Visual surveillance in maritime port facilities

Mikel D. Rodriguez Sullivan and Mubarak Shah

University Of Central Florida, Orlando FL

ABSTRACT

In this work we propose a method for securing port facilities which uses a set of video cameras to automatically detect various vessel classes moving within buffer zones and off-limit areas. Vessels are detected by an edge-enhanced spatio-temporal optimal trade-off maximum average correlation height filter which is capable of discriminating between vessel classes while allowing for intra-class variability. Vessel detections are cross-referenced with e-NOAD data in order to verify the vessel's access to the port. Our approach does not require foreground/background modeling in order to detect vessels, and therefore it is effective in the presence of the class of dynamic backgrounds, such as moving water, which are prevalent in port facilities. Furthermore, our approach is computationally efficient, thus rendering it more suitable for real-time port surveillance systems. We evaluate our method on a dataset collected from various port locations which contains a wide range of vessel classes.

Keywords: Maritime port security, computer vision, MACH filters

1. INTRODUCTION

Currently the maritime system in the United States includes more than 300 ports which hold more than 3,700 cargo and passenger terminals. Every year approximately 6,000 commercial ships (most of which are foreign owned and operated) make 60,000 U.S. port calls.[1] Port areas as well as ships docked in ports are vulnerable to numerous hazardous scenarios and are targets for terrorist attacks. Given the vast size of the perimeters of most port areas, manual inspection of all potential landside points of entry is infeasible.

Surveillance of port areas is further complicated by the fact that some ports are located immediately adjacent to dense urban areas. Additionally, ports provide numerous opportunities for potential attackers due to the large numbers of trucks that move in and out of the port area.[2] Many ports in the US also harbor small fishing and recreational boats which share access to a body of water.[3]

In addition to the risks associated with the port facility, commercial cargo ships are also at risk, given that they are often stationary at ports, and those moving through port do so at slow speeds, making them easy to intercept by a fast-moving boat.

Currently, a number of port facilities are equipped with video surveillance systems. However, most of these systems are used post-factum, as legal tools used to track possible transgressors or assist in reconstructing the chain of actions that lead to a particular incident. Additionally, a number of surveillance systems present in port facilities are controlled by human operators which are responsible for monitoring numerous video feeds simultaneously.

In this work we propose a computer vision system that aids the process of securing port facilities by using video cameras to automatically detect and classify various vessel classes as they approach buffer zones in the port. Vessel detections are time-stamped and compared with notice of arrivals received by the port. The system enables ports to obtain an early warning of unauthorized or unannounced vessels in the port area.

The organization of the paper is as follows: In the next section we present an overview of MACH filters, which form the base of our vessel detector module. In section 3 we present our vessel recognition approach and describe the early warning system. We describe the experiments that we performed and their corresponding results in section 4, and we conclude by discussing future work in section 5.

Further author information: (Send correspondence to Mikel D. Rodriguez Sullivan.)
E-mail: mikel-at-cs.ucf.edu , Telephone: 321 947 8132

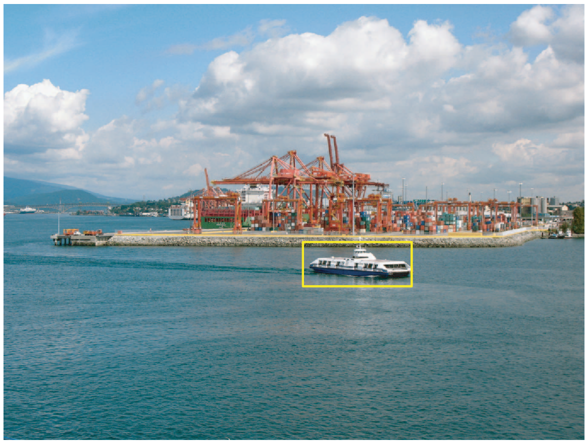


Figure 1. A typical view from the port shore. A vessel approaches the port dock and is detected by the system (yellow bounding box).

2. MACH FILTER

MACH filters have been employed in a range of application domains such as object classification, palm print identification,[4] and aided target recognition problems.[5–7] Given a series of instances of a class, a MACH filter combines the training images into a single composite template by optimizing four performance metrics: the Average Correlation Height (ACH), the Average Correlation Energy (ACE), the Average Similarity Measure (ASM), and the Output Noise Variance (ONV).

This procedure results in a two-dimensional template that may express the general shape or appearance of an object. Templates are then correlated with testing sequences in the frequency domain via a FFT transform, resulting in a surface in which the highest peak corresponds to the most likely location of the object in the frame. In this section we review the basic concepts which form the basis of the classification modules of the system.

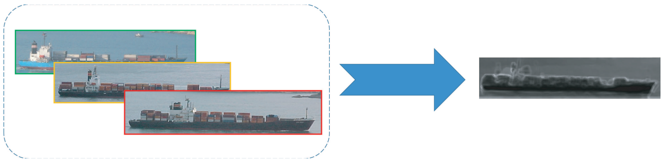


Figure 2. MACH filters combine a collection of training images (left) into a single composite template (right) by optimizing a set of metrics.

2.1 OT-MACH Filter

In our experiments we employ an optimal trade-off(OT) MACH filter,[8] a filter class which improves upon the traditional MACH filter. In the frequency domain a OT-MACH filter is given by:

$$h = \frac{m_x^*}{\alpha C + \beta D_x + \gamma S_x}, \tag{1}$$

where α, β and γ represent OT parameters which control the behavior of the filter. m_x is the average of the training image vector $(x_1, x_2, ...x_N)$ in the frequency domain. C represents the diagonal power spectral density matrix of additive input noise (which in our experiments is set to the white noise covariance matrix). D_x is the diagonal average power spectral density of the training images:

$$D_x = \frac{1}{N} \sum_{i=1}^{N} X_i^* X_i, \tag{2}$$

where X_i is the diagonal matrix of the i^{th} training image. S_x denotes the similarity matrix of the training images:

$$S_x = \frac{1}{N}\sum_{i=1}^{N}(X_i - M)^*(X_i - M_x), \tag{3}$$

where M_x is the average of X_i.

The different values of the OT parameters control the behavior of the MACH filter, when $\alpha = 0$ and $\gamma = 0$ the filter generally exhibits sharp peaks and good clutter suppression, but is more sensitive to target intra-class variability. However, when both $\alpha = 0$ and $\beta = 0$ the filter provides higher tolerance for intra-class variability but is less discriminative in general.

3. VISUAL SURVEILLANCE IN MARITIME PORT FACILITIES

3.1 Vessel Detector

In this subsection we describe the process of synthesizing MACH filters to recognize various vessel classes. In our experiments we focus on some of the most common vessel types, including: tanker ships, container ships, tugboats, speed boats, fishing boats, and cruise ships. A collection of typical examples of the set of vessels which we attempt to recognize is depicted in Figure 3.

The first step of our method is to detect various vessel classes in a video frame by matching appearance templates trained on different instances of the vessel classes. Templates are created for each vessel class using the OT-MACH filters as described in section 2.

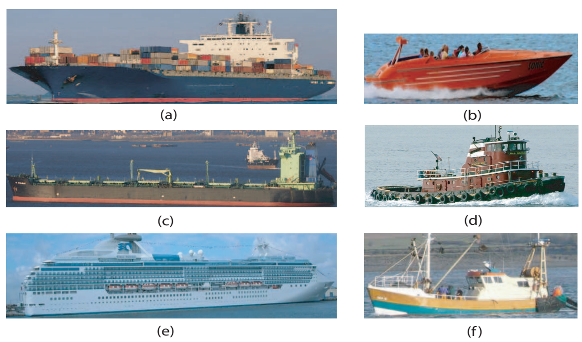


Figure 3. Our dataset and experiments focus on some of the most common vessel types, these include: container ships (a),speed boats (b), tanker ships(c), tugboats (d), cruise ships (e), and fishing boats (f).

Given the layout of most maritime ports, approaching vessels are for the most part oriented in approximately the same pose: heading towards the camera (Figure 6). We synthesize vessel templates using a collection of cropped instances of vessels as they approach the port, resulting in a template that expresses the general shape of a vessel class. We achieve color invariance by using edge-enhanced images instead of the original color frames. The OT-MACH filter generated for the container ship class is shown in Figure 2.

Having trained a set of mach filters, we proceed to apply each of the templates to incoming video frames by performing cross-correlation in the frequency domain. Correlation results are made robust to varying lighting conditions in the frame by normalizing the correlation output. The process of correlating the vessel template with a frame of the video results in a surface in which the highest peak corresponds to the most likely location of a vessel in the frame. When a vessel is visible in the scene, the correlation surface presents a sharp and distinct peak at that location in the frame (Figure 4). If there is no peak in the surface with a height greater than a threshold, then we determine that there is no vessel in the scene at that instance in time.

The most significant change in the appearance of vessels resides in the change of scale as they approach the port and the field of view of the cameras. In order to account for this change in appearance we resize the training images to an average size specific to a particular vessel class. Additionally, we perform detection with several re-scaled versions of the original template. Having computed the correlation at each scale, the scale that produces the maximum correlation peak is selected.

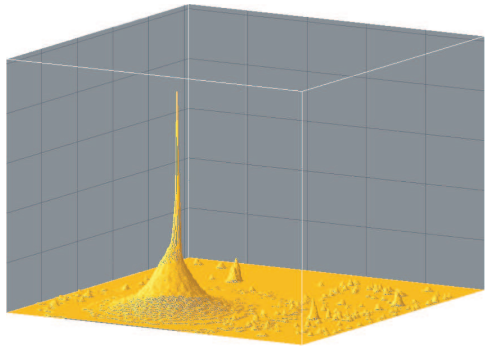


Figure 4. Correlating a vessel template with a frame of the video results in a surface as depicted in this figure. High peak values correspond to the most likely location of a vessel in the frame.

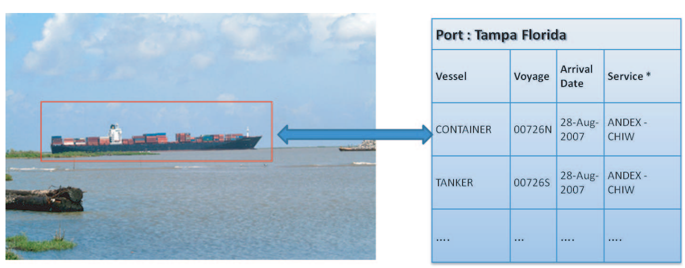
Once a specific vessel class instance is detected in the scene, a Kalman filter[9] tracker is initialized with the corresponding location in the image where the vessel was detected. Figure 9 depicts a series of frames and the corresponding vessel detections.

3.2 Port Access Verification

Currently the Master of an arriving foreign ship must submit a 96 hour pre-arrival notice,[10] which includes information about the ship's flag state administration, owner, operator, cargo type, classification society, last port of call, next port of call, U.S. port state control boarding history, etc.

This notice of arrival is received by a centralized database, the Ship Arrival Notification System (SANS) and its electronic component e-NOAD (electronic notice of arrival/departure). Information from e-NOAD is distributed to the arrival port's Captain of the Port's office for assessment. However, most ports are currently limited to manual inspection and cross-referencing of individual notice of arrivals (NOAs).

In our system, detections of vessels which are approaching the port are cross-referenced automatically with the e-NOAD database, thereby providing a means for early warning signs of unauthorized access to the port area. Each vessel detection is time-stamped, given the vessel type and time of arrival the e-NOAD database is queried for an entry matching the data for the port (Figure 5). A vessel entry in e-NOAD which matches the detected vessel class and whose announced arrival time is within a margin of the time-stamped detection is green-flagged, whereas a vessel detection with no corresponding e-NOAD entry is red-flagged and a red-bounding box is placed around the vessel (Figure 10).



Port : Tampa Florida

Vessel	Voyage	Arrival Date	Service *
CONTAINER	00726N	28-Aug-2007	ANDEX - CHIW
TANKER	00726S	28-Aug-2007	ANDEX - CHIW
....	...	....	....

(a) (b)

Figure 5. Each vessel detection is cross-referenced with entries in a e-NOAD database. Vessel class and arrival time fields are matched. Unannounced vessels are highlighted as potential port intruders.

4. EXPERIMENTS AND RESULTS

We tested the main components of our system by collecting a set of videos from the port of Tampa and the port of Saint Petersburg in Florida. Figure 6 depicts the camera layout for the Saint Petersburg experiments (a similar layout was implemented in the port of Tampa). Cameras were placed such that they were facing vessel entry and exit points, thereby capturing numerous instances of a wide range of vessel classes.

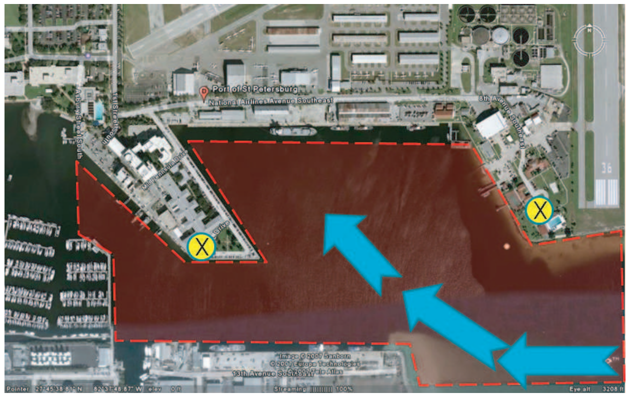


Figure 6. Aerial photograph of the port of Saint Petersburg, overlayed is the placement of the cameras in the port.

4.1 Vessel Classifier Tests

Our first round of experiments were aimed at evaluating the performance of the vessel classifier. For this purpose, video sequences containing various instances of vessel types were collected from numerous sources, including two maritime ports in Florida, as well as a collection of video sequences from the US coast guard [*] and stock footage data [†] . Video sequences were divided into training and testing sets using K-fold cross-validation framework[11] to partition the dataset into K sub-samples. From these sets we randomly pick a single sub-sample as the validation set for testing our approach, and the remaining $K-1$ sub-samples are used as training data to synthesize vessel templates. This process is repeated K times (5 folds in our tests), with each of the sub-samples used exactly once as the validation data. We report the mean of the results in a confusion matrix in Figure (Figure 7). We achieve an overall mean accuracy of 88.1%.

As can be appreciated in the confusion matrix, most of the misclassifications can be categorized into two clusters. One cluster corresponds to large vessels such as container ships, which at a distance can be confused with tanker ships or cruiser ships. A second cluster corresponds to smaller vessels such as speed boats and fishing boats.

A second round of experiments evaluated the effect of the OT parameters (α, β, and γ) on classification results. By changing the values of the OT parameters, the behavior of the MACH filter is altered to be more or less sensitive to intra-class variation and/or clutter. A plot of the mean detection rate and false alarm rate for each of the OT parameter combinations is depicted in Figure 8. The overall optimal combination of OT parameters was found to be $\alpha = 0.6$, $\beta = 0.5$, and $\gamma = 0.8$.

4.2 Port Access Verification Tests

A third set of experiments focused on testing the vessel access verification module. Based on the database schema provided in the official e-NOAD application development package, a set of tables with the appropriate schemas was implemented in a relational database management system.

In order to test the false alarm rate of the module, tables in the database were populated with data from the various testing video sequences (mainly vessel types and port entrance times). For each test iteration a set of tuples corresponding to different vessels in the database were masked-out, thereby effectively removing the vessels from the roster of authorized

[*] http://www.uscg.mil/
[†] www.bbcmotiongallery.com/
[‡] Found at http://www.nvmc.uscg.gov

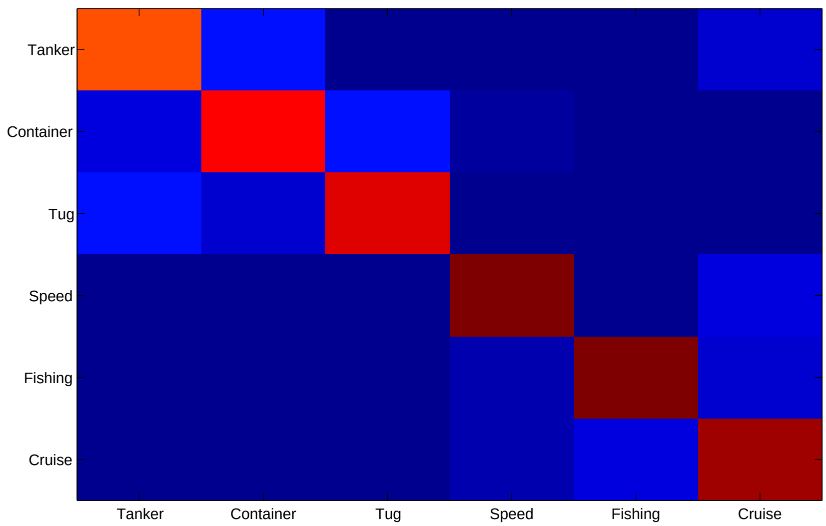


Figure 7. Confusion matrix for the vessel detection and classification module. Acuracy=88.1%

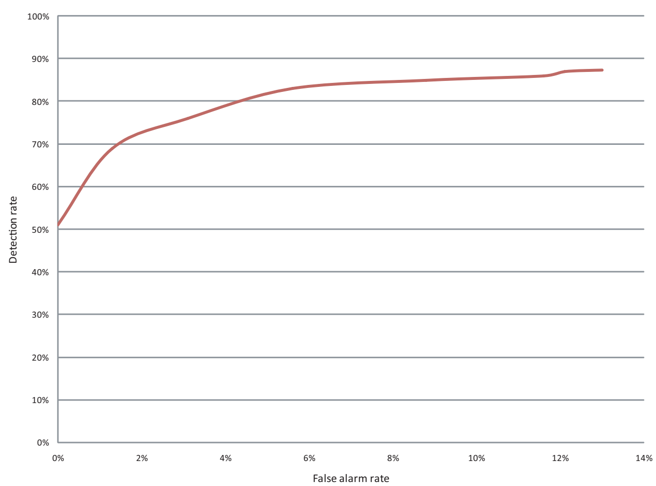


Figure 8. Receiver operating characteristic curve for the classification rate obtained by varying the OT parameters.

arrivals. The port access verification module is then expected to highlight only the corresponding set vessels which have been masked-out as potential intruders.

Additionally, in order to test the robustness of the approach, scheduled arrival times for the remaining tuples were modified by adding or subtracting a random number of minutes within the search window's range (+/- 15 minutes). This process was repeated 10 times, each time by randomly selecting a different set of vessels.

Based on this round of experiments, the mean false alarm rate was found to be 5.12%, whereas the mean false negative rate was 3.8%. Most of the false alarms and false negatives were due to incorrectly classified vessels.

5. CONCLUSION

We have proposed a vision-based system that can aid the process of securing port facilities by detecting and classifying various vessel types as they approach the port. Vessel detections are flagged as being potential intruders if no corresponding

Figure 9. A series of frames from a testing sequence. A tugboat is detected and tracked over a series of frames.

entries in an arrival notification database are found. The system is aimed at acting as an early warning source of information which can potentially mitigate unauthorized access incidents.

Our initial experiments and tests of the system suggest that a computer vision approach can be the basis of an effective means of port security and a more efficient use of existing surveillance camera resources.

REFERENCES

[1] Frittelli, J., [*Port and Maritime Security Background Issues for Congress*], Congressional Research Service, Library of Congress (2005).

[2] Harrald, J., Stephens, H., and vanDorp, J., "A Framework for Sustainable Port Security," *Journal of Homeland Security and Emergency Management* **1**(2) (2004).

[3] Barnes, P. and Oloruntoba, R., "Assurance of security in maritime supply chains: Conceptual issues of vulnerability and crisis management," *Journal of International Management* **11**(4), 519–540 (2005).

[4] Vijaya Kumar, B., Savvides, M., Xie, C., Venkataramani, K., Thornton, J., and Mahalanobis, A., "Biometric Verification with Correlation Filters," *Applied Optics* **43**(2), 391–402 (2004).

[5] Mahalanobis, A., Kumar, B., Song, S., Sims, S., and Epperson, J., "Unconstrained correlation filters," *Applied Optics* **33**(33), 3751–3759 (1994).

[6] Zhou, H. and Chao, T., "MACH Filter Synthesizing for Detecting Targets in Cluttered Environment for Gray-Scale Optical Correlator," (1999).

[7] Mahalanobis, A. and Kumar, B., "Optimality of the maximum average correlation height filter for detection of targets in noise," *Optical Engineering* **36**, 2642 (1997).

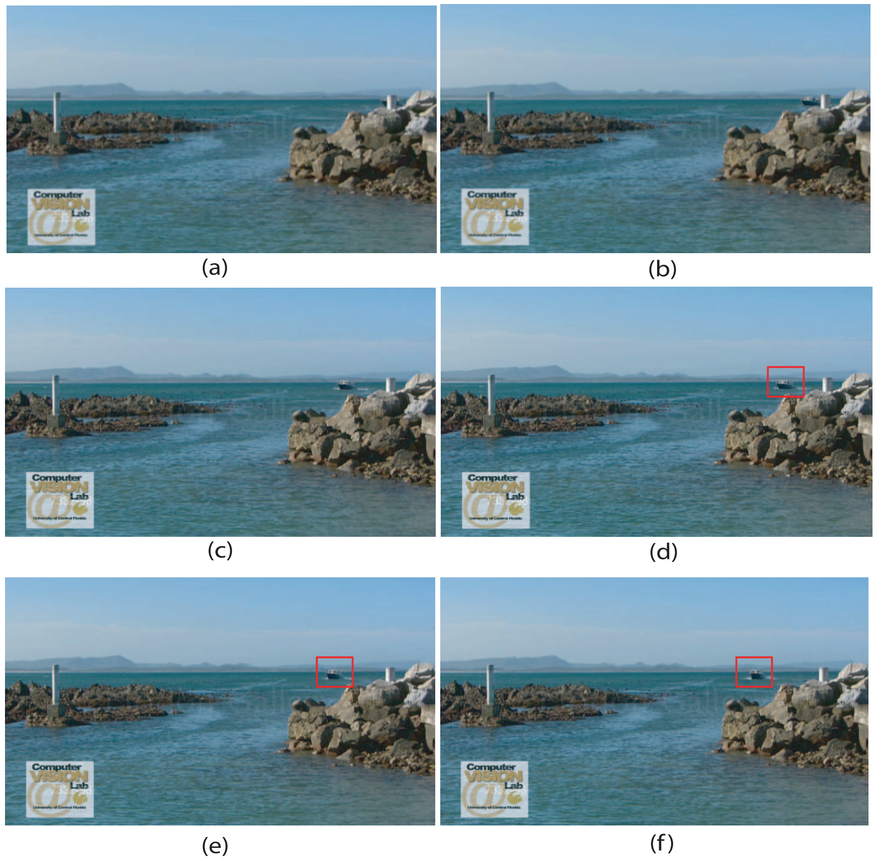


Figure 10. A series of frames depicting the detection of an unauthorized/unannounced vessel. In frames $a - c$ the vessel enters the field of view of the camera. In frame d a fishing boat is detected and the e-NOAD database is queried, upon retrieving no tuples matching the detected vessel class and arrival time a red bounding box is placed around the vessel.

[8] Refregier, P., "Optimal trade-off filters for noise robustness, sharpness of the correlation peak, and Horner efficiency," *Applied Optics* **16**(11), 829–831 (1991).

[9] Kalman, R., "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering* **82**(1), 35–45 (1960).

[10] Pruitt, T., "Maritime Homeland Security for Ports and Commercial Operations: The Case for Floating Port Security Barriers; Providing Necessary Waterside Protection for Marine Facilities and Vessels," *Sea Technology* **45**(11), 20–24 (2004).

[11] Kohavi, R., "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence* **2**, 1137–1145 (1995).