



In-Sensor Computing for Rapid Image Focusing during Liquid Lens Oscillation

Yanan Liu^{a,b}, Zhiguo Yang^a, Huixin Zhong^c, Rui Fan^d, Ruidong Ma^e, Xiupeng Shi^{a,*}, Jianhua Zhang^{a,b}

^a School of Microelectronics, Shanghai University, Jiading, Shanghai, 201800, China

^b Shanghai Key Laboratory of Chips and Systems for Intelligent Connected Vehicle, Shanghai University, Shanghai, 201800, China

^c HeXie Academy & College of Industry-Entrepreneurs, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, Jiangsu, China

^d College of Electronics & Information Engineering, Tongji University, Jiading, Shanghai, 201804, China

^e School of Computing and Digital Technologies, Sheffield Hallam University, Sheffield, S9 2AA, UK

ARTICLE INFO

Keywords:

In-situ Vision Processor

Liquid lens

Autofocus

In-Sensor computing

Embedded vision system

ABSTRACT

Focus-tunable liquid lenses are a cornerstone in modern optical systems, enabling rapid and energy-efficient focus adjustments. However, the inherent fluidic dynamics of these lenses induce transient oscillations following focus adjustments, necessitating stabilization periods that introduce latency and hinder real-time autofocus performance. To overcome these limitations, this study presents a novel compact autofocus system leveraging an In-situ Vision Processor (IVP) coupled with a focus-tunable liquid lens. Our method introduces in-sensor parallel sharpness measurement and post-processed focused image acquisition during the liquid lens's oscillation phase directly on the sensor, eliminating the need for the liquid lens to stabilize before capturing images in most traditional liquid lens based vision systems. Consequently, the system is able to capture optimal in-focus images for both static and high-speed dynamic objects due to its high capture-process sampling speed. The 'global focus' image can also be generated for spatial objects while the lens sweeps between diopter extremes, enhancing the system's suitability for diverse applications. This technology operates with low power consumption (around 2.0 W), low latency (<2 ms for single image capturing and analysis), and low-cost (\$10), making it suitable for mobile and portable applications.

1. Introduction

Focus-tunable liquid lenses represent a pivotal innovation in optical technology, offering the ability to rapidly adjust focus at various distances. This adaptability is crucial for maintaining image sharpness in various scenarios, thereby ensuring the generation of high-quality images critical for accurate visual analysis and subsequent automated operations. In the realm of instrumentation and measurement, the ability to precisely and rapidly control focus is paramount for applications ranging from microscopy and industrial inspection to consumer electronics [1–3]. These lenses are known for their minimal power requirements and quick response, which are vital for energy-efficient and high-speed applications [4]. However, lens oscillation following an external stimulus is an inherent challenge in liquid lens technology due to the fluidic dynamics and the elastic properties of the lens materials. These oscillations arise from the interplay between the liquid's inertia, membrane elasticity, and damping forces, making it difficult for the lens to achieve a stable focus instantaneously [5–8]. Consequently,

existing liquid-lens-based autofocus systems typically rely on waiting for the liquid lens to settle into a stable state before capturing images and proceeding with further processing. However, this waiting period not only introduces latency but also limits the system's ability to perform rapid and continuous autofocus adjustments, which are essential for applications requiring real-time imaging and quick focus shifts. To address above-mentioned limitations, our work introduces the integration of a focus-tunable liquid lens with a high-speed in-sensor parallel computing architecture, utilizing a In-situ Vision Processor (IVP). This integration allows for simultaneous image capture and sharpness evaluation during the lens's short transient oscillatory phase. By leveraging the parallel computational capabilities of the IVP, our system can enhance image feature extraction and analysis in real-time, even while the lens is still adjusting its focus. This approach mitigates the adverse effects of oscillations by enabling continuous image processing without waiting for the lens to fully stabilize.

* Corresponding author.

E-mail address: sxp@shu.edu.cn (X. Shi).

<https://doi.org/10.1016/j.sna.2025.116868>

Received 7 March 2025; Received in revised form 28 May 2025; Accepted 6 July 2025

Available online 25 July 2025

0924-4247/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

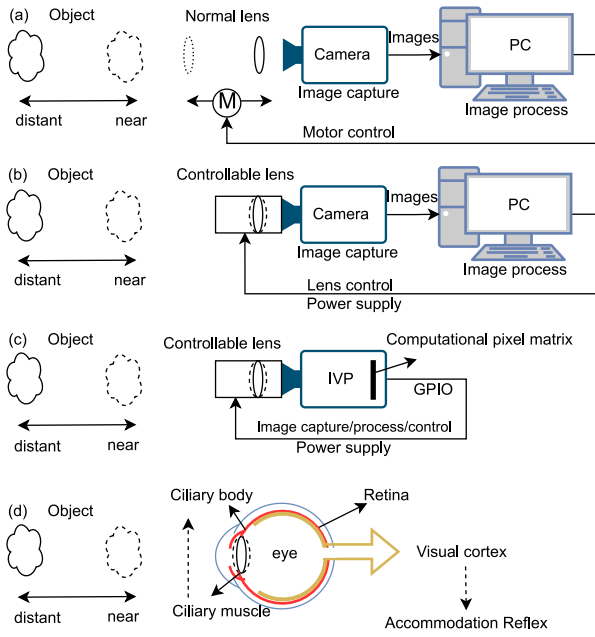


Fig. 1. Comparison of image-based autofocus systems. (a) Autofocus system employing motor control for adjusting the camera's focus relative to object distance [12,13]. (b) Autofocus system using a controllable liquid lens to adjust focus [14]. (c) Our proposed bio-inspired autofocus system integrating a compact IVP with a tunable-focus liquid lens, enabling image capture, processing, and focus control within one single compact module. (d) Schematic of the human eye accommodation system demonstrating the biological counterpart of autofocus involving the ciliary muscle and accommodation reflex [15].

Recent progress in focus-tunable lens spans a broad performance envelope. Kang et al. provide a survey of variable optical elements showing response times that range from tens of milliseconds down to the microsecond regime [9]. Among these platforms, the ultrasonic Tunable-Acoustic-Gradient (TAG) lens achieves outstanding temporal resolution: operating near its 70 kHz resonance. Such bandwidth has enabled millisecond-scale 3-D topography via encoded-search focal scanning (ESFS) [10] and single-shot extended-depth-of-field imaging in the commercial TAGLENS-T1 module [11]. This capability, however, is accompanied by a continuous ultrasonic drive of 20 W with extra illumination light source (≈ 20 W), external controller—constraints that remain prohibitive for mobile, UAV, and wearable systems [11]. Our present work addresses a complementary bottleneck: computational latency in low-power electrowetting lenses that already dominate embedded cameras. By relocating the focus-metric evaluation from controller onto IVP, we compress the system delay to 2 ms.

As depicted in Fig. 1(a), conventional vision-based autofocus systems, rely on a camera to capture images and a computer to process these images to evaluate focus quality. These systems typically control the lens position via motors through software APIs, a process that often results in significant system latency. This latency is caused by several factors: the mechanical adjustment of the lens position, the time required for image capture and transfer to the processing unit, and the slow communication between the computer and the lens system. Image-based autofocus methods with a liquid lens, like the one illustrated in Fig. 1(b), iteratively change the lens curvature to enhance specific image features, particularly image sharpness, under the condition that the scene remains static. This requires capturing multiple frames at different focus positions, usually requiring 4–7 different levels determined by a contrast-based search algorithm [16,17]. Each change in focus

level demands a pause, allowing the liquid lens to stabilize—a period that can last from 10 to 100 ms, varying with the type of liquid lens used and contributing further to the overall system delay. As can be summarised from Fig. 1 (a) (b), traditional autofocus imaging systems often incur high power costs and increased weight due to their complex setups, impeding their deployment in low-cost, energy-efficient applications like portable and embedded systems. These factors highlight limitations in instrumentation design for applications with stringent size, weight, and power (SWaP) constraints [18]. In contrast, the human visual system achieves real-time autofocus effortlessly through the ciliary body's curvature adjustments controlled by the accommodation reflex, a process independent of conscious cognitive effort [19] (Fig. 1(d)). Drawing inspiration from this biological mechanism, we propose a bio-inspired, compact, low-cost, in-sensor real-time autofocus system. This system processes all information on the IVP's CPM, bypassing the need for external computing resources or specialized software that could introduce extra time delay, akin to retinal cells processing visual data before sending the processed useful information to visual cortex (Fig. 1 (c)). Furthermore, unlike traditional autofocus, which often requires multiple steps to adjust the lens for optimal focus [4,14,20], resulting in slower response times and increased power consumption, our approach needs only one single trigger to the liquid lens. The autofocus system capitalises on the lens's oscillatory behaviour to capture and evaluate images in rapid succession, harnessing the IVP's high-speed, parallel processing capabilities. Consequently, this contribution not only dramatically accelerates focused image capturing but also simplifies the hardware architecture, enabling the real-time capture of sharply focused images within a power-efficient and lightweight embedded system.

The main contributions of this work are summarised as follows: (1) High-speed image capture and process for focus during lens oscillation: We introduce a new approach for IVP-based high-speed autofocus system that captures and processes images during the oscillatory phase of a liquid lens, triggered by a single input step (shown in Eq. (1) where the $V_{initial}$ and V_{final} represent the voltage before and after the trigger at time t_0 , respectively).

$$V(t) = \begin{cases} V_{initial} & \text{for } t < t_0 \\ V_{final} & \text{for } t \geq t_0 \end{cases} \quad (1)$$

This enables the acquisition of sharp edge images within the short oscillation interval (usually within 50 ms according to the type of the liquid lens). In addition, we developed a 'global focus' technique that assembles an in-focus composite image for spatial objects by stacking individual focus responses obtained during the lens' oscillatory motion. (2) In-Sensor computing method for focus evaluation: Our research presents an innovative, high-speed method for parallel in-sensor image contrast measurement for real-time focus evaluation, requiring only 0.3 ms per evaluation. This facilitates the precise tracking of rapid contrast fluctuations throughout the lens' oscillation cycle. This methodology bypasses the traditional requirement for the lens to stabilise before image capture and processing, which is a common limitation in conventional machine vision systems. (3) Compact, energy-efficient, and low-cost autofocus system: This study proposes a compact, high-speed, and low-cost autofocus system that seamlessly integrates a IVP with a liquid lens. The resulting system is not only lightweight and energy-efficient but also outperforms conventional autofocus mechanisms in speed and portability. Our in-sensor processing operates at over 500 fps (frames per second), with a latency below 2 ms, including the time taken and transfer for image display. The image capture and processing are rapid enabling high-speed sampling of the lens' dioptric adjustment in response to electrical stimuli. Moreover, with a modest power consumption of around 2.0 W, the system sets a standard in energy efficiency for autofocus.

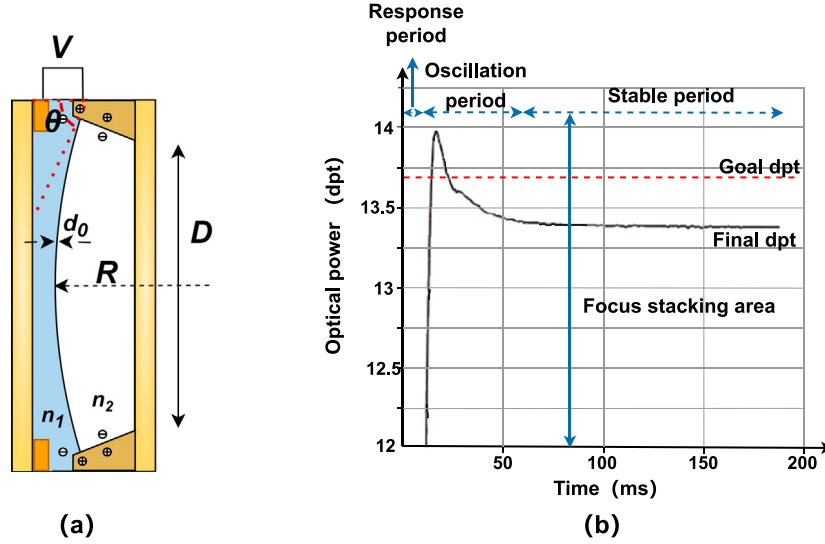


Fig. 2. (a) The mechanical structure of the dual-fluid liquid lens. (b) The process of liquid lens adjusts to the final diopter after stimulus. The focal power (diopter) is changing rapidly from initial diopter to its maximum, in which process the goal diopter associated with a focused image is passed. The whole time for the liquid lens to settle consists of response time and oscillation time. The Goal dpt corresponds to the optimal optical power required for achieving a focused image.

2. Design and working principle

Liquid lens oscillations, induced by rapid voltage changes [21], can transiently degrade image quality due to focus fluctuations. While typically a drawback requiring a stabilization period before image capture, we exploit this phenomenon to accelerate autofocus. Our system, leveraging rapid in-sensor imaging and computation, analyzes image contrast during these oscillations. This allows us to identify and capture the optimally focused image much faster than waiting for stabilization, effectively transforming a limitation into a key feature. Our approach is grounded in the principle that well-focused images exhibit higher high-frequency content compared to defocused images. Specifically, we utilize edge-based sharpness metrics, benefiting from established, efficient algorithms readily implementable in our hardware. This ensures both high performance and accurate sharpness measurement, critical for our application. Therefore, we propose a high-speed, parallel, in-sensor sharpness measurement method to directly identify optimal focus during lens curvature adjustment.

2.1. Focus-tunable liquid lens for autofocus

The liquid lens employed in this study is a dual-fluid zoom lens based on the electrowetting effect [22], with its structure illustrated in Fig. 2 (a). Inside the resonant cavity (shown in Fig. 2 (a)), two different liquids with refractive indices of n_1 and n_2 are present. By controlling the interface curvature of the dual liquids with voltage, the focus f of the liquid lens can be adjusted. The relationship between the focal length and the working voltage can be described by combining the Young-Lippmann equation [23] (Eq. (2)) and the lens fabrication equation [24] (Eq. (3)), as Eq. (4):

$$\cos \theta = \cos \theta_0 + \frac{\epsilon \epsilon_0}{2d_0 \gamma_{12}} \quad (2)$$

$$\frac{1}{f} = (n_1 - n_2) \frac{1}{R} \quad (3)$$

$$f = \frac{D}{2(n_1 - n_2) \cos \theta_0 + \frac{\epsilon \epsilon_0 (n_1 - n_2)}{d_0 \gamma_{12}} \cdot V^2} \quad (4)$$

where, D represents the clear aperture of the liquid lens; n_1 and n_2 are the refractive indices of the conductive liquid and the insulating liquid, respectively; θ denotes the contact angle between the liquids, with θ_0

being the initial contact angle when no voltage is applied; R is the radius of curvature of the interface; ϵ and ϵ_0 are the dielectric constants of the insulating film and vacuum, respectively; d_0 is the thickness of the insulating film; γ_{12} is the interfacial tension; and V is the applied voltage.

Liquid lens oscillatory model during focusing: The behaviour of the liquid lens under applied voltages which modifies its curvature and thus focal length, correlates with the oscillations observed in the lens's dioptric power over time, demonstrating the lens's dynamic focus adjustment in response to electrical stimuli. Fig. 2 (b) illustrates the dynamic response of a liquid lens when an external stimulus (such as voltage change) is applied. The response period is the time between the application of the stimulus and the beginning of the first significant change in focal power. This delay could be due to the time it takes for the electric field to affect the fluid, causing it to change the curvature of the lens. The dynamic process of the liquid lens triggered by voltage change can be modelled using a damped oscillation, which can be described by a function of the Eq. (5):

$$P(t) = P_0 + A e^{-\gamma t} \cos(\omega t + \phi) \quad (5)$$

where P_0 is the initial optical power (the starting point of the oscillation), A is the amplitude of the oscillation which is related to the voltage change, γ is the damping coefficient (which represents how quickly the oscillations decay), ω is the angular frequency of the oscillation, ϕ is the phase shift. Fig. 2(b) presents a temporal profile of the lens's focal power (diopter), demonstrating oscillatory/damping behaviour following an electrical stimulus. The focal power is shown to rapidly vary from its initial state, reaching a peak before settling towards the final diopter within a short period of time. This fluctuation, represented by the black curve, occurs around a red dotted line that signifies the goal diopter, indicating moments when the system has the potential to capture in-focus images. Fig. 2(b) conceptualises the lens's focus adjustment trajectory as it transitions from the nearest focus point to the farthest. The blue straight lines depict the range of the focus stacking area, converging on the goal focus position. This demonstrates the lens's capacity to transition through multiple focal points, enabling a composite image to be constructed from the sharpest segments captured during the oscillation, a process essential for achieving a 'global focus' image in spatial object imaging. The principle of this work assumes that during the transition from minimum

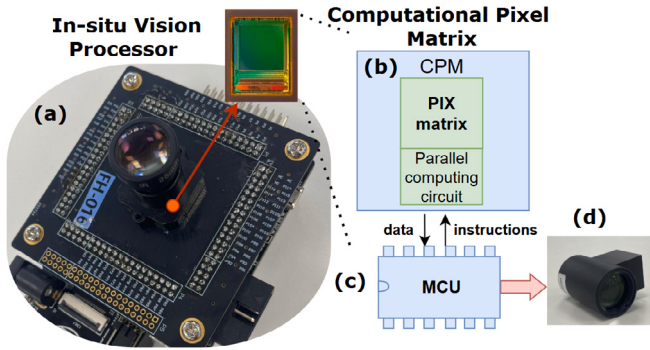


Fig. 3. IVP system and the liquid lens. (a) Qianxun IVP vision system. (b) The computational pixel matrix chip. (c) The micro-controller. (d) The focus-tunable liquid lens.

to maximum focal power, an optimal focal power exists at which a focused image can be obtained, provided that the IVP operates at a sufficiently high speed to capture and process the image effectively. It is important to note that the amplitude A of this oscillation, and thus the range of optical powers traversed, is dependent on the magnitude of the applied voltage step. Consequently, the range of object distances over which autofocus can be achieved in a single triggered event is determined by the specific parameters of this driving signal.

2.2. In-situ vision processor

IVP represents a significant advancement in imaging and processing technology, enabling processing signals directly on the sensor. This integration minimizes data transfer and external computation, greatly reducing latency and power usage. Notable IVP architectures like Qianxun (Fig. 3), High-Speed Vision Chip [25], and Eye-RIS [26] vision system have demonstrated parallel processing, improving speed and efficiency for tasks such as edge detection and object tracking [27] [28]. As depicted in Fig. 3, the Computational Pixel Matrix (CPM) is configured with 320×320 pixels, constituting the core of the Qianxun vision chip. Beyond the CPM, the architecture of the Qianxun chip integrates a Microcontroller Unit (MCU) dedicated to command instruction management and sequence data processing within a System-on-a-Chip (SoC) framework. This integration facilitates streamlined control and efficient data handling, enhancing the overall system performance and scalability. The SoC design ensures that the IVP can manage complex tasks autonomously, reducing the need for external processing units and minimizing latency. Qianxun IVP system achieves a power consumption of approximately 1.5 W, while delivering a computational efficiency of 1 Tera Operations Per Second per Watt (1 TOPS/W). This high efficiency underscores the system's suitability for power-constrained applications, making it an ideal solution for portable and mobile imaging devices. The advanced manufacturing process not only enhances performance but also contributes to the system's compactness and reliability.

2.3. High-speed in-sensor parallel sharpness measure

2.3.1. In-sensor parallel image edge extraction

Algorithm 1 Parallel Sobel Filter Implementation

Require: I : Input image
Require: T : Threshold value
Ensure: $focus_measure$: Focus evaluation output

- 1: // Compute vertical gradient G_x in parallel for each pixel:
- 2: $V \leftarrow I_{i-1,j} + 2 \times I_{i,j} + I_{i+1,j}$
- 3: $G_x \leftarrow V_{i,j-1} - V_{i,j+1}$
- 4: $|G_x| \leftarrow |G_x|$
- 5: // Compute horizontal gradient G_y in parallel for each pixel:
- 6: $H \leftarrow I_{i,j-1} + 2 \times I_{i,j} + I_{i,j+1}$
- 7: $G_y \leftarrow H_{i-1,j} - H_{i+1,j}$
- 8: $|G_y| \leftarrow |G_y|$
- 9: $Edge \leftarrow |G_x| + |G_y|$
- 10: // Apply threshold and compute focus measure
- 11: If $Edge > T$, then mark as edge (e.g., set to 1), else 0
- 12: $focus_measure \leftarrow$ sum of all edge marks

The autofocus mechanism, a quintessential component in modern imaging systems, leverages edge detection algorithms to ascertain the optimal focus. This process is predicated on the premise that the sharpness of an image is at its zenith when the edges within the image are rendered with maximum clarity and distinction. Therefore, contrast detection is one of useful criteria to determine whether an image is well focused. The Sobel operator is a discrete differentiation operator that computes an approximation of the gradient of the image intensity function, which can be parallelly implemented on the IVP for high-speed edge extraction, making it an ideal choice for efficient hardware implementation in our in-sensor computing approach. It is worth noting that while other edge detection operators, such as the Laplacian operator commonly used in focus variation metrology, were considered during our preliminary investigations, the Sobel operator was selected for implementation. This decision was based on its superior balance of computational efficiency for parallel execution on our specific IVP architecture (achieving edge map generation in 0.013 ms) and its demonstrated robustness in providing a distinct sharpness peak for focus evaluation across various imaging conditions encountered in our experiments. The Sobel operator's inherent smoothing characteristics also contribute to better noise resilience compared to second-order derivative operators like the Laplacian, which is advantageous for high-speed image processing during lens oscillation.

The Sobel operator uses two 3×3 convolutional matrices, G_h and G_v , to detect edges in the horizontal and vertical directions respectively. The matrices are as follows:

$$G_h = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad G_v = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (6)$$

The convolution of an image I with the Sobel matrices is performed to find the gradients $G_h(I)$ and $G_v(I)$ at each pixel:

$$G_h(I(x, y)) = I(x, y) * G_h, G_v(I(x, y)) = I(x, y) * G_v \quad (7)$$

where $*$ denotes the convolution operation. The gradient magnitudes at each pixel are then computed using:

$$G(x, y) = \sqrt{G_h(I(x, y))^2 + G_v(I(x, y))^2} \quad (8)$$

The high-speed parallel implementation of the above-mentioned algorithms on the IVP can be seen from Alg. 1, which is mainly realised by parallelly operating on the parallel computing circuit for parallel computing including shifting, addition, subtraction, and taking absolute value. By leveraging in-sensor parallel computing techniques, the sobel edge can be efficiently obtained using just 0.013 ms in-sensor. We define a sharpness metric that evaluates image sharpness in Eq. (9) operating directly on the CPM, thus reducing latency and power consumption while achieving real-time focus.

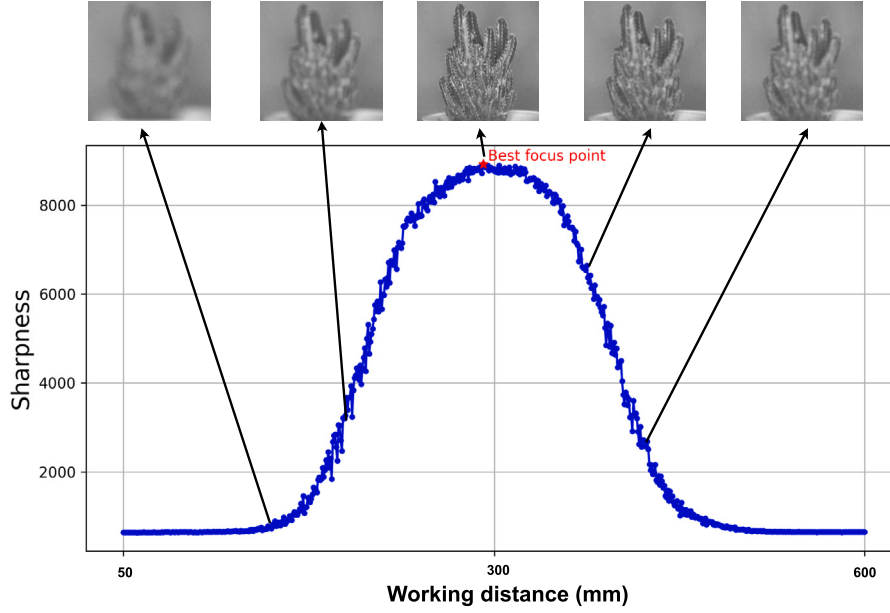


Fig. 4. Sharpness profile of an autofocus system with a liquid lens: this graph plots the sharpness against the working distance, revealing the optimal focus point where sharpness peaks. This figure shows sample images at various working distances, demonstrating the progression from blurry to sharp and back to blurry, corresponding to the near, best focus, and far points on the curve, respectively. The units of the horizontal and vertical coordinates are millimetres and quantity, respectively.

$$M(t) = \sum_{x=0}^W \sum_{y=0}^H G_i(x, y, t) \quad (9)$$

where $M(t)$ represents the focus evaluation function, $t \in [0, T]$ and T denote the total adjustment time for the liquid lens. The focus evaluation function $M(t)$ is derived from the intensity gradient of the image, which measures edge strength. This metric is directly proportional to the output of edge detection algorithms, as higher gradient values correspond to sharper transitions in pixel intensities. As shown in Fig. 4, the sharpness curve rises to a peak where the edges in the image are most pronounced, indicating optimal focus. At suboptimal focus distances, the intensity gradients are reduced due to blurring, which attenuates high-frequency components in the image, thereby lowering the edge output and sharpness metric. This relationship has been experimentally validated by comparing edge detection results and the corresponding sharpness values across multiple focus distances. The results confirm that sharpness values increase proportionally with edge strength, reinforcing the validity of using the intensity gradient-based metric for focus. Based on the sharpness evaluation, we can determine the optimal focus by solving Eq. (10), where t_f is the optimal time for image acquisition that associated with a focused image.

$$t_f = \arg \max_t M(t) \quad (10)$$

2.3.2. Focus measure and evaluation

The fundamentals for autofocusing is the Gaussian distribution of high-frequency image features, such as sharpness, which can determine the optimal focus by maximising the used features via focus tuning as shown in Fig. 4.

The IVP leverages advanced parallel processing techniques to achieve efficiency and speed in focus measurement. At the core of its computational method is a parallel summation approach, which enables in-sensor computation of the focus measure by summing the edge pixels directly. This parallel processing approach allows the IVP to complete focus measurements in just 0.32 ms, showcasing its high-speed capabilities. The parallel nature of the IVP's architecture is

Algorithm 2 In-Sensor computing for focused image during lens oscillation

Require: AM_{pix} : the raw image; DM_{edge} : edge image; DM_{focus} : focused image
Require: t : recorded time; T_{stable} : liquid lens stable time without oscillation (usually lasts 40-60 ms); M : current Sobel filter response; M_{max} : Maximum Sobel filter response; $T_{exposure}$: photo detector exposure time

- 1: *Start*
- 2: trigger lens to adjust diopter from smallest to largest {lens starts to oscillate}
- 3: **while** $t < T_{stable}$ **do**
- 4: Record t {record current time}
- 5: Reset AM_{pix}
- 6: photo-sensors exposure for $T_{exposure}$
- 7: Capture AM_{pix}
- 8: $DM_{edge} \leftarrow f_{filter}(AM_{pix})$ {Compute image edge using Sobel filter}
- 9: $M \leftarrow f_{pixelsum}(DM_{edge})$ {Compute contrast sharpness using in-sensor parallel summation method}
- 10: **if** $M > M_{max}$ **then**
- 11: $M_{max} \leftarrow M$ {Store max filter response}
- 12: $DM_{focus} \leftarrow DM_{edge}$ {Store the edge image}
- 13: **end if**
- 14: **end while**
- 15: display AM_{focus}

further exploited in implementing multiple algorithms simultaneously on the CPM. These algorithms include image edge extraction, focus measure calculation, and evaluation, all of which occur concurrently during the brief period of lens focus adjustment. This parallelism is key to the system's efficiency, as it eliminates the need for sequential processing and reduces overall computational time. The IVP's parallel processing capabilities allow for rapid computation of this function across multiple focal points simultaneously, enabling quick determination of the maximum sharpness and, consequently, the optimal focus. The computational efficiency of the IVP is further enhanced by its ability to perform these complex calculations directly on the sensor. This in-sensor computing approach minimizes data transfer overhead, a common bottleneck in traditional image processing systems. By processing data at its source, the IVP significantly reduces latency and

Table 1
Focus evaluation performance comparison (ms).

Platforms	Imaging	Sobel edge	Focus measure	In total
IVP	0.1	0.013	0.32	0.433
FPGA	0.612	0.52	1.1	2.232
CPU	0.612	15	20	35.612
ARM	2.2	150	55	207.2

increases overall system speed. The detailed algorithms for in-sensor computing during lens oscillation are outlined in Alg. 2. The parallel implementation of these algorithms on the IVP's CPM exemplifies how the system's computational theory translates into practical speed and efficiency gains in real-time autofocus applications.

3. Experiments and results

3.1. Algorithm performance comparison among various platforms

To present the performance of the proposed parallel algorithms based on the IVP, this section provides a comparison with three widely-used computing platforms: FPGA, CPU, and ARM Cortex, using identical algorithm configurations with the same-size image input. It is important to note that the 'Imaging' category in Table 1 includes the exposure time, image capturing, and image transmission durations. Specifically, the platforms utilized in this study include the FPGA Xilinx Virtex-7, CPU Intel Core i7-9700K, and ARM Cortex-A53. In conclusion, when compared to other prevalent platforms, the proposed method demonstrates superior performance in terms of computational efficiency, power consumption, and cost-effectiveness (Table 1).

3.2. Hardware system setups

Our compact in-sensor focusing system mainly includes two key components:

1. Qianxun IVP system: The core computational unit, the IVP, processes image data directly *in situ*, facilitating rapid and accurate focus monitoring without the need for external processing power.
2. Liquid Lens: LK-10M357.1Y01 from [29]: This liquid lens features a variable curvature, adjustable via serial communication, enabling quick refocusing in response to changes in the visual field. It consumes approximately 0.5 W of power and offers various continuously tunable focus levels. The lens has a diagonal field of view (FOV) of 29.8°, and an aperture f/7.1. With a minimum working distance of >150 mm, it is well-suited for high-speed embedded vision applications and seamlessly integrates with the IVP-based autofocus system. Operating Temperature: -10°C to 50°C. The liquid lens is integrated into a lens module that also includes a 35 mm primary fixed-focus lens. The liquid lens component provides the focus tunability for this module. The effect of the liquid lens's tuning results in an overall focal length adjustment for the entire lens module from approximately 22.9 mm to 42.4 mm. This range allows our system to achieve autofocus over various working distances.

As depicted in our system setup (Fig. 5), Our system setup is designed to replicate conditions found in embedded automation environments. By presenting objects at varying distances, we demonstrate the system's adaptability and efficiency in real-time machine vision tasks. This setup allows us to test the system's performance in high-speed, dynamic scenarios. The IVP communicates with liquid lens via a serial RS232 communication protocol. This protocol enables the IVP to send control signals to the liquid lens, directing it to adjust its curvature as required to maintain or change focus. The synchronization of these components is critical, allowing the system to achieve rapid refocusing capabilities essential for high-speed and high-accuracy imaging tasks.

Operating at a Baud Rate of 115200, the protocol achieves an exceptionally low instruction transmission latency of approximately 0.087 ms between the IVP and the liquid lens. This minimal delay is crucial for maintaining continuous and smooth focus transitions, particularly in applications requiring rapid and precise image adjustments.

Fig. 6 illustrates the commonly-used test objects selected for validating the performance of our in-sensor computing system's autofocus functionality: (a) a QR code, presenting a common pattern in daily life to evaluate focal precision with size of 68 × 68 mm. (b) an alphabet series, employed for their ubiquity in evaluating textual clarity and sharpness with size of 105 × 105 mm. (c) a USAF resolution test chart, the established benchmark for resolving power assessment with size of 107 × 137 mm. This array encapsulates a spectrum of testing environments to thoroughly check our system's image processing capabilities across diverse visual patterns with different sizes.

In the initial phase of our experiments, we presented three different patterns at varying distances in front of our autofocus system. Upon initiating the test, the IVP dispatched signals to the liquid lens to commence its focus adjustment from the nearest to the farthest point. Concurrently, the IVP begins to capturing images and computing real-time sharpness.

Fig. 7 demonstrates the liquid lens's autofocus response across a variety of test objects, marked by distinctive peaks indicating the moments of optimal focus. These peaks, a result of the lens's damped sinusoidal behaviour and the Gaussian distribution of the sharpness along focus, represent the global maxima in sharpness. These data points are recorded before and after the command (CMD) switching from focus level 40 to 10 corresponding to the change of liquid lens curvature. CMD switch means a rapid voltage change to the liquid lens. Noteworthy is the lens response latency between command initiation and lens response, which our system precisely monitors in real-time, efficiently analyzing the sharpness of the captured images. The consistent pattern of sharp peaks – displayed across test objects such as a QR code, alphabet, and USAF test chart – confirms the system's ability to swiftly monitor focus from the lowest to the highest levels and to accurately detect and respond to varying focal conditions. Fig. 7(d) presents the focus measurement values for dynamic objects, showing similarities to the results obtained from the static scene. The fluctuations in Fig. 7(d) primarily result from the continuous motion of the dynamic object during the autofocus process. Unlike static scenes in (a) - (c), where the liquid lens response is primarily influenced by its internal stabilization dynamics, the autofocus system for dynamic objects must account for additional variability introduced by object motion. These fluctuations arise as the system attempts to compensate for real-time positional changes in the object, requiring rapid adjustments to the lens curvature.

Moreover, the edge quality is evaluated using following two methods to test the focus accuracy of our proposed systems. (1) Intersection over Union (IoU), is a statistic used for gauging the similarity and diversity of sample sets. The Jaccard Index measures the similarity between two sets of data. For binary image segmentation, it can be defined as the size of the intersection divided by the size of the union of the segmented regions. It is formulated as: $J(A, B) = \frac{|A \cap B|}{|A \cup B|}$, where A , B represents the groundtruth edge image and the focused edge image, respectively. (2) The Dice coefficient (DSC) quantifies the similarity between two sets of data. In image segmentation, it evaluates the accuracy of a segmentation by measuring the overlap between the predicted segmentation Y and the ground truth X , calculated as $DSC = \frac{2|X \cap Y|}{|X| + |Y|}$.

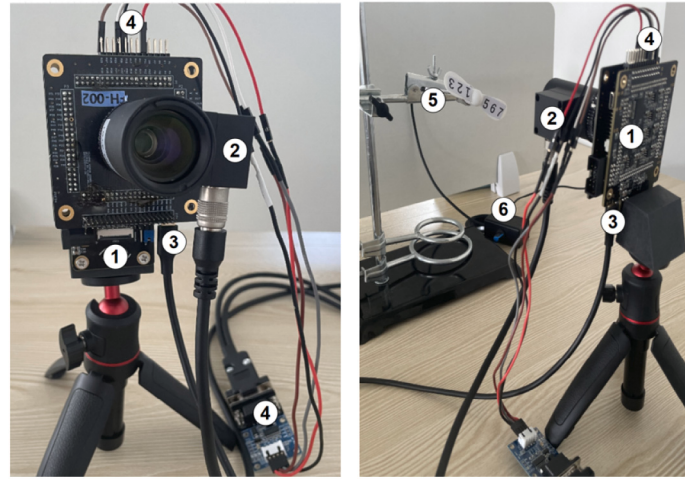


Fig. 5. The compact autofocus system: (1) IVP, (2) liquid lens, (3) converter and cable from IVP to the liquid lens, (4) display interface, (5) a miniature fan for dynamic test, (6) speed controller. Signals are synchronized by serial RS232 communication through UART interface.

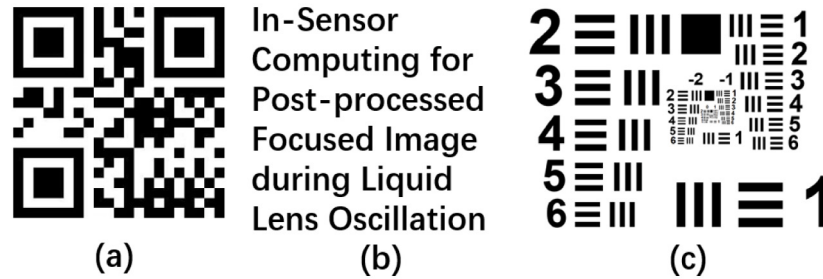


Fig. 6. Three different test patterns. (a) QR code, (b) Alphabet, (c) USAF test chart.

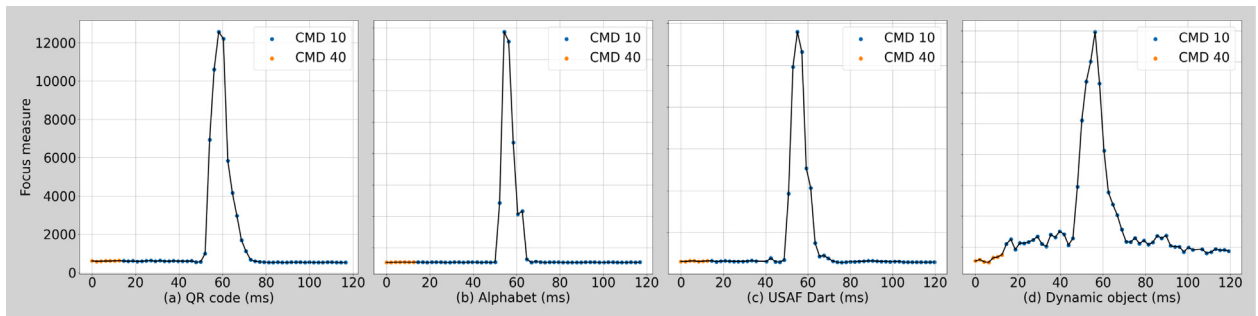


Fig. 7. The liquid lens response measured by the IVP for four different objects after a control command switch. (a) to (c): static scene, (d) dynamic scene.

Fig. 8 presents the experimental focus results for a series of static objects, tracking the progression from initial capture to a focused edge image, and ultimately comparing these results to the ground truth. The sequence captures the IVP's real-time edge detection capabilities at various time intervals, demonstrating the system's ability to refine the image detail as it approaches the optimal focus. The QR code, a text sample, and the USAF resolution test chart serve as subjects to challenge the system's focus under static conditions. The focused edge images reveal a contrast against the ground truth, validating the precision and effectiveness of the in-sensor computing during liquid lens oscillation and its suitability for applications demanding high levels of image processing speed in static image focus.

3.3. Autofocus for dynamic object

Given the high-speed imaging and processing capacity of our proposed autofocus system and associated algorithms, we formulated an

experiment focusing on responsive measurements for dynamic objects. The experimental configuration, depicted in Fig. 5, involves the utilization of a standard USB interface miniature fan (size: 96×8 mm) connected to a power source via a fan speed controller. Throughout the experiment, we configured the fan speed to 192 rpm and rotated it clockwise. Subsequently, we modulated the liquid lens focal length from its minimum to maximum extent by giving a command switch, recording the corresponding focus measurement values throughout the procedure to substantiate the system's precision concerning dynamic objects.

Fig. 8 (bottom row) illustrates the process of capturing dynamic moving objects: a series of four images arranged from left to right, demonstrating the adjustment of the liquid lens focal length from its lowest to farthest level. The ground truth image was captured with a fixed focal length while the fan was stationary. During this experiment, the visual system IVP recorded image data at intervals of around 2 ms including image capturing, processing and outputting for display. In

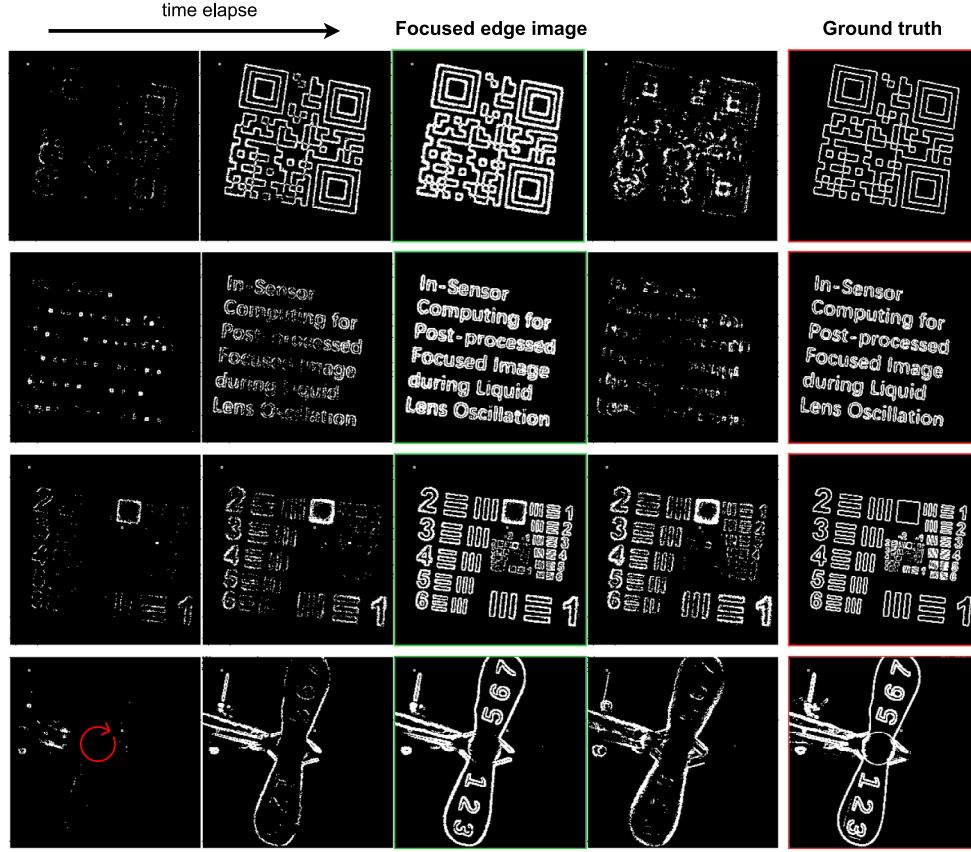


Fig. 8. Experimental focus results for static objects (first three rows) and a dynamic object (fourth row).

Table 2
Sharpness measure accuracy and focused time.

Patterns	Object distance (mm)	IoU (%)	Dice (%)	Focus time (ms)
QR code	320	75.7	86.2	45.7
Alphabet	320	75.2	85.8	41.9
Test chart	300	78.2	87.7	42.9
Rotating fan	300	70.1	82.0	44.2
Spatial objects	130, 320, 820	–	–	<42.0, 45.6, 49.1

Fig. 8, the progression from initial capture to focused edge images is visually compared against the ground truth, demonstrating the system's capability to refine image detail as it approaches optimal focus. This qualitative validation is further supported by the quantitative results in Table 2, which shows high Intersection over Union (IoU) and Dice coefficients across different patterns, confirming the accuracy of the focused edge images. The short focused times, as indicated in the table, also highlight the effectiveness of the in-sensor computing system in achieving rapid focus adjustments, validating its suitability for applications requiring high-speed image processing. The experimental findings confirm that our proposed autofocus system achieves practical performance when applied to both dynamic and static objects.

3.4. Autofocus for spatial objects

Our research advances the domain of autofocus in imaging systems by employing an innovative method that harnesses lens oscillation for rapid focus adjustment of spatial objects. Unlike conventional systems that depend on multiple lens positions, which are typically slow and power-hungry, our approach utilizes the natural oscillation of a liquid lens coupled with high-speed in-sensor sampling to swiftly adjust focus, resulting in efficient power consumption and processing. The proposed system operates by sweeping the liquid lens across various focus levels

Algorithm 3 High-Speed In-Sensor ‘Global Focus’ for Spatial Objects

Require: AM_1, DM_L, DM_m : image, edge response, ensembled edge image

Require: AM_p : the image acquisition array (the photo-sensors)

Require: ΔT : exposure control

```

1: while Lens Oscillation Period do
2:    $AM_1 \leftarrow AM_{pix}$  {Load current image}
3:    $AM_{pix} \leftarrow f(AM_p)$  {Start light integration for the next image}
4:    $DM_L \leftarrow f_{filter}(AM_1)$  {Edge image process}
5:   if  $R_L < \theta$  then
6:      $DM_L \leftarrow 0$  {Discard pixels with low maxima}
7:   end if
8:    $DM_m = OR(DM_m, DM_L)$ 
9:   wait  $\Delta T$  {Control  $AM_{pix}$  exposure}
10: end while
11: trigger depth-frame readout of  $DM_m$ 

```

and capturing a series of local focus images. These images are then aggregated to form a single, globally-focused image. By executing high-speed image contrast measurements within the sensor, our system identifies and synthesizes the sharpest edges for each frame in milliseconds, significantly accelerating the autofocus process.

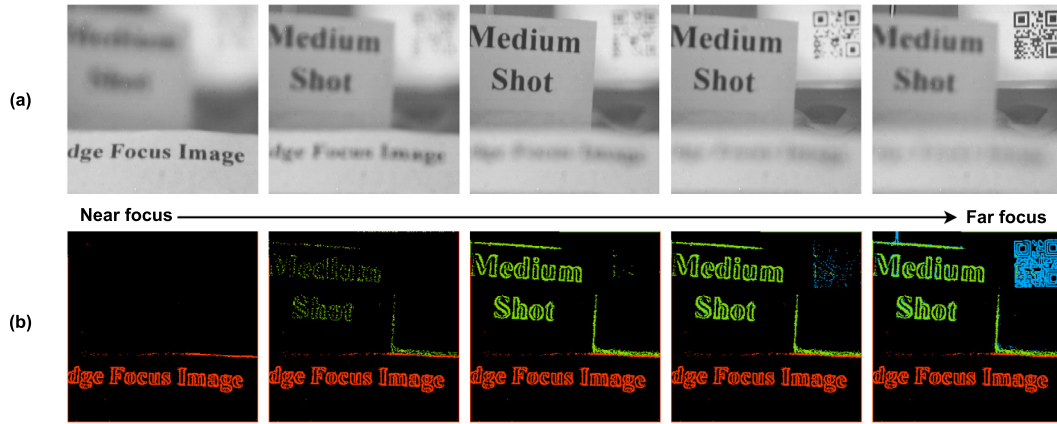


Fig. 9. Edge Image Progression through Liquid Lens Oscillation. (a) The transition of focus from near to far within a single oscillation period of the lens, demonstrating the dynamic focus adjustment capability of the system. (b) The stages of in-sensor edge detection, highlighting the system's efficiency in edge refinement without requiring stabilization of the lens. The varied colours represent different object distances, with changes depicted as the focus shifts.

Table 3
Autofocus system performance comparison.

Work	Hardware setup	Method	Autofocusing speed	System power
[12]	PC with Intel i9, RTX2080, motor	Gaussian standard deviation	500 ms	>100 W
[30]	PC with Intel Core2, motor	look-up table	126.2 ms	>50 W
[13]	PC with Intel i5, motor	ARSM	>1000 ms	>20 W
[31]	PC with Xeon, high-speed camera	Brenner gradient	15.8 ms	>100 W
[32]	CompactRIO-9049, high-speed camera	LSTM	150 ms	>60 W
[33]	PC with Intel i7, high-speed camera	histogram feature	–	>20 W
[14]	PC with AMD R7	CNN and DFF	251.8 ms	>65 W
[34]	PC with Intel i7, RTX2080ti, motor	CNN	>250 ms	>100 W
[10]	TAG lenses with controller and light source	ESFC	ms-level	40 W
Ours	IVP system	In-Sensor sharpness measure	<50 ms	2 W

Fig. 9 illustrates the proposed autofocus system in action. A set of objects, including an 'Edge Focus Image', 'Medium Shot' sign and a QR code, are positioned at varying distances from the IVP. As the liquid lens oscillates, it fluctuates the focus from the nearest object to the farthest, sequentially bringing each object into sharp relief. This dynamic process results in a comprehensive capture of focused edge information across multiple focal planes. Alg. 3 details the autofocus process for spatial objects: it leverages the DM_L to temporarily store the edge information. Through logical OR operations, these edge details are accumulated in the DM_m . The culmination of this process is a 'global focus' image, which conveys focused edges from the entire range of oscillation. The entire cycle, from initiating the lens movement to achieving a global edge-focused image, is completed within a mere 30 ms excluding lens response time, showcasing the speed and efficiency of our system. θ is a user-defined parameter to reduce influence from noise. Table 2 summarises the autofocus system performance in terms of edge accuracy and time. The proposed compact autofocus system is able to obtain the comparatively high-quality edge image within a short period of time. Additionally, Table 3 presents a comprehensive performance comparison of the proposed autofocus system with related systems, evaluating methods, efficiency, and power consumption. Notice that, the focused time include the liquid lens response time and oscillation time as shown in Fig. 2. With a quicker responsive liquid lens, a shorter focus time can be expected.

4. Discussion

The primary performance bottleneck of the proposed compact autofocus system is the response latency of the liquid lens, measured at approximately 15–20 ms. This latency significantly constrains overall system responsiveness. While the latency introduced by digital frame output could be mitigated through faster readout circuitry (e.g.,

a USB 3.0 interface), further performance gains, approaching near-instantaneous optimal focus capture, would require a liquid lens with a reduced response time. Furthermore, in well-lit outdoor environments, the integrated light source may be unnecessary, simplifying system configuration.

It is important to note that while our method efficiently finds focus within a single triggered lens oscillation (as described in Alg. 2 and 3), if the object of interest lies outside the dioptic range covered by that specific oscillation, it will not be brought into focus in that attempt. This is a general consideration for focusing systems. A complete autofocus implementation based on our principle would incorporate a higher-level control strategy to manage such scenarios, for example, by initiating subsequent triggers with different voltage step parameters to explore a wider focal range until a satisfactory sharpness peak is detected.

In addition, like other autofocus systems relying on contrast or edge-based sharpness metrics, the performance of our proposed system can be challenged in scenes with very few discernible edges or extremely low overall contrast. In such conditions, the reduced gradient information would lead to a less distinct sharpness peak, potentially affecting the accuracy and reliability of the autofocus process. The system is designed to excel in environments where sufficient edge or texture information is available for robust sharpness evaluation.

Regarding the 'global focus' image, while effective for static spatial objects, significant object motion during the lens oscillation period (<50 ms) would result in a composite image that captures an envelope of the object's sharp features over its trajectory, rather than a single, motion-frozen all-in-focus snapshot.

5. Conclusions

This work has developed first efficient in-sensor autofocus system that integrates a IVP with a liquid lens using proposed parallel image

processing methods. This innovative system delivers over 500 *fps* with a latency of less than 2 *ms*, covering image capture, processing, and useful data transfer. Notably, it maintains energy efficiency with around 2 watts of power consumption, marking an advancement over traditional autofocus technologies. Uniquely, it utilizes the liquid lens oscillation period for rapid imaging and in-focus image processing, setting it apart from traditional autofocus technologies. The compactness, rapid processing speed, cost-effectiveness, and low power consumption make our system suitable for portable and embedded applications, especially in scenarios requiring the detection of fast-moving objects.

In the future, we aim to integrate the proposed system with a mobile platforms, enabling machine vision systems with the capability of effective imaging and processing despite vibrations. Furthermore, key improvements will focus on higher resolution sensors, advanced readout circuit designs, and a liquid lens with a quicker response time to enable more research and practical possibilities.

CRedit authorship contribution statement

Yanan Liu: Project administration – original draft, Methodology, Conceptualization. **Zhiguo Yang:** Writing – original draft, Visualization, Methodology. **Huixin Zhong:** Writing – original draft, Formal analysis. **Rui Fan:** Writing – review & editing. **Ruidong Ma:** Writing – review & editing, Visualization. **Xiupeng Shi:** Supervision, Writing – review & editing. **Jianhua Zhang:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Yanan Liu reports financial support was provided by Natural Science Foundation of China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Natural Science Foundation of China (grant 62202285) and The Special Funds for Promoting High-quality Industrial Development in Shanghai (grant JJ-ZDHYLY-01-23-0004).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.sna.2025.116868>.

Data availability

Data will be made available on request.

References

- [1] H. Cheng, L. Wang, S. Tabata, Y. He, Y. Hu, J. Liu, High-speed all-in-focus 3D imaging technology based on the liquid lens focus scanning, in: *Optoelectronic Imaging and Multimedia Technology X*, vol. 12767, SPIE, 2023, pp. 81–90.
- [2] S. Huang, L. Wang, Y. Huang, Y. He, S. Bai, Measurement method of virtual image distance for a head-mounted display based on a variable-focus liquid lens, *Appl. Opt.* 63 (15) (2024) 4175–4181.
- [3] C. Liu, Y. Zheng, R.-Y. Yuan, Z. Jiang, J.-B. Xu, Y.-R. Zhao, X. Wang, X.-W. Li, Y. Xing, Q.-H. Wang, Tunable liquid lenses: emerging technologies and future perspectives, *Laser Photonics Rev.* 17 (11) (2023) 2300274.
- [4] S. Pasinetti, I. Bodini, G. Sansoni, F. Docchio, M. Tinelli, M. Lancini, A fast autofocus setup using a liquid lens objective for in-focus imaging in the macro range, in: *AIP Conference Proceedings*, vol. 1740, (1) AIP Publishing, 2016.
- [5] Y. Marrakchi, X. Barcala, E. Gamba, I. Martinez-Ibarburu, C. Dorronsoro, L. Sawides, Experimental characterization, modelling and compensation of temperature effects in optotunable lenses, *Sci. Rep.* 13 (1) (2023) 1575.
- [6] D. Schipf, *Electrowetting Liquid Lens Oscillations for Optical Applications* (Ph.D. thesis), 2019.
- [7] N. Xie, H. Zhang, R.-q. Xu, Frequency characteristics of basic mode oscillation at liquid-liquid interface of electrowetting liquid lens, *Optik* 212 (2020) 164753.
- [8] C.A. López, A.H. Hirs, Fast focusing using a pinned-contact oscillating liquid lens, *Nat. Photonics* 2 (10) (2008) 610–613.
- [9] S. Kang, M. Duocastella, C.B. Arnold, Variable optical elements for fast focus control, *Nat. Photonics* 14 (9) (2020) 533–542, <http://dx.doi.org/10.1038/s41566-020-0684-z>.
- [10] N. Vilar, R. Artigas, M. Duocastella, G. Carles, Fast topographic optical imaging using encoded search focal scan, *Nat. Commun.* 15 (2024) 2065, <http://dx.doi.org/10.1038/s41467-024-46267-y>.
- [11] Mitutoyo Corporation, TAGLENS Product Overview, 2023, <https://www.mitutoyo.com/taglens/>. (Accessed 09 May 2025).
- [12] P. DiMeo, L. Sun, X. Du, Fast and accurate autofocus control using Gaussian standard deviation and gradient-based binning, *Opt. Express* 29 (13) (2021) 19862–19878.
- [13] Q. Hao, Y. Xiao, J. Cao, Y. Cheng, C. Sun, Improving the performances of autofocus based on adaptive retina-like sampling model, *Opt. Commun.* 410 (2018) 269–276.
- [14] Z. Liu, H. Hong, Z. Gan, K. Xing, Bionic vision autofocus method based on a liquid lens, *Appl. Opt.* 61 (26) (2022) 7692–7705.
- [15] J. Kremers, L.C.L. Silveira, N.R. Parry, D.J. McKeefry, The retinal processing of photoreceptor signals, *Hum. Color. Vis.* (2016) 33–70.
- [16] Y. Wang, H. Feng, Z. Xu, Q. Li, Y. Chen, M. Cen, Fast auto-focus scheme based on optical defocus fitting model, *J. Modern Opt.* 65 (7) (2018) 858–868.
- [17] C. Wang, Q. Huang, M. Cheng, Z. Ma, D.J. Brady, Deep learning for camera autofocus, *IEEE Trans. Comput. Imaging* 7 (2021) 258–271.
- [18] E. Huang, J. Thomas, D. Hibberd, V. Luong, R. Lott, M. Oleson, A. Hood, M.H. MacDougall, Small pixel MWIR sensors for low swap applications, in: *Infrared Technology and Applications XLVII*, vol. 11741, SPIE, 2021, pp. 123–129.
- [19] F. Toates, Accommodation function of the human eye, *Physiol. Rev.* 52 (4) (1972) 828–863.
- [20] T. Zhang, K. Shimazaki, I. Ishii, A. Namiki, High-magnification object tracking with ultra-fast view adjustment and continuous autofocus based on dynamic-range focal sweep, *Sensors* 24 (12) (2024) 4019.
- [21] E.J. Micles, W.Y. Lim, O.D. Supekar, M. Zohrabi, J.T. Gopinath, V.M. Bright, Axisymmetrical resonance modes in an electrowetting optical lens, *Appl. Phys. Lett.* 122 (20) (2023) 201106, <http://dx.doi.org/10.1063/5.0141787>.
- [22] C.-P. Chiu, T.-J. Chiang, J.-K. Chen, F.-C. Chang, F.-H. Ko, C.-W. Chu, S.-W. Kuo, S.-K. Fan, Liquid lenses and driving mechanisms: a review, *J. Adhes. Sci. Technol.* 26 (12–17) (2012) 1773–1788.
- [23] S. Kuiper, B.H. Hendriks, Variable-focus liquid lens for miniature cameras, *Appl. Phys. Lett.* 85 (7) (2004) 1128–1130.
- [24] I.S. Park, Y. Park, S.H. Oh, J.W. Yang, S.K. Chung, Multifunctional liquid lens for variable focus and zoom, *Sensors Actuators A: Phys.* 273 (2018) 317–323.
- [25] T. Yamazaki, H. Katayama, S. Uehara, A. Nose, M. Kobayashi, S. Shida, M. Odahara, K. Takamiya, S. Matsumoto, L. Miyashita, et al., A 1ms high-speed vision chip with 3d-stacked 140gops column-parallel pes for spatio-temporal image processing, in: 2017 IEEE International Solid-State Circuits Conference (ISSCC), IEEE, San Francisco, CA, USA, 2017, pp. 82–84, <http://dx.doi.org/10.1109/ISSCC.2017.7870271>.
- [26] Á. Rodríguez-Vázquez, R. Domínguez-Castro, F. Jiménez-Garrido, S. Morillas, J. Listán, L. Alba, C. Utrera, S. Espejo, R. Romay, The eye-ris cmos vision system, in: *Analog circuit design*, Springer, New York City, 2008, pp. 15–32.
- [27] Á. Zarándy, *Focal-plane sensor-processor chips*, Springer Science & Business Media, New York City, 2011.
- [28] Y. Liu, R. Fan, J. Guo, H. Ni, M.U.M. Bhutta, In-sensor visual perception and inference, *Intelligent Computing* 2 (2023) 0043.
- [29] Ioipshare, Ioipshare, 2025, <https://en.ioipshare.com/>. (Accessed 03 March 2025).
- [30] P. Sripathan, P. Mittrapiyanuruk, P. Keawtrakulpong, A high speed autofocus system for micro system applications, *J. Electron. Sci. Technol.* 14 (1) (2016) 73–79.
- [31] H. Oku, M. Ishikawa, High-speed liquid lens for computer vision, in: 2010 IEEE International Conference on Robotics and Automation, IEEE, 2010, pp. 2643–2648.
- [32] J. Yan, P. DiMeo, L. Sun, X. Du, LSTM-based model predictive control of piezoelectric motion stages for high-speed autofocus, *IEEE Trans. Ind. Electron.* 70 (6) (2022) 6209–6218.
- [33] C. Guo, Z. Ma, X. Guo, W. Li, X. Qi, Q. Zhao, Fast auto-focusing search algorithm for a high-speed and high-resolution camera based on the image histogram feature function, *Appl. Opt.* 57 (34) (2018) F44–F49.
- [34] S.A.R. Farnes, D.-M. Tsai, W.-Y. Chiu, Autofocus measurement for electronic components using deep regression, *IEEE Trans. Compon. Packag. Manuf. Technol.* 11 (4) (2021) 697–707.

Yanan Liu obtained a Ph.D. in Robotics and Autonomous Systems from the University of Bristol, UK. He is a Lecturer at the School of Microelectronics, Shanghai University, China. He researches in the areas of in-sensor computing, embedded AI, Robotics, machine learning and applications.

Zhiguo Yang received the Bachelor degree in Internet of Things Engineering from Yantai University at Yantai, China. He is currently pursuing the M.Sc. degree in Electronic and Information Engineering at the School of Microelectronics, Shanghai University at Shanghai, China. His researches in the areas of in-sensor computing and machine vision.

Huixin Zhong obtained a Ph.D. at the Center for Doctoral Training in Accountable, Responsible, and Transparent AI at the University of Bath, UK. She is an Assistant Professor at the College of Industry Entrepreneurs, Xi'an Jiaotong-Liverpool University at Suzhou, China. She researches in the areas of AI ethics and governance, human-AI collective decision-making and machine psychology.

Rui Fan received the Ph.D. degree in Electrical and Electronic Engineering from the University of Bristol. He worked as a Research Associate at the Hong Kong University of Science and Technology from 2018 to 2020 and a Postdoctoral Scholar-Employee at the University of California San Diego between 2020 and 2021.

He began his faculty career as a Full Research Professor in the College of Electronics & Information Engineering at Tongji University in 2021. His research interests include computer vision, deep learning, and robotics, with a specific focus on humanoid visual perception under the two-streams hypothesis.

Ruidong Ma received a Ph.D. in Robotics from the University of Sheffield. He is currently a research fellow at Sheffield Hallam University, United Kingdom. His research interests include robot learning, robot perception, and human-robot collaboration.

Xiupeng Shi obtained a Ph.D. in Artificial Intelligence from the Nanyang Technological University, Singapore. He is currently an Associate Professor at the School of Microelectronics, Shanghai University, China. His research focuses on AI algorithm-chip co-design, smart mobility, intelligent manufacturing, and AI for Science.

Jianhua Zhang received the Ph.D. degree in mechanical engineering from Shanghai University at Shanghai, China. She is currently a Professor with the School of Microelectronics, Shanghai University, where she is also the Director of the Key Laboratory of Advanced Display and System Applications. Her main research interests include digital chip design and system integration, digital video and stereoscopic display technology, and artificial intelligence algorithms and chip design.