Acta Oceanol. Sin., 2024, Vol. 43, No. 9, P. 136–144
https://doi.org/10.1007/s13131-024-2359-y
http://www.aosocean.com
E-mail: ocean2@hyxb.org.cn

Estimation of peak wave period from surface texture motion in videos

Haipeng Yu[1], Xiaoliang Chu[1]*, Guang Yuan[1]

[1] College of Physics and Optoelectronic Engineering, Ocean University of China, Qingdao 266100, China

Received 19 March 2024; accepted 24 June 2024

Abstract

Wave information retrieval from videos captured by a single camera has been increasingly applied in marine observation. However, when the camera observes ocean waves at low grazing angles, the accurate extraction of wave information from videos will be affected by the interference of the fine ripples on the sea surface. To solve this problem, this study develops a method for estimating peak wave periods from videos captured at low grazing angles. The method extracts the motion of the sea surface texture from the video and obtains the peak wave period via the spectral analysis. The calculation results captured from real-world videos are compared with those obtained from X-band radar inversion and tracking buoy movement, with maximum deviations of 8% and 14%, respectively. The analysis of the results shows that the peak wave period of the method has good stability. In addition, this paper uses a pinhole camera model to convert the displacement of the texture from pixel height to actual height and performs moving average filtering on the displacement of the texture, thus conducting a preliminary exploration of the inversion of significant wave height. This study helps to extend the application of sea surface videos.

Key words: low grazing angle, sea surface texture, video, peak wave period, significant wave height, image matching

Citation: Yu Haipeng, Chu Xiaoliang, Yuan Guang. 2024. Estimation of peak wave period from surface texture motion in videos. Acta Oceanologica Sinica, 43(9): 136–144, doi: 10.1007/s13131-024-2359-y

1 Introduction

Wave information is crucial for coastal and ocean engineering, marine energy and other fields (Villas Bôas et al., 2019; Ahn et al., 2020; Arefin et al., 2022; Blenkinsopp et al., 2022; Li et al., 2022). Specifically, peak wave period and significant wave height are two important parameters to be considered when designing ocean engineering structures (Ti et al., 2019; Malliouri et al., 2021; Afzal and Kumar, 2022). Wave parameters are usually obtained using buoys, where wave motion information is obtained by measuring the buoy motion (Ardhuin et al., 2019). The results of this *in situ* measurement technology are more accurate. However, the buoys cannot provide a wide range of wave information, and they are susceptible to damage by natural factors and human activities, which is costly. In recent years, remote sensing technology has developed significantly. This technology does not need to be immersed in the water, thus avoiding the interference of wave measurement and enriching the means of obtaining wave information. At present, a variety of remote sensing wave measurement techniques have been developed, such as satellite altimeter, Synthetic Aperture Radar (SAR) (Chaturvedi, 2019), high-frequency surface wave radar, X-band radar (Wu et al., 2022), and video-based wave analysis technology (Osorio et al., 2019; Guimarães et al., 2020).

Among these technologies, video-based wave analysis technology is noteworthy for its high spatial and temporal resolution, low cost, and ease of deployment, and it has been utilized to obtain wave parameters including peak wave period and significant wave height (Spencer et al., 2006). Furthermore, upgrading and networking existing sea surface video-based monitoring systems, initially deployed for various marine operations, could enable real-time sea wave monitoring. Such enhanced systems would offer expanded capability beyond the confines of single-point measurements provided by buoys and other instruments, thereby supplying substantial data for wave-related marine research.

Stereo imaging, a technique that utilizes the geometric relationships between the left and right cameras to reconstruct the ocean surface (Kim et al., 2022), has seen significant advancements in recent years as a prevalent method for wave measurement. A representative system called Wave Acquisition Stereo System (WASS) was developed based on the pyramidal pixel-based correlation method (Benetazzo, 2006), and its first open-source version was released in 2017 (Bergamasco et al., 2017). Unlike traditional methods that analyze a single stereo image pair, WASS reconstructs surface heights from a sequence of images, and it has been widely used in wave research (Alberello et al., 2022; Davison et al., 2022; Malila et al., 2022). With the development of this system, it was upgraded to WASSfast, which significantly improved the performance of 3D reconstruction of the sea surface and enabled near real-time optical measurements of ocean waves (Bergamasco et al., 2021). Based on this open-source software, Vieira et al. (2020) implemented wave observation through a stereo imaging system comprising two smartphones, simplifying the design and reducing costs. However, the stereo imaging technique requires at least two cameras and image matching, excluding a large amount of video data captured

Foundation item: The Key R&D Program of Shandong Province under contract No. 2023CXPT101.
*Corresponding author, E-mail: xlchu@ouc.edu.cn

by many single-video systems and thus limiting the video sources.

In contrast to stereo imaging, single-camera video captures images using a single sensor. This video analysis technique is suitable for a broader range of data sources and application scenarios and shows tremendous potential in video-based wave research. The Argus program (Holman and Stanley, 2007) and the Coastal Imaging Research Network group (Palmsten and Brodie, 2022) are dedicated to gaining a better fundamental understanding of the coastal environments derived from optical images. These research projects monitor the coast region to analyze beach morphology, estimate coastal depths, etc. In addition, other researchers have also carried out some work related to single-camera video. Osorio et al. (2019) proposed a method that employs time-stack images and signal processing techniques to estimate wave spectra parameters from beach videos and validate the peak wave period. Spencer et al. (Spencer and Shah, 2004; Spencer et al., 2006) determined the real-world scale and marine environment parameters using videos from an uncalibrated camera. They combined spatial and temporal frequency spectra extracted from these videos with knowledge of the physics of water waves to derive wave heights, wind speeds, and sea state. Almar et al. (2021) developed an independent regression model that derives sea state from drone and satellite imagery. In the mentioned coastal studies, wave speed decreases and wave height increases as waves approach the shore due to the shallowing of water depths. This change in water depth results in the wave direction being perpendicular to the shoreline while the wave crests become more prominent or even break, which provides a key feature for video-based wave analysis. However, wave propagation on flat-bottomed shores exhibits different characteristics than on shores with sloping bottoms. The crests of the waves will not be as prominent as they would be while propagating toward the beach. Additionally, the direction of wave propagation will neither be constant nor always perpendicular to the shoreline, necessitating a different strategy for studying waves. In the work of Spencer et al. (2006), images captured by a camera at an angle of 80° (10° below the horizon) were converted to a vertical view through a perspective transformation. This angle is almost the limit of observation. However, in specific video scenarios, such as harbor surveillance and fisheries management, the camera tends to capture the sea surface in a horizontal view. When observing at this low grazing angle (no more than 5° below the horizon), the texture in the image is mainly from waves propagating parallel to the line-of-sight direction, while waves propagating perpendicular to the line-of-sight direction are not apparent (Perugini et al., 2019). Moreover, the features in the images are primarily from high-frequency waves rather than from the dominant waves (Stilwell, 1969). This scenario poses some challenges in extracting wave parameters from the video, especially for waves propagating perpendicular to the line-of-sight direction, as they are difficult to distinguish.

In this article, a method of peak wave period estimation from sea surface videos taken at low grazing angle is proposed. The method utilizes the texture presented by the ripples of the sea surface in the image as a marker of wave motion. The texture motion in the video is analyzed using image matching technique. Subsequently, peak wave periods are determined through spectral analyses. This research contributes to expanding the regions and modalities of wave observation and provides a broader range of avenues for wave research.

The paper is structured as follows. Section 2 presents the method for recognizing wave motion using image texture and video data processing. Section 3 shows the computational results of applying this method to a field case. Section 4 discusses some factors affecting the experimental results. Section 5 concludes this paper.

2 Materials and methods

2.1 *Data and instrumentation*

The data used in this paper were obtained from a video system deployed in Zhoushan City, Zhejiang Province, China. The video system features a field of view of 20° × 10°, an image resolution of 2 560 pixel × 1 440 pixel, and a frame rate of 25 frames per second. Meanwhile, an X-band radar is deployed at the same site. The radar system possesses a range resolution of 5 m, a detection range of up to 3 000 m and an angular resolution of 0.1°. Data are collected at 10-min intervals, with each acquisition containing 64 frames and taking about 2 min. The distances necessary for the calculations are obtained from pre-calibrated X-band radar. The experimental results are evaluated using the average values of wave parameters, which are obtained through the inversion of radar data collected concurrently with the video.

2.2 *Algorithms*

Capillary waves on the ocean surface show discontinuous texture features in the images (Pan et al., 2020; Falcon and Mordant, 2022). This texture is particularly noticeable at low grazing angles. Capillary waves on the sea surface rise and fall with the waves and are modulated by the waves. The texture presented in the video image also reflects the wave motion, allowing the crests and troughs to be clearly distinguished in the video. Figure 1 gives a frame from the video to illustrate the relationship between texture variations in the video and wave propagation. Observing the footage shows that the wave propagates from right to left. The whitecap of the wave is visible in the figure as the crest of the wave. A region to the left of the crest (white box in Fig. 1) was selected as an observation sub-image to analyze the texture changes caused by the propagation of the wave to this region. To further illustrate the characteristics of the texture transformation over time here, the sub-images of this region at six moments selected from the video are given in Fig. 2. Among them, Figs 2a–c are the images before the wave reaches the region, while Figs 2d–f are the images after the wave passes through the region. It can be seen that the propagation of the wave causes a change in the texture distribution in the image. Selecting one of the dark stripes (the region surrounded by the dotted line in the figure) from

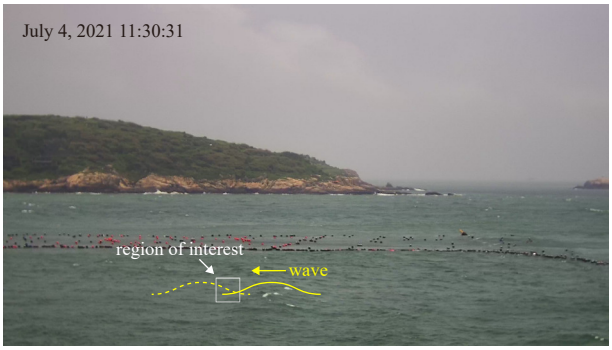

Fig. 1. Sea surface image captured by the video system. The yellow arrow in the figure points to the direction of wave propagation, and the white box indicates the sampling region used to identify the waves. Time in UCT+8.

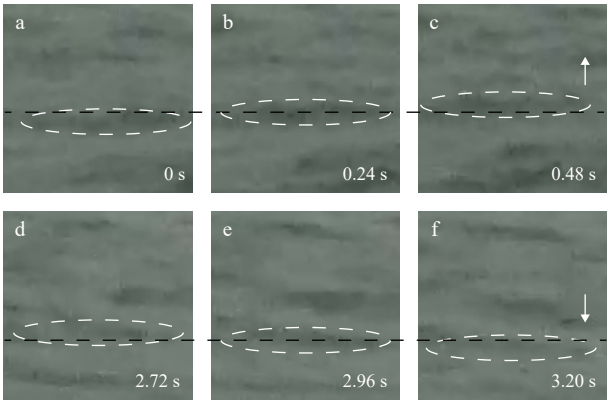

Fig. 2. Sub-images of six moments taken from the video. The white dashed ellipse indicates the position of the ripple texture in the image, and the black dashed line serves as a reference baseline for observing the movement of the texture. The textures in the image show upward (a, b, c) and downward (d, e, f) movement as the waves pass by.

Figs 2a–c, it is found that the position of the stripe in the image gradually moves upwards with the increase of time. Similarly, observing one of the dark stripes in Figs 2d–f, it is found that the stripe moves downwards with time. This change indicates that the texture in the video image moves up and down regularly with the propagation of waves, and its movement has the same oscillation period as the waves. Therefore, these textures can serve as markers of sea surface movement similar to buoys, allowing wave parameters to be extracted by analyzing their movement period.

In computer vision, the task of tracking objects within images has been extensively studied, and this approach is commonly known as optical flow (Horn and Schunck, 1981; Lowe, 2004; Bay et al., 2008; Goncalves et al., 2011). Observing the textures produced by the waves revealed that these textures existed in the observed area for a short period, characterized by the rapid disappearance of old textures and the continuous emergence of new ones. Not only that, the texture of the sea surface makes it often difficult to present a clear outline in the image. Given the features of wave texture, this paper utilizes a template matching approach to extract motion information from videos. This method does not track a particular texture but involves every sub-image pixel in the matching process. It can effectively track the motion of the texture, thus avoiding tracking failures due to texture disappearance. Through the application of a Fourier transform to the displacement of the texture, the peak wave period can be derived. The flow diagram is shown in Fig. 3.

The purpose of adjacent sub-image matching is to determine the position of the texture from the previous frame within the current frame. In this study, image matching is performed by exploiting the correlation between adjacent sub-images, a method that has been widely studied in the field of image processing

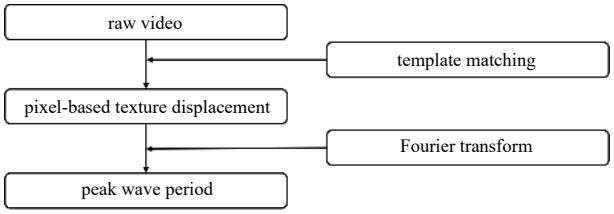

Fig. 3. Flow diagram of video analysis.

(Kim et al., 2019; Ye et al., 2019; Cui et al., 2022). The sub-image from the previous frame serves as the template image. Meanwhile, the sub-image from the current frame, which is slightly larger in size than the template image, serves as the target image. The image matching operation is implemented by the Open Computer Vision Library (OpenCV). It should be noted that in this study, the propagation of waves is considered the cause of vertical texture changes. Therefore, the deviation distance between the matched position and the center pixel in the vertical direction is taken as an indicator of texture movement. Figure 4 provides an example illustrating the acquisition of texture displacement. The template image and the target image are sized at 20 pixel × 20 pixel and 30 pixel × 30 pixel, respectively. The right panel of Fig. 4 illustrates the distribution of matching correlations across different positions, comprising an 11 × 11 grid of values. The red box denotes the center of the image, whereas the black box indicates the position of the maximum value obtained from the matching process, representing the best match position. Based on these, it can be concluded that the texture has shifted upwards by 1 pixel in the time interval between the two frames.

Each frame in the video was matched with the previous frame according to this method. The cumulative displacement of the texture for each frame was determined by accumulating the number of pixels shifted up to that frame (with upward movement denoted as positive and downward movement as negative), thus obtaining the texture displacement over time. This matching method, which only returns integer results, inherently introduces a loss of precision. To enhance accuracy, the image is enlarged using bicubic interpolation before matching, ultimately achieving sub-pixel resolution. Figure 5 shows the matching res-

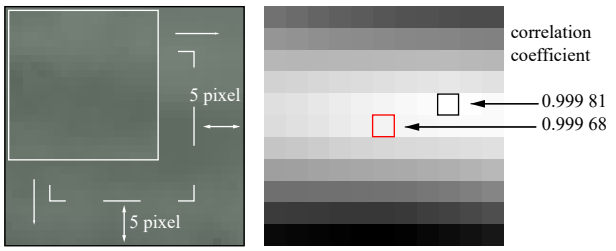

Fig. 4. Schematic of sub-image matching (left) and matching results (right). Lighter grey scale indicates larger values; the black box in the figure indicates the position of the maximum value obtained from the matching process; the red box denotes the center of the image.

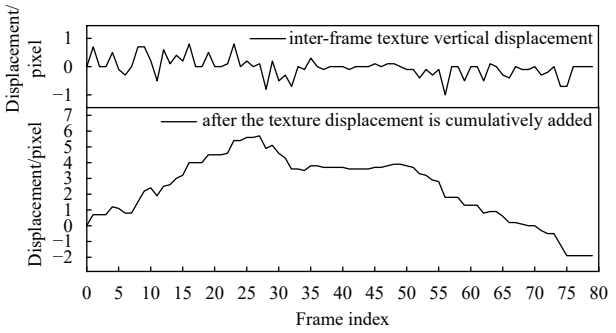

Fig. 5. Texture displacements were derived by matching between every two adjacent frames before and after the wave passes the observation position (upper curve), and continuous texture displacements were obtained by accumulating the displacement data (lower curve).

ults for 80 consecutive frames as a wave passes. During the matching process, the images were enlarged by a factor of 10, achieving a matching accuracy of 0.1 pixels. The time-varying texture displacement reflects the periodicity of the sea surface undulation, and by performing a Fourier transform on these displacement data, the wave spectrum is obtained, from which the wave period can be derived. To be able to obtain more accurate peak wave periods, the Wave Analysis for Fatigue and Oceanography (WAFO) toolkit is used in this paper (Brodtkorb et al., 2000).

3 Results

The first video was recorded on July 3, 2021, from 16:30 to 16:50 (UCT+8). The video spans 20 minutes, encompassing a total of 30 000 frames. Three regions labelled A, B, and C from the video were selected for extracting the wave data to verify the spatial stability of the method, as shown in Fig. 6.

Based on observations, the height of textures in the video ranged from approximately 10 pixel to 20 pixel. Therefore, in this study, sub-images of 20 pixel × 20 pixel were cropped to ensure that they encompassed at least one texture. The texture displacement obtained by applying the image matching method to Region A in the video is shown in Fig. 7. Subsequently, utilizing the WAFO toolbox, the wave spectrum based on these texture displacements was plotted, as shown in Fig. 8. Notably, Fig. 8 reveals a tendency for long-period offsets in the texture displacements, resulting in an unusually elevated wave spectrum in the low-frequency portion (below 0.8 rad/s). Osorio et al. (2019) faced a comparable noise challenge during pixel intensity processing. They tackled it by using cutoff frequencies to remove noisy signals. Consequently, disregarding the low-frequency part of the spectrum does not compromise the spectral analysis. Within the spectral region exceeding 0.800 rad/s, the peak is at 0.982 1 rad/s, corresponding to a period of 6.40 s (2π/0.982 1 = 6.40).

Fig. 6. Frame from the video showing three selected regions (A, B, and C) for wave parameter extraction.

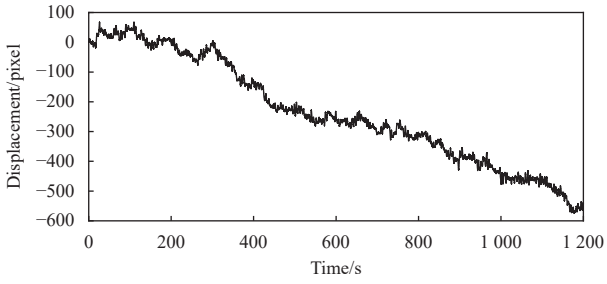

Fig. 7. Texture displacement in Region A.

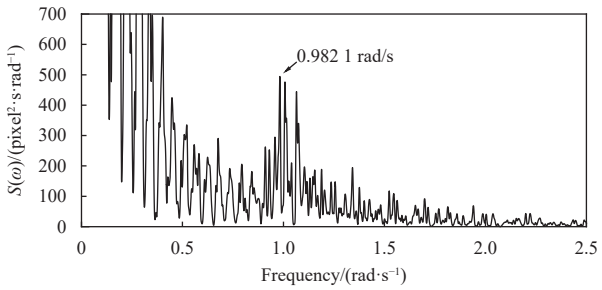

Fig. 8. Spectrum of texture displacement in Region A.

In order to assess the accuracy of the resulting peak wave period, this paper employs two methods for data validation. The first method involves using an X-band radar deployed in the field to invert the peak wave period from radar data. The peak wave period inverted by the X-band radar is 6.06 s, which differs by only 0.34 s from the peak wave period obtained through the image matching method. The second method uses the buoy movement to obtain the peak wave period. Kim (2005) and Kim and Cho (2005) have verified that buoy motion observed from video can reflect wave characteristics. In their study, the video-based results demonstrated an 8% deviation in wave period compared to measurements from a bottom pressure gauge (Kim et al., 2008). Figure 6 identifies a buoy in the vicinity of Region A, which will be utilized to estimate the peak period of waves. The buoy motion in the video is tracked using the Channel and Spatial Relatibility Tracking (CSRT) tracker, capturing its vertical movement data. The WAFO toolkit is employed to transform the buoy motion data into a wave spectrum, and the peak period is extracted from this spectrum, as shown in Fig. 9. The spectral peak is located at 1.103 rad/s, which differs from the spectral peak shown in Fig. 8 by 0.120 9 rad/s, corresponding to a difference of 0.7 s in the peak wave period as determined by the two methods. This deviation is not substantial, and as observed in Fig. 9, a higher numerical distribution is present within the range of 0.925 8 rad/s to 1.103 0 rad/s. This phenomenon is attributed to the complexity of nearshore waves, where various wave components superimpose to form this wave spectrum.

The calculated results for Regions A, B, and C labeled in Fig. 6 are shown in Table 1. It can be observed that the peak wave peri-

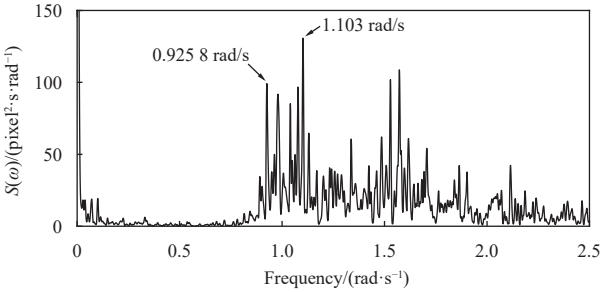

Fig. 9. Wave spectra calculated from buoy motion data.

Table 1. Peak wave periods T_A, T_B, and T_C corresponding to Regions A, B, and C in Fig. 6, respectively, along with the radar-derived peak wave period T_{radar} and the peak wave period T_{buoy} extracted from buoy motion, during 16:30 to 16:50 (UCT+8) on July 3, 2021

T_A/s	T_B/s	T_C/s	T_{radar}/s	T_{buoy}/s
6.40	6.52	5.82	6.06	5.70

ods derived from the three observation regions are relatively consistent, with a maximum deviation of 8% compared to the radar inversion results. When compared to the results derived from buoy motion, the maximum deviation is 14%. These findings indicate that this video analysis method has the capability to invert peak wave periods.

To further validate the video analysis method proposed in this study, additional video data from seven time intervals throughout July 3 and four time intervals from various dates were analyzed. The sub-images for analysis were selected from Region A in Fig. 6, a location closer to the buoy. The peak wave periods derived from image matching are shown in Table 2. Compared to the radar inversion results, the mean deviation is 3%, with a maximum deviation of 7%. Compared to the results derived from buoy motion, the mean deviation is 5%, with a maximum deviation of 11%.

Table 2. Comparison of peak wave periods derived from image matching method T_{image}, radar inversion T_{radar}, and buoy motion analysis T_{buoy}, and the deviations of T_{image} relative to T_{radar} (D_{radar}) and T_{buoy} (D_{buoy})

Date and time (UCT+8)	T_{image}/s	T_{radar}/s	D_{radar}	T_{buoy}/s	D_{buoy}
2021-07-03 08:30–08:50	6.04	6.46	6.5%	6.04	0.0%
2021-07-03 10:30–10:50	6.01	5.97	0.7%	6.02	0.2%
2021-07-03 11:30–11:50	5.72	5.96	4.0%	5.34	7.1%
2021-07-03 12:30–12:50	5.72	6.04	5.3%	6.43	11.0%
2021-07-03 13:30–13:50	6.18	6.20	0.3%	6.04	2.3%
2021-07-03 15:30–15:50	6.13	6.20	1.1%	6.04	1.5%
2021-07-03 18:30–18:50	6.21	5.91	5.1%	6.15	1.0%
2021-06-02 15:50–16:10	6.21	6.63	6.3%	7.01	11.4%
2021-06-09 09:50–10:10	6.02	6.12	1.6%	5.50	9.5%
2021-06-15 15:00–15:20	6.37	6.73	5.3%	6.26	1.8%
2021-07-04 17:00–17:20	6.39	6.35	0.6%	6.76	5.5%

4 Discussion

4.1 *Selection of observation region positions*

Significant differences exist between the near and far views in videos captured from a horizontal perspective. Given that the calculation of wave parameters intrinsically depends on the texture within the image serving as a pivotal feature, it is necessary to select a region where the texture is distinctly visible. This requires avoiding selecting regions so far away that the texture is indistinguishable and choosing relatively closer regions. It is not only because the texture is clearer at closer distances but also because pixels at different distances represent different sizes of the sea surface. For sub-images of the same size, farther away scenes correspond to larger regions of the sea surface, resulting in textures in the image that do not effectively serve as markers reflecting wave motion.

An assessment of this effect was conducted by selecting a distant region from the video recorded during 16:30 and 16:50 (UCT+8) on July 3, 2021, as shown in Fig. 10. The distances shown in this figure are calculated based on a perspective projection matrix, and the region depicted spans a range from 690 m to 776 m. Figure 11 illustrates the wave spectrum derived from the texture displacement extracted at this position. The spectral density in the figure decreases monotonically with increasing frequency, making it impossible to identify the position of the peak from the spectrum. Therefore, to obtain accurate data, it is advisable to select regions with clear textures that are closer.

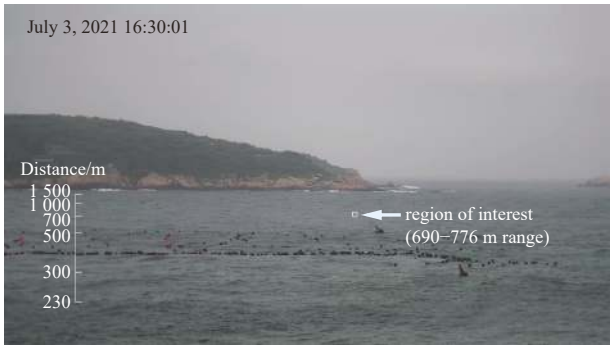

Fig. 10. Selected region for assessing the impact of distance.

4.2 *Potential for estimating significant wave height*

Through observation of the video, it was found that the vertical displacement of texture is closely correlated with wave height, where larger waves correspond to greater movement distances. Based on this, the present study further explored the potential of extracting significant wave heights from the videos. Specifically, mapping the pixel movement of texture to real-world distances, followed by a smoothing technique to derive the sea surface elevation. Finally, the significant wave height was calculated from the derived data. The pinhole camera model can be used to establish a mapping relationship between pixel heights and actual heights (Hao et al., 2023; Yoo and Kim, 2023), as shown in Fig. 12.

In Fig. 12, y is the height of a single pixel, O is the optical center of the camera, and Z is the distance between the camera and the observation point. Y represents the actual height of the sea surface ripple change corresponding to a one-pixel movement of texture and can be determined using the equation below:

$$Y = \frac{Z \times \tan\frac{\alpha}{2}}{\frac{n}{2}}, \tag{1}$$

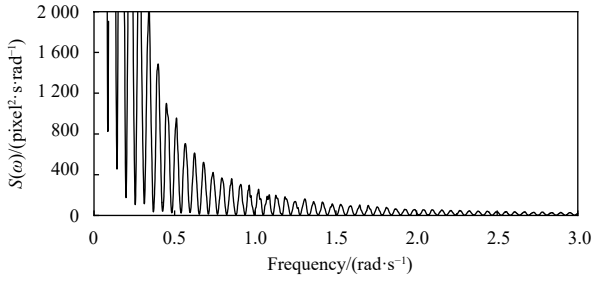

Fig. 11. Spectrum derived from texture displacement extracted at a distant location.

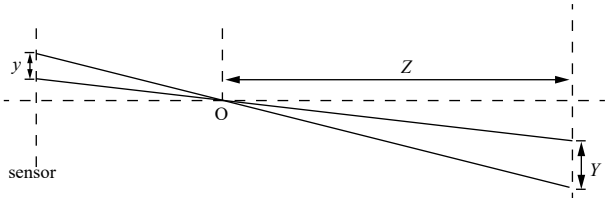

Fig. 12. Relationship between height in the image and real-world height.

where α is the vertical field of view of the camera, and n is the vertical resolution. By multiplying the calculated height of a single pixel by the number of pixels corresponding to texture displacement, the unit can be converted from "pixel" to "m". However, the obtained texture displacement is unsuitable for direct use as sea surface height data. The sea surface can be affected by many factors. Although waves are the main factor in the motion of the texture, other environmental factors, such as current and wind, affect the movement of the texture of the sea surface. These factors give the texture in the image an additional component of motion in the vertical direction, resulting in a bias in the displacement of the texture extracted from the video. For example, in Fig. 5, the accumulated texture displacement starts from 0 pixel and moves up to 5.7 pixel and then down to –1.9 pixel instead of returning to the initial 0 pixel. Such bias accumulates gradually over time, leading to a significant deviation. This data deviation is a long-period variation that a moving average filter can separate. The filtered result h_{smoothed} reflects the accumulation of displacement deviation, and the difference between this result and the original displacement h_{original} is the real sea surface elevation h_{result}, as in the following:

$$h_{\text{result}} = h_{\text{original}} - h_{\text{smoonthed}}. \tag{2}$$

The wave height data obtained after filtering can be used to calculate the significant wave height using the zero up-crossing method (Pierson Jr., 1954; Rattanapitikon, 2008).

The vertical resolution of the video system is 1 440, and the vertical field of view is 10°. The distance to Region A is measured by the radar to be 245 m. Substituting these values into Eq. (1) yields a height of 0.03 m per pixel. The distance-based texture displacement can be calculated by taking the product of this height and the texture displacement of Region A. The obtained new displacement curve is morphologically consistent with the original curve (Fig. 7), with the only difference being the change in displacement units from "pixel" to "m". To address the long-period offset present in the displacement, a moving average filter is applied to isolate the sea surface elevation from the data. Based on the wave spectrum presented in Fig. 8, a cutoff frequency of 0.8 rad/s was chosen to separate the high and low frequencies. To achieve this, a filter of length 197, equivalent to a temporal window of 7.84 s (25 frames per second), was defined as the convolution kernel for filtering the data. After filtering, the data fluctuated around zero with values less than 1 m, as shown in Fig. 13. The filtered data can be regarded as the sea surface elevation, and the significant wave height $H_{1/3}$ can be calculated to be 0.98 m by the zero up-crossing method.

The filtered sea surface elevation data shown in Fig. 13 have eliminated low-frequency interference, and the wave spectrum plotted by WAFO is shown in Fig. 14. The dominant peak frequency of the wave spectrum is 1.078 0 rad/s, which differs by only 0.095 9 rad/s from the 0.982 1 rad/s shown in Fig. 8. The

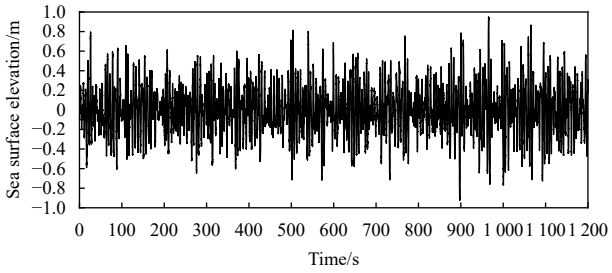

Fig. 13. Sea surface elevation after filtering.

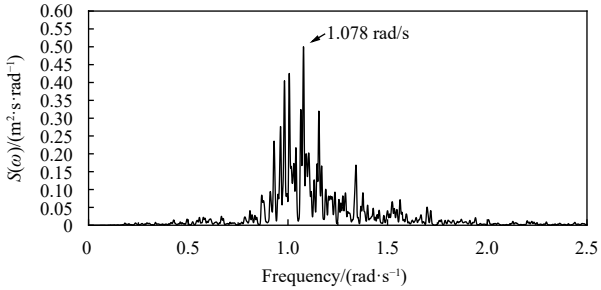

Fig. 14. Wave spectrum plotted from filtered sea surface elevation.

zeroth moment m_0 can be calculated from the wave spectrum by spectral analysis and the significant wave height H_{m0} can be further calculated:

$$H_{m0} = 4\sqrt{m_0}, \tag{3}$$

$$m_0 = \int_0^\infty S(\omega)\mathrm{d}\omega. \tag{4}$$

The zeroth moment m_0 is extracted from the wave spectrum as 0.068 5, from which the significant wave height H_{m0} is further calculated to be 1.05 m. This result is close to the significant wave height $H_{1/3}$ of 0.98 m obtained by the zero up-crossing method, but there is a slight difference. The existence of this difference is reasonable because the wave heights from statistical and spectral analyses will only be consistent if the wave heights follow the Rayleigh distribution (Rattanapitikon and Shibayama, 2013). However, for real data, $H_{1/3}$ is usually not equal to H_{m0} (Battjes and Groenendijk, 2000; Goda, 2009).

The significant wave height obtained from the simultaneous spatio-temporal inversion of X-band radar data is 1.22 m. In comparison, the result from the zero up-crossing method is underestimated by 0.24 m, and that calculated through the zeroth moment is underestimated by 0.17 m. It should be noted that the size of the sub-images is closely related to the significant wave height, which will be discussed in the following section. Although the accuracy of the significant wave height requires further validation with more data, the currently obtained significant wave height can reflect the state of the sea waves to a certain extent, indicating the potential of the video analysis method presented in this paper for inverting significant wave height.

4.3 *Selection of sub-image size*

The size of the sub-image cropped from the video is an issue to be considered. A size that is too small to contain enough texture information and too large to contain multiple textures with

different motion trends can lead to inaccurate matching results. The sub-images ranging from 10 pixel × 10 pixel to 100 pixel × 100 pixel were extracted from Region A of Fig. 6 for illustrative comparison. Figure 15 shows the texture displacement results for 10 pixel × 10 pixel and 20 pixel × 20 pixel. Examination of the data found some anomalous data in the matching results for the 10 pixel × 10 pixel size. For instance, between 335.84 s and 335.88 s, the texture is moved down by up to 5 pixels. However, it is clear that the texture of the sea surface does not move 5 pixels in 0.04 s time. This data implies that the texture was moving very fast, which is clearly false. In fact, the texture usually moves no more than one pixel in a two-frame interval, as shown in Fig. 5. Within 20 min of data comprising 30 000 frames, matching results for texture displacements exceeding one pixel were recorded 260 times in the 10 pixel × 10 pixel images. In contrast, such instances were recorded only 6 times in the 20 pixel × 20 pixel data, which may be attributed to incidental factors such as camera shake.

Selecting a larger size for the sub-image, which encompasses more textures, will likewise preclude the accurate determination of the displacement of the textures. The time interval between the two frames is brief relative to the movement of the textures, where most of the textures that do not move more than one pixel are considered to be not displaced. Only a very small percentage of the textures are recognized as displaced. However, the results of image matching reflect the correlation between the entire sub-images, implying that the displacement of individual textures is difficult to capture. It can also be seen from the figure that the amplitude of texture displacement decreases as the size increases. Table 3 presents the results of the calculations for these ten sizes, where the significant wave height deviates noticeably as the size of the sub-image increases. The significant wave height is reduced by 14% compared to 20 pixel × 20 pixel when the size is increased to 100 pixel × 100 pixel. Therefore, the criterion for selecting the sub-image size is to choose the smallest size possible while ensuring the rare occurrence of anomalous data. From the above analysis, it can be seen that the selection of the region mainly affects the amplitude of the motion of the textures and has little effect on the motion period. This is consistent with the calculated results listed in Table 3, where the peak wave period of each region is observed to be nearly unaffected, exhibiting good consistency. In practical processing, the size of the sub-image can be gradually increased, and the required sub-image size can be determined once the resulting peak wave periods tend to stabilize.

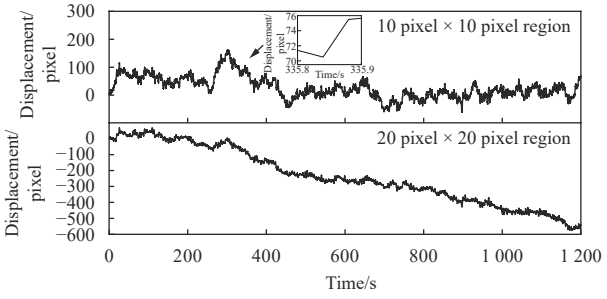

Fig. 15. Comparison of texture displacement between 10 pixel × 10 pixel and 20 pixel × 20 pixel.

Table 3. Peak wave period T_p and significant wave height $H_{1/3}$ corresponding to pixel sizes from 10 × 10 to 100 × 100 for sub-images

Size	T_p/s	$H_{1/3}$/m
10 pixel × 10 pixel	5.83	0.92
20 pixel × 20 pixel	6.40	0.98
30 pixel × 30 pixel	6.40	0.96
40 pixel × 40 pixel	6.40	0.93
50 pixel × 50 pixel	6.40	0.92
60 pixel × 60 pixel	6.40	0.91
70 pixel × 70 pixel	6.40	0.89
80 pixel × 80 pixel	6.40	0.87
90 pixel × 90 pixel	5.90	0.86
100 pixel × 100 pixel	6.19	0.84

To further elucidate the variation in sub-image sizes under different sea state conditions, and considering the limited availability of real data, this study employed Tessendorf's ocean simulation technique for analysis (Tessendorf, 2001). This simulation method combines inverse Fourier transform technique with relevant oceanographic knowledge to create sea surface scenes, thus enabling the generation of videos under various sea state conditions. Building upon this, the sea surface optical video images are calculated using the principles of light reflection, and are generated using the Open Graphics Library (OpenGL). It is equivalent to placing a virtual camera in the simulation process to mimic the real camera imaging. For all simulated videos, the viewport was set to 300 pixel× 300 pixel, with a 2° field of view (equivalent to a 1 500 pixel × 1 500 pixel resolution at a 10° field of view). The viewpoint was positioned 15 m above the sea surface and directed towards the sea at a 3° depression angle, as shown in Fig. 16. The Fourier grid was set to 256 × 256 in size, and each video had a duration of 20 min at 25 frames per second. The model used to generate the sea surface video is based on the Phillips spectrum (Tessendorf, 2001), which is parameterized by wind speed. There is a relationship between wind speed and peak wave period as follows (Tessendorf, 2001; Spencer et al., 2006):

$$T = \frac{2\pi V\sqrt[4]{2}}{g}, \tag{5}$$

where V is the wind speed, g is the acceleration of gravity (9.8 m/s^2), and T is the peak wave period.

In the generated videos, the wind speeds were set at 3 m/s, 7 m/s, and 12 m/s, propagating parallel to the line of sight. The peak wave periods derived from video analysis and the theoretical values calculated according to Eq. (5) are summarized in Fig. 17. At a wind speed of 3 m/s, the peak wave periods derived from small-sized sub-images exhibited some fluctuation and stabilized starting from a 60 pixel × 60 pixel size. At wind speeds of 7 m/s and 12 m/s, the peak wave periods derived from video analysis are relatively stable, except for significant deviations at the

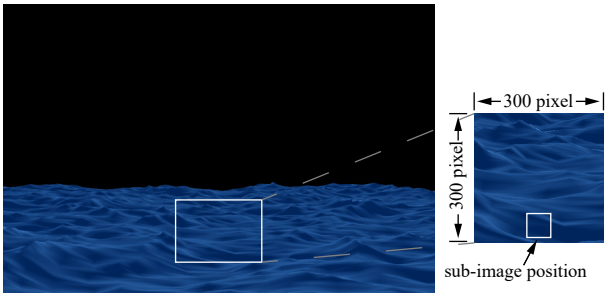

Fig. 16. Simulated sea surface video image.

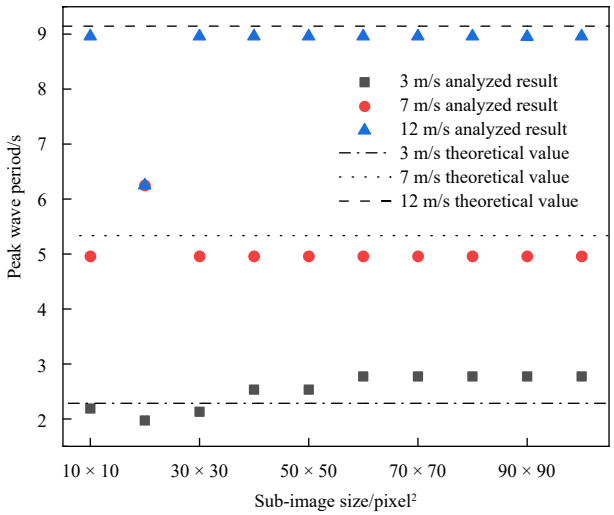

Fig. 17. Comparison between the peak wave periods calculated from the sub-image sizes ranging from 10 pixel × 10 pixel to 100 pixel × 100 pixel and the theoretical values, under wind speed conditions of 3 m/s, 7 m/s, and 12 m/s.

20 pixel × 20 pixel. According to the analysis results, the size of the sub-images should not be too small, which may be due to the lack of fine textures in the simulated videos. Changes in wind speed result in variations in the size of the selected sub-image. However, when the size of the sub-image is within a certain range, the peak wave period exhibits relatively high stability, which is consistent with the calculation results from real data.

5 Conclusions

This paper proposes a method for extracting peak wave period from sea surface video. The method first identifies the motion of sea surface textures through image matching and then calculates the peak wave period from these data. These extracted peak wave periods are compared with X-band radar inversion results and values calculated from buoy motion tracking. This method demonstrates good stability in inverting the peak wave period across varying sizes of sub-images. However, it encounters limitations in distinguishing peak positions when the distance becomes excessive. Hence, when selecting regions, it is advisable to choose those that are neither too large nor too distant, to ensure clear visibility of texture features within the observation area. Under these conditions, the peak wave periods derived from the method presented in this study exhibited a maximum deviation of 8% compared to the results from X-band radar inversion, and a maximum deviation of 14% compared to the results from tracking buoy motions. Although the method still has some limitations in selecting areas, it can provide a cost-effective option for near-shore surface wave monitoring, especially in areas where conventional instruments are difficult or costly to deploy. Moreover, this study employed ocean simulation technique to validate the effectiveness of this method under various sea state conditions, as well as the size of sub-images corresponding to each condition. For future work, this method needs more real data with various sea state conditions to be verified and improved, thereby elevating its applicability and utility in engineering applications.

References

Afzal M S, Kumar L. 2022. Propagation of waves over a rugged topography. Journal of Ocean Engineering and Science, 7(1): 14–28, doi: 10.1016/j.joes.2021.04.004

Ahn S, Haas K A, Neary V S. 2020. Wave energy resource characterization and assessment for coastal waters of the United States. Applied Energy, 267: 114922, doi: 10.1016/j.apenergy.2020.114922

Alberello A, Bennetts L G, Onorato M, et al. 2022. Three-dimensional imaging of waves and floes in the marginal ice zone during a cyclone. Nature Communications, 13(1): 4590, doi: 10.1038/s41467-022-32036-2

Almar R, Bergsma E W J, Catalan P A, et al. 2021. Sea state from single optical images: A methodology to derive wind-generated ocean waves from cameras, drones and satellites. Remote Sensing, 13(4): 679, doi: 10.3390/rs13040679

Ardhuin F, Stopa J E, Chapron B, et al. 2019. Observing sea states. Frontiers in Marine Science, 6: 124, doi: 10.3389/fmars.2019.00124

Arefin M A, Saeed M A, Akbar M A, et al. 2022. Analytical behavior of weakly dispersive surface and internal waves in the ocean. Journal of Ocean Engineering and Science, 7(4): 305–312, doi: 10.1016/j.joes.2021.08.012

Battjes J A, Groenendijk H W. 2000. Wave height distributions on shallow foreshores. Coastal Engineering, 40(3): 161–182, doi: 10.1016/S0378-3839(00)00007-7

Bay H, Ess A, Tuytelaars T, et al. 2008. Speeded-up robust features (SURF). Computer Vision and Image Understanding, 110(3): 346–359, doi: 10.1016/j.cviu.2007.09.014

Benetazzo A. 2006. Measurements of short water waves using stereo matched image sequences. Coastal Engineering, 53(12): 1013–1032, doi: 10.1016/j.coastaleng.2006.06.012

Bergamasco F, Benetazzo A, Yoo J, et al. 2021. Toward real-time optical estimation of ocean waves' space-time fields. Computers & Geosciences, 147: 104666, doi: 10.1016/j.cageo.2020.104666

Bergamasco F, Torsello A, Sclavo M, et al. 2017. WASS: An open-source pipeline for 3D stereo reconstruction of ocean waves. Computers & Geosciences, 107: 28–36, doi: 10.1016/j.cageo.2017.07.001

Blenkinsopp C E, Bayle P M, Martins K, et al. 2022. Wave runup on composite beaches and dynamic cobble berm revetments. Coastal Engineering, 176: 104148, doi: 10.1016/j.coastaleng.2022.104148

Brodtkorb P A, Johannesson P, Lindgren G, et al. 2000. WAFO—A Matlab toolbox for analysis of random waves and loads. Paper presented at The Tenth International Offshore and Polar Engineering Conference. Washington, DC, USA. ISOPE 2000-GFC-02

Chaturvedi S K. 2019. A case study of tsunami detection system and ocean wave imaging mechanism using radar. Journal of Ocean Engineering and Science, 4(3): 203–210, doi: 10.1016/j.joes.2019.04.005

Cui He, Chen Jianyu, Cao Zhenyi, et al. 2022. A novel multi-candidate multi-correlation coefficient algorithm for GOCI-derived sea-surface current vector with OSU tidal model. Remote Sensing, 14(18): 4625, doi: 10.3390/rs14184625

Davison S, Benetazzo A, Barbariol F, et al. 2022. Space-time statistics of extreme ocean waves in crossing sea states. Frontiers in Marine Science, 9: 1002806, doi: 10.3389/fmars.2022.1002806

Falcon E, Mordant N. 2022. Experiments in surface gravity-capillary wave turbulence. Annual Review of Fluid Mechanics, 54(1): 1–25, doi: 10.1146/annurev-fluid-021021-102043

Goda Y. 2009. A performance test of nearshore wave height prediction with CLASH datasets. Coastal Engineering, 56(3): 220–229, doi: 10.1016/j.coastaleng.2008.07.003

Goncalves H, Corte-Real L, Goncalves J A. 2011. Automatic image registration through image segmentation and SIFT. IEEE Transactions on Geoscience and Remote Sensing, 49(7): 2589–2600, doi: 10.1109/TGRS.2011.2109389

Guimarães P V, Ardhuin F, Bergamasco F, et al. 2020. A data set of sea surface stereo images to resolve space-time wave fields. Scientific Data, 7(1): 145, doi: 10.1038/s41597-020-0492-9

Hao Yang, Tang Tao, Gao Chunhai. 2023. Train distance estimation in turnout area based on monocular vision. Sensors, 23(21):

8778, doi: 10.3390/s23218778
Holman R A, Stanley J. 2007. The history and technical capabilities of Argus. Coastal Engineering, 54(6/7): 477–491, doi: 10.1016/j.coastaleng.2007.01.003
Horn B K P, Schunck B G. 1981. Determining optical flow. Artificial Intelligence, 17(1–3): 185–203, doi: 10.1016/0004-3702(81)90024-2
Kim B O. 2005. Photography aided determination of video camera orientation in coastal environments. Journal of Coastal Research, (42): 352–362
Kim B O, Cho H Y. 2005. Image processing for video images of buoy motion. Ocean Science Journal, 40(4): 213–220, doi: 10.1007/BF03023521
Kim B O, Cho H Y, Lim D I, et al. 2008. Nearshore wave measurement using single-video images of buoy motions. Journal of Coastal Research, 246: 1481–1486, doi: 10.2112/07-0850.1
Kim J I, Hyun C U, Han H, et al. 2019. Evaluation of matching costs for high-quality sea-ice surface reconstruction from aerial images. Remote Sensing, 11(9): 1055, doi: 10.3390/rs11091055
Kim M, Lee S, Hong J W. 2022. Empirical estimation of the breaker index using a stereo camera system. Ocean Engineering, 265: 112522, doi: 10.1016/j.oceaneng.2022.112522
Li Jiangxia, Pan Shunqi, Chen Yongping, et al. 2022. Assessment of combined wind and wave energy in the tropical cyclone affected region: An application in China seas. Energy, 260: 125020, doi: 10.1016/j.energy.2022.125020
Lowe D G. 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2): 91–110, doi: 10.1023/B:VISI.0000029664.99615.94
Malila M P, Thomson J, Breivik Ø, et al. 2022. On the Groupiness and intermittency of oceanic whitecaps. Journal of Geophysical Research: Oceans, 127(1): e2021JC017938, doi: 10.1029/2021JC017938
Malliouri D I, Memos C D, Soukissian T H, et al. 2021. Assessing failure probability of coastal structures based on probabilistic representation of sea conditions at the structures' location. Applied Mathematical Modelling, 89: 710–730, doi: 10.1016/j.apm.2020.08.001
Osorio A F, Montoya-Vargas S, Cartagena C A, et al. 2019. Virtual BUOY: A video-based approach for measuring near-shore wave peak period. Computers & Geosciences, 133: 104302, doi: 10.1016/j.cageo.2019.07.006
Palmsten M L, Brodie K L. 2022. The coastal imaging research network (CIRN). Remote Sensing, 14(3): 453, doi: 10.3390/rs14030453
Pan Hailang, Gao Peilin, Zhou Huicheng, et al. 2020. Roughness analysis of sea surface from visible images by texture. IEEE Access, 8: 46448–46458, doi: 10.1109/ACCESS.2020.2978638
Perugini E, Soldini L, Palmsten M L, et al. 2019. Linear depth inversion sensitivity to wave viewing angle using synthetic optical video. Coastal Engineering, 152: 103535, doi: 10.1016/j.coastaleng.2019.103535
Pierson Jr W J. 1954. An interpretation of the observable properties of 'sea' waves in terms of the energy spectrum of the Gaussian Record. Eos, Transactions American Geophysical Union, 35(5): 747–757, doi: 10.1029/TR035i005p00747
Rattanapitikon W. 2008. Verification of significant wave representation method. Ocean Engineering, 35(11/12): 1259–1270, doi: 10.1016/j.oceaneng.2008.03.008
Rattanapitikon W, Shibayama T. 2013. Verification and extension of goda formulas for computing representative wave heights transformation. Coastal Engineering Journal, 55(3): 1350009, doi: 10.1142/S0578563413500095
Spencer L, Shah M. 2004. Water video analysis. In: Proceedings of International Conference on Image Processing. Singapore: IEEE, 2705–2708, doi: 10.1109/ICIP.2004.1421662
Spencer L, Shah M, Guha R K. 2006. Determining scale and sea state from water video. IEEE Transactions on Image Processing, 15(6): 1525–1535, doi: 10.1109/TIP.2006.871102
Stilwell Jr D. 1969. Directional energy spectra of the sea from photographs. Journal of Geophysical Research, 74(8): 1974–1986, doi: 10.1029/JB074i008p01974
Tessendorf J. 2001. Simulating ocean water. In: Simulating nature: realistic and interactive techniques. SIGGRAPH, 3-1–3-26
Ti Zilong, Zhang Mingjin, Li Yongle, et al. 2019. Numerical study on the stochastic response of a long-span sea-crossing bridge subjected to extreme nonlinear wave loads. Engineering Structures, 196: 109287, doi: 10.1016/j.engstruct.2019.109287
Vieira M, Guimarães P V, Violante-Carvalho N, et al. 2020. A low-cost stereo video system for measuring directional wind waves. Journal of Marine Science and Engineering, 8(11): 831, doi: 10.3390/jmse8110831
Villas Bôas A B, Ardhuin F, Ayet A, et al. 2019. Integrated observations of global surface winds, currents, and waves: Requirements and challenges for the next decade. Frontiers in Marine Science, 6: 425, doi: 10.3389/fmars.2019.00425
Wu Lichung, Doong Dongjiing, Lai Jianwu. 2022. Influences of nononshore winds on significant wave height estimations using coastal X-band radar images. IEEE Transactions on Geoscience and Remote Sensing, 60: 1–11, doi: 10.1109/TGRS.2021.3077903
Ye Yuanxin, Bruzzone L, Shan Jie, et al. 2019. Fast and robust matching for multimodal remote sensing image registration. IEEE Transactions on Geoscience and Remote Sensing, 57(11): 9059–9070, doi: 10.1109/TGRS.2019.2924684
Yoo S, Kim N. 2023. Coarse alignment methodology of point cloud based on camera position/orientation estimation model. Journal of Imaging, 9(12): 279, doi: 10.3390/jimaging9120279