Estimating Coral Reef Slope or Camera Pitch from Video

Meryl Regine Algodon*, Aaron Hilomen , Maricor Soriano
National Institute of Physics
Instrumentation Physics Laboratory
University of the Philippines, Diliman, Quezon City
malgodon@nip.upd.edu.ph*

***Abstract*— Area estimation of coral colonies from images can be improved by eliminating the distorting effect of camera rotation angle or reef slope. We propose a method that estimates the angle between reef and camera plane using the tracks formed by the movement of image features in the video. Estimates obtained from experimental setup show the effectiveness of the technique, with acceptable error for small angles. The technique can be used on benthos monitoring setups to provide good area estimates without any additional instrumentation.**

***Keywords*— Coral Reef Slope, Camera Rotation, Coral Assessment**

I. Introduction

Coral reefs are among the most important natural resources in the world. They are important as they sustain highly biodiverse ecosystems which are relevant to the fisheries and tourism industries. The high biodiversity of coral reefs also means that they harbor a great potential for new medicines, materials and other scientific developments. Coral reefs also play a part in coastal protection by providing a buffer from wave action and coastal erosion.

As with any important natural resource, there is a need to be able to quantify and manage these resources. One of the means of quantifying coral reefs is in terms of coral cover or area. The latest assessment methods that have been developed allow for the rapid assessment of coral reefs using underwater cameras. The underwater camera is fixed facing downwards on a towed underwater platform, such as Teardrop [1]. The platform is towed behind a boat and the camera records video. Using image processing software, such as Kiko and Stitch [1], it is possible to stitch the video into an image showing a plan view of the coral. These methods allow the image capture of kilometers worth of coral reefs in a single day.

Area estimation of coral reefs is possible using these images but there are many sources of error. One of the sources of error would be towed platforms are not steady. When towed through the water the platform exhibits pitch and roll which is translated to camera rotation. This complicates area estimation as the area being measured is distorted in the image since it is not perpendicular to the camera. Even with an ideal platform that is both stable and static, error also arises as the sea floor more often than not is sloped or undulating, this causes a difference in the apparent area seen in the image and the actual area being measured.

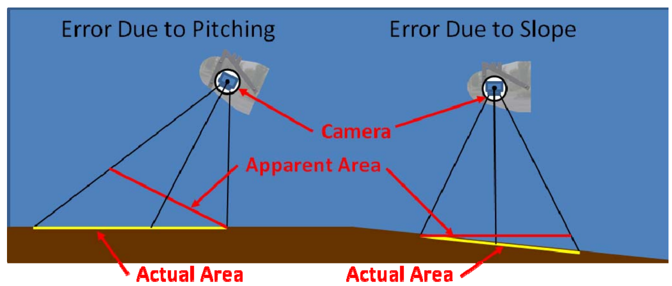

Fig. 1. Errors Caused by Camera Rotation

To correct this error caused by camera rotation, either by camera platform movement or sea floor slope, it is possible through the use of video to be able to measure the apparent camera rotation with respect to the sea floor.

Unlike other camera calibration techniques that use sophisticated pan-tilt camera setups [2] or calibration objects for vanishing point extraction [3], we do not require any other instrument except an echosounder as an average depth gauge. Yuan et al. [4] proposed a camera motion estimation based on the camera's forward translation. In our technique, the assumed camera movement is an upward translation relative to a static scene.

II. THEORETICAL DEVELOPMENT

Teardrop [1] is an underwater tool for coral reef assessment that records a video of the benthos as it is towed by a boat. While it is possible to obtain coral coverage from the image and camera altitude, camera rotation and change in coral slope may lead to erroneous estimates. Calculation of the camera rotation angle from the image could help correct this error. Our goal is to develop an approach for estimating the orientation of the camera relative to the slope of the benthos using the apparent movement of coral features across the image frame.

Consider a pinhole camera model where the world coordinate system z axis coincides with the optical axis of camera coordinate system. The two-dimensional projection of the world coordinates system is the image plane represented by the xy-plane. The camera focal length f is the distance of the view point O to the origin of the camera coordinate system. The rotation matrix R of the camera coordinate system relative to world coordinate system is given by:

$$\boldsymbol{R} = \begin{bmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{32} & r_{33} \end{bmatrix} \tag{1}$$

The camera, submerged in water, has three rotational degrees of freedom: one along the optical axis (yaw), and one for each of the axes in the image plane (pitch and roll).

Image features flow parallel but in opposite direction to camera movement along the axis. For example, radial outward movement of image features from the focus of expansion implies camera movement along the optical axis towards the object, and inward movement of features imply camera movement away from the object. Image features translate downward along the image frame if the camera is moving upward the image axis, and vice versa. This transformation is directed as the transpose of camera rotation. Thus, we may express image feature transformation from one point to another, $\boldsymbol{x}$ to $\boldsymbol{x}'$ as:

$$\boldsymbol{x}' = \boldsymbol{R}^T \boldsymbol{x}. \tag{2}$$

where

$$\boldsymbol{x}' = \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} \text{ and } \boldsymbol{x} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \tag{3}$$

Multiple feature points could also be used to form a line projected in the image plane [5], written as:

$$Ax + By + C = 0. \tag{4}$$

Similar to the transformation due to point translation, the line after rotation can be written as:

$$A'x + B'y + C' = 0, \tag{5}$$

where line coefficients are given as:

$$A' : B' : C' = r_{11}A + r_{21}B + {r_{31}C}/{f} : r_{12}A + r_{22}B + {r_{32}C}/{f} : f(r_{13}A + r_{23}B) + r_{33}C. \tag{6}$$

Line transformation due to rotation can be expressed in matrix form as:

$$\begin{bmatrix} A' \\ B' \\ C'/f \end{bmatrix} = \boldsymbol{R}^T \begin{bmatrix} A \\ B \\ C/f \end{bmatrix}. \tag{7}$$

Using equation (7), it is possible to calculate the camera rotation angle along a single axis when the camera rotation matrix and the initial equation of the line are known.

Suppose that the camera is initially aligned such that the image plane is parallel to the object plane. If, during camera movement, the image plane and object plane remains parallel to each other, any feature from the image tracked as a point will trace a straight line oriented opposite to the direction of movement of the camera.

Suppose further that the movement of the camera is along the y-axis, where multiple features tracked from the image trace parallel vertical lines $y_i = k_i$. By rotating the camera along the x-axis, the slope of the lines will change such that these lines will converge into a vanishing point and the transformed line is given by the expression $y' = mx' + b$ where m is the slope of the line and b is the y-intercept. The parallel vertical lines can also be written in matrix form as

$$\boldsymbol{V} = \begin{bmatrix} 1 \\ 0 \\ k/f \end{bmatrix} \tag{10}$$

Using rotation matrix for rotation along x-axis $\boldsymbol{R_x}$ given as

$$\boldsymbol{R_x}(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\,\alpha & \sin\alpha \\ 0 & -sin\,\alpha & cos\,\alpha \end{bmatrix}, \tag{11}$$

it is possible to calculate the rotation angle α along the x-axis by substituting the equations (10) and (11) to equation (7),

$$R_x^T(\alpha) \begin{bmatrix} 1 \\ 0 \\ k/f \end{bmatrix} = \begin{bmatrix} -m \\ 1 \\ -b/f \end{bmatrix}. \tag{12}$$

Where camera pitch angle is derived as

$$\alpha = tan^{-1}\frac{f}{b}. \tag{13}$$

III. ALGORITHM DEVELOPMENT

To estimate camera pitch angle from actual coral videos, lines projected on the image plane must be extracted from the image features as it moves across the image frame. In our algorithm, the image feature is tracked using image correlation. The image frame is cropped vertically into seven overlapping image sections each with sizes of 270 pixels by 1080 pixels (270x1080). For a series of consecutive image sections, the initial image frame is correlated to the succeeding image frames until the feature reach the edge of the image frame. Correlation is the expression measuring the similarity between two images $f(x,y)$ and $g(x,y)$ and is given by the equation

$$f \odot g = IFT\{F^* G\} \tag{14}$$

where F and G are the Fourier Transforms of f and g, respectively, '*' is conjugation, and IFT is the inverse Fourier Transform. From the output of two correlated images, the maximum correlation (denoted by the bright peaks) represents the region on the image with highest probability of similarity between the images. In some instances, due to signal noise, there is a possibility of multiple white peaks found on the resulting image. Hence, a mask is applied to the succeeding image where the region of interest is limited to only a circular region of 100 pixel diameter where the center is located to the position of the maximum correlation from previous correlated image frame as shown in Figure 2. Masking the rest of the image prevents the algorithm from tracking high correlation peaks caused by image noise. This method assumes that the tracked feature is moving at a speed such that it is always within the region of interest. The position of maximum correlation corresponds to movement of the feature in that image section.

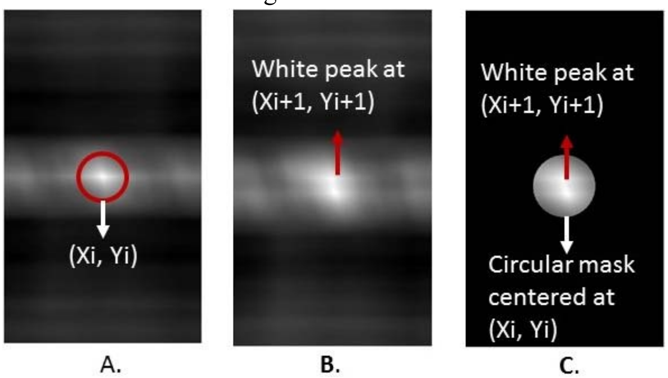

Fig. 2. A. Output image from correlation of the first frame to the second frame. B. Output image from correlation of the second frame to the next frame. C. A mask is applied to correlated image frame of B where the circular regions of interest is centered to the maximum correlation detected on A.

The track of the image feature from all the seven images sections are extracted and combined into a single plot. Linear regression is applied to data points gathered from each image section to get the equation of the line and its corresponding y-intercept. The multiple estimated angles are calculated from the y-intercept values and, after plotting the angles into a histogram, Normal Density Function is fitted on the data. This is used to calculate the mean camera pitch angle and its error. The flowchart of steps for the algorithm is shown in Figure 3.

IV. Methodology

A laboratory experiment was devised to determine the accuracy at which the algorithm is able to estimate the camera rotation or the apparent slope. The experimental setup can be seen in Figure 4. It consists of a long flume tank filled with water. The bottom of the flume tank has been lined with tarpaulin printed with a coral pattern which the algorithm will identify and track features in order to estimate the angle the camera has been rotated. A track is mounted on top of the flume tank to allow a trolley to traverse the entire length of the flume tank. The trolley is equipped with a dipping rod to allow the camera to be lowered and angled as needed. A GoPro Hero 3+ Silver Edition Camera, calibrated using MATLAB Calibration toolbox [5], was used in the experiment and is attached to the dipping rod as depicted in Figure 5. The angle of the camera was measured from the horizontal using the bubble level app for iPhone 4s.

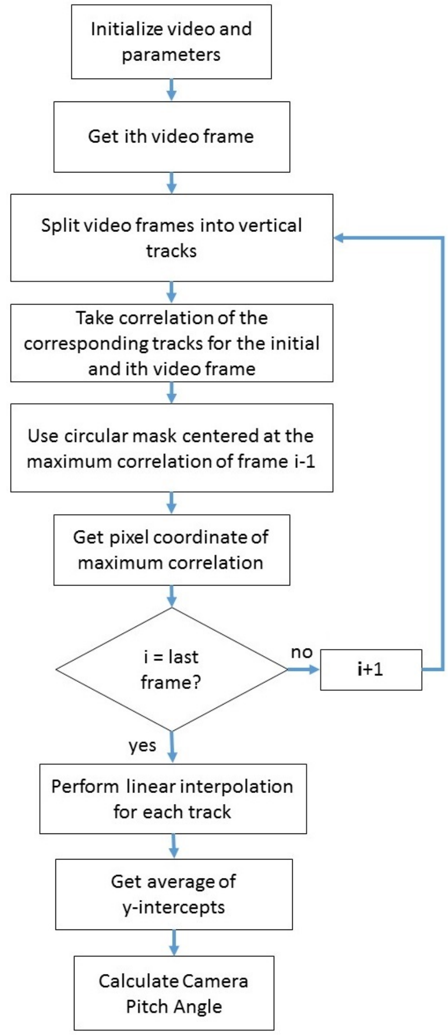

Fig. 3. Flow Chart for Camera Pitch Estimation using videos

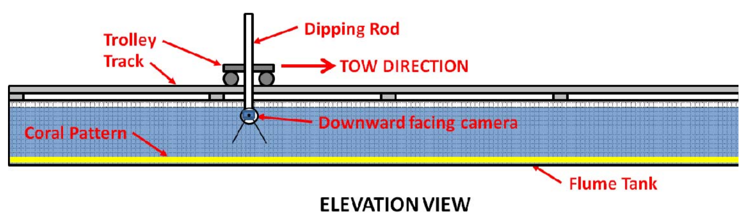

Fig. 4. Experimental Setup for Camera Rotation and Slope Measurement

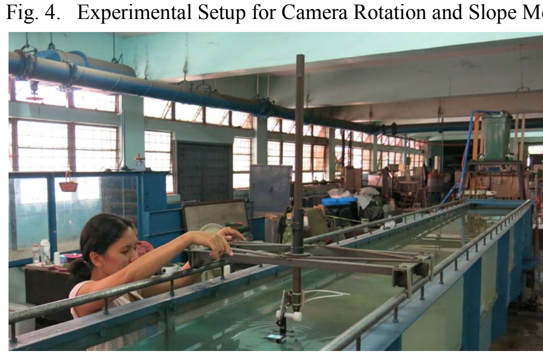

Fig. 5. Trolley and Camera Setup

Using the setup the camera is then dipped into the water at a known pitch angle. The camera is then traversed back and forth along the tank as it records video data. Camera pitch angles tested were 5°, 10° and 16°. The recorded video was then processed using the algorithm to come up with an estimate of camera angle. The actual camera angle and the algorithm estimated camera angle are then compared to validate the accuracy of the algorithm.

V. Results and Discussion

Using the algorithm, about a thousand image frames extracted from video data obtained in the experiment was processed to estimate the camera pitch for each angle. For every fifty frames, the image features moved almost halfway across the frame so the feature track points were linearly interpolated before the tracked image features go beyond the image boundary. The images were partitioned to produce 7 feature tracks across the image frame, comprising of approximately 140 linearly interpolated lines for each angles. From the line tracks, the y-intercepts were extracted and the frequency distribution of the estimated angles were plotted. Normal density function was fitted on each histogram plot to get the mean of the data and measure the standard deviation. Figure 6 shows sample video frames at different angles, their corresponding interpolated track plots and histogram of each measure angles as estimated using the technique.

Using the video tracks from the experiments, the estimated camera pitch results were $4° \pm 2°$ for measured angle of 5°, $8° \pm 3°$ for measured angle of 10° and $13° \pm 5°$ for measured angle of 16°. The mean of the estimated angles are relatively close to the measured plots as shown on Figure 6, where there is at most 3° difference from the actual angles. It is also observed that the standard deviation and error increases with increasing angles. This may be attributed to the method of image partitioning before finding the track movements of the image segment, since increasing the angles limit the movement of the track within the segment. It can be stated that the algorithm works better for predicting small angles.

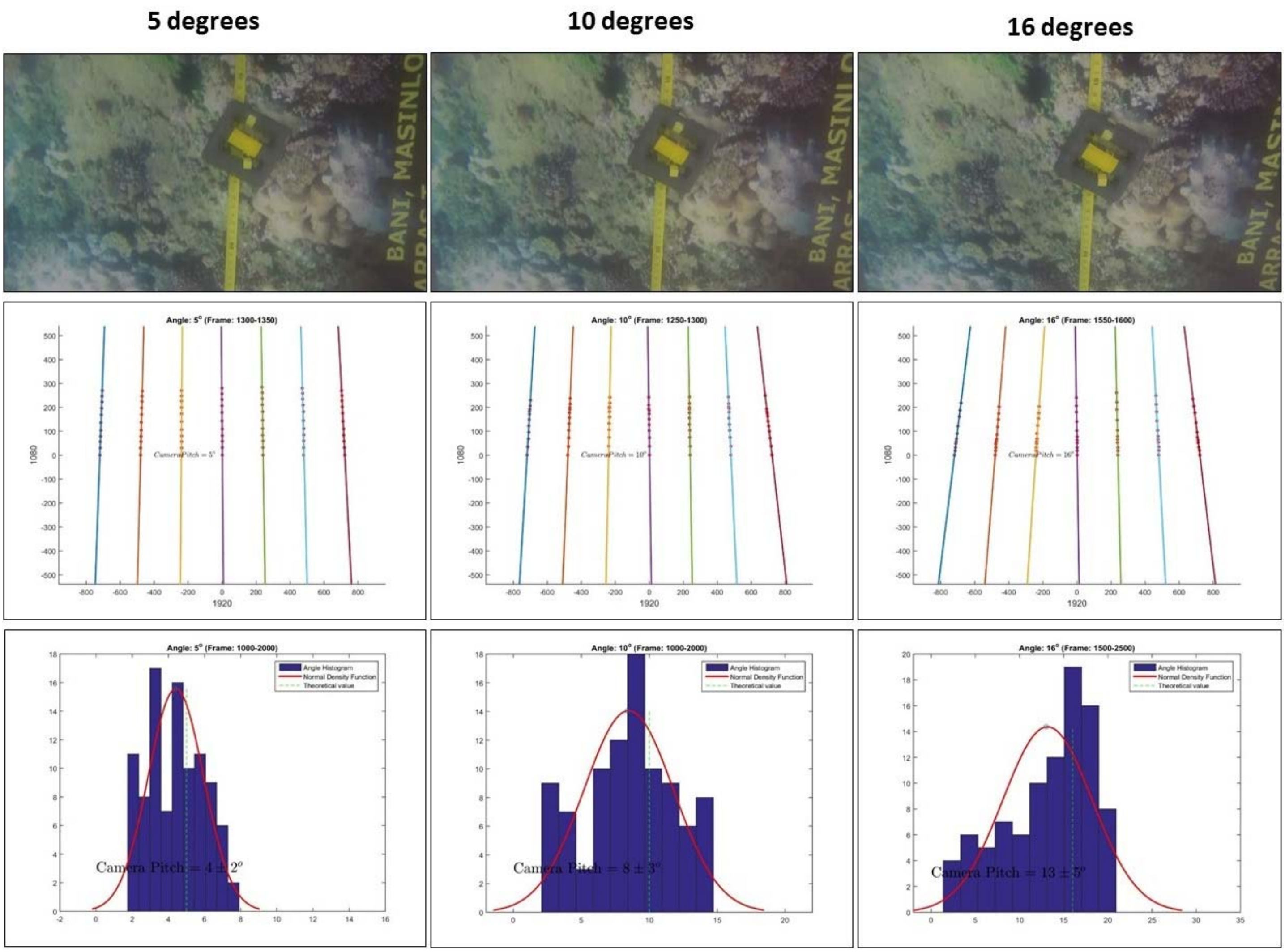

Fig. 6. (top row) Sample video frame extracted from the simulated coral reef videos at angles 5°, 10° and 16°. (middle row) Linearly interpolated feature tracks from 50 video frames. (bottom row) Frequency distribution of the estimated angles with fitted normal distribution function to estimate the mean of the data and get the standard deviation.

VI. Conclusion

A camera estimation technique using apparent motion of the reef that is recorded from videos was presented. The results indicate that the proposed algorithm is a reliable technique for estimating camera rotation angle without the use of sophisticated setups or gadgets. From the increasing standard deviation for increasing angles, it can be concluded that the technique works better for small angles due to image partitioning method that limits the image tracking movement across the segment.

Assuming that the camera's image plane is parallel to the water's surface and its optical axis is perpendicular to the benthic floor, the algorithm could be used to measure the coral reef slope. This study can be extended such that multiple rotation along the camera axis could be extracted from the videos alone.

References

[1] Corpuz, F.J.; Naval, P.; Capili, E.; Jauod, J.; Judilla, R.J.; Soriano, M., "Coral reef mosaicking using Teardrop and Fast Image Labeling," in OCEANS, 2012 - Yeosu , vol., no., pp.1-6, 21-24 May 2012

[2] Doyle D.D., Jennings A.L., Black J.T., "Optical flow background estimation for real-time pan/tilt camera object tracking", Measurement vol 48, pp 195-207, February 2014

[3] He B.W., Li Y.F., "Camera calibration from vanishing points in a vision system", Optics & Laser Technology vol 40, pp. 555-561, April 2008

[4] Yuan D., Liu M., Yun J., Hu J., "Camera motion estimation through monocular normal flow vectors", Pattern Recognition Ltters vol 52, pp. 59-64, 2015

[5] Kanatani K., "Camera Rotation Invariance of Image Characteristics", Computer Vision, Graphics, and Image Processing vol. 39, pp. 328-354, 1987

[6] Jean-Yves Bouguet, "Camera Calibration Toolbox for Matlab" (December 2nd, 2013) < http://www.vision.caltech.edu/bouguetj/calib_doc/ >