Measurement 235 (2024) 114908

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Marker-assisted cross-scale measurement for robotic macro–micro manipulation utilizing computer microvision

Sheng Yao [a,*], Xianmin Zhang [b], Sergej Fatikow [c]

[a] Guangdong Province Engineering Laboratory for Medical Imaging and Diagnostic Technology, Guangdong Provincial Key Laboratory of Medical Image Processing, and School of Biomedical Engineering, Southern Medical University, Guangzhou, 510515, China

[b] Guangdong Key Laboratory of Precision Equipment and Manufacturing Technology, School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou, 510640, China

[c] Division of Microrobotics and Control Engineering, Department of Computing Science, University of Oldenburg, Oldenburg, 26129, Germany

ARTICLE INFO

Keywords:
Microvision
Cross-scale measurement
Robotic macro-micro manipulation
Marker-assisted
Global map construction

ABSTRACT

Cross-scale measurement techniques are pivotal for seamless coordination between robotic macro and micromanipulation. While computer microvision offers non-contact and multi-degree-of-freedom capabilities, its limitations in measurement range and narrow field-of-view (FOV) prove unsatisfactory for robotic macro-micro manipulation. This study presents a marker-assisted microvision-based method to minimize errors during measurement range extension and ensure precise cross-scale motion measurement for robotic macro-micro manipulation. The microvision-based system utilizes the microstructural pattern designed for global map construction. Additionally, it employs a learning-based approach to efficiently estimate the FOV position for macro-motion measurement, followed by the integration of micro-motion tracking within the FOV for precise micro-motion measurements. Experimental investigations showcase a measurement accuracy of 9 μm in an extended range of 3.11 mm for macro-motion estimation and 219 nm for micro-motion measurement, which highlight the microvision-based method in achieving effective cross-scale motion measurement, potentially paving the way for the automation of robotic macro-micro manipulation.

1. Introduction

Macro-micro mechanical architecture has shown great advantages in robotic manipulation systems as the intrinsic capability to enable the long-stroke, high-speed, and high-precision performances [1]. Switching between fine and coarse positioning with such robotic systems is a hot research topic, and it often involves integrating several measurement techniques. For instance, the use of capacitive sensor sensing techniques has been employed to measure the motion of a nanopositioning robot [2], while laser displacement sensing techniques have been applied to track the macro-positioning robot [3]. Nevertheless, capacitive sensing techniques or laser techniques inherently offer single degree-of-freedom (DOF) linear measurement for one sensor. Achieving multi-DOF measurement often requires cumbersome hardware, resulting in significant drawbacks such as narrowing the workspace of robotic manipulation and the necessity for customized mechanical structure design [4]. As a result, these measurement techniques would boost the costs and complicate the system setup due to the necessity of customized mechanical structure design.

On the other hand, by combining microscopy and computer vision techniques, microvision-based motion measurement has become one of the most promising approaches for automation in robotic manipulation, with desirable features such as direct visualization, easy integration, multi-DOF sensing ability, and tremendous available information [5]. In recent studies, a microvision-based measurement method using template matching was proposed by Li et al. [6] for robotic nanopositioning. A phase correlation algorithm was designed by Marturi et al. [7] for 5-DOF motion measurement of nano-positioning stages. An intelligent feature-matching approach for robotic micromanipulation was proposed by Qin et al. [8]. An optical-flow-based algorithm for full-field motion measurement of the robotic micro-positioning stage was developed by Yao et al. [9]. Zhao et al. [10] proposed a marker-assisted microvision-based method employing feature matching to achieve sub-microscale measurement. While these microvision-based methods have demonstrated nanometer resolution, their measuring range is typically confined to the microscale due to the narrow field of view (FOV) of the microvision systems, which significantly restricts the robotic workspace.

To address the aforementioned issue, a few solutions have been investigated for microvision systems. For instance, switching to a lower

* Corresponding author.
E-mail addresses: yaosheng@smu.edu.cn (S. Yao), zhangxm@scut.edu.cn (X. Zhang).

https://doi.org/10.1016/j.measurement.2024.114908
Received 24 December 2023; Received in revised form 16 April 2024; Accepted 13 May 2024
Available online 16 May 2024

Table 1
Comparison of the microvision-based methods for cross-scale measurement.

Methods	Characteristics
Laurent et al. [13]	LFSR sequence encoding, phase-based
Clévy et al. [15]	QR code encoding, phase-based
Proposed approach	Periodic and fiducial pattern, learning-based

magnification objective lens could enlarge the FOV of the microvision systems, but it would sacrifice the measurement precision. To extend the measuring range while maintaining high accuracy, Potsaid et al. [11] proposed a customized adaptive objective lens. However, this objective lens exhibited low stability and repeatability, imposing significant limitations on its applications. Generally, the microvision systems with this kind of adaptive magnification rely on mechanically moving the solid lenses, which inevitably increases the weight and complicates the system configuration. Moreover, the requirement to displace the optical elements makes these microvision systems vulnerable to mechanical vibration and misalignment. Besides, the response time increases with the displacement, making fast switching between different magnifications difficult [12].

Therefore, microvision systems with a fixed objective lens are much more cost-efficient, compact, and stable for motion measurement. Accordingly, a microvision-based cross-scale measurement method with a fixed objective lens was developed by Laurent et al. [13] to achieve a large range-to-resolution utilizing phase correlation. However, this method required encryption, including sophisticated bit encoding and decoding, i.e. pseudo-random binary sequences obtained by linear feedback shift register (LFSR), and it could be affected by occlusion from the manipulated objects on the encoded micro-patterns [14]. Recent efforts by Clévy et al. [15] aimed to simplify the marker-assisted measurement method by incorporating periodic and fiducial patterns for enhanced versatility. Nevertheless, this approach still requires the use of QR codes for accurate measurement. As a result, we are motivated to develop a more intuitive and practical method capable of simultaneously achieving an extensive measuring range and high accuracy for cross-scale measurements, with the ultimate goal of automating robotic tasks such as macro-micro manipulation.

The main contributions of this paper are as follows: (1) This study facilitates cross-scale motion measurement through the microvision-based method without additional encryption. (2) A high-resolution microstructural marker is meticulously designed and manufactured based on the periodic and fiducial pattern, which significantly minimizes errors in measurement range extension and ensures the precise construction of the global map. (3) A coordination strategy based on a fixed magnification of the microvision system is devised to both align the requirements of robotic macro-micro manipulation and simplify the system setup.

The characteristics comparison of the proposed cross-scale measurement method with similar microvision-based methods is listed in Table 1. The proposed cross-scale measurement method not only can ensure switching between fine and coarse positioning during robotic macro-micro manipulation but also can be applied to cell manipulation and biomedical characterization in the future [16–18]. The remainder of this article is organized as follows. Section 2 presents the methodology of the cross-scale microvision-based method. In Section 3, an experimental system is established, and the versatility of the proposed method is verified by other dedicated instruments in precision engineering. Finally, Section 4 concludes this article.

2. Methodology of the cross-scale microvision measurement system

In this paper, we introduce a novel cross-scale measurement system based on computer microvision within the framework of robotic macro-micro manipulation, as depicted in Fig. 1. The process of the proposed cross-scale measurement for robotic macro-micro manipulation is outlined in Fig. 2, encompassing two main components: global map construction and motion measurement. The global map construction is serving as an initialization for motion measurement in robotic macro-micro manipulation.

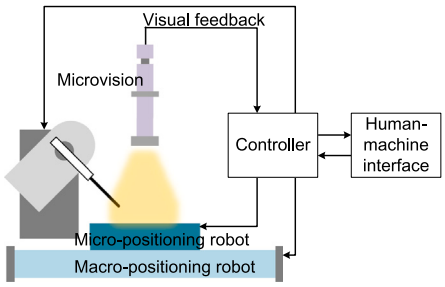

Fig. 1. Framework of the robotic macro-micro manipulation system using the computer microvision.

2.1. Design of the microstructural marker and global map construction

To facilitate the measurement of macroscale motion by the microvision system, it is imperative to establish a global map that extends the measuring range. This is accomplished by capturing microscopic images of distinct contiguous regions through robotic manipulation. The global map of the robot workspace can subsequently be constructed from regional microscopic images, employing an appropriate image stitching technique [19]. This approach effectively extends the sensing range of the computer microvision.

A high-resolution marker is devised based on the Archimedean spiral architecture to mitigate motion measurement errors arising from global map construction. Significantly, the Archimedean spiral exhibits the property that any ray originating from the origin intersects successive turns of the spiral at points with a constant separation distance. Therefore, the mapping relationship between the one-dimensional parameter ω and the two-dimensional coordinates (x, y) can be established using the parametric equation as follows:

$$\begin{cases} x = (a + b\omega) \cdot cos\omega \\ y = (a + b\omega) \cdot sin\omega \end{cases} \tag{1}$$

where the constant separation distance is represented by $2\pi b$, with a denoting the distance between the fixed point and the coordinate center and ω governing the size of the Archimedean spiral. Motivated by this appealing property, the designed marker integrates the Archimedean spiral with equidistant grids. This fusion ensures that each region in the global map can be perceived as a distinctive pattern for visual recognition, as illustrated in Fig. 3.

Fig. 2(a) shows the overall procedure for global map construction. After placing the high-resolution microstructural marker onto the macro-micro robot and the subsequent movement to acquire different FOV microscopic images, measuring range extension is accomplished through image stitching incorporating the regionally unique pattern design. Specifically, image stitching relies on registration in the overlapping regions to resolve the transformation of multiple microscopic images. Distinguished by its high efficiency and robustness, the DAISY descriptor stands out for its ability to be efficiently computed for dense feature extraction with high matching accuracy [20]. Employing a circular neighborhood and the Gaussian filter function for convolution, the gradients are accumulated into histograms in the direction of convolution. Assuming the convolution of the Gaussian filter at pixel (μ, ν) in the jth direction in the overlapping region of the microscopic images, it can be obtained:

$$G_{\sigma}^{\Sigma}(\mu, \nu) = G_{\Sigma} * (\frac{\delta I(\mu, \nu)}{\delta o})^{+} \tag{2}$$

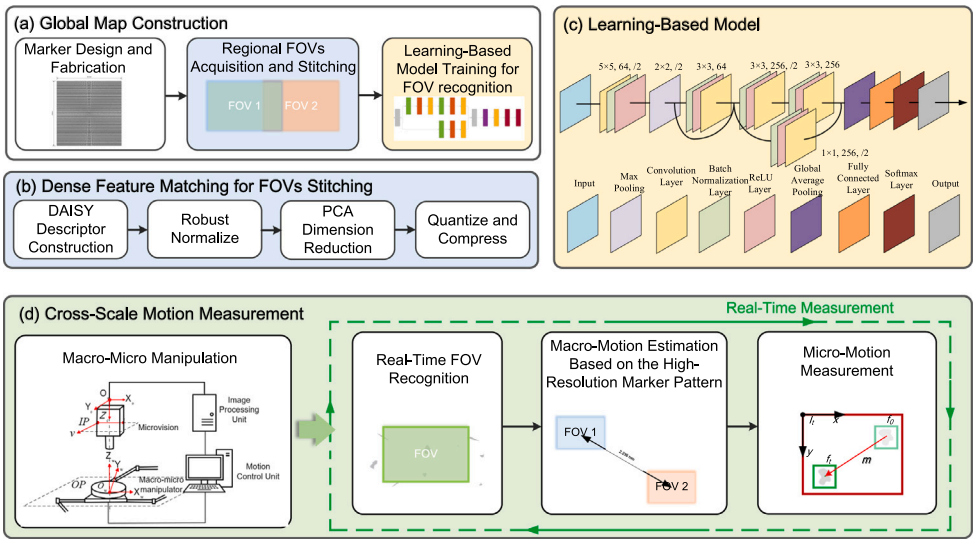

Fig. 2. Cross-scale motion measurement method based on computer microvision. (a) Global map construction. (b) Dense feature matching for FOVs stitching. (c) Learning-based model for FOV recognition. (d) Real-time cross-scale motion measurement for robotic macro-micro manipulation.

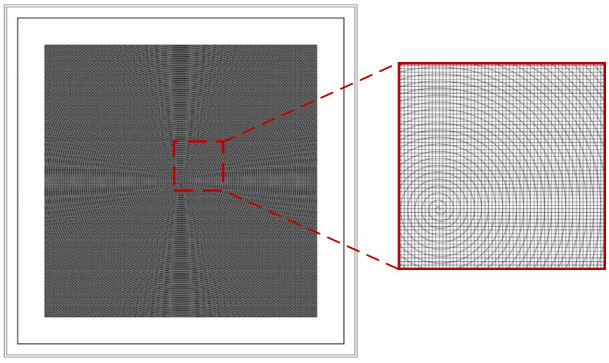

Fig. 3. Design of the marker for cross-scale motion measurement.

$$H_{\sum}(\mu,\nu) = [G_1^{\sum}(\mu,\nu), G_2^{\sum}(\mu,\nu), \ldots, G_j^{\sum}(\mu,\nu)] \tag{3}$$

where $o = 1, 2, \ldots, j$ denotes each gradient direction, $(\cdot)^+$ represents the calculation $(a)^+ = max(a, 0)$, $G_{\sum}$ corresponds to the Gaussian kernel, and $\sum$ controls the size of circular region. Therefore, the histogram of the Gaussian convolution gradient for each pixel is obtained as $H_{\sum}$.

In terms of the neighborhood structure, the descriptor can be constructed as:

$$d(\mu,\nu) = \begin{bmatrix} H^T_{\sum 1}(\mu,\nu) \\ H^T_{\sum 1}(l_1(\mu,\nu,r_1)), \ldots, H^T_{\sum 1}(l_n(\mu,\nu,r_n)) \\ H^T_{\sum 2}(l_1(\mu,\nu,r_1)), \ldots, H^T_{\sum 2}(l_n(\mu,\nu,r_n)) \\ \vdots \\ H^T_{\sum m}(l_1(\mu,\nu,r_1)), \ldots, H^T_{\sum m}(l_n(\mu,\nu,r_n)) \end{bmatrix} \tag{4}$$

where m denotes the layer number, n represents the direction of each layer, and $l_i(\mu,\nu,r_j)$ denotes the jth neighborhood location of pixel (μ,ν) in the ith direction with a distance r. Given the separability of the Gaussian filter functions, recalculation of the convolution histograms is not necessary. For image registration, parameters $n = 8$ and $m = 2$ are chosen to attain the highest accuracy [21].

An enhancement procedure is employed after extracting feature information from regional microscopic images, as illustrated in Fig. 2(b). This procedure encompasses the following steps: normalizing the vectors to enhance invariance to illumination, reducing the dimension of the feature descriptor through principal component analysis to further enhance efficiency, and compressing the storage occupancy of the descriptor through dynamic quantization for multi-image stitching.

After constructing the feature descriptors for every pixel in the overlapping regions, feature matching is performed employing the Euclidean distance to assess the similarity:

$$D(d_i, d_j) = \sqrt{\sum_{i=1}^{n}(d_i - d_j)^2} \tag{5}$$

where d_i and d_j are descriptors in different microscopic images. Once image registration is finalized in the overlapping regions, the transformations among multiple microscopic images are computed. The global map is ultimately derived from regional microscopic images through an image compositing algorithm [22].

2.2. Real-time FOV recognition using the learning-based approach

By moving the macro-micro robot and automatically capturing the microscopic image sequence, the global map can then be constructed using the aforementioned method, extending the range of microvision-based sensing. Thus, the macro-motion measurement remains real-time FOV recognition within the global map.

Robotic systems have significantly benefited from the implementation of learning-based approaches [23], particularly in the domain of precise microvision-based measurement. To attain high accuracy and efficiency in FOV recognition, learning-based approaches are considered, with a particular focus on utilizing the deep convolutional neural network (DCNN) for image recognition. DCNN demonstrates an exceptional performance owing to several factors: (1) they facilitate effective dimension reduction to handle large amounts of image data, thereby reducing the number of required parameters; (2) they preserve original image features, ensuring recognition accuracy; (3) they operate in a manner akin to the human visual system, progressively abstracting the sensing image; and (4) they enable data-driven automatic extraction of image features, eliminating the need for specific image feature selection in model design.

Consequently, a real-time FOV recognition method is developed based on the DCNN approach. To further augment the performance of the DCNN model, a residual learning architecture [24] is incorporated into the model design. Including shortcut connections in the neural network enhances gradient propagation and model generalization,

culminating in creating a residual neural network model specifically designed for estimating the FOV location within the global map.

The designed DCNN model is comprised of five convolutional layers and one fully connected layer. Initially, a convolutional layer with 64 filters, a substantial filter size of 5 × 5, and a step size of 2 is devised to downsample the microscopic images from the input layer. To expedite convergence and enhance the stability of the model, a batch normalization module [25] has been integrated. ReLu activation [26] follows to introduce nonlinear factors and enhance the expressive ability of the learning model. Following this, through maxpooling, the maximum characteristic response point in the 2-neighborhood is selected, with the step size set to 2 to avoid redundant scanning of the same area. This approach reduces the model calculations while retaining the main texture features of the image. The first designed residual block includes 64 convolutional filters with a size of 3 × 3, batch normalization, and a ReLu layer. It is noteworthy that the batch normalization and the ReLu layer precede the convolutional layer to further diminish model error. In the second residual block, the number of filters is increased to 256 to adequately extract image features. The branch contains filters with a size of 1 × 1 to balance the dimensions. Subsequently, the global average pooling layer is connected to reduce network parameters and mitigate over-fitting. A fully connected layer serves as the classifier, with its output corresponding to the total FOV image number. Finally, the Softmax function is employed to rank the data from the fully connected layer, and the region estimation is ultimately provided in the output layer. Fig. 2(c) illustrates the detailed structure of the designed DCNN model, where the colors denote different types of layers.

The aforementioned DCNN model is then trained using the labeled regional FOV images, which can be obtained by labeling and scanning the workspace of the macro-micro robot using the computer microvision. Thus, the trained DCNN model would contain the global map information for location estimation of the real-time FOV.

2.3. Cross-scale motion estimation for robotic macro-micro manipulation

During the cross-scale motion measurement, the macro-robot equipped with the microstructural marker moves automatically first. Subsequently, the stationary microvision system captures a sequence of microscopic images of regional FOVs, which are later stitched together using the previously mentioned dense feature matching technique. A global map is then constructed using the designed pattern from the microstructural marker. This map is subsequently fed into the DCNN for real-time FOV recognition, serving as a crucial step in the overall motion measurement procedure.

After constructing the DCNN model, the computer microvision system can measure cross-scale motion in robotic macro-micro manipulation. Once the real-time FOV image is input into this learning-based model, the FOV's location is recognized based on the global map. Following coarse positioning, the displacement of the centers of different FOVs is calculated using equidistant grids on the high-resolution marker to measure the movement of the macro-positioning robot. Using Otsu's method, binary thresholding is performed, followed by image closing to obtain clear patterns of real-time FOV images and the global map. Based on the known fixed distance between the grids, the macro-motion across different FOVs is obtained. During fine positioning and rotation, an automatic feature-to-phase motion tracking method from our latest work [27] is employed to precisely measure micro-motion. The process of real-time cross-scale motion measurement is shown in Fig. 2(d). Consequently, the proposed method achieves motion measurement with high efficiency, an extensive range, and high precision for robotic macro-micro manipulation using computer microvision.

3. Experiments and discussions

3.1. Experimental setup

A microvision-based motion measurement system was developed in our ISO Class 7 cleanroom. The microvision system was composed of a focus motor (TSA150-ABZ, Zolix, China), a camera (Genie TS M2048, Tendency Dalsa, Canada), a zoom lens (1-6010, Navitar, USA), an objective lens (M Plan Apo 10x, Mitutoyo, Japan), a sample stage (DZTSA300GAB-SVO1, Zolix, China), an illuminator (LMI-6000, Dolan-Jenner, USA) and a vibration isolation table (Huaan Precision Metrology, China). Specifically, the total magnification of the microvision system after calibration was 16.25, and the corresponding pixel intensity was 0.355 μm/pixel. As presented in Fig. 4(a), a macro-positioning robot based on the 3-PRR parallel mechanism (P and R stand for prismatic and resolute joints, respectively) and a micro-positioning robot based on 3-RRR compliant mechanism were designed and manufactured in our previous work [28,29], which were driven by ultrasonic motors (U-264, PI, Germany) and piezoelectric actuators (GmbH P-841.2B, PI, Germany), respectively. Both the robots and the microvision system were programmed and controlled from a PC (Intel Core i7-4770M, CPU 2.40 GHz, RAM 8 GB). Additionally, the capacitive sensor (CS) (GmbH D-E 20.200, PI, Germany) for micro-motion measurement [30] and the laser displacement sensor (LDS) (LK-H050, Keyence, Japan) for macro-motion measurement [31] were employed in the experimental setup as comparisons to the proposed method. The experimental system is shown in Fig. 4(b).

3.2. Real-time FOV recognition test

In the experiments, the marker for cross-scale motion measurement was manufactured on a glass substrate using photolithography (Hongcheng Optical, China) with 1 μm precision and 3 mm thickness. Considering the FOV size of the experimental microvision system, the microstructural pattern contains the Archimedean spiral with the constant separation distance of 200 μm, and the distance between horizontal and vertical lines was 100 μm, while all the line widths were 20 μm. During the experiments, the microstructural marker was affixed to the center of the robot, and the pattern was utilized for real-time FOV recognition and cross-FOV macro-motion measurement, while the transparent background provided a natural texture for micro-motion tracking.

As the first step of the online experiment, global mapping was conducted using the method proposed in Section 2.1. To obtain various regional microscopic images, the robot moved equidistantly while the microvision position was fixed. A zigzag path was strategically planned to enhance the efficiency of scaning different FOV, as indicated by the red arrow in Fig. 5. Corresponding to the 1200 × 1200 image resolution and the 30% overlap percentage, the step size of the macro-positioning robot for switching FOVs was set to 298.2 μm. After acquiring the microscopic image sequence, image stitching was performed to construct the global map, where the process diagram and the resulting diagram are also shown in Fig. 5. Consequently, the effective sensing range of the microvision was extended from 426 × 426 μm to 3.11 × 3.11 mm.

Real-time FOV recognition tests were conducted to validate the effectiveness of the proposed learning-based model. Based on the labeled FOV images, the residual neural network designed in Section 2.2 was trained. Specifically, the microscopic images of each area were randomly cropped into 100 images with a 1000 × 1000 pixel resolution for training data augmentation. Random rotations from −5 to 5 degrees and rescaling from 0.9 to 1.1 were applied to enhance the robustness of the model. Twenty percent of the training data was allocated for validation. The Adam optimization algorithm was employed during the training process with a 0.001 initial learning rate and a 128 mini-batch size. A piece-wise learning rate strategy was implemented, with a 0.1 learning rate drop factor every 10 epochs. The training was conducted

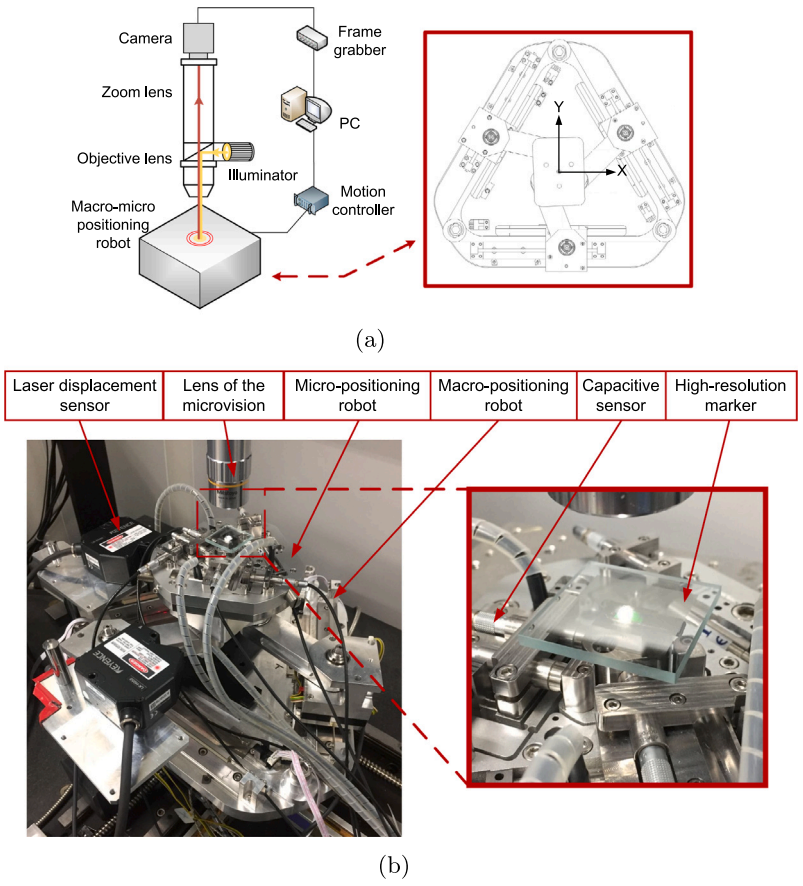

Fig. 4. Experimental setup. (a) Schematic diagram of the microvision system and the macro-positioning robot; (b) View of the established experimental system.

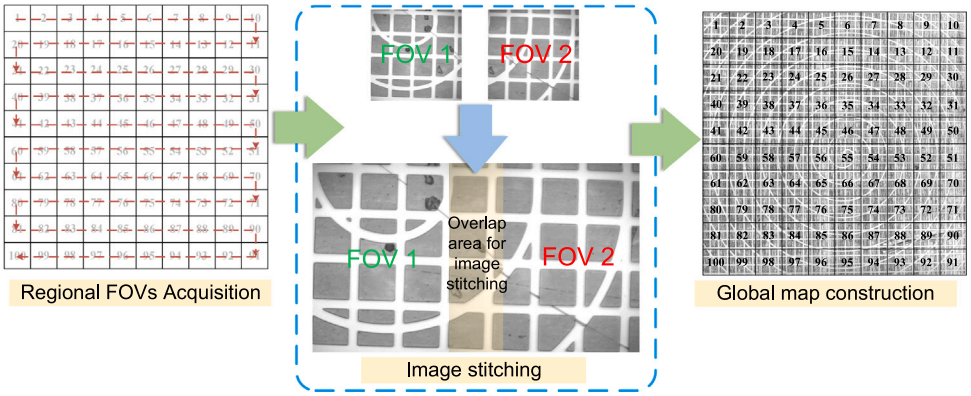

Fig. 5. The process of global map construction.

with GPU acceleration (GeForce GTX 1080Ti, NVIDIA, USA), resulting in a total training time of 14 min and 33 s. The validation loss decreased from 15.34 to 0.01 throughout the iteration, ultimately achieving a 100% validation accuracy after training.

To verify the effectiveness of the trained residual neural network for real-time FOV recognition, 200 tests were conducted while the macro-positioning robot moved randomly within the constructed global map. Additionally, as representatives of feature matching and template matching, the SURF-based and SSD-based microvision methods [32,33] were employed for comparison with the proposed approach. The results of the real-time FOV recognition tests are listed in Table 2. Both the proposed learning-based approach and template matching achieved 100% recognition accuracy in the tests, indicating the effectiveness of the microstructural marker design in making the pattern distinguishable enough for different regions in the global map. However, the speed of template matching was more than 108 times slower than the proposed approach. Although feature matching was faster than

Table 2
Results of real-time FOV recognition.

Methods	Performance	
	Accuracy (%)	Speed (s)
Template matching	100	22.87
Feature matching	88.5	12.98
Proposed approach	100	0.21

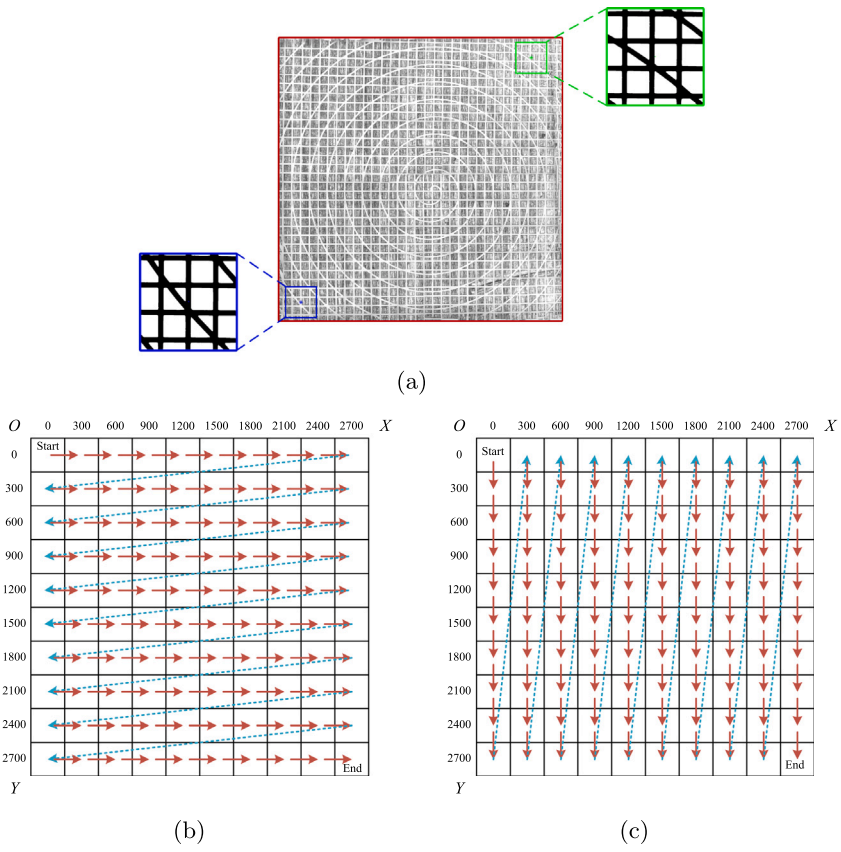

Fig. 6. Evaluation of macro-motion measurement. (a) The location of the FOVs in the global map before and after diagonal positioning of the robot, where the sub-images are the real-time FOV after binarization and morphological processing. Displacement of the macro-positioning robot in (b) *x*-direction and (c) *y*-direction for robustness evaluation.

Table 3
Experimental results of the proposed method for motion measurement of macro-positioning robot.

Direction	Microvision (mm)	LDS (mm)	Deviation (mm)	Linearity (%)
X	2.6526	2.6471	0.009	0.34
Y	2.6799	2.6856	0.007	0.26

template matching in real-time FOV recognition, the accuracy was relatively low. One reason for this is that feature matching is based on feature descriptor extraction, and there are many feature mismatches without effective algorithms to remove those outliers. The proposed approach enjoys better FOV recognition accuracy and speed compared to conventional matching methods, thus it can rapidly and precisely locate the real-time FOV in the global map.

3.3. Evaluation of macro-motion measurement

To verify the measurement performance of the proposed cross-FOV measurement method for macro-motion, a LDS system with two laser displacement sensors was constructed. The LDS had a measuring range of 10 mm, a linearity of ±0.02%, and a repeatability of 0.025 μm. During the experiment, the macro-positioning platform was controlled to travel on the diagonal of the constructed global map, thoroughly verifying the range, as illustrated in Fig. 6(a). Real-time FOV images were acquired before and after the macro-positioning robot movement, and the motion was estimated based on the equidistant grid on the global map. Simultaneously, the LDS measured the motion as a benchmark to evaluate the proposed method. Ten groups of experiments were conducted, and the maximum measurement deviation of the microvision system and the LDS in the *x*-direction was 0.009 mm with a linearity of 0.34% in the large-stroke movement of 2.6526 mm. The maximum measurement deviation in the *y*-direction was 0.007 mm with a linearity of 0.26% in the large-stroke movement of 2.6856 mm, as summarized in Table 3.

To assess the robustness of the macro-motion measurement, the macro-positioning robot was moved along both the x- and *y*-directions over the constructed global map. In Figs. 6(b) and 6(c), the red arrows illustrate the displacement of the robot. The robot was moved in steps of 300 μm from its starting position at 0 to the final position of 2700 μm. Measurement data from both the microvision and the LDS were collected, and the measurement errors of the proposed method were calculated. The experimental results of 200 measurements are presented in Fig. 7(a). The average error of the cross-FOV motion measurement in the *x*-direction was 2.3318 μm, with a standard deviation of 1.9549 μm. The measurement errors were mainly distributed between −5 to 5 μm, as shown in Fig. 7(b). In the *y*-direction, the average position measurement error was 2.4514 μm, and the standard deviation was 2.1439 μm. The measurement errors followed a normal distribution with the mean near 0, as presented in Fig. 7(d). These results indicate that the measurement errors of the motion of the macro-positioning robot were primarily random errors in the experiments. The errors are acceptable for macro-motion measurement because the manufacturing accuracy of the marker in the experiment was 1 μm. The measurement accuracy for macro-motion can be further increased by adopting a marker with higher precision.

Subsequently, the macro-rotation measurement capability of the microvision system was tested using the macro-positioning robot. Specifically, the macro-positioning robot was controlled to perform a reciprocating rotation, which was measured in real-time by the microvision system using the method described in Section 2.3. The experimental result is presented in Fig. 7(c), where the blue line represents the measurement result of the microvision system, and the red dashed line represents the planned trajectory of the robot. The experimental

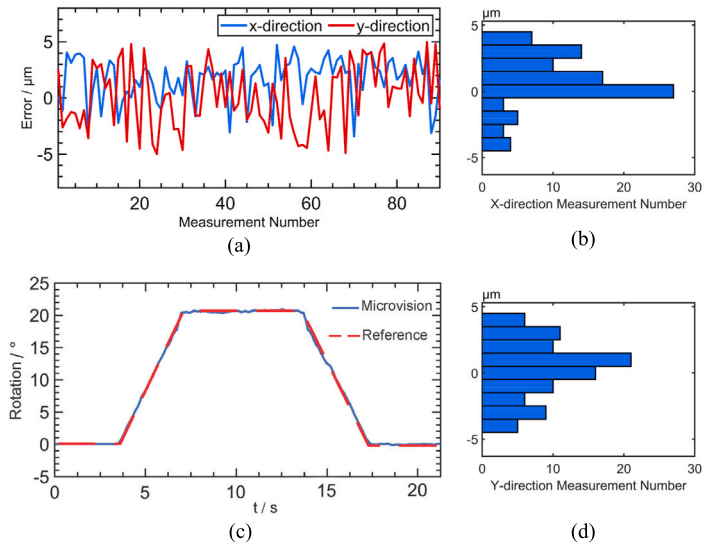

Fig. 7. Macro-motion measurement by the computer microvision and LDS. (a) Measurement errors in translation. (b) Error distribution in the x-direction. (c) Rotation by robotic macro-positioning. (d) Error distribution in the y-direction.

result shows the measurement and the planned trajectory are highly consistent, thus verifying the macro-rotation measurement capability of the proposed method.

3.4. Evaluation of micro-motion measurement

A CS system was employed as the reference to validate the micro-motion measurement ability of the proposed microvision-based method. A circular trajectory with a 10 μm radius for the micro-positioning robot was used to evaluate the displacement measurement performance. The micro-motion tracking results from the computer microvision $\boldsymbol{m}_v$ and the CS $\boldsymbol{m}_c$ are presented in Fig. 8, where a total of 955 measurements were recorded. It can be observed that the measurement results of both microvision and CS coincide closely on the x–y plane, as shown in Fig. 8(a). Furthermore, a smooth tracking trajectory from the microvision was recorded, as shown in Fig. 8(b), which is precisely aligned with the measurement of CS in terms of time. These results indicate that the proposed computer microvision can accurately estimate positions in space and timely track micro-motion, including curves and straight trajectories.

The micro-motion measurement error by the microvision can be calculated as $e_{max} = \max|\boldsymbol{m}_v - \boldsymbol{m}_c|$, where the maximum displacement measurement error e_{max} is defined by the maximum deviation between $\boldsymbol{m}_v$ and $\boldsymbol{m}_c$. As presented in Fig. 8(d), the maximum position measurement error was recorded as 0.219 μm. The root-mean-square error, which often indicates the deviation between the observed and real values, was also calculated and recorded as 0.065 μm. The angular measurement capability of the microvision system was tested by rotating the micro-positioning robot from 0 to 3 mrad in steps of 1 mrad. The microvision and CS simultaneously tracked the micro-rotation, as illustrated in Fig. 8(c). The experimental results from the microvision were consistent with those from the CS. As presented in Fig. 8(e), the maximum angular measurement error of the entire micro-rotation was found to be 0.191 mrad. The experimental results prove that the precision of the proposed method is similar to that of the CS. Moreover, the proposed microvision-based method enjoys high precision for micro-motion measurement and it has the potential to replace the expensive CS, which can simplify the system setup for robotic macro-micro positioning.

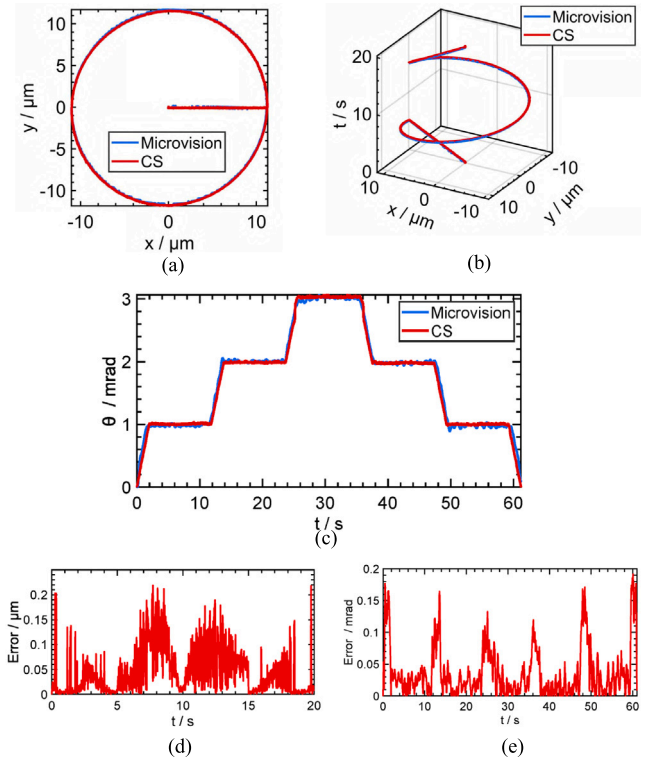

Fig. 8. Micro-motion measurement results by the computer microvision and CS in (a) x–y plane and (b) x–y–t space. (c) Rotation by robotic micro-positioning. (d) Position measurement error of the microvision system. (e) Angular measurement error of the microvision system.

3.5. Discussion

The designed microstructural marker in the global map construction serves a dual purpose—it enriches pattern texture for efficient image stitching and ensures controllable measurement errors. In particular this marker facilitates accurate measurement of cross-FOV macro-motion. The marker-assisted global map construction transforms the issue of image stitching errors in cross-scale motion measurement into

manufacturing errors of the marker, a task easily managed by the photolithography process. Consequently, the proposed microvision-based measurement method enhances the accuracy of measuring cross-FOV macro-motion.

Although the tracking frequency of the microvision-based method under the current experimental setup is lower than CS due to the computation of image processing, with the advancement of the computer hardware, utilization of parallel processing procedures, conversion from MATLAB to a more efficient programming language such as C++, the tracking frequency can be significantly improved. Moreover, the microvision-based motion measurement enjoys the following attractive properties: (1) direct visualization result; (2) multi-DOF motion measurement; (3) easy installation for robotic systems; (4) tremendous information for further development. As a result, the microvision-based method can be potentially used for micro-part assembly, cell manipulation, closed-loop control of robotics, micro-force sensing, full-field strain and stress measurement.

4. Conclusion

This paper introduces a marker-assisted microvision system to enable cross-scale measurement for robotic automation tasks, particularly in the context of macro-micro manipulation. A microstructural marker is designed and manufactured to extend the sensing range while maintaining measurement accuracy. A learning-based approach is proposed for macro-motion measurement, and a micro-motion tracking approach is integrated to ensure compatibility with robotic macro-micro manipulation. The key performances of the proposed method are evaluated through several experimental studies using the established setup. Under a fixed magnification, the microvision system exhibits high accuracy in cross-scale measurement, i.e., 9 μm for macro-motion, 219 nm for micro-motion in translation, and 0.191 mrad in rotation. This microvision-based method can offer effective visual feedback, enabling potential closed-loop control applications for robotic macro-micro manipulation, assembly, or sample mechanical characterization.

CRediT authorship contribution statement

Sheng Yao: Writing – original draft, Visualization, Validation, Software, Methodology, Funding acquisition. **Xianmin Zhang:** Supervision, Resources, Funding acquisition, Conceptualization. **Sergej Fatikow:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant 52130508 and Grant 52305023, the Guangzhou Science and Technology Planning Project under Grant 2023A04J2039. The authors gratefully acknowledge the above support agencies.

References

[1] J.A. Marvel, R. Bostelman, J. Falco, Multi-robot assembly strategies and metrics, ACM Comput. Surv. 51 (1) (2018) 1–32.
[2] R. Wang, H. Wu, Design and performance of a spatial 6-R RRR compliant parallel nanopositioning stage, Micromachines 13 (11) (2022) 1889.
[3] Q. Xu, Design and smooth position/force switching control of a miniature gripper for automated microhandling, IEEE Trans. Ind. Inform. 10 (2) (2013) 1023–1032.
[4] M. Torralba, J.A. Yagüe-Fabra, J.A. Albajez, J.J. Aguilar, Design optimization for the measurement accuracy improvement of a large range nanopositioning stage, Sensors 16 (1) (2016) 84.
[5] S. Yao, H. Li, S. Pang, B. Zhu, X. Zhang, S. Fatikow, A review of computer microvision-based precision motion measurement: Principles, characteristics, and applications, IEEE Trans. Instrum. Meas. 70 (2021) 1–28, Art no. 5007928.
[6] H. Li, X. Fang, Z. Zhu, W. Fu, C. Zhao, The approach of nanoscale vision-based measurement via diamond-machined surface topography, Measurement (2023) 112814.
[7] N. Marturi, B. Tamadazte, S. Dembélé, N. Piat, Image-guided nanopositioning scheme for SEM, IEEE Trans. Autom. Sci. Eng. 15 (1) (2016) 45–56.
[8] F. Qin, D. Xu, D. Zhang, W. Pei, X. Han, S. Yu, Automated hooking of biomedical microelectrode guided by intelligent microscopic vision, IEEE/ASME Trans. Mechatronics 28 (5) (2023) 2786–2798.
[9] S. Yao, H. Li, S. Pang, L. Yu, S. Fatikow, X. Zhang, Motion measurement system of compliant mechanisms using computer micro-vision, Opt. Express 29 (4) (2021) 5006–5017.
[10] C. Zhao, J. Xiang, C.F. Cheung, Sub-microscale precision repeatability position measurement using integrated polar microstructure and feature extraction method, Measurement (2023) 113254.
[11] B. Potsaid, Y. Bellouard, J.T. Wen, Adaptive Scanning Optical Microscope (ASOM): A multidisciplinary optical microscope design for large field of view and high resolution imaging, Opt. Express 13 (17) (2005) 6504–6518.
[12] Y. Cheng, J. Cao, X. Tang, Q. Hao, Optical zoom imaging systems using adaptive liquid lenses, Bioinspiration Biomim. 16 (4) (2021) 041002.
[13] A.N. André, P. Sandoz, B. Mauzé, M. Jacquot, G.J. Laurent, Sensing one nanometer over ten centimeters: A microencoded target for visual in-plane position measurement, IEEE/ASME Trans. Mechatronics 25 (3) (2020) 1193–1201.
[14] A.N. André, P. Sandoz, B. Mauzé, M. Jacquot, G.J. Laurent, Robust phase-based decoding for absolute (x, y, Θ) positioning by vision, IEEE Trans. Instrum. Meas. 70 (2020) 1–12.
[15] A.N. André, O. Lehmann, J. Govilas, G.J. Laurent, H. Saadana, P. Sandoz, V. Gauthier, A. Lefèvre, A. Bolopion, J. Agnus, V. Placet, C. Clévy, Automating robotic micro-assembly of fluidic chips and single fiber compression tests based-on $XY\Theta$ visual measurement with high-precision fiducial markers, IEEE Trans. Autom. Sci. Eng. (2022) 1–14.
[16] W. Hu, Y. Ma, Z. Zhan, D. Hussain, C. Hu, Robotic intracellular electrochemical sensing for adherent cells, Cyborg Bionic Syst. (2022).
[17] W. Hu, Y. Ma, Z. Zhan, D. Hussain, C. Hu, Robotic intracellular electrochemical sensing for adherent cells, Cyborg Bionic Syst. (2022).
[18] X. Guo, Y. Zhang, M. Cao, Q. Shu, A. Knoll, H. Jiang, Y. Ying, M. Zhou, Mechanical force characterization of living cells based on needle deformation, Adv. Mater. Interfaces 10 (20) (2023) 2300293.
[19] Z. Wang, Z. Yang, Review on image-stitching techniques, Multimedia Syst. (2020) 1–18.
[20] E. Tola, V. Lepetit, P. Fua, Daisy: An efficient dense descriptor applied to wide-baseline stereo, IEEE Trans. Pattern Anal. Mach. Intell. 32 (5) (2009) 815–830.
[21] S. Winder, G. Hua, M. Brown, Picking the best daisy, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2009, pp. 178–185.
[22] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, M. Cohen, Interactive digital photomontage, in: ACM SIGGRAPH 2004 Papers, 2004, pp. 294–302.
[23] O.I. Abiodun, A. Jantan, A.E. Omolara, K.V. Dada, N.A. Mohamed, H. Arshad, State-of-the-art in artificial neural network applications: A survey, Heliyon 4 (11) (2018) e00938.
[24] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[25] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: International Conference on Machine Learning, PMLR, 2015, pp. 448–456.
[26] X. Glorot, A. Bordes, Y. Bengio, Deep sparse rectifier neural networks, in: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, in: JMLR Workshop and Conference Proceedings, 2011, pp. 315–323.
[27] S. Yao, X. Zhang, B. Li, L. Yu, S. Fatikow, A microvision-based motion measurement system for nanopositioners using the feature-to-phase method, IEEE Trans. Instrum. Meas. 72 (2023) 1–11, Art no. 5001411.
[28] L. Xie, Z. Qiu, X. Zhang, Experimental research of loading effect on a 3-DOF macro-micro precision positioning system, in: International Conference on Intelligent Robotics and Applications, Springer, 2017, pp. 777–787.

[29] R. Wang, X. Zhang, Optimal design of a planar parallel 3-DOF nanopositioner with multi-objective, Mech. Mach. Theory 112 (2017) 61–83.
[30] R. Wang, X. Zhang, Parameters optimization and experiment of a planar parallel 3-DOF nanopositioning system, IEEE Trans. Ind. Electron. 65 (3) (2017) 2388–2397.
[31] L. Xie, Z. Qiu, X. Zhang, Development of a 3-PRR precision tracking system with full closed-loop measurement and control, Sensors 19 (8) (2019) 1756.
[32] H. Li, X. Zhang, B. Zhu, S. Fatikow, Online precise motion measurement of 3-DOF nanopositioners based on image correlation, IEEE Trans. Instrum. Meas. 68 (3) (2019) 782–790.
[33] P. Kumar, P. Devanand, G. Alexander, Saritha, K. Sujathan, R. Deepak, Automated matching of pixel of interest between two digital images from two different microscope imaging devices, in: 2020 IEEE Recent Advances in Intelligent Computational Systems, RAICS, 2020, pp. 96–100.