

Image-based Motion Stabilization for Maritime Surveillance

Daniel D. Morris, Brian R. Colonna and Franklin D. Snyder

General Dynamics Robotic Systems, 1501 Ardmore Blvd, Pittsburgh, PA 15221*

ABSTRACT

Robust image-based motion stabilization is developed to enable visual surveillance in the maritime domain. The algorithm developed is neither a dense registration method nor a traditional feature-based method, but rather it captures the best aspects of each of these approaches. It avoids feature tracking and so can handle large intra-frame motions, and at the same time it is robust to large lighting variations and moving clutter. It is thus well-suited for challenges in the maritime domain. Advantage is taken of the maritime environment including use of the horizon and shoreline, and fused data from an inexpensive inertial measurement unit. Results of real-time operation on an in-water buoy are presented.

Keywords: Optical stabilization, feature registration, maritime, buoy surveillance, computer vision, inertial measurement unit, water surface, fusion, horizon, shoreline.

1. INTRODUCTION

The first step in surveillance is usually stabilization for camera motion [3,4,9,11], and that is the focus of this paper. While there has been much work in ground and aerial surveillance, the maritime domain has not yet drawn significant attention. Yet the maritime domain presents a number of unique challenges and opportunities for motion stabilization. Floating sensors must deal with constant motion of their own and of the water surface around them. Waves may generate high-texture features, but their motion is non-rigid and their appearance changes rapidly. Maritime lighting typically has greater contrast than on land due to direct sunlight reflections with a resulting loss of detail for cameras with limited dynamic range. Rapid changes in lighting can be countered by an auto-iris, but this leads to changes in object appearance and details. Finally water spray can lead to droplets and salt deposits on the camera protective cover, causing distortions and clutter in the imagery.

Here algorithms are developed for robust image stabilization on a floating buoy. While it is possible to use a gyro-based sensor to do stabilization, this requires a high-precision device which is typically expensive and bulky. Directly using images enables stabilization down to sub-pixel precision, which is precisely the level needed for motion-based detection algorithms.

There are two general approaches to image stabilization: dense techniques and feature-based techniques. In dense techniques images patches are directly warped onto each other or correlated with each other using motion models, for example see [8,9]. On the other hand, feature-based techniques rely on tracking features between images and directly calculating motion from them, see [2,3,4,10,12]. Advantages of dense methods include more pixels being used and requiring explicit feature correspondences is avoided. But these methods are typically not robust to many of the challenges in maritime environments such as rapidly moving specular reflections and large changes in gain and contrast due to lighting changes. Feature-based methods can be quite robust to lighting variations as well as to moving clutter in the scene, however they face the challenge of maintaining feature tracks especially when motion is large. They also have difficulties if the scene does not have stable corner features.

Here a hybrid technique for motion stabilization is presented. It achieves robustness to lighting variations and moving clutter in a similar way as feature techniques, but at the same time, like the dense techniques, it does not require finding corner features or tracking features between frames. It can thus work with very large intra-frame motion and with large lighting variations and moving clutter. The geometry of the maritime domain is leveraged including use of the horizon and shoreline as well as an inertial measurement unit (IMU) if available. Stabilization was implemented in a real-time system onboard a buoy. Sample results are illustrated.

* Work was performed in part at Northrop Grumman Corp, 1501 Ardmore Blvd, Pittsburgh PA 15221.

The paper is organized as follows. First our assumptions are stated in Section 2. Then stabilization of the vertical axis is described in Section 3, followed by stabilization of heading in Section 4. Filtering to combine inertial and image measurements is described in Section 5. Finally results are given and discussed in the conclusion.

2. ASSUMPTIONS

Stabilization is performed using a 360-degree field of view, 5-camera array that has been calibrated such that for each pixel in each camera, (u_i, v_i) there is a known unit 3-vector, \hat{p}_i , giving the direction of the ray incident on that point, namely:

$$\hat{p}_i = f(u_i, v_i), \quad (1)$$

where $f(u_i, v_i)$ is the calibration function.

Since the cameras are close together compared to the distance to detected objects, their optical centers are approximated as being coincident. By working in spherical coordinates, data from all cameras can be treated uniformly. It is assumed that over short time periods the buoy translation is negligible compared to the objects it observes.

Furthermore it is assumed that the horizon or distant shoreline is visible for vertical stabilization, and that shoreline features are available for heading estimation.

3. VERTICAL STABILIZATION

Image motion can be explained as a rotation of the platform, and hence stabilization is achieved by estimating the platform rotation. The approach here is to divide rotation estimation into two sequential steps: first vertical axis estimation and then heading estimation. Vertical axis estimation is described in this section.

The key property of the maritime domain that aids stabilization is that the visible horizon or distant shoreline defines a horizontal plane. Figure 1 illustrates a buoy rotated by ${}^w_B \mathbf{R}$ in world coordinates with respect to a reference frame on the horizontal plane, H . The third row of ${}^w_B \mathbf{R}$, is ${}^B \mathbf{z}_w^T$, the transpose of the vertical axis in the buoy reference frame. In Euler coordinates this is $[-\sin(p), \cos(p)\sin(r), \cos(p)\cos(r)]$, see 5 pg. 46, where r and p are the roll and pitch around the world x and y axes respectively. Thus finding the vertical vector, ${}^B \mathbf{z}_w$, or equivalently the horizontal plane, in buoy coordinates is sufficient to determine the buoy roll and pitch.

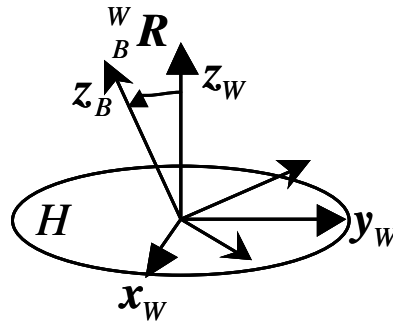


Figure 1 The world coordinate system is defined such that \mathbf{z}_w is perpendicular to the horizontal plane H . The buoy coordinate system is rotated with respect to this by ${}^w_B \mathbf{R}$.

It is assumed that the horizon (or a sufficiently distant shoreline) is visible as a contrast change in part of the 360-degree image. However the images will typically be highly textured with waves, clouds and shoreline providing clutter from which the horizon must be extracted. The following robust technique was used to determine the horizontal plane.

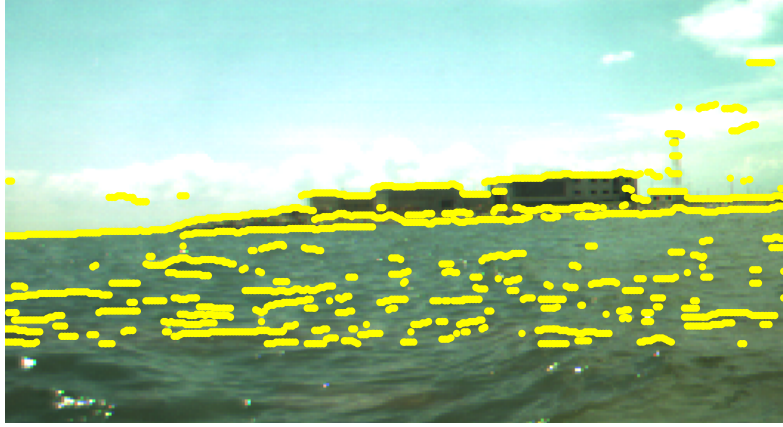


Figure 2 Peaks and troughs of the vertical gradients are found in each image column in the region around the horizon. These points include the horizon, distant shoreline and other clutter. The horizon is found by determining the plane that explains the most peaks in all simultaneous images.

For each column of the smoothed vertical gradient images of all the cameras, all the local maxima and minima in the region around the predicted horizon and above a small threshold are found. A subset of these will correspond to the horizon. Each of the maxima and minima is mapped to a point on the unit sphere, \hat{p}_k , using Eq. (1). Now the horizon points will all lie on the plane H , and at the same time it is unlikely that there will be any other plane through the origin generating a large number of points \hat{p}_k . Thus a robust technique, such as RANSAC [6], is used to find the best plane fitting these points. Pairs of points are sufficient to define the perpendicular to a plane through their cross product:

$$\mathbf{v} = \hat{p}_j \times \hat{p}_k \quad (2)$$

The vector \mathbf{v} with the most inliers is the initial estimate for ${}^B\mathbf{z}_w$, and a least squares estimate using the inliers can be obtained as the eigenvector corresponding to the minimum eigenvalue of:

$$A = \sum_{i \in \text{inliers}} \mathbf{p}_i \mathbf{p}_i^T. \quad (3)$$

Examples of stabilization are shown in Figures 8 and 9.

4. HEADING STABILIZATION

The next step is to determine the change in heading. Our approach to achieving this is to temporally align vertical features on the visible shoreline. We want these features to be robust to lighting variation and we want to avoid tracking individual features. A technique that achieves this is based on vertical curves.

Vertical curves trace the vertical edges of objects and landmarks on the shoreline, see Figure 3. They are obtained with sub-pixel accuracy and are invariant to brightness and contrast changes. First the horizontal gradients of all the images are found. The maxima and minima of the gradient for each row are parabolically fit to sub-pixel accuracy, and roughly vertical contour-based curves are created by connecting close-by maxima and minima between adjacent horizontal rows as illustrated in Figure 3.

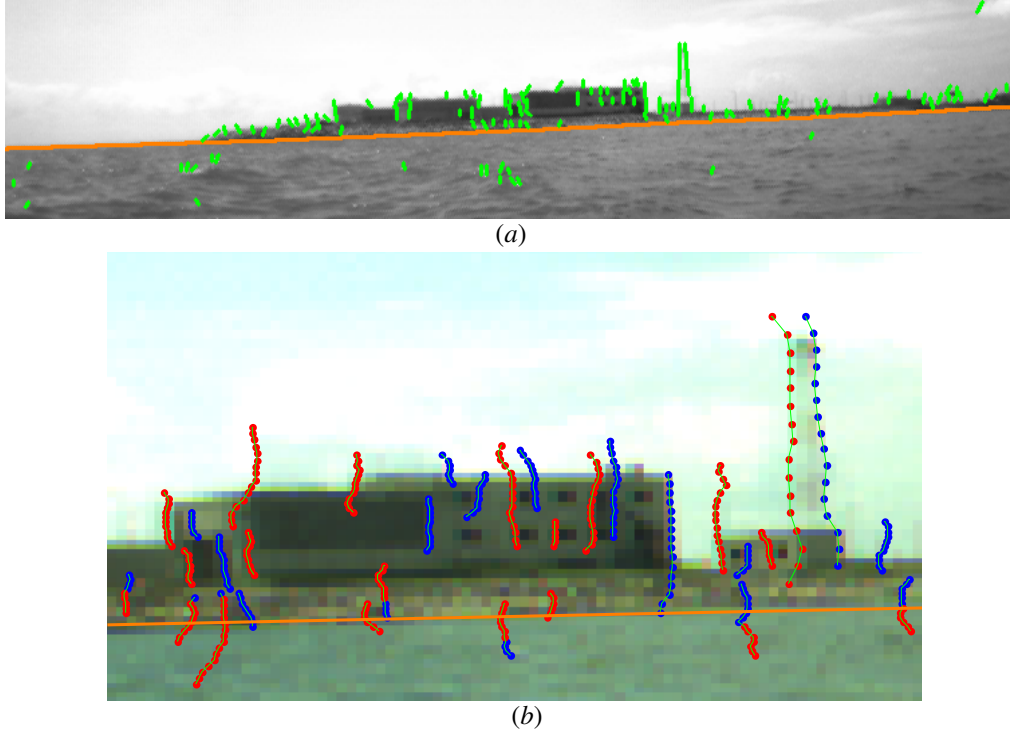


Figure 3 (a) Curves found on objects on the shoreline. (b) Close-up showing the peaks and troughs of the horizontal gradient (red and blue dots), and the curves traced through them.

Curves are built in image space, but then transformed onto the unit sphere using Eq. (1). They are then stabilized with respect to roll and pitch by rotating by ${}^w_B \mathbf{R}^T$ with zero heading. After this the elevation and azimuth angles, (φ_i, θ_i) , for each transformed point, $\hat{\mathbf{p}}_i$, on each curve can be calculated up to an unknown overall heading:

$$\varphi_i = \arctan\left(\sqrt{p_{xi}^2 + p_{yi}^2}, p_{zi}\right) \quad (4)$$

$$\theta_i = \arctan(p_y, p_x) \quad (5)$$

It is assumed that a portion of shoreline is visible above the horizon, and this is used for finding heading change as follows. A series of horizontal slices, $s_t(\varphi_k; \theta)$, at a set of elevations φ_k and time t of the transformed curves are made, and for each slice the locations, θ , of all curve intersections are recorded with a ± 1 pixels depending on the sign of the curve at that point, see Figure 4. Corresponding slices are made through subsequent images, and the relative heading found by a circular convolution of these slices with the slices at previous times. Circular convolutions can be efficiently calculated with the use of fast Fourier transforms and their inverses: FFTs and IFFTs respectively. If $S_t(\varphi_k; \theta)$ is the FFT of $s_t(\varphi_k; \theta)$ and $\tilde{S}_t(\varphi_k; \theta)$ the complex conjugate, then the sum of the circular convolutions with slices at time t_m is given by:

$$w_m(\theta) = \sum_{slices} \text{IFFT}\left\{S_{im}(\varphi_k; \theta) \cdot G(\theta) \cdot \tilde{S}_{i0}(\varphi_k; \theta)\right\} \quad (6)$$

where $G(\theta)$ is the FFT of a Gaussian added for smoothing and to reduce sensitivity to calibration imprecision. There is no need for zero padding as the convolution is circular. The location of the peak of $w_m(\theta)$ gives the heading with respect to time t_0 . This is robust to moving objects and spurious curves, since curves that do not have a match do not contribute to the result. Hence a robust and precise heading is obtained.

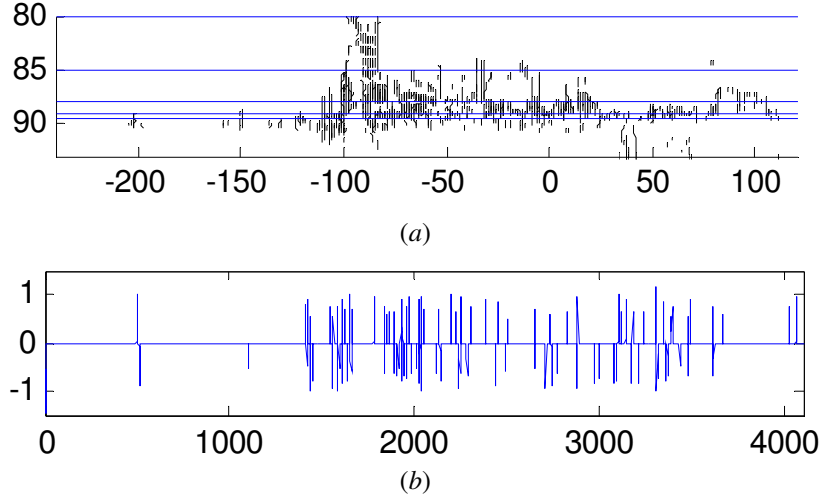


Figure 4 (a) Using the roll and pitch estimated previously, vertical curves are transformed into the (φ, θ) space. A series of horizontal slices cut through these at various elevations above the horizon. The interpolated location of these curves on one of these slices is shown in (b).

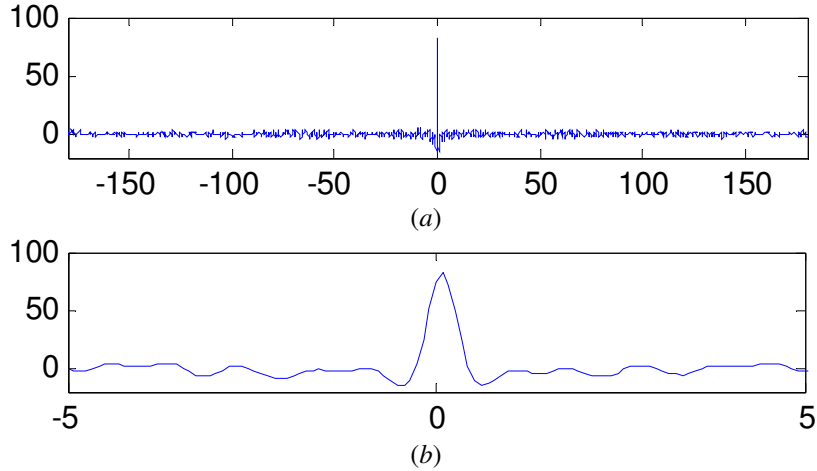


Figure 5 (a) Convolution $w_m(\theta)$ from Eq. (6) of the slices in Figure 4. The lower chart is a close-up showing a small change in heading at the maximum in degrees.

5. FILTERING

Rotation estimates can be improved by filtering. This enables the inclusion of a buoy dynamic model and motion estimates from the IMU. A standard Kalman filter was implemented to achieve this with the following particular properties. The buoy dynamics were modeled as independent damped harmonic oscillators in roll and pitch, and damped angular speed model in heading. With state vector, $\mathbf{x} = [r, \dot{r}, p, \dot{p}, h, \dot{h}]^T$, containing roll, pitch, heading and their time derivatives, the equation of motion is:

$$\dot{\mathbf{x}} = \mathbf{F}\mathbf{x} + \mathbf{u}(t). \quad (7)$$

The component of \mathbf{x} for roll is $[r \ \dot{r}]^T$ and its dynamics are described by:

$$\mathbf{F}_{roll} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -\gamma \end{bmatrix} \quad (8)$$

where ω and γ are the characteristic buoy angular speeds and damping respectively. Analogous expressions apply to pitch and heading (with $\omega = 0$ for the latter). Together they form a block-diagonal \mathbf{F} , and state transfer function:

$$\Phi = \exp(\mathbf{F} \Delta t). \quad (9)$$

The driving term, $\mathbf{u}(t)$, is the action of the waves and is unknown, and so is modeled as system noise, \mathbf{Q} . It acts as a continuous acceleration term and so to integrate it into our discrete formulation the following equation was used for roll noise (and similar equations for pitch and heading):

$$\dot{\mathbf{Q}}_{roll} = \mathbf{F}_{roll} \mathbf{Q}_{roll} + \mathbf{Q}_{roll} \mathbf{F}_{roll}^T + \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2 \end{bmatrix}, \quad (10)$$

where σ^2 is the measure of the acceleration from wave motion. This can be re-written in the form:

$$\dot{\mathbf{q}} = \mathbf{M} \mathbf{q} + \begin{bmatrix} 0 \\ 0 \\ \sigma^2 \end{bmatrix} \quad (11)$$

where

$$\mathbf{q} = \begin{bmatrix} q_{11} \\ q_{12} \\ q_{22} \end{bmatrix}, \quad \mathbf{Q}_{roll} = \begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix}, \quad \text{and} \quad \mathbf{M} = \begin{bmatrix} 0 & 2 & 0 \\ -\omega^2 & -\gamma & 1 \\ 0 & -2\omega^2 & -2\gamma \end{bmatrix}.$$

The solution is given by:

$$\mathbf{q} = \mathbf{M}^{-1} (\exp(\mathbf{M} \Delta t) - \mathbf{I}) \begin{bmatrix} 0 \\ 0 \\ \sigma^2 \end{bmatrix}, \quad (12)$$

from which \mathbf{Q}_{roll} is obtained. \mathbf{Q}_{pitch} and $\mathbf{Q}_{heading}$ are obtained in a similar manner and together form the block-diagonal \mathbf{Q} .

The measurement matrix, \mathbf{H} , defined by $\mathbf{H}\mathbf{x} = \mathbf{z}$ where \mathbf{z} contains the measured quantities, is simple to calculate. The measurements include $r, \dot{r}, p, \dot{p}, h, \dot{h}$ directly from the IMU, and $r, p, \Delta h$ from the images, where Δh is the heading relative to the previous image. The covariance, \mathbf{R} , on the measurements was chosen to have much larger terms on the IMU components than the image components. All these terms are plugged into standard discrete Kalman filter equations and produce a fused image and inertial stabilization system. For completeness these equations are summarized as follows:

$$\begin{aligned} \mathbf{x}_t^{(-)} &= \Phi_t \mathbf{x}_{t-1}^{(+)} \\ \mathbf{P}_t^{(-)} &= \Phi_t \mathbf{P}_{t-1}^{(+)} \Phi_t^T + \mathbf{Q}_{t-1} \\ \mathbf{K}_t &= \mathbf{P}_t^{(-)} \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_t^{(-)} \mathbf{H}_t^T + \mathbf{R}_t)^{-1} \\ \mathbf{x}_t^{(+)} &= \mathbf{x}_t^{(-)} + \mathbf{K}_t (\mathbf{z}_t - \mathbf{H}_t \mathbf{x}_t^{(-)}) \\ \mathbf{P}_t^{(+)} &= \mathbf{P}_t^{(-)} - \mathbf{K}_t \mathbf{H}_t \mathbf{P}_t^{(-)} \end{aligned} \quad (13)$$

6. RESULTS

Stabilization through horizon-finding in five cameras turned out to be very robust. Examples are shown in Figures 7, 8 and 9. Using an inexpensive inertial measurement device significantly sped up the computation by reducing the search space as well as the percentage of outliers to be dealt with by the RANSAC operation. Relative heading estimation was fast, requiring only 1D FFTs and gave pixel-level precision, see Figure 5. A comparison of IMU and image-stabilization as well as fused results is given in Figure 6. Distant ships can help in short-term stabilization, although over longer term their

motion relative to the shoreline can be detected. Stabilization was easily calculated with a Pentium III processor with data from 5 1024x768 cameras at 7 frames per second.

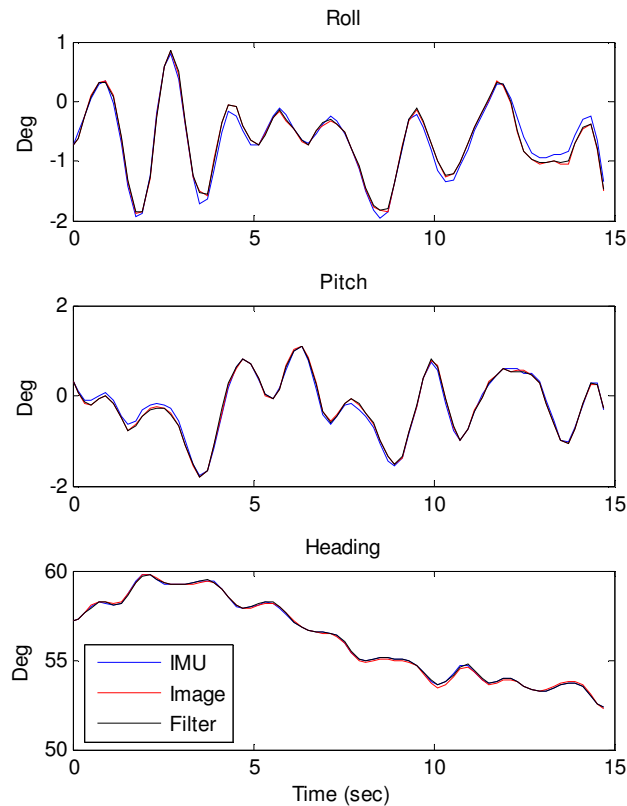


Figure 6 Comparison rotation estimates from IMU and the image-based algorithm and a filtered approach that fuses these. The IMU is rated to accurate to a RMS accuracy of 2° whereas the image-based heading is accurate down to at least pixel resolution, which for these cameras is $\pm 0.2^\circ$. In this example of low acceleration, the IMU accuracy is much higher than its rated value and comparable to the image technique. For larger motions, however, the IMU accuracy degrades as can be seen in Figure 9, whereas the image-based technique maintains high accuracy.

7. CONCLUSION

An efficient and robust image-based motion stabilization technique was developed for the maritime domain. The use of features based on local maxima and minima of the gradient images gave robustness to rapid lighting changes. The use of curves rather than corners is more appropriate to maritime environments where corner features may be rare. Avoiding the need to track features makes the technique robust to large intra-frame motion. Using spherical coordinates enables measurements from cameras pointing in all directions to be used, greatly reducing ambiguities in horizon-finding that occur in single-camera solutions. The next step is to use the motion stabilization to achieve moving object detection. The curved features developed here can be used for this. Obtaining the absolute vertical, as this does, rather than simply an incremental change in rotation, enables precise surveillance applications that search for objects close to the horizon.

A limitation of image-based techniques is that they depend on environmental conditions. Fog or haze can obstruct the view of the horizon or shoreline. In these cases it may be necessary to rely on inertial measurements which can easily be done with our filtering approach. A useful advantage of the image-based approach is that accuracy can be increased simply by adding more cameras with higher resolution.

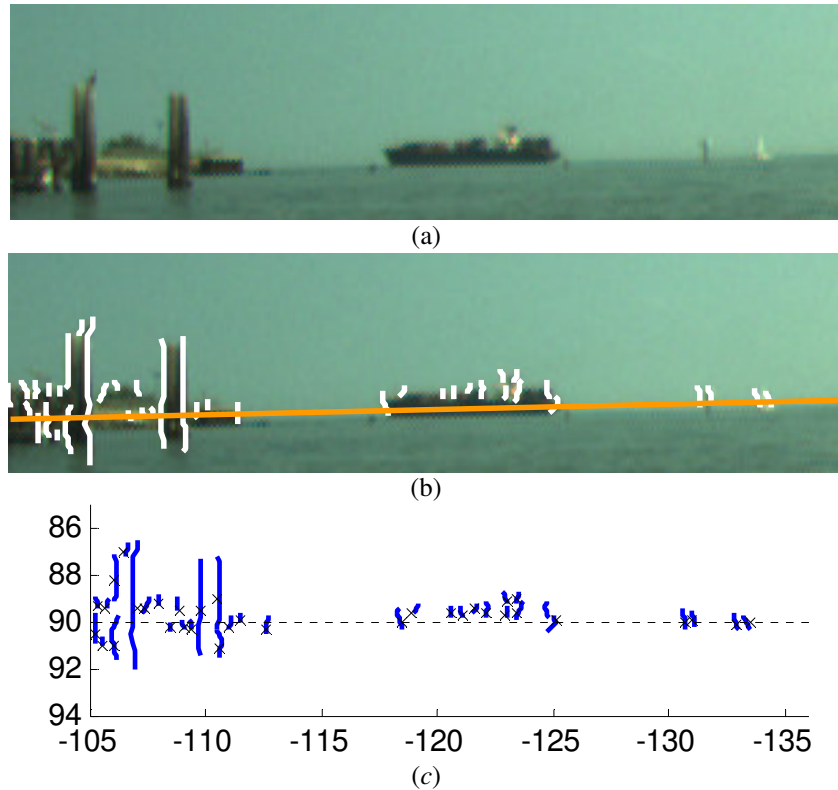


Figure 7 (a) A portion of the field of view of the camera array including shoreline and watercraft. (b) Horizon and curves are shown overlaid. (c) Curves are transformed into stabilized, spherical coordinates.

ACKNOWLEDGEMENT

This work was performed under the sponsorship of ONR FNC-AO-IA, PM Marc Steinberg, #N00421-03-C-0027 and # DAAD19-01-0066.

REFERENCES:

1. H. Asada, M. Brady, "The Curvature Primal Sketch," IEEE PAMI 8(1):2-14, 1986.
2. D. Burschka, G. Hager, "Vision-based control of mobile robots," in Proc. International Conference on Robotics and Automation, pages 1707-1713, 2001.
3. A. Censi, A. Fusiello, and V. Roberto, "Image stabilization by features tracking," in Proc. 10th Int. Conf. Image Analysis and Processing, Sep 1999, pp. 665–667.
4. I. Cohen, G. Medioni "Detecting and tracking moving objects in video surveillance," Proc IEEE Conf Computer Vision and Pattern Recognition 1999; II: 319–325
5. J.J. Craig, "Introduction to Robotics Mechanics and Control," Second Edition, Addison Wesley Longman, 1999.
6. M.A. Fischler, R.C. Bolles. "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," in Comm. of the ACM, volume 24, pages 381--395, 1981.
7. C. Harris, M. Stephens, "A combined corner and edge detector", in Alvey Vision Conf., 1988, pp. 147-151.
8. Irani, M., Rousso, B., and Peleg, S., "Recovery of Ego-Motion Using Image Stabilization," CVPR 1994, pp. 454-460.
9. C.D. Kuglin, D.C. Hines. "The phase correlation image alignment method," IEEE Conference on Cybernetics and Society, p163-165, 1975.
10. A. Litvin, J. Konrad, W.C. Karl "Probabilistic video stabilization using Kalman filtering and mosaicking" Image and Video Communication and Processing 2003, SPIE Vo. 5022, pp.663-674.

11. D.G. Lowe, "Object recognition from local scale-invariant features," International Conference on Computer Vision, Corfu, Greece (September 1999), pp. 1150-1157.
12. L. Marcenaro, G. Vernazza, C.S. Regazzoni, "Image stabilization algorithms for video-surveillance applications," in Proc. Int. Conf. Image Processing 2001, Vol 1, pp. 349-352.
13. E. Shilat, M. Werman, Y. Gdalyahu, "Ridge's corner detection and correspondence," IEEE Proc. CVPR, 1997, pp. 976-981.

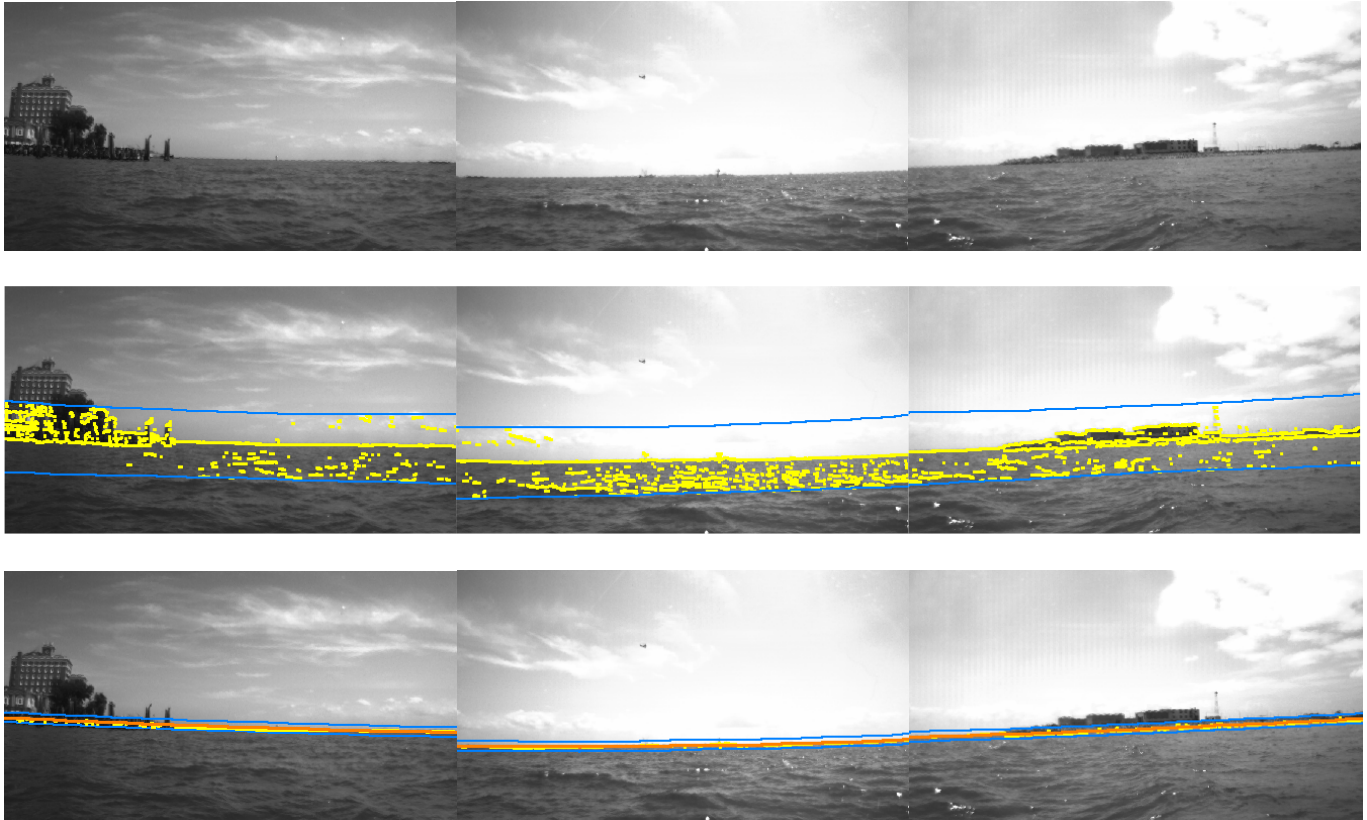


Figure 8 Horizon stabilization shown on 3 of 5 images. Top row shows original images. The vertical offsets and radial distortions are unimportant as they are accounted for in the calibration, Eq (1). Blue lines in center row show region bounded by uncertainty of the IMU. In this region the yellow dots show all the vertical gradient peaks. Using RANSAC a subset of these are determined to be inliers to the horizon and these are plotted in the bottom row with the blue lines being the bounds on the inliers. The final estimate of the horizon is shown by the orange line and is obtained as a least squares fit to the inliers, see Eq. (3).

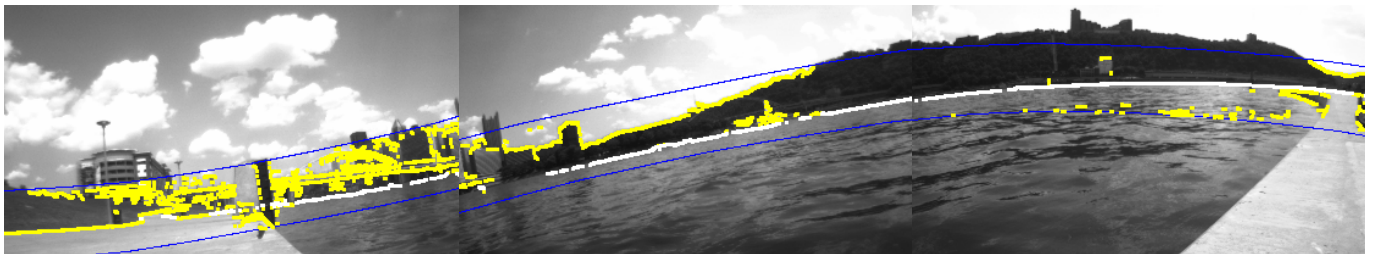


Figure 9 Another example of horizon fitting with very large roll and pitch shown in 3 of the 5 cameras. The two blue curves are bounds of the search region given by the IMU and centered around its estimate which shows significant error compared to the final image-estimate. Gradient peaks are highlighted in this region. Application of RANSAC finds inliers to the horizontal plane, here marked as white.