11}^A v^A] \\ [L_1^B - L_9^B u^B] & [L_2^B - L_{10}^B u^B] & [L_3^B - L_{11}^B u^B] \\ [L_5^B - L_9^B v^B] & [L_6^B - L_{10}^B v^B] & [L_7^B - L_{11}^B v^B] \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} [u^A - L_4^A] \\ [v^A - L_8^A] \\ [u^B - L_4^B] \\ [v^B - L_8^B] \\ \vdots \end{bmatrix} \quad (9)$$

where $(L^A, u^A, v^A)$ and $(L^B, u^B, v^B)$ represent the DLT coefficients and image coordinates for cameras A and B, respectively [23]. The ellipses signify the possibility of additional cameras to yield a further overdetermined system.

The importance of stereometric intersection is that ground coordinates of any identifiable object can be determined given pixel positions. Yet the main difficulty in employing stereo methods in the analysis of imagery from field experiments is the identification of corresponding features in multiple views. With respect to nearshore oceanography, easily identifiable point objects (the tip of a buried shell, for example) are relatively uncommon. However, Holland and Holman [24] recognized that the intensity contrast between the foreshore surface and the edge of incoming swash bores is particularly sharp and could be used as a feature to map foreshore topography using stereometric intersection (9). Since this edge is restricted to the sand level, measurement of the world coordinates of the moving edge over the first half of a swash period effectively defines the foreshore surface (Fig. 7). Further details on the implementation and accuracy of this method are given in [24].

## VII. Conclusion

We have presented an approach for the quantification of physical processes using video imagery from nearshore oceanographic field studies. The camera model and techniques described were derived from often elaborate methods used in the fields of photogrammetry and computer vision, but were simplified and adapted to allow for greater flexibility in our specific field applications. Most of this flexibility results from solving for the various camera parameters in two, time-separated, steps. In the first step, we experimentally determine certain intrinsic parameters in the laboratory. These intrinsic parameters are those that typically require a large number of accurately measured control point observations and are generally invariant for a specific lens, camera, and image processing system. After positioning the cameras in the field, the extrinsic parameters (geometric variables describing the camera orientation) and the remaining intrinsic parameter (the effective focal length) are determined. The separation into lab and field-specific steps allows the transformation variables to be determined accurately over a range of field applications with a minimum of difficulty. This computation requires as few as two ground coordinates and the camera position to be surveyed. The solution method is receptive to various other combinations of known and unknown parameters (for example, knowledge of the focal length, but not the camera position), as long as the minimum number of GCP measurements are made.

Results from field tests indicate that this approach is both accurate and feasible under the conditions typically encountered during field studies. However, the ultimate utility of the approach lies not in the model derivation, but in the application of the model to the study of physical processes. This application typically involves either the transformation of 3-D world coordinates to 2-D image coordinates or vice versa. The latter transformation (2-D to 3-D) requires either knowledge of one spatial variable *a priori* (such as the vertical coordinate of a horizontal plane) or stereometric measurements. We describe several applications illustrating some of the many possibilities for using our approach, specifically those that have proved fruitful in the study of nearshore fluid processes and bathymetric features. We feel certain that this same approach can be applied toward the study and monitoring of oceanographic processes in other geophysical settings.

## Appendix A

$$\begin{aligned} L &= -(x_c m_{31} + y_c m_{32} + z_c m_{33}) \\ L_1 &= (u_0 m_{31} + f m_{11})/(\lambda_u L) \\ L_2 &= (u_0 m_{32} + f m_{12})/(\lambda_u L) \\ L_3 &= (u_0 m_{33} + f m_{13})/(\lambda_u L) \\ L_4 &= -(L_1 x_c + L_2 y_c + L_3 z_c) \\ L_5 &= m_{31}/L \\ L_6 &= m_{32}/L \\ L_7 &= m_{33}/L \\ L_8 &= (v_0 m_{31} + f m_{21})/(\lambda_v L) \end{aligned}$$

$$L_9 = (v_0 m_{32} + f m_{22})/(\lambda_v L)$$
$$L_{10} = (v_0 m_{33} + f m_{23})/(\lambda_v L)$$
$$L_{11} = -(L_8 x_c + L_9 y_c + L_{10} z_c).$$

## APPENDIX B

Coefficients describing the partial derivatives of $F$ and $G$ scaled by $q$ used in (8):

$$b_{11} = \frac{\partial F}{\partial \phi} = \frac{u^*}{q}(\Delta x m_{32} - \Delta y m_{31}) + \frac{f}{q}(\Delta x m_{12} - \Delta y m_{11})$$

$$b_{12} = \frac{\partial F}{\partial \tau} = \frac{u^*}{q}(dx \cos\tau \sin\phi + dy \cos\phi \cos\tau + dz \sin\tau) + \frac{f}{q}(-\Delta x \sin\phi \sin\tau \sin\sigma - \Delta y \cos\phi \sin\tau \sin\sigma + \Delta z \cos\tau \sin\sigma)$$

$$b_{13} = \frac{\partial F}{\partial \sigma} = \frac{f}{q}(\Delta x m_{21} + \Delta y m_{22} + \Delta z m_{23})$$

$$b_{14} = \frac{\partial F}{\partial f} = o/q$$

$$b_{15} = \frac{\partial F}{\partial x_c} = \frac{u^*}{q} m_{31} + \frac{f}{q} m_{11}$$

$$b_{16} = \frac{\partial F}{\partial y_c} = \frac{u^*}{q} m_{32} + \frac{f}{q} m_{12}$$

$$b_{17} = \frac{\partial F}{\partial z_c} = \frac{u^*}{q} m_{33} + \frac{f}{q} m_{13}$$

$$b_{21} = \frac{\partial G}{\partial \phi} = \frac{v^*}{q}(\Delta x m_{32} - \Delta y m_{31}) + \frac{f}{q}(\Delta x m_{22} - \Delta y m_{21})$$

$$b_{22} = \frac{\partial G}{\partial \tau} = \frac{v^*}{q}(dx \cos\tau \sin\phi + dy \cos\phi \cos\tau + dz \sin\tau) + \frac{f}{q}(-\Delta x \sin\phi \sin\tau \cos\sigma - \Delta y \cos\phi \sin\tau \cos\sigma + \Delta z \cos\tau \cos\sigma)$$

$$b_{23} = \frac{\partial G}{\partial \sigma} = \frac{-f}{q}(\Delta x m_{11} + \Delta y m_{12} + \Delta z m_{13})$$

$$b_{24} = \frac{\partial G}{\partial f} = p/q$$

$$b_{25} = \frac{\partial G}{\partial x_c} = \frac{v^*}{q} m_{31} + \frac{f}{q} m_{21}$$

$$b_{26} = \frac{\partial F}{\partial y_c} = \frac{v^*}{q} m_{32} + \frac{f}{q} m_{22}$$

$$b_{27} = \frac{\partial F}{\partial z_c} = \frac{v^*}{q} m_{33} + \frac{f}{q} m_{23}.$$

## REFERENCES


[1] J. R. Apel, "Ocean science from space," *EOS Trans. Amer. Geophys. Union*, vol. 57, pp. 612–624, 1976.

[2] P. G. McCurdy, *Manual of Photogrammetry*. New York: Pitman, 1944, pp. 1–841.

[3] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE J. Robot. Automat.*, vol. RA-3, pp. 323–344, 1987.

[4] T. C. Lippmann and R. A. Holman, "Quantification of sand bar morphology: A video technique based on wave dissipation," *J. Geophys. Res.*, vol. 94, pp. 995–1011, 1989.

[5] K. W. Wong, "Mathematical formulation and digital analysis in close-range photogrammetry," *Photogramm. Eng. Remote Sens.*, vol. 41, pp. 1355–1373, 1975.

[6] W. Faig, "Calibration of close-range photogrammetric systems: Mathematical formulation," *Photogramm. Eng. Remote Sens.*, vol. 41, pp. 1479–1486, 1975.

[7] D. C. Brown, "Close-range camera calibration," *J. Amer. Soc. Photogrammetry*, vol. 37, pp. 855–866, 1971.

[8] Y. I. Abdel-Aziz and H. M. Karara, "Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry," in *Proc. ASP/UI Symp. Close-Range Photogrammetry*, Urbana, IL, 1971, pp. 1–18.

[9] H. Hatze, "High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified DLT approach," *J. Biomechanics*, vol. 21, pp. 533–538, 1988.

[10] S. K. Ghosh, *Analytical Photogrammetry*, 2nd ed. New York: Pergamon, 1988, pp. 1–308.

[11] R. K. Lenz and R. Y. Tsai, "Techniques for calibration of the scale factor and image center for high accuracy 3-d machine vision metrology," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 713–720, 1988.

[12] H. M. Karara and Y. I. Abdel-Aziz, "Accuracy aspects of nonmetric imageries," *Photogramm. Eng. Remote Sens.*, vol. 40, pp. 1107–1117, 1974.

[13] A. H. Beyer, "Accurate calibration of CCD cameras," in *IEEE Computer Soc. Conf. Computer Vision and Pattern Recognition*, Champaign, IL, 1992, pp. 96–101.

[14] M. A. Penna, "Camera calibration: A quick and easy way to determine the scale factor," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 1240–1245, 1991.

[15] R. G. Willson and S. A. Shafer, "What is the center of the image," *J. Opt. Soc. Amer.*, vol. 11, pp. 2946–2955, 1994.

[16] J. Weng, "Camera calibration with distortion models and accuracy evalution," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 965–980, 1992.

[17] R. A. Holman, A. H. Sallenger, Jr., T. C. Lippmann, and J. W. Haines, "The application of video image processing to the study of nearshore processes," *Oceanography*, vol. 6, pp. 78–85, 1993.

[18] T. C. Lippmann and R. A. Holman, "Phase speed and angle of breaking waves measured with video techniques," in *Coastal Sediments, '91*, N. Kraus, Ed. New York: ASCE, 1991, pp. 542–556.

[19] R. A. Holman and R. T. Guza, "Measuring run-up on a natural beach," *Coastal Eng.*, vol. 8, pp. 129–140, 1984.

[20] T. Aagaard and J. Holm, "Digitization of wave run-up using video records," *J. Coastal Res.*, vol. 5, pp. 547–551, 1989.

[21] K. T. Holland, B. Raubenheimer, R. T. Guza, and R. A. Holman, "Runup kinematics on a natural beach," *J. Geophys. Res.*, vol. 100, pp. 4985–4993, 1995.

[22] R. A. Holman, T. C. Lippmann, P. V. O'Neill, and K. Hathaway, "Video estimation of subaerial beach profiles," *Marine Geology*, vol. 97, pp. 225–231, 1991.

[23] J. S. Walton, "Close-range cine-photogrammetry: A generalized technique for quantifying gross human motion," Pennsylvania State University, State College, PA, 1981, pp. 1–626.

[24] K. T. Holland and R. A. Holman, "Video estimation of foreshore topography using trinocular stereo," *J. Coastal Res.*, vol. 13, pp. 81–87, 1997.



**K. Todd Holland** received the B.S. degree in geology from Duke University, Durham, NC, in 1986 and the M.S. and Ph.D. degrees in marine geology from Oregon State University, Corvallis, OR, in 1995. His dissertation is entitled "Foreshore Dynamics: Swash Motions and Topographic Interactions on Natural Beaches."

He is presently a Project Scientist at the Naval Research Laboratory (NRL) studying surfzone mine behavior and nearshore morphologic change using video techniques. Prior to his joining NRL, he also worked as a Postdoctoral Researcher at Stennis Space Center, MS, and for five years as an Oceanographer with the U.S. Geological Survey in Reston, VA.

Dr. Holland is a member of the American Geophysical Union.

**Robert A. Holman** received the Ph.D. degree from Dalhousie University, Halifax, Canada, in 1979.

He is currently a Professor in the College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, studying the world's beaches and nearshore processes. The range of scales of interest on Oregon's local beaches slowly forced him to adopt video remote sensing techniques to studies of the nearshore, in contrast to the traditional *in situ* sensor approach. With his students, he has continued to develop these techniques to a quantitative science.

**Thomas C. Lippmann** received the B.S. degree from Linfield College, McMinnville, OR, and the M.S. and Ph.D. degrees from Oregon State University, Corvallis.

He is presently an Assistant Research Oceanographer at the Scripps Institution of Oceanography, San Diego, CA, studying a variety of topics involving nearshore video image processing. He worked as a National Research Council Post-Doctoral Fellow at the Naval Postgraduate School from 1992 to 1995.

Dr. Lippmann is a member of the American Geophysical Union.

**John Stanley** (A'95) received the B.S. degree in biochemistry and the M.S. degree in analytial chemistry from Michigan State University, Ann Arbor, in 1978 and 1982, respectively. His research involved the measurement of atomic excited state lifetimes in atmospheric flames.

He worked as a Senior Software Engineer at New Methods Research, Inc., East Syracuse, NY, for five years. He is currently a Faculty Research Assistant in the Coastal Imaging Lab, College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, where he manages computing and video resources and designs and implements field data collection systems.

Mr. Stanley is a member of ACM, SPIE, the American Chemical Society, and the Society for Applied Spectroscopy.

**Nathaniel Plant** was born in San Francisco, CA, in 1966. He received the B.S. and M.S. degrees in earth science from the University of California at Santa Cruz (UCSC) in 1989 and 1990, respectively, studying coastal erosion.

After graduating from UCSC, he worked part time for the U.S. Geological Survey (USGS), studying regional coastal geology and erosional processes. In 1992, he left the USGS to start a Ph.D. program with Rob Holman at Oregon State University, Corvallis, where he presently remains.

Mr. Plant is a member of the American Geophysical Union.