Multiple Views Tracking of Maritime Targets

Thomas Albrecht, Geoff A.W. West, and Tele Tan
Curtin University of Technology
Western Australia

Thanh Ly
Defence and Technology Organisation
Australia

Abstract

This paper explores techniques for multiple views target tracking in a maritime environment using a mobile surveillance platform. We utilise an omnidirectional camera to capture full spherical video and use an Inertial Measurement Unit (IMU) to estimate the platform's ego-motion. For each target a part of the omnidirectional video is extracted, forming a corresponding set of virtual cameras. Each target is then tracked using a dynamic template matching method and particle filtering. Its predictions are then used to continuously adjust the orientations of the virtual cameras, keeping a lock on the targets. We demonstrate the performance of the application in several real-world maritime settings.

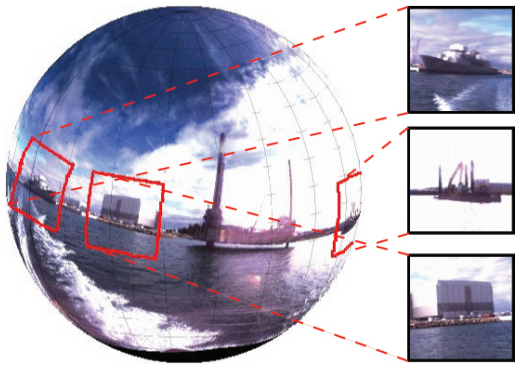


Figure 1. Three Virtual Cameras extracted from the omnidirectional camera video. Note that the virtual camera images look horizontally flipped as they represent the view from within the sphere.

1. Introduction

In high-risk, hazardous, inaccessible, or remote areas surveillance systems play an important role to minimise security threats. Fixed and pan-tilt-zoom (PTZ) cameras, which have been traditionally used for surveillance, have only limited fields of view (FOV) at any one time thus require multiple installations to cover larger areas. Using mobile platforms for field surveillance minimises the need for significant infrastructure. Furthermore, in remote or inaccessible areas, their use is essential for full situational awareness as they are able to navigate around obstacles and cover areas where the line of sights are restricted. Coastal and port surveillance can benefit from mobile maritime platforms as they are capable of exploring the dynamic and fast changing environment specific to this type of area. Once potential security threats have been identified, their tracking and mitigation is essential to ensure safety and security. As there can easily be more than one object of interest, the use of omnidirectional cameras is advantageous as they provide continuous 360° video footage, which fixed or PTZ cameras cannot match. When using omnidirectional cameras, the correspondence between the cameras needs to be fused into a single geometric domain. A natural representation is the projection onto a unit sphere with the omnidirectional camera situated in the centre of the sphere [1]. With this, it is possible to extract a multitude of target regions of interest. This effectively forms a set of *virtual cameras* (Figure 1) that can be used to create focused views onto target objects.

On a mobile maritime platform the intended movement is overlaid by disturbances caused by swell and waves. The overall motion of the platform is called *ego-motion*. Due to its erratic characteristics, ego-motion is not predictable and hence needs to be measured. The rotational component can be measured effectively using gyroscopes contained in an Inertial Measurement Unit (IMU), which is rigidly attached to the camera system. Applying the estimated change of orientation to the set of virtual cameras results in an rotation stabilised view for each camera. However, additional effort is necessary to create focused views: the calibration between camera system and IMU needs to be taken into account [2], the translational component of the platform's ego-motion needs to be estimated, and finally, the trajectory of the target objects needs to be computed. This way, the operator is provided with a stabilised view onto the target objects and assisted in evaluating the potential threat.

In this paper, we use the same camera system, equipped with an omnidirectional camera and an IMU, as in [3] on

978-0-7695-4271-3/10 $26.00 © 2010 IEEE
DOI 10.1109/DICTA.2010.59

a mobile surveillance platform. We utilise the platform for multi view target tracking in a maritime environment. For each target object, we extract a region with limited FOV from the omnidirectional video, forming a set of virtual cameras. We make use of the IMU to form estimates of the platform's ego-motion. We then use a particle filtered template matching approach to track target objects in a global coordinate space and continuously update the orientations of the virtual cameras to keep each of the targets in view. The probabilistic approach provides a suitable framework to combine the measurements from the different sensors over time and makes the tracker robust to noise, occlusions, and synchronisation issues.

This paper is organised as follows: section 2 discusses related work, whilst in section 3 notation and coordinate systems are introduced and the tracking system is described. The experimental setup and the results are discussed in section 4 with conclusions given in section 5.

2. Related Work

Being able to selectively choose where to look attracted much attention in the late 1980s and early 1990 and went under the general term *Active Vision* [4]. Many systems were built using motor-powered and servo-controlled cameras, but it was recognised that high performance was needed for such systems to function, e.g. a rotational speed of $500°s^{-1}$ and an acceleration of $5000°s^{-2}$ [5], which was expensive to obtain using hardware. Research has continued in this field as PTZ cameras became widely available. A tracking system that uses a PTZ camera has been developed by Kumar et al. [6], their system mechanically adjusts the orientation of the camera platform according to the position of the target object. As a change of orientation might cause a second target to drop out of view, the system is not capable of robustly tracking multiple objects. A tracking system that uses a wide FOV to detect moving objects and a high resolution PTZ camera with a narrow FOV to perform the tracking was reported by Bashir and Porikli [7]. In [8], Kang et al. demonstrated a system that fuses video streams of stationary and PTZ cameras using an adaptive background model for tracking.
Using an omnidirectional camera overcomes the mechanical and FOV constraints as the camera does not have any moving parts that restrict the angular velocity and provides instantaneous $360°$ view. From the omnidirectional view, a virtual camera representing a region of interest can be extracted. This concept was used by [3] to create a stabilised window showing a region of interest within an omnidirectional video, while the camera is subject to significant egomotion. Sun et al. [9] used a virtual camera to detect and track a person in an indoor environment using a in a single wide angle panoramic camera.

Note that the focus of this paper is the tracking of target objects using multiple views. However, multiple target tracking algorithms like [10] or [11] could be applied to the presented camera system as well.

3. System Description

We utilise the Ladybug 2, an omnidirectional camera manufactured by Point Grey Research and an IMU, MTi, manufactured by Xsens. The Ladybug 2 camera consists of six individual cameras each capturing 1024×768 pixels at 30 frames per second. Five cameras are horizontally aligned in a ring, with the sixth pointing upwards. In this setup the system can capture about 80% of the whole sphere. It is precalibrated, and the geometry between the cameras is provided by the manufacturer. This allows for fast and precise spherical mapping, negating the need to register all images. The MTi IMU has acceleration, gyroscopic, and magnetic sensors that are fused using a hardware-based Kalman filter. It outputs calibrated measurements of acceleration and angular velocity, as well as a drift-free 3D orientation with a static accuracy of $\leq 1.0°$, at a maximum sample rate of 100Hz.

3.1. Notation

The following notation is used in this paper: p^{A} denotes a point in a coordinate system A, $\mathbf{p}^{\mathrm{A}}$ the local vector from the origin of A towards p^{A}. The 3×3 rotation matrix $\mathbf{R}_{\mathrm{B}}^{\mathrm{A}}$ denotes the orientation of A w.r.t. another coordinate system B. The 3×1 vector $\mathbf{t}_{\mathrm{B}}^{\mathrm{A}}$ is the translational offset of A w.r.t. B. Both are combined using homogeneous coordinates resulting in the 4×4 transformation matrix $\mathbf{T}_{\mathrm{B}}^{\mathrm{A}}$

$$\mathbf{T}_{\mathrm{B}}^{\mathrm{A}} = \begin{pmatrix} \mathbf{R}_{\mathrm{B}}^{\mathrm{A}} & \mathbf{t}_{\mathrm{B}}^{\mathrm{A}} \\ 0 & 1 \end{pmatrix}. \tag{1}$$

$\mathbf{T}_{\mathrm{A}}^{\mathrm{B}}$ subsequently defines the inverse transformation $(\mathbf{T}_{\mathrm{B}}^{\mathrm{A}})^{-1}$.

3.2. Flat Earth

The shape of earth is unique. Thus, exact computations on the earth's surface can be complex. In this paper, we are dealing with close range distances within the line of sight, we adopt a flat earth approximation. The vicinity of a fixed reference point $(\Phi_0\,, \lambda_0)$ on the earth's surface can be approximated using a planar projection, resulting in a mapping where the circles of latitude and the lines of longitude are equidistant, straight and cross at right angles [12]. As the circumference of the circles is dependent on Φ_0, the length of a radian $r'(\Phi_0)$ and the radius of the curvature $r''(\Phi_0)$ are computed as functions of the reference latitude [12]. For the parameters for the equatorial radius and flattening the earth, the World Geodetic System (WGS84) [13]

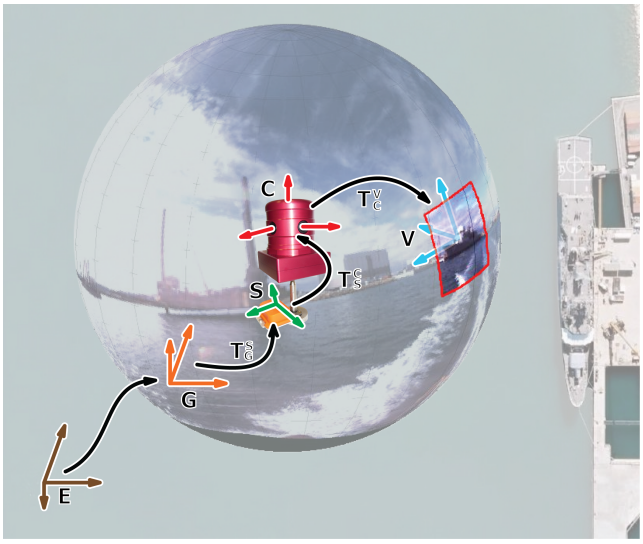


Figure 2. The global coordinate system G represents an upright unit-sphere projection at the current position of the mobile platform within the earth coordinate system E. The orientation of the sensor coordinate system w.r.t the global coordinate system G at any one time step t, as measured by the IMU, is denoted as $\mathbf{R}^{\mathrm{S}}_{\mathrm{G},t}$, while camera C and sensor S coordinate systems are rigidly connected, denoted by $\mathbf{T}^{\mathrm{C}}_{\mathrm{S}}$. Note that all coordinate systems are right-handed.

is used. This maps a point $(\Phi\,,\lambda)$ at sea level altitude into local Cartesian coordinates $(x\,,y\,,z)$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r'(\Phi_0) & 0 \\ 0 & r''(\Phi_0) \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Phi - \Phi_0 \\ \lambda - \lambda_0 \end{pmatrix}. \tag{2}$$

3.3. Coordinate Systems

The following five coordinate systems (Figures 1 and 2) are used in this paper:

1. The earth coordinate system (E) uses the model of the flat earth as described in section 3.2, with the origin located at the reference point $(\Phi_0\,,\lambda_0)$.

2. The global coordinate system (G) is a projection of earth coordinates $(\Phi\,,\lambda)$ onto the unit sphere at the current position $(\Phi_{t=0}\,,\lambda_{t=0}) \rightarrow p^{\mathrm{G}}_{t=0}$ of the platform. The y-axis of G is aligned with the line of longitude at $(\Phi_{t=0}\,,\lambda_{t=0})$ and pointing towards North.

3. The sensor coordinate system (S) is defined w.r.t. G. Its orientation is measured by the IMU at every time step t, denoted as the homogeneous transformation $\mathbf{T}^{\mathrm{S}}_{\mathrm{G},t}$.

4. The camera coordinate system (C) is defined with the origin in the centre of the omnidirectional camera. The transformation of C w.r.t. S is denoted as $\mathbf{T}^{\mathrm{C}}_{\mathrm{S}}$. Note that $\mathbf{T}^{\mathrm{C}}_{\mathrm{S}}$ is constant as the camera and IMU are rigidly connected; we use the method of [2] to compute an estimate.

5. The virtual camera coordinate systems ($\mathrm{V}_{1\ldots N}$), where N is the number of virtual cameras used, are defined at the centre of each virtual camera w.r.t. C. $\mathbf{T}^{\mathrm{V}_n}_{\mathrm{C},t}$ describes the transformation from C to the virtual camera V_n at time step t. See section 3.4.

The transformation of a point p^{C} into p^{G} at time step t is thus given as

$$\mathbf{p}^{\mathrm{G}} = \mathbf{T}^{\mathrm{G}}_{\mathrm{S},t}\ \mathbf{T}^{\mathrm{S}}_{\mathrm{C}}\ \mathbf{p}^{\mathrm{C}}. \tag{3}$$

3.4. Virtual Camera

A virtual camera (Figure 1) is a sub-window extracted from a full spherical view. The virtual camera is defined by its orientation $\mathbf{R}^{\mathrm{V}_n}_{\mathrm{C},t}$ w.r.t. C and the FOV $\alpha_{n,t}$ at time step t. Applying the perspective projection with z_1 and z_2 as the projection's far and near clipping respectively yields the transformation from C to the virtual camera coordinates V_n as $\mathbf{T}^{\mathrm{V}_n}_{\mathrm{C},t}$ at time step t

$$\mathbf{T}^{\mathrm{V}_n}_{\mathrm{C},t} = \begin{pmatrix} \cot\frac{\alpha_{n,t}}{2} & 0 & 0 & 0 \\ 0 & \cot\frac{\alpha_{n,t}}{2} & 0 & 0 \\ 0 & 0 & \frac{z_1+z_2}{z_2-z_1} & \frac{2\,z_1\,z_2}{z_2-z_1} \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{R}^{\mathrm{V}_n}_{\mathrm{C},t} & 0 \\ 0 & 1 \end{pmatrix}. \tag{4}$$

A point p^{V_n} can thus be transformed into global coordinates as

$$\mathbf{p}^{\mathrm{G}} = \mathbf{T}^{\mathrm{G}}_{\mathrm{S},t}\ \mathbf{T}^{\mathrm{S}}_{\mathrm{C}}\ \mathbf{T}^{\mathrm{C}}_{\mathrm{V}_n,t}\ \mathbf{p}^{\mathrm{V}_n}. \tag{5}$$

3.5. Particle Filter Framework

We make use of a particle filter framework [14] to track the target objects in global coordinates G and use the predicted position $p^{\mathrm{G}}_{t|t-1}$ to adjust the orientation of the virtual camera (see section 3.6). The state-space model of the particle filter can be described with the state-vector $\mathbf{x} = (x\,,y\,,z\,,\dot{x}\,,\dot{y}\,,\dot{z})^T$ containing the global 3D-projection onto the unit sphere and velocity of the target object.
The Ladybug 2 camera does not possess a hardware trigger interface, thus data from the camera and other sensors are not exactly synchronised. Following [3], we model this as uncertainty of the position estimation.

3.6. Object Tracking

We manually initialise the tracking process by creating virtual cameras $\mathrm{V}_{n=1..N}$ centred on each of the $p^{\mathrm{C}}_{n=1..N}$ target objects and extract image templates $\mathbf{J}_{n=1..N}$. The initial

orientations $\mathbf{R}_{\mathrm{C},t=0}^{\mathrm{V}_{n=1..N}}$ can be computed using Rodrigues' rotation formula

$$\mathbf{R}_{\mathrm{C},t=0}^{\mathrm{V}_{n=1..N}} = \mathbf{R}_{\tilde{\mathbf{\Omega}}}(\beta) = \mathbf{I}_3 + \sin\beta \cdot \tilde{\mathbf{\Omega}} + (1 - \cos\beta) \cdot \tilde{\mathbf{\Omega}}^2, \tag{6}$$

where the skew symmetric matrix $\tilde{\mathbf{\Omega}}$ is defined as

$$\tilde{\mathbf{\Omega}} = \begin{pmatrix} 0 & -\tilde{\omega}_3 & \tilde{\omega}_2 \\ \tilde{\omega}_3 & 0 & -\tilde{\omega}_1 \\ -\tilde{\omega}_2 & \tilde{\omega}_1 & 0 \end{pmatrix}, \text{ with } \tilde{\omega} = \mathbf{u}_0^{\mathrm{C}} \times \mathbf{p}_{n,t=0}^{\mathrm{C}}, \tag{7}$$

and β is the angle between the C-unit vector and the centre of the virtual camera

$$\beta = \parallel \mathbf{u}_0^{\mathrm{C}} \times \mathbf{p}_{t=0}^{\mathrm{C}} \parallel_2. \tag{8}$$

In subsequent frames, $t+1$, we compute for each target a normalised correlation coefficient between the corresponding template $\mathbf{J}_n$ and the image of the virtual camera V_n, as described in [15]. For the search space, we select the region of the predicted position of the target object within the virtual camera. The position of the match is then transformed into global coordinates, in which the required change of orientation for the virtual camera is computed to keep the target in the centre of the view. In global coordinates, the change of position between frames can be described as

$$\mathbf{p}_{n,t+1|t}^{\mathrm{G}} = \mathbf{R}_{\tilde{\Omega}}\, \mathbf{p}_{n,t|t}^{\mathrm{G}}, \tag{9}$$

where $\mathbf{R}_{\tilde{\mathbf{\Omega}}}$ can be computed using eqs. (6) with the parameters for the skew symmetric matrix (7)

$$\tilde{\omega} = \mathbf{p}_{n,t+1|t}^{\mathrm{G}} \times \mathbf{p}_{n,t|t}^{\mathrm{G}} \tag{10}$$

and

$$\beta = \parallel \mathbf{p}_{n,t+1|t}^{\mathrm{G}} \times \mathbf{p}_{n,t|t}^{\mathrm{G}} \parallel_2. \tag{11}$$

This results in the updated orientation for each virtual camera as

$$\mathbf{R}_{\mathrm{C},t+1|t}^{\mathrm{V}_n} = \mathbf{R}_{\tilde{\mathbf{\Omega}}}\, \mathbf{R}_{\mathrm{C},t|t}^{\mathrm{V}_n}. \tag{12}$$

To take the change of perspective or environmental conditions into account, we update $\mathbf{J}_n$, once the matching quality has dropped below a threshold of 95%, which proved to be a resonable value. The new template is created at the predicted position, $\mathbf{p}_{n,t+1|t}^{\mathrm{V}_n}$, thus this method is robust to outlier measurements at the update time step.

4. Experiments

The camera system, consisting of an omnidirectional camera rigidly attached to the IMU and GPS receiver, was mounted at height of about 2.5m at the stern of a 6m boat. We captured sequences with full resolution omnidirectional video data at 30fps, inertial data sampled at 90Hz, and GPS data sampled at 1Hz. The goal of the experiments was to qualitatively evaluate the stability and robustness of the multi view target tracking using the camera system in a real-world scenario. In each sequence, we tracked the position of the recording platform using the GPS. We created a virtual camera out of the omnidirectional video for each of the target objects, as described in section 3, to keep the target objects in the centre of the view. Figure 3 shows selected frames of the raw omnidirectional video while the recording platform is subject to substantial ego-motion as well as two virtual cameras tracking a stationary and a moving target respectively.

For visualisation, we geo-register the recording platform's trajectory onto a satellite image. We then present the results of the tracker as the orientation of the virtual cameras as lines of bearing from the current position of the recording platform. For a static target object, the lines of bearing intersect at the position of the target object, while in case of a moving target, local intersections can aid in position estimation. As a reference, the position of the target objects (and their trajectory, in case of a moving target object) has been manually estimated as we do not have GPS-position data for the target objects.

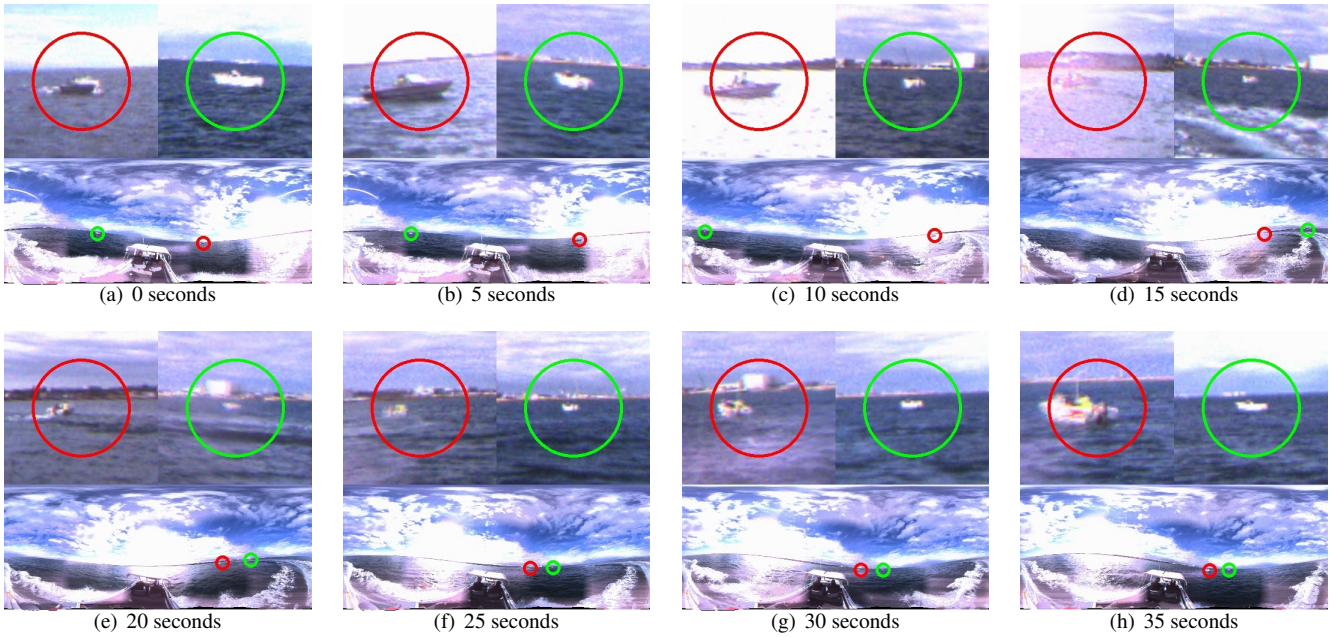
The following three experiments will be discussed in the following:

1. A 35 second recording in the bay with an average speed of $6.1ms^{-1}$ while keeping track of a stationary and a moving boat (Figure 4).

2. A 60 second recording in the marina with an average speed of $2.5ms^{-1}$ while tracking a moored up ship and a moving ship exiting the marina (Figure 5).

3. A 60 second recording near the port facility with an average speed of $6.4ms^{-1}$ while tracking two moored ships (Figure 6).

In the first recording, we selected a stationary and a moving boat as target objects. The recording platform was navigated between the boats and made a turn after passing the moving boat. The trajectory of the platform and the manually estimated positions of the two boats are shown in Figure 4. A frame of the raw omnidirectional image and the images of the virtual cameras foveating towards the target objects are shown in Figure 3. In the raw omnidirectional image, the distortions, caused by the ego-motion of the recording platform are clearly visible, while the virtual camera views are stabilised onto the targets. At each time instance, we compute bearings out of the orientation of the virtual cameras and plot it as a ray emitted by the recording platform.



(a) 0 seconds (b) 5 seconds (c) 10 seconds (d) 15 seconds

(e) 20 seconds (f) 25 seconds (g) 30 seconds (h) 35 seconds

Figure 3. Raw panoramic image (bottom) and stabilised virtual cameras tracking moving (left) and stationary boat (right) simultaneously.

Given that the virtual camera keeps the target object centred in the image, the point of intersection of the bearings is a measurement on the quality of the tracker. Figure 4(a) shows the bearings intersecting at the position of the stationary boat, and Figure 4(b) shows the bearings for the moving boat. Note that as the boat is moving, the rays do not intersect in a single point, but form a path of local intersections.

The recording platform was moving slowly in the second recording and due to the sheltered marina, only moderate disturbances were present. We selected a moored ship and a ship exiting the marina as target objects. The trajectory of the recording platform and the bearings towards both targets are shown in Figure 5. While the rays of bearing towards the moored ship intersect at the actual position of the target object, the intersection of rays towards the moving ship are much closer than the actual position of the ship, indicating that the ship is in fact moving.

In the third recording, two moored ships near a port site were tracked. Due to no speed restrictions in that area, the recording platform was navigated at a higher speed. Figure 6 shows the trajectory of the platform and the rays of bearings towards the target objects. Because both targets are stationary, the rays of bearing intersect at the actual positions of the ships, indicating that the orientations of the virtual cameras was computed correctly to keep the target objects in the centre of the view.

5. Conclusion

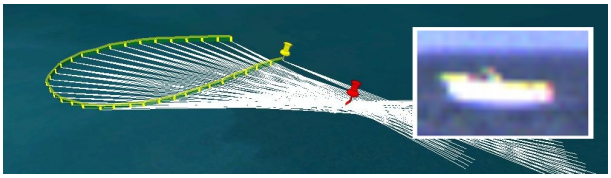
In this paper, we demonstrated a multiple view camera system consisting of an omnidirectional camera, an IMU, and a GPS for a maritime surveillance platform. We captured several data sequences in real-world maritime scenarios. In a qualitative evaluation, we have shown that the probabilistic integration of IMU and omnidirectional camera provides a robust multi target tracking system that is able to compensate for the erratic ego-motion of a maritime platform and to robustly track multiple maritime targets. Our future work includes an extended quantitative evaluation using multiple GPS equipped target objects and the deployment of the camera system on a dedicated mobile maritime surveillance platform.

Supplementary Material

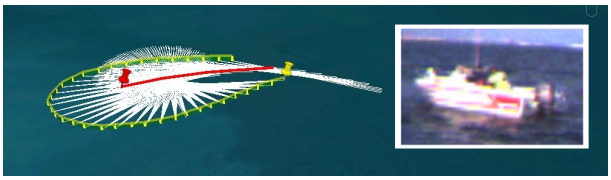
Visualisiation of the recorded footage in Google Earth and videos demonstrating the tracking are available online at `http://www.computing.edu.au/~14133369/multiviewtracking`.

Acknowledgment

Figures 4(a), 4(b), 5, and 6 created using Google Earth with image data ©2010 Google, DigitalGlobe, Cnes/Spot Image.
We specially thank Skipper Graeme Muller for taking us out on his boat. This research is supported by the DSTO, Australia PhD scholarship program.



(a) Trajectory of the platform (yellow) and bearings (white) towards stationary target based on orientation of the virtual camera with the image template of the target shown on the right.



(b) Trajectory of the platform (yellow) and bearings (white) towards moving target based on orientation of the virtual camera with the image template of the target shown on the right.

Figure 4. **Bay**. Trajectory of the platform and bearings towards the moving and stationary target objects. The position of the stationary target, 4(a), and the trajectory of the moving target, 4(b), have been registered manually. Note that 4(a) and 4(b) show the same frame and time instance, the visualisation has been broken into two figures for clarity.

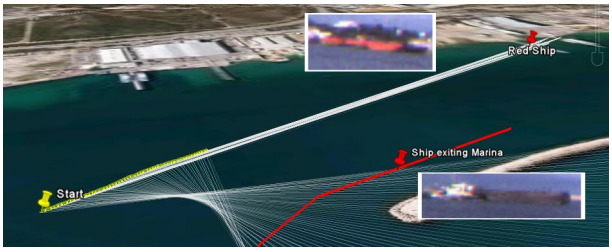


Figure 5. **Marina**. Trajectory of the platform (yellow path) and bearings (white lines) towards the stationary (red ship) and moving (black ship) target objects. The position of the stationary target and the trajectory of the moving target have been registered manually.

Figure 6. **Port Facility**. Trajectory of the platform (yellow path) and bearings (white lines) towards the stationary target objects. The positions of the stationary targets have been registered manually.

References

[1] Shree K. Nayar. Catadioptric omnidirectional camera. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 0:482, 1997.

[2] J.D. Hol, T.B. Schon, and F. Gustafsson. Modeling and Calibration of Inertial and Vision Sensors. *The International Journal of Robotics Research*, 29(2-3):231–244, 2010.

[3] T. Albrecht, T. Tan, G. A. W. West, and L. Thanh. Omnidirectional Video Stabilisation On a Virtual Camera Using Sensor Fusion. *International Conference on Control, Automation, Robotics and Vision*, 2010.

[4] A. Blake and A. Yuille. *Active vision*. MIT Press Cambridge, MA, USA, 1993.

[5] D.W. Murray, K.J. Bradshaw, P.F. McLauchlan, I.D. Reid, and P.M. Sharkey. Driving saccade to pursuit using image motion. *International Journal of Computer Vision*, 16(3):205–228, 1995.

[6] Pankaj Kumar, Anthony Dick, and Tan Soo Sheng. Real time target tracking with pan tilt zoom camera. *Digital Image Computing: Techniques and Applications*, 0:492–497, 2009.

[7] Faisal Bashir and Fatih Porikli. Collaborative tracking of objects in eptz cameras. *Visual Communications and Image Processing 2007*, 6508(1), 2007.

[8] J. Kang, I. Cohen, and G. Medioni. Continuous tracking within and across camera streams. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings*, volume 1, 2003.

[9] Xinding Sun, Jonathan Foote, Don Kimber, and B. S. Manjunath. Region of interest extraction and virtual camera control based on panoramic video capturing. *IEEE Transactions on Multimedia*, 7:981–990, 2005.

[10] K. Okuma, A. Taleghani, N. Freitas, J.J. Little, and D.G. Lowe. A boosted particle filter: Multitarget detection and tracking. *Computer Vision-ECCV 2004*, pages 28–39, 2004.

[11] C. Yang, R. Duraiswami, and L. Davis. Fast multiple object tracking via a hierarchical particle filter. In *Tenth IEEE International Conference on Computer Vision, 2005. ICCV 2005*, pages 212–219, 2005.

[12] J.P. Snyder. *Map projections–a working manual*. USGPO, 1987.

[13] Department of Defense. World Geodetic System. Technical report, National Imagery and Mapping Agency, 2000. Third Edition.

[14] A. Doucet, S. Godsill, and C. Andrieu. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and computing*, 10(3):197–208, 2000.

[15] D. Ballard and Brown C. *Computer Vision*. Prentice Hall, 1982.