Journal of Physics: Conference Series

A Ship Monitoring Framework Based on Multimodal Remote Sensing Data

To cite this article: Z Y Yin *et al* 2023 *J. Phys.: Conf. Ser.* **2486** 012018

A Ship Monitoring Framework Based on Multimodal Remote Sensing Data

Z Y Yin[1], Y Q Tang[1,*], Y Z Chen[1] and Y Y Zhang[1]

[1] School of Geosciences and Info-Physics, Central South University, Changsha, 41000, China

Email: yqtang@csu.edu.cn

Abstract. Due to the wide monitoring range, remote sensing satellites have more advantages than ground monitoring in large-scale monitoring. In particular, satellite network observations make rapid and frequent ground monitoring possible. In this paper, an all-day and all-weather marine ship monitoring framework based on multimodal remote sensing data was established. Scene recognition method was first used to segment sea areas. Then, we analyzed the ship characteristics of different data and used them for ship detection. Finally, the motion state of the ship was judged and the dynamic ships in the video were tracked. To prove the proposed framework, the data of Sentinel-1/2 and Jilin-1 data were used for verification. The experimental results demonstrated the advantages of the proposed framework for ship monitoring, which achieved the purpose of ship detection and tracking.

1. Introduction

Ship monitoring is critical for monitoring piracy, smuggling, military security, espionage, and more. Generally, ship monitoring systems mainly rely on ground radar and satellite positioning. However, these systems suffer from limited spatial coverage and the need to install a large number of sensors. Such as, Coastal Automatic Identification System (AIS) only covers an area 40 km from the coast and requires AIS which is not turned off [1]. Using remote sensing technology to automatically monitor ships plays an important role in marine safety, maritime transportation, and ship traffic monitoring. In the past decade, research on ship monitoring using high temporal and spatial resolution remote sensing imagery has gained attention. The joint application of multimodal data can give full play to the advantages of different types of sensors. Synthetic aperture radar (SAR) has all-day and all-weather observation capabilities. Multispectral satellites can acquire spectral information of targets. Video satellites have the ability to capture short-term dynamic information.

Sentinel satellites under European Space Agency (ESA) Copernicus program provide users freely available multispectral and SAR imagery down to 10 m resolution [2]. Ships only have a dozen to hundreds of pixels in the images. Compared with tens of millions of background pixels, ship detection is a small target detection problem. Lack of texture information will increase the difficulty of ship detection. Video satellites can obtain short-term 1m resolution ground dynamic information by pushing sweeping or gazing [3]. When the ship is in motion, the white wake generated by the propeller mixing water and air affects the tracking accuracy.

Establishing a large-scale, diverse, and reliable dataset can effectively test the generality and reliability of the algorithm. Huang et al. built a dataset called OpenSARShip for monitoring ships of Sentinel-1 [4]. The dataset provides 11 346 SAR ship objects with AIS messages. According to the different backscattering properties of sea and land pixels, the sea and land categories can be separated by threshold segmentation [5]. Then, ships were detected in the sea mask area. Grosso et al. detected ships by identifying the ship's wake of sentinel-1 by linear detection [6]. The direction and speed of the ship's movement can be estimated by identifying the ship's wake.

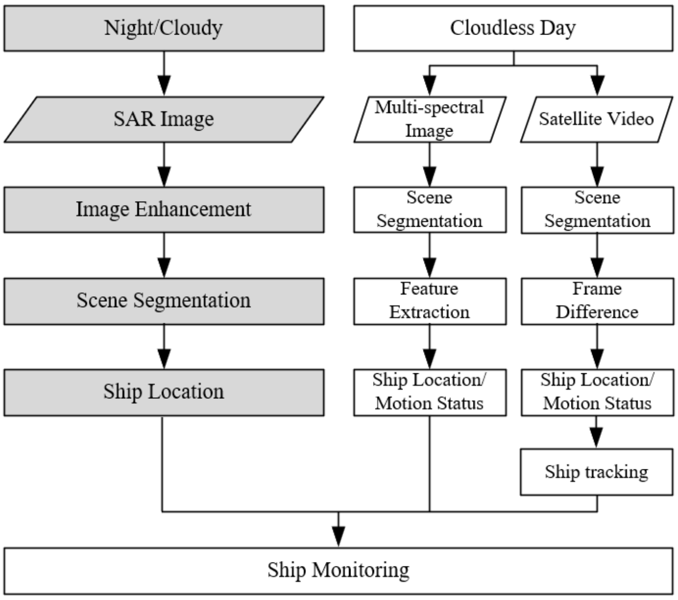


Figure 1. Flowchart of the proposed framework.

Although it is difficult for Sentinel-2 to obtain the texture information of ship, ship can be identified through important ship parameters, such as spectral information, contour features, ship wake, etc. Heiselberg distinguished ships, turbulent wakes, and the Kelvin waves [7]. And the proposed method could determine the ship's length and breadth. Deep learning algorithms have been widely used in remote sensing image processing. Ciocarlan et al. proposed a ship detection method in Setinel-2 images with self-supervised learning [8]. And it was based on a U-Net architecture with a ResNet-50 backbone to produce binary segmentation.

Recently, video satellites have received a lot of attention in the fields of traffic monitoring and military monitoring. Video satellites are a valuable addition to ground monitoring systems due to their wide monitoring range. Li et al. proposed a method to automatically detect and track moving ships of different sizes [9]. The algorithm detected ships of different sizes through motion compensation and dynamic multiscale saliency mapping (DMSM), and used important factors such as centroid distance, area ratio, and histogram distance to match and track ships in different frames.

In summary, ship monitoring using remote sensing imagery has been developed. But the problem of low image resolution and noise is still a challenge. In this paper, a new framework was proposed which ships are monitored using SAR, optical and satellite video data.

2. Methodology

In this section, the proposed framework is described in detail. The proposed ship monitoring framework includes three parts: scene segmentation, target detection and tracking. Firstly, the ocean area is divided according to the spectral or polarization characteristics of land and ocean. Then, the ship and wake are segmented according to statistics. Finally, the optimized multi-frame difference method is used on satellites video data for dynamic ship recognition and tracking. figure 1 illustrates the flowchart of the proposed framework. An all-weather automatic marine ship monitoring framework realized by the combination of SAR, multi-spectral, and satellite video data. In this framework, the network observation of different types of earth observation satellites is extremely important.

2.1. Sentinel-1

The Sentinel-1 images contain polarimetric SAR backscatter for horizontal (H) and vertical (V) polarizations. Because there is a lot of noise which is very similar to the ship in the sentinel1 image, it is necessary for image enhancement. In this paper, image enhancement is carried out by fusion of two polarized images.

$$I_{SAR} = VV + VH \tag{1}$$

where I_{SAR} represents the fused image of Sentinel-1. VV and VH represent the intensity information of the two polarization modes.

The image is segmented using an adaptive threshold. Morphological and connected area analysis was then used to remove noise. And sea masks was obtained. Ship detection in ocean mask areas can improve computational efficiency and reduce noise. Finally, ship detection based on intensity information.

2.2. Sentinel-2

Sentinel-2 has a total of 13 spectral bands with different resolutions. The proposed algorithm uses four 10m resolution bands for ship detection, which are B2 (Blue), B3 (Green), B4 (Red), and B8 (NIR), respectively. By comparing different water index algorithms, Deng et al. found that using Normalized Difference Water Index (NDWI) for Sentinel-2 images could achieve higher producer accuracy [10]. Therefore, this paper used NDWI to extract sea regions.

$$NDWI = \frac{Green - NIR}{Green + NIR} \tag{2}$$

Then, the spectral future of the ship is converted into redness (RN) [2]. Ships reflect more red and infrared light than sea water. And ship detection was based on an empirical threshold.

$$RN = \frac{I_{B4} + I_{B8}}{I_{B2} + I_{B3} + I_{B4} + I_{B8}} \tag{3}$$

Although water absorbs near-infrared, the ability to absorb near-infrared is reduced due to the mixing of air in ship wakes. According to the statistical analysis of the ship wake in B8 band image, the ship and wake can be extracted by threshold segmentation. However, RN could only detect the ships. The motion state of the ship could be judged by extracting the ratio of the length and width of the patch and the overlap of the two results. The position of the region extracted by the two methods for stationary ships is almost the same. The moving ship extracted by B8 band has a large ratio of length to width and the area is larger than the RN extraction result.

2.3. Satellite Video Data

Satellite video data have great advantages in large-scale dynamic monitoring. However, it only has single-band or visible light video. Jilin-1 video satellites can shoot 90-second video with the resolution of 1m. Low contrast and lighting variations increase the difficulty of monitoring ships with satellite video data.

As with sentinel 1/2 data, the ocean scene is firstly segmented. The brightness image of each frame was calculated as the maximum brightness of each pixel and morphological building index (MBI) was used to automatically extract buildings and bright objects [11].

$$I_{max} = \max_{1 \le k \le K} (I_k) \tag{4}$$

where I_{max} is the maximum brightness image, and I_k is the k_{th} spectral band. K is the number of spectral bands per frame in satellite video data.

Then, a green-blue-band difference vegetation index (GBVI) was used for vegetation detection to supplement MBI. MVI was established based on the spectral characteristics of vegetation in RGB color imagery.

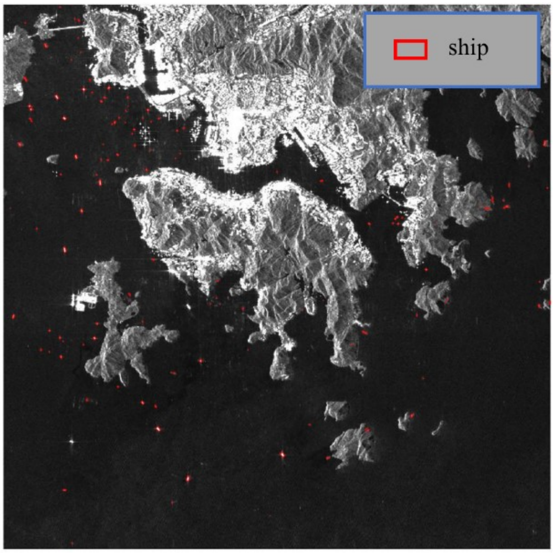


Figure 2. The ship monitoring results of Sentinel 1 image, which the red box indicates the position of the ship.

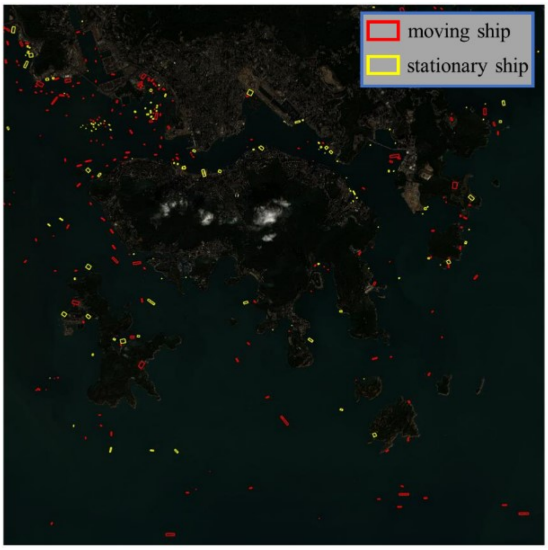


Figure 3. The ship monitoring results of Sentinel 2 image, which the red and yellow box indicates the position of the moving and stationary ship, respectively.

Table 1. The ship detection result.

Data	***Re***	***Pre***	***F1***
SAR	0.75	0.76	0.76
Multi-spectral	0.92	0.89	0.90
Satellite video	0.93	0.88	0.90

$$I_{GBVI} = \frac{G - B}{G + B} \quad (5)$$

where I_{GBVI} is the vegetation image. G and B are the green and blue band of satellite video data, respectively.

The sea area can be extracted by morphological processing of the fusion result of MBI and GBVI. Compared with the ocean background, the ship shows high brightness features on the image. The ships were extracted through the MBI algorithm. The motion state of the ship was judged by the inter-frame difference method. Finally, the detection result of the current frame and the previous frame are associated with the same target by the nearest neighbour search method, and the motion trajectory was obtained. This method could achieve the purpose of multi-target monitoring.

3. Experiment and Analysis

To verify the effectiveness of the proposed framework, the data of sentinel 1/2 and jilin1 were used for experiments. Sentinel 1/2 images monitor the area of more than 700 square kilometres. The study area included land, islands, and ocean scenes. Since it was challenging to coordinate satellite transit times, the multimodal data times chosen in this paper were not the same. The test results show the performance of the framework in overall ship detection, motion state judgment and tracking. The quantitative and qualitative results are shown in Table 1 and figure 2 to figure 6.

Some measures are used to quantitatively evaluate the results, including the true positives (TP), false positives (FP), false negatives (FN), precision (Pre), recall (Re) and F1-score ($F1$). If one pixel of the detected area overlaps with the real ship, the detected area is considered the true positive. Pre, Re and $F1$ are calculated as:

$$Re = \frac{TP}{TP + FN} \quad (6)$$

Figure 4. The ship detection results from satellite video data.

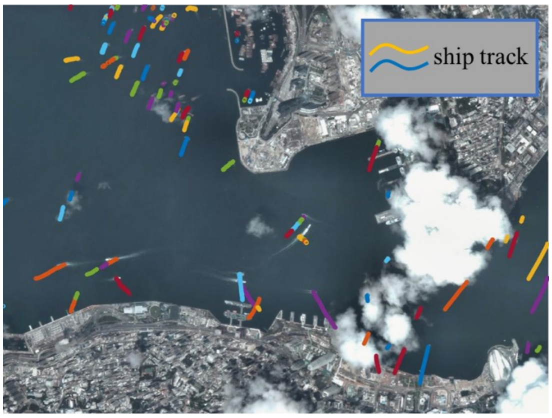
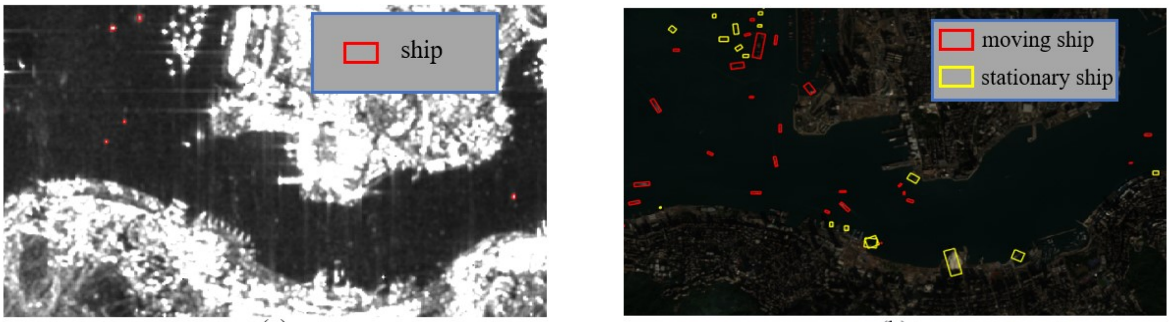


Figure 5. The ship tracking results from satellite video data.



(a) (b)

Figure 6. The zoomed in area of the Sentinel 1/2 image is similar to the range of the video satellite data.

$$Pre = \frac{TP}{TP + FP} \tag{7}$$

$$F1 = 2 * \frac{Pre * Re}{Pre + Re} \tag{8}$$

For quantitative evaluation, we manually checked all ship detection results to distinguish whether they were TP or FP. In order to accurately judge the ship, we counted the FN with a certain distance from the coast. Considering the influence of target size and image resolution, there might be slight errors between the statistical results and the actual situation. As shown in Table 1, most of the detected areas are ships, and the false detection is less than 0.25. Due to the high noise of SAR image, the detection effect is not as good as that of optical images.

The ship labelling result is illustrated in figure 2-figure 4. The local area of the Sentinel 1 data result is shown in figure 2, which the red box indicates the position of the ship. The local area of the Sentinel 2 data result is shown in figure 3, which the red and yellow box indicates the position of the moving and stationary ship, respectively. This paper used the Sentinel 1 image in the evening for verification. In Sar image, most ships were detected, especially large ships. Most of the detected ships are at a certain distance from the shore because these ships occupy dozens of pixels. Small ships are difficult to detect because of the effect of spatial resolution. Ships near the shore often have only a few to a dozen pixels and are easily affected by the scattering characteristics of land. On the other hand, the contrast between the intensity information of the boat and the ocean is low, which increases the difficulty of detection. However, the intensity information of ship wake waves is similar to that of sea water, so it is difficult to identify ship wake waves. It means that is difficult to judge the state of ship motion.

Compared with Sentinel 1/2's ship detection results perform better. Due to the strong absorption of water in the near-infrared band, ships are easily identified. Not only ships in the ocean are detected, but ships close to shore can also be identified. By the size of the ships, it can be found that large ships are farther from the shore than small ships, and the wake waves produced are also shorter. Unfortunately, some piers were wrongly detected because their spectral and shape characteristics were similar to ships. And some ships that are close together will also increase the difficulty of distinguishing. The ship's motion direction can be judged by the ship wake wave, and the ship's behaviour can be analyzed by clustering. Longer ship wakes mean faster motion. The wakes of small ships close to land were several times longer than small ships. The exact outline of the ship can be obtained by mathematical morphological processing. In addition, it can be classified by the pixel area, color, length, width, and other characteristics of the ship.

The ship monitoring results of Jilin-1 data are shown in figure 4 and figure 5, which figure 4 is the ship detection result and figure 5 is the ship track result. The spatial resolution of satellite video data is about 1m. And the coverage area of Jilin-1 video data is about 3km*4km. A large number of ships are detected by the algorithm proposed in this paper. 52 correct ship targets and 7 incorrect targets were detected. The brightness characteristics of the ship can easily distinguish itself from the ocean background. Due to the influence of clouds, false detections mainly originate from small areas of clouds. Ship matching through the minimum distance threshold can perform ship tracking at the same time of detection, and this method can realize multi-target trajectory extraction. However, the development time of video satellites is short, and the technology is immature, which leads to jitter in the obtained images. This is the main reason for the inaccurate judgment of the ship's motion state. Some short trajectories are false motions due to image drift.

The purpose of the framework proposed in this paper is to combine SAR, multispectral and video satellite data for all-day, all-weather ship monitoring. The temporal resolution of observations can be improved through the networked observations of different types of observation satellites. It is also beneficial to make full use of massive remote sensing data. In order to further compare the monitoring results of the three data, we highlighted an area which the Sentinel 1/2 image is similar to the range of the video satellite data, as shown in figure 6. It is not difficult to find that the detection results of SAR are not perfect in narrow water channels. Because the ships in this area are small and have low contrast, they are not obvious on the SAR image. On the other hand, buildings close to the coast generated strong scattered signal that interfered with ship detection. In the multispectral imagery and satellite video data, the wake of a dynamic ship can make the ship more visible.

4. Conclusion

In this paper, we performed a new framework for marine ship monitoring using multimodal remote sensing data, which includes SAR image, multi-spectral image, and satellite video data. Experimental results show that the proposed method is useful and effective. We provide a new idea for the application of ship monitoring and remote sensing data. The application of satellites in ocean monitoring has great prospects. In the future, we will focus on longer-term ship monitoring, which means being able to analyse ship behaviour. The illegal activity of ships will be identified by AIS, remote sensing data and the class of the ship. Furthermore, through long-term ship monitoring, the range of activities of illegal ships will be predicted.

5. References

[1] S. Brusch, S. Lehner, T. Fritz, M. Soccorsi, A. Soloviev, B. Schie, “Ship Surveillance With TerraSAR-X,” J. IEEE Transactions on Geoscience & Remote Sensing, 2011, 49(3): pp. 1092-1103.

[2] P. Heiselberg, H. Heiselberg, “Ship-Iceberg Discrimination in Sentinel-2 Multispectral Imagery by Supervised Classification,” J. Remote Sensing, 2017, 9(11):1156.

[3] Z. Yin and Y. Tang, "Analysis of Traffic Flow in Urban Area for Satellite Video," IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, 2020, Waikoloa, HI, USA, pp. 2898-2901.
[4] L. Huang, B. Liu, B. Li, W. Guo, W. Yu, Z. Zhang, W.Yu, "OpenSARShip: A Dataset Dedicated to Sentinel-1 Ship Interpretation," J. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 11(2): pp. 195-208.
[5] R. Pelich, M. Chini, R. Hostache, P. Matgen, C. López-Martínez, "Coastline Detection Based on Sentinel-1 Time Series for Ship- and Flood-Monitoring Applications," J. IEEE Geoscience and Remote Sensing Letters, 2021, 18(10): pp. 1771-1775.
[6] E. Grosso, R. Guida, "A New Automatic Ship Wake Detection for Sentinel-1 Imagery," IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, 2020, Waikoloa, HI, USA, pp. 1259-1262.
[7] H. Heiselberg, "A Direct and Fast Methodology for Ship Recognition in Sentinel-2 Multispectral Imagery," J. Remote Sensing, 2016, 8(12):1033.
[8] A. Ciocarlan, A. Stoian. "Ship Detection in Sentinel 2 Multi-Spectral Images with Self-Supervised Learning," Remote Sensing, 2021, 13(21):4255
[9] H. Li, L. Chen, F. Li, M. Huang, "Ship detection and tracking method for satellite video based on multiscale saliency and surrounding contrast analysis," J. Journal of Applied Remote Sensing, 2019, 13(2):026511.
[10] K. Deng, C. Ren, "Water extraction model of multispectral optical remote sensing image," J. Acta Geodaetica et Cartographica Sinica, 2021, 50(10): pp.1370-1379.
[11] X. Huang, L. Zhang, "A Multidirectional and Multiscale Morphological Index for Automatic Building Extraction from Multispectral GeoEye-1 Imagery," J. Photogrammetric Engineering & Remote Sensing, 2011, 77(7): pp.721-732.

Acknowledgment
The authors would like to thank the 2022 International Conference on Information Processing in Ocean Science and Technology (ICIPOST 2022) Organizing Committee. This work was supported by the National Natural Science Foundation of China under Grant 41971313.