



OPEN Research on motion target detection based on infrared biomimetic compound eye camera

Linhan Li^{1,2}, Xiaoyu Wang^{1,2}, Teng Lei^{1,2}, Juan Yue¹, Sili Gao¹✉, Yang Yu¹✉ & Haifeng Su³✉

The cooled mid-wave infrared biomimetic compound eye camera has wide range of applications, such as industrial inspection, military project, and security. Due to the low resolution of individual eyes and the large field view of the imaging system, existing motion target enhancement and detection algorithms cannot effectively detect all potential targets. To address this issue, we propose an improved elementary motion detector model that combines a double-layer ON/OFF channel and a cross-type computational architecture, which is able to suppress a stationary background and enhance moving targets. In order to further reduce the missed detection, we designed the spatial and temporal consistency detection methods of the compound eye structure, which further improved the accuracy and stability of the detection results. The experimental results show that our method can fully utilize the features of the image and can be applied to the enhancement and detection of moving objects in complex scenes, and the detection efficiency is significantly improved.

Motion detection is a crucial aspect of the biological visual system as it enables organisms to perceive and interpret the world by detecting motion in their surrounding environment. This ability is present in many organisms, including humans and various animals. Organisms receive light signals through photoreceptor cells, convert them into nerve signals, transmit them to the brain for processing, and exhibit sensitivity to different motion signals through the arrangement of different neurons. In nature, many insects or crustaceans perceive an enlarged field of view through the structure of compound eyes. Thousands of closely spaced small-eyes can provide local responses at different positions on the retina, which can be used to calculate the orientation and distance of the organism itself with other objects in the environment, and to perceive the shape and relative size of the target. Multiple photosensitive cells working synchronously can process multiple local information in parallel, which is more favorable for compound-eyed insects in making faster judgments and responses. Due to the use of parallel processing of information and the flicker effect^{1–3}, compound eye organisms possess a sensitivity to moving objects. The characteristics of the compound eye visual system of these organisms have also inspired many researchers. For many years, numerous scholars have been engaged in studying bionic compound eye imaging systems and motion detection methods based on biological visual mechanisms.

German biologist Reichardt conducted theoretical analysis on the relative motion neural computation of beetles by analyzing their visual behavior and proposed Elementary motion detector (EMD)⁴. The EMD model relies on the information of different locations in the visual scene for temporal difference analysis, that is, the signal difference generated by the time delay is used to obtain motion information. This model demonstrates sensitivity to the speed and direction of moving objects, exhibiting relative stability even in low-light conditions. Santen proved that the model proposed by Reichardt is also applicable to the visual system of mammals. In the following decades, many scholars have successively proposed several variant models of EMD. Pallus et al. introduced a novel architecture for the EMD model, known as the correlation-based EMD model⁵. Researchers have also proposed a dual-channel EMD model⁶, which processes the input signal through separate channels based on changes in brightness. Considering the estimation of motion direction and velocity, Wang et al. proposed a lateral inhibition model based on the EMD theory to improve the performance of small target detection in surveillance cameras⁷. On this basis, they further improved the model for detecting dynamic small targets, including setting up feedback and integrating multi-field of view visual features^{8,9}. Wang and his team developed a motion detector based on the theory of EMD, which uses the shallow visual neural pathway of fruit flies for three-dimensional object detection¹⁰. Many scholars have been studying the visual neurophysiological model of compound eyes in order to apply it to motion detection in complex backgrounds, and they have made

¹Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China. ²University of Chinese Academy of Sciences, Beijing 100049, China. ³Shanghai University, Shanghai 200444, China. ✉email: gaosili@mail.sitp.ac.cn; yuyang@mail.sitp.ac.cn; hfsu@shu.edu.cn

good progress^{11–15}. This provides a biological basis for finely elucidating the visual neural computation process of target motion perception.

Meanwhile, many scientists study the structure of compound eyes based on the visual characteristics of compound eye organisms. In contrast to the single-eye structure, compound eyes are composed of many independent imaging units called ommatidia. These ommatidia are arranged on a curved surface to collect light information from various directions in the scene. This means gaining a broader perspective and more accurate visual information. In 1891, Exner proposed the theory of overlapping visual fields in compound eyes, while the research findings of scientists such as H.B. Barlow, M.F. Land, G.A. Horridge have provided insights for the development of new optical sensors and image processing technologies. Over the years, various artificial compound eye structures have been developed, including planar compound eyes, curved compound eyes, camera array compound eyes, and others¹⁶.

The algorithm in our work relies on a self-developed cooled mid-wave infrared biomimetic compound eye camera^{17,18}. By simulating the physiological structure of overlapping compound eyes, a fixed structure multi-lens array is used to map to a single sensor for imaging. The imaging performance and image characteristics will be detailed in the subsequent chapters. Compared to a multi-sensor structure, its advantages include a more convenient calibration process and smaller imaging errors. Our infrared biomimetic compound eye camera has advantages, such as a large field of view, small size, and high sensitivity similar to a compound eye camera. It also has the advantages of cooled medium wave infrared photoelectric equipment, such as working all day, having better performance through fog and dust, better environmental adaptability and low altitude detection ability, can fully sense the surrounding environment and capture potential targets. It can be used in various fields such as industrial inspection, military, and security.

In this work, we combine the image characteristics of infrared biomimetic compound eye imaging system and the sensitivity of fruit fly vision to fast-moving objects. First, we conduct a simulation analysis of the Elementary Motion Detector model to verify its effectiveness in infrared biomimetic compound eye image sequences. Then, we make improvements based on the image features. Finally, we test the algorithm using common evaluation metrics on real image sequences captured by the infrared biomimetic compound eye camera and simulated data obtained through the mechanical structure and imaging characteristics.

Methods

In 1961, Reichardt constructed the famous Elementary Motion Detector model (EMD)¹⁹. In 1987, Reichardt and Egelhaaf conducted a comprehensive theoretical analysis of a one-dimensional first-order approximation of fly motion detection model²⁰. The simplified model of 1D EMD is shown in Fig. 1(A). Set the spatial distance between two adjacent photo sensors as $\Delta\phi$. The inputs are the signals received by A and B. The two input signals are respectively multiplied by the time delay unit τ and then subtracted from the other input signal to obtain the output signal out of EMD.

We can use a sine grating as a graphical input to simulate one-dimensional EMD. Let the coordinates of the two light sensors AB be x_1, x_2 , with a spatial distance of $\Delta\phi$. The input graph moves at a continuous velocity $(ds(t))/dt$, where $s(t)$ represents the displacement of the temporal domain graph, that is, $(ds(t))/dt = C$ represents the instantaneous velocity of the image. The input mode can be represented as:

$$F(x, t) = I + \Delta I \sin\left(\frac{2\pi}{\lambda}(x + s(t))\right) \quad (1)$$

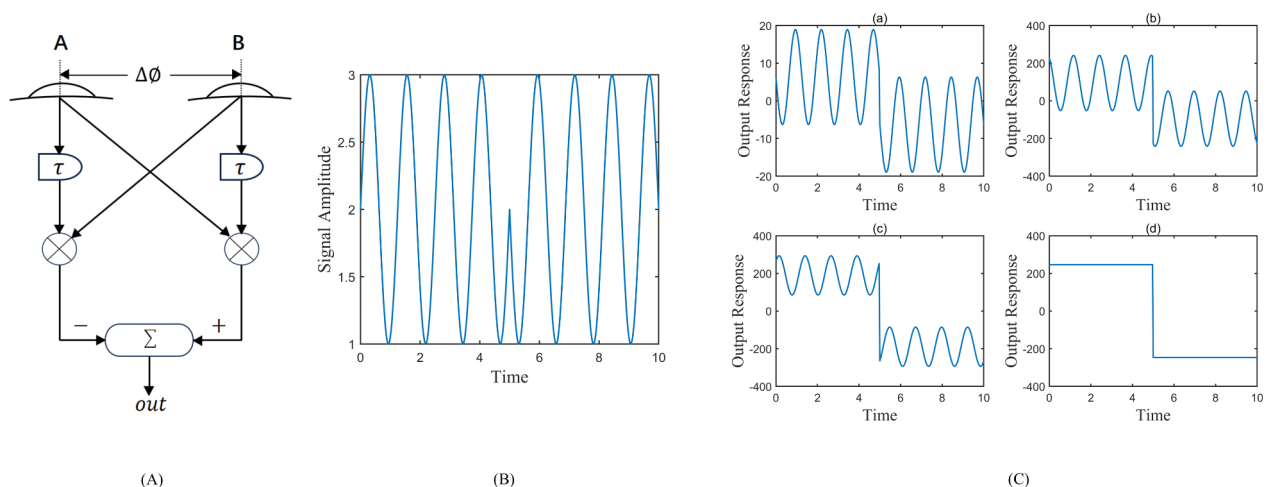


Fig. 1. EMD simulation. (A): primary motion detection EMD model. (B): input signal. (C): output of the EMD detector. (a), (b), (c), and (d) represent the response results of integrated outputs from 1, 13, 30, and 40 motion detectors respectively.

λ represents the spatial wavelength of the grating, I represents the average light intensity, and ΔI represents the modulation of light intensity. The output can be represented as the graphic input moving from left to right at a speed of $-V$, with a time delay of τ :

$$R(x, t) = F[x + s(t)] * F[x + \Delta \varnothing + s(t - \tau)] - F[x + \Delta \varnothing + s(t)] * F[x + s(t - \tau)] \quad (2)$$

$$R(x, t) = \tau * V * \left(\frac{2\pi}{\lambda}\right)^2 * \Delta I * \left[\Delta I + I * \sin\left(\frac{2\pi}{\lambda}(x - Vt)\right)\right] \quad (3)$$

We set up a group of input signals: $\lambda = 3$, $I = 2$, $\Delta I = 1$. The signal initially moves uniformly with a velocity of $V = 4$. At $t = 5$, it then reverses its motion. The graph of this input signal is shown in Fig. 1(B). The simulation output of EMD is shown in Fig. 1(C), covering 40 motion detectors within a spatial period. Each detector is separated by $1/40$ of a spatial wavelength. It can be seen that the EMD output signal can indeed detect local motion, and the direction of the input signal's motion is represented by the sign of the output signal, while its magnitude is also proportional to the speed of the motion. Furthermore, it can be observed from (d) that by integrating 40 detectors, the influence of time modulation can be eliminated.

Based on the above theoretical results, the EMD can be extended to two-dimensional image motion detection, improving the output response through two-dimensional spatial integration. The simplified model is shown in Fig. 2(A). We also simulated the two-dimensional movement of the targets in space: in Fig. 2(B), We simulated two different sizes of moving and static targets. Within a single image, the circular region is the imaging of a single small-eye. Among them, S1, S2 and S3 are three small-eyes, S1 and S2 are arranged horizontally, S1 and S3 are arranged vertically. The three kinds of targets were small square, small cross and large cross. In the time range from T0 to T1, the small square remained stationary, and the two kinds of cross moved from the upper right to the lower left in the field of view of the three small-eyes.

Based on the above simulation results, it can be seen that two-dimensional EMD can suppress stationary targets and enhance moving targets. With simple binarization processing, motion target detection can be achieved. However, considering the practical application, if we calculate using the closest small-eyes and shorten the distance between the photoreceptors, it may not fully enhance all target pixels for larger moving targets. In the application of biomimetic compound eye camera, due to the lower number of pixels in individual small-eyes, many of the target in the entire compound eye image belong to the category of small targets. We can improve the 2D EMD based on the camera structure and image features.

Infrared biomimetic compound eye imaging system

Our work relies on the innovative equipment of a cooled mid-wave infrared biomimetic compound eye camera. Based on the physiological structure of overlapping compound eyes, small lenses are arranged in a circular shape on a spherical structure, and imaging is achieved on a single plane detector through a relay optical structure.

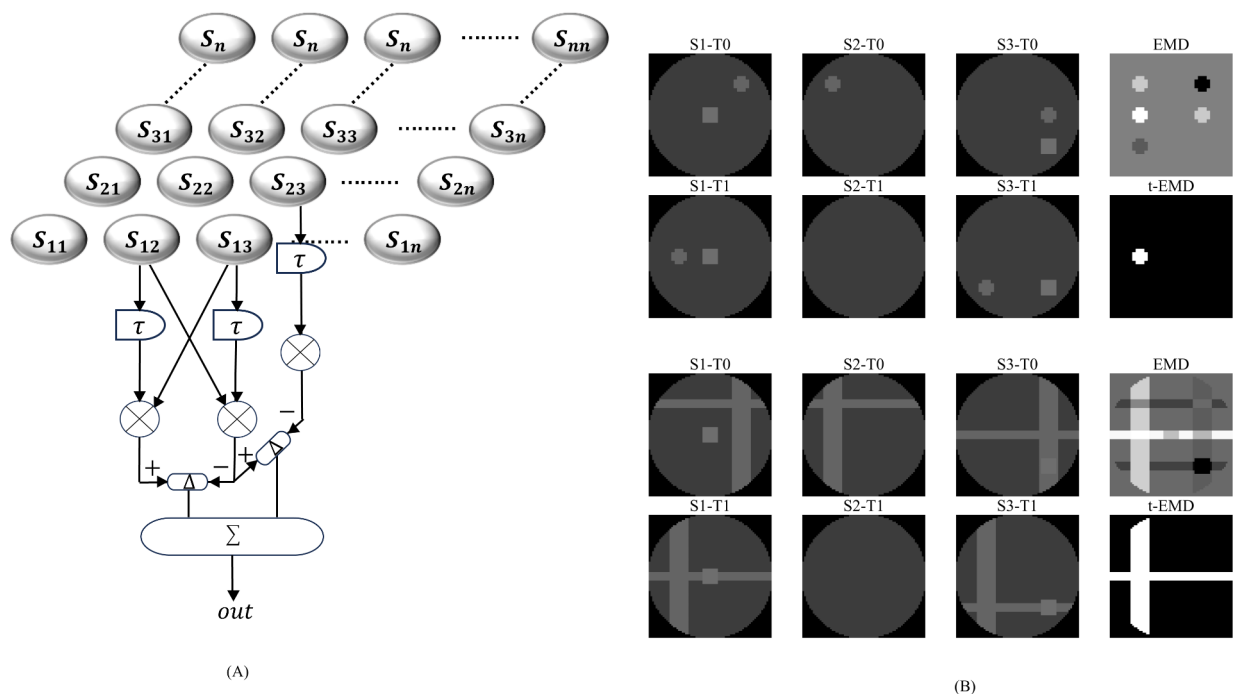


Fig. 2. EMD simulation results of the motion target. (A): two-dimensional EMD diagram. (B): EMD simulation results, including the moving target and stationary target.

In our camera, the optical axis of the small lens is perpendicular to the installation structure's spherical surface. The primary optical axis of the compound eye is the optical axis of the central small lens. 4 circles of small lenses are extended outward at intervals of 10 degrees. The small lenses on each circle are evenly distributed, with quantities of 6, 12, 18, and 24, respectively. Including the central small lens, there are a total of 61 small lens mounting holes, forming a field of view of $108^\circ \times 108^\circ$. The distortion at the edge of the small lens is approximately 4–5%, achieving large field of view imaging with minimal distortion at the edge. The schematic diagram of its imaging effect is shown in Fig. 3. Considering the actual imaging effect and the data collected in the laboratory, this work mainly considers the data collected by the central three annuli of smaller eyes to study the moving target enhancement. Figure 3(C) shows the imaging result of the central three annuli of smaller eyes on the character A target at a distance of 1 m^{17,18}.

Taking into account the actual situation, the data obtained from the prototype is mostly laboratory data. When considering the effectiveness of validating algorithms, it is necessary to test various scenario images. Based on the distribution of small-eyes' positions, the radius, and the coordinates of the center small-eye, we can simulate an infrared biomimetic compound eye image sequence by statistically analyzing the imaging results of distant target. Simulated data has the following characteristics: (1) Compared to fast-moving targets, the background tends to be more static or have slower movements. Therefore, image sequences of small targets with different motion speeds are generated in dynamic backgrounds or intricate backgrounds. (2) Image sequences of targets with different contrast levels compared to the background's grayscale distribution. In this paper, the introduction of simulated data will be discussed in the result section.

Motion target enhancement algorithm based on two-dimensional EMD

The overall algorithm is shown in Fig. 4(A), we obtain the corresponding numbered small-eye images from the whole image. In fly visual cells, ON-type input signals result in an increase in the voltage difference between cells,

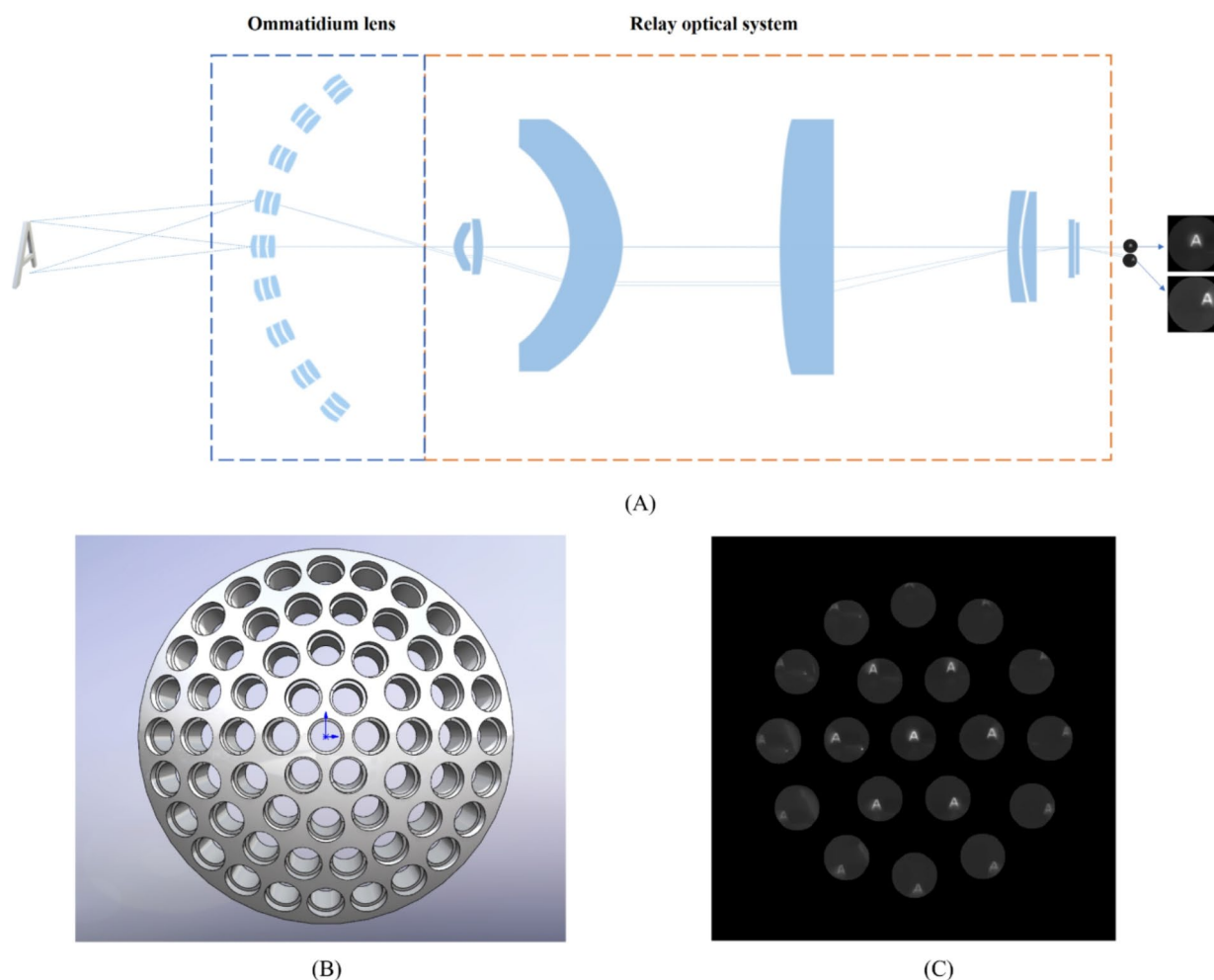


Fig. 3. Illustration of the image formation of a refrigeration type mid-wave infrared biomimetic compound eye. (A): simplified optical diagram, including a lens array and a relay optical system. (B): schematic diagram of the lens array shell structure. (C): schematic diagram of the actual imaging effect.

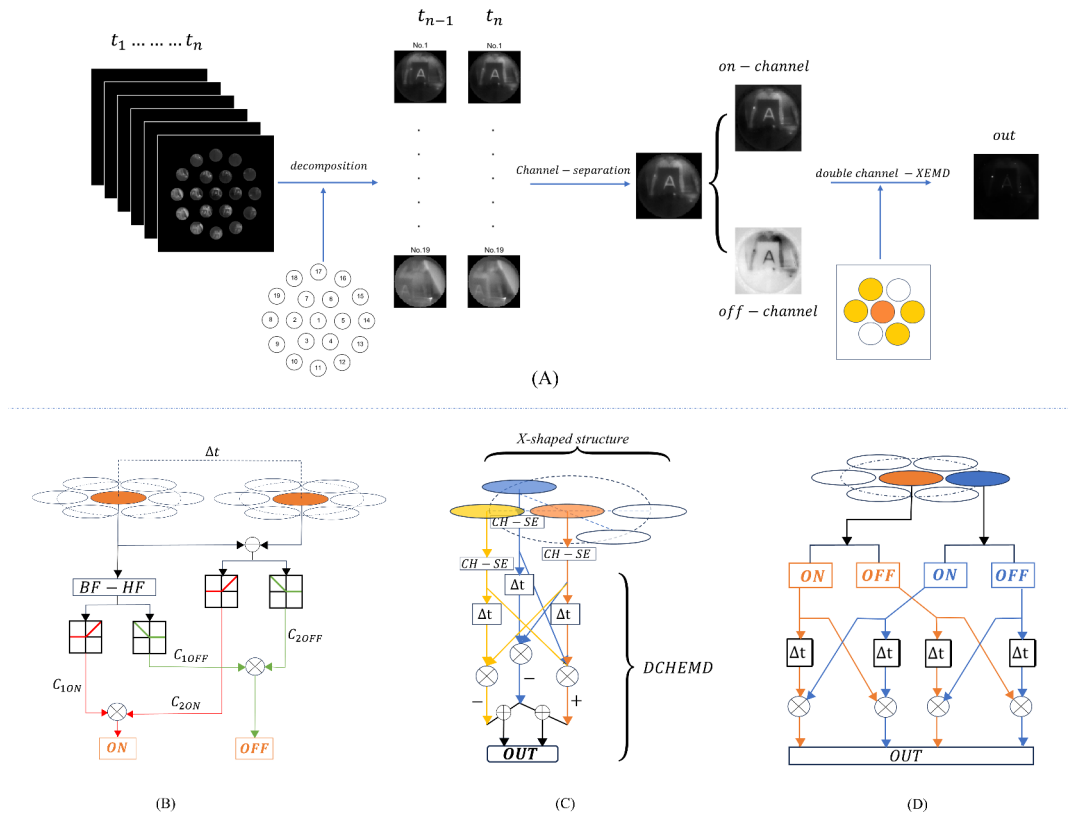


Fig. 4. Algorithm flow diagram. (A): Simple flow diagram of the algorithm. (B): channel separation process (CH-SE). (C): X-shaped structure EMD(XEMD). (D): double channel EMD(DCHEMD).

while OFF-type signals result in a decrease in the voltage difference between cells⁶. The input signal is divided into two channels based on changes in brightness. When the light intensity increases, it is considered as ON channel signal, and when the light intensity decreases, it is considered as OFF channel signal. Therefore, dividing the input signal into two channels can obtain a more accurate motion region. This article uses half-wave rectification to calculate the first stage ON_OFF channel. Firstly, we perform segmentation on the entire compound eye sequence image. Assuming the input signals of all adjacent segmented small-eyes are I_1, I_2, \dots, I_n , the channel separation calculation is performed for all small-eyes, here, I_1 is taken as an example, the segmentation of the ON and OFF channels of the I_1 small-eye can be expressed using the following mathematical model:

$$C_{10N} = \frac{|diff fer_{I_1}| + diff fer_{I_1}}{2} \quad (4)$$

$$C_{10FF} = \frac{|diff fer_{I_1}| - diff fer_{I_1}}{2} \quad (5)$$

$diff fer_{I_1}$ refers to all motion information obtained by comparing the sequence delays. In the study, it was found that regardless of whether the background is still or moving, as long as there is a brightness difference, it will have a negative impact on the enhancement of motion results, resulting in some still backgrounds being mistakenly identified as moving targets. Therefore, in order to avoid the influence of errors caused by dynamic changes in high brightness backgrounds on the final detection results, a butterfly filter is introduced to suppress these high bright backgrounds. Butterfly-shaped high-pass filter(BF-HP) can better preserve the detailed features of the target, while enhancing small targets and suppressing smooth backgrounds. The butterfly filter is:

$$p = \begin{bmatrix} -1.5 & -0.5 & -1.5 \\ -0.5 & 8 & -0.5 \\ -1.5 & -0.5 & -1.5 \end{bmatrix} \quad (6)$$

Similarly, the secondary ON_OFF channel of eye I_1 can be represented by the following mathematical model:

$$C_{20N} = I_1 * \frac{(sgn(I_1) + 1)}{2} \quad (7)$$

$$C_{2OFF} = I_1 * \frac{(1 - \text{sgn}(I_1))}{2} \quad (8)$$

The channel separation process (CH-SE) is shown in Fig. 4(B). Equations (5) and (6) for channel segmentation of a sequence of images of a single small-eye, differ refers to all the motion information obtained by comparing the sequence time delays. The results of the first ON-OFF channel are obtained by the computation of (5) and (6). Equations (8) and (9) are the results of the second ON-OFF channel computed using the sign function for the channel segmentation of the filtered result of a single single-frame small-eye image after (7). Among them, the result of the first level ON-OFF channel contains more motion information, including the motion information of the target and the slow changing information of the brighter background. The result of the second level ON-OFF channel contains more information about small targets in the current frame image. When associating the two-level ON-OFF calculations, a method of mutual filtering is used to obtain more accurate results. The specific process can be simplified as follows: the pixels with the same or similar normalized values obtained from the calculations in both ON channels are preserved more, while the larger values with significant differences are considered as error points introduced by the calculation. After obtaining the two-level ON-OFF result for a single small-eye, we use this result with the two-level ON-OFF result for the neighboring small-eye for EMD calculation, as shown in the Fig. 4.

In the previous section of the one-dimensional EMD mathematical analysis and simulation results, it can be concluded that spatial integration can eliminate the effect of time modulation and improve the output response. Meanwhile, the simulation results of the two-dimensional EMD indicate that the response is affected by the motion of the target in different small-eye's image. Combining the small-eye geometric structure of the device in this article with the imaging overlap rate between the small-eyes, we improved the 2D EMD structure to an X-shaped structure EMD(XEMD). As shown in Fig. 4(C), it consists of two symmetrical two-dimensional EMD structures, which correspond to the hexagonal ring structure expanding from the inside to the outside of the device lens. In this structure, the central small-eye's image of the X-shaped structure are connected to the double-channel output results of the neighboring small-eye's image and input into the double-channel EMD calculation model (DCHXEMD), as shown in Fig. 4(D). Ultimately, a motion enhancement matrix is obtained, where pixels in the motion region are assigned higher weights, while pixels in the static or slowly changing background region have minimal weights.

We use an adaptive threshold to segment objects. In this letter, the adaptive threshold Th is denoted by

$$Th = \mu + \lambda * \delta$$

where μ and δ are the mean value and the standard deviation of the result, respectively. λ is an adjustable parameter, which ranges from 0 to 30.

Consistency detection of target position based on system architecture

For the global motion weights obtained from the above process, we will further conduct spatiotemporal consistency inspection to ensure higher credibility. The consistency here refers to the characteristic that different adjacent small-eye in the infrared biomimetic compound eye imaging system have the same or similar motion information. We can perform consistency detection from both the temporal and spatial perspectives.

According to the overlap of the field of view, there is structural consistency in the imaging of a single small eye on the target scene. As shown in Fig. 3(C), the whole image contains images of 19 small-eye's image, in the geometric structure of the small-eye's image, the 1st, 2nd, 5th, 8th, and 14th small-eye's image are uniformly distributed along the horizontal line of the sphere's shell structure. The 2nd and 5th small-eye images are symmetrical in structure, as are the 8th and 14th small-eye images. The 1st, 11th, and 17th small-eye images are uniformly distributed along the vertical line of the sphere's shell structure, while the rest of the small-eyes are uniformly distributed on the sphere's shell. Therefore, in an ideal situation, we can obtain the same scene image at the given position of different small eyes with adjacent structures. As shown in Fig. 5(A), the imaging situation of the "A" character in front of the camera is as follows: all 19 small lenses in three circles can capture the image of the entire or parts of the "A" character. Among them, the imaging consistency of the 1st small lens can be calculated by the surrounding 6 small lenses, and they can also verify each other. We mainly improve the detection effect by comparing the target pixel coordinates in the adjacent ommatidium images.

According to the results from DCHXEMD, the global motion weight can be calculated, and it can also represent the motion direction of the moving target. In Fig. 5(B), assuming a simple point target moves a certain distance in the up-left direction within a time interval Δt in the first small-eye image, the DCHXEMD results can be computed for this set of input sources to represent the motion direction.

$$\tan\theta = \frac{v_E - v_S}{u_E - u_S} \quad (9)$$

(u, v) represents calculations in the pixel coordinate system, $(u_E, v_E) = \max\{\text{DCHXEMD}\}$, $(u_S, v_S) = \min\{\text{DCHXEMD}\}$. E and S are used only for mathematical calculation identification in order to distinguish calculation results. Assuming that at time t and $t + \Delta t$, the target appears in all small-eye's images, the direction calculation for all targets in small-eye's images should yield consistent results. In all the small eyes that capture the moving target, the coordinates of the target can be obtained. Due to the field of view overlap between adjacent small eye clusters of the device, the same target will appear in multiple small eye clusters, so the pixel coordinates of the target in these small eye clusters are correlated, and the difference of coordinates obtained by different small eye is fixed. The error detection results are filtered by comparing all coordinate

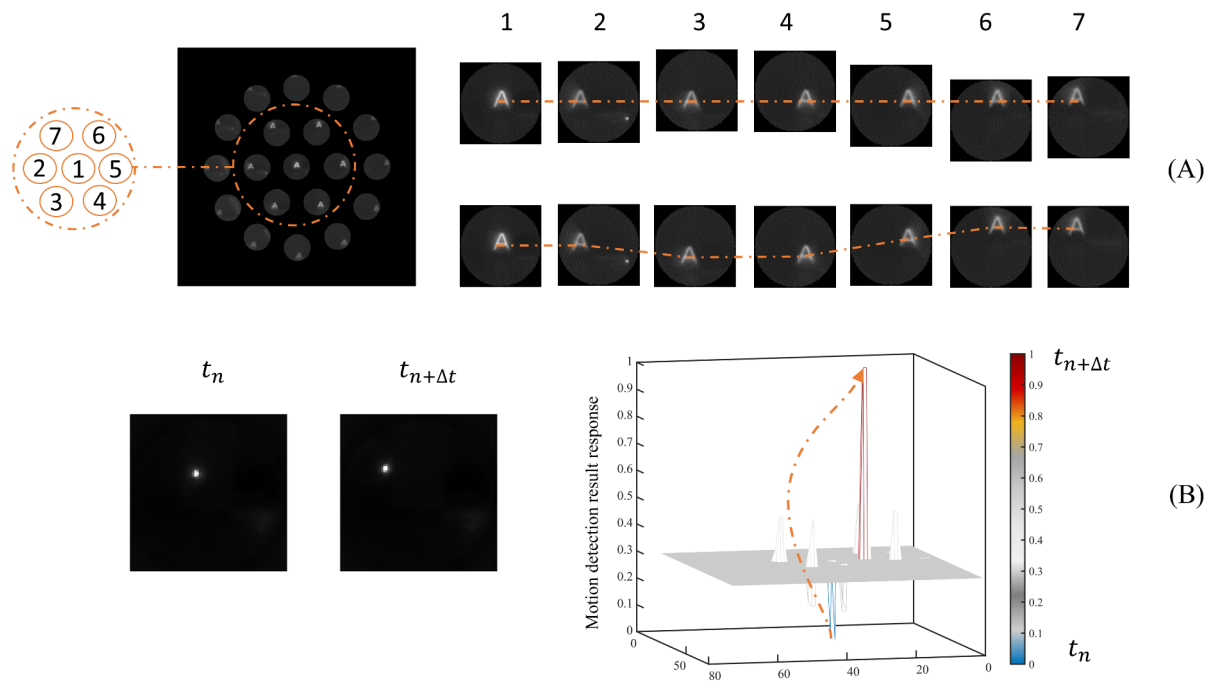


Fig. 5. The schematic diagram of small-eye consistency detection. (A): spatial detection based on small-eye geometric structure consistency detection. (B): temporal detection based on motion direction calculation consistency detection.

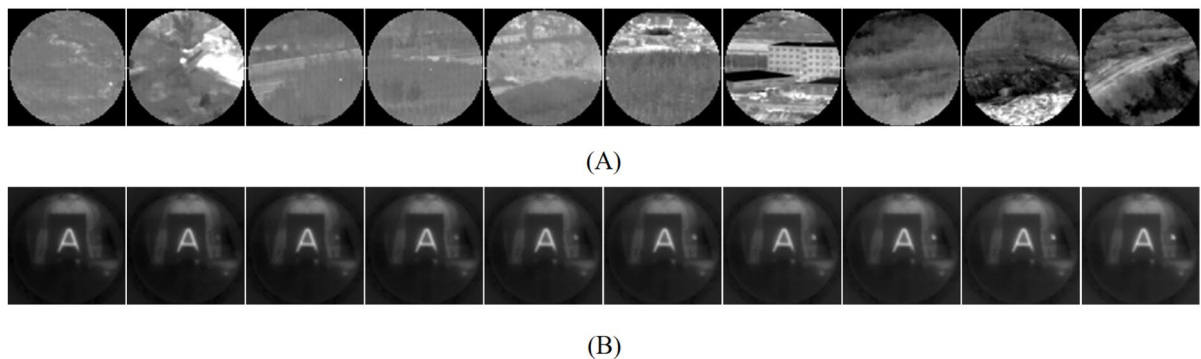


Fig. 6. Partial experimental data. (A) 1st eye for some simulated data with different scene background. (B) 1st eye for laboratory real-shot images with different target contrasts.

information. Also because of the overlapping field of view, the adjacent small eyes also have the same Angle of target motion. The false detection results are further filtered out by comparing the motion angles of all potential targets.

Results

Based on the self-developed infrared bionic compound eye camera, at this stage, we mainly obtain the data in the laboratory and some outdoor static scene data sets. In order to analyze the performance of the algorithm and validate its effectiveness, we use both real data and simulated data for analysis. Various motion target data from different scenarios are selected, as well as data with different types of interference.

Figure 6(A) shows the 1st eye for some simulated data. The scene includes different backgrounds such as trees, grass, mountains, buildings, and roads. Additionally, due to the original data being captured on a mobile

platform, the scene includes dynamic backgrounds. Figure 6(B) shows the 1st eye for laboratory real-shot images with different target contrasts.

From an objective quantitative evaluation perspective, we consider background suppression factors, signal-to-noise ratio, signal-to-noise ratio gain, and contrast gain as evaluation indicators for algorithm performance^{21,22}.

Global Background Suppression Factor (BSF). By conducting statistical analysis on the entire image, it measures the difference between the target and the background, calculates the prominence of the target and the ability to suppress the background. BSF is defined as $BSF = \delta_{in} / \delta_{out}$, where δ_{in} and δ_{out} respectively represent the standard deviations of the entire background region in the input image and the processed image.

The Signal-to-Clutter Ratio Gain (SCRG) is a metric used to evaluate the performance of image processing algorithms by comparing the signal to noise and distortion ratio before and after processing. It is related to the Signal-to-Clutter Ratio (SCR), which is defined as $SCR = |\mu_t - \mu_b| / \sigma_b$. Where μ_t and μ_b respectively represent the average pixel values of the target area and the surrounding background area, where σ_b represents the variance of the pixel values in the background region surrounding the target. SCRG is defined as $SCRG = SCR_{out} / SCR_{in}$. Among them, the SCR_{in} represents the SCR of the input image, and SCR_{out} represents the SCR of the processed image.

Contrast gain CG is an indicator used to evaluate the grayscale contrast enhancement ability of image processing algorithms in target and background, and can be defined as $CG = CON_{out} / CON_{in}$. Where CON is defined as $CON = |\mu_t - \mu_b|$. The larger SCRG, BSF and CG, the better the performance of the algorithm.

The detection probability and false alarm rate are used to measure the performance of an algorithm in target detection tasks. The detection probability refers to the probability of the algorithm correctly detecting a target. It can be represented by the ratio of the number of times the target is correctly identified as a target to the number of times the target actually exists. It can be defined as $P_d = DT / AT$. DT represents the number of detected target, AT represents the number of targets present in the image sequence. In addition, the false alarm rate describes the probability of the algorithm incorrectly reporting non-target samples as targets. The image size of our single small-eye is 75*75. In each of the small-eye's imaging area, each pixel has two possible classifications, target pixel region and background pixel region. Outside the imaging area of each small-eye, all pixels do not contain imaging information. In this paper, considering the low resolution of the image, the number of pixels is used to calculate the false alarm rate, so the calculation index of paper is referred to²². False alarm rate can be defined as $F_a = FP / NP$. In this case, FP represents the number of pixels in the false alarm region, NP represents the total number of pixels in the image sequence. The ROC curve is plotted with the detection probability as the y-axis and the false alarm rate as the x-axis. By calculating the area between the curve and the coordinate axes, we obtain the AUC. A higher AUC value indicates that the algorithm has better discrimination, allowing for better distinction and classification between the target and background.

As shown in Fig. 7 (A), for the enhanced results on the motion objectives, we calculated the correlation indexes for a total of 1000 frames of real captured image sequence datasets with 10 different target contrasts and simulated image datasets with 10 different scenes. In order to test the effectiveness of our improvement, we calculated the index of object enhancement and compared the results of only performing 2D image EMD calculation. The effectiveness of our added channel separation process (CH-SE) step and X-shaped structure EMD(XEMD) model is demonstrated. Meanwhile, we validate the effectiveness of the proposed algorithm for small target detection in infrared biomimetic compound eye imaging systems and compares it with other small target detection algorithms in Fig. 7 (B)(a-b), including double-layer local contrast measure (DLCM)²³, multiscale patch-based contrast measure (MPCM)²⁴, the AADCDD which using absolute average difference weighted by cumulative directional derivatives²⁵, the directional approach is used to develop a novel algorithm called ADMD²⁶, Robinson-Guard filter and pixel convergence(ERG)²⁷, local mutation weighted information entropy(LMWIE)²⁸, Local Component Uncertainty Measure With Consistency Assessment(ELUM)²⁹, anisotropy filter bank and modified with a point spread function(AFB-PSF)³⁰. By analyzing the ROC and AUC, it was found that when applying other detection methods to the detection of individual small-eyes, classical methods with fewer calculations can achieve better detection results due to the fewer pixels in a single small-eye. We also compare several infrared moving small target detection algorithms: the novel spatial-temporal local difference measure (STLDM) algorithm to detect a moving IR small target³¹, the anisotropic spatial-temporal fourth-order diffusion filter (ASTFDF)³², the low-rank and sparse spatial-temporal tensor representation learning model based on local binary contrast measure (STRL-LBCM)³³, a simple but powerful spatial-temporal local contrast filter (STLCF)³⁴. The result shown in Fig. 7 (B)(c-d).

Therefore, we compared the results of several infrared small target detection methods that also have good effects after spatial consistency detection and direction consistency detection, and calculated the detection results in real shooting datasets with different target and background brightness.

Discussion

Through analysis, we can conclude that the existing methods cannot adapt well to infrared biomimetic compound eye images, but our method has good detection performance. It is attributed to the use of a two-dimensional primary motion detection improved model and spatiotemporal consistency detection based on the curved compound eye camera structure, which ensures that our algorithm does not miss any targets. Our method also verifies the sensitivity of the compound eye structure to the moving target. Combined with the advantages of the cooled mid-wave infrared detector, our algorithm will have a broad application scene. Our method is applied to the moving target detection task of the medium-wave cooled infrared biomimetic compound eye system, and it can complete the output of all the hidden targets with small-eyes. At the same time, according to the number of small-eye of the target, we can roughly estimate the bearing, distance and movement direction of the target, so our camera and method can be applied in military detection or fully automatic security tasks. If further

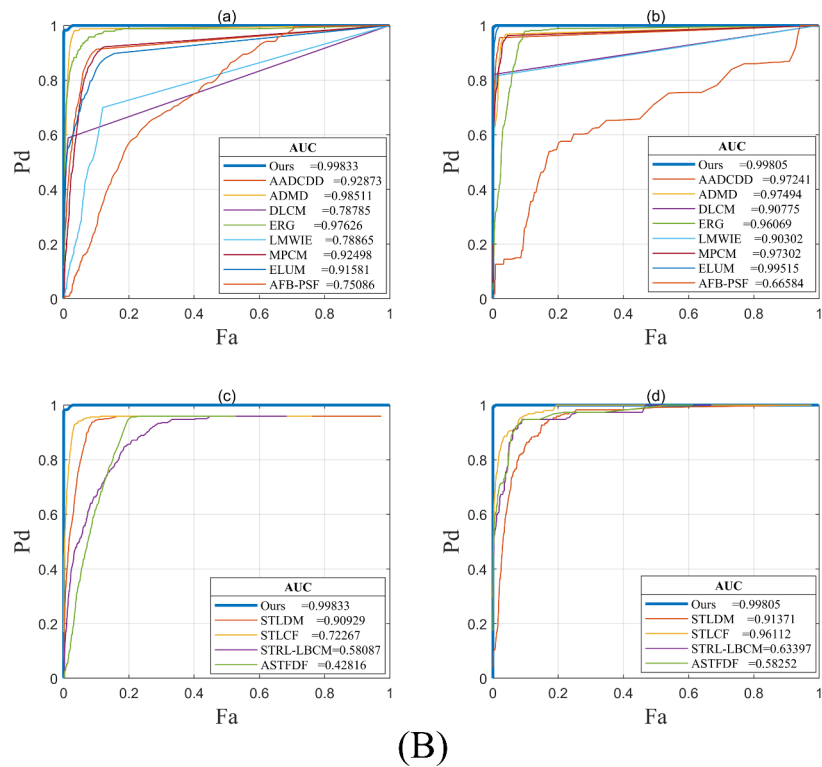
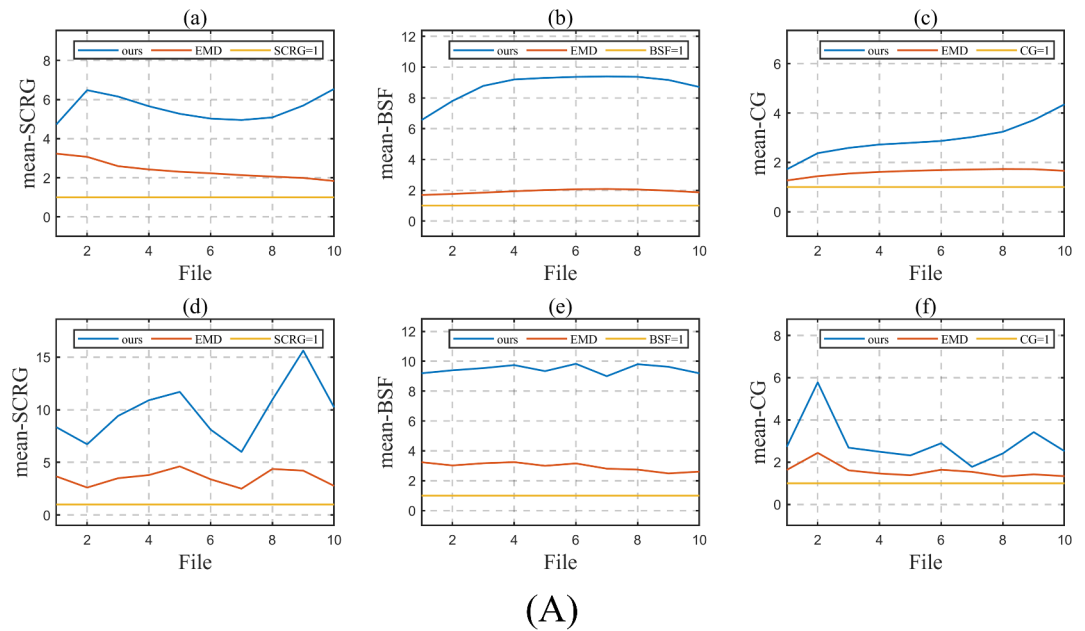


Fig. 7. Ours algorithm performance evaluation results. (A): the enhanced index for motion targets, (a-c) is the result of laboratory real-shot images, (d-f) is the result of simulated data. (B): the ROC curve of target detection results for different algorithms, (a)(c) is the result of simulated data, (b)(d) is the result of laboratory real-shot image.

combined with insect neural structures and information processing processes or more advanced multi-view image processing techniques^{35,36}, our camera will have a wider range of uses.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Code availability

The codes are available from the main developer (lilinh@sitp.ac.cn) upon reasonable request. We also uploaded the code on GitHub(<https://github.com/xiaolibao123/DCHXEMD.git>).

Received: 8 February 2024; Accepted: 4 November 2024

Published online: 11 November 2024

References

- Zhang, L., Zhan, H., Liu, X., Xing, F. & You, Z. A wide-field and high-resolution lensless compound eye microsystem for real-time target motion perception. *Microsyst. Nanoeng.* **8**, 83 (2022).
- Qu, P. et al. A simple route to fabricate artificial compound eye structures. *Opt. Express* **20**, 5775–5782 (2012).
- Zhang, B., Chen, G., Cheng, M.M.-C., Chen, J.C.-M. & Zhao, Y. Motion detection based on 3D-printed compound eyes. *OSA Continuum* **3**, 2553–2563 (2020).
- Hassenstein, B. & Reichardt, W. System theoretical analysis of time, sequence and sign analysis of the motion perception of the snout-beetle *Chlorophanus*. *Z. Naturforsch. B* **11**, 513–524 (1956).
- Pallus, A. C., Fleishman, L. J. & Castonguay, P. M. Modeling and measuring the visual detection of ecologically relevant motion by an *Anolis* lizard. *Journal of Comparative Physiology A* **196**, 1–13 (2010).
- Eichner, H., Joesch, M., Schnell, B., Reiff, D. F. & Borst, A. Internal structure of the fly elementary motion detector. *Neuron* **70**, 1155–1164 (2011).
- Wang, H., Peng, J. & Yue, S. Bio-inspired small target motion detector with a new lateral inhibition mechanism. in *2016 International Joint Conference on Neural Networks (IJCNN)* 4751–4758 (2016). <https://doi.org/10.1109/IJCNN.2016.7727824>.
- Wang, H., Peng, J. & Yue, S. An improved LPTC neural model for background motion direction estimation. in *2017 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* 47–52 (2017). <https://doi.org/10.1109/DEVLRN.2017.8329786>.
- Wang, H., Peng, J. & Yue, S. A feedback neural network for small target motion detection in cluttered backgrounds. in *Artificial Neural Networks and Machine Learning – ICANN 2018* (eds. Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L. & Maglogiannis, I.) 728–737 (Springer, 2018).
- Wang, L. et al. Drosophila-inspired 3D moving object detection based on point clouds. *Information Sciences* **534**, 154–171 (2020).
- Maisak, M. S. et al. A directional tuning map of *Drosophila* elementary motion detectors. *Nature* **500**, 212–216 (2013).
- Winding, M. et al. The connectome of an insect brain. *Science* **379**, eadd9330 (2023).
- Borst, A., Haag, J. & Mauss, A. S. How fly neurons compute the direction of visual motion. *Journal of Comparative Physiology A* **206**, 109–124 (2020).
- James, J. V., Cazzolato, B. S., Grainger, S. & Wiederman, S. D. Nonlinear, neuronal adaptation in insect vision models improves target discrimination within repetitively moving backgrounds. *Bioinspiration & Biomimetics* **16**, 066015 (2021).
- Wu, Z. & Guo, A. Bioinspired figure-ground discrimination via visual motion smoothing. *PLOS Computational Biology* **19**, e1011077 (2023).
- Bae, B. et al. Stereoscopic artificial compound eyes for spatiotemporal perception in three-dimensional space. *Science Robotics* **9**, eadl3606 (2024).
- Yu, Y. et al. Design of cooled infrared bionic compound eye optical system with large field-of-view. in *Earth and Space: From Infrared to Terahertz (ESIT 2022)* vol. 12505 125050L (2023).
- Wang, X. et al. Research on key technology of cooled infrared bionic compound eye camera based on small lens array. *Scientific Reports* **14**, 11094 (2024).
- Sun, B., Sang, N., Wang, Y. & Zheng, Q. Motion detection based on biological correlation model. in *Advances in Neural Networks—ISNN 2010* (eds. Zhang, L., Lu, B.-L. & Kwok, J.) 214–221 (Springer, 2010).
- Egelhaaf, M. & Reichardt, W. Dynamic response properties of movement detectors: Theoretical analysis and electrophysiological investigation in the visual system of the fly. *Biological Cybernetics* **56**, 69–87 (1987).
- Liu, F. et al. Infrared small and dim target detection with transformer under complex backgrounds. *IEEE Transactions on Image Processing* **32**, 5921–5932 (2023).
- Hu, Y., Ma, Y., Pan, Z. & Liu, Y. Infrared dim and small target detection from complex scenes via multi-frame spatial-temporal patch-tensor model. *Remote Sensing* **14**, 66 (2022).
- Pan, S., Zhang, S., Zhao, M. & An, B. Infrared small target detection based on double-layer local contrast measure. *Acta Photonica Sinica* **49**, 184–192 (2020).
- Wei, Y., You, X. & Li, H. Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recognition* **58**, 216–226 (2016).
- Aghaziyarati, S., Moradi, S. & Talebi, H. Small infrared target detection using absolute average difference weighted by cumulative directional derivatives. *Infrared Physics & Technology* **101**, 78–87 (2019).
- Moradi, S., Moallem, P. & Sabahi, M. F. Fast and robust small infrared target detection using absolute directional mean difference algorithm. *Signal Processing* **177**, 107727 (2020).
- Lou, C., Zhang, Y. & Yin, J. Small target detection method based on Robinson–Guard filter and pixel convergence. *Acta Optica Sinica* **40**, 1504001 (2020).
- Qu, X., Chen, H. & Peng, G. Novel detection method for infrared small targets using weighted information entropy. *Journal of Systems Engineering and Electronics* **23**, 838–842 (2012).
- Zhao, E., Zheng, W., Li, M., Sun, H. & Wang, J. Infrared small target detection using local component uncertainty measure with consistency assessment. *IEEE Geoscience and Remote Sensing Letters* **19**, 1–5 (2022).
- Zhao, E. et al. A fast detection method using anisotropic guidance for infrared small target under complex scenes. *IEEE Geoscience and Remote Sensing Letters* **20**, 1–5 (2023).
- Du, P. & Hamdulla, A. Infrared moving small-target detection using spatial-temporal local difference measure. *IEEE Geoscience and Remote Sensing Letters* **17**, 1817–1821 (2020).
- Zhu, H., Guan, Y., Deng, L., Li, Y. & Li, Y. Infrared moving point target detection based on an anisotropic spatial-temporal fourth-order diffusion filter. *Computers & Electrical Engineering* **68**, 550–556 (2018).
- Luo, Y., Li, X., Yan, Y. & Xia, C. Spatial-temporal tensor representation learning with priors for infrared small target detection. *IEEE Transactions on Aerospace and Electronic Systems* **59**, 9598–9620 (2023).
- Deng, L., Zhu, H., Tao, C. & Wei, Y. Infrared moving point target detection based on spatial-temporal local contrast filter. *Infrared Physics & Technology* **76**, 168–173 (2016).
- Xu, G. J. W., Guo, K., Park, S. H., Sun, P. Z. H. & Song, A. Bio-inspired vision mimetics toward next-generation collision-avoidance automation. *The Innovation* **4**, 100368 (2023).
- Wang, Q. et al. Large-scale generative simulation artificial intelligence: The next hotspot. *The Innovation* **4**, 100516 (2023).

Acknowledgements

This research was funded by the Shanghai Institute of Technical Physics, and the National Pre-research Program during the 14th Five-Year Plan (No. 514010405).

Author contributions

L.L. performed the simulations and wrote the manuscript. X.W., Juan.Yue and T.L. supervised the project. L.L. and X.W. designed the work. S.G., Y.Y. and S.H. reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.G., Y.Y. or H.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024