Adaptive Background Modeling for Land and Water Composition Scenes

Jing Zhao[1], Shaoning Pang[1], Bruce Hartill[2], and AbdolHossein Sarrafzadeh[1]

[1] Unitec Institute of Technology, Private Bag 92006, Auckland 1020, New Zealand
[2] National Institute of Water and Atmospheric Research, Private Bag 99940, Auckland 1149, New Zealand

Abstract. In the context of maritime boat ramps surveillance, this paper proposes an Adaptive Background Modeling method for Land and Water composition scenes (ABM-lw) to interpret the traffic of boats passing across boat ramps. We compute an adaptive learning rate to account for changes on land and water composition scenes, in which the portion of water changes over time due to tidal dynamics and other environmental influences. Experimental comparative tests and quantitative performance evaluations of real-world boat-flow monitoring traffic sequences demonstrate the benefits of the proposed algorithm.

Index terms— Background modeling, moving object detection, marine traffic, land and water composition scene, dynamic learning rate.

1 Introduction

Background modeling has been studied for traffic surveillance in a variety of situations including: motorways [1–3], road intersections [4–6], car parks [7–10], swimming pools [11, 12], and water channels [13, 14], etc. In general, we categorize different types of scenes into two groups: land scenes and water scenes, as background dynamics in these contexts differ markedly. On land, the background is usually static, with little or no change in topography. In contrast, water scenes are intrinsically dynamic, as water is a reflective surface that moves continuously, often to varying degrees. The reflection of the sun on water, coupled with the unpredictability of waves caused by the wind, moving vessels (wakes) and tidal flows in the maritime environment creates situations where background modeling is far more challenging.

In the context of maritime boat ramps surveillance, we consider background modeling for a dynamic maritime environment, at the interface between the land and the sea. Fig. 1 shows an example of maritime boat ramp which our case study is based on. As seen, the region-of-interest (ROI), identified as a polygon contains both areas of land and water. The boundary between the water and land changes over time with the rise and fall of the tide. As a consequence, the distribution of

water and land varies over time in the ROI, which makes background modeling in this case extremely difficult given the varying area of water and amount of light reflected from this water at differing times of day and sun angle.

In this paper, we propose an Adaptive Background Modeling for Land and Water composition scenes (ABM-lw), for performing real-time traffic surveillance at maritime boat ramps. The proposed ABM-lw dynamically classify areas of an image as either land or water, given ancillary tidal height data so that different strategies can be adopted to model backgrounds on land and on the water, respectively. The impact of sunrise and sunset is also specifically considered by proposed ABM-lw, to allow for changes in outdoor luminance. In particular, the use of dynamic learning rate and intelligent updating rules for areas of land and water, respectively, significantly increases the robustness of ABM-lw method. We apply the ABM-lw to real 24-hour boat-flow analysis and counting system and compare it against existing methods for background modeling, the empirical results show that the proposed ABM-lw achieves better performance.

2 Proposed Background Modeling Algorithm

To cope with this extremely challenging land and water composition scene, we classify areas of each image as either land or water, given ancillary model data on predicted tidal height, so that different strategies can be adopted to model the background on land and on the water, respectively. The proposed ABM-lw has four main components:

(1) Separation of areas of land and water: determining the optimized boundary between areas of land and water given ancillary tidal height model data.
(2) Background learning rate calculation: computing background learning rates for areas of land and water, respectively.
(3) Accounting for the influence of sunrise and sunset: adjusting current learning rates by applying the sunrise/sunset pattern.
(4) Background updating: computing background separately for land and water area, and updating the modeled background accordingly.

Before addressing the above main components, we introduce firstly the base model that we use to derive the proposed background modeling. Given $\{\boldsymbol{I}_t\}_{t=1\ldots T}$ as current set of images in observation, and polygon R as the Region Of Interest (ROI). We introduce the following signum function of image pixels as,

$$\boldsymbol{I}_t(i,j) = \begin{cases} \boldsymbol{I}_t(i,j) & \text{if } (i,j) \in R, \\ \text{-1} & \text{otherwise.} \end{cases} \tag{1}$$

By (1), we are able to cast any image (or the same size matrix) and related calculations into the ROI block.

According to [15], the next background $\boldsymbol{B}_{t+1}$ is generated using the weighted average of the instantaneous background $\boldsymbol{IB_t}$ and the current background $\boldsymbol{B}_t$:

$$\boldsymbol{B}_{t+1} = \alpha \boldsymbol{IB_t} + (1-\alpha)\boldsymbol{B}_t. \tag{2}$$

Here $\boldsymbol{B_1}$ is initialized as an image of the background without any moving objects present. α is the learning rate which determines the updating speed of background. In practice, α should be big enough so that background modeling algorithms can adapt rapidly to changes in the background, but small enough so that they are not sensitive to momentary changes. Because the background is influenced by changing luminance, different weather conditions, etc., it is essential for the background modeling algorithm to adopt a dynamic learning rate to optimize performance. We have therefore adopted a dynamic learning rate in the proposed method, rather than a static rate in [15].

The instantaneous background $\boldsymbol{IB_t}$ is generated from the current image but with detected objects removed, and their regions are filled with the current background. Specifically, for each incoming image $\boldsymbol{I}_t$, we calculate its differences to current background $\boldsymbol{B}_t$, and threshold the resulting difference image to obtain a binary object mask,

$$\boldsymbol{O}_t(x,y) = \begin{cases} 0 & \text{if } |\boldsymbol{I}_t(x,y) - \boldsymbol{B}_t(x,y)| < \lambda \ , \\ 1 & \text{otherwise} \end{cases} \tag{3}$$

where λ refers to the luminance threshold for object detection. The current instantaneous background $\boldsymbol{IB}_t$ can then be calculated as,

$$\boldsymbol{IB}_t(x,y) = \begin{cases} \mathrm{B}_t(x,y) & \text{if } \boldsymbol{O}_t(x,y) = 1, \\ \mathrm{I}_t(x,y) & \text{otherwise.} \end{cases} \tag{4}$$

Note that the above calculations are all in terms of gray-level intensity. In the case of color image, we simply transform the image to gray-level before any calculation.

2.1   Land and Water Area Separation

The goal of separation is to segment the ROI into areas of land and water. In the image coordinate system shown in Fig.1, the distinction between areas of land and water can be simplified as a geometric problem, which is to find/fix a straight line as,

$$n = km + b \tag{5}$$

where k and b refer to the slope and intercept of straight line, respectively. Thus to determine the shoreline, the task is to search for optimal values of k and b.

Physically, we look the sea as a large container, with the amount of water determined the position of tidal boundary. In this sense, for a specific ramp, we are able to determine the optimal b^* given tidal height data H provided by an ancillary model using interpolation methods such as linear interpolation, polynomial interpolation, or spline interpolation, etc. As a result, we have the revised shoreline function as,
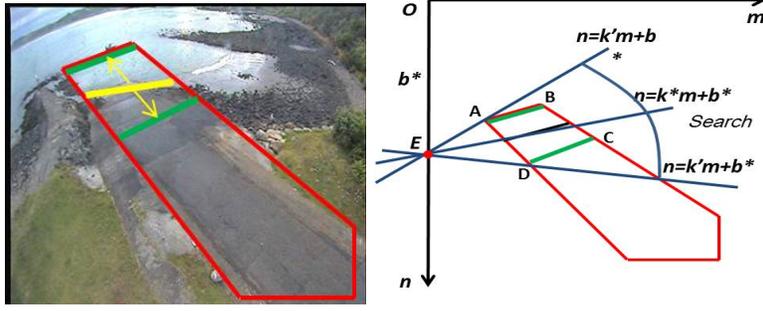
$$n = km + b^* \tag{6}$$

Fig. 1: Left: an example of maritime boat ramp. The region-of-interest (ROI) includes areas of land and water as seen inside the red polygon, and the boundary between water and land is shown by the yellow line. Right: an illustration of searching optimal boundary between water and land.

However, the slope k varies over time, as the direction of the boundary between the land and the water is not only determined by the shape of container, but also by the prevailing weather conditions such as the wind direction. Here, the proposed solution is to classify all ROI pixels into land and water area, then we seek the optimal slope k^* by a searching process described below.

Let $\boldsymbol{D}_t$ denote a land-water distribution matrix of current image $\boldsymbol{I}_t$, we can find the land water border line by accurately classifying every pixel as covering either land or water. $\boldsymbol{D}_t$ can be obtained by a binary pixel classification, which can be formulated as a convex optimization problem, i.e. the task of finding a minimizer of a convex function f that depends on a variable vector $\boldsymbol{\omega}$. Formally, we formulate this as an optimization problem, where the objective function is of the form

$$f(\boldsymbol{\omega}) = \frac{1}{2}\boldsymbol{\omega}^T\boldsymbol{\omega} + C\sum_{i=1}^{l}\max(1 - y_i\boldsymbol{\omega}^T\boldsymbol{x}_i, 0), \qquad (7)$$

Here the vectors $x_i \in \mathcal{R}^d$ are the training data examples, for $1 \leq i \leq r$, and $y_i \in [-1, 1]$ are their corresponding labels, which we want to predict. Consequently with the SVM trained, every pixel in $\mathcal{S}$ is classified as either land or water. It is not difficult to model a line $n = gm + l$ that gives a pixel level shoreline approximation regardless of tide change.

Consider shoreline approximation in Fig. 1, by (6) we have point E that the actual land/water boundary should have gone through, and its distance to O is b^*. Without loss of generality, we can define for every ramp in surveillance a maximum margin for all possible shorelines. In the example of Fig. 1, rectangle $ABCD$ is the margin area which we denote hereafter as $\mathcal{S}$. To find the optimal slope k^*, we rotate line (6) around E by trying every possible slope k' that directs the line going through margin $\mathcal{S}$. For each test line, we calculate its angle to the land-water border line comes from (7) for land and water pixel classification.

Thus, we have optimized slope k^* calculated as,

$$k^* \leftarrow \underset{k' \in \mathcal{S}}{\operatorname{argmin}}(\arctan \frac{|k' - g|}{1 + k'g}) \tag{8}$$

2.2 Background learning rate calculation

Under the condition of land and water area distinction, we are able to model land and water backgrounds separately by applying (2) to land and water image block as,

$$\begin{aligned} \boldsymbol{B}_{t+1}^l &= \alpha_l \boldsymbol{IB}_t + (1 - \alpha_l)\boldsymbol{B}_t, \\ \boldsymbol{B}_{t+1}^w &= \alpha_w \boldsymbol{IB}_t + (1 - \alpha_w)\boldsymbol{B}_t, \end{aligned} \tag{9}$$

where α_l and α_w represent the learning rate for land and water scenes respectively.

The proposed method adopts a dynamic learning rate for land and water area respectively rather than a static one in [15]. In practice, we maintain a 24-hour learning rate buffer, in which a pair of land and water learning rates (α_l, α_w) are stored at every minute. The process for computing learning rates is described below.
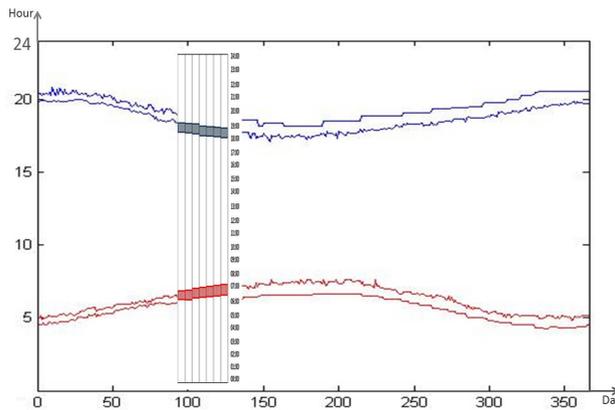

Fig. 2: Learning rate patterns for sunrise and sunset on different days

Consider in our case that, the objective is traffic analysis; more specifically, to count the number of boats/cars passing through the ramps. Thus, the number of objects is the ground truth of our background modeling. When determining the learning rate at time t, we search the optimal rates in the criterion of minimizing the error of objects counting as,

$$(\alpha_l^*, \alpha_w^*) \leftarrow \underset{\alpha_l, \alpha_w \in [0,1]}{\operatorname{argmin}} \left\{ \|(\psi_l - \eta_l) + (\psi_w - \eta_w)\|^2 \right\}, \tag{10}$$

where η_l and η_w are the predicted number and ψ_l and ψ_w are the actual number of objects in land and water area respectively. Here ψ_l and ψ_w can be obtained by manually viewing each frame. In practice, this is a time-consuming process. For simplicity, we implement (10) by counting the total number of objects using currently computed background regardless of water or land scene.

2.3 Accounting for the Influence of Sunrise and Sunset

In complex outdoor scenes, the level of luminance is easily influenced by several factors, such as time of day, cloud cover, time of year, available street lighting, etc. The rate of luminance change is obviously higher during sunrise and sunset than that of other time. It is therefore necessary to specifically consider the influence of sunrise and sunset as part of our background modeling.

Sunrise and sunset occur during two brief periods every 24-hours. Publicly available sunrise/sunset data provided by Land Information New Zealand (www.linz.govt.nz) gives an accurate estimate of when sunrise and sunset occur, which changes throughout the year. In practice, the influence of sunrise and sunset on the learning rate follows a specific pattern. We assume the pattern gradually changes throughout the year. Fig. 2 gives the learning rate patterns for sunrise and sunset on different days, where the top and bottom curves are for sunset and sunrise respectively.

Accordingly we refresh the learning rate for every minute during a 24-hour loop (i.e., learning rate buffer) as Fig. 2, given the shifting timing of sunrise and sunset during the year. For every incoming image, we search in the buffer suitable rates by time t, if t falls during the expected period of sunrise or sunset, the pattern learning rates are assigned to α_w and α_l.

2.4 Background Updating

Our strategy for updating background consists of two steps: firstly update land and water background respectively as,

$$
\begin{aligned}
\boldsymbol{B}_{t+1}^l &= \alpha_l^* \boldsymbol{IB}_t + (1 - \alpha_l^*)\boldsymbol{B}_t, \\
\boldsymbol{B}_{t+1}^w &= \alpha_w^* \boldsymbol{IB}_t + (1 - \alpha_w^*)\boldsymbol{B}_t.
\end{aligned}
\tag{11}
$$

Then merge obtained land and water backgrounds into the next background,

$$
\boldsymbol{B}_{(t+1)} = \boldsymbol{B}_{t+1}^w \cup \boldsymbol{B}_{t+1}^l.
\tag{12}
$$

Note that $\boldsymbol{B}_{(t+1)}$ here is an ROI image. To have the entire background image, we can simply merge $\boldsymbol{B}_{(t+1)}$ with the ROI residual image which can be obtained by applying a reversed function (1) on $\boldsymbol{I}_t$.

3   Experimental Results

New Zealand's National Institute of Water and Atmospheric Research (NIWA) has established a network of web cameras overlooking key boat ramps, on behalf

of the Ministry for Primary Industries, to monitor trends in recreational fishing effort over time. In this monitoring system, one image is captured per minute for each web camera, providing 1440 images of a monitored ramp on each day. These images are viewed in series by a technician who manually interprets these images and records a count of returning boats for that day. In our experiments, we used 2010-2012 image series captured at Waitangi, Takapuna and Manu Bay boat ramp. We compared backgrounds generated by the ABM-lw algorithm with those generated by the SABS method [15], which is initially designed for detecting vehicles in a terrestrial situation, such as on a highway, and which is proposed ABM-lw based on. The parameter settings for the SABS algorithm were made according to the authors' recommendations [15].

3.1 Robustness to Changes in Luminance at Sunrise and Sunset

In this comparative study, we demonstrate algorithm robustness to sunrise and sunset luminance changes. As we know, the periods of sunrise or sunset last for approximately 30 minutes each day, for which a corresponding 30 frame images are collected by each web camera system. Starting from the first frame, we select frames of minutes with an interval of six and observe algorithm performance in the whole procedure of luminance changes. Fig. 3 gives the comparison matrix, in which the first column shows input frames, and the second and third column show the backgrounds modelled using the SABS and proposed ABM-lw algorithms, respectively.

As would be expected, levels of illumination increased rapidly in the sunrise sequence, being darkest in T1 to brightest in T31; whereas the reverse occurred at sunset. The background models generated by both the ABM-lw and SABS approaches adapt to changes in levels of illumination at either end of the day, but the proposed ABM-lw approach performs much more rapidly and accurately than the SABS approach. The background model luminance generated by the ABM-lw approach closely matches that of the actual image at the time, but the background images generated by the SABS approach do not track the actual change in level of luminance seen in the first column. By T31 the background model generated by the proposed ABM-ls approach is very different from that generated by the less accurate SABS method.

3.2 Robustness to Changes in Tidal Height

Background modeling in coastal situations also needs to consider changes in tidal height. In this section we compare the performance of the proposed ABM-lw and existing SABS approaches at differing tidal states at Waitangi. The results of these comparisons are shown in Fig. 4, in which SABS background and the background from proposed ABM-lw are given in the middle and right column, and the ellipses in red color highlight the shoreline area of each predicted background.

Since the moving objects presented in previous images have not yet been completely forgotten, the ghost cars and boats are evident in the land water boundary area of the SABS backgrounds, and the shoreline is blurred. This

suggests that the SABS approach has difficulty with modeling backgrounds in areas where the movement of water is highly variable, whereas the dual area ABM-lw approach can readily account for this variability and generates more reliable background image for each frame. Results of moving objects detection provided by the ABM-lw approach should therefore be more accurate.
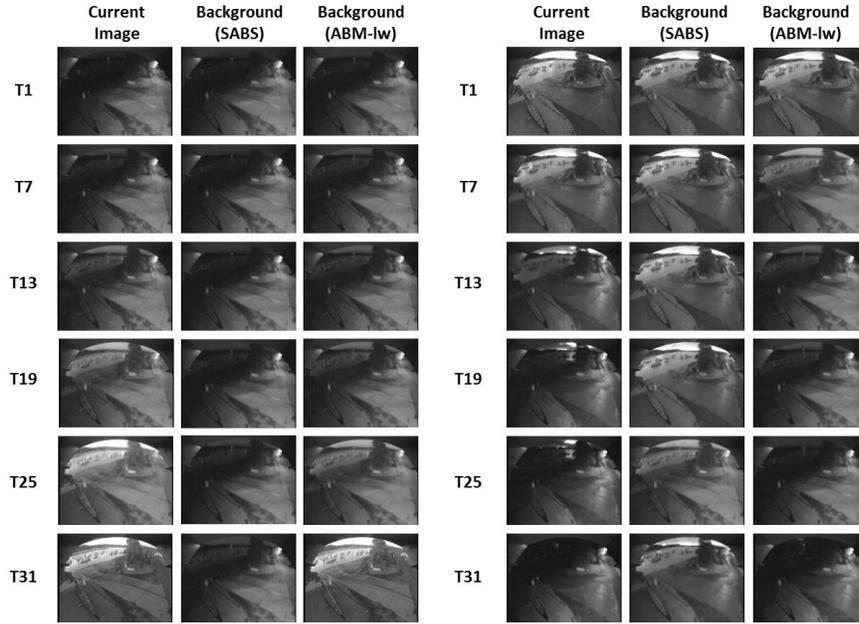

Fig. 3: Comparisons of SABS and proposed ABM-lw modelled backgrounds for a sunrise(left) and sunset(right) sequence of images.

3.3 Quantitative Performance Evaluation on Real Data

The overall relative performance of the SABS and ABM-lw background modeling approaches is also evaluated quantitatively. We use the alternative backgrounds generated by the two algorithms when detecting moving objects appearing at all three boat ramps, at different times of day, under differing weather conditions, and at different tidal states. For performance evaluation, we measure object detection accuracy as the ratio of the number of objects correctly detected by the algorithm, against the manual number of objects.

Table 1 gives the comparison results. As seen from the table, the proposed ABM-lw gives in general above 90% accuracy for all three boat ramps moving objects detection, whereas the accuracy of the SABS approach is in the range of 60% to 77%. The superiority of the ABM-lw approach relative to the SABS
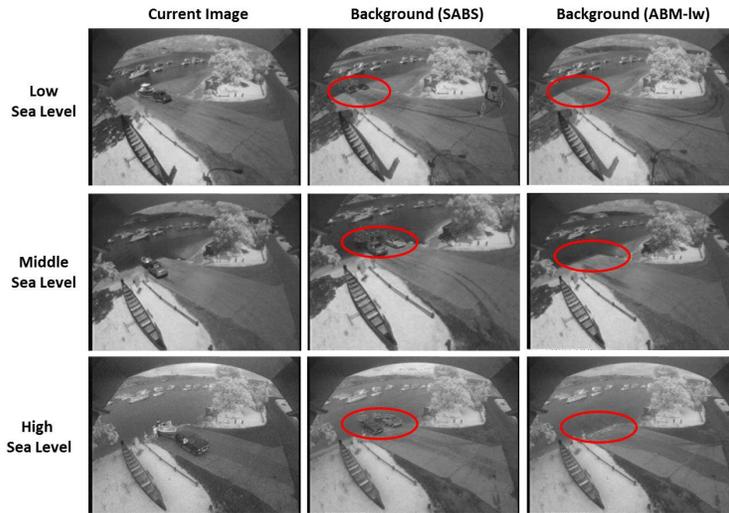
Fig. 4: Comparisons of SABS and proposed ABM-lw modelled backgrounds at different tidal states at Waitangi. (Left to right) current image, background from SABS, and background from proposed ABM-lw.

method is as high as 25%. Also, the average standard deviation of proposed ABM-lw approach is about 2%, which is three times lower than SABS's 6.5%. This suggests that, the ABM-lw algorithm generates a more stable background than the SABS algorithm in this context, under all conditions. Both algorithms perform worst at Takapuna which is a busier four lane ramp with a longer water line, and this causes that the background modeling for Takapuna is more difficult in practice.

4 Conclusions

In this paper, we propose a new background modeling algorithm (ABM-lw) intended to be used at maritime boat ramps where areas of both land and water are in frame. Background modeling in this context is especially challenging. The proposed ABM-lw approach attempts to classify each image into areas of land and water, and uses different strategies to model background of land and water scenes. Experimental tests and evaluations of its performance have been presented on a real 24-hour boat-flow analysis and counting system, where the proposed ABM-lw approach is compared with an existing method. These tests demonstrate that a much more stable background model is obtained by the ABM-lw algorithm in this context. The optimal learning rate α is mostly determined by the luminance threshold λ which is used in (3) for objects detection. Future work will therefore focus on how to estimate the optimal value for this luminance threshold.

Table 1: Object Detection Accuracy on Waitangi, Takapuna, and Manu Bay, respectively.

Waitangi		Detection Accuracy(%)		Takapuna		Detection Accuracy(%)	
		SABS	ABM-lw			SABS	ABM-lw
Time	Sunrise	65.48(±6.81)	**91.62(±2.31)**	Time	Sunrise	62.39(±5.83)	**90.37(±2.12)**
	Sunset	64.29(±7.12)	**91.67(±2.12)**		Sunset	61.46(±6.84)	**90.52(±2.33)**
	Daytime	75.01(±5.52)	**91.78(±2.52)**		Daytime	72.12(±7.02)	**90.73(±1.98)**
	Night	68.01(±6.96)	**91.69(±2.81)**		Night	65.34(±5.97)	**90.69(±2.03)**
Weather	Rainy	66.17(±6.37)	**91.58(±2.12)**	Weather	Rainy	63.34(±5.39)	**90.57(±2.26)**
	Foggy	63.19(±7.39)	**91.37(±2.91)**		Foggy	60.32(±6.23)	**90.21(±2.57)**
	Windy	65.01(±7.53)	**91.63(±2.13)**		Windy	62.34(±6.58)	**90.45(±2.25)**
	Sunny	74.82(±4.62)	**92.01(±1.95)**		Sunny	71.91(±5.67)	**90.96(±1.89)**
Tide	Low	66.48(±6.51)	**91.39(±2.13)**	Tide	Low	63.62(±6.72)	**90.22(±2.48)**
	Mid	71.29(±5.92)	**91.61(±2.01)**		Mid	68.42(±6.02)	**90.56(±2.36)**
	High	75.21(±4.34)	**91.92(±1.89)**		High	72.32(±3.39)	**90.91(±1.93)**

Manu Bay		Detection Accuracy(%)	
		SABS	ABM-lw
Time	Sunrise	67.56(±5.76)	**92.59(±2.01)**
	Sunset	66.44(±6.18)	**92.55(±2.23)**
	Daytime	77.23(±5.63)	**92.84(±1.78)**
	Night	70.75(±7.01)	**92.82(±1.92)**
Weather	Rainy	68.43(±5.65)	**92.66(±2.32)**
	Foggy	65.38(±7.02)	**92.25(±2.51)**
	Windy	67.32(±6.59)	**97.75(±2.33)**
	Sunny	76.98(±5.23)	**93.11(±1.83)**
Tide	Low	68.43(±7.03)	**92.28(±2.32)**
	Mid	73.45(±6.08)	**92.69(±2.17)**
	High	77.19(±5.33)	**93.08(±1.88)**

Acknowledgment

References

1. L. Unzueta, M. Nieto, A. Cortes, J. Barandiaran, O. Otaegui, and P. Sanchez, "Adaptive multicue background subtraction for robust vehicle counting and classification," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 527–540, 2012.

2. W. Wang, J. Yang, and W. Gao, "Modeling background and segmenting moving objects from compressed video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 5, pp. 670–681, 2008.

3. R. Cucchiara, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts, and shadows in video streams," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1337–1342, Oct 2003.

4. S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi, "Traffic monitoring and accident detection at intersections," *IEEE Trans. Intell. Transp Syst.*, vol. 1, no. 2, pp. 108–118, 2000.

5. A. Ottlik and H. H. Nagel, "Initialization of model-based vehicle tracking in video sequences of inner city intersections," *Int. J. Comput. Vis.*, vol. 80, no. 2, pp. 211–225, 2008.

6. H. Veeraraghavan, O. Masoud, and N. Papanikolopoulos, "Vision-based monitoring of intersections," in *Proc. IEEE 5th Int. Conf. Intell. Transp Syst.*, 2002, pp. 7–12.

7. K. Park, D. Lee, and Y. Park, "Video-based detection of street-parking violation," in *Int. Conf. Image Process. CVPR*, 2007.

8. S. Lin, Y. Chen, and S. Liu, "A vision-based parking lot management system," in *Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference*, 2006, pp. 2897–2902.

9. K. Choeychuen, "Automatic parking lot mapping for available parking space detection," in *Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference*, 2013, pp. 117–121.

10. ——, "Available car parking space detection from webcam by using adaptive mixing features," in *Computer Science and Software Engineering (JCSSE), 2012 International Joint Conference*, 2012, pp. 12–16.

11. H. Eng, J. Wang, A. Kam, and W. Yau, "Novel region-based modeling for human detection within highly dynamic aquatic environment," in *CVPR'04 Proceedings of the 2004 IEEE computer society conference on Computer vision and pattern recognition*, 2004, pp. 390–397.

12. P. Nuno, C. Nuno, C. Jorge, T. Adriano, and M. Jose, "A segmentation approach for object detection on highly dynamic aquatic environments," in *Industrial Electronics, 2009. IECON '09. 35th Annual Conference of IEEE*, 2009, pp. 1985–1989.

13. D. Bloisi and L. Iocchi, "Argos - a video surveillance system for boat traffic monitoring in venice," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 23, no. 07, pp. 1477–1500, 2009.

14. D. Bloisi, A. Pennisi, and L. Iocchi, "Background modeling in the maritime domain," *Machine Vision and Applications*, vol. 25, no. 5, pp. 1257–1269, 2014.

15. S. Gupte, O. Masoud, R. Martin, and N. Papanikolopoulos, "Detection and classification of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 1, pp. 37–47, March 2002.