Real-Time Detection of Small Surface Objects Using Weather Effects

Baojun Qi, Tao Wu, Hangen He, and Tingbo Hu

Institute of Automation, College of Mechatronics Engineering and Automation,
National University of Defense Technology, Changsha 410073, P. R. China

Abstract. Small surface objects, usually containing important information, are difficult to be identified under realistic atmospheric conditions because of weather degraded image features. This paper describes a novel algorithm to overcome the problem, using depth-aware analysis. Because objects-participating local patches always contain low intensities in at least one color channel, we detect suspicious small surface objects using the dark channel prior. Then, we estimate the approximate depth map of maritime scenes from a single image, based on the theory of perspective projection. Finally, using the estimated depth map and the atmospheric scattering model, we design spatial-variant thresholds to identify small surface objects from noisy backgrounds, without contrast enhancement. Experiments show that the proposed method has real-time implementation, and it can outperform the state-of-the-art algorithms on the detection of distant small surface objects with only a few pixels.

1 Introduction

Small surface objects always contain important information. Radar system, however, is ineffective against small objects close to the vessel [6]. In addition, for objects without hot parts present (e.g., buoys, surface rocks), near infrared images get unsuitable for detection either [16]. Therefore, detecting small surface objects from visual images is highly desired.

There has been growing interest in surface targets identification, as most standard computer vision algorithms for traffic scene are ineffective for maritime conditions [2,16]. To avoid collisions of surface vessels, Sanderson et al. have proposed some maritime targets identification algorithms [16,17], using statistical characteristics of the sea and motions of the targets. Sullivan et al. [18] use an optimal trade-off MACH filter to detect vessels from maritime surveillance videos. But the mach filter for various targets needs to be trained beforehand. Gupta et al. [6] provides an approach for maritime objects recognition through case-based statistical relational learning.

Algorithms, described above, require robust detection of image features. Under realistic atmospheric conditions, however, images of outdoor scenes are usually affected by scattering medium [11,12] (e.g., small aerosols, water-droplets). The affection, usually leading to contrast lost or color infidelity, increases exponentially with the distances of scene points from the sensor [13]. To eliminate

the depth-dependent degradation, many haze removal algorithms are provided to preprocess the hazy images [4, 7, 13, 19], which may lead to time-consuming detection algorithms. Thus, we intent to use the spatial-variant weather effects for surface objects detection without preprocessing hazy images.

In this paper, we describe an algorithm for small surface objects detection, using depth-aware analysis of image features by the atmospheric scattering model. Our purpose is giving early warning sources of information to avoid collisions or terrorist attacks. The main contributions of the paper are as follows. Firstly, we locate suspicious small surface objects with the dark channel prior, proposed by He et al. [7] from the statistics of hazy free outdoor images. The prior is based on the observation that objects-participating local patches have very low intensities in at least on color channel because of shadows, high-colored or dark objects. Secondly, we detect the horizon with dyadic cubic spline wavelet transforms [10], and estimate the scaled depth map of the sea surface based on the theory of perspective projection [5]. Lastly, using the atmospheric scattering model and the estimated depth map, we obtain spatial-variant thresholds and real-time detection of small surface objects, without hazy removal manipulations beforehand. Note that, our algorithm can distinguish objects from moving waves effectively, and it can outperform the state-of-the-art techniques [1,9,15] for detecting small objects with only a few pixels. In addition, when sequential information is used for the estimation of the depth map, the computational complexity for detecting objects from an $m \times n$ RGB image is $7mn$ times comparisons.

Our approach does have limitations. It becomes invalid for objects that are inherently similar to the sea surface and without any shadows casted on. However, we believe that the techniques used in this paper (e.g., designing depth-aware thresholds according to the weather effects, detecting small objects by local filters) can provide useful indications for other computer vision algorithms.

2   Weather Effects on Vision

In participating medium, light reflected from an object gets scattered by atmospheric particles that have significant size and concentration, leading to weather degraded images [19]. As the atmosphere on the sea always has high quantities of hygroscopic particles (e.g., sea salt), weather effects must be considered on surface objects detection under realistic weather conditions.

2.1   Atmospheric Scattering Model

In computer vision, the widely used model to describe scattering effects is:

$$E(x) = R(x)t(x) + E_\infty \left[1 - t(x)\right] \tag{1}$$

where x is the scene point, E is its observed intensity, R is its scene radiance, E_∞ is the horizon brightness, and t is its transmission describing the percent of

the light reaching the observer. For homogenous medium, the transmission can be expressed as:

$$t(x) = e^{-\beta d(x)} \tag{2}$$

Here, d is the depth of the scene point from the observer. β is the scattering coefficient, related to the meteorology visibility V as $\beta = \frac{3.912}{V}$ for homogeneous medium. Generally, the first term on the right hand side of Eq. (1) is called *attenuation*, and the second term is called *airlight*, denoted as A in the paper.

2.2   Degradation Analysis

According to Eq. (1), the observed contrast for two adjacent points i and j is:

$$\left| \frac{E_i - E_j}{E_i + E_j} \right| = \frac{|R_i - R_j|}{(R_i + R_j) + 2E_\infty (e^{\beta d} - 1)} \tag{3}$$

This means that the differences of foreground and background degrade exponentially with the depths of scene points. Thus, many haze removal algorithms have been proposed recently for robust detection of image features [4, 7, 12–14, 19].

Real-time vision systems always require single image based contrast enhancement techniques. As depth estimation from a single image is an ill-posed problem, many algorithms estimate the airlight based on some kind of prior first and then obtain the depth map [4, 7, 19]. The dark channel prior [7] indicates that the dark channel R^{dark} tends to be zero for most object-participating local patches,

$$R^{dark}(x) = \min_{x \in \Omega(x)} \left(\min_{c \in \{r,g,b\}} (R^c(x)) \right) \tag{4}$$

where $\Omega(x)$ is the local patch of x. If we assume the airlight of the points in a local patch to be identical, it can be estimated by $A = \min_{\Omega} \left(\min_c (E^c) \right) - t(x) R^{dark}$. In addition, the scaled depths of the objects-participating patches can be estimated according to Eq. (1) and Eq. (4) as:

$$\beta d = -\ln t = -\ln \left(1 - \min_{\Omega} \left(\min_c \frac{E^c}{E_\infty^c} \right) \right) \tag{5}$$

That is, we can estimate the degradation of hazy images and the 3D structure of the scene using the dark channel prior.

3   Detection Using Weather Effects

3.1   Dark Channel Prior for Surface Objects

The rationality of the dark channel prior has been verified by He et al. [7], based on the statistical analysis of large numbers of haze-free outdoor images. The prior has made great success in haze removal [7] as well as the 3D structure estimation from a single image [3,7]. However, the prior becomes invalid for maritime scenes

Fig. 1: Left: input image. Right: the dark channel image with 5×5 rectangular kernels.

which have large regions similar to the atmospheric light. Therefore, we directly detect surface objects without applying haze removal algorithms.

Figure 1 shows an example of the dark channel image of a sea surface image, obtained by local minimum filters with 5×5 rectangular kernels [8]. Note that, the dark channel image has the following characteristics. First, small surface objects with only a few pixels are enhanced and enlarged in the dark channel image. Second, for moving waves, only those near the sensor have low intensity, which may interfere with distant small objects detection. However, weather effects analysis can help us remove those waves as they are not dark enough if considering the additive airlight. Last, because the mountains are farther than surface objects, the dark channel of the mountains is with higher intensities according to Eq. (1). The observation can distinguish surface objects from the mountains far away.

In the rest of the paper, we first estimate the depth map of the sea surface. And then, we compute the scattering effects by the estimated depth map, followed by surface objects detection with spatial-variant thresholds.

3.2   Depth Map from a Single Image

Our purpose here is estimating each pixel's rough depth with respect to (w.r.t.) its row index. Based on the theory of the perspective projection, we have the following results.

Proposition 1. *Denote the pixel's row index as v. Assuming pixels in the same row with an identical depth d, then the depths of other pixels can be obtained by:*

$$d = d_0 \frac{v_0 - c}{v_0 - h} \cdot \frac{v - h}{v - c} \qquad (6)$$

here, c is the horizon, $v \neq c$, h is the height of the image, and d_0 is the depth of a special pixel with $v = v_0$.

Proof. According to the theory of the perspective projection, the homogeneous coordinate of a point $U_{img} = (u, v, 1)^T$ in image plane can be expressed with:

$$Z \cdot U_{img} = M_{3 \times 4} X_w \qquad (7)$$

where $X_w = (x_w, y_w, z_w)^T$ is the homogeneous coordinate in world, and $M_{3 \times 4}$ is the projection matrix. We only consider the depth map in the horizontal plane, and assume the pixels in the same row have the same depth. Then, $x_w = 0$ and $y_w = 0$ in Eq. (7). A scene point's depth w.r.t. its row index can be derived as:

$$v = \frac{a}{z_w + b} + c' \tag{8}$$

where $a = \frac{m_{24}}{m_{33}} - \frac{m_{23} m_{34}}{m_{33}^2}$, $b = \frac{m_{34}}{m_{33}}$, and $c' = \frac{m_{23}}{m_{33}}$. We assume the size of the image is $h \times w$. The parameters in Eq. (8) are estimated as follows.

(1) Let $z_w = \infty$, then $v = c'$. That is, c' is the vanishing line or the horizon of the maritime images, denoted as c in the following.
(2) Assuming the row index h having the minimum depth d_{min}, we obtain $h = \frac{a}{d_{min} + b} + c$. Thus, $a = (h - c)(d_{min} + b)$.

Assume that we have estimated the depth z_2 of some pixel, then the depths of other pixels can be derived from Eq. (8) as:

$$\frac{z_1}{z_2} = \frac{v_2 - c}{v_1 - c} \cdot \frac{a - b(v_1 - c)}{a - b(v_2 - c)} \tag{9}$$

Substituting estimated a into Eq. (9) and assuming $d_{min} \approx 0$, we obtain

$$\frac{z_1}{z_2} = \frac{v_2 - c}{v_1 - c} \cdot \frac{v_1 - h}{v_2 - h} \tag{10}$$

Denote the depth as d instead of z in Eq. (10), we finally get

$$d = d_0 \frac{v_0 - c}{v_0 - h} \cdot \frac{v - h}{v - c} \tag{11}$$

where v and d are the row index and the depth of a scene point respectively, c is the horizon, and d_0 is the depth of a reference point. Equation (6) indicates that once the horizon is estimated, the depth map of the maritime scenes can be computed from a reference pixel with known depth.   □

Corollary 1. *When misestimate of c occurs, images with smaller c have more robust estimations of depths.*

Proof. Assume the estimated horizon to be $c' = c + \varepsilon$. According to Eq. (11), the estimated error is:

$$\left| \frac{d'}{d} - 1 \right| = \left| \frac{\varepsilon(v_0 - v)}{(v - c - \varepsilon)(v_0 - c)} \right| \tag{12}$$

Note that, the points of the sea surface usually satisfy $v > c$ in the image plane. When the detection error ε of the horizon cannot be avoided, Equation (12) shows that larger $|v_0 - c|$ implies more robust estimation of d. As the reference point v_0 is on the sea surface, larger $|v_0 - c|$ indicates smaller c.   □
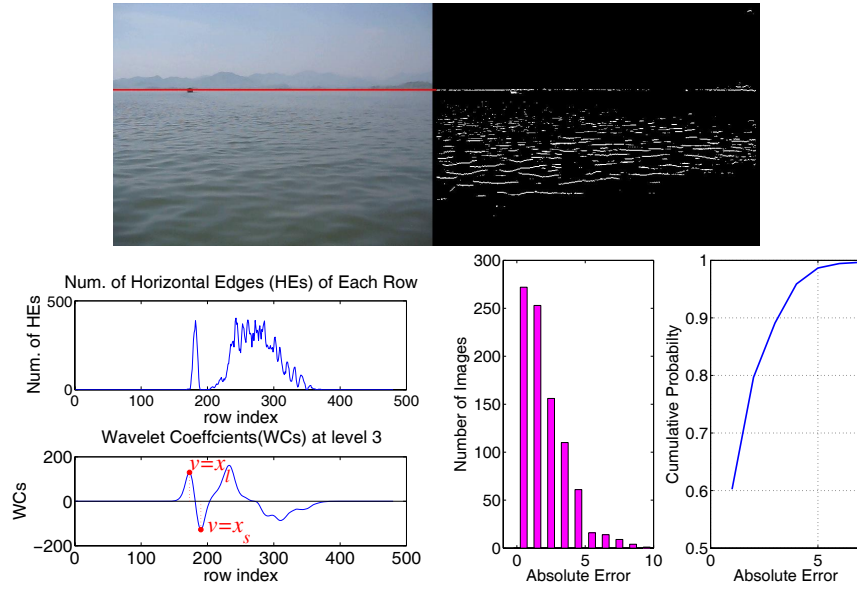
Fig. 2: Detection of the horizon. Topleft: detection results of the horizon. Topright: detection of the horizontal edges. Bottomleft: the number of detected horizontal edges for each row and its corresponding wavelet coefficients at level 3. Bottomright: histogram of the absolute detection error and its corresponding cumulative distributions.

3.3 Detecting the Horizon

Note that, estimating the depth map needs detection of the horizon. Additionally, surface objects are usually around the horizon due to the perspective projection. Therefore, this subsection mainly discuss how to detect the horizon.

To obtain robust detection, we apply the wavelet transforms based singularity analysis techniques. The wavelet we used is the orthogonal cubic dyadic spline wavelet transforms, whose coefficients for decomposition are given in Table 1. The advantage of the method is that it can detect various types of edges through multi-resolution analysis [10]. Additionally, the wavelet is translation invariant. If the low-pass filters and the high-pass filters for decomposition are $\{h_m\}$ and $\{g_m\}$ respectively, the wavelet coefficients at level j for input signal a^0 are:

$$d_n^j = (a^{j-1} * g^{j-1})_n = \sum_m g_m a_{n-2^{j-1}m}^{j-1} \qquad (13)$$

$$a_n^j = (a^{j-1} * h^{j-1})_n = \sum_m h_m a_{n-2^{j-1}m}^{j-1} \qquad (14)$$

For an input image, we detect the horizontal edges by labeling pixels with larger magnitudes of wavelet coefficients than some threshold. Then, we compute each row's number of edges and construct a 1D vector. Moving waves, shown as the top right of Fig.2, are hindrances to the horizon detection. However, we

Table 1: Coefficients of orthogonal cubic dyadic spline wavelet. h_n, g_n are the low-pass filters and the high-pass filters respectively for decomposition, where $h_{-n}=h_n$, $g_{-n}=-g_n$ for $n \leq 4$, and $h_n=0$, $g_n=0$ for $|n| > 4$.

n	0	1	2	3	4
$h_n/\sqrt{2}$	0.3750	0.2500	0.0625	0	0
$g_n/\sqrt{2}$	0	0.59261	0.10872	0.01643	0.00008

discover that there is a peak around the horizon and the horizontal edges of waves only existing nearby. Thus, we analyze the wavelet coefficients of the 1D signal at large scale space, e.g. Level 3. We detect points with the smallest wavelet coefficients (denoted as $v=x_s$), and points with the largest wavelet coefficients ($v=x_l$) between $v=1$ and $v=x_s$. The horizon is assumed to be $\frac{x_l+x_s}{2}$ (shown in Fig.2). For randomly selected 600 maritime images which have land on the horizons or have sea/sky horizons, we manually label the ground truth horizons. Our algorithm is tested on the data set and obtains a performance with 0.5 pixel mean error and 2.7 pixels stand deviation. Figure 2 shows the histogram of the absolute detection error and its corresponding cumulative distributions.

3.4   Spatial-Variant Thresholds

According to the subsection 3.2, the depth-map of the sea surface can be computed with the horizon and some reference point with known depth. In this subsection, we firstly discuss how to select the reference point and estimate its scaled depth. Then, we describe the method for surface objects detection from the dark channel image, using depth-variant thresholds which are computed from the estimated depth map in subsection 3.2 according to the scattering model.

Selecting the reference point . As discussed in Section 3.2, the depth of the reference point should be computable. Thus, the reference point must be selected from the objects-participating local patches which satisfy the dark channel prior. When multiple objects share the same row index, we choose the furthermost object's location as the reference point. The procedure is as follows.

(1) Compute the dark channel image dch for the input image img, the local minimum operators we used is $m \times n$ rectangular kernels;
(2) For dch, compute the histogram of the horizon centered region R_{int}, denoted as H_{dch}(Left of Fig. 3). Because objects-participating local patches tend to have pixels with lower intensities, we estimate the reference point's intensity A_0 as follows.
   - Search the first i satisfying $H_{dch}(i)<\varepsilon$ (e.g. $\varepsilon=1$)from $i=\arg\max_{j}\{H_{dch}(j)\}$ to $i=0$;
   - For some δ, look for the first ii satisfying $H_{dch}(ii)>\delta$ from $ii=i$ to $ii=0$, which is just the A_0 we needed.
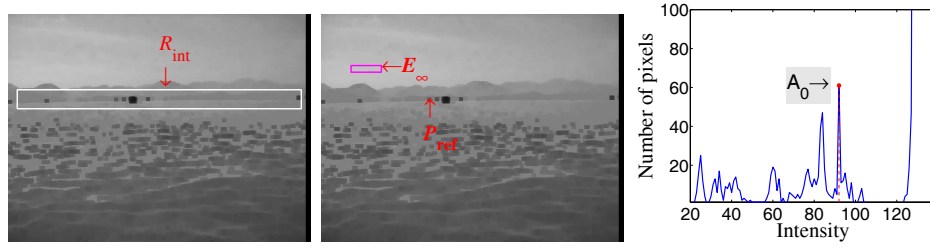
Fig. 3: Selecting the reference point P_{ref} from the dark channel image. Left: region of interest R_{int} for selecting P_{ref}, bounded by the white rectangle. Middle: Estimation of E_∞ and the local patch contains P_{ref}. Right: the estimated airlight of P_{ref} from the histogram of R_{int}.

(3) As only the row index and the depth of the reference point are concerned, we label the pixels with intensities A_0 and select the one with largest row index v_0 as the reference point. Considering the translation of v_0 caused by local minimum filters, we adjust v_0 to $v_0 \leftarrow v_0 - \frac{m}{2}$.

Then we can compute the depth of the reference point from Eq. (5) as:

$$d_0 = \frac{1}{\beta} \ln \left(\frac{E_\infty}{E_\infty - A_0} \right) \tag{15}$$

Where β is the scattering coefficient, and E_∞ is the horizontal brightness.

Estimating the airlight . Using Eq. (5), we can estimate the transmission of other pixels as:

$$\frac{\ln t}{\ln t_0} = \frac{d}{d_0} \implies t = t_0^{d/d_0} \tag{16}$$

where t_0 is the transmission of the reference point. Additionally, the airlight of other pixels can be estimated by:

$$\frac{A}{A_0} = \frac{E_\infty(1-t)}{E_\infty(1-t_0)} \implies A = \frac{A_0}{1-t_0} \left(1 - t_0^{d/d_0} \right) \tag{17}$$

Estimating E_∞ with the method described in [7] and substituting Eq. (10) into Eq. (17), we compute each pixel's airlight w.r.t. its row index v by:

$$A = \frac{A_0}{1-t_0} \left(1 - t_0^{\frac{(v_0-c)(v-h)}{(v_0-h)(v-c)}} \right) \tag{18}$$

where h is the height of the input image, and $t_0 = 1 - \frac{A_0}{E_\infty}$. Figure 3 shows the selection of the reference point and the estimation of E_∞.
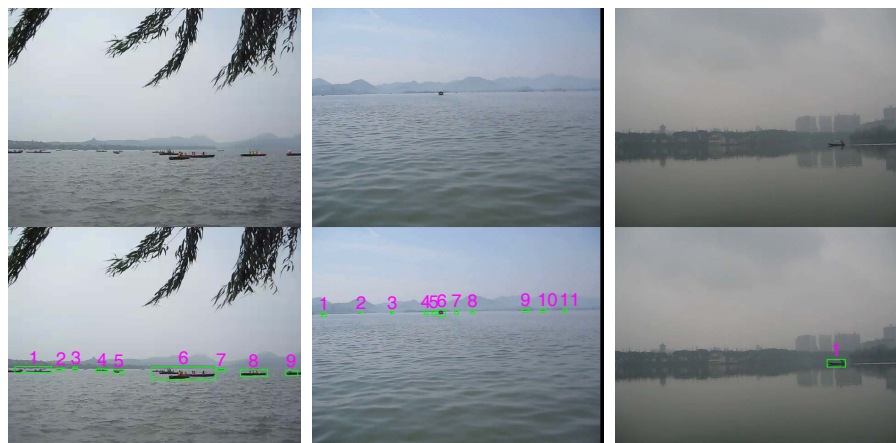
Fig. 4: Surface objects detection for surveillance videos of the lake scenes. Top: Original frames of different video clips. The first two images contain objects at various distances on a sunny day, while the last one is captured on a rainy day. Bottom: detection results of our method.

Detecting surface objects . According to the statistical report of He et al. [7], 90% of the pixels have intensities less than 25 for objects-participating local patches in the dark channel image. However, the intensities are often larger due to weather effects (e.g., small aerosols, haze). By quantitative description of the weather effects with the airlight A, we can design depth-variant thresholds by $A+\delta'$ ($\delta'>25$), and detect surface objects from the dark channel image by labeling pixels whose intensities are less than their corresponding thresholds.

4   Experimental Results

In our experiments, the dark channel image is computed using Marcel van Herk's fast local minimum operator [8] with 8×8 rectangular kernels. To obtain robust estimation of the airlight, we use $\hat{c}=c-\sigma$ ($\sigma=20$) instead of c in our experiments, according to the Corollary 1. However, the estimated airlight will shrink due to the adjustment of c. To overcome the problem, we design the thresholds with the estimated airlight plus a constant δ (e.g. $\delta=10$).

Figure 4 shows the results of our algorithm on the surveillance videos of the lake. As can been seen, our algorithm can achieve good performance for objects at various distances on both sunny days and rainy days. Especially, the method works for far objects with only a few pixels (shown in the middle of Fig. 4).

Furthermore, we test our algorithm on various maritime surveillance videos, such as objects with various depths, objects in low contrast regions, and objects in noisy backgrounds. Robust detections are shown in Fig. 5.

To evaluate the performance of our algorithm, we compare our algorithm with other detection methods, which are based on background modeling [20]
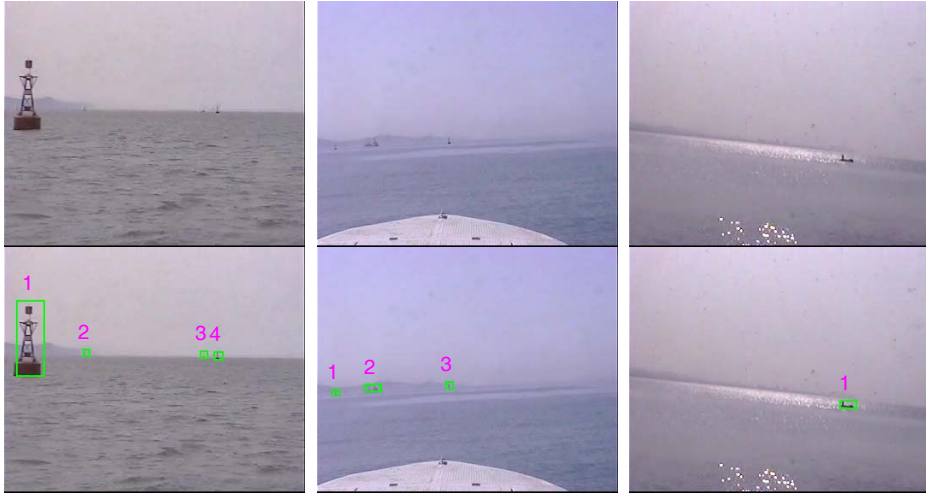
Fig. 5: Results on maritime surveillance videos. Top: original. Bottom: results.

and saliency analysis [1, 9, 15]. Different algorithms are tested on large numbers of images captured on the lake and the sea surface. Quantitative evaluation of the algorithms are analyzed by making the precision/recall (PR) curves, which are obtained by comparing the detection results with the manually labeled objects. The PR curves for different videos are shown in Fig. 6. As can been seen, both background modeling based method and saliency detection based methods have many false detections at large recall values. However, our algorithm make high detection precision for high recall levels (over 90%) as shown in Fig. 6. Thus, our algorithm can outperform the methods described in [1, 9, 15, 20] for small surface objects detection.

Computational Complexity Analysis. In our experiment, computing the dark channel image with Marcel van Herk's method needs $6mn$ times comparisons for an $m \times n$ image, independent of the size of the kernels. The detection of the horizon needs $8mm + 18m$ multiplications and $8mn + 14m$ additives. The estimation of the thresholds needs about $4m$ multiplications and $6m$ additives respectively. Therefore, the computational complexity of our algorithm is linear with the number of pixels. We implement our method in C++ and it takes about $30 \sim 40$ms to process a 352×288 image on a PC with 2.6GHZ Intel Pentium-V processor and 1G memory. The algorithm can be accelerated by using sequential information for the estimation of the horizon.

5   Discussions and Conclusions

This paper has described a novel algorithm for small surface objects detection using weather effects. We first compute the depth map of the sea surface from a single image according to the theory of the perspective projection. Then, using
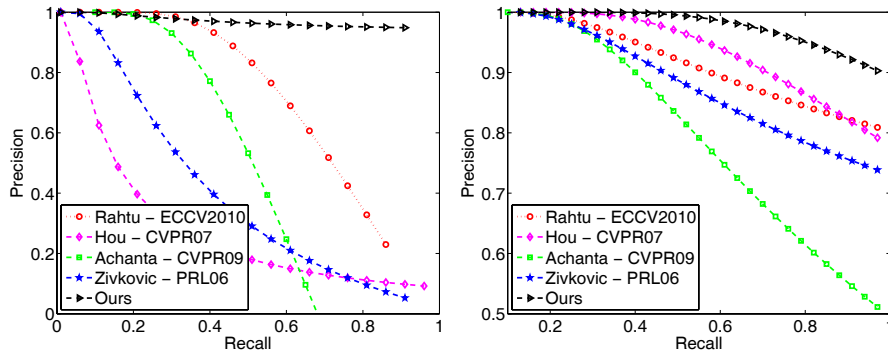
Fig. 6: Comparison of our algorithm with other methods for different videos. Left: results on the surveillance videos with multiple objects at various distances and moving waves presented. Right: results on the maritime surveillance videos with small objects near the horizon.

Fig. 7: Failure case. Left: input image. Right: the dark channel image.

the estimated depth map, we compute the spatial-variant airlight by the atmospheric scattering model, and design depth-aware thresholds for surface objects detection from the dark channel image. As shown in Fig. 6, our algorithm has more than 90% true positive rate for high recall values (recall > 0.9). Especially, the proposed method outperforms other algorithms described in [1, 9, 15, 20] for detecting faraway objects with only a few pixels.

However, for objects which are inherently similar to the atmospheric light and no shadow is casted on them, our algorithm cannot work because of the invalidation of the dark channel prior, which is just a kind of statistic. We intend to integrate other local descriptors of surface objects with the dark channel prior to improve the performance of our algorithm in the future.

References

1. Achanta, R., Hemami, S., Estrada, F., Süsstrunk, S.: Frequency-tuned salient region detection. In: CVPR. (2009)
2. Chan, A.B.: Beyond Dynamic Textures: a Family of Stochastic Dynamical Models for Video with Applications to Computer Vision. PhD thesis, University of California San Diego (2008)
3. Cozman, F., Krotkov, E.: Depth from scattering. In: CVPR. (1997) 801–806
4. Fattal, R.: Single image dehazing. In: SIGGRAPH. (2008) 1–9
5. Forsyth, D.A., Ponce, J.: Computer Vision: A Modern Approach. Prentice Hall (2003)
6. Gupta, K.M., Aha, D.W., Moore, P.: Case-based collective inference for maritime object classification. In: ICCBR. (2009) 434–449
7. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. In: CVPR. (2009) 1956 –1963
8. Herk, M.V.: A fast algorithm for local minimum and maximum filters on rectangular and octogonal kernels. Pattern Recognition Letters **13** (1992) 517–521
9. Hou, X., Zhang, L.: Saliency detection: a spectral residual approach. In: CVPR. (2007)
10. Mallat, S., Hwang, W.: Singularity detection and processing with wavelets. IEEE Transactions on Information Theory **38** (1992) 617–643
11. McCartney, E.J.: Optics of the Atmosphere–Scattering by Molecules and Particles. John Wiley and Sons Inc. (1976)
12. Narasimhan, S.G., Nayar, S.K.: Vision and the atmosphere. International Journal of Computer Vision **48** (2002) 233–254
13. Narasimhan, S.G., Nayar, S.K.: Contrast restoration of weather degraded images. IEEE Trans. Pattern Anal. Mach. Intell **25** (2003) 713–724
14. Nayar, S.K., Narasimhan, S.G.: Vision in bad weather. In: ICCV. (1999) 820–827
15. Rahtu, E., Kannala, J., Salo, M., Heikkilä, J.: Segmenting salient objects from images and videos. In: ECCV. (2010)
16. Sanderson, J., Teal, M., Ellis, T.: Target identification in complex maritime scenes. In: the Sixth International Conference on Image Processing and Its Applications. Volume 2. (1997) 463–467
17. Smith, A., Teal, M., Voles, P.: The statistical characterization of the sea for the segmentation of maritime images. In: the 4th EC-VIP-MC. Volume 2. (2003) 489–494
18. Sullivan, M.D.R., Shah, M.: Visual surveillance in maritime port facilities. In: Visual Information Processing XVII. (2008) 1–8
19. Tan, R.: Visibility in bad weather from a single image. In: CVPR. (2008) 1–8
20. Zivkovic, Z., van der Heijden, F.: Efficient adaptive density estimation per image pixel for the task of background subtraction. Pattern Recognition Letters **27** (2006) 773–780