

<https://doi.org/10.1038/s44172-024-00270-9>

Multi-view neural 3D reconstruction of micro- and nanostructures with atomic force microscopy

Check for updates

Shuo Chen ^{1,3}, Mao Peng ^{2,3}, Yijin Li ¹, Bing-Feng Ju², Hujun Bao¹, Yuan-Liu Chen ² ✉ & Guofeng Zhang ¹ ✉

Atomic Force Microscopy (AFM) is a widely employed tool for micro- and nanoscale topographic imaging. However, conventional AFM scanning struggles to reconstruct complex 3D micro- and nanostructures precisely due to limitations such as incomplete sample topography capturing and tip-sample convolution artifacts. Here, we propose a multi-view neural-network-based framework with AFM, named MVN-AFM, which accurately reconstructs surface models of intricate micro- and nanostructures. Unlike previous 3D-AFM approaches, MVN-AFM does not depend on any specially shaped probes or costly modifications to the AFM system. To achieve this, MVN-AFM employs an iterative method to align multi-view data and eliminate AFM artifacts simultaneously. Furthermore, we apply the neural implicit surface reconstruction technique in nanotechnology and achieve improved results. Additional extensive experiments show that MVN-AFM effectively eliminates artifacts present in raw AFM images and reconstructs various micro- and nanostructures, including complex geometrical microstructures printed via two-photon lithography and nanoparticles such as poly(methyl methacrylate) (PMMA) nanospheres and zeolitic imidazolate framework-67 (ZIF-67) nanocrystals. This work presents a cost-effective tool for micro- and nanoscale 3D analysis.

The investigation of the three-dimensional (3D) structure plays a vital role in nanotechnology research, encompassing areas like nanofabrications^{1,2}, nanorobots^{3,4}, and nanomedicines^{5,6}, given its critical relevance to the functional properties of micro- and nanoscale objects. The scanning electron microscope (SEM)⁷ is a prevalent tool for providing qualitative information about the geometry of micro- and nanostructures. This technique involves irradiating the sample with an electron beam and capturing a 2D image by detecting the intensity of secondary electrons emitted from the sample surface. Despite its widespread use and multiple advantages, SEM is a destructive method⁸, requires a vacuum environment, and cannot provide accurate height information in the images. In contrast, the atomic force microscope (AFM)⁹ acquires precise height information of the sample surface through the forces between its probe and the sample. Moreover, AFM can operate in various environments, is insensitive to the sample material, and is non-destructive.

Nonetheless, conventional AFM comes with its own set of challenges. One primary limitation is that conventional AFM can only capture 2.5D¹⁰ information instead of a complete 3D representation of the sample because the position feedback in conventional AFM systems is confined to the

vertical direction⁹. Another challenge is the issue of tip-sample convolution¹¹. This phenomenon arises from geometrical interactions between the AFM tip and the surface features of the sample (Supplementary Fig. 1). These interactions often lead to artifacts^{11–13} in the scanning results that are inherently difficult to differentiate from the actual sample geometry. Such limitations impede the effective use of AFM to investigate intricate 3D micro- and nanostructures and catalyze the development of advanced AFM technologies, i.e., 3D-AFM.

The advancement of 3D-AFM technology predominantly follows two distinct trajectories. The first approach involves the design of specialized probe shapes aimed at enabling the measurement of structures that are inaccessible with conventional AFM scanning. As an example, critical dimension AFM (CD-AFM)^{14,15}, currently a prevalent method for semiconductor structures, manipulates the lateral dithering of a flared tip to capture lateral geometric information of the sample. These designs equip the AFM with the capability to image not only vertical but also undercut sidewall structures. Additionally, there are other creative designs of probes, such as the introduction of hinge structures¹⁶, orthogonal cantilevers¹⁷, and probes made of carbon nanotube with high aspect ratios¹⁸. However, the

¹State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China. ²State Key Lab of Fluid Power&Mechatronic Systems, Zhejiang University, Hangzhou, China.

³These authors contributed equally: Shuo Chen, Mao Peng. ✉e-mail: yuanliuchen@zju.edu.cn; zhangguofeng@zju.edu.cn

extra cost and complexity of manufacturing special probes and customized AFM scanning systems present substantial challenges to the widespread adoption of these methods.

Another common technique for 3D-AFM is the practice of tilting either the probe^{19–23} or the sample^{24,25} to scan micro- and nanostructures from multiple directions. These methods avoid the necessity for specialized flared tips, instead relying on integrating multiple scans into a complete 3D model. As a result, the effectiveness relies on the precision of the data stitching of multiple scans. Historically, previous methods^{19–25} predominantly apply to simple, well-defined structures, such as gratings, for their CD metrology. The grating's relatively straightforward structure facilitates the manual removal of artifacts resulting from tip-sample convolution while only requiring two tilt scans towards both sidewalls to obtain complete geometric information. This greatly simplifies the problem of the tilting method. However, substantial challenges arise when the tilting method is applied to micro- and nanostructures of unknown and intricate shapes, such as those created by Two-photon Lithography (TPL)^{26–29} or comprised of diverse nanoparticles^{5,6,30}—a scenario frequently encountered in nanotechnology research. Firstly, the complex surface geometries of these structures make it difficult to manually identify and remove artifacts from the AFM images. Secondly, the existence of these unremoved artifacts in the scans impedes the accuracy of the stitching process. Thirdly, owing to the complex overlapping relationships among multi-view data, simply stitching of these data is insufficient for the construction of a clear and accurate 3D model.

In this study, we propose MVN-AFM, a framework that is able to reconstruct the 3D surface model of a wide range of complex-shaped micro- and nanostructures without any specially shaped probes or costly modifications to the AFM system. Our framework leverages the concept of tilting samples, but we extend its application for complex structures beyond the limitations of existing methods. Specifically, we propose an iterative optimization algorithm to automatically remove AFM artifacts and improve the alignment accuracy of multi-view data from intricate micro- and nanostructures. Compared with previous tilt methods that only consider two tilt scans, our method can handle multi-view AFM data from more tilt angles, eight in our experiments. Subsequently, in order to reconstruct the 3D model of these structures by multi-view AFM data, we draw inspiration from multi-view depth fusion techniques^{31–36} in computer vision. We introduce the neural implicit surface reconstruction methods^{37–42}, the recent advance in this field, to utilize a neural network to represent the 3D model of micro- and nanostructures. By employing differentiable volume rendering to train the neural implicit function with multi-view AFM data supervision, we fuse the multi-view scanning results into an accurate and comprehensive 3D model. Furthermore, we conduct extensive experiments to evaluate the capabilities of the MVN-AFM framework. In detail, we utilize the TPL technique to fabricate various 3D microstructures with distinct geometrical characteristics and prepare specimens of commonly used nanoparticles, including poly(methyl methacrylate) (PMMA) nanospheres^{43–45} and zeolitic imidazolate framework-67 (ZIF-67) nanocrystals^{46–49}. MVN-AFM effectively eliminates artifacts present in raw AFM images and successfully reconstructs not only the overall shape but also specific hidden details that are not discernible in conventional AFM scans. The ability of MVN-AFM to provide detailed and accurate 3D reconstructions of a broad spectrum of micro- and nanostructures, coupled with its low implementation cost, positions it as a potentially valuable tool in nanotechnology research.

Results

Pipeline of MVN-AFM

The objective of MVN-AFM is to construct a generalized process that can be used for 3D reconstruction of unknown-shaped complex micro- and nanostructures based on multi-view AFM scanning data, relying only on conventional AFM systems and standard probes. MVN-AFM consists of three main steps (Fig. 1): multi-view AFM scanning, data alignment and mask solving, and neural implicit surface reconstruction.

The step of Multi-view AFM Scanning (Fig. 1a) captures multi-view AFM images of 3D micro- and nanostructures, providing essential geometric information for the subsequent reconstruction process. Previous tilting methods^{19–25} acquire complete geometric information with only two AFM scans towards each sidewall of the grating. However, in nanotechnology research, the prior knowledge of the sample's shape and orientation is often unknown. To address this, we design a multi-view AFM scanning process that is independent of the sample's shape. This process aims to comprehensively acquire the surface geometric information of unknown and complex-shaped structures. We use the sample-tilting approach to collect multi-view data, thus avoiding modifications to the mechanics of conventional AFM. For this purpose, we designed a rotatable stage with a tilt angle (Supplementary Fig. 2). In our experiment, the tilt angle is set to 30°, the largest tilt angle that prevents the cantilever of the AFM probe from hitting the sample (see Methods for further details). We carefully designed the size of the whole stage so that it can be used in the limited activity space of a commercial AFM without any collision. The sample is placed on a turntable in the center of the stage so that multi-view scans around the sample can be acquired as it rotates to different directions. We performed eight tilt scans for each sample, and these scan directions are evenly distributed around 360° of the sample's surroundings. Together with a scan without sample tilt, a total of nine multi-view scans are acquired per sample.

In the step of Data Alignment and Mask Solving (Fig. 1b), we iteratively align multi-view AFM data to a unified coordinate system and remove the artifacts in AFM images. A critical process in the tilting method involves establishing the spatial relationship among multi-view data. This requires determining a coordinate transformation, denoted as pose T , to align data from different views within the same coordinate system (Supplementary Fig. 3). To achieve this, some methods^{19,22,23} employ designs with high-cost components to enable precise control of probe scanning direction, which allows direct access to T . Others^{20,24,25} utilize the Iterative Closest Point (ICP) algorithm⁵⁰ to solve T by minimizing the distance between AFM data points. The ICP algorithm relies heavily on data free from artifacts that do not represent the actual sample shape. Consequently, these previous methods^{20,24,25} manually remove highly recognizable artifacts from AFM data of simple structures before using the ICP algorithm. However, for multi-view images obtained by a conventional AFM system on intricate structures, there are two challenges: eliminating artifacts and solving for the pose. Here, we define the label of whether each data point is an artifact as a latent variable, mask M . It determines whether the measurement from each AFM pixel is utilized in the reconstruction process. To simultaneously solve for T and M , we propose an iterative EM-like algorithm^{51,52}. Initially, we consider all AFM data artifact-free, i.e., M_0 is all zeros, and directly apply the ICP algorithm to obtain a set of coarse poses, T_0 . As ICP also considers AFM artifacts when minimizing the point set distance between multi-view data, T_0 is coarse and contains many errors. In the E-step, we project multi-view data onto each other using T_{i-1} from the previous iteration $i-1$. We then conduct cross-validation of the projected data to identify areas of inconsistency in multi-view data. These regions are then labeled as ones, and we obtain the updated M_i . The motivation for the cross-validation is that artifacts vary with the probe-sample angle, so they are inconsistent in multi-view data. In contrast, the sample's geometric surface remains consistent across different views, regardless of the probe scanning directions. In the M-step, we erase the artifact through M_i and apply the ICP algorithm again to compute the updated T_i . Iterating the EM steps, the data filtered out of most artifacts by M yields a more precise T . The improved T also makes the cross-validation accurate. Two steps are iteratively performed to enhance each other. After k iterations, we obtain the accurate pose T_k and the artifact mask M_k for each viewpoint of the AFM data.

The step of Neural Implicit Surface Reconstruction (Fig. 1c) utilizes the aligned and masked AFM data to train an implicit function represented by a neural network and extract the final 3D surface model of micro- and nanostructures from the network. Specifically, we follow previous work⁵³ and model the geometry surface of the sample as a neural network encoded

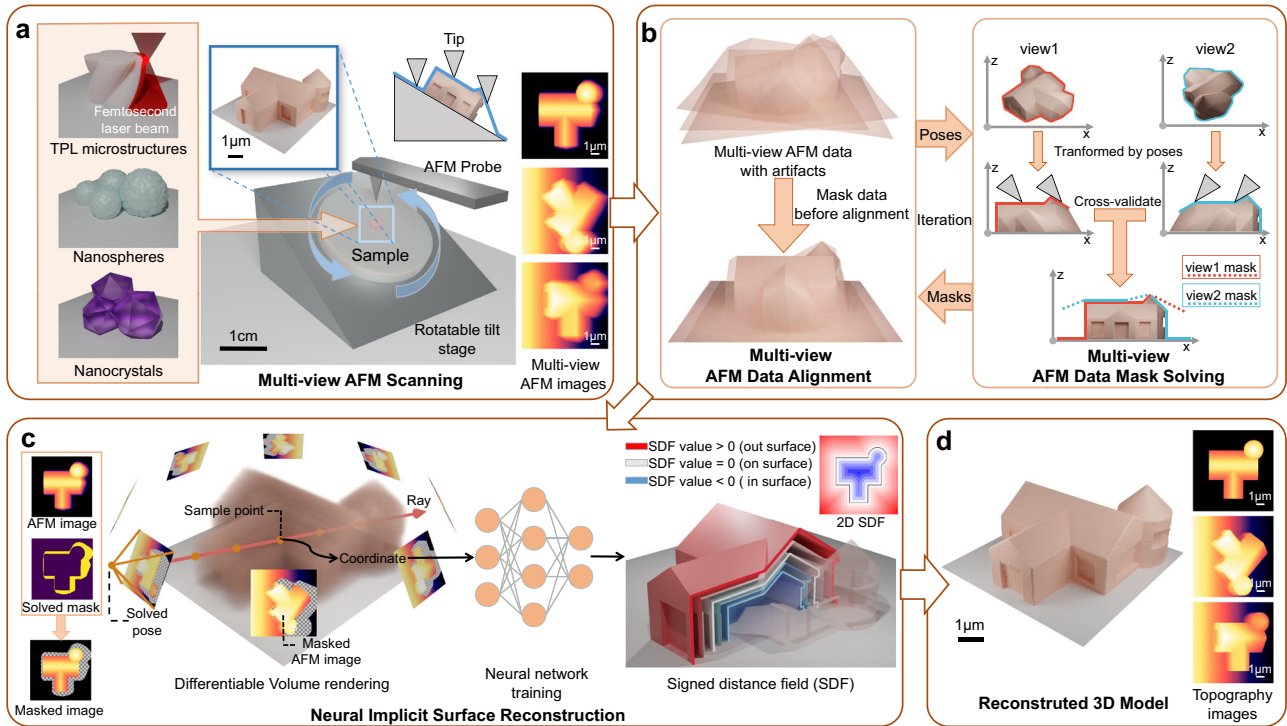


Fig. 1 | The pipeline of MVN-AFM. **a** First, we place various micro- and nanostructures on a rotatable tilt stage, including two-photon lithography (TPL) microstructures and some commonly used nanoparticles. Second, we rotate the turntable and measure the vertical heights by conventional atomic force microscopy (AFM), resulting in a set of multi-view AFM images with many artifacts. **b** Input raw AFM images with artifacts and iterate two sub-steps. In the data alignment process, data judged as artifacts are eliminated before alignment, and the poses of multi-view

images are updated. In the mask-solving process, the solved poses transform the multi-view data, and the data consistency is cross-validated to solve the mask of artifacts. **c** The posed and masked multi-view AFM images are used to train a neural network representing a signed distance field in space by the differentiable volume rendering technique (Supplementary Fig. 5). **d** The 3D surface model extracted from the signed distance field, and corresponding topography images without artifacts.

Signed Distance Field (SDF): $s(x; \theta) : \mathbb{R}^3 \rightarrow \mathbb{R}$, where x denotes a 3D position and θ is the parameters of a Multilayer Perceptron (MLP). The SDF defines a scalar field where each point in space is associated with the shortest distance to a surface. This distance is positive if the point is outside the surface and negative if it is inside. Previously, the neural implicit surface reconstruction methods⁴² were developed for posed images from cameras in the macroscopic world, not for nanotechnology and AFM data applications. To adapt this method to our reconstruction process, we convert AFM images into depth maps as captured by virtual orthogonal cameras (Supplementary Fig. 4). Each pixel in AFM images transformed by pose T and filtered by mask M represents a sample ray. The loss function is the disparity between the AFM data and the depth value derived from differentiable volume rendering along the ray. We then optimize the MLP network parameters θ through back propagation⁵⁴. Moreover, we also use the multi-resolution hash encoding technique⁵⁵ to accelerate the training process. Upon completing the training, we can query the SDF value of any spatial point by inferring the network. Based on the fact that the zero set of SDF represents the structure surface, the Marching Cubes algorithm⁵⁶ is finally utilized to extract the 3D surface model of the micro- and nanostructures (Fig. 1d).

Reconstruction of two-photon lithography structures

In this section, we evaluate the proposed MVN-AFM on microstructures printed by TPL technology. The TPL technology, which focuses a femtosecond laser into tiny voxels in a photosensitive resist, enables 3D printing of a given computer-aided design (CAD) model with sub-100 nm resolution through the two-photon polymerization (TPP) process⁵⁷. To fully demonstrate the performance of MVN-AFM on complex 3D microstructures, we printed a set of samples with different geometrical features. Specifically, we printed six structures (Supplementary Fig. 6): cylinder, undercut, spiral,

gear, monkey, and house. For centrosymmetric structures, we incorporated three small cones around each microstructure to indicate their orientation, as depicted in the first four rows of Fig. 2. This step is unnecessary for non-centrosymmetric structures, such as the monkey and the house. The height of these microstructures varies between 2 μm and 3.5 μm . We performed AFM scans in tapping mode, with a scan size of 10 $\mu\text{m} \times 10 \mu\text{m}$ and 256 lines of 256 points for each AFM image, corresponding to a pixel size of 39 nm. Because the minimum feature size of structures printed by the commercial photoresist in our experiment is 160 nm (see “Methods” section for further printing details), the pixel density we choose is sufficient to capture the geometry feature of the TPL samples.

The cylinder (Fig. 2a) is a representative structure that challenges conventional AFM scanning⁵⁸ and previous tilting methods. Unlike grating structures, the vertical annular sidewall cannot be divided into distinct left and right sections. Next, the undercut (Fig. 2b) is a prevalent structural feature in semiconductor manufacturing⁵⁹. This structure differs from the cylinder by having a sloped sidewall. We further constructed the gear (Fig. 2c), a mechanical structure frequently encountered in Micro-Electro-Mechanical systems (MEMS)⁶⁰. The spiral (Fig. 2d) is distinguished by an intricate array of rotating curved concave and convex structures on its sidewall. Furthermore, we also conducted tests using the Suzanne Monkey (Fig. 2e), a standard model in computer graphics⁶¹. Unlike the previous columnar structures, this model poses unique challenges due to its curved features and the indistinct boundary between its top surface and sidewalls. We finally designed a house structure (Fig. 2f) that included shapes with both planar and curved features, along with detailed elements like grooves on the sidewalls to represent doors and windows.

In conventional AFM scanning, the results are a mixture of incomplete surface geometry and artifacts, which do not accurately represent the sample surface. As illustrated in Fig. 2g and h, despite the vast difference in sidewall

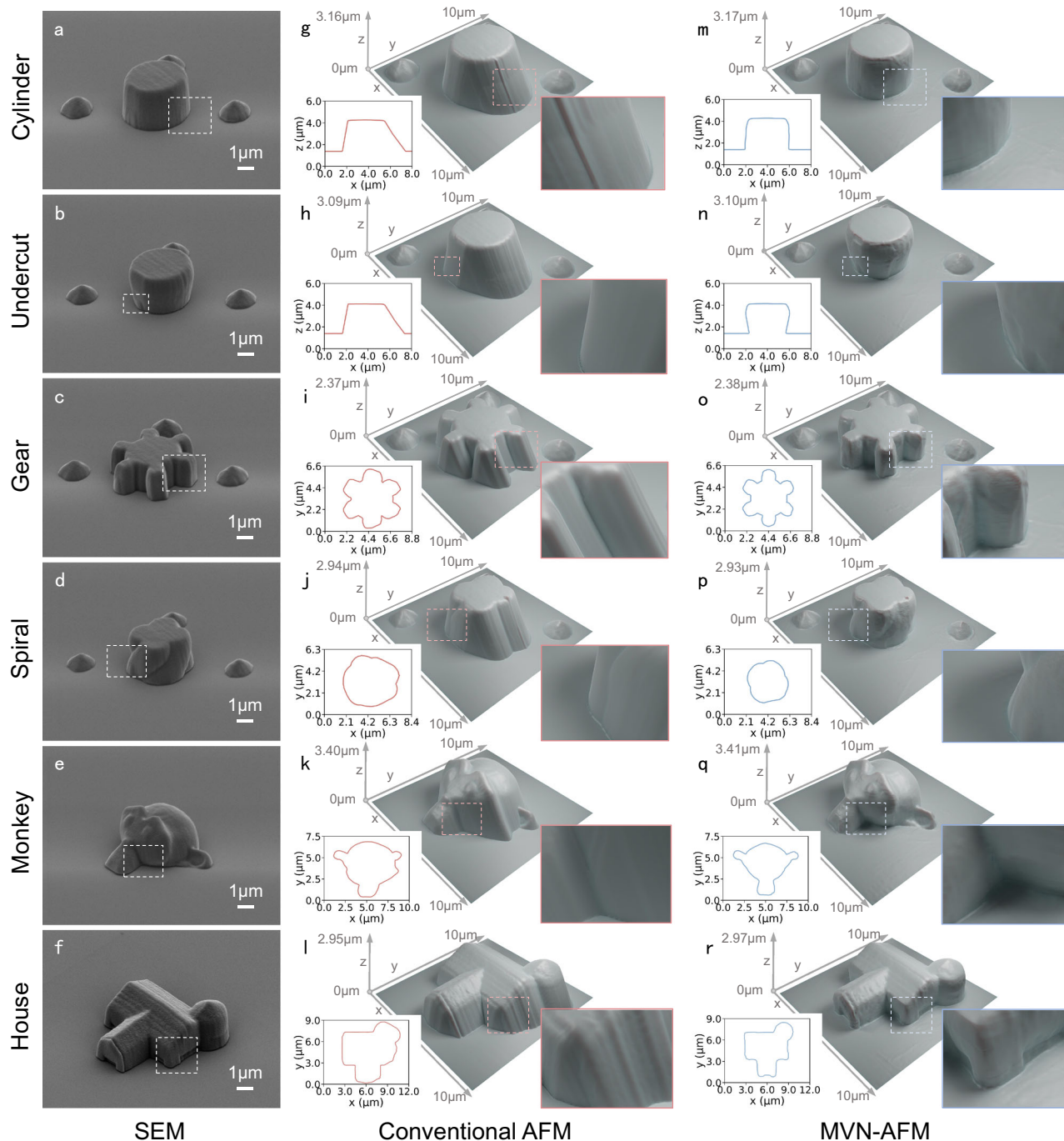


Fig. 2 | MVN-AFM reconstructs the surface models of two-photon lithography (TPL) microstructures. a–f Scanning Electron Microscope (SEM) images of TPL microstructures. **g–l** 3D models of TPL microstructures’ conventional AFM scanning data. **m–r** 3D models of TPL microstructures reconstructed by MVN-AFM.

g, h, m, n Include the cross-section profiles in the x-z plane. **i–l, o–r** Include the cross-section profiles in the x-y plane. More visualizations can be found in Supplementary Fig. 7 and Supplementary Movie 1.

geometry, the scanning result of the undercut is indistinguishable from that of the cylinder model. Some detailed features are also virtually invisible, such as the doors and windows in the house model (Fig. 2l). The cross-sectional profiles reveal considerable distortion of these scanning results, which may lead researchers to misjudge the actual shape of these samples. Moreover, it is obvious that manually separating artifacts from the AFM scans of these intricate structures is almost impractical.

In contrast, the proposed MVN-AFM framework effectively eliminates artifacts while precisely merging geometric information from multi-view AFM scanning into accurate and comprehensive 3D models. These reconstructed models align consistently with SEM images and

demonstrate the surface of these samples. These models clearly differentiate between the cylinder (Fig. 2m) and undercut (Fig. 2n) structure, precisely reconstruct the gear’s teeth (Fig. 2o), and capture the correct orientation of the spiral threads (Fig. 2p) and the monkey’s subtly inward-curving side faces (Fig. 2q). Even the minutely detailed grooves (Fig. 2r) on the house sidewalls are observable.

Reconstruction of nanoparticles

To further demonstrate the generalization of MVN-AFM on structures with smaller sizes and different geometry features, we selected some widely used nanoparticles, including PMMA nanospheres and ZIF-67 nanocrystals.

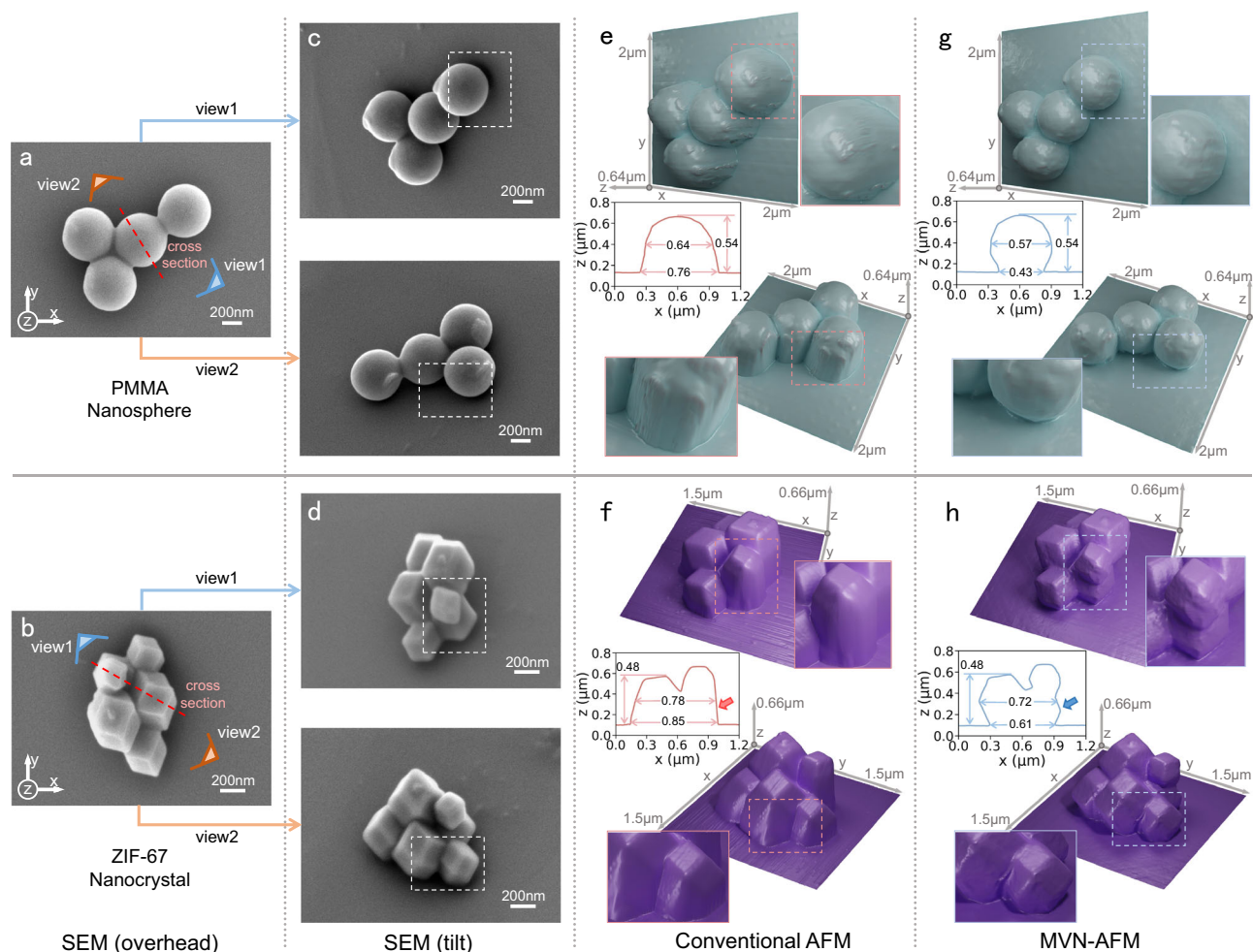


Fig. 3 | MVN-AFM reconstructs the surface models of poly(methyl methacrylate) (PMMA) nanospheres and zeolitic imidazolate framework-67 (ZIF-67) nanocrystals. a, b SEM images of nanoparticles in the overhead view. **c, d** SEM images of nanoparticles in the tilt view. **e, f** 3D models and cross-sections of nanoparticles' conventional AFM scanning data. **g, h** 3D models and cross-sections of nanoparticles reconstructed by MVN-AFM. More visualizations can be found in Supplementary Movie 2.

Nanospheres are a common type of nanoparticle with a wide range of applications⁶². The characteristics of nanospheres depend greatly on their size and surface structure⁶, making accurate 3D reconstruction valuable for their research. To evaluate the effectiveness of our proposed method on spherical structure, we chose PMMA^{43–45} with a diameter of about 500 nm, a widely used type of polymeric nanosphere. In Fig. 3e, it is evident that the artifacts in the conventional AFM scanning data are seamlessly connected with the top curved surface of the nanospheres, and the overall shape does not exhibit a spherical appearance. Under these circumstances, manually distinguishing artifact boundaries in AFM scanning as in previous methods becomes unachievable, and the details on the sides of the nanospheres are entirely lost, posing a challenge for researchers to accurately determine the size and structure of these nanospheres. In contrast, MVN-AFM demonstrates its advanced capabilities by accurately reconstructing several adherent nanospheres, each mirroring the shape observed in SEM photographs (Fig. 3c). The reconstruction result of our method is the deepest penetration made by the tip. The multi-view scanning enables the tip to approach the sample's surface closely, reducing the errors in the results. As shown in the cross-sections in Fig. 3e, g, the model of MVN-AFM has a center width very close to the height, which is consistent with the characteristics of a sphere. However, it is impossible to get the geometric information of structures that the tip cannot touch, such as the region near the bottom of the nanosphere that is tangential to the substrate plane. Therefore, the bottom width of the reconstructed model by MVN-AFM is smaller than the conventional AFM result but not close to zero.

Next, we selected ZIF-67, a cubic symmetric nanocrystal, as a representative crystal-like nanoparticle to assess the effectiveness of our method. ZIF-67⁴⁶ and its derivatives exhibit various excellent properties, leading to their extensive attention and research^{47–49}. The morphological characteristics and size of nanocrystals can be tailored by manipulating experimental conditions during synthesis, leading to variations in their properties^{63–65}. Therefore, obtaining accurate 3D surface models of nanocrystals is of paramount importance. In the SEM images (Fig. 3d), the ZIF-67 nanocrystals exhibit a distinct polyhedral shape, ranging in size from about 100 nm to 500 nm. However, conventional AFM results (Fig. 3f) only partially demonstrate the top surface of the crystals, resulting in an overall blurred representation of the particles' shape. Furthermore, in scenarios where multiple crystals aggregate, as illustrated in our example, only the uppermost crystal in the stack is visible in the conventional AFM scanning (Fig. 3f, the arrow in cross-section), with the underlying crystal completely obscured by the top crystal and associated probe artifacts. On the contrary, the surface model reconstructed by our method (Fig. 3h) accurately captures the polyhedral shape of the ZIF-67 crystals, delineating their side planes and edges with precision. Even in cases where the particles are stacked up, the MVN-AFM method successfully reveals the bottom crystal (Fig. 3h, the arrow in cross-section), typically obscured in conventional scans, and accurately represents the arrangement of the particles in the stack, aligning with the SEM photograph (Fig. 3d).

In our nanoparticle experiments, PMMA nanospheres and ZIF-67 nanocrystals differ markedly from the previous TPL microstructures in

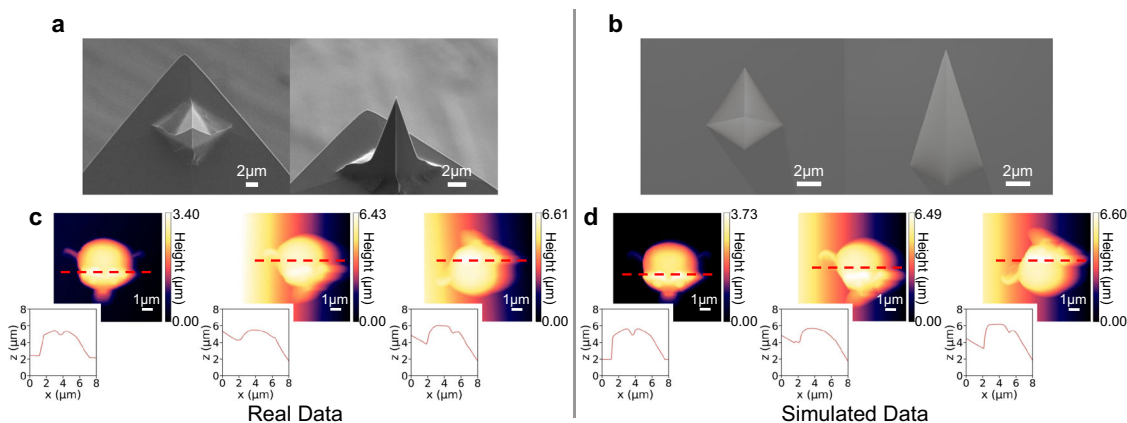


Fig. 4 | The consistency between our simulated AFM data and the real experiment data. a SEM images of the real AFM probe used in the two-photon lithography experiment. **b** The probe model in the simulation environment. **c** The

real AFM images captured by the multi-view tilt scanning of the monkey structure in the two-photon lithography experiment. **d** Simulated multi-view AFM images constructed by simulating the collision between the probe and the structure surface.

terms of material compositions, geometric features, and particle sizes. MVN-AFM precisely reconstructs these diverse samples by the exact same procedure and parameters, showcasing its generalizability and potential applicability in a broad spectrum of micro- and nanostructure research.

Evaluation of simulated data

To complement the previous qualitative comparisons on real experimental data, we embarked on quantitative evaluations using a set of simulated AFM data. We generated these data based on the CAD models of structures in the TPL experiment. The simulation environment allows for the precise determination of the spatial relationships between multi-view AFM data and access to an accurate surface model of the sample, a feat challenging to achieve in real-world experiments. To ensure that the simulated data closely mimics real AFM scanning conditions, we developed a simulated probe model. This model is based on the quadrilateral pyramid probe (Fig. 4a) utilized in our TPL experiments. Considering the nanoscale curvature of the actual AFM probe is negligible compared to the microscale dimensions of the TPL samples, we simplified the probe representation into a pyramid shape (Fig. 4b). The simulation of AFM scanning was then carried out by modeling the rigid body collision⁶⁶ between the probe and the sample models. As depicted in Fig. 4c, d, the simulated data exhibit a high degree of similarity to the real AFM data in terms of the overall shape and the presence of artifacts.

First, we focus on showcasing the enhancements MVN-AFM brings to the alignment accuracy of multi-view AFM data. In the alignment process, we estimate the pose T_{est} for each viewpoint. We also acquire the corresponding accurate pose T_{gt} in our simulation environment. Here, we respectively calculate the absolute pose error (APE) for the rotation and translation components between the estimated and ground truth poses of multi-view AFM data (see Methods for further details of error calculation). As illustrated in Fig. 5, we present a comparative analysis between the alignment method in MVN-AFM and the direct ICP alignment of raw AFM data, which includes artifacts. The analysis reveals that MVN-AFM achieves a substantial improvement in alignment accuracy, evidenced by an impressive average reduction of 46% in rotation errors and 27% in translation errors. These results not only demonstrate the negative impact of artifacts present in AFM data on the precision of data alignment but also highlight the efficacy of MVN-AFM in mitigating these challenges.

In the subsequent analysis, we compare the models reconstructed by two prominent multi-view depth fusion techniques: the neural implicit method and Truncated Signed Distance Function (TSDF) Fusion³¹. Our method interprets AFM images as depth images from virtual orthogonal cameras, framing the challenge as the depth fusion problem in computer vision. Depth fusion techniques are categorized into traditional^{31–33} and neural implicit methods^{34–36,41,42}. The TSDF Fusion is a widely used

traditional method that efficiently fuses multi-view depth data by dividing the 3D space into weighted discrete voxels and updating these weights according to the depth information along the pixel ray. However, multi-view AFM scanning of micro- and nanostructures presents unique challenges, particularly the uneven sampling density (Supplementary Fig. 9a) due to restricted tilt angles and limited viewpoints during the scanning process. This limitation often leads to regions with sparse sampling, such as the sidewall grooves of the spiral model (Fig. 6c). In the context of TSDF Fusion (Supplementary Fig. 9b), unintersected voxels in sparsely sampled regions demonstrate as voids in the reconstructed model, a limitation evident in Fig. 6a. Conversely, the neural implicit method, which represents the 3D model as a continuous neural network, exhibits its ability to construct a smooth and complete surface model, even with limited sample points, as depicted in Fig. 6b. This capability of the neural implicit method to effectively handle sparse data and reconstruct intricate surfaces makes it more suitable for the 3D reconstruction of multi-view AFM data in the MVN-AFM framework.

Finally, we evaluated the accuracy of the topography in the 3D surface models reconstructed by MVN-AFM. Our simulation environment enables the capture of precise surface topography unaffected by the probe's shape. The difference between accurate surface topography and the AFM images reveals substantial artifacts (Fig. 6e), particularly around the edges and at the sharper geometric features of the structure in the raw AFM data. These results underscore the complexity of artifacts in AFM images of intricate structures and highlight the challenges associated with their manual removal. The visualization of the difference between the topography images from 3D models of MVN-AFM and the accurate topography images (Fig. 6f) clearly indicates that MVN-AFM is highly effective in eliminating the artifacts present in the AFM data. Moreover, it successfully integrates accurate surface geometric information from various viewpoints, significantly diminishing the surface topography error. To quantify these improvements, we calculated the average of the absolute pixel error values across multiple viewpoints for each model. As summarized in Table 1, this analysis reveals that MVN-AFM achieved an exceptional average reduction of 94% in topography error for each structure, affirming the high accuracy of the 3D models reconstructed by MVN-AFM.

Discussion

In this work, we introduce MVN-AFM, a framework for 3D surface reconstruction of intricate micro- and nanostructures using multi-view AFM scanning data. We propose an iterative optimization method to simultaneously align the multi-view data and remove artifacts in the AFM image, achieving higher alignment accuracy. Moreover, we utilize the neural implicit surface reconstruction technique to the field of nanotechnology, which enables fusing spatially overlapping multi-view AFM data into an

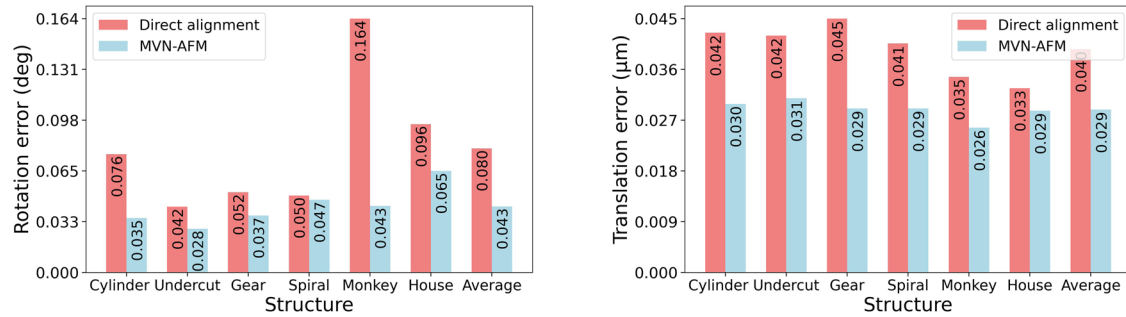


Fig. 5 | Evaluation of MVN-AFM’s improvement in multi-view data alignment accuracy. Comparison of the absolute pose error, that is, rotation and translation error (lower is better), of direct alignment of raw AFM data and the alignment method of MVN-AFM.

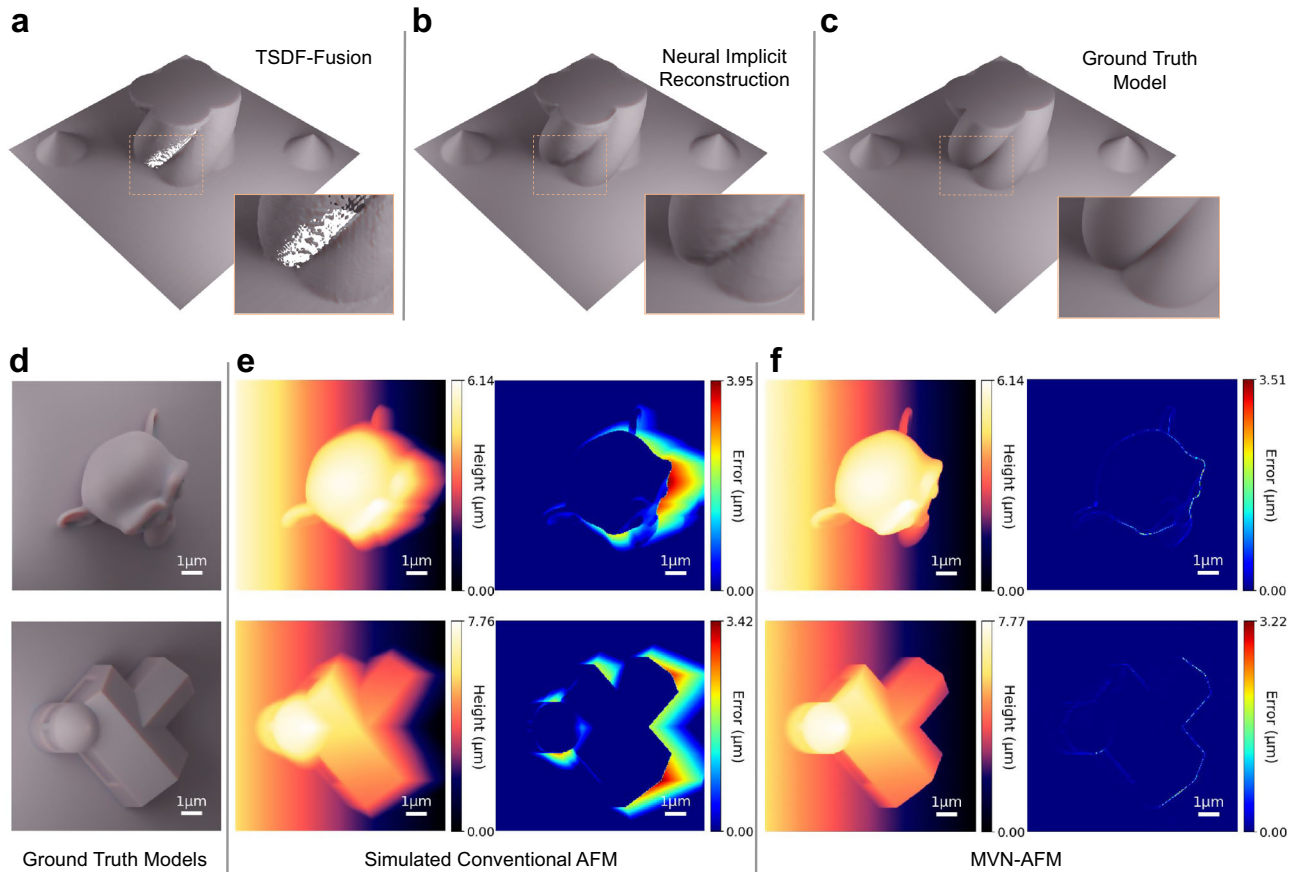


Fig. 6 | Evaluation of MVN-AFM’s improvement in reconstructed 3D models. **a** Given a set of masked and posed multi-view AFM images, the spiral model reconstructed by the Truncated Signed Distance Function (TSDF) Fusion method. The local zoom shows the voids on the reconstructed model. **b** Given the same set of masked and posed multi-view AFM images, the spiral model reconstructed by the

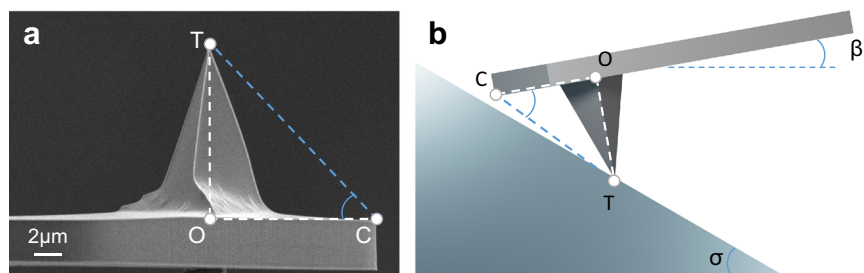
neural implicit surface reconstruction. **c** The ground truth model of the spiral. **d** The images of the monkey and house models in the given viewpoints. **e** The simulated AFM images and corresponding height error maps. **f** The topography images of 3D models reconstructed by MVN-AFM and corresponding height error maps. More visualizations can be found in Supplementary Fig. 8.

Table 1 | The error comparison of conventional AFM images and MVN-AFM

MAE (μm)	Cylinder	Undercut	Gear	Spiral	Monkey	House	Average
Conventional AFM	0.2587	0.2642	0.1750	0.2456	0.2043	0.2414	0.2315
MVN-AFM	0.0131	0.0137	0.0204	0.0128	0.0122	0.0173	0.0149

The mean absolute error (lower is better) of the input conventional AFM images and the topography images from 3D models reconstructed by MVN-AFM.

Fig. 7 | The limitation of tilt angle. **a** The SEM image of the TESPA-V2 AFM probe in our experiment. **b** The image of AFM tilt scanning.



accurate 3D model. MVN-AFM shows considerable practical value. Extensive experiments demonstrate the superior capability of MVN-AFM on diverse micro- and nanostructures, including microstructures printed by TPL, PMMA nanospheres, and ZIF-67 nanocrystals. The 3D models reconstructed by MVN-AFM provide researchers with a more comprehensive representation of micro- and nanostructures than what is achievable with conventional AFM scanning and 2D SEM images. The success of MVN-AFM across these varied samples, each with distinct geometries, types, and sizes, robustly affirms its effectiveness and broad applicability in nanofabrication, nanoparticles, and many other fields. Importantly, MVN-AFM only requires a conventional AFM system and a standard AFM probe to achieve these results. This aspect makes MVN-AFM a more accessible and cost-effective option for researchers to analyze intricate 3D micro- and nanostructures.

Our framework is efficient and flexible. While multi-view AFM data provides more surface information, it also increases the time for AFM scanning. We tested the effect of reconstruction using different numbers of tilted AFM data (Supplementary Fig. 10 and Supplementary Table 1). We found that for the structures in the TPL experiment, the reconstruction quality converged with only eight tilt scans. The scanning time for one AFM image is about 4.5 minutes, and considering the time required to switch to different scanning directions, multi-view scanning takes about 2 hours for a single structure. Next, given a set of multi-view AFM data, our framework takes about 10 minutes to complete the 3D reconstruction. Notably, the basis of our algorithm is the multi-view consistency of the accurate surface topography and the multi-view inconsistency of image artifacts in multi-view AFM data. Because our algorithm does not take parameters such as scanning number, tilt angle, and probe shape as prior information, users have the flexibility to adjust these parameters according to their requirements. Here are some guidelines for the setting of these parameters. Firstly, users can choose the scanning number by starting with a small number of scans and then gradually increasing it until the reconstruction results converge. Eight tilt scans performed well in our experiments. Secondly, users should maximize the tilt angle while preventing the probe cantilever from hitting the sample, as discussed in detail in Methods. Lastly, our method does not require specially shaped probes. However, if available, using probes with high aspect ratios benefits obtaining more geometric information during multi-view scanning.

Here, our study underscores the immense potential of integrating nanotechnology with neural implicit representations^{67–69}, an emerging and rapidly evolving field in computer vision. Specifically, we employ neural implicit surface reconstruction methods, where a neural network effectively represents a continuous SDF in space. Because of the continuous nature of neural networks, it is more suitable for representing geometric surface models that are inherently continuous in space than traditional discrete methods, as demonstrated in numerous recent works^{34–36}. Our research further reveals the successful application of this technique in the reconstruction of 3D micro- and nanostructures with multi-view AFM data.

Our methodology's foundational assumption is that the sample remains static during the multi-view AFM scanning process because the multi-view AFM data alignment step depends on the consistency of geometric features across different views. Therefore, our method is unsuitable for dynamic samples, such as living cells or samples prone to deformation

during scanning. Precisely reconstructing the deformation process of nanostructures is widely demanded in many research, which points to a promising direction for future work. One possible solution is applying our method to high-speed atomic force microscopy (HS-AFM)⁷⁰, which allows observing the dynamic action of nano-objects.

Methods

Hardware and software requirements of MVN-AFM

In our experiments, all the code of MVN-AFM was run on a computer with an Intel i9-13900KF CPU, an Nvidia RTX4090 GPU, 64 gigabytes of RAM, and a Linux operation system with a 5.15.0 kernel version, which is a typical configuration of the current lab workstation computer. In order to run the code of MVN-AFM properly, it requires at least one graphics card with memory larger than 12 gigabytes. We use Open3D⁷¹ 0.17.0, an open-source Python library, to handle the 3D data. Our implementation of neural implicit surface reconstruction is based on an open-source repository⁷² of hash encoding⁵⁵ and NeuS⁴², and the network is built on the deep learning framework PyTorch⁷³ 1.13.1.

Multi-view AFM scanning

All of the multi-view AFM images in our experiments were acquired through a commercial AFM (Dimension ICON, Bruker) operated in ‘Tapping Mode’ in the air, with 256 lines of 256 points, at a scan rate of 1 Hz, using a TESPA-V2 probe (Bruker) with a resonance frequency of 320 kHz and nominal spring constant of 37 Nm^{-1} . The setpoint amplitude was set to ~70% of the free air amplitude. The AFM probe has a height of $15 \mu\text{m}$, an overall shape of a quadrilateral pyramid, a front angle of 25° , a back angle of 17.5° , and a side angle of 20° . The stage has a $24 \text{ mm} \times 24 \text{ mm}$ square bottom, and the height is 16 mm with a 30° tilt angle. The turntable can hold a $4 \text{ mm} \times 4 \text{ mm}$ sample. The whole stage can be placed directly into a commercial AFM and does not collide with any part of the AFM during scanning. In our experiments, we rotate 45° each time between two adjacent scans and obtain eight tilt scans around the sample. Together with a conventional scan without sample tilt, nine scans are acquired per sample (Supplementary Movie 3).

We set the tilt angle to 30° to ensure that the AFM cantilever does not touch the sample surface during the scanning process while making the tilt angle as large as possible. For many types of AFM probes, the AFM tip is not located at the very end of the cantilever. As shown in Fig. 7a, the length of line segment OC of the TESPA-V2 probe used in our experiments is not negligible relative to the height of the tip, line segment OT. If the tilt angle is too large, point C may contact the sample before point T during the AFM probe engagement, causing the scan to fail. As depicted in Fig. 7b, to ensure that the tip contacts the sample, the tilt angle σ must satisfy $\sigma < \angle OCT - \beta$, where β is the inherent angle of the AFM holder, which is 11° . For our probe TESPA-V2, the angle $\angle OCT$ is approximately 45° . As the tilt angle increases, the AFM probe can access more structures inaccessible to conventional scanning without sample tilt, such as undercut sidewalls, providing more comprehensive surface geometry information for better 3D reconstruction. Therefore, we chose a 30° tilt angle in all our experiments to achieve the largest possible tilt angle while satisfying the constraints. We also included a slight margin in the tilt angle to ensure the cantilever would not touch the sample.

Precisely localizing the identical region across multi-view AFM scans is a critical step in the data capture process. The methodologies for achieving this localization are diverse and can be tailored to the unique characteristics of the experiment sample. In our TPL experiments, we utilized polymer grid markers printed around the sample to assist in localization by the optical microscope in the AFM system. For experiments of nanoparticles, we constructed scored markers on mica bases. These are just examples of the various strategies that can be adopted for localization, with other available methods including the use of a Transmission Electron Microscopy (TEM) index grid^{74,75} or the creation of noticeable artificial markers⁷⁶. The common destination of these techniques is to ensure that the specific structure for 3D reconstruction can be precisely and efficiently located within the AFM system.

Data alignment and mask solving

First, we claim some basic concepts in this step. Each AFM image is equivalent to a set of 3D points under an AFM coordinate system (Supplementary Fig. 3), with the z-axis being the position feedback direction of AFM and the x-y plane being the probe scanning plane. We define the AFM coordinate system of the scan result without sample tilt as the destined sample coordinate system. Moreover, we define a corresponding virtual orthogonal depth camera for each AFM image (Supplementary Fig. 4). With a given set of raw AFM data, we convert the AFM height information h into a depth value d for a virtual orthogonal camera parallel to the x-y plane, $d = \alpha - h$, where α is the assumed height of the camera. The value of α is simply ensured all d to be positive. Each AFM data point is treated as a ray $r = o + d\vec{v}$, originating from the pixel position o on the imaging plane and extending along the direction \vec{v} of the camera to the depth d .

In the initialization and the M-step, we applied the point-to-plane ICP algorithm⁵⁰ to align the data points filtered by mask M of AFM images and get a set of transformations T . In the E-step, We compute the artifact mask M of each AFM image by a cross-validation method. In detail, we first transform each camera ray r to the sample coordinate system to obtain the ray r' by the currently solved $T = \{(R, t) | R \in \mathbb{R}^{3 \times 3}, t \in \mathbb{R}^3\}$, where $r' = o' + d\vec{v}'$, $o' = Ro + t$, and $\vec{v}' = R\vec{v}$. Subsequently, we generate n sets of meshes by connecting spatial points corresponding to neighboring pixels in n AFM images within the sample coordinate system. Next, we compute the intersection of each ray with these meshes and obtain n depth values, $D = \{d_1, d_2, \dots, d_n\}$. Due to a basic fact, the artifacts of tip-sample convolution cause an expansion of the overall topography¹¹, resulting in the height value of AFM scanning being larger than the actual sample height, equivalent to the smaller depth value. Therefore, we consider a pixel as an artifact when $D_{\max} - d > \phi$, where d is the measured depth of each pixel, and ϕ is set to 3% of the AFM scan size initially and linearly reduced to 1% with iteration, a value determined experimentally and applied consistently across all our experiments. The tiny threshold is set to make the algorithm robust to noises in the AFM data and inaccurate T during the iteration process. These iterative EM-steps reinforce each other. After a fixed number of iterations, five in our experiments, the resolved poses T , and masks M are saved for subsequent steps. For a set of multi-view AFM data, this process takes about 2 minutes.

Neural implicit surface reconstruction

In the neural implicit surface reconstruction step, we train a multi-resolution hash table with learnable parameters and an MLP neural network named the SDF network by the aligned and masked multi-view AFM data (Supplementary Fig. 5). In the training process, we sample 3D points along the ray r' of pixels filtered by mask M . First, the 3D point coordinate is encoded by multiresolution hash technology⁵⁵. Here, we use 16 resolution levels, each obtaining a 2-dimensional feature vector. Concatenating the hash encoding and the 3D coordinate, we get a 35-dimensional feature as input of the SDF network. The SDF network is a one-layer MLP network with 64 hidden sizes and ReLU activation, which maps the input feature to an SDF value at that 3D point. The SDF value of each point is converted to a density value through the unbiased and occlusion-aware weight function proposed by NeuS⁴². Then, the density values of sample points along the ray

are accumulated by the differentiable volume rendering method to obtain the depth value \hat{d} of that ray. The loss function L consists of a depth error term L_{depth} and a regularization term L_{reg} :

$$L_{\text{depth}} = \frac{1}{b} \sum_p^b (\hat{d}_p - d_p)^2, \tag{1}$$

$$L_{\text{reg}} = \frac{1}{bm} \sum_{p,q}^{b,m} (\|\mathbf{n}_{pq} - 1\|)^2. \tag{2}$$

L_{depth} is the mean square error (MSE) between the rendering depth value of each pixel and the AFM data supervision, where b is the batch size and p is the index. The regularization term⁵³ is used to constrain the SDF field represented by the network, where \mathbf{n} is the normal of the sample point, m is the number of sample points along a ray, and q is the index. L_{reg} facilitates a smooth and natural surface, commonly used in SDF-based neural implicit surface reconstruction methods.

$$L = L_{\text{depth}} + \lambda L_{\text{reg}}. \tag{3}$$

The weight λ is 0.1 in our experiment. During the network training, the Adam optimizer updates the network parameters with a learning rate of 0.001 to minimize the loss and perform 20,000 iterations. In one iteration, we randomly select 256 rays with 1024 sample points along the ray. Notably, Each set of network parameters can only represent a 3D model of one structure, so multi-view AFM images for different samples need to be trained from scratch. The whole training time is about 8 minutes on an Nvidia RTX4090 GPU. To visualize the 3D model, we divide the space into $256 \times 256 \times 256$ voxels. Subsequently, the SDF values for each voxel are obtained through neural network inference, followed by the extraction of meshes using the Marching Cube algorithm⁵⁶. Unlike the traditional discrete voxel-based representation³¹, which requires the prior determination of a voxel division resolution, neural implicit surface representations do not have a resolution limitation. The network can infer SDF values at any location, enabling the generation of a mesh representation with arbitrary resolution.

Constructing the two-photon lithography structures

We used a commercial photoresist IP-Dip2 (Nanoscribe GmbH) as our material. The IP-Dip2 was dropped on a glass substrate with a thickness of 170–190 μm (Borosilicate substrates, Nanoscribe GmbH) for fabricating the structures. We used a commercial Direct Laser Writing setup (Photonic Professional GT2, Nanoscribe GmbH) equipped with a 780 nm femtosecond laser (a repetition rate of 80 MHz, a pulse duration of 80–100 fs) and a 63 \times , numerical aperture (NA) = 1.4 oil immersion objective to print the microstructures. We imported the STL files into Describe 2.7 (Nanoscribe GmbH) to generate the executable job files. We set the slice and hatching distances to 0.1 μm for microstructures, the highest accuracy this machine can achieve. These distances were set to 0.3 μm for grid markers because they are only used for optical microscope localization, which has no high requirement for printing accuracy. The printing parameters were set to 30 mW of laser power and 10,000 $\mu\text{m/s}$ scanning speed. Then, we imported the executable job files to Nanowrite 1.8 (Nanoscribe GmbH) to start the job. After the printing process, the printed structures were developed with propylene glycol methyl ether acetate (PGMEA) for 20 minutes and isopropyl alcohol (IPA) for 5 minutes to wash out the unpolymerized resists at room temperature and leave the microstructures on the substrate.

Constructing the nanoparticle samples

PMMA nanosphere dispersion (500 nm) was purchased from the Jiangsu Zhichuan Technology Co., Ltd (China). ZIF-67 powder (300 nm) was purchased from the Nanjing Xianfeng Nano Co., Ltd (China). We used ethanol to dilute these nanoparticles, sonicated them for 10 minutes, and then deposited the suspension onto a 4 mm \times 4 mm mica base. We

performed multi-view localization of the particles of interest based on the markers of the mica surface around the region. We used the same view number, tilt angle, AFM scanning mode, and AFM probe as in the TPL experiment. The AFM scan size was $2\ \mu\text{m} \times 2\ \mu\text{m}$ for PMMA nanospheres and $1.5\ \mu\text{m} \times 1.5\ \mu\text{m}$ for ZIF-67 nanocrystals. We obtained SEM images of these micro- and nanostructures by sputter-coating samples with platinum by sputtering apparatus (MCIOO, Hitachi) and then observing them with a field-emission scanning electron microscope (GeminiSEM 300, ZEISS).

Constructing the simulated data

The simulated data was generated using the 3D design software Blender⁶¹ 3.3. Within Blender, we constructed models of the structures as well as the AFM probe. To mimic the conditions of our real-world multi-view AFM scanning, we set up orthogonal cameras within the software positioned to align with the scanning directions in our real experiment. Furthermore, to replicate the real-world experimental setup more accurately, we rotated the probe model by 11° . This adjustment accounts for the inherent angle between the working cantilever of the AFM holder and the scanning plane in a real AFM system^{11,58}. Next, we orthogonally projected the surface model onto these cameras and performed a convolution of the probe shape to generate simulated AFM images. In the quantitative evaluation of the accuracy of solved poses, we extracted the precise poses of these cameras directly from Blender and evaluated the absolute pose error (APE) by an open-source tool EVO⁷⁷. The camera pose T at each viewpoint consists of a rotation matrix R and a translation vector t :

$$T = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4}, R \in \mathbb{R}^{3 \times 3}, t \in \mathbb{R}^3. \quad (4)$$

The relative pose E_i between the estimated camera pose $T_{\text{est},i}$ and the ground truth pose $T_{\text{gt},i}$ for viewpoint i is calculated:

$$E_i = T_{\text{est},i}^{-1} T_{\text{gt},i} \in \mathbb{R}^{4 \times 4}. \quad (5)$$

For n viewpoints, the translation and rotation components of the relative pose E are extracted and averaged respectively to obtain the translation and rotation errors:

$$\text{APE}_{\text{translation}} = \frac{1}{n} \sum_{i=1}^n \|\text{trans}(E_i)\|, \quad (6)$$

$$\text{APE}_{\text{rotation}} = \frac{1}{n} \sum_{i=1}^n |\text{angle}(\log(\text{rot}(E_i)))|, \quad (7)$$

where $\text{trans}(\cdot)$ extracts the translation vector from the pose matrix, and the final translation error is in units of micrometers, $\text{rot}(\cdot)$ extracts the rotation matrix from the pose matrix, $\log(\cdot)$ is the inverse operation of Rodrigues' rotation formula⁷⁸, transforming the rotation matrix into a rotation vector, and $\text{angle}(\cdot)$ then calculates the rotation angle from the rotation vector, in units of degrees. We implemented the TSDF Fusion method based on an open source repository⁷⁹ and added support for the orthogonal camera projection model that allows for the fusion of multi-view AFM data. We divided the space into $256 \times 256 \times 256$ voxels to keep consistent with the setup of the mesh model extraction step in the neural implicit surface reconstruction. When it comes to evaluating the topography images of the 3D models reconstructed by MVN-AFM, we employ the Mean Absolute Error (MAE) as our metric:

$$\text{MAE} = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n |\bar{h}_{ij} - h_{ij}|, \quad (8)$$

where n denotes the number of multi-view images, m is the pixel number in each image, \bar{h} is the accurate height value of a pixel, and h denotes the value of pixels in raw AFM images or topography images from MVN-AFM.

Data availability

A test data case is available in our GitHub repository⁸⁰. Data underlying the results presented in this paper are not publicly available but may be obtained from the authors upon reasonable request.

Code availability

The source code of MVN-AFM is available at <https://github.com/zju3dv/MVN-AFM>⁸⁰.

Received: 29 April 2024; Accepted: 19 August 2024;

Published online: 12 September 2024

References

- Gates, B. et al. New approaches to nanofabrication: molding, printing, and other techniques. *Chem. Rev.* **105**, 1171–1196 (2005).
- Quake, S. & Scherer, A. From micro- to nanofabrication with soft materials. *Science* **290**, 1536–1540 (2000).
- Douglas, S. M., Bachelet, I. & Church, G. M. A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **335**, 831–834 (2012).
- Li, S. et al. A DNA nanorobot functions as a cancer therapeutic in response to a molecular trigger *in vivo*. *Nat. Biotechnol.* **36**, 258+ (2018).
- Gratton, S. E. A. et al. The effect of particle design on cellular internalization pathways. *Proc. Natl Acad. Sci. USA* **105**, 11613–11618 (2008).
- Wang, J., Byrne, J. D., Napier, M. E. & DeSimone, J. M. More effective nanomedicines through particle design. *Small* **7**, 1919–1931 (2011).
- Seiler, H. Secondary-electron emission in the scanning electron-microscope. *J. Appl. Phys.* **54**, R1–R18 (1983).
- Egerton, R., Li, P. & Malac, M. Radiation damage in the TEM and SEM. *Micron* **35**, 399–409 (2004).
- Binnig, G., Quate, C. F. & Gerber, C. Atomic force microscope. *Phys. Rev. Lett.* **56**, 930–933 (1986).
- Tian, F., Qian, X. & Villarrubia, J. S. Blind estimation of general tip shape in afm imaging. *Ultramicroscopy* **109**, 44–53 (2008).
- Golek, F., Mazur, P., Ryszka, Z. & Zuber, S. AFM image artifacts. *Appl. Surf. Sci.* **304**, 11–19 (2014).
- Velegol, S., Pardi, S., Li, X., Velegol, D. & Logan, B. AFM imaging artifacts due to bacterial cell height and AFM tip geometry. *Langmuir* **19**, 851–857 (2003).
- Westra, K., Mitchell, A. & Thomson, D. Tip artifacts in atomic-force microscope imaging of thin-film surfaces. *J. Appl. Phys.* **74**, 3608–3610 (1993).
- Martin, Y. & Wickramasinghe, H. K. Method for imaging sidewalls by atomic-force microscopy. *Appl. Phys. Lett.* **64**, 2498–2500 (1994).
- Orji, N. G. & Dixon, R. G. Higher order tip effects in traceable CD-AFM-based linewidth measurements. *Meas. Sci. Technol.* **18**, 448–455 (2007).
- Thiesler, J., Tutsch, R., Fromm, K. & Dai, G. True 3D-AFM sensor for nanometrology. *Meas. Sci. Technol.* **31**, 074012 (2020).
- Geng, J., Zhang, H., Meng, X., Rong, W. & Xie, H. Sidewall imaging of microarray-based biosensor using an orthogonal cantilever probe. *IEEE Trans. Instrum. Meas.* **70**, 1–8 (2021).
- Nguyen, C. et al. Carbon nanotube scanning probe for profiling of deep-ultraviolet and 193 nm photoresist patterns. *Appl. Phys. Lett.* **81**, 901–903 (2002).
- Cho, S.-J. et al. Three-dimensional imaging of undercut and sidewall structures by Atomic Force Microscopy. *Rev. Sci. Instrum.* **82**, 23707 (2011).
- Kizu, R., Misumi, I., Hirai, A., Kinoshita, K. & Gonda, S. Development of a metrological atomic force microscope with a tip-tilting mechanism for 3D nanometrology. *Meas. Sci. Technol.* **29**, 075005 (2018).
- Xie, H., Hussain, D., Yang, F. & Sun, L. Development of three-dimensional atomic force microscope for sidewall structures imaging

- with controllable scanning density. *IEEE/ASME Trans. Mechatron.* **21**, 316–328 (2016).
22. Wu, J.-W. et al. Effective tilting angles for a dual probes AFM system to achieve high-precision scanning. *IEEE/ASME Trans. Mechatron.* **21**, 2512–2521 (2016).
23. Xie, H., Hussain, D., Yang, F. & Sun, L. Atomic force microscope caliper for critical dimension measurements of micro and nanostructures through sidewall scanning. *Ultramicroscopy* **158**, 8–16 (2015).
24. Zhao, X., Fu, J., Chu, W., Nguyen, C. & Vorburger, T. V. An image stitching method to eliminate the distortion of the sidewall in linewidth measurement. In *Metrology, Inspection, and Process Control for Microlithography XVIII*, vol. 5375, 363–373 (2004).
25. Pan, S.-P., Liou, H.-C., Chen, C.-C. A., Chen, J.-R. & Liu, T.-S. Precision measurement of sub-50 nm linewidth by stitching double-tilt images. *Jpn J. Appl. Phys.* **49**, 06GK06 (2010).
26. Kawata, S., Sun, H., Tanaka, T. & Takada, K. Finer features for functional microdevices - micromachines can be created with higher resolution using two-photon absorption. *Nature* **412**, 697–698 (2001).
27. Jaiswal, A. et al. Two decades of two-photon lithography: materials science perspective for additive manufacturing of 2D/3D nano-microstructures. *Iscience* **26**, 106374 (2023).
28. Li, J. & Pumera, M. 3D printing of functional microrobots. *Chem. Soc. Rev.* **50**, 2794–2838 (2021).
29. Dabbagh, S. R. et al. 3D-printed microrobots from design to translation. *Nat. Commun.* **13**, 5875 (2022).
30. Jun, Y.-w, Choi, J.-s & Cheon, J. Shape control of semiconductor and metal oxide nanocrystals through nonhydrolytic colloidal routes. *Angew. Chem. Int. Ed.* **45**, 3414–3439 (2006).
31. Izadi, S. et al. Kinectfusion: Real-time 3D reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, 559–568 (2011).
32. Curless, B. & Levoy, M. A volumetric method for building complex models from range images. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, 303–312 (1996).
33. Nießner, M., Zollhöfer, M., Izadi, S. & Stamminger, M. Real-time 3D reconstruction at scale using voxel hashing. *ACM Trans. Graph. (ToG)* **32**, 1–11 (2013).
34. Xie, Y. et al. Neural fields in visual computing and beyond. *Computer Graph. Forum* **41**, 641–676 (2022).
35. Weder, S., Schonberger, J. L., Pollefeys, M. & Oswald, M. R. NeuralFusion: Online depth fusion in latent space. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3162–3172 (2021).
36. Li, K., Tang, Y., Prasad, V. A. & Torr, P. H. BNV-fusion: dense 3D reconstruction using bi-level neural volume fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6166–6175 (2022).
37. Mildenhall, B. et al. NeRF: representing scenes as neural radiance fields for view synthesis. *Commun. ACM* **65**, 99–106 (2021).
38. Oechsle, M., Peng, S. & Geiger, A. UNISURF: unifying neural implicit surfaces and radiance fields for multi-view reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 5589–5599 (2021).
39. Yariv, L. et al. Multiview neural surface reconstruction by disentangling geometry and appearance. *Adv. Neural Inf. Process. Syst.* **33**, 2492–2502 (2020).
40. Yariv, L., Gu, J., Kasten, Y. & Lipman, Y. Volume rendering of neural implicit surfaces. *Adv. Neural Inf. Process. Syst.* **34**, 4805–4815 (2021).
41. Sucar, E., Liu, S., Ortiz, J. & Davison, A. J. iMAP: Implicit mapping and positioning in real-time. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 6229–6238 (2021).
42. Wang, P. et al. NeuS: Learning neural implicit surfaces by volume rendering for multi-view reconstruction. *Adv. Neural Inf. Process. Syst.* **34**, 27171–27183 (2021).
43. Bettencourt, A. & Almeida, A. J. Poly(methyl methacrylate) particulate carriers in drug delivery. *J. Microencapsul.* **29**, 353–367 (2012).
44. Tang, E., Cheng, G., Pang, X., Ma, X. & Xing, F. Synthesis of nano-ZnO/poly(methyl methacrylate) composite microsphere through emulsion polymerization and its UV-shielding property. *Colloid Polym. Sci.* **284**, 422–428 (2006).
45. Zhu, A., Shi, Z., Cai, A., Zhao, F. & Liao, T. Synthesis of core-shell PMMA-SiO₂ nanoparticles with suspension-dispersion-polymerization in an aqueous system and its effect on mechanical properties of PVC composites. *Polym. Test.* **27**, 540–547 (2008).
46. Zhong, G., Liu, D. & Zhang, J. The application of ZIF-67 and its derivatives: adsorption, separation, electrochemistry and catalysis. *J. Mater. Chem. A* **6**, 1887–1899 (2018).
47. Qian, J., Sun, F. & Qin, L. Hydrothermal synthesis of zeolitic imidazolate framework-67 (ZIF-67) nanocrystals. *Mater. Lett.* **82**, 220–223 (2012).
48. Wang, L. et al. Flexible solid-state supercapacitor based on a metal-organic framework interwoven by electrochemically-deposited PANI. *J. Am. Chem. Soc.* **137**, 4920–4923 (2015).
49. Yang, J. et al. Hollow Zn/Co ZIF particles derived from core-shell ZIF-67@ZIF-8 as selective catalyst for the semi-hydrogenation of acetylene. *Angew. Chem. -Int. Ed.* **54**, 10889–10893 (2015).
50. Rusinkiewicz, S. & Levoy, M. Efficient variants of the ICP algorithm. In *Proceedings third international conference on 3-D digital imaging and modeling*, 145–152 (2001).
51. Do, C. B. & Batzoglu, S. What is the expectation maximization algorithm? *Nat. Biotechnol.* **26**, 897–899 (2008).
52. Moon, T. The expectation-maximization algorithm. *IEEE Signal Process. Mag.* **13**, 47–60 (1996).
53. Gropp, A., Yariv, L., Haim, N., Atzmon, M. & Lipman, Y. Implicit geometric regularization for learning shapes. *Proceedings of the 37th International Conference on Machine Learning* 3789–3799 (2020).
54. Hecht-Nielsen, R. Theory of the backpropagation neural network. In *Neural networks for perception*, 65–93 (1992).
55. Müller, T., Evans, A., Schied, C. & Keller, A. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph. (ToG)* **41**, 1–15 (2022).
56. Lorensen, W. E. & Cline, H. E. Marching cubes: a high resolution 3D surface construction algorithm. In *Seminal Graphics: Pioneering Efforts That Shaped The Field*, 347–353 (1998).
57. Selimis, A., Mironov, V. & Farsari, M. Direct laser writing: principles and materials for scaffold 3D printing. *Microelectron. Eng.* **132**, 83–89 (2015).
58. Shen, J., Zhang, D., Zhang, F.-H. & Gan, Y. AFM tip-sample convolution effects for cylinder protrusions. *Appl. Surf. Sci.* **422**, 482–491 (2017).
59. Lee, J. H. et al. Electrically pumped sub-wavelength metallo-dielectric pedestal pillar lasers. *Opt. Express* **19**, 21524–21531 (2011).
60. Chaubey, S. K. & Jain, N. K. State-of-art review of past research on manufacturing of meso and micro cylindrical gears. *Precis. Eng.* **51**, 702–728 (2018).
61. Community, B. O. *Blender - a 3D Modelling And Rendering Package* (Blender Foundation, Stichting Blender Foundation, 2018).
62. Reis, C. P., Neufeld, R. J., Ribeiro, A. J. & Veiga, F. Nanoencapsulation I. Methods for preparation of drug-loaded polymeric nanoparticles. *Nanomed. Nanotechnol. Biol. Med.* **2**, 8–21 (2006).
63. Saliba, D., Ammar, M., Rammal, M., Al-Ghoul, M. & Hmadeh, M. Crystal growth of ZIF-8, ZIF-67, and their mixed-metal derivatives. *J. Am. Chem. Soc.* **140**, 1812–1823 (2018).
64. Nordin, N. A. H. M., Ismail, A. F., Mustafa, A., Murali, R. S. & Matsuura, T. The impact of ZIF-8 particle size and heat treatment on CO₂/CH₄ separation using asymmetric mixed matrix membrane. *RSC Adv.* **4**, 52530–52541 (2014).

65. Xia, Y., Xiong, Y., Lim, B. & Skrabalak, S. E. Shape-controlled synthesis of metal nanocrystals: simple chemistry meets complex physics? *Angew. Chem. Int. Ed.* **48**, 60–103 (2009).
66. Amyot, R. & Flechsig, H. BioAFMviewer: an interactive interface for simulated AFM scanning of biomolecular structures and dynamics. *PLoS Comput. Biol.* **16**, e1008444 (2020).
67. Sitzmann, V., Martel, J., Bergman, A., Lindell, D. & Wetzstein, G. Implicit neural representations with periodic activation functions. *Adv. Neural Inf. Process. Syst.* **33**, 7462–7473 (2020).
68. Chen, Y., Liu, S. & Wang, X. Learning continuous image representation with local implicit image function. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 8628–8638 (2021).
69. Pumarola, A., Corona, E., Pons-Moll, G. & Moreno-Noguer, F. D-NeRF: neural radiance fields for dynamic scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10318–10327 (2021).
70. Uchihashi, T., Kodera, N. & Ando, T. Guide to video recording of structure dynamics and dynamic processes of proteins by high-speed atomic force microscopy. *Nat. Protoc.* **7**, 1193–1206 (2012).
71. Zhou, Q.-Y., Park, J. & Koltun, V. Open3D: a modern library for 3D data processing. *arXiv* <https://doi.org/10.48550/arXiv.1801.09847> (2018).
72. Guo, Y.-C. Instant neural surface reconstruction. *GitHub* <https://github.com/bennyguo/instant-nsr-pl> (2022).
73. Paszke, A. et al. PyTorch: an imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* **32**, 8024–8035 (2019).
74. Markiewicz, P. & Goh, M. Identifying locations on a substrate for the repeated positioning of AFM samples. *Ultramicroscopy* **68**, 215–221 (1997).
75. Abu Quba, A. A., Schaumann, G. E., Karagulyan, M. & Diehl, D. A new approach for repeated tip-sample relocation for AFM imaging of nano and micro sized particles and cells in liquid environment. *Ultramicroscopy* **211**, 112945 (2020).
76. Liu, Z. et al. Mechanically engraved mica surface using the atomic force microscope tip facilitates return to a specific sample location. *Microsc. Res. Tech.* **66**, 156–162 (2005).
77. Grupp, M. evo: Python package for the evaluation of odometry and SLAM. *GitHub* <https://github.com/MichaelGrupp/evo> (2017).
78. Dai, J. S. Euler-rodriques formula variations, quaternion conjugation and intrinsic connections. *Mech. Mach. Theory* **92**, 144–152 (2015).
79. Zeng, A. et al. Volumetric TSDF Fusion of RGB-D images in python. *GitHub* <https://github.com/andyzeng/tsdf-fusion-python> (2017).
80. Chen, S. et al. Multi-view neural 3D reconstruction of micro- and nanostructures with atomic force microscopy. *GitHub* <https://github.com/zju3dv/MVN-AFM> (2024).

Acknowledgements

We thank A. Ren, L. Ma, and C. Wu for assistance in the two-photon lithography experiment. We are grateful to Y. Wang, J. Tang, and M. Duan for helpful discussions. We also thank the staff of the Analysis Center of Agrobiolgy and Environmental Sciences, Zhejiang University, for their

support in SEM imaging. This work was partially supported by the National Natural Science Foundation of China (No.61932003 received by G.Z., No.51975522 and No.U22A20207 received by Y.-L.C.).

Author contributions

S.C. and M.P. conceived the idea and jointly proposed this project; S.C. and M.P. performed experiments; S.C. wrote code and processed data; S.C. and Y.L. wrote the draft of the manuscript, and all co-authors proofread and revised the manuscript; G.Z., Y.-L.C., H.B., and B.-F.J. provided valuable suggestions including the experiment design and writing. G.Z. and Y.-L.C. supervised this project, including the framework design and improvement.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44172-024-00270-9>.

Correspondence and requests for materials should be addressed to Yuan-Liu Chen or Guofeng Zhang.

Peer review information : *Communications Engineering* thanks Ankita Ray and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Anastasia Vasylenkova and Saleem Denholme. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

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

© The Author(s) 2024