Optics Letters

Efficient autofocusing in digital holography using the Tanimoto coefficient

Jie Liu,[1,†] Ye Liu,[1,†] Haiyan Ou,[1,2,*] and Edmund Y. Lam[3]

[1]Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, Shenzhen 518110, China
[2]School of Physics, University of Electronic Science and Technology of China, Chengdu 611731, China
[3]Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong SAR, China
[†]These authors contributed equally to this work.
*ouhaiyan@uestc.edu.cn

Received 4 June 2025; revised 23 July 2025; accepted 24 July 2025; posted 24 July 2025; published 22 August 2025

This manuscript proposes a novel, to the best of our knowledge, autofocusing method for digital holography using the Tanimoto coefficient (TC) to assess similarity between dual-channel reconstructions (real and imaginary components). Additionally, connected component analysis is used to extract regions of interest (ROI), improving autofocus accuracy. Unlike existing similarity-based methods, this dual-channel TC approach eliminates the need for RGB information and offers over 90 times faster computation than the TR-MUSIC method. The method's effectiveness is validated through simulations and experimental trials. © 2025 Optica Publishing Group. All rights, including for text and data mining (TDM), Artificial Intelligence (AI) training, and similar technologies, are reserved.

https://doi.org/10.1364/OL.569530

Digital holography (DH) is a non-invasive three-dimensional (3D) imaging technique that encodes the wavefront information of 3D objects in the form of two-dimensional (2D) holograms [1]. This technique has applications across various fields, including biomedical imaging [2], micrometer-sized particles imaging [3], and 3D imaging using a single-pixel detector [4].

To extract the wavefield of an object from a hologram, reconstruction is essential. A key challenge in this process is the accurate determination of the object's spatial axial position. To address this issue, various autofocusing algorithms have been proposed, including sharpness metric-based methods [5–9] and deep learning approaches [10–14].

In this manuscript, we propose a novel method that leverages the Tanimoto coefficient (TC) for autofocusing. This approach separates the hologram into real and imaginary components and uses the Tanimoto coefficient to compare reconstructed images from both channels to locate the focus position where similarity peaks. Additionally, the connected component method mitigates defocus noise, improving autofocus result reliability.

Various optical systems can be employed to record holograms of objects. This manuscript focuses on optical scanning holography (OSH), which utilizes active 2D scanning to capture holograms of 3D objects [15]. OSH is widely applied in fields such as microscopy [16], remote sensing [17], and image encryption [18].

Figure 1(a) illustrates the specific configuration of OSH system. Two beams of light with frequencies ω_0 and $\omega_0 + \Omega$ pass through their respective pupils and lenses before being combined at the second beam splitter (BS2). The synthesized light then scans the object using an XY scanning mirror. The light carrying the object's information is converted into an electrical signal by a photodetector (PD). This electrical signal is subsequently modulated and stored on a computer. In OSH, the two pupil functions are typically set as $p_1 = 1$ and $p_2 = \delta(x, y)$. Under these assumptions, the system's spatial impulse response can be modeled as $h(x, y; z) = \left[(-jk_0)/2\pi z\right] \cdot exp\left[jk_0(x^2 + y^2)/2z\right]$. Here, x, y, and z are spatial coordinates, j denotes the imaginary unit, and $k_0 = 2\pi/\lambda$ is the wavenumber, with λ representing the wavelength.

Assuming the 3D object has a complex amplitude distribution denoted by $\Gamma(x, y, z)$, the hologram recorded by the OSH system can be expressed as

$$\begin{aligned} H(x, y) &= \int_{-\infty}^{\infty} (|\Gamma(x, y; z)|^2 \otimes h(x, y; z))dz \\ &= \sum_{i=1}^{N} (|\Gamma(x, y; z_i)|^2 \otimes h(x, y; z_i)), \end{aligned} \tag{1}$$

where $\otimes$ represents convolution and $|\cdot|$ denotes the modulus operation. As illustrated in Eq. (1), the object can be discretized into N axial layers, with z_i indicating the position of the i-th layer. Following standard OSH convention [16], the real and imaginary components correspond to the sine and cosine holograms, respectively, representing quadrature and in-phase interference components.

After recording, the reconstruction process is performed to retrieve the 3D structure of the object. Based on Eq. (1), the reconstructed image at the k-th depth layer can be represented as

$$\begin{aligned} R(x, y; z_k) &= H(x, y) \otimes h^*(x, y; z_k) \\ &= |\Gamma(x, y; z_k)|^2 + \sum_{i \neq k}^{N} |\Gamma(x, y; z_i)|^2 \otimes h(x, y; z_i) \\ &\quad \otimes h^*(x, y; z_k), \end{aligned} \tag{2}$$

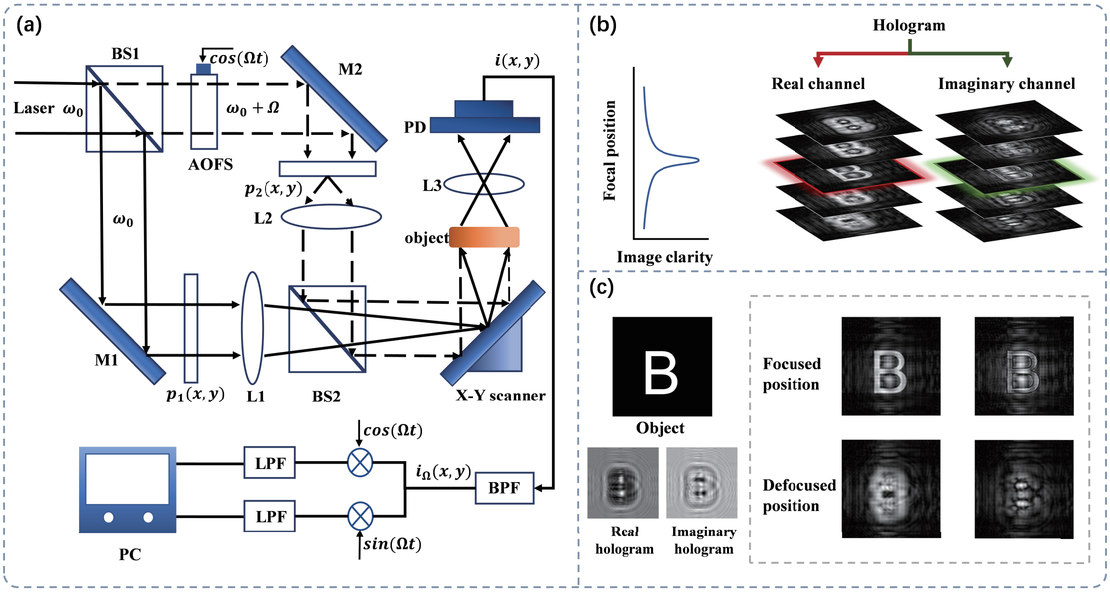


Fig. 1. (a) OSH system. AOFS, acousto-optic frequency shifter; BS, beam splitter; BPF, band-pass filter; L, lens; LPF, low-pass filter; M, mirror; $p(x,y)$, pupil; PD, photodetector. (b) Schematic diagram of the sharpness evaluation curve based on dual-channel separated holograms. (c) Reconstruction results of dual-channel holograms at both focused and defocused positions.

where $*$ represents the conjugation operation. The equation contains two terms: the first is the in-focus contribution from the target layer at depth z_k, and the second is defocus noise from contributions of other layers.

From Eq. (2), it is clear that identifying the axial position of the object is a crucial step in the reconstruction process. To address this, we propose the TC method. The Tanimoto coefficient, which can be regarded as a generalization of the Jaccard index, serves as a similarity metric [19,20]. Given two vectors A and B, the Tanimoto coefficient is computed as

$$\mathrm{TC}(A,B)=\frac{A\cdot B}{\|A\|^2+\|B\|^2-A\cdot B}, \tag{3}$$

where $\|\cdot\|^2$ denotes the squared norm of a vector. The Tanimoto coefficient offers several advantages, including a simple underlying principle and ease of computation, making it widely utilized in fields such as virtual screening [21], antenna selection [22], and species co-occurrence evaluation [23]. Furthermore, it can be employed for autofocusing by analyzing the pixel differences among the RGB channels of a color image [24].

Since holograms are recorded as grayscale images without RGB channels, directly applying traditional TC-based autofocusing methods becomes nontrivial. To overcome this, we propose treating the real and imaginary parts of a complex hologram as two independent channels, analogous to RGB components, enabling the use of the Tanimoto coefficient for grayscale holographic autofocusing.

Figure 1(b) shows the schematic diagram of the sharpness evaluation curve based on dual-channel separated holograms. Figure 1(c) presents the reconstruction results of dual-channel holograms at both focused and defocused positions. As the reconstruction approaches the focal plane, edge sharpness, high-frequency detail visibility, and inter-channel similarity all increase. At the focal plane ($z=z_k$), the reconstructed dual-channel components (R_r and R_i) have the highest structural similarity because: (1) The in-focus term dominates, resulting in coherent structural features in both real and imaginary channels; (2) The defocus noise is minimized, reducing random interference that lowers inter-channel similarity; (3) The phase alignment ensures constructive interference in both channels.

As defocus increases ($z\neq z_k$), the growing phase mismatch causes progressive divergence between R_r and R_i, thereby reducing their Tanimoto coefficient:

$$\mathrm{TC}(R_r,R_i)=\frac{R_r\cdot R_i}{\|R_r\|^2+\|R_i\|^2-R_r\cdot R_i}\xrightarrow{z\to z_k}\text{maximum}. \tag{4}$$

This theoretical expectation is validated by the peak in Fig. 1(b) and the reconstruction contrast in Fig. 1(c).

With increasing defocus, the phase mismatch between the propagation function and the object's complex amplitude intensifies, causing the reconstructed real and imaginary channel images to diverge increasingly in their structural features.

Single-slice object holograms can be directly autofocused using TC, whereas multi-layer structures introduce defocus noise that impairs performance. To address this issue, connected component (CC) analysis is employed to extract the region of interest (ROI). CC analysis is typically applied to binary images, where connectivity is determined by checking whether neighboring pixels belong to the foreground (usually assigned a value of 1). It is a fundamental technique widely used in object detection, segmentation, and image analysis [25,26].

Figure 2 presents the flowchart of the proposed TC-based autofocusing method, which includes the following steps:

(1) Hologram Generation: Generate the hologram $H(x,y)$.
(2) Channel Separation: Decompose complex hologram $H(x,y)$ into its real part $H_r(x,y)$ and imaginary part $H_i(x,y)$ to form dual channels.
(3) Images Reconstruction: Perform image reconstruction to obtain $R_r(x,y;z_k)$ and $R_i(x,y;z_k)$ from the real and imaginary components, respectively, at each candidate depth z_k.

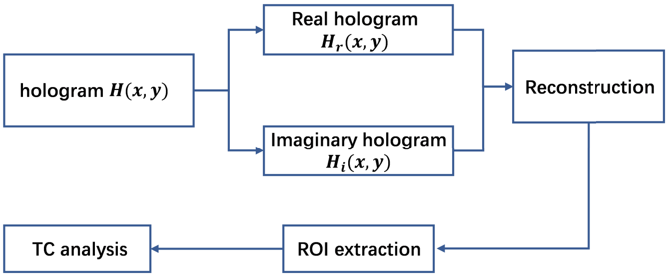


Fig. 2. Flowchart of the TC algorithm.

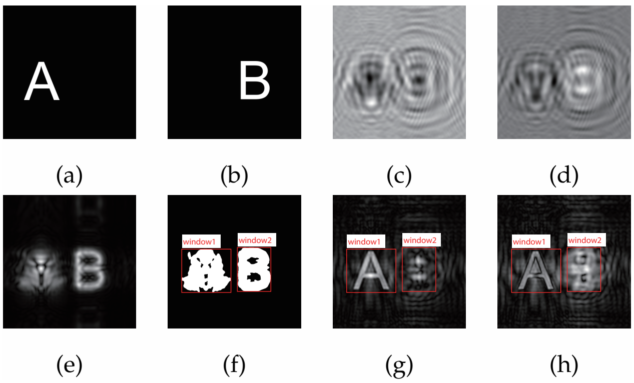


Fig. 3. (a), (b) First and second sections of the simulated object. (c), (d) Real and imaginary parts of the simulated hologram. (e) Hologram reconstruction image. (f) ROI detection results. (g), (h) Illustrations of the segmentation positions in the real and imaginary part reconstruction images.

(4) ROI Extraction: Apply connected component (CC) analysis to the reconstructed complex-valued image to determine the region of interest (ROI) and extract the corresponding regions from R_r and R_i.

(5) TC Analysis: Binarize the ROI-extracted images and compute the Tanimoto coefficient to evaluate the similarity between channels.

Iteratively applying Steps (3)–(5) over a defined axial range calculates TC at each depth, forming an autofocusing curve; the axial position of its maximum is the object's optimal focus.

A double-layer sliced object is simulated to illustrate the proposed method. The object consists of two sections located at depths z_a and z_b, as shown in Figs. 3(a) and 3(b). A hologram is generated according to Eq. (1), with its real and imaginary parts presented in Figs. 3(c) and 3(d). The CC method is applied to the reconstructed image, as shown in Fig. 3(e), to detect ROIs. In this method, threshold 1 is employed for image binarization, and threshold 2 is utilized to eliminate small connected components. The detected ROIs are marked by red bounding boxes in Fig. 3(f). These ROIs are then used to segment the dual-channel (real and imaginary) reconstruction images, as shown in Figs. 3(g) and 3(h). Window 1 fully encompasses Character A in both reconstructed images, demonstrating successful ROI extraction. The segmented regions are further binarized using threshold 3, and the resulting sub-images are evaluated using the Tanimoto coefficient to quantify their similarity.

The feasibility of the proposed method is verified through simulations. In the setup, the laser wavelength is set to 632.8 nm, and both lenses L1 and L2 have focal lengths of 50 mm.

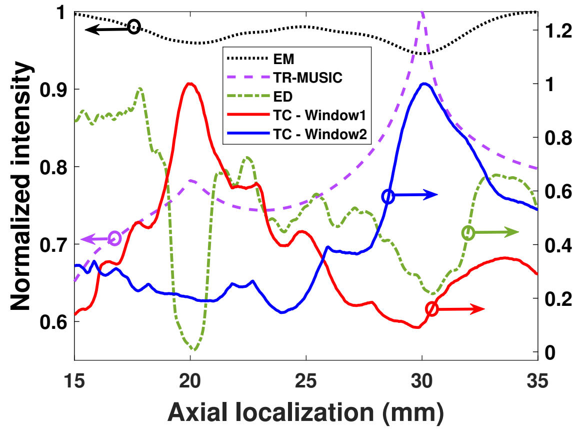


Fig. 4. Parameters calculated along the z-axis obtained by different methods.

Table 1. Autofocusing Results in Simulation

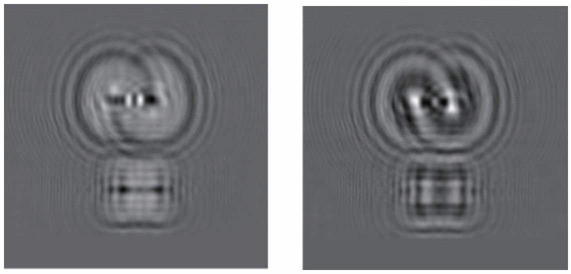
Method	RFP	EM	TR-Music	ED	TC
z_1 (mm)	20	20.35	20.00	20.10	20.05
z_2 (mm)	30	30.00	30.00	30.35	30.05
Time (s)	/	3.4	934.3	10.8	7.6

The simulated double-layer sliced object, corresponding to Figs. 3(a) and 3(b), comprises two sections located at axial depths of $z_a = 20$ mm and $z_b = 30$ mm, respectively. The reconstructed position is determined to be 27.6 mm. During ROI extraction, two thresholds are employed: threshold1 $= \{\max(R(x, y; z_k)) - \min(R(x, y; z_k))\}/5$, where $\max(\cdot)$ and $\min(\cdot)$ represent the maximum and minimum pixel values in the image; and threshold2 $= 0.1 \times s_{max}$, with s_{max} denoting the area of the largest connected component. After ROI extraction, a fixed threshold of threshold3 $= 0.4$ is applied for binarization in the subsequent TC analysis. Autofocusing is conducted over the axial range of 15–35 mm with a step size of 0.05 mm, based on the parameters defined above.

Figure 4 presents the autofocusing results obtained using four methods: Time-Reversal Multiple Signal Classification (TR-MUSIC) [9], Entropy Minimization (EM) [27], Edge Detection (ED) [28], and the proposed Tanimoto Coefficient (TC)-based method. For better comparison, all methods autofocus over the same axial range with uniform step size: TR-MUSIC and TC use local maxima for detected positions, EM and ED use local minima. Localization results and computation times are in Table 1 (RFP, real focus position). TR-MUSIC has higher accuracy but longer computation time; among the other three, TC is slightly more accurate than EM and ED.

The feasibility of the proposed method is further validated through experiments using data sourced from Zhang *et al.* [28]. The experiment employs a laser wavelength of 632.8 nm. Two sections of the object, each sampled at 500×500 pixels, are located at distances of $z_1 = 87$ cm and $z_2 = 107$ cm. Figures 5(a) and 5(b) display the real and imaginary parts of the hologram, respectively. The threshold settings used in the experiment are consistent with those applied in the preceding simulation. All four methods perform autofocusing over the range of 80–115 cm with a step size of 0.05 cm.

Autofocus results of all experimental methods are shown in Fig. 6 and Table 2: TR-MUSIC has high localization accuracy



(a) (b)

Fig. 5. (a), (b) Real part and imaginary part of the recorded hologram.

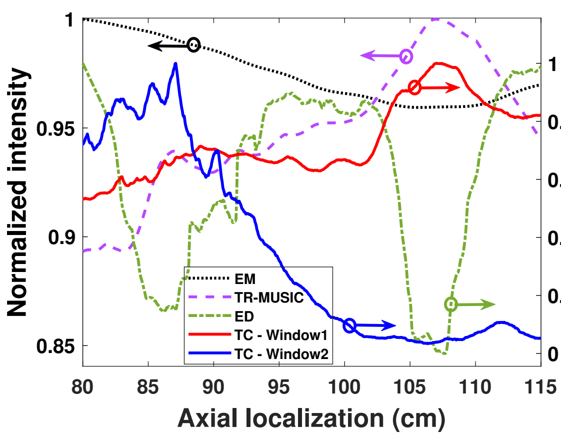


Fig. 6. Parameters calculated along the z-axis obtained by different methods.

Table 2. Autofocusing Results in Experiment

Method	RFP	EM	TR-Music	ED	TC
z_1 (cm)	87	*Failed*	86.95	86.30	87.10
z_2 (cm)	107	105.75	107.04	107.65	107.05
Time (s)	/	4.5	1416.9	14.6	15.2

but high computation time; EM detects only a single slice position; among the remaining methods, TC has better accuracy than ED.

The TC method utilizes window information derived from the connected domain to extract the ROI prior to autofocusing. The resulting focal positions are 87.10 cm and 107.05 cm, respectively. Furthermore, the window information can be leveraged to suppress defocus noise in the reconstructed images by setting pixels outside the identified ROIs to zero. This process yields the reconstructed images shown in Figs. 7(c) and 7(d). The distinct visibility of the characters in the reconstructed image highlights the accuracy of the autofocusing method. Simultaneously reconstructed at the location obtained by the ED method, the reconstruction results are shown in Figs. 7(a) and 7(b). The two methods yield nearly identical reconstruction outcomes.

In conclusion, the proposed method presents a novel application of the TC to the autofocusing problem in grayscale holography by decomposing holograms into real and imaginary components. Additionally, connected component analysis is integrated to isolate ROIs and effectively reduce the impact of defocus noise on the autofocusing results. It should be noted

(a) (b) (c) (d)

Fig. 7. Reconstructed results at the predicted focal positions using different methods: (a), (b) ED; (c), (d) TC.

that the proposed method is applicable to various holographic configurations capable of recording complex-valued holograms. Although effective for amplitude-modulated objects, the method is not directly applicable to pure phase samples due to its reliance on intensity-derived binarization. While the TC method achieves slightly lower accuracy than the TR-MUSIC, it offers a computational speed-up of over 90 times. This demonstrates its potential as a highly efficient strategy for tackling autofocusing challenges in digital holography.

Funding. Basic and Applied Basic Research Foundation of Guangdong Province (2024A1515510017).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data is available upon request to the corresponding author.

REFERENCES

1. B. Javidi, A. Carnicer, A. Anand, *et al.*, Opt. Express **29**, 35078 (2021).
2. B. K. Goud, D. Shinde, D. Udupa, *et al.*, Opt. Lasers Eng. **114**, 1 (2019).
3. M. J. Berg, J. Jacquot, J. M. Harris, *et al.*, Opt. Lett. **49**, 2653 (2024).
4. L. Martínez-León, P. Clemente, Y. Mori, *et al.*, Opt. Express **25**, 4975 (2017).
5. M. Liebling and M. Unser, J. Opt. Soc. Am. **21**, 2424 (2004).
6. F. Dubois, C. Schockaert, N. Callens, *et al.*, Opt. Express **14**, 5895 (2006).
7. Z. Ren, N. Chen, and E. Y. Lam, Opt. Lett. **42**, 1720 (2017).
8. Y. Zhang, H. Wang, Y. Wu, *et al.*, Opt. Lett. **42**, 3824 (2017).
9. H. Ou, Y. Wu, E. Y. Lam, *et al.*, Opt. Express **26**, 3756 (2018).
10. Y. Wu, Y. Rivenson, Y. Zhang, *et al.*, Optica **5**, 704 (2018).
11. T. Pitkäaho, A. Manninen, and T. J. Naughton, Appl. Opt. **58**, A202 (2019).
12. L. Huang, T. Liu, X. Yang, *et al.*, ACS Photonics **8**, 1763 (2021).
13. S. Cuenat, L. Andréoli, A. N. André, *et al.*, Opt. Express **30**, 24730 (2022).
14. N. Madali, A. Gilles, P. Gioia, *et al.*, Opt. Express **31**, 4199 (2023).
15. T.-C. Poon, J. Opt. Soc. **13**, 406 (2009).
16. E. Y. Lam, X. Zhang, H. Vo, *et al.*, Appl. Opt. **48**, H113 (2009).
17. T. Kim, T.-C. Poon, and G. Indebetouw, Opt. Eng. **41**, 1331 (2002).
18. L.-Z. Zhang, X. Zhou, D. Wang, *et al.*, Opt. Eng. **59**, 093102 (2020).
19. M. Levandowsky and D. Winter, Nature **234**, 34 (1971).
20. D. Bajusz, A. Rácz, and K. Héberger, J. Cheminformatics **7**, 20 (2015).
21. P. Willett, Biochem. Soc. Trans. **31**, 603 (2003).
22. Y. Haifen and L. Guangjun, J. Syst. Eng. Electron. **19**, 624 (2008).
23. N. C. Chung, B. Miasojedow, M. Startek, *et al.*, BMC Bioinformatics **20**, 644 (2019).
24. X. Zhou, P. Xiong, D. Chi, *et al.*, Opt. Lett. **47**, 3752 (2022).
25. H. Lifeng, C. Yuyan, and S. Kenji, IEEE Trans. Pattern Anal. Mach. Intell. **20**, 2122 (2011).
26. L. He and Y. Chao, IEEE Trans. Image Process. **24**, 2725 (2015).
27. Z. Ren, N. Chen, A. Chan, *et al.*, in *Digital Holography and Three-Dimensional Imaging* (2015), paper DT4A–4.
28. X. Zhang, E. Y. Lam, T. Kim, *et al.*, Opt. Lett. **34**, 3098 (2009).