2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Real-Time Particle Cluster Manipulation with Holographic Acoustic End-Effector under Microscope

Siyuan An, Chengxi Zhong, Mingyue Wang, Shudong Wang, Haojian Lu, Jiaqi Li, Youfu Li, and Song Liu

***Abstract*—Non-contact particle cluster manipulation holds significant promise in the realms of advanced manufacturing, chemistry, and pharmacy. However, achieving precise and dynamic control over the spatial kinematics of particle clusters remains a significant challenge, necessitating real-time and accurately programmable robotic end-effector. To this end, we develop an innovative non-contact, precise particle cluster manipulation system with ultrasonic phased array transducer (PAT) under microscope. This system combines a physics-based deep learning algorithm for real-time calculation of phase-only holograms (POHs), supporting PAT to dynamically form acoustic fields, namely holographic acoustic end-effector (HAE). Leveraging the dynamically and accurately generated HAEs by our system, kinematics control of particle clusters including aggregation, rotation, and translation is yielded. The extensive experiments well demonstrated the effectiveness of proposed system for particle cluster manipulation.**

I. Introduction

Non-contact particle cluster manipulation has emerged as a promising technology spanning distinct fields like physics [1], chemistry [2], medicine [3], and pharmaceutical testing [4]. Hitherto, this technology has been explored using various physical mechanisms, such as electromagnetics [5], optics [6], chemical reactions [7], electroosmosis [8], and acoustics [9]. Among them, acoustics mechanism offers prominent merits of deep penetration ability [9], robust loading capacity [10], and independence of pre-add compounds [11]. However, the acoustics-actuated non-contact particle cluster manipulation necessitates precise and dynamic spatiotemporal kinematic control, which is still in an infancy. Thus, this underscores the significant importance of ongoing study regarding this field.

Dynamically spatiotemporal kinematic control of particle clusters by acoustics actuation involves generating three-dimensional (3D) acoustic fields precisely and in real time, whereby the resultant acoustic radiation force (ARF) [12] conducted on particle clusters induces motion energies. Thus far, there are two main manners to dynamically generating 3D acoustic fields, including ultrasound phased array transducer (PAT) [13, 14] and acoustic metasurface combined single transducer [12]. The former individually modulates the phase of emitted acoustic waves from transducers, while the later pre-fabricates a metasurface by encoding phase shift as pixelated thickness using 3D printing technique, which is subsequently used to regulate phases of acoustic wave emitted from the single transducer. Although the metasurfaces impart high-resolution feature to generated acoustic fields, the dynamic modulation capability required for dynamic spatiotemporal kinematic control over particle clusters is lack due to its prefabricated structure. On the contrary, the PAT is characterized with dynamic modulation ability and acoustic field generation versatility. However, achieving phase profile for PAT given arbitrary acoustic field corresponding to specific manipulation task can be challenging due to its inherent nonlinearity and high dimensionality [15], which makes phase calculation be viewed as an inverse kinematics problem from perspective of robotics.

The required phase profile for PAT is termed phase-only hologram (POH). For its calculation, there are two dominant algorithms including traditional iterative optimization algorithms like IASA [12], IB [16], diff-PAT [17], and advanced deep learning algorithms like PhysNet [18] and HU-Net [19]. The iterative optimization algorithms embrace underlying physical constraints by iteratively operating two symmetrical processes, i.e., acoustic field forward propagation and backward time reverse. Nevertheless, they face accuracy limitation and unstable real-time performance that highly influenced by the complexity of desired acoustic fields [12]. Alternatively, deep learning which has been well demonstrated in computer vision and natural language process was induced to figure out POH. Its high-dimensional mapping and parallel computing abilities facilitate the accurate POH solving in real time. However, the reported deep learning algorithms mainly focus on the improvement of reconstruction quality from algorithmic perspective. The implementation of effective algorithm incorporated with developed practical system benefits enhancing potential in real world scenarios.

In this work, we develop an innovative non-contact particle cluster manipulation system under microscope. With the objective of real-time and accurate POH calculation, we further incorporate a physics-based deep learning approach to the system, which simultaneously combines underlying physical constraints, high-dimensional mapping ability, and parallel computation. By the developed system empowered by proposed deep learning approach, we can precisely and dynamically generate acoustic fields, which is vividly termed as holographic acoustic end-effector (HAE). The experiments regarding aggregation, rotation, and translation validate the system's reliability and effectiveness.

This work was supported by National Natural Science Foundation of China under Grant 62303321 and in part by Hong Kong Research Grant Council under CityU11206122 and in part by State Key Laboratory of Electrical Insulation and Power Equipment (EIPEEIPE24212). (S. An and C. Zhong contribute equally to this work.) (Corresponding author: Song Liu)

S. An, C. Zhong, M. Wang, J. Li, S. Liu are with the School of Information Science and Technology, ShanghaiTech University, Shanghai, China. (e-mail: ansy2022, zhongchx, wangmy1, lijq1, liusong@shanghaitech.edu.cn).

S. Wang is with the School of Electrical Engineering, Xian Jiaotong University, and with State Key Laboratory of Electrical Insulation and Power Equipment, Xi an, (email: Shudong_wang@xjtu.edu.cn).

H. Lu is with the School of the College of Control Science and Engineering, Zhejiang University, China, and with the Institute of Cyber-Systems and Control, Hangzhou, (email: luhaojian@zju.edu.cn).

Y. Li is with the Department of Mechanical Engineering, City University of Hong Kong, Kowloon, HK, (email: meyfli@cityu.edu.hk)

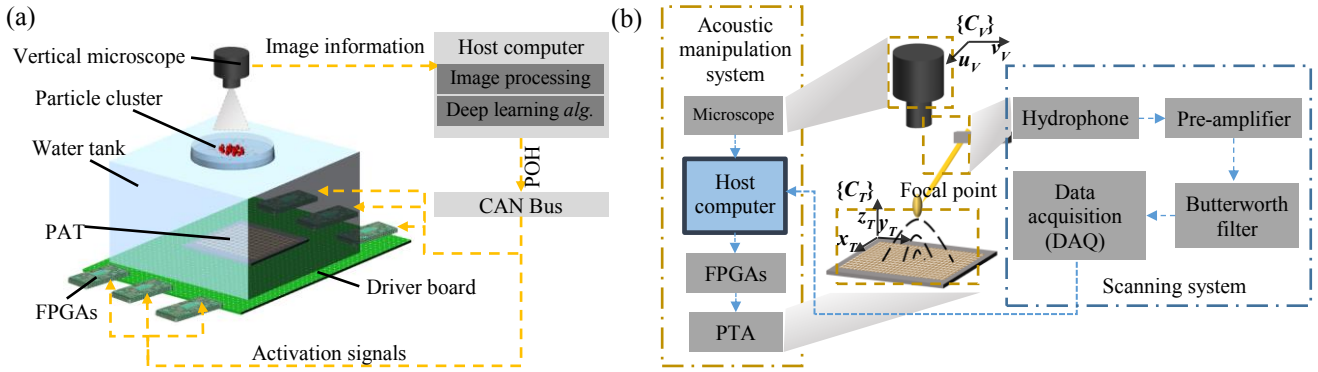

Fig. 1. The schematic diagrams of (a) acoustic manipulation system and (b) its hand-eye relationship calibration.

II. System Configuration and Calibration

A. System Configuration

Our developed acoustic manipulation system is shown in Fig. 1(a). Specifically, a PAT is positioned in the central bottom of a deionized water filled tank. Petri dishes placed on the water surface with a fixed vertical height act as working space to control particle clusters by virtue of generated HAE. Above the tank, a vertical microscope captures images of the particle clusters manipulated by HAE, which is fed into a host computer executing images processing and our proposed deep learning algorithm (which will be described in the next section). Afterward, the POH is achieved and transmitted to the PAT in the form of electric signal with different time delays relying on FPGAs to form expected HAE in the working space, whereby the specific particle cluster manipulation is accomplished. We mainly manipulate particle clusters with positive acoustic impedance in water to align with cellular in in-vivo practical applications. Through the specific HAE, these kind of particle clusters will be trapped in the region with low local acoustic pressures.

B. Calibration

Before robotic manipulation using HAE, the calibration of HAE is foundational and crucial, which is illustrated in Fig. 1(b). The Cartesian space coordinate $\{C_T\}$ is established at PTA plane whose origin sitting in PTA's right-up corner, while the image coordinate $\{C_V\}$ is established at the microscope plane. The needle hydrophone used to scan non-visible HAE is characterized with simultaneous physical position in $\{C_T\}$ and image position in $\{C_V\}$. Then, a reference point is randomly selected. Since the particles for manipulation are confined to the XOY plane, the HAE's movement along the z-axis can be neglected, which makes image Jacobian matrix [20] denoted $\boldsymbol{J}_v^*$ for calibration be quantified with a two-dimensional matrix formulated by

$$\begin{bmatrix} x_{ptc} \\ y_{ptc} \\ z_{ptc} \end{bmatrix} = \begin{bmatrix} x_{ref} \\ y_{ref} \\ z_{ref} \end{bmatrix} + \begin{bmatrix} \boldsymbol{J}_v^* \\ 0 \quad 0 \end{bmatrix} \begin{bmatrix} u_{ptc} - u_{ref} \\ v_{ptc} - v_{ref} \end{bmatrix} \tag{1}$$

where $[x_\cdot, y_\cdot, z_\cdot]^T$ and $[u_\cdot, v_\cdot]^T$ refer the coordinates in $\{C_T\}$ and $\{C_V\}$, respectively. The subscript *ptc* and *ref* denote hydrophone's needle point and selected reference point. Notably, the neglected z-axis let z_{ptc} equal z_{ref}.

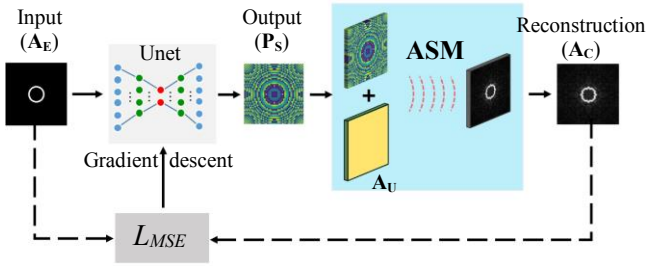

Fig. 2. Pipeline of the proposed physics-based unsupervised learning method.

III. Physics-based Unsupervised Learning for Particle Cluster Manipulation

To generate desired HAE for a specific particle cluster manipulation task, it's crucial to solve the ill-posed inverse kinematics problem denoted as f^{-1} from HAE to corresponding POH. To this regard, this paper introduces a novel physics-based unsupervised learning approach illustrated as Fig. 2. Specifically, a classic convolutional network, namely U-Net [19], takes charge of predict POHs denoted as $\mathbf{P_S}$ given desired HAE denoted as $\mathbf{A_E}$. Afterward, the predicted $\mathbf{P_S}$ coupled with a fixed uniformly distributed amplitude denoted as $\mathbf{A_U}$ are fed into a well-established physical model named Angular Spectrum Method (ASM) [15] to yield reconstructed HAE denoted as $\mathbf{A_C}$. Subsequently, the mean square error (MSE) between $\mathbf{A_E}$ and $\mathbf{A_C}$ is used as network error expressed as

$$L_{MSE} = \frac{1}{N}\sum(\mathbf{A_E} - \mathrm{ASM}(f^{-1}(\mathbf{A_E})))^2 \tag{2}$$

where N is the pixel amount used for quantification. Thereby, error minimization through the gradient descent algorithm leads to the optimized network parameters to support off-line POH calculation given desired HAE for specific particle cluster manipulation tasks. Significantly, the incorporation of ASM with the learning process facilitate network training without dependence of expensive annotated dataset since the POH does not explicitly feature in the loss function and shows a great generalizability due to the embedded underlying physical constraints. Once convergence is achieved, the well-trained U-Net can be employed to predict POH for arbitrary desired HAE. Leveraging the parallel calculation capabilities of the network, our method supports real-time POH calculation for dynamic manipulation.

IV. Manipulation Task Specification

Particle cluster manipulations encompass kinematic and morphological manipulations, diverging from traditional particle manipulations which typically focus on individual particle [21] or uniform approaches [22]. Such traditional methods cannot support particle swarm operations merely depends on dynamically modulated HAE under particles with the same acoustic characteristics. In this work, authors focus on the investigation of kinematic manipulations of particle clusters including aggregation, translation, and rotation. These different manipulation tasks rely on the dynamic HAE forming empowered by the previous introduced deep learning approach and integrated control strategies. The possible unstable velocity is compensated by meticulous adjustment of the step size and time intervals, which permits us to disregard excessive external disturbances that might occur over successive operational cycles, allowing a concentrated focus on the HAE forming process.

A. Particle Aggregation

Particle aggregation represents the foundational and most crucial task in particle cluster manipulations, with the extent of cluster directly impacting the reliability of these operations. Any loss of particles during aggregation is viewed as a failure manipulation. Here, particle aggregation is accomplished by modulate the size of HAE with a specific shape, like circle and square. As the size of HAE gradually decreases, the particle clusters gradually aggregate into the desired shape. As illustrated in Fig.3 (a), the aggregation starts with manual selection of a physical position and the expected size of HAE. Using these initializations, the corresponding HAE can be formed, whereby we check whether the particle clusters sitting in the region of HAE. The expected size will be continuously updated till the particle clusters indeed sit in the region of HAE. Afterward, particle cluster aggregation is completed by changing size of HAE. The real-time closed-loop feedback based on the captured images from vertical microscope is used to adequately ensure all particle clusters is confined within region of HAE during performing aggregation. The precise HAE modulation and real-time closed-loop feedback facilitate robust and adaptive aggregation of particle clusters.

B. Particle Cluster Translation

To facilitate the controlled spatial distribution of particle clusters, it is significant to impart particle clusters with translation mobility. As illustrated in Fig.3 (b), particle cluster translation starts with randomly setting a desired trajectory, which is subsequently discretized into a series of positions. In turn, we form corresponding HAE for these positions to drive the movement of the particle cluster along the desired trajectory. For each position, we ensure it has an adequately small position error regarding a pre-defined threshold. If the position error is unfortunately larger than the threshold, we will conduct PID control which expressed as

$$x_{n+1} = x_n + k_p\left(x_n - x^*_n\right) \tag{3}$$

where x is the physical position, the subscript represents the iteration index, the superscript *star* sign represents the actual position, and k_I is the integral control coefficient. This control compensates for the position error between the desired position of particle clusters and the actual position captured from microscope, circumventing the deviation of manipulated

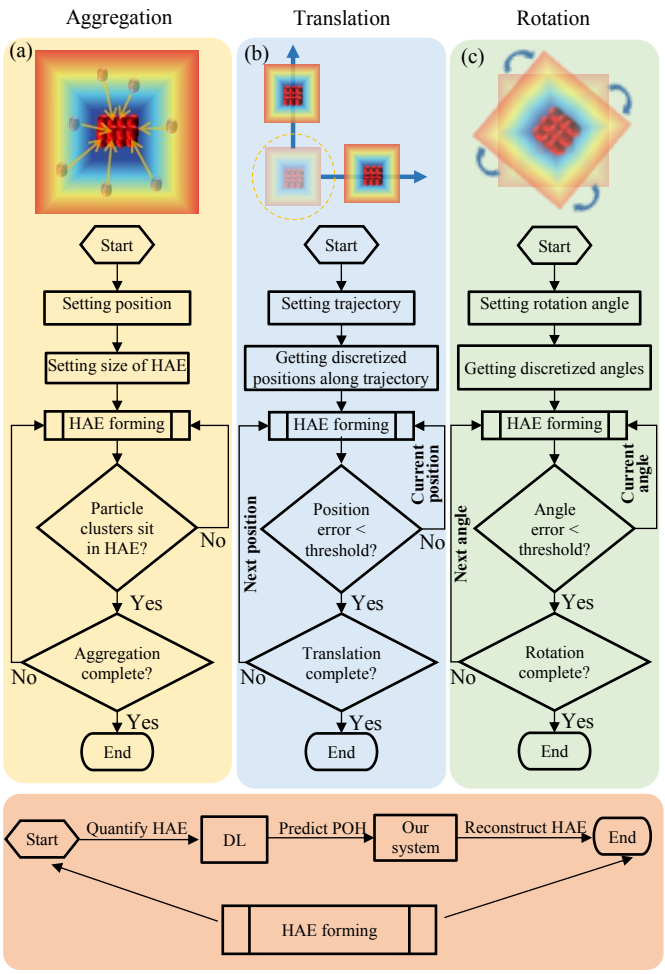

Fig. 3. Flow chats for different manipulation tasks covering (a) aggregation, (b) translation, and (c) rotation.

particle clusters from desired trajectory. Even the particle cluster might lose some particle elements, the real-time closed-loop feedback in aggregation will be executed for re-aggregation of lose particle elements. The aforementioned phenomenon of loss and trajectory deviation typically faced by non-contact manipulation is impermissible in practice due to their undetermined effects to surround environment, which illustrates the necessity of mechanism for translation. In a work, this translation process empowered by control mechanism facilitates to mitigate the occurrence of particle loss, ensuring that the integrity of the particle cluster is maintained throughout the operational process. This proactive strategy of immediate re-aggregation not only preserves the operational efficacy but also enhances the reliability of the manipulation tasks by promptly addressing any disruptions in cluster manipulation continuity.

C. Particle Cluster Rotation

As shown in Fig. 3(c), particle cluster rotation is realized with another efficient closed-loop control concerning rotation angle. Specifically, our desired rotation angle is firstly represented into a series of subangles, which subsequently are used to gradually complete rotation by corresponding formed HAE. For each subangle, we guarantee its angle error

compared with desired value be smaller than a pre-defined threshold. Nevertheless, rotation process inevitably induces disturbance. To mitigate this disturbance, we will conduct PID control given as

$$\theta_{n+1} = \theta_n + k_p\left(\theta_n - \theta^*_n\right) \tag{4}$$

where θ represents the rotation angle, the subscript denotes the iteration index of PID control, the superscript *star* sign denotes actual angle, while desired angle is signed without superscript, and k_I is the integral control coefficient. The successful implementation of this task signifies the system's ability to achieve precise rotational control over particle clusters, marking a significant advancement in the manipulation and orientation of clusters for specific operational requirements.

V. Experiment and Results

A. Experimental Details

System Setup. The acoustic Non-Contact Particle Manipulation (NPM) platform's physical prototype, as detailed in Section II-A, is depicted in Fig. 4. This prototype features a phased array system, crafted using piezoelectric ceramics. Specifically, a 50 × 50 channel Phased Transducer Array (PTA) was meticulously constructed through lithography and wet etching techniques on a 1mm thick bulk lead zirconate titanate ceramic. This ceramic is integrated with nickel electrodes on either side. The array measures 50mm by 50mm and comprises 2500 distinct channels, each separated by a 1mm transducer spacing. This configuration allows for the generation of controllable acoustic field with an operating frequency of 2.3 MHz. A driver board, containing six Field-Programmable Gate Arrays (FPGAs) operating at 400 MHz, is employed to precisely control the HAE parameters. Positioned above the PZT, a transparent water tank made of acrylic, filled with deionized water, serves as the primary operating medium. The system's imaging capabilities are enhanced by a Prosilica GC2450 microscope camera, coupled with a Navitar zoom lens, which captures images at 15 frames per second. This setup is critical for both visual feedback and HAE forming. The system's operational workspace spans approximately 30 × 30 mm² at a designated height, offering a substantial area for conducting intricate particle manipulation tasks within the experimental framework.

Jacobian Calibration. The calibration is performed by unifying the actual space and the image controls by calibration using the image Jacobian matrix in Section II-B. After taking the feature points several times, the average reference point was obtained as $[x_{ref}, y_{ref}, z_{ref}]^T = [24.384, 23.384, 57.000]^T$ mm. The image Jacobian matrix was calibrated as

$$J_V^* = \begin{bmatrix} 15.2370 & -0.3242 \\ 0.3242 & 15.1907 \end{bmatrix} \mu m / pixel \tag{5}$$

The error of the cluster manipulation objective is evaluated by comparing the calculated position with the actual position. The results show that the error is 12 μm, which can meet the accuracy of particle cluster manipulation experiments.

Network Architecture and Training. The physics-based self-supervised learning algorithm we developed is implemented using PyTorch and trained on a Tesla M40 24GB GPU. We set the acoustic frequency at 2.32 MHz,

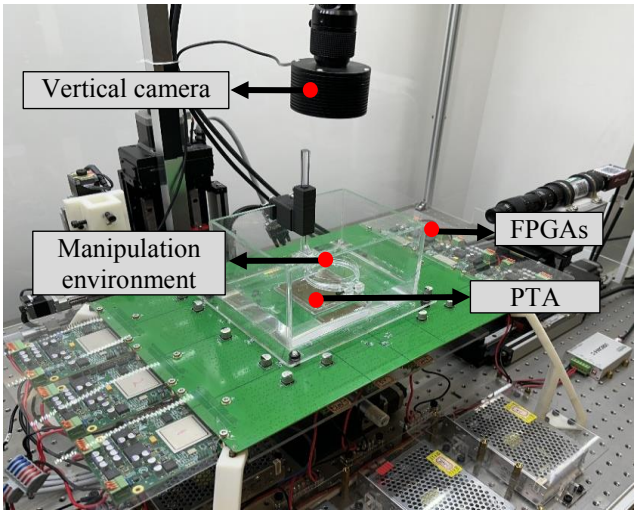

Fig. 4. The experimental platform consists of an PTA, a microscopic camera, and FPGAs.

mirroring real-world experimental conditions. The axial distance between source and target planes is fixed at 50 mm, aligning with our actual physical experiments. Our dataset has 6000 samples depicting variously sized square quantify HAE positioned arbitrarily in space, which is split into a training set and testing set with a ratio of 8:2. To ensure the physical relevance of the output, the predicted POH undergoes processing by a sigmoid function, which confines the predicted **P**s values within the [0, 1] range. This range is then rescaled to [0, 2π] to fit the acoustic phase range. The U-Net parameters are initialized using the Kaiming method. The batch size was set to 16. For the self-supervised learning aspect, we employed the Adam optimizer with an initial learning rate of $1e^{-3}$, along with default values for other hyper-parameters. The learning process spanned 200 epochs to ensure network convergence while preventing overfitting.

B. Accuracy and Real-time Performance of HAE Generation

In our experimental process, the reconstruction of the HAE proved to be the most time-consuming step, making it an ideal metric for assessing the algorithm's real-time capabilities. We benchmarked our method against established comparison methods, including IASA, IB, and DIFF-PAT, to gauge its performance. As highlighted in Table I, our proposed method demonstrates significant advantages in real-time performance compared to these methods. Once the neural network is fully trained, it can directly compute the source hologram. This capability indicates that our method is an effective tool for enabling dynamic and precise real-time operations, potentially useful in robotic applications. Quality of the reconstructed HAE is another critical metric, directly influencing the experiment's success rate in practical scenarios. The comparison in Table I reveals that our method achieves a higher Peak Signal-to-Noise Ratio (PSNR) compared to alternative algorithms, with an average PSNR of 22.64. This superior PSNR indicates greater accuracy and efficiency of our approach. In practical experiments, we

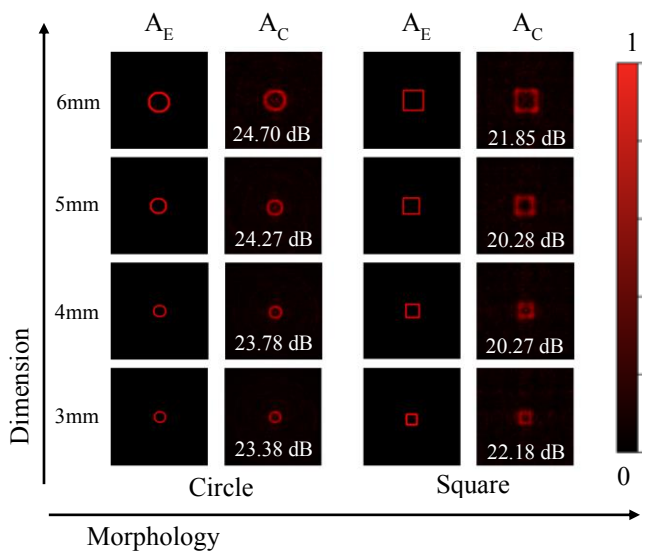

Fig. 5. The reconstruction results consist of the target amplitude hologram A_E and the reconstructed amplitude hologram A_C. The range of the amplitude is (0, 1), respectively. The reconstructed PSNR values are shown under the corresponding pictures. The reconstruction results can adjust the HAE according to morphology and dimensions.

selected the necessary components—target amplitude hologram $\mathbf{A_E}$, reconstructed amplitude hologram Ac, and predicted phase hologram POH—as demonstrated in Fig. 5. The highest recorded PSNR value was 24.70. By scanning the HAE using a hydrophone, we confirmed that the HAE generated by our method is conducive to completing real-world operations.

TABLE I. COMPARISON WITH SOTA ALGORITHMS ON REAL-TIME PERFORMANCE

Method	IASA	IB	DIFF-PAT	Ours
PSNR	19.53	20.07	21.54	**22.64**
Time (ms)	2.29	2.45	0.75	**0.008**

C. Aggregating, Rotating, and Translating Experiment

In order to verify the reliability and task feasibility of the deep learning algorithm, 200-400um polystyrene (PS) balls were randomly distributed on the bottom of the petri dish for the experiment.

In the aggregation experiment, particles often display a square shape spontaneously, which is related to the interaction force between particles in the liquid environment. In order to prove the effectiveness of the method more rigorously, we expect the final shape of particle clusters to be square, select our desired position, and generate the corresponding HAE to complete the aggregation manipulation task. First, set the initial HAE side length to 6 mm. By decreasing the side length of the HAE by 3 mm, the particles are collected, and a closed loop is introduced, so that more particles can be collected in a preset position. The closed-loop collection process is shown in Fig. 6(a). By comparing the open-loop and closed-loop collection processes, as shown in Table II, it can be clearly found that more expected particles can be collected through the closed-loop process.

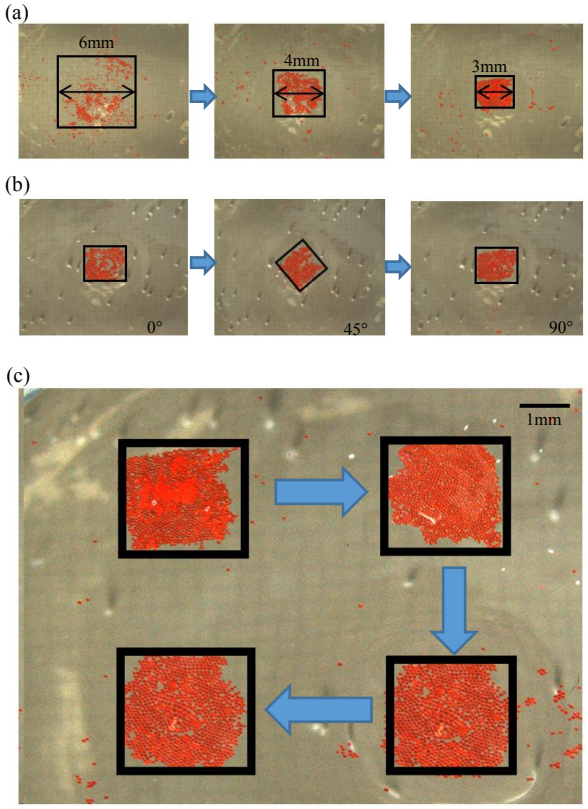

Fig. 6. (a), (b) and (c) respectively represent the process of particle cluster aggregation and rotation and translation, where the black line represents the actual position and shape of the HAE.

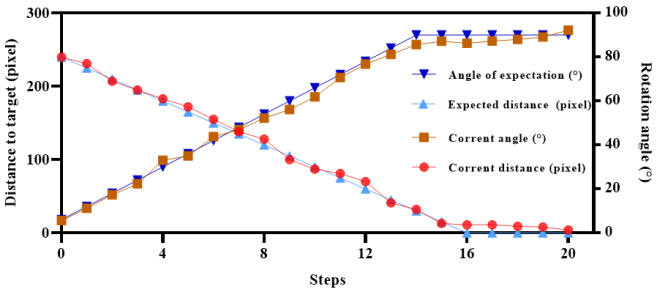

Fig. 7. The experimental results of rotation and translation show that the rotation angle error and positioning error converge with the number of cycles.

Fig. 6(b) and Fig. 6(c) respectively show the rotating and translating capabilities of particle clusters. The PID loop closure is introduced to better make the position and angle converge to the desired The best integration coefficient KI was found to be 0.2 through experiments. The experimental results are shown in Fig. 7. With the introduction of the closed loop, the position error and angle error continue to decrease until our expectation is reached. Among them, the termination error of single positioning can be as small as 64 μm, and the angle error after rotation is less than 2.2°, which meets the requirements of particle clusters. This margin of error is permissible. The completion of these tasks shows that by combining deep learning methods and selecting good control strategies, the designed system can achieve accurate positioning and angle control of particle clusters.

TABLE II. PARTICLE AGGREGATION EXPERIMRNT RESULTS

Samples	Open-loop aggregate rate (%)	Close-loop aggregate rate (%)
1	77	100
2	82	100
3	82	100
4	75	100
5	74	100

VI. CONCLUSION

This study develops a regulatable and versatile HAE system empowered with a customized deep learning approach and a closed-loop control algorithm, enabling the dynamic manipulation of particle clusters at the micron level, including their aggregation, rotation, and translation. Additionally, the visual feedback from microscope system enhances manipulation stability and robustness in actual environments. The extensive experiments affirm the system's ability to facilitate particle aggregation, translation, and rotation. These parameters fulfill the precision requirements for spatial manipulations of particle clusters by micro-nanorobots. This research lays the groundwork for the future application of acoustic-driven particle cluster across various domains. Moving forward, the focus on refined acoustic manipulation of the kinematics and morphology of particle clusters. The objective is to widen the scope of this technology's applications, extending its benefits to diverse areas within biological and medical engineering, including but not limited to cell analysis and targeted drug delivery.

REFERENCES

[1] A. G. El-Gazzar, L. E. Al-Khouly, A. Klingner, S. Misra, and I. S. Khalil, "Non-Contact manipulation of microbeads via pushing and pulling using magnetically controlled clusters of paramagnetic microparticles," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 778-783, 2015.

[2] C. Gan, S. Liang, F. Wang, Y. Cao, Y. Ji, J. Lina, and L. Feng, "Non-contact massively parallel manipulation of micro-objects by optoelectronic tweezers," *WRC Symposium on Advanced Robotics and Automation*, pp. 7-12, 2021.

[3] J. A. Seon, Z. Cenev, and Q. Zhou, "Automatic noncontact extraction and independent manipulation of magnetic particles using electromagnetic needle," *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 2, pp. 931-941, 2019.

[4] X. Chen, K. H. Lam, R. Chen, Z. Chen, X. Qian, J. Zhang, and Q. Zhou, "Acoustic levitation and manipulation by a high-frequency focused ring ultrasonic transducer," *Applied Physics Letters*, vol. 114, no. 5, 2019.

[5] A. Ashkin, and J. M. Dziedzic, "Optical trapping and manipulation of viruses and bacteria," *Science*, vol. 235, no. 4795, pp. 1517-1520, 1987.

[6] I. D. Vlaminck, and C. Dekker, "Recent Advances in magnetic tweezers," *Annu. Rev. Biophys*, vol. 41, pp. 453-472, Jan. 2012.

[7] X. Chen, C. Zhou, Y. Peng, Y. Q. Wang, and W. Wang, "Temporal light modulation of photochemically active, oscillating micromotors: Dark pulses, mode switching, and controlled clustering," *ACS applied materials & interfaces*, vol. 12, no. 10, pp. 11843-11851, 2020.

[8] H. Miyazako, K. Mabuchi, and T. Hoshino, "Spatiotemporal control of electrokinetic transport in nanofluidics using an inverted electron-beam lithography system," *Langmuir*, vol. 31, no. 23, pp. 6595-6603, 2015.

[9] Y. Yang, Y. Yang, D. Liu, Y. Wang, M. Lu, Q. Zhang, and H. Zheng, "In-vivo programmable acoustic manipulation of genetically engineered bacteria," *Nature Communications*, vol.14, no. 1, p. 3297, 2023.

[10] Y. Yang, T. Ma, Q. Zhang, J. Huang, Q. Hu, Y. Li, and H. Zheng, "3D acoustic manipulation of living cells and organisms based on 2D array," *IEEE Transactions on Biomedical Engineering*, vol.69, no. 7, pp. 2342-2352, 2022.

[11] S. Liang, C. Gan, Y. Dai, C. Zhang, X. Bai, S. Zhang, and L. Feng, "Interaction between positive and negative dielectric microparticles/microorganism in optoelectronic tweezers," *Lab on a Chip*, vol. 21, no. 22, pp. 4379-4389, 2021.

[12] K. Melde, A. G. Mark, T. Qiu, and P. Fischer, "Holograms for acoustics," *Nature*, vol. 537, no. 7621, pp. 518-522, 2016.

[13] C. W. Shields Iv, C. D. Reyes, and G. P. López, "Microfluidic cell sorting: a review of the advances in the separation of cells from debulking to rare cell isolation," *Lab on a Chip*, vol. 15, no. 5, 1230-1249, 2015.

[14] C. Zhong, Y. Jia, D. C. Jeong, Y. Guo, and S. Liu, "Acousnet: A deep learning based approach to dynamic 3d holographic acoustic field generation from phased transducer array," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 666-673, 2021.

[15] C. Zhong, Q. Lu, T. Li, H. Su, and S. Liu, "Real-time acoustic holography with physics-reinforced contrastive learning for acoustic field reconstruction," *Journal of Applied Physics*, vol. 135, no. 1, 2024.

[16] A. Marzo, and B. W. Drinkwater, "Holographic Acoustic Tweezers," *Proc. Natl. Acad. Sci. (PNAS)*, vol. 116, no. 1, pp. 84-89, Jan. 2019.

[17] T. Fushimi, K. Yamamoto, and Y. Ochiai, "Acoustic hologram optimisation using automatic differentiation," *Scientific reports*, vol. 11, no.1, p. 12678, 2021.

[18] Q. Lu, C. Zhong, Q. Liu, T. Li, H. Su, and S. Liu, "Ultrafast Acoustic Holography with Physics-Reinforced Self-Supervised Learning for Precise Robotic Manipulation", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2673-2678, 2023.

[19] M. H. Lee, H. M. Lew, S. Youn, T. Kim, and J. Y. Hwang, "Deep learning-based framework for fast and accurate acoustic hologram generation," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 12, pp. 3353-3366, 2022.

[20] C. Cai, N. Somani, and A. Knoll, "Orthogonal image features for visual servoing of a 6-DOF manipulator with uncalibrated stereo cameras," *IEEE transactions on Robotics*, vol. 32, no. 2, pp.452-461, 2016.

[21] Y. Yang, T. Ma, Q. Zhang, J. Huang, Q. Hu, Y. Li, and H. Zheng, "3D acoustic manipulation of living cells and organisms based on 2D array," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 7, pp. 2342- 2352, 2022.

[22] K.Yiannacou, and V. Sariola, "Controlled manipulation and active sorting of particles inside microfluidic chips using bulk acoustic waves and machine learning," *Langmuir*, vol. 37, no. 14, pp. 4192-4199, 2021.