A Review of Computer Microvision-Based Precision Motion Measurement: Principles, Characteristics, and Applications

Sheng Yao, Hai Li, Shuiquan Pang, Benliang Zhu, Xianmin Zhang, and Sergej Fatikow

***Abstract*— Microengineering/nanoengineering is an emerging field that enables engineering and scientific discoveries in the microworld. As an effective and powerful tool for automation and manipulation at small scales, precision motion measurement by computer microvision is now broadly accepted and widely used in microengineering/nanoengineering. Unlike other measurement methods, the vision-based techniques can intuitively visualize the measuring process with high interactivity, expansibility, and flexibility. This article aims to comprehensively present a survey of microvision-based motion measurement from the collective experience. Working principles of microvision systems are first introduced and described, where the hardware configuration, model calibration, and motion measurement algorithms are systematically summarized. The characteristics and performances of different microvision-based methods are then analyzed and discussed in terms of measurement resolution, range, degree of freedom, efficiency, and error sources. Recent advances of applications empowered by the developed computer microvision-based methods are also presented. The review can be helpful to researchers who engage in the development of microvision-based techniques and provide the recent state and tendency for the research community of vision-based measurement, manipulation, and automation at microscale/nanoscale.**

***Index Terms*— Microvision, micromanipulation/nanomanipulation, microscopy, motion measurement, review.**

I. Introduction

COMPUTER microvision-based precision measurement has drawn remarkable attention in recent years, which contributes to multiple research fields, such as microdisplacement analysis, microstructure fabrication, transportation, characterization, and particle motion control. With regard to the size of measured objects, Fig. 1 sketches the main scale of precision motion measurement, which covers from 0.1 to 1 mm. Aside from the contact-based measurement techniques, various noncontact methods, including electric field techniques, such as capacitive sensing [1] and inductive sensing [2], encoder-based techniques, such as optical gratings [3], and laser-based techniques, such as laser interferometry [4] and laser displacement sensing [5], have been developed and applied in micromotion/nanomotion systems. Electric field techniques utilize the principle that certain electrical quantities (capacitance, inductance, and so on) are proportional to the distance between the target and the sensor so that the displacement can be measured. However, such measurement methods are only capable of measuring a single-degree-of-freedom (DOF), and it would introduce Abbe errors and accumulation errors during the sensor fusion for multiple-DOF tasks. Commercial optical encoders provide positional and directional information based on the Moire phenomenon, but only 1-DOF motion is estimated. Laser-based techniques require a coherent light source and utilize the properties of light to calculate the distance of target before and after the motion. The laser path can be easily blocked, and the installation requirement of laser-based devices is strict. To address these limitations, the researchers have been exploring the technologies that possess the high precision and easy accessibility in precision engineering.

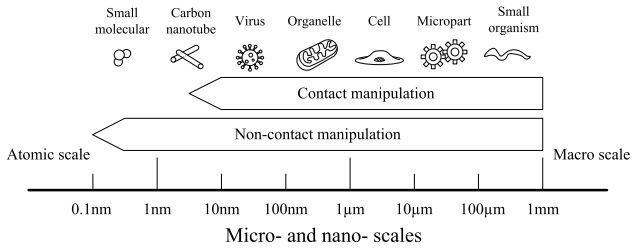


Fig. 1. Scales for microengineering/nanoengineering and manipulated targets, including CNTs, organelles, cells, microparts, and small organisms.

Vision-based techniques generate images and measure the target by digital image processing through the computer. With high flexibility and developability, they have been gradually accepted and rapidly grown as a prime solution for precision motion measurement. By image acquisition, computing in the image space, and then coordinate conversion, vision-based techniques can intuitively provide the visualized measurement results that can enable numerous applications. The computer microvision is inseparable from the development of multiple technologies, particularly image sensors and microscopes. Inventions of image sensors at Bell Laboratories empowered digital image processing [6], [7]. Concretely, the

Manuscript received January 8, 2021; accepted March 1, 2021. Date of publication March 10, 2021; date of current version March 24, 2021. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 51820105007 and Grant 51905176 and in part by the Pearl River Nova Program of Guangzhou under Grant 201906010061. The Associate Editor coordinating the review process was Emanuele Zappa. *(Corresponding authors: Hai Li; Xianmin Zhang.)*

Sheng Yao, Hai Li, Shuiquan Pang, Benliang Zhu, and Xianmin Zhang are with the Guangdong Key Laboratory of Precision Equipment and Manufacturing Technology, South China University of Technology, Guangzhou 510640, China (e-mail: me_syao@mail.scut.edu.cn; lihai@scut.edu.cn; mepang.sq@mail.scut.edu.cn; meblzhu@scut.edu.cn; zhangxm@scut.edu.cn).

Sergej Fatikow is with the Division of Microrobotics and Control Engineering, Department of Computing Science, University of Oldenburg, 26129 Oldenburg, Germany (e-mail: sergej.fatikow@uni-oldenburg.de).

Digital Object Identifier 10.1109/TIM.2021.3065436

metal–oxide–semiconductor (MOS) technology developed in 1959 [8] is regarded as the basis for modern image sensors. Later in 1969, the charge-coupled device (CCD) was invented [9] and became the most widely used image sensors for decades, until recently being surpassed by the complementary-MOS (CMOS) active-pixel sensor [10]. With the computer innovations in the 1970s, digital images can be cheaply and efficiently processed in real time, which initialized a new era for vision-based technology. Such a revolutionary breakthrough also greatly benefits the field of microscopy. The microscopy technologies, such as the optical microscope (OM) and scanning electron microscope (SEM), have been extensively investigated and significantly improved for achieving high resolution, increasing efficiency, and reducing equipment costs during the last few decades. The first commercial digital microscope was sold in Japan in 1986, allowing observation to be further recorded and processed via a computer. Moreover, the invention of USB microscopes facilitates vision-based techniques adapting in a broad range of applications at low cost via the convenient USB port [11]. To exploit the advantages of each microscopy and overcome corresponding limitations, many researchers have even customized and combined different types of microscope systems. For instance, optical microscopy has been integrated into the SEM [12], [13] so that the advantages of high-speed imaging from the optical microscopy and ultrahigh-resolution from the SEM can be efficiently achieved. Besides, combining SEM with the atomic force microscope (AFM) can simultaneously perform scanning and manipulation tasks, which also raised the interest [14]–[16]. In this review, instead of customized systems with hybrid microscopes, we focus on precision motion measurement by general microscope platforms since they are more affordable for most laboratories and have been used worldwide for vision-based measurement at small scales. A review based on the general microscope platforms can also benefit the readers who want to develop their own customized systems from scratch.

There are mainly three categories of microscopies that are well-known and commonly used: optical, electron, and scanning probe microscopies. A scanning probe microscope (SPM), including AFM, can obtain high-resolution observation using a physical probe that scans the target [17]. Even though high-speed SPMs have been recently reported with a tradeoff of the imaging quality [18], [19], the current commercial SPMs could not meet the frame rate in a large field of view (FOV) for some high-demanding tasks, such as real-time visual servo for micromanipulation/nanomanipulation. Consequently, in this article, we mainly review precision motion measurement under OM and electron microscope (EM).

Compared with the other noncontact techniques employed for precision motion measurement at microscale/nanoscale, the computer microvision-based motion measurement can offer the following attractive properties.

1) *Direct Visualization Result:* Although motion measurement can be conducted without any visualization, it is much more accurate to measure geometrical information, such as position, length, and distance, by vision-based sensors with high reliability [20].
2) *Multiple-DOF Motion Measurement:* Instead of the single DOF measurement by many traditional methods in precision engineering, such as capacitive sensing and optical grating sensing, microvision-based sensing can simultaneously measure multiple DOFs at low cost.
3) *Easy Installation for Integrated Systems:* When measurement devices need to be combined with specific operating systems for certain applications, according to different sensor techniques, the system structure may need to be designed or customized in advance for the integration [21], [22]. In contrast, such systems can be easily installed under either the OM or SEM, thereby simplify the experimental setup and preparation, and enhance the accessibility and practicality of vision-based measurement.
4) *Tremendous Information From the Obtained Data:* Microvision-based sensing can provide enormous visual information of microworld/nanoworld for human operators or computers. Object recognition and their relationship can be qualified and quantified, which is particularly important due to uncertainty at the microscale/nanoscale, such as deformation of the micro-objects, size effects, and parasitic forces during microengineering/nanoengineering [23]. Sufficient information can also enable machine learning to join the research community.

Furthermore, the iteration of computers will continuously boost the computing capability of the computer microvision systems, and the efficiency of motion measurement will keep enhanced. Thus, it can be said that the computer microvision-based technique is currently the most attractive method for precision motion measurement.

Although the literature referring to vision-based displacement detection and motion estimation at macroscale can be found [24]–[26], there still lacks a review paper that focuses on computer microvision-based precision motion measurement at small scales. This article aims to comprehensively review the advances of microvision-based motion measurement methods for microengineering/nanoengineering with a clear taxonomy. The remainder of this article is organized as follows. In Section II, principles of precision motion measurement based on computer microvision that is broadly used for micromotion/nanomotion measurement will be categorized and described in detail. Section III will analyze the characteristics of different motion measurement methods in terms of their capabilities and performances, where comparisons of state-of-the-art research will also be presented and discussed. Recent advances of applications based on precision motion measurement will be introduced in Section IV. Summary and outlook will be provided in Section V.

II. Working Principles of Precision Motion Measurement Based on Computer Microvision

Since objects at the microlevel/nanolevel are beyond our naked eye, microvision systems play an important role in bringing the microworld to our observations. After digital image sequences are captured through the microvision system,

TABLE I
COMMERCIAL MICROVISION SYSTEMS FOR PRECISION MOTION MEASUREMENT

Micro-vision systems		Magnification	Imaging speed	Price	Major manufacturer
Optical micro-vision	Conventional lens	4× - 1200×	Fast	Low	Nikon, Zeiss, Olympus
	Telecentric lens	2× - 10×	Fast	Low	Opto, VST, Edmund
	Confocal	4× - 1200×	Slow	High	Nikon, Zeiss, Olympus
Electron micro-vision		20× - 400000×	Slow	High	Zeiss, FEI, Hitachi

image processing algorithms for precision motion measurement are able to track the targets at the microscale/nanoscale. Hence, the microvision hardware together with the motion tracking algorithms is the key component for vision-based precision motion measurement, where researchers have many options to configure and customize their own computer microvision-based systems.

A. Typical Hardware Configuration of the Microvision Systems

Two types of imaging systems are widely accessible and used for precision motion measurement at microscale/nanoscale, namely, imaging systems based on the optical and electron microvision systems. Table I lists the information for typical commercial microvision systems. The principles of these microvision systems and their derivatives are described in this section.

1) Optical Microvision Systems:

a) Optical imaging system with conventional lenses: Optical microvision systems are constructed based on OMs, which are designed to observe targets through transmitted light. All OMs consist of the same basic optical path components, namely, visible light source, focus knobs, stage, revolver, and optical lenses system. Combined with image sensors, digital image sequences can be obtained and then sent to the computer for further processing. There are mainly four advantages of OMs: 1) real-time imaging; 2) high levels of observational quality; 3) not affected by electromagnetic fields and radiation-free; and 4) low cost for maintenance.

As illustrated in Fig. 2, OM platforms can be divided into two main categories in terms of use: the industry-oriented microscope platform and the biomedical-oriented microscope platform. The industry-oriented microscope platform with an intensive light source can be used not only for manipulation and fabrication at microscale/nanoscale but also for observation of solid microstructure, such as metals, ceramic, and polymeric materials. Meanwhile, the targets in biological tasks, such as cells and organisms, are normally transparent. Therefore, the light source and the objective lens are on the different sides in the biomedical-oriented microscope platform so that the transmitted light can pass through the targets and the glass slide on the stage. Such construction has a larger load capacity and better expandability, leaving more space on the top of the stage and enabling multiple updates of microscopes for biomedical applications, such as a large container, focus motors, and manipulators.

Horizontally, the imaging resolution of conventional OMs is limited by the optical principle [27], which can be given as

$$R_{\text{opt}} = \frac{\lambda}{2\text{NA}} \tag{1}$$

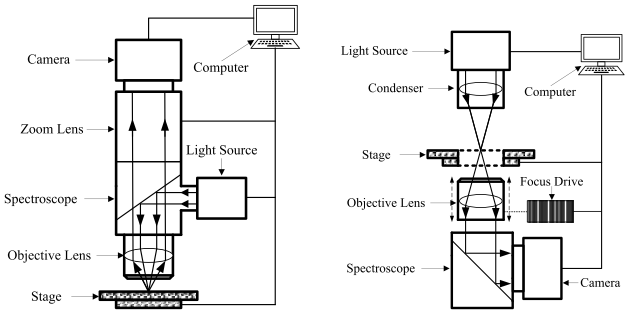


Fig. 2. Typical setup of OM platforms. (a) Industry-oriented microscope setup with the reflected illumination. An upright microscope is often included. Since most of the mechanical parts are not transparent, the light source of the industry-oriented microscope system is set at the same side of the object lens, and the light path is reflected from the stage to the digital camera. (b) Biomedical-oriented microscope setup with the transmitted illumination. A biomedical-oriented microscope platform normally contains the inverted microscope, which has the light source and condenser above the stage on the top and the objective lens below the stage pointing up.

where the imaging resolution is proportional to the wavelength of light λ, and NA is the objective numerical aperture. Vertically, the depth of focus of OMs can be calculated as

$$D = \frac{n_r}{M \cdot \text{NA}}\omega + \frac{\lambda \cdot n_r}{\text{NA}^2} \tag{2}$$

where M denotes the total magnification, n_r represents the refractive index of the medium between the objective lens and the targets, λ is the light source wavelength, and ω is the resolving power of the sensor.

b) Optical imaging system with telecentric lenses: To alleviate the shallow depth of field, the telecentric lenses can be applied to the microvision systems to meet the needs of precision measurement. The telecentric lens is a combination of conventional lens and pinhole imaging principle, which has the entrance pupil or exit pupil at infinity distance and provides an orthographic view of the targets [28]. According to different configurations based on the aperture stop position, telecentric lens systems can be classified into three types: the object-space telecentric lens system, the image-space telecentric lens system, and the double telecentric lens system [29].

Generally, the telecentric measurement systems enjoy the following features [30]–[32].

1) *Less Perspective Error:* The measurement systems with the telecentric lens can minimize the perspective error. No matter how the target moves forward or far from the lens, the projection on the image plane remains the same size. This is particularly useful for precise measurements.
2) *Large Depth of Field:* Telecentric lens systems have a larger depth of field than conventional lens systems.

Even though the target is not at the best focus position, a precise measurement can still be achieved with high contract images and effective algorithms.

3) *Low Distortion Value:* Conventional lenses may have distortions greater than 1%–2%, which would seriously affect the measurement accuracy. Telecentric lenses, however, have a very small amount of distortion, which is normally lower than 0.1% [33].

With these appealing properties, applying telecentric lenses into the microvision system is a recommended option to meet the needs of precision motion measurement.

c) Other optical imaging systems: In recent years, optical microscopy has been rapidly developed for better performances, and many commercial products are available in the market. One research direction is exploring optical microscopy from planar imaging to 3-D space so that more detailed information of the targets can be obtained and, thereby, enables many applications that require 3-D information.

A fluorescent microscope utilizes fluorescence, that is, the light emitted from targets to generate an image. Unlike reflected and transmitted light microscopy techniques, fluorescence microscopes can only observe specific structures labeled with fluorescence. Classical components of a fluorescence microscope are a light source, the excitation filter, optical lenses, the dichroic mirror, and the emission filter. After the absorption of light illuminated from the light source, the fluorescence can occur within nanoseconds and the captured by the digital camera [34]. Using the optical sectioning techniques [35], the fluorescent microscopes can generate image information in 3-D. To remove the out-of-focus blurs and noise that contaminate the images, the 3-D point spread function (PSF) can be applied to the image deconvolution process [36], [37].

The confocal microscope is a special type of fluorescent microscope with apertures placed at the confocal positions. Instead of all parts of the target excited at the same time in fluorescent microscopes, a confocal microscope uses the point illumination principle and the pinhole aperture to block out-of-focus light in image formation. Confocal microscopes offer several advantages, including: 1) elimination of out-of-focus signal in the background; 2) the capability to control depth of field; and 3) the ability to obtain serial optical sections from thick targets. With laser illumination, which is a high-intensity monochromatic light source, the laser scanning confocal microscope (LSCM) can achieve thin optical sectioning with high resolution along the z-axis [38]. Specifically, the development of high-speed and high-accuracy LSCM with a 1-nm resolution further enables precision motion measurements [39]–[41].

2) Electron Microvision Systems: Although OMs are widely used for precision motion measurement, due to the optical diffraction limit, the commercial OMs are not able to realize the visualization of nano-objects. The electron microvision system contains the EM that uses the shaped magnetic field to form an electron–optical lens system and an accelerated electron beam as the illumination source to generate images. Since the wavelength of the electron can be up to 100 000× shorter than that of the visible light, EMs possess higher resolution than conventional OMs and can be used for motion measurement and manipulation of nanoscale objects inside the chamber [42]. Typical types of EM family in microengineering/nanoengineering application include transmission electron microscope (TEM) and SEM. Extremely thin sections of the target are required (mostly around 100 nm) under TEMs, and the specimen chamber of TEM is too narrow to contain manipulators with complex functions; the electron beam penetration may damage the sample [43]. Thus, TEMs have not been commonly used in automated micromanipulation/nanomanipulation and characterization applications, and there are few existing motion measurement and object tracking algorithms based on TEMs. Images generated by SEMs contain information about samples from a depth of up to 3 μm from the sample surface with the highest resolution of 1–3 nm. At present, the commercial SEMs are divided into field emission SEMs and conventional SEMs. Generally, the field emission SEM has a higher resolution and larger magnification range. In the working space of SEM, the increasing of magnification means the higher the resolution, which opens the door for vision-based cross-scale manipulation and measurement. In addition, the SEM possesses the advantages of short imaging time, wide FOV, wide magnification range, and continuous zoom. Therefore, it is an ideal global sensor for substituting on-board position sensors in microengineering/nanoengineering and measurement tasks [44], [45]. Fig. 3 gives a brief introduction of the typical SEM architecture and imaging principle.

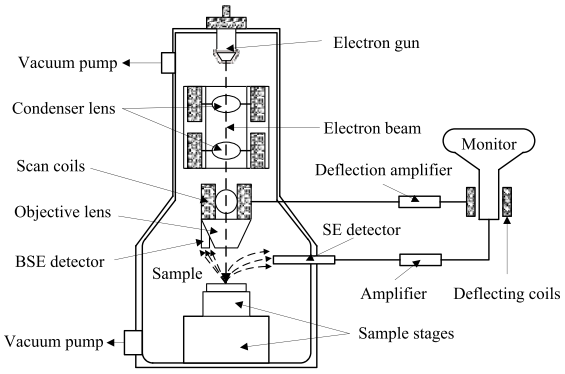


Fig. 3. Schematic of SEM setup with various components. SEM contains an illumination source called the electron gun. First, free electrons generated by the electron gun are accelerated by an electric field to form a focused electron beam with nanoscale resolution through a condenser lens. The smaller the diameter of the electron beam, the higher the resolution. Then, the scan coils are used to control the electron beam to scan the sample's surface dot by dot in a rectangular area. The focus plane on the sample is defined through the objective lens. The sample is generally mounted on the positioning stages in the vacuum chamber of SEM. When the electron beam strikes the sample surface, several types of interaction information about the topography of materials are generated. Different types of detectors can be used to collect the information as needed. The commonly used types for collecting interaction signals are the SE detector and backscattered electron (BSE) detector. Finally, digital images are generated by using the collected topographical data through the monitor.

While OMs can achieve magnifications up to 1200×, SEMs achieve a maximum magnification of more than 400 000×, which means that SEMs are capable of ultrahigh precision motion measurement. Since the magnification can be adjusted continuously, SEMs can provide a wide range of position

sensing information, so it has high flexibility for motion measurement. As the sample is scanned point by point in the sequence, there is a delay time in the process of SEM image generation [46]. By increasing the scanning speed, the delay time can be reduced, but the corresponding defect occurs that introduces a lot of noise into SEM images [47]. Thus, there is a tradeoff between the image quality and imaging time. As a result, high noise robustness is needed in motion measurement and object tracking algorithms. In the visual servo application of SEMs, many efforts focus on the accurate recognition, and tracking of the target under high scanning speed is always an important problem to be solved [48]. Corresponding solutions are summarized in Sections III-C and III-D.

The high vacuum environment brings many limitations to SEM, which is mainly manifested in that SEM is incapable of observing wet or oil-bearing samples. Besides, since electrons cannot dissipate in a high vacuum environment after striking the sample surface, conventional SEM is unable to generate clear images for nonconductive samples. As an evolution of the SEM, the environmental SEM (ESEM) aims to overcome the aforementioned issues. The first commercial ESEM was developed in the late 1980s, which is capable of observing samples in low-pressure environments and high relative humidity [49]. This is made by separating the sample chamber and electron gun chamber of the low vacuum SEM and creating a high vacuum environment in the electron gun chamber and a low-pressure environment in the sample chamber [50]. The decrease in the vacuity of the sample chamber makes the invalidity of conventional secondary electron (SE) detector, so ESEM has developed its own signal collection detectors and obtained new performance advantages. ESEMs inherit all the advantages of SEM, such as high resolution, large depth of field, and high magnification. Besides, ESEMs can still provide high-resolution images in the environment of high pressure and high temperature. It is concluded that ESEMs have better performance than SEM regardless of the difficulty of instrument manufacturing and maintenance cost. In the application of precision motion measurement, ESEMs are mainly used for the manipulation and characterization of biological samples. Since the imaging principle of ESEMs is the same as that of SEMs, the problems of noise interference and time delay still exist, which also challenges the speed and noise robustness of motion measurement algorithms at small scales.

B. Geometric Models and Calibration Methods of the Typical Microvision Systems

Before a microvision system is used for motion measurement, the geometric model should be accurately calibrated. The calibration process aims to obtain the parameters of the geometric model that defines the relationship between the world coordinate of the target and its projection in the image plane. There are two types of classical projection models, namely, perspective projection and parallel projection, which has been proven to be sufficient for most optical and electron microvision systems [51]. In general, different microvision systems can be classified into either perspective projection model or parallel projection model, which are listed in Table II.

TABLE II
GEOMETRIC MODELS FOR MICROVISION SYSTEMS

Micro-vision systems	Subconfiguration	Projection model
Optical micro-vision	Conventional lens	Perspective projection
	Telecentric lens	Parallel projection
	Confocal	Perspective projection
Electron micro-vision	Low magnification	Perspective projection
	High magnification	Parallel projection

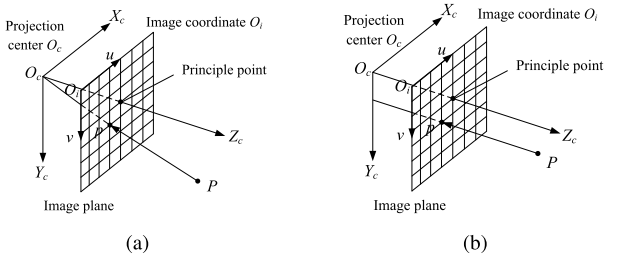


Fig. 4. Projection models of microvision systems. (a) Perspective projection. (b) Parallel projection.

1) Perspective Projection Model: Fig. 4(a) illustrates the diagram of perspective projection model. Assuming that $P = (X, Y, Z)$ is a point on the target in the world coordinate, $p = (u, v)$ is its projection on the image plane in the image coordinate, and the perspective projection can be expressed as

$$\begin{bmatrix} \mu \\ \nu \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & \gamma & \mu_0 \\ 0 & \beta & \nu_0 \\ 0 & 0 & 1 \end{bmatrix} [\boldsymbol{R}, \boldsymbol{T}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3}$$

where the intrinsic parameters include the principle point (μ_0, ν_0), the scale factors of each direction α and β, and the skew γ, which can denote the manufacturing error of sensor array. $[\boldsymbol{R}, \boldsymbol{T}]$ are the extrinsic parameters that consist of the rotation matrix and the translation vector. The intrinsic parameters of the projection model depend on the hardware, while extrinsic parameters are related to the position of the image sensor and the calibration pattern.

The geometric model of OMs with conventional lenses can be simplified as the perspective projection model. However, optical distortions always make the model nonlinear. The calibration itself is the main procedure that can reduce model errors. Based on different calibration patterns, several calibration methods for optical microvision systems with conventional lenses have been proposed. In [52], a practical calibration method considering the distortion compensation was proposed using a line pattern, while work [53] employed the grid pattern for calibrating the conventional OMs. In [54], a 2-D dot array pattern was applied in the calibration process of a special geometric model of the microscope. The works [55], [56] presented so-called virtual-image-based or pattern-free calibration methods, which utilized the mechanical setups themselves from the micromanipulation/nanomanipulation as the calibration objects, such as micromanipulators. The learning-based methods can also be applied in calibration to handle the residual error in FOV for different systems [57]. These methods can ensure high system

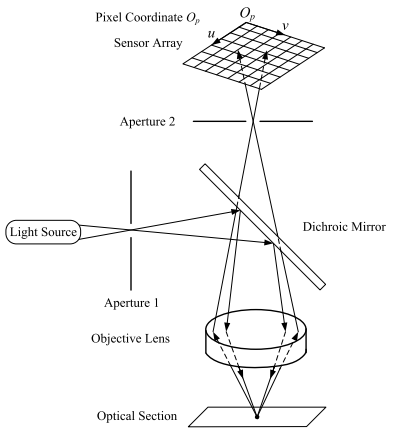


Fig. 5. Schematic of the confocal projection model. The coherent light from the light source passes through an aperture set at the conjugate plane. After reflected by the dichroic mirror and transmitted through the objective lens, a scanning point of the targeted is illuminated. The second aperture is placed in front of the sensor. When the entire target in a defined focal plane is scanned, the secondary fluorescence emitted from the point on the target will return through the dichroic mirror and focus on a confocal point at the second aperture, which is placed in front of the sensor. Then, the image of an optical section on the focal plane is captured by moving either the specimen stage or the light beam.

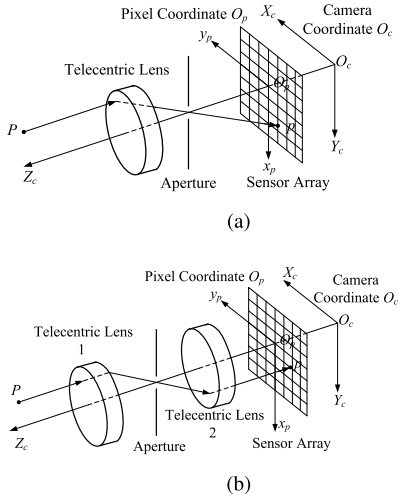


Fig. 6. Schematic diagrams of telecentric projection models. (a) When the aperture stop is placing at the rear focal plane, the entrance pupil at infinity is maintained, and it is defined as the object-space telecentric lens system. Only parallel light is allowed to reach the image plane. The image magnification is independent of the distance or position of the object in the FOV according to the geometrical relationship. (b) Placing another aperture stop at the front focal plane, one can make a double telecentric system. The double telecentric system enjoys the properties that the captured image is free of both the axial target shift and the image plane shift. Since the principal light intercept position on the image plane does not change, the double telecentric lens has a more accurate magnification and, thereby, allows precise measurement of targets regardless of the position.

measurement accuracy by modeling the imaging geometry and optical lens distortion.

The projection model of the confocal microscope can also be seen as the perspective projection model, which is diagrammatically presented in Fig. 5. Capturing multiple 2-D images at different depths in a target allows the reconstruction of 3-D structures within an object [58]. Compared to the conventional OMs, LSCMs are able to characterize micro-objects in 3-D. Thus, many efforts of LSCM calibration have specifically focused on the z-dimension. Wang *et al.* [59] employed the Monte Carlo method to address the measurement uncertainty issue and calibrate LSCMs in the z-dimension. Martínez and Oliva [60] proposed a practical calibration procedure of LSCMs, which can be applied in both research and industrial departments. Consider that more measurement conditions can further improve the calibration accuracy. Mudrak *et al.* [61] took into account the refractive medium for calibrating the LSCMs, and higher measurement accuracy is achieved.

2) Parallel Projection Model: In the parallel projection model, the projection center is located at the infinite, and the projection rays are parallel, as presented in Fig. 4(b). Since there is no longer a principle point in parallel projection, the general expression of parallel projection is

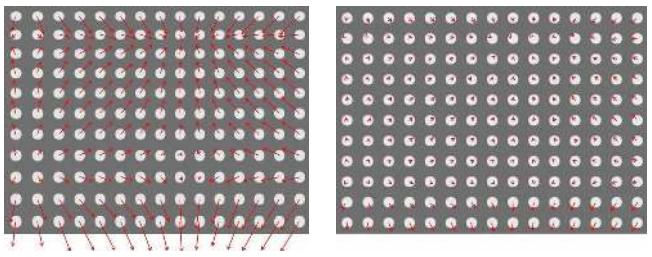
$$\begin{bmatrix} \mu \\ \nu \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & \gamma & 0 \\ 0 & \beta & 0 \\ 0 & 0 & 1 \end{bmatrix} [\boldsymbol{R}, \boldsymbol{T}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{4}$$

where α, β, and γ are the intrinsic parameters, and $[\boldsymbol{R}, \boldsymbol{T}]$ are the extrinsic parameters.

The telecentric lens imaging systems perform parallel projection, which is illustrated in Fig. 6. Similar to the conventional lens, the precise projection models of telecentric lenses can be obtained by the calibration process. Quan [62] developed a self-calibration method for the telecentric lens imaging system. However, this type of method is appropriate for a moving camera, and it may not be flexible enough for a microvision system due to the difficulty of focusing on the microscale workspace at different positions. When it comes to the method with calibration patterns as the calibrator, there are 3-D and planar patterns for the telecentric lens imaging system. The factorization approach [33] works well in the telecentric lens systems, but solving the mirror ambiguity of the pose of the camera may be problematic. In work [28], the planar pattern was moved along the z-direction of the world coordinate with a precision positioning stage, so the extrinsic parameter can be determined by comparing the reprojection with two poses of camera. In contrast, the complete position of the mark points is known in the 3-D-pattern method [63], in which the camera parameters can be estimated by a direct linear transformation.

For SEMs, both projection models have been used in terms of magnifications [64]. At low magnification (lower than 500×), the perspective projection model is selected, as the scanning electron beam, FOV, and the angle of the observation field are large. At high magnification (higher than 500×), the parallel projection model is applied instead since the projection center is assumed to be infinite. Different researchers have different views about the magnification boundaries for selecting the perspective model and the parallel model [65].

Different from OMs, special multiscale calibration patterns are needed for calibrating SEMs under variational magnifications. For instance, multiscale chessboard grids are used to determine the internal and external parameters of SEMs in the magnification range of 300× to 10 000× by iteratively nonlinear minimizing the registration error [66]. Multiscale



(a) (b)

Fig. 7. Example of calibration using circular patterns. (a) Nonlinearity distortions with red arrows. (b) Residual errors have been significantly reduced after calibration. Reproduced from [71] with the permission from IOP Publishing.

TABLE III

CATEGORIES OF MICROVISION-BASED MOTION MEASUREMENT ALGORITHMS

Measurement techniques	Target characteristics
Template matching	Targets with the high contrast region
Feature matching	Targets with salient features
Phase correlation	High precision, fringe-texture targets
Optical flow estimation	Every pixel location within the target region
Silhouette-based tracking	shape-various, irregular or line-type targets

circular patterns were designed to obtain the perspective matrix based on the general imaging model under magnifications from 20× to 500× [64], [67]. Trapezoidal-shape markers were fabricated to calibrate SEMs in 2-D directions under a certain magnification [68], [69]. Distinctively, pattern-free calibration methods [70] were also developed to increase the operational flexibility. Concretely, these calibration methods aim to obtain an accurate mathematical projection model or the pixel scale factors for the SEM. As can be seen in Fig. 7, the deviations of magnification and nonlinearity cause significant distortion in the FOV. With the proper calibration process, the spatial distortions and nonlinearities normally can be very much decreased.

C. Image Processing Algorithms for Precision Motion Measurement

Through the aforementioned computer microvision systems, motion measurement is primarily enabled by different kinds of motion/displacement tracking techniques. Generally, the motion measurement techniques are expected to have the capabilities of high accuracy, high speed, and multi-DOF measurement. According to the corresponding characteristics of the measured target, the existing motion measurement algorithms can be classified into five types: template matching, feature matching, phase correlation, optical flow, and silhouette-based method, as shown in Table III.

1) Template Matching: Template matching is a technique for finding small parts of a new image, which matches the predefined templates in the reference image; thereby, motion measurement can be achieved by tracking the templates in the image sequences. Since the template is usually a rectangle area and the matching process is on searching image intensities, this technique has been given different names in literature, such as block matching [72], area-based matching [73], and image correlation [74]. Nevertheless, in this review, this technique is

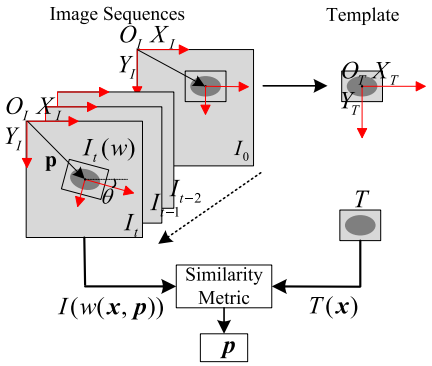


Fig. 8. Schematic of the template matching.

called template matching due to the intuitive concept and the popularity in the majority of the published literature.

The procedure of template matching is illustrated in Fig. 8 in terms of a common case: 3-DOF motion measurement. I_0 is the intensity matrix of the reference image, and $T(\boldsymbol{x})$ is an extracted target template from I_0 with the template coordinates $\boldsymbol{x} = (x, y)$, where Ω is the set of all pixel coordinates in T. Assuming that, when the image I_t is captured at time t in the image sequences, the template T moves to a new location, the warp function $w(\boldsymbol{x}; \boldsymbol{p})$ aims to map T to the new area $I_t(w(\boldsymbol{x}; \boldsymbol{p}))$, where $\boldsymbol{p} = (t_x, t_y, \theta)$, (t_x, t_y) is the translation and θ is the rotation. Then, the problem of estimating the parameters vector $\boldsymbol{p}$ and the problem of template matching can be expressed as

$$\boldsymbol{p} = \arg\ \max/\min \sum_{\boldsymbol{x}\in\Omega} S(I_t(w(\boldsymbol{x}; \boldsymbol{p})), T(\boldsymbol{x})) \tag{5}$$

where S is a certain kind of similarity metric that $\boldsymbol{p}$ can be computed by maximizing the correlation or minimizing the difference. Thus, the motion measurement of the target can be achieved by tracking the template T in the image sequences, wherein the similarity metric gets involved for searching the area that is most closely resembling the template T and then updates the warp function $w(\boldsymbol{x}; \boldsymbol{p})$.

Similarity metrics for template matching can be further classified into two categories: cross correlation (CC) criteria and the sum of squared difference (SSD) criteria. CC criteria include CC, normalized cross correlation (NCC), and zero-normalized cross correlation (ZNCC), while SSD criteria include SSD, the normalized sum of squared differences (NSSD), and the zero-normalized sum of squared differences (ZNSSD). The definitions of these similarity metrics are listed in Tables IV and V, where $T(\boldsymbol{x})$ represents the value of image intensity in a specific location in the template with a size $M \times N$ in the reference image; $I(w(\boldsymbol{x}; \boldsymbol{p}))$ is the image intensity in the candidate area with the same size of the template in the new image; T_m and I_m are the mean intensity value of the template and the candidate area; T_{rss} and I_{rss} are the root sum squares of the image intensity values; and T_{rtss} and I_{rtss} denote the root of total sum squares of $T(\boldsymbol{x})$ and $I(w(\boldsymbol{x}; \boldsymbol{p}))$. While CC and SSD are sensitive to illumination variation, NCC and NSSD are less affected by the linear scale in illumination lighting but still sensitive to the offset of illumination. Thus, ZNCC and ZNSSD are regarded as robust

TABLE IV
CC CRITERIA

CC correlation criterion	Definition
Cross-correlation (CC)	$S = \sum_{i=-\frac{M}{2}}^{\frac{M}{2}} \sum_{j=-\frac{N}{2}}^{\frac{N}{2}} T(\boldsymbol{x}) I(w(\boldsymbol{x};\boldsymbol{p}))$
Normalized cross-correlation (NCC)	$S = \sum_{i=-\frac{M}{2}}^{\frac{M}{2}} \sum_{j=-\frac{N}{2}}^{\frac{N}{2}} \frac{T(\boldsymbol{x}) I(w(\boldsymbol{x};\boldsymbol{p}))}{T_{rss} I_{rss}}$
Zero-normalized cross-correlation (ZNCC)	$S = \sum_{i=-\frac{M}{2}}^{\frac{M}{2}} \sum_{j=-\frac{N}{2}}^{\frac{N}{2}} \frac{[T(\boldsymbol{x}) - T_m][I(w(\boldsymbol{x};\boldsymbol{p})) - I_m]}{T_{rtss} I_{rtss}}$

TABLE V
SSD CRITERIA

SSD correlation criterion	Definition
Sum of squared differences (SSD)	$S = \sum_{i=-\frac{M}{2}}^{\frac{M}{2}} \sum_{j=-\frac{N}{2}}^{\frac{N}{2}} [T(\boldsymbol{x}) - I(w(\boldsymbol{x};\boldsymbol{p}))]^2$
Normalized sum of squared differences (NSSD)	$S = \sum_{i=-\frac{M}{2}}^{\frac{M}{2}} \sum_{j=-\frac{N}{2}}^{\frac{N}{2}} \left[\frac{T(\boldsymbol{x})}{T_{rss}} - \frac{I(w(\boldsymbol{x};\boldsymbol{p}))}{I_{rss}}\right]^2$
Zero-normalized sum of squared differences (ZNSSD)	$S = \sum_{i=-\frac{M}{2}}^{\frac{M}{2}} \sum_{j=-\frac{N}{2}}^{\frac{N}{2}} \left[\frac{T(\boldsymbol{x}) - T_m}{T_{rtss}} - \frac{I(w(\boldsymbol{x};\boldsymbol{p}))) - I_m}{I_{rtss}}\right]^2$

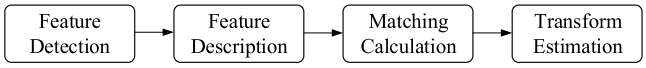


Fig. 9. General pipeline of feature matching.

criteria with the best noise-proof performance [75], [76]. When the candidate area reaches the highest similarity to the template by searching in the image, the location can be selected as the new location for the target and, thereby, achieve displacement measurement.

In addition to the template selection, template matching is easy to implement and use with few user interventions. By directly operating on the pixel values, it does not require any form of specific target pattern, and it is relatively robust to image noise since highly redundant information is involved. Hence, template matching approaches can work well even with few image features, thereby, are popular, and have been widely adopted for motion measurement in microvision systems [77]–[80]. However, the high computational cost might limit the applications for real-time motion tracking. Therefore, many improvements focus on the efficiency of template matching [81]–[83]. The details can be found in Section III-C.

2) Feature Matching: Feature matching is achieved by tracking some features, such as lines, points, shapes, and textures, in the image sequence with known geometrical information for motion measurement. This technique focuses on the extraction of image feature to be matched in the new image and puts more emphasis on the detection of distinctive regions. As shown in Fig. 9, feature matching starts with interest feature detection, which should be covariant to a type of transformation. An invariant feature representation, also called the descriptor, is then constructed for each detected interest

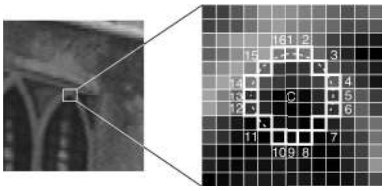


Fig. 10. Feature detection in an image with the popular detector FAST. Reproduced from [84] with the permission from IEEE.

feature. Once the feature and the descriptor are successfully extracted, the matching between the reference image and a new image can be performed by a matching criterion. Then, the transformation of the target between two images, so as the displacement, can be estimated.

Feature detection is the first step, and it is expected to have the following properties: accuracy, robustness, repeatability, efficiency, and generality. A good feature detector should also be independent of scaling, rotation, shifting, deformations, compression artifacts, and noise. Image features can be categorized into three types, namely, point, line, and region. For motion measurement, key point features and blob features are mostly selected. Key point detectors aim to extract point-like features in the image, which are known as key points. Some articles also use the term corners referring to this kind of detection since a few algorithms detect the feature by finding the rapid change in a direction, which is a corner in the traditional sense. The Moravec detector [94], for instance, is a corner detection algorithm to find the key point that has low self-similarity. As an improvement, Harris and Stephens [89] proposed a more desirable point detector by introducing intensity variation, where autocorrelation is used for image feature detection. To cut the computational cost, the SUSAN

TABLE VI
CATEGORIES OF POPULAR FEATURE MATCHING ALGORITHMS

Descriptor categories	Descriptor names	Feature detector	Matching criteria
Distribution-based descriptors	Scale-invariant feature transform (SIFT) [85] Speeded-up robust features (SURF) [86]	Difference-of-Gaussian(DoG) [85] Determinant of the Hessian (DoH) [87]	Euclidean distance
Moment-based descriptors	Zernike moment [88]	Harris [89] [90]	Tables III and IV
Binary string-based descriptors	Oriented FAST and rotated BRIEF (ORB) [91] Fast retina keypoint descriptor (FREAK) [92]	Features from accelerated segment test (FAST) [84] Adaptive and generic accelerated segment test (AGAST) [93]	Hamming distance

Detector [95] uses a low-level image processing technique to compute key points in an image instead of using image intensity variation. Similarly, features from the accelerated segment test (FAST) detector [84] take a circle of 16 pixels around the candidate pixels for calculation. If the candidate pixel is brighter or darker than a specific threshold value of a set of n contiguous neighbor pixels within the circle, then the candidate pixel can be considered as the key point feature c, as shown in Fig. 10. Many efforts have been put into improving the FAST detector. The adaptive and generic accelerated segment test (AGAST) detector [93] is one of the outstanding modifications of the FAST detector that can efficiently detect the key points by the optimal decision tree algorithm. Blob detectors, on the other hand, use more information about the image structure by detecting regional features. A blob is a region in the image containing some properties of interests that are approximately constant. There are roughly two types of blob detectors: 1) differential-based detectors, such as Laplacian of Gaussian (LoG) and 2) local extrema-based detectors, such as maximally stable extremal regions (MSERs). LoG [96] is a common blob detector and the linear combination of second derivatives. Although the LoG detector is invariant to rotation and well adapted to find scale-invariant regions, the computational cost is still relatively high. Difference-of-Gaussian (DoG) detector [85] was then developed to optimize computational efficiency. A close approximation to LoG is given in DoG based on scale-space extrema in the Gaussian pyramid so that the DoG function $D(x, y, \sigma)$ can be calculated by subtracting two nearby scale levels of the Gaussian pyramid without convolution in operation, and the feature detection can speed up. Nevertheless, using scale-space extrema of the determinant of the Hessian (DoH) for blob detection has been proven to outperform LoG and DoG [87]. Another approach for blob detection is to combine the bright or dark blob with each local extrema in the intensity. To overcome noise interference, local extrema is normally detected with the extent at multiple scales in scale-space according to watershed analogy. Therefore, MSER [97] was developed to detect local extrema to such extent. MSER has several advantages, including high stability, multiscale detection, and invariance to affine transformation in image intensity.

Once the interest points or regions have been successfully detected, their structure needs to be encored to differentiate one feature from another by a suitable descriptor. The descriptors also ought to be capable of discriminatingly matching and insensitive to local image deformations. Three major categories of the existing descriptors are given as follows.

1) Distribution-based descriptors, including scale-invariant feature transform (SIFT) [85], speeded-up robust features (SURFs) [86], and gradient location-orientation histogram (GLOH) [98] descriptors. SIFT is one of the most popular descriptors with excellent performance under different scales, rotations, and lighting. Around each detected key point, the SIFT descriptor samples the gradient magnitude and orientation in a 16×16 region, which is weighted by a Gaussian window, as shown with a circle in Fig. 11. Then, the sample accumulated into a set of orientation histograms over a 4×4 subregion, and a compact feature vector with 128 elements is formed to contain values of all the orientation histograms. Decreasing the elements in the feature vector to 64, the SURF descriptor is then developed to have a faster processing speed. The GLOH descriptor is also an extension of SIFT to enhance robustness against illumination changes by a log-polar coordinate.
2) Moment-based descriptors, such as geometric moments [99], Zernike moments [88], and pseudo-Zernike moments [100]. Zernike moments are robust to scale, noise, rotation, and capable of describing various shapes of patterns with the multilevel representation.
3) Binary string-based descriptors, such as fast retina keypoint descriptor (FREAK) [92], binary robust independent elementary features (BRIEFs) [101], and its derivative oriented FAST and rotated BRIEF (ORB) [91], encode pairs of local intensity differences into binary vectors. The construction and following the matching process of this type of descriptors are faster than others, which is a significant advantage and binary string-based descriptors have become popular in recent years.

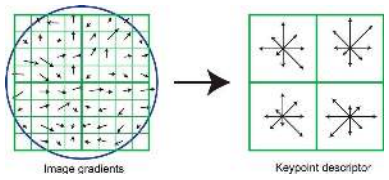


Fig. 11. One of the most popular feature descriptors, i.e., SIFT descriptor, converts the key point information into the orientation histograms. Reproduced from [85] with the permission from Springer Nature.

Once the feature descriptor is constructed, the matching of feature points or regions between two images is performed by a similarity measure. For the moment-based descriptors, the similarity metrics from Tables IV and V can be applied [102]. For distribution-based descriptors, the most common similarity measure is the Euclidean distance in feature space

$$\text{Euclidean}(I_1, I_2) = \sqrt{\sum_{i=1}^{n} (d_i(I_1) - d_i(I_2))^2} \tag{6}$$

where I_1 and I_2 are the image pair, and $d_i(I_1)$ are the feature descriptors in the images. A threshold can be set, and the descriptor distance within the threshold can be considered as the matched descriptors. Although the nearest-neighbor matching algorithms can be used to minimize the Euclidean distance between the image descriptors, the optimal matching algorithm depends on the specific data set characteristics. The randomized k-d forest and the fast library for approximate nearest neighbors (FLANNs) have been recognized most efficiently for handing high-dimensional features.

Binary string-based descriptors can be measured by the Hamming distance

$$\text{Hamming}(I_1, I_2) = \sum_{i=1}^{n} b_i(I_1) \otimes b_i(I_2) \tag{7}$$

where $b_i(I_1) \in (0,1)^n$ and $b_i(I_2) \in (0,1)^n$ denote the binary descriptors of the image pair I_1 and I_2, and $\otimes$ represents the bitwise XOR operation. The random sample consensus (RANSAC) method [103] can also be used to remove the false matching target. Eventually, with the matched target, the motion in the image sequence is tracked. Table VI summarizes the most popular feature matching algorithms that have been used in precision motion measurement.

Feature matching is a high-efficiency technique that handles sparse points rather than the entire area in template matching. Local descriptors, instead of the raw image intensity, are used to represent the image information, which results in low sensitivity to variations in illumination, shape, and scale. The feature matching process can break down into different steps and, thereby, allow different combinations of detectors, descriptors, and matching measures. Improvements in any step of Fig. 9 are valuable. Thus, researchers can have a high degree of flexibility to design and customize their own motion measurement algorithms based on feature matching. Recent advances of feature matching approaches can be found in [104]–[109].

3) Phase Correlation: The phase correlation technique is based on a linear relationship, namely, the phase-to-displacement relationship, in which the target displacement can be encoded in the phase shift of the Fourier transform and can be written as

$$\mathcal{F}(I(x-\delta)) = \mathcal{F}(I(x)) \cdot e^{-2\pi i \delta \xi} \tag{8}$$

where $\mathcal{F}$ denotes the Fourier transformation, x is the spatial coordinate, $I(x)$ stands for the pixel intensity, δ represents the spatial displacement, and ξ is the transform variable of x. The phase difference matrix is then obtained as

$$G(x) = \frac{\mathcal{F}(I(x-\delta))}{\mathcal{F}(I(x))} = e^{-2\pi i \delta \xi}. \tag{9}$$

The displacement δ can be retrieved in either the spatial or frequency domains. For example, δ can be recovered from the spatial phase correlation function $g(x) = \mathcal{F}^{-1}(G(x))$ by locating the maximum value of it

$$\delta = \arg\max\{g(x)\}. \tag{10}$$

Rotation can be also estimated by aligning the magnitudes of the polar Fourier transforms on the images [110].

While the phase processing considering the whole Fourier spectrum is time-consuming, the pattern that only involves one or a few spectral components will greatly improve the algorithm efficiency and allow real-time target tracking. Therefore, applying a periodic pattern on the target can significantly decrease the computational cost and increase the signal-to-noise ratios (SNRs). The target displacement can then be formulated as

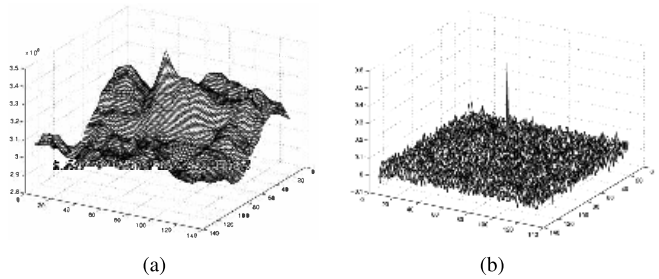
$$\delta = \frac{\phi \cdot P}{2\pi} + aP \tag{11}$$

where $\phi = 2\pi\delta\xi$ is the phase shift in the frequency domain, P stands for the period, and a denotes the entire number of periods. However, the method based on a single periodic pattern has a limited measurement range due to phase ambiguity. To enlarge the measurement range, different periodic patterns have been specifically designed, such as the pseudoperiodic pattern [111] and the fringe pattern [112]. With additional information from the second periodic pattern as a reference, the phase ambiguities can be removed. Thus, a twin-scale pattern combining two fringe sets with different periods can form a larger synthetic pattern [113], and (9) turns into

$$\delta = \frac{\phi_1 \cdot P_1}{2\pi} + bP_1 + c\frac{P_1 \cdot P_2}{|P_1 - P_2|} \tag{12}$$

where ϕ_1 denotes the phase shift for the smallest pattern set P_1, P_2 is the period of the second pattern set, b represents the number of periods obtained from the new synthetic phase, and c is the number of new periods, which stands for the new ambiguity range. With actual periods P_1 and P_2 accurately known from the manufacturer, a size reference in images is provided, enabling the direct metric transformation from pixels to meters for measurement. This self-calibrating property also allows the method to have a low sensitivity to experimental parameters. Moreover, the spectral filter introduced in the phase measurement process significantly decreases the spatial noises in images.

With the aforementioned process, only integer pixel displacement is estimated in the image sequence by locating the peak coordinates of $g(x)$ in (10). To further increase the measurement accuracy, subpixel phase correlation methods have been developed, which can be classified into two categories. In the first category, interpolation approaches have been used to precisely estimating the peak location of $g(x)$ of the normalized cross-power spectrum. Simple approaches based on fitted functions, such as the sinc function, quadratic function, and the Gaussian function, take account of the main integer peak and its two neighboring side peaks for interpolation [114]. However, such fitted function-based approaches are easily influenced by noises. According to the coherent peak theory, an analytically demonstrable method using the sinc function for the Dirichlet kernel approximation was proposed in [115]. A matrix-multiply Fourier transform approach was also developed with high efficiency [116]. In the second category, the displacement is estimated from the linear phase difference in the Fourier domain. The least-squares estimate can be employed to fit a 2-D plane that can derive the phase shift angle [117]. In [118], a straightforward approach was proposed using singular value decomposition (SVD) to



(a) (b)

Fig. 12. Corresponding standard (a) CC and (b) phase correlation from two same images with a displacement. Reproduced from [115] with the permission from IEEE.

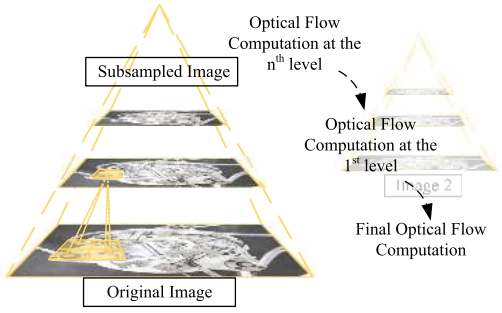


Fig. 13. Pyramidal implementation of optical flow for large motion measurement.

find the dominant approximation of the spectrum matrix, followed by a phase angle unwrapping process and the subpixel shift estimation with the least-squares estimate. Replacing the least-squares estimate with the RANSAC algorithm can further enhance precision and robustness [119].

The most remarkable advantage of the phase correlation method is high matching accuracy and effectiveness by detecting the precise peak of the correlation function, as shown in Fig. 12. Moreover, compared to other spatial-domain techniques, the phase correlation method is resilient to noise, occlusions, and other imaging defects. New developments of phase correlation-based methods can be found in [45] and [120]–[123].

4) Optical Flow Estimation: The optical flow method measures motions by detecting the brightness shift of all pixels in the target region. The optical flow itself can be defined as the apparent velocity of movement of a target formed by the relative motion between two consecutive images in the image sequence. Optical flow reflects the change of two images caused by the movement in the minute time interval to determine the movement direction and movement rate on the image point. By optical flow, the direction and rate of motion on the image point can be determined, and clues are provided to restore target movement. The optical flow method is based on the brightness constancy [124]: the pixel intensity of the target in the image does not change between successive frames. By this rule, the general equation of the optical flow method can be expressed as

$$\nabla I^T \cdot \boldsymbol{v} + I_t = 0 \tag{13}$$

where $\boldsymbol{v} = (\mu, \nu)$ denotes the optical flow with the horizontal velocity μ and vertical velocity μ, $\nabla I = (I_x, I_y)$ is the spatial gradient with $I_x = (\delta I/\delta x)$, $I_y = (\delta I/\delta y)$, and $I_t = (\delta I/\delta t)$ represents the temporal derivative at time t.

Therefore, the intensity differential at the same location is the dot product of the spatial differential and the velocity. For estimating the optical flow in (13), the gradient-based approaches have been widely studied. Horn and Schunck [125] introduced the smoothness constraint that turns the problem into minimizing the energy function

$$E_{HS}(\boldsymbol{v}) = \int_I (\nabla I^T \cdot \boldsymbol{v} + I_t)^2 + \alpha |\nabla \boldsymbol{v}|^2 = 0 \tag{14}$$

where α is a regularization constant of the smoothness term. Lucas and Kanade [126], however, proposed the spatial constraint: adjacent pixels have coherent motions so that the neighbor pixels share the same flow. Then, we can obtain

$$\begin{bmatrix} \mu \\ \nu \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n I_{xi}^2 & \sum_{i=1}^n I_{xi} I_{yi} \\ \sum_{i=1}^n I_{xi} I_{yi} & \sum_{i=1}^n I_{yi}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^n I_{xi} I_{ti} \\ \sum_{i=1}^n I_{yi} I_{ti} \end{bmatrix} \tag{15}$$

where $i = (1, \ldots, n)$ denotes the pixel number inside the neighborhood. In this way, a series of tracking points is obtained and used to create the optical flow field, and (15) can be solved when the matrix is invertible. After discarding the wrong motion estimation value by the least-squares method or RANSAC, the image motion can be estimated. Many techniques have also been proposed to estimate the optical flow, including the ones introduced in Section II-C, such as block-based methods [127] and the phase correlation-based method [128].

When the moving speed of the target is too fast, the spatial constraint is difficult to be satisfied, which will cause a large error in the final optical flow result. By reducing the image size, the relative speed in the image can be decreased. Thus, combining a multiscale pyramid [129] with the optical flow method can further increase the robustness for large motion measurement. As illustrated in Fig. 13, two pyramids are built for each of two consecutive images by their successive subsampled images. Assuming that the original image is at level 0, the top-level is level n, and the upper level in the pyramid is obtained by subsampling the image in the lower level with a factor m, which means that a pixel in the upper level can stand for m pixels in the lower level. In the iterative process, the optical flow and affine transformation are first calculated on the images of the top level with the lowest resolution from two pyramids. The calculation result of the previous level is then passed as the initial value to the image of the lower level. This process continues until the information is passed to the last level, which contains the original two images. The optical flow and affine transformation calculated by the bottom level are used as the final optical flow result. Repeating the same process to the consecutive images from the image sequence, large motions of the target can be measured.

In recent years, the optical flow method has been greatly empowered with the large displacement measurement capability [130] and high-efficiency [131]. Integrating optical flow with feature matching is another research direction so that feature descriptors instead of pixel intensity have been used in the optical flow computation. For precision motion measurement, the optical flow method allows tracking targets without the need for artificial markers and monitoring the motions in

the full FOV. It can provide fast and accurate results under a controlled environment. Optical flow not only represents the motion velocities of moving objects but also detects moving objects without knowing any information about the scene context. Consequently, optical flow estimation has excellent versatility and can be extensively employed for computer microvision systems [132]–[134].

5) Silhouette-Based Tracking: Silhouette-based tracking capitalizes the typical shape of the target to conduct motion tracking. This technique aims to find the target region in each image by a target model generated from the previous images of the image sequence. This model may have many forms of shapes, including edges or outlines. In general, the silhouette-based methods can be divided into two main categories. The first one is shape-based tracking, which searches for a specific target model in the image. The second one is contour tracking that evolves an initial contour to a new position.

Shape-based tracking directly searches the target silhouette in the current image. This process is performed by computing the similarity of the target with the associated model obtained from the hypothesized shape according to the previous consecutive images or the preexisting knowledge of the target. The target outlines are first extracted by edge-based methods, and the target edges are then matched and fitted to those in the hypothesized model by distance metrics, such as the Hausdorff metric [135]. Another method is rendering the model first, which can be constructed by extracting targets from real images or using computer-aided-design (CAD) models based on the known geometrical structures. Then, a sparse 1-D search is conducted to find the matching edges in the image, where the computation cost will go down, and real-time performance can be achieved [136]. This kind of CAD model-based tracking also improves the system precision and robustness for motion tracking with implicit 3-D information [137]. Note that since the silhouette is expected to translate from one image to the other, only motions of rigid targets can be handled in this approach.

Contour tracking, on the other hand, can properly process the nonrigid targets by contour evolution. Once some parts of the target in the previous image are overlapped with the target in the consecutive image, the target silhouette in the previous image can be used as the initial contour for the next image. This contour evolution is based on two types of models, namely, the state-space model and the active contour model.

The state-space model is a kind of probabilistic model that describes the probability dependence between the latent state and observed measurement. Thus, it can also be called probabilistic tracking. According to the state-space model, targets are tracked considering their state, which can be defined by the shape and the motion parameters. The state is updated by maximizing the posterior probability of the target contour. As a result, the distance between the contour and the observed outline would be minimized. Then, the Kalman filter [138] and the particle filter [139] are employed to update the state. Many enhancements of the tracking methods using the state-space models have also been proposed [140]–[142], which enables this approach to have capabilities of multiple target tracking, occlusion, and pose variation handling.

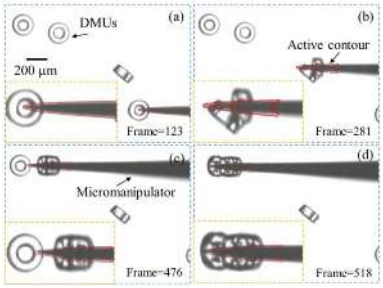


Fig. 14. Contour tracking results in red with the occlusion situation for the moving target (tip of the micromanipulator). Reproduced from [143] with the permission from IOP Publishing.

The active contour model evolves the contour shape and motion to the target outline by minimizing the contour energy using gradient descent or energy distribution. With the initial contour from the previous image in the image sequence, the contour evolution function of the current image can be expressed as

$$\begin{aligned} E &= E_{\text{in}}(\boldsymbol{c}) + E_{\text{out}}(\boldsymbol{c}) \\ &= \int_{\text{in}(\boldsymbol{c})} |I(x,y) - c_i|^2 dxdy + \int_{\text{out}(\boldsymbol{c})} |I(x,y) - c_o|^2 dxdy \end{aligned} \tag{16}$$

where $\boldsymbol{c}$ is the contour of the target, $E_{\text{in}}(\boldsymbol{c})$ stands for the internal energy that inside the contour $\boldsymbol{c}$, $E_{\text{out}}(\boldsymbol{c})$ denotes the external energy outside the contour $\boldsymbol{c}$, $I(x, y)$ represents the pixel value corresponding to image coordinates (x, y), and c_o and c_i are average pixel values outside and inside the contour, respectively. Thus, the evolution contour is driven to the boundary of the target in the current image by the internal and external energy forces [144]. An alternative to the explicit contour representation $\boldsymbol{c}$ is the implicit contour representation level set $\phi(x, y)$, which can be defined by

$$\phi(x,y) = \begin{cases} d(x,y,\boldsymbol{c}) & \text{if } (x,y) \text{ is outside contour } \boldsymbol{c} \\ -d(x,y,\boldsymbol{c}), & \text{if } (x,y) \text{ is inside contour } \boldsymbol{c} \end{cases} \tag{17}$$

where d is the shortest distance between coordinate (x, y) to the contour $\boldsymbol{c}$. In terms of (17), the explicit representation can be transformed into the implicit representation. Thus, the contour is implicitly encoded as $\phi(x, y) = 0$ on a spatial grid in the level set representation. The contour evolution is then driven by updating each grid value based on the energy computation in (17), which updates a new zero crossing $\phi'(x, y) = 0$ and the new contour in the implicit form. A valuable property of the level set method is the flexibility of representing the shape and topology changes of the target. Consequently, the silhouette of the nonrigid topology-varying target can be precisely computed from one image to another in the image sequence. Multiple improvements based on active contour models have also been developed, which can greatly enhance the accuracy, efficiency, and robustness [143], [145], [146]. After the target silhouette is matched in the image sequence, the optical flow estimation can be applied, and the

silhouette motion is confirmed by the dominant flow according to the calculation.

Silhouette-based tracking is useful when the target contour is changing, or motions of the entire region of the targets are required. Silhouette-based methods provide an accurate and flexible shape description for the moving targets, which leads to a great advantage of dealing with object deformation, split, and merge. Occlusion handling ability is another advantage of the silhouette-based methods, as shown in Fig. 14. These merits contribute to the practicalities of the silhouette-based technique [147]–[149].

III. Characteristics of Computer Microvision-Based Precision Motion Measurement

Aiming at different performances, intensive studies about motion measurement at microscale/nanoscale have been conducted. Many works have reached outstanding achievements in certain aspects, which shows the strong capabilities of vision-based microscopic systems for micro-engineering/nanoengineering. Nevertheless, few motion measurement systems can meet all expectations according to resolution, range, efficiency, and accuracy. In this section, the capabilities of motion measurement systems are reviewed and summarized. To provide a comprehensive perspective, measurement challenges, as well as current solutions, are also discussed.

A. Resolution Versus Range

In practical circumstances, one crucial concern for the vision-based motion measurement systems is their accurate measurement ability. Resolution is an important metric for any kind of measuring instrument, which stands for the smallest change that can be measured. It is noteworthy that the resolution that we discuss in this section is the motion measurement resolution, which is different from the optical resolution. Optical resolution is known as the shortest distance between two points on a specimen that can still be distinguished by the sensor as separate entities. The optical resolution depends on the wavelength and numerical aperture, as listed in (1), which is normally around 200 *nm*. The motion measurement resolution, however, refers to the smallest motion to which a change can be theoretically detected, which is higher than the optical resolution. As illustrated in Fig. 15, suppose that there are two parallel lines with a distance smaller than the optical resolution; due to the optical diffraction, these two lines cannot be recognized by using most of the conventional OMs. After imaged by the microscope, the real image of these two lines may deform into a small region. Although the two lines cannot be recognized, a small movement of the region (less than the distance between the two lines) is detectable by using image processing algorithms. Hence, with a proper motion tracking technique, the motion resolution can also reach nanoscale even based on the conventional optical microvision system. The motion measurement resolution can be formulated as

$$R_m = \frac{S}{M} R_{\text{agr}} \tag{18}$$

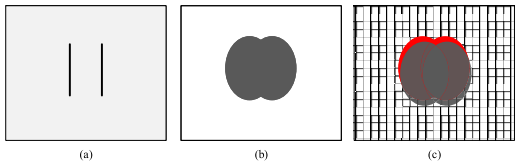


Fig. 15. Explanation for the optical resolution and motion measurement resolution of an optical microscopic vision system. (a) Two lines with the distance of dozens of nanometers. (b) After imaging by a common OM, the two lines become a line region (the dark area). (c) Line region has a small movement (the red and dark areas represent the line region before and after movement, respectively).

where R_{agr} represents the resolution of the image processing algorithm, S is the pixel size of the imaging sensor, and M stands for the total magnification ratio of the imaging system.

Range, on the other hand, represents the amount or extent of a value that can be measured. In the case of the microvision systems, the measurement range can also refer to as the effective FOV. The larger the measurement range is, the wider the observation space and acquired information can be in the motion measurement system. Thus, the measurement range is also an important performance for motion measurement. Since the range and resolution are a pair of contradictory indicators, there is often a tradeoff between achieving a large range and high resolution. As the scales and environments of different motion measurement tasks vary, only using resolution or range for performance evaluation might be one-sided. Besides, the range-to-resolution ratio can evaluate the multi-scale measurement ability of a method. Therefore, in Fig. 16, we review and summarize different state-of-the-art motion measurement methods [31], [45], [77], [82], [105], [109], [113], [122], [123], [149]–[165] by the range-to-resolution ratio in terms of (18) and different motion measurement algorithms, which can more comprehensively investigate the motion measurement performances.

Although the required measurement resolution or range is specifically determined by the different microengineering/nanoengineering tasks, most works achieve the range-to-resolution ratio above 10^3, which can be regarded as a fundamental requirement for modern microvision motion measurement systems. Concretely, microvision systems are designed for large working range tasks in works [158], [162], which shows different ratios and further demonstrate that the range is an independent performance to resolution. The work [156] exploits optical imaging systems with telecentric lenses, indicating that the telecentric lens structure is unrelated to the range-to-resolution ratio. SEMs have been used in works [45], [151]–[155], [163], which have relatively a smaller measurement range but yield a fine resolution, due to the properties of the SEMs. Particularly, it can be seen that, with specially customized patterns as the tracking markers, the phase correlation methods [45], [113], [122], [123] can achieve ultrahigh precision, while the need for specifically designed markers may lessen this advantage.

In works [45], [77], [82], [113], [122], [155], [160], [161], [165], subpixel algorithms have been applied, which significantly improves the measurement accuracy and contributes to a high range-to-resolution ratio with a tradeoff of measurement

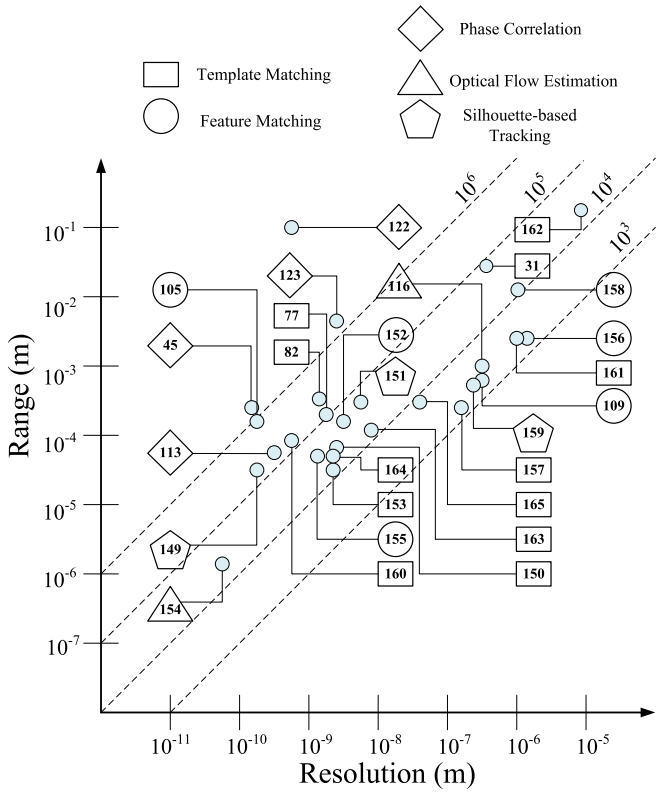


Fig. 16. Range-to-resolution performance of selected state-of-the-art works. Dashed line denotes the ratio of range and resolution. Numbers correspond to the references of this review.

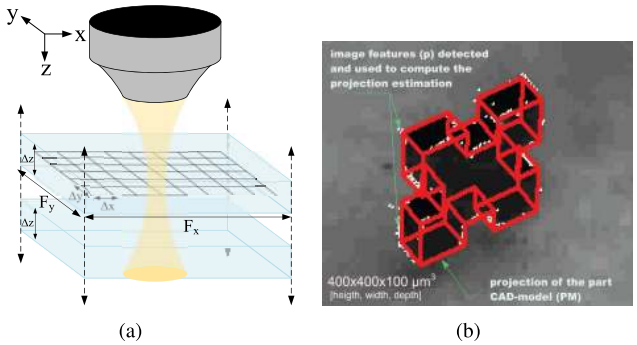


Fig. 17. Diagram of 3-D motion measurement. (a) Optical sections with different focal planes along z-axis, Δx and Δy denote the projective pixels, and the FOV is shown in blue with the range F_x, F_y, and depth of field Δz. (b) CAD model-based method for 3-D motion tracking, reproduced from [172] with the permission from SAGA Publishing.

efficiency. Since the smallest unit of an image is one pixel, displacements are normally measured at integer-pixel resolution. Therefore, the coarse-to-fine strategy can be adapted to different image processing algorithms for precision motion measurement: 1) preliminarily locate the target in the pixel accuracy; 2) select a small area of suitable size near the target location as the initial search area for fine locating; and 3) choose an appropriate subpixel approach to complete fine locating in subpixel accuracy. For the phase correlation method [45], [113], [122], [165], subpixel accuracy can be achieved by locating the peak of the CC, as introduced in Section II-C3. For other motion measurement algorithms, mainly three types of subpixel approaches can be employed as the fine locating process, namely, geometric approach, moment approach, and interpolation approach. The geometric approach [155], including the centroid method and gray-gravity method, utilizes the geometric characteristics of a target in the image to obtain subpixel measurement data. The moment approach constructs the image intensity into the moment feature descriptors [77], and it has a strong antinoise ability, as described in Section II-C2. The interpolation approach is also a popular choice. Many subpixel interpolation algorithms have been combined in template matching, including bilinear interpolation [166], bicubic spline interpolation [167], and B-spline interpolation [168].

In general, to achieve a high range-to-resolution ratio, the following approaches are recommended: 1) incorporate subpixel algorithms; 2) employ a large magnification; 3) use a high-resolution sensor for image acquisition; and 4) apply a small pixel size of the sensor to obtain a small displacement-to-pixel ratio. One should also note that though many subpixel algorithms can theoretically achieve very high accuracy and enhance the range-to-resolution ratio for the microvision systems, the realistic performance is still limited and related to the dynamic range of the acquired image sequence [169].

B. Multi-DOF Capability

From the measuring instrument point of view, the multi-DOF capability can remarkably benefit motion estimation and facilitate practical applications.

At present, lots of microvision-based systems have focused on 2-D measurement, where in-plane x/y translation and rotation are commonly involved. Only a limited number of works for 1-DOF visual measurement are proposed, which mainly exploring high-frequency rate in one dimension [170]. The 2-DOF measurement, on the other hand, has been widely studied [72], [77], [113], [171], where in-plane position information is sufficient for the task requirement. All the tracking techniques mentioned in Section II-C are able to easily perform motion measurement for in-plane displacement by locating the target in the image sequence. The 3-DOF measurement, however, is a bit more demanding that requires handling the in-plane rotation estimation. This can be achieved by adopting isometrics into the warp function in (5) for template matching [81], [83], rotation-invariant feature descriptors for feature matching [105], [155], or suitably designed patterns for phase correlation [122]. With the rotation estimation, the in-plant motions can be completely tracked and measured. The 3-DOF measurement can serve and satisfy a wide range of applications and tasks in the laboratory, including planar micromanipulation/nanomanipulation.

Nevertheless, a study [173] reported that the in-plant measurement suffers an error if an out-of-plane motion of the tracked target occurs. Moreover, expanding measuring dimensions toward 3-D is a trend that can fully realize the visual measuring potentials. Therefore, numerous efforts have been made for 3-D motion measurement based on conventional OMs and SEMs. The 4-DOF measurement can be regarded as an extension of the 3-DOF measurement along the depth axis or the z-axis. Consequently, depth from defocus information

can be utilized. Depth estimation is performed to calculate the distance between the target plane and the focal plane [174] and, thereby, achieve 4-DOF motion measurement (translation along the x-, y-, and z-axes, rotation around the z-axis). As illustrated in Fig. 17(a), the optical sectioning technique can also be applied to overcome the shallow depth of field by adjusting the focal plane along the z-axis and obtaining z-stack images, where the PSF can be incorporated to restore clear image stacks [175], [176]. The z-coordinate of the target can be located by precisely estimating the focus level in the image stacks [157], [177]. By recording the focal plane location in the z-stack image, the z-coordinate of the target can be confirmed, where the measurement resolution along the z-axis depends on the number of sections and the depth of field. The 6-DOF measurement introduces extra rotation estimation along the x- and y-axes. Based on the prior geometric knowledge of the target, the CAD model-based method is reliable to precisely track 6-DOF motions of rigid body targets [151], [172], as shown in Fig. 17(b). Without the CAD model, one can still estimate the rotations by the geometric approach [178] or the Fourier-based approach [45], [150]. Customized microvision-based systems are also developed for 3-D motion measurement as well. In [150], an interferometer-equipped OM system achieves real-time 6-DOF visual tracking from a single camera. Another alternative is stereo imaging, which includes adjusting the optical light path, such as LSCM-based approach [179], tilting the electron beam of SEMs [180], or deploying additional cameras to create the stereovision system for additional observed views [156]. Particularly, the stereovision system consists of a combination of multiple lenses with standalone light paths. It is a straightforward way to measure 3-D microscale targets in vertical and horizontal directions and avoid blind spots. A deep depth of focus of the stereo microvision systems can be achieved, and one can also adjust the magnification through a zoom lens [30], [181].

To sum up, the 3-D motion measurement under OMs and SEMs is expected to draw more research attention in the future, while the 2-D motion measurement is still sufficient and remains high efficiency in a wide range of applications at present.

C. Efficiency

Measurement efficiency is influenced by many aspects, such as hardware limitation, speed of matching algorithms, code efficiency, and environmental conditions. The measurement time consists of image acquisition time and matching processing time. In order to obtain high measurement accuracy, it is common that sophisticated algorithms are incorporated, which may sacrifice efficiency. For instance, the subpixel algorithms in the matching techniques involve a complex interpolation process, which can be time-consuming. There are still a large number of studies in recent years that remain off-line and postprocessing due to the heavy computational cost [105], [160], [182]. Nevertheless, real-time and fast-speed motion measurement is always preferred. High efficiency of visual feedback is also required for many online applications. For example, low tracking frequencies of measurement would significantly limit the response speed of visual servoing for micromanipulation. Therefore, pursuing a high-tracking frequency for motion measurement is meaningful and beneficial.

For conventional optical microvision systems, the image acquisition speed is normally high with the utilization of high-performance sensors. Thus, many efforts have focused on improving the efficiency of matching algorithms. Concretely, to accelerate template matching, many methods have been developed to efficiently search for the template in the image sequence. The searching strategies can be divided into the exhaustive searching method and the iterative searching method. For the exhaustive searching method, by approximating the template or the image [183], [184], the computation complexity decreases, and the searching efficiency can speed up. One can also adapt the coarse-to-fine procedure to find out the potential regions where the template might be [182]. Notably, heuristic approaches, such as particle swarm optimization [72] and artificial neural networks [185], have been developed rapidly in recent years and employed in the coarse search, which shows a promising performance to reduce the searching time. For the iterative searching method, Gauss–Newton and Levenberg–Marquardt algorithms outperform other gradient descent approximations in convergence speed and accuracy [186]. The inverse compositional algorithm [187] is strongly recommended as the numerical optimization method with the perfect balance between high efficiency and accuracy. High-speed feature matching has been investigated as well. A recent benchmark [188] reports that binary descriptors enjoy high efficiency, and deep descriptors [189], which exploit the deep learning technique, can further boost the speed for extracting target features. Among phase correlation-based methods, apart from decreasing the spectral components, as described in Section II-C3, the efforts have also been put into the subpixel motion extraction. By optimizing the initial upsampled estimate for the location of the phase correlation peak, computational time can be greatly reduced without sacrificing the measurement accuracy [116], [190]. For contour tracking, many works concentrate on fast contour evolution and smart initialization, which can be found in [145], [191], and [192].

Scanning-based microscopes, such as LSCMs and SEMs, have a nature bottleneck of comparatively low imaging efficiency. With the hardware innovation, nowadays, the state-of-the-art commercial LSCMs and SEMs can present above 20-Hz frame rates with a tradeoff of decreasing resolution and image quality. Sparse imaging can further increase the frame rate. In the work [193], using a smaller region of interest, image acquisition time can be significantly reduced, and the tracking frequency is improved. Nevertheless, the image acquisition time is still longer than the matching calculation time in practice. Hence, to maintain and cooperate a fast scanning speed, the measurement methods are required to have strong robustness to the bad image quality. Motion measurement techniques can be combined with image enhancement approaches, in order to obtain a high scanning speed. The work [194] proposed a real-time SIFT-based feature tracking for an LSCM, in which ten successive frames were coadded to

TABLE VII
PERFORMANCE COMPARISON OF DIFFERENT MICROVISION-BASED MOTION MEASUREMENTS

Microscopy	DOF	Matching technique	Resolution	Range	Efficiency	Example	Year
OM	1	Phase correlation	5 nm	168 μm	1390 Hz	[170]	2014
OM	2	Phase correlation	0.5 nm	221 μm	30 Hz	[113]	2016
OM	3	Template matching	3.33 nm	560 μm	300 Hz	[82]	2019
OM	3	Feature matching	0.227 nm	200 μm	2.6 Hz	[105]	2020
OM	3	Phase correlation	1 nm	110 mm	5 Hz	[122]	2020
OM	6	Feature matching	1.875 μm	41 μm	25.4 Hz	[156]	2016
OM	6	Silhouette-based	0.3 μm	570 μm	10 Hz	[159]	2016
SEM	3	Template matching	2.71 nm	42.24 μm	13 Hz	[153]	2012
SEM	3	Optical flow	0.075 nm	1.6 μm	10 Hz	[154]	2014
SEM	4	Template matching	0.27 μm	200 μm	10.7 Hz	[174]	2016
SEM	5	Phase correlation	0.2 nm	396 μm	2.2 Hz	[45]	2018
SEM	6	Silhouette-based	11.1 nm	111 μm	4 Hz	[151]	2009

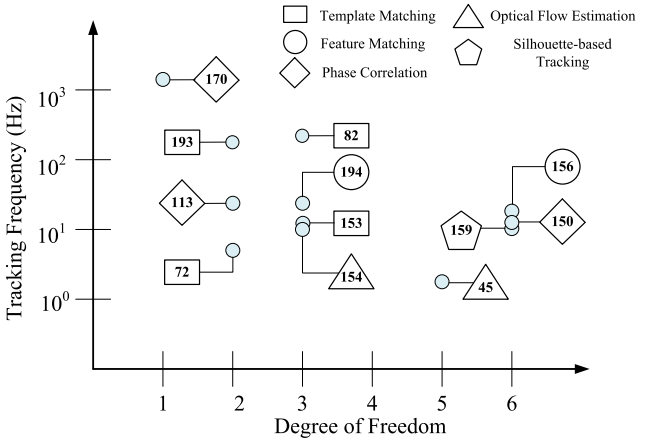


Fig. 18. Tracking frequency of different efficient motion measurement methods in terms of the DOF. Numbers correspond to the references of this review.

improve the image quality. In [195], a state-space model was applied with PSF to efficiently track the fluorescent particles under LSCMs. The work [153] exploited morphological image processing to turn the SEM images into binary images for template matching and achieves visual servoing. In [154], robust visual tracking was performed under a conventional SEM with real-time image denoising and drift compensation process. Fig. 18 illustrates the state-of-the-art works in terms of the DOF, which reported high efficiency based on different motion measurement techniques.

Regardless of the computational cost, real-time motion measurement is always preferable. To achieve this purpose, the following procedures can be taken: decreasing the number of tracking targets, lowering the imaging resolution, setting a small template and region of interest, adopting high-performance sensors and microscopes, and optimizing the matching algorithms.

D. Error Source Analysis

Measurement errors of vision-based systems are caused by multiple sources, which cannot be totally eliminated. Thus, efforts have been employed to minimize the measurement errors as possible. For OMs, inaccurate projective model, optical distortion, matching techniques, pose estimation model, marker manufacturing error, system resolution, system installation error, pixel manufacturing error, data transmission, or nonuniform refractive medium can result in the measurement error. Many effective calibration methods have been proposed with certain patterns for microvision systems to obtain an accurate projective model, as to introduce in Section II-B. For other procedures, in work [196], errors of in-plane 3-DOF motion measurement were evaluated under different pose estimation models, and the result demonstrated that the tilt angle between the image plane and motion plane is insignificant to measurement results if the angle is under a certain value. The study [31] conducted the error analysis considering out-of-plane motions for 2-D measurement and the thermal influence of imaging sensors and suggested using telecentric lenses and adopting a preheating process to achieve better performances. In [123], the measurement accuracy was limited because of the manufacturing error of tracking markers, which is called for high-precision markers to be used.

Apart from the spatial distortion, which can be alleviated by the calibration process, SEMs have some additional measurement error sources due to the characteristics, particularly image noises and drifts. Image noises of SEMs are generated by primary emission, secondary emission, scintillator, photocathode, and photomultiplier [197], and they can be quantified by the image SNR and, thereby, monitor the image quality. According to the requirement of a specific task, image noises can be controllable to a certain level by optimizing the parameters of the SEM. The study [198] developed a monitoring method and reported that a relatively long scan time, low magnification, in-focus target, long stabilization preparation time could contribute to a high SNR and low image noises. Thus, monitoring the SNR of images can ensure the image quality is sufficient for accomplishing the vision-based task. Applying filters on the images [199], or improving the circuit control [200], can also benefit the noise reduction. Another common issue for SEMs is image drifts, which may cause a temporal measurement error with around 10 nm/min. There are several factors that may cause the drifts: 1) charged nonconductive materials in the FOV deflect the electron beam of SEMs; 2) temperature variations of the system components; 3) mechanical vibration; and 4) fluctuations of the sample-stage control voltages. This drift effect is inevitable in SEMs, especially noticeable under high magnification. In order to correct the nonlinear time-varying drifts, numerous methods have been proposed. The main idea of these methods is to estimate the drifts in the image sequence. By comparing

adjacent pairs of images, the drift can be measured using matching techniques. Concretely, template matching [201], feature matching [46], and phase correlation [202] can be employed in this process. To model the drift-time relationship, the drift velocity at pixel positions in the FOV is commonly fitted with the B-spline function in time. Then, the removal of drifts can be completed with drift correction strategies, such as direct compensation [46], Kalman filter estimation [203], and image composition [202].

In terms of DOF, resolution, range, and efficiency, Table VII lists a performance comparison of different microvision-based motion measurements from the selected works. It can be seen that different measurement systems focus on different aspects, and the research direction is mostly application-oriented due to practical limitations. So far, one single technique is hardly found to surpass others in all aspects. Nevertheless, research efforts have been pushing the motion measurement potentials toward more DOFs, higher resolution, larger range, and faster speed in order to be capable of the growing demands of microengineering/nanoengineering.

IV. Cutting-Edge Applications

Precision motion measurement by computer microvision has been widely applied in microengineering/nanoengineering, which is multidisciplinary that has drawn intense efforts in the last two decades. This section introduces the recent advances of applications enabled by computer microvision-based methods, including the precision displacement measurement of compliant mechanisms, automated micromanipulation, force sensing, full-field strain, and stress measurement.

A. Precision Displacement Measurement of Compliant Mechanisms

Compliant mechanisms have been commonly used as the basic mechanical structures for micromanipulation systems and microelectromechanical systems (MEMSs) [207], [208]. This kind of flexure-hinge-based mechanism driven by piezoelectric actuators enjoys the features that are suitable for nanorobotic systems, such as high precision, fast response, long life, and compact structure, which also raises the demand for precise displacement measurement. Since compliant mechanisms are usually small and compact, they normally require the extra dedicated design for the installation of traditional instruments, such as the laser interferometer, the capacitive displacement sensor, the autocollimator, the grating ruler, and the strain gauge. Therefore, the precision displacement measurement of compliant mechanisms is essential for microvision-based systems, which is expected to outperform and take the place of the traditional instruments. With the advantages of low-cost, noncontact, and multi-DOF capabilities, computer microvision-based methods have served a vital role in detecting and measuring displacements in precision engineering, especially for the compliant mechanism-based platforms, as shown in Fig. 19.

Wu *et al.* [77], [182] proposed displacement measurement methods for compliant mechanisms using the exhaustive-searching-based template matching and moment-based feature matching, respectively. The experimental results showed that the microvision-based methods process the same accuracy and stability as the laser interferometer, while the microvision system is more flexible, operable, and simpler in practice. Chen *et al.* [205] measured the displacements of a microgripper by the optical flow method under the OM. The optical ow-based motion analysis was presented, which validated the grasping performance of the microgripper. Zhang *et al.* [160] applied a rotation-invariance template matching for displacement measurement of compliant positioning stage. The precision, stability, and adaptability of the computer microvision system were also proven. Li *et al.* [81]–[83] presented iterative-based template matching for precise displacement measurement of 3-DOF compliant nanopositioners. The efficiency of template matching methods was optimized, and the results demonstrated that the microvision-based method could compete with capacitive displacement sensors in accuracy, repeatability, and measurement range. Guelpa *et al.* developed a phase correlation method based on the stripe pattern, extended the method from 1-D [170] to 2-D [113], and then applied it to SEM [206]. Andre *et al.* [122] achieved precision displacement measurement of the complaint positioning stage in a long-range by attaching and tracking an encoded marker on the target surface.

B. Automated Micromanipulation

According to different robotic systems, contact-based micromanipulation is conducted by end-effectors such as microgrippers, while noncontact-based micromanipulation uses remote physical fields, including optical, magnetic, electric, acoustic, and fluid fields. In the past, micromanipulation was completed manually by human operators under microscopes, which is costly to train and skill-dependent. The computer microvision techniques empower robotic systems to automatically, individually, and delicately control and handle the microtarget/nanotarget, thereby creating a practical and feasible tool to shape the microworld. Fig. 20 shows the representative works in the last 20 years, which can be divided into two categories in terms of the manipulated targets and the application purposes: industrial use and clinical use.

The industrial use of automated micromanipulation, including assembly of MEMS and microfabrication, has significant potentials to improve nanotechnology and boost the performance of MEMS with a more compact and exquisite structure. Via vision-based motion measurement for providing the target position and pose information as feedback, visual guidance and full closed-loop control for manipulation can be achieved with robotic systems. In the micromanipulation systems, positioners play a key role as the basic component for carrying the manipulated targets in the robotics systems. Hence, visual servoing of the positioning stage in automated micromanipulation/nanomanipulation is essential. Ralis *et al.* [209] proposed a framework of high-precision visual servoing for positioners. The optical flow method and depth estimation were used for motion measurement, and the positioning accuracy was up to 2.2 μm by servomotors. To increase the positioning accuracy, piezoelectric positioners have been introduced to the manipulation systems [221], [222]. Marturi *et al.* [45]

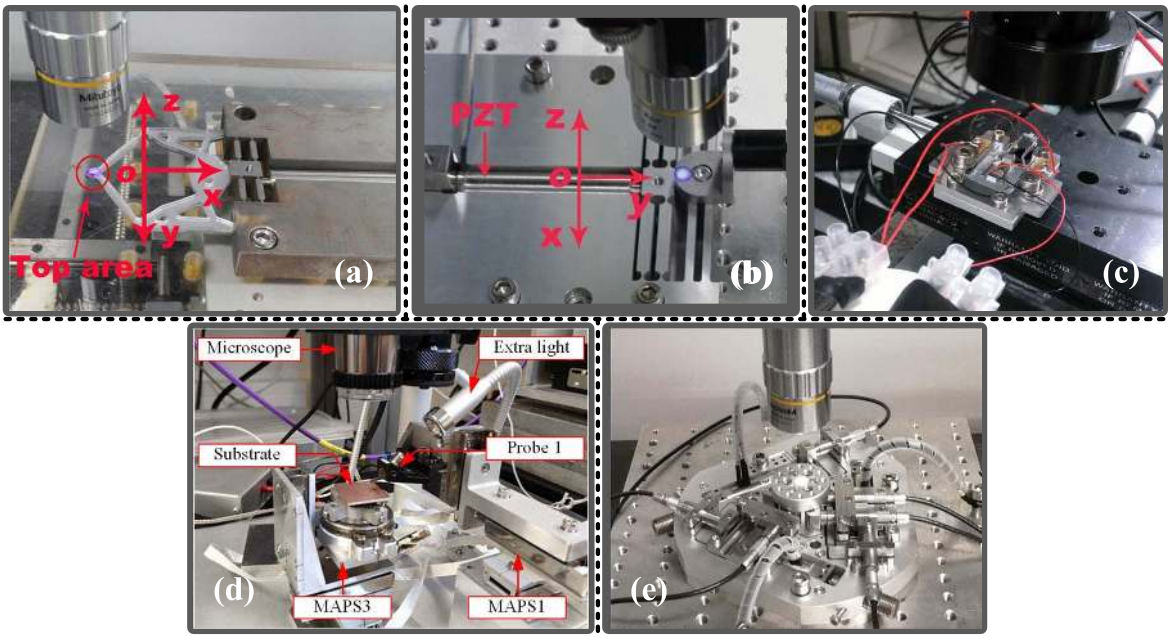


Fig. 19. Recent advances of high-precision displacement measurement by computer microvision for compliant mechanism-based systems. (a) Micromotion detection for the displacement inverters. Panel adapted from [182] with the permission from Elsevier. (b) Real-time displacement tracking of the compliant nanopositioner. Panel adapted from [81] with the permission from IEEE. (c) Motion tracking of the positioning stage and robotic probe. Panel adapted from [204] with the permission from IEEE. (d) Motion analysis of the compliant microgripper using the microvision-based method. Panel adapted from [205] with the permission from AIP Publishing. (e) Displacement measurement under SEM by phase correlation with a complaint structure pattern. Panel adapted from [206] with the permission from IEEE. (f) Large-range high-precision displacement measurement of positioning stages using an encoded marker. Panel adapted from [122] with the permission from IEEE.

presented an automatic 3-D nanopositioning scheme for piezoelectric nanopositioners with an accuracy up to 0.04 μm. The 5-DOF marker-free motion measurement under SEM was achieved by the phase correlation-based method. Li *et al.* [223] developed a high-speed visual servo control of compliant mechanisms for nanopositioning. The motion tracking was demonstrated with high tracking frequency under OM.

Micropart assembly is a direct way to manufacture MEMS by constructing separate microparts. With precise motion tracking and manipulation capabilities, the automation of assembly of microdevice/nanodevice can be enabled. Ren *et al.* [56] implemented visual servoing for MEMS microassembly under OM using template matching. The 2-D micrograsping task was facilitated with a 6-DOF microassembly robot. Tamadazte *et al.* [172] delivered a silhouette-based tracking for 3-D MEMS microassembly based on the CAD model of the microparts. Real-time performance and practicality were validated under SEMs. Chen *et al.* [156] developed a telecentric stereo microvision system for optical fiber assembly. The 3-D optical fiber alignment was conducted with five DOFs, and subpixel accuracy was achieved in the motion measurement. He *et al.* [213] presented a nanoalignment system based on compliant mechanism and computer microvision. An improved template matching method was designed and enabled accurate alignment for advanced chip packaging.

The ability to conduct pick-and-place tasks is also important for micromanipulation/nanomanipulation systems since prefabricated structure blocks are often required to be constructed into the MEMS devices. Mazerolle *et al.* [210] brought up a concept of Lab-in-SEM and provided a framework for nanoscale target manipulation. Semiautomation was presented for the manipulation of carbon nanotubes (CNTs). Ye *et al.* [211] automatically picked and placed the single nanowires on the MEMS device with vision-based motion control. The motions of the end-effector tip and a stationary feature against the image drift were measured by template matching. Zimmermann *et al.* [212] enabled automation of pick-and-place of colloidal particles under SEM. Both feature matching and template matching techniques were used to track the motion of microgrippers and small particles with 1-μm diameter.

Tasks of probing allow material characterization for understanding the properties and mechanisms of certain structures. Ru *et al.* [224] presented automated nanoprobing of nanowires under SEM. Both probes and nanowires were visually recognized and tracked by template matching. Gong *et al.* [154] proposed SEM-based visual servoing for nanoprobing of nanostructures by optical flow. The automated nanoprobing system possessed a large working space with high precision. Different works of vision-based industry-oriented micromanipulation are summarized in Table VIII.

Automated micromanipulation of biomicroparticles, including small organisms, cells, and organelles, benefits life science by system cell engineering and becomes popular in recent years. The nonrigid and time-varying properties of biomicroparticles are also challenging to be measured and manipulated. Biomicroparticles also require a proper environment to conduct manipulation, which promoted the development of ESEM. On the other hand, high requirements provide a good opportunity to verify the performances of motion measurement and robotic systems.

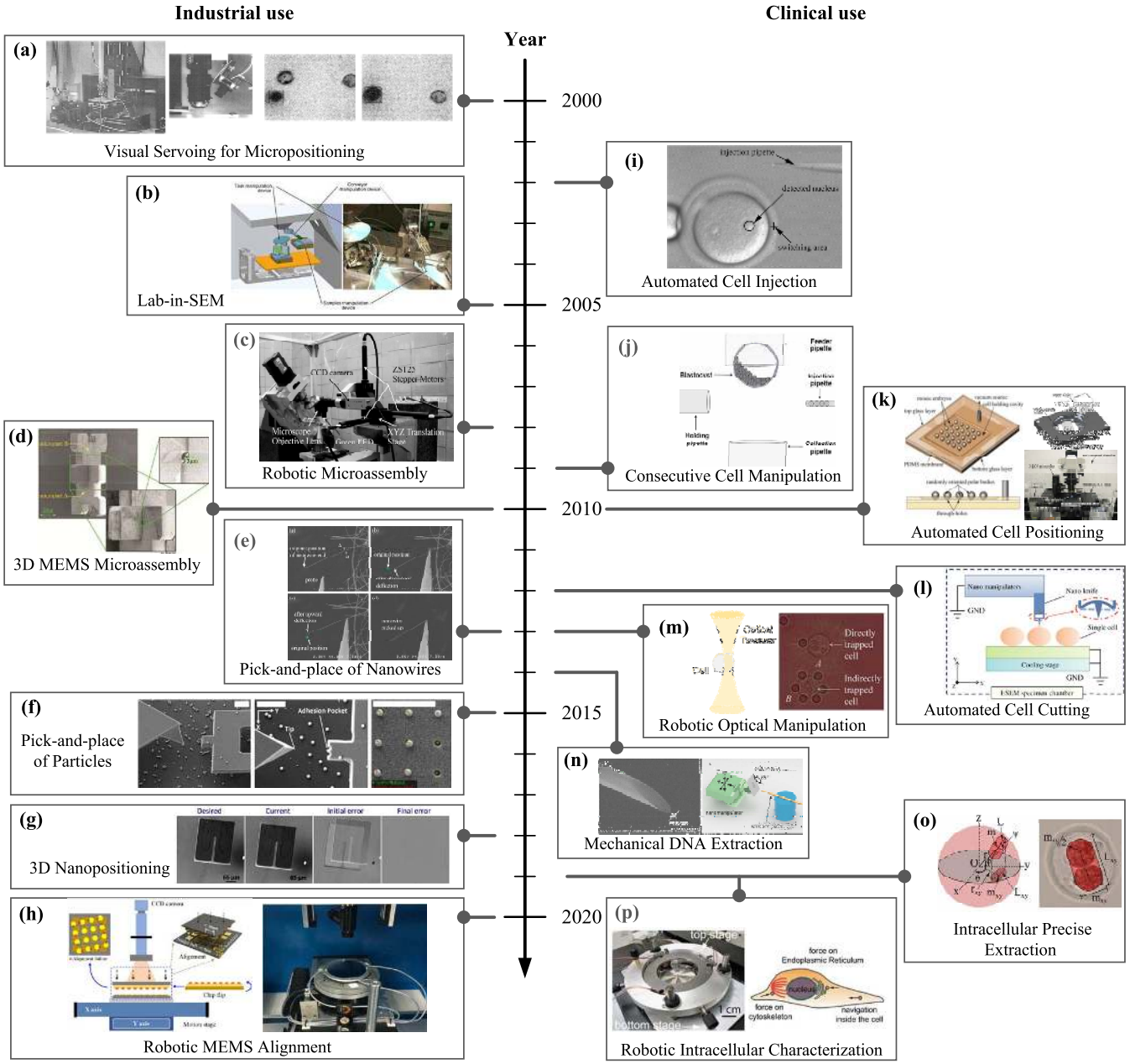


Fig. 20. Representative works of vision-based automated micromanipulation. (a) Visual servoing for micropositioning under OM. Panel adapted from [209] with the permission from IEEE. (b) Lab-in-SEM concept. Panel adapted from [210] with the permission from Elsevier. (c) Automatic microassembly for MEMS. Panel adapted from [56] with the permission from IEEE. (d) CAD-based real-time 3-D tracking for microassambly under SEM. Panel adapted from [172] with the permission from SAGA Publishing. (e) Automated pick-and-place of nanowires. Panel adapted from [211] with the permission from IEEE. (f) Robotic manipulation of particles using vision-based control. Panel adapted from [212] with the permission from IEEE. (g) 3-D image-guided nanopositioning under SEM. Panel adapted from [45] with the permission from IEEE. (h) Robotic MEMS alignment at the nanoscale. Panel adapted from [213] with the permission from IEEE. (i) Automated cell injections with a 100% success rate. Panel adapted from [214] with the permission from SAGA publishing. (j) Consecutive cell manipulation with multiple micromanipulators. Panel adapted from [215] with the permission from IEEE. (k) High-throughput cell position and orientation control under OM. Panel adapted from [216] with the permission from IEEE. (l) Precision cell cutting under ESEM. Panel adapted from [217] with the permission from IEEE. (m) Robotic manipulation by optical tweezers. Panel adapted from [218] with the permission from IEEE. (n) Intracellular DNA extraction. Panel adapted from [219] with the permission from John Wiley and Sons. (o) Real-time 3-D rotation tracking inside the cell. Panel adapted from [159] with the permission from IEEE. (p) 3-D intracellular navigation using a magnetic tweezers system. Panel adapted from [220] with the permission from the American Association for the Advancement of Science.

Position and orientation controls of biomicroparticles are significant in biomicromanipulation. Liu *et al.* [216] developed a high-throughput cell pose adjustment method based on a motorized stage under OM. The embryos were immobilized by a microfabricated device, and 3-DOF closed-loop control was performed with visual feedback. Except for mechanical tools, noncontact automated manipulation has also been investigated. Xie *et al.* [225] introduced an automated noncontact-based cell manipulation method by laser optical tweezers. The optical flow was employed to extract the motion field of cells, and 3-DOF motion control was achieved. Chowdhury *et al.* [218] indirectly trapped cells by optical tweezers. Dielectric beads

TABLE VIII
SUMMARY OF AUTOMATED INDUSTRY-ORIENTED MICROMANIPULATION

Task	Year	Techniques	Features	References
Stage micropositioning	2000	OM, optical flow	Robust positioning	[209]
Stage nanopositioning	2018	SEM, phase correlation	3D visual servoing	[45]
	2020	OM, template matching	High-speed visual servoing	[213]
Microassembly	2008	OM, template matching	High-efficiency micrograsping	[56]
	2010	SEM, silhouette-based	3D real-time microassembly	[172]
	2016	OM, feature matching	Telecentric stereo microvision	[156]
	2020	OM, template matching	Large stroke, high accuracy	[213]
Pick-and-place CNTs	2005	SEM, template matching, feature matching	Concept demonstration	[210]
Pick-and-place nanowires	2013	SEM, template matching	High precision	[211]
Pick-and-place particles	2015	SEM, template matching, feature matching	High repeatability with excellent precision	[212]
Nanoprobing	2010	SEM, template matching	High success rate	[224]
	2014	SEM, optical flow	High efficiency	[154]

TABLE IX
SUMMARY OF AUTOMATED BIOMEDICAL-ORIENTED MICROMANIPULATION

Task	Year	Techniques	Features	References
Cell positioning	2010	OM, template matching	High throughput	[216]
	2013	OM, silhouette-based	Non-contact-based	[218]
	2016	OM, optical flow	High accuracy	[225]
Cell injection	2002	OM, template matching	100% success rate	[214]
	2009	OM, template matching	High efficiency	[215]
	2013	OM, silhouette-based	High robustness	[226]
Organism injection	2017	OM, template matching	100% target survival rate	[227]
DNA extraction	2014	SEM, feature matching	High accuracy	[219]
Intracellular extraction	2019	OM, silhouette-based	Minimize the damage to cells	[228]
Cell characterization	2010	OM, silhouette-based	Non-invasive, topological changes	[229]
	2013	LSCM, optical flow	Robustness, complex backgrounds	[134]
	2014	OM, silhouette-based, template matching	3D micromanipulation	[230]
	2019	LSCM, template matching	Intracellular characterization	[220]

were used as grippers to push or hold the cells forward, where the photodamage from the laser beam can be avoided.

Precision robotic surgery at a small scale is a large group of biomedical micromanipulation tasks that have been widely studied to understand biological phenomena and discover new treatments for human diseases. Precision robotic surgery involves a few primary operations, such as injection, extraction, and cutting. The technologies that are able to transport materials in and out of cells play a crucial role in cell biology. Sun and Nelson [214] automated the cell injections by tracking the tip of the micropipette, which was mounted on a 3-DOF high-precision microrobot. The autofocus algorithm was utilized in the automation process under OM. Mattos *et al.* [215] designed a four-pipette robotic system for consecutive blastocyst microinjection. SSD template matching was employed to track the pipettes, and system efficiency was improved by multiple micromanipulators. Karimirad *et al.* [226] presented an active contour-based tracking method for deformable cells during the microinjection. It also provided a precise description of the deformable cells. Liu *et al.* [147] proposed a fast autolocating method for end-effector tips in 3-D under OM. The active contour was also experimentally demonstrated as the most effective method to track the end-effector tip compared to others. Zhuang *et al.* [227] developed a visual servo microinjection system for small organisms. The heart of zebrafish larva was targeted for injection, and experiments showed a high success rate and survival rate of zebrafish. Chen *et al.* [219] performed the SEM-based visual servoing for intracellular DNA extraction. The cells were mounted on the nanopositioner, and the extraction was conducted by a nanospatula. Wong and Mills [228] proposed a blastomere extraction method by silhouette-based tracking. Using the geometrical model of blastomeres [159], the 3-D rotation tracking in the cell was achieved inside the embryos in real time, and the damage from cell surgery was minimized by optimizing the cell surgery zone, where the zona breaching was performed by the laser drilling. Shen *et al.* [217] conducted automatic cell cutting with a nanorobotic manipulator under ESEM. Real-time images were utilized as the visual feedback, and a yeast cell was put by a tungsten probe inside ESEM.

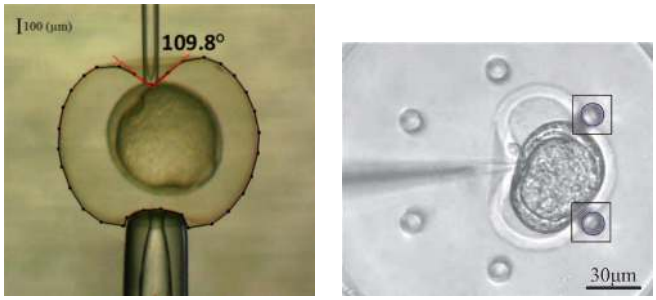
Cell characterization provides a quantitative method to study the behaviors and health conditions of cells from the biological and mechanical properties. Automated contact-based cell characterization can be achieved with robotic systems, which is one of the latest development trends for biomedical-oriented micromanipulation. Liu *et al.* [230] investigated the dye transfer among adherent cells and characterized cell communication by the automated robotic system. 3-D motions of the micromanipulator were tracked by the active contour, and cells were located by template matching, respectively. Wang *et al.* [220] presented 3-D intracellular manipulation using magnetic tweezers. LSCM was used to provide 3-D visual feedback, and a bead was tracked and navigated inside the human bladder cancer cell, which can probe the subcellular structures and organelles and contribute to intracellular characterization and disease diagnosis. Noncontact and noninvasive cell characterization and monitoring have also been embedded with motion measurement techniques. Wong *et al.* [229] used the state-space model to track the early-stage embryos with the topology-varying characteristic. The growth of embryos

was estimated and predicted by accurately tracking the blastomeres. Guo *et al.* [134] performed red blood cell tracking by optical flow, and multiple targets were tracked simultaneously under LSCM. Table IX summarizes and compares vision-based biomedical-oriented robotic systems with different tasks.

C. Force Sensing

With prior knowledge of deformable materials, forces can be estimated by visually tracking structural deformations. Greminger and Nelson [231] demonstrated a vision-based force measurement method by template matching in 2004. The Dirichlet to Neumann map was established, and then, the force distribution applied to linearly elastic targets was recovered through the displacement field, which was measured from the image sequence. The experiments proved that the target boundary information was sufficient to estimate the force, and a resolution of 2.8 nN was achieved in real time. In 2006, Anis *et al.* [232] determined the relationship between displacements and forces by finite-element modeling, where template matching was also employed to obtain the displacements. In 2007, Liu *et al.* [233] proposed an indirect vision-based force measurement method by switching the measured targets to objects with known elastic properties. The obtained forces were then converted back to the target that interacts with the objects. Karimirad *et al.* [234] developed a vision-based force measurement by the active contour method in 2014, as presented in Fig. 21(a). The motions were tracked by the active contour first, and the elastic nonlinearity of the target was accurately modeled by the artificial neural network to recover the applied forces. Cappelleri *et al.* [235] designed a force sensor for microrobotics by combining microvision-based tracking technique and the compliant mechanism. It further indicates that, with known force-deflection characteristics, microvision-based force sensing can achieve excellent performances.

Force sensing can improve the performance of robotic systems by preventing damages. Since a microvision system has all the configurations that force sensing needs, microvision-based force sensing can be naturally integrated into vision-based robotic systems. Lu *et al.* [236] surveyed different strategies of microforce control and the integration of force, vision, and position information for micromanipulation. Ferreira *et al.* [237] developed a vision-based position/force control for MEMS microassembly. By modeling the deformations of the microparts, force feedback was acquired for visual servoing. Liu *et al.* [238] performed force-controlled microinjection and micrograsping by the indirect force estimation, as shown in Fig. 21(b). A resolution of 3.7 nN was obtained through subpixel template matching, and a tracking frequency of 15 Hz was achieved in the closed-loop control of gripping forces. Xie *et al.* [239] proposed a microvision-based force control framework for robotic microinjection, which contained two control loops. With the direct force feedback from the manipulated cells, the penetration force during the injection can be regulated to follow the desire force trajectory. In addition to cell surgery, force control can be also applied to



(a) (b)

Fig. 21. Force estimation based on motion measurement. (a) Microvision-based force measurement by the active contour. Reproduced from [234] with the permission from Elsevier. (b) Indirect cell force measurement by matching the holders. Reproduced from [238] with the permission from SAGA Publishing.

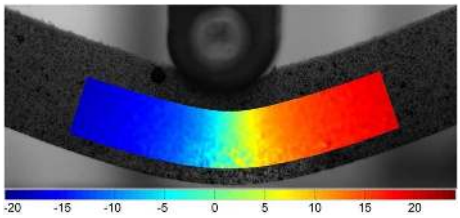


Fig. 22. Motion measurement for strain estimation. The color distribution represents the rotation angle of the deformable structure. Reproduced from [242] with the permission from IOP Publishing.

cell positioning. Zhao *et al.* [240] proposed a force-controlled cell reorientation method with rotation force. The deformation and rotation angle of the cell by each poke were visually tracked, and the relationship between the rotation angle and the operating force was precisely quantified with the calibrated mechanical properties of the manipulated cell.

D. Full Field Strain and Stress Measurement

By tracking the same points of interest of deformable surface in the images under microscopes, different stages of material and structure deformation can be extracted and then converted into full-field displacement and strain results on the target area, which is valuable for material and microstructure stress analysis. Microvision-based methods can also visualize the deformation and monitor the structure's health. Therefore, motion measurement techniques have been employed for estimating the microscale surface deformations and strain in the field of solid mechanics. For example, Pan *et al.* [74] summarized different template matching methods for 2-D strain field estimation in solid mechanics. Sutton *et al.* [241] investigated SEM-based strain measurement by template matching. A novel drift and spatial distortion correction method was proposed for SEM images, and elastic and elastic-plastic deformations were accurately measured. Wu *et al.* [242] proposed a full-field strain measurement method considering undergo arbitrary rotations based on microvision. The feature matching technique and the subpixel algorithm were employed, and the method was robust under nonuniform rotations. Dong *et al.* [243] developed a duel-FOV telecentric imaging system for strain measurement. The strain resolution was greatly enhanced to 3.55 $\mu\varepsilon$. Gao *et al.* [244] proposed a fast-speed 3-D strain field

measurement based on stereoimaging. The inverse compositional Gauss–Newton algorithm was also reported to possess high efficiency and high accuracy in template matching. Zhang and He [245] investigated the stress intensity of the specimen by template matching, and the interpolation approach was also applied in the coarse-to-fine strategy. Fig. 22 presents an example of microvision-based strain field measurement, where the deform is fully visualized.

V. Conclusion

The development of microscopy and tracking techniques has greatly promoted computer microvision-based motion measurement, which is still a relatively young field but has already become a powerful tool for understanding and engineering the microworld. The goal of this review is to guide other researchers who are interested in computer microvision-based motion measurement but without a strong background and to provide a useful reference with recent advances and performance comparisons for experienced scholars.

In this review, practical and applicable microvision systems and their general principles are first reviewed. The advantages and limitations of different configurations are introduced. Different motion measurement algorithms are systematically summarized and categorized, namely, template matching, feature matching, phase correlation, optical flow, and silhouette-based methods. Emphases are specifically placed on the performance of state-of-the-art motion measurement methods with regard to measurement resolution, range, DOF, efficiency, and error reduction, where their capabilities and challenges are compared and discussed. A wide range of applications and research efforts established on computer microvision-based motion measurement during the last two decades is showcased and analyzed in the research field of microengineering/nanoengineering. Many remarkable works are presented, and their methodologies and features are highlighted.

Despite the much attention that has been drawn and achievements that have been done, there remain a few issues and challenges that need to be further addressed in the future.

1) Computer microvision-based force sensing can be simultaneously obtained and combined with the position feedback for robotic systems, which can minimize the risks of tools and sample breakage in micromanipulation/nanomanipulation. Thus, it is worthy of putting more effort into vision-based force control for micromanipulation/nanomanipulation.
2) Multiscale motion measurement is needed in cross-scale engineering, thereby raising the need for more large range-to-resolution measurement methods, which is a meaningful research direction.
3) Only a few works have provided accessible software or open-source codes of microvision-based measurement methods, which limits the practicality and accessibility. A sharing and educational environment is called in the research community, which can contribute to rapid development and boom this research field.
4) Many works were task-dependence and the corresponding motion measurement techniques were specifically designed. A uniform framework for general tasks is absent. This can be problematic to effectively select the appropriate motion measurement technique. Hence, the method generalization needs to be enhanced.
5) In order to fit in 3-D micromanipulation/nanomanipulation, which has been more and more proposed in recent years, motion measurement and pose estimation methods are expected to be more effectively extended to 3-D.

With the unique advantages of expansibility, programmability, and measurement capabilities, microvision-based systems are irreplaceable in the development microengineering/nanoengineering. Any improvement in either hardware innovation or algorithm aspect is welcome, and the research works of computer microvision-based motion measurement would be pushed forward in better measurement accuracy, range, DOF, speed, and robustness. Therefore, this article provides valuable viewpoints of microvision-based motion measurement to promote more talents to this field, awaited more exciting and transformative development to come.

Acknowledgment

The authors gratefully acknowledge the support agencies.

References

[1] R. Wang and X. Zhang, "Parameters optimization and experiment of a planar parallel 3-DOF nanopositioning system," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2388–2397, Mar. 2018.
[2] S. Fericean and R. Droxler, "New noncontacting inductive analog proximity and inductive linear displacement sensors for industrial automation," *IEEE Sensors J.*, vol. 7, no. 11, pp. 1538–1545, Nov. 2007.
[3] X. Li, W. Gao, H. Muto, Y. Shimizu, S. Ito, and S. Dian, "A six-degree-of-freedom surface encoder for precision positioning of a planar motion stage," *Precis. Eng.*, vol. 37, no. 3, pp. 771–781, Jul. 2013.
[4] L. Clark, B. Shirinzadeh, Y. Tian, and D. Oetomo, "Laser-based sensing, measurement, and misalignment control of coupled linear and angular motion for ultrahigh precision movement," *IEEE/ASME Trans. Mechatronics*, vol. 20, no. 1, pp. 84–92, Feb. 2015.
[5] L.-B. Xie, Z.-C. Qiu, and X.-M. Zhang, "Development of a 3-PRR precision tracking system with full closed-loop measurement and control," *Sensors*, vol. 19, no. 8, p. 1756, Apr. 2019.
[6] A. Rosenfeld, "Picture processing by computer," *ACM Comput. Surv.*, vol. 1, no. 3, pp. 147–176, 1969.
[7] J. B. Williams, *The Electronics Revolution: Inventing the Future*. Berlin, Germany: Springer, 2017.
[8] D. Kahng, "Silicon-silicon dioxide field induced surface devices," in *Proc. Solid State Device Res. Conf.*, Pittsburgh, PA, USA, Jun. 1960.
[9] W. S. Boyle and G. E. Smith, "Charge coupled semiconductor devices," *Bell Syst. Tech. J.*, vol. 49, no. 4, pp. 587–593, Apr. 1970.
[10] E. R. Fossum and D. B. Hondongwa, "A review of the pinned photodiode for CCD and CMOS image sensors," *IEEE J. Electron Devices Soc.*, vol. 2, no. 3, pp. 33–43, May 2014.
[11] B. Wijnen, E. E. Petersen, E. J. Hunt, and J. M. Pearce, "Free and open-source automated 3-D microscope," *J. Microsc.*, vol. 264, no. 2, pp. 238–246, Nov. 2016.
[12] F. Schmoeckel, S. Fahlbusch, J. Seyfried, A. Buerkle, and S. Fatikow, "Development of a microrobot-based micromanipulation cell in a scanning electron microscope (SEM)," in *Microrobotics Microassembly II*, vol. 4194. Bellingham, WA, USA: International Society for Optics and Photonics, 2000, pp. 129–140.
[13] M. Nakajima, T. Hirano, M. Kojima, N. Hisamoto, M. Homma, and T. Fukuda, "Direct nano-injection method by nanoprobe insertion based on E-SEM nanorobotic manipulation under hybrid microscope," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 4139–4144.
[14] I. Joachimsthaler, R. Heiderhoff, and L. J. Balk, "A universal scanning-probe-microscope-based hybrid system," *Meas. Sci. Technol.*, vol. 14, no. 1, p. 87, 2002.

[15] U. Mick, V. Eichhorn, T. Wortmann, C. Diederichs, and S. Fatikow, "Combined nanorobotic AFM/SEM system as novel toolbox for automated hybrid analysis and manipulation of nanoscale objects," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2010, pp. 4088–4093.

[16] C. Shi *et al.*, "Recent advances in nanorobotic manipulation inside scanning electron microscopes," *Microsyst. Nanoeng.*, vol. 2, no. 1, p. 16024, Dec. 2016.

[17] E. Meyer, H. J. Hug, and R. Bennewitz, *Scanning Probe Microscopy: The Lab on a Tip*. Springer, 2003, pp. 2–14.

[18] A. Bazaei, Y. K. Yong, and S. O. R. Moheimani, "Combining spiral scanning and internal model control for sequential AFM imaging at video rate," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 1, pp. 371–380, Feb. 2017.

[19] H. Xie, Y. Wen, X. Shen, H. Zhang, and L. Sun, "High-speed AFM imaging of nanopositioning stages using H_∞ and iterative learning control," *IEEE Trans. Ind. Electron.*, vol. 67, no. 3, pp. 2430–2439, Mar. 2020.

[20] S. Fatikow, *Automated Nanohandling by Microrobots*. Berlin, Germany: Springer, 2007.

[21] R. Wang and X. Zhang, "Optimal design of a planar parallel 3-DOF nanopositioner with multi-objective," *Mechanism Mach. Theory*, vol. 112, pp. 61–83, Jun. 2017.

[22] H. Bettahar, A. Caspar, C. Clévy, N. Courjal, and P. Lutz, "Photorobotic positioning for integrated optics," *IEEE Robot. Autom. Lett.*, vol. 2, no. 1, pp. 217–222, Jan. 2017.

[23] D. B. Dusenbery, *Living at Micro Scale: The Unexpected Physics of Being Small*. Cambridge, MA, USA: Harvard Univ. Press, 2009.

[24] D. Feng and M. Q. Feng, "Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection—A review," *Eng. Struct.*, vol. 156, pp. 105–117, Feb. 2018.

[25] Y. Xu and J. M. W. Brownjohn, "Review of machine-vision based methodologies for displacement measurement in civil structures," *J. Civil Struct. Health Monitor.*, vol. 8, no. 1, pp. 91–110, Jan. 2018.

[26] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Comput. Surv.*, vol. 38, no. 4, p. 13, 2006.

[27] H. Kawata, J. M. Carter, A. Yen, and H. I. Smith, "Optical projection lithography using lenses with numerical apertures greater than unity," *Microelectron. Eng.*, vol. 9, nos. 1–4, pp. 31–36, May 1989.

[28] Z. Chen, H. Liao, and X. Zhang, "Telecentric stereo micro-vision system: Calibration method and experiments," *Opt. Lasers Eng.*, vol. 57, pp. 82–92, Jun. 2014.

[29] M. Watanabe and S. K. Nayar, "Telecentric optics for focus analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 12, pp. 1360–1365, Dec. 1997.

[30] H. W. Schreier, D. Garcia, and M. A. Sutton, "Advances in light microscope stereo vision," *Experim. Mech.*, vol. 44, no. 3, pp. 278–288, Jun. 2004.

[31] B. Pan, L. Yu, and D. Wu, "High-accuracy 2D digital image correlation measurements with bilateral telecentric lenses: Error analysis and experimental verification," *Experim. Mech.*, vol. 53, no. 9, pp. 1719–1733, Nov. 2013.

[32] Y. Hu, Q. Chen, S. Feng, T. Tao, A. Asundi, and C. Zuo, "A new microscopic telecentric stereo vision system–calibration, rectification, and three-dimensional reconstruction," *Opt. Lasers Eng.*, vol. 113, pp. 14–22, Feb. 2019.

[33] D. Li and J. Tian, "An accurate calibration method for a camera with telecentric lenses," *Opt. Lasers Eng.*, vol. 51, no. 5, pp. 538–541, May 2013.

[34] J. W. Lichtman and J.-A. Conchello, "Fluorescence microscopy," *Nature Methods*, vol. 2, no. 12, pp. 910–919, Dec. 2005.

[35] L. Schermelleh, R. Heintzmann, and H. Leonhardt, "A guide to super-resolution fluorescence microscopy," *J. Cell Biol.*, vol. 190, no. 2, pp. 165–175, Jul. 2010.

[36] W. Gao, A. Shakoor, L. Zhao, Z. Jiang, and D. Sun, "3-D image reconstruction of biological organelles with a robot-aided microscopy system for intracellular surgery," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 231–238, Apr. 2019.

[37] P. Sarder and A. Nehorai, "Deconvolution methods for 3-D fluorescence microscopy images," *IEEE Signal Process. Mag.*, vol. 23, no. 3, pp. 32–45, May 2006.

[38] S. W. Paddock and K. W. Eliceiri, "Laser scanning confocal microscopy: history, applications, and related optical sectioning techniques," in *Confocal Microscopy*. Berlin, Germany: Springer, 2014, pp. 9–47.

[39] W. Zhao, Y. Sun, Y. Wang, L. Qiu, R. Shao, and H. Cui, "Three-dimensional super-resolution correlation-differential confocal microscopy with nanometer axial focusing accuracy," *Opt. Exp.*, vol. 26, no. 12, pp. 15759–15768, 2018.

[40] S. Choi, P. Kim, R. Boutilier, M. Y. Kim, Y. Lee, and H. Lee, "Development of a high speed laser scanning confocal microscope with an acquisition rate up to 200 frames per second," *Opt. Exp.*, vol. 21, no. 20, pp. 23611–23618, 2013.

[41] R. Lima, T. Ishikawa, Y. Imai, M. Takeda, S. Wada, and T. Yamaguchi, "Measurement of individual red blood cell motions under high hematocrit conditions using a confocal micro-PTV system," *Ann. Biomed. Eng.*, vol. 37, no. 8, pp. 1546–1559, 2009.

[42] C. Zhou *et al.*, "A closed-loop controlled nanomanipulation system for probing nanostructures inside scanning electron microscopes," *IEEE/ASME Trans. Mechatronics*, vol. 21, no. 3, pp. 1233–1241, Jun. 2016.

[43] R. Keyse, *Introduction to Scanning Transmission Electron Microscopy*. Evanston, IL, USA: Routledge, 2018.

[44] H. Ding *et al.*, "Visual servoing-based nanorobotic system for automated electrical characterization of nanotubes inside SEM," *Sensors*, vol. 18, no. 4, p. 1137, Apr. 2018.

[45] N. Marturi, B. Tamadazte, S. Dembélé, and N. Piat, "Image-guided nanopositioning scheme for SEM," *IEEE Trans. Autom. Sci. Eng.*, vol. 15, no. 1, pp. 45–56, Jan. 2018.

[46] N. Marturi, S. Dembélé, and N. Piat, "Fast image drift compensation in scanning electron microscope using image registration," in *Proc. IEEE Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2013, pp. 807–812.

[47] J. P. Buban and S.-Y. Choi, "Auto-encoders for noise reduction in scanning transmission electron microscopy," *Microsc. Microanal.*, vol. 23, no. S1, pp. 130–131, Jul. 2017.

[48] D. Jasper, "High-speed position tracking for nanohandling inside scanning electron microscopes," in *Proc. IEEE Int. Conf. Robot. Automat.*, Kobe, Japan, 2009, pp. 508–513, doi: 10.1109/ROBOT.2009.5152489.

[49] G. D. Danilatos, "Foundations of environmental scanning electron microscopy," *Adv. Electron. Electron Phys.*, vol. 71, p. 109–250, Jan. 1988.

[50] G. Danilatos, "Theory of the gaseous detector device in the environmental scanning electron microscope," in *Advances in Electronics and Electron Physics*, vol. 78. Amsterdam, The Netherlands: Elsevier, 1990, pp. 1–102.

[51] L. Cui and É. Marchand, "Scanning electron microscope calibration using a multi-image non-linear minimization process," *Int. J. Optomechtron.*, vol. 9, no. 2, pp. 151–169, Apr. 2015.

[52] H. Li, X. Zhang, H. Wu, and J. Gan, "Line-based calibration of a micro-vision motion measurement system," *Opt. Lasers Eng.*, vol. 93, pp. 40–46, Jun. 2017.

[53] H. Li, X. Zhang, and B. Zhu, "Single grid image based calibration of an optical microscope," in *Proc. Int. Conf. Manipulation, Autom. Robot. Small Scales (MARSS)*, Jul. 2017, pp. 1–6.

[54] Y. Zhou and B. J. Nelson, "Calibration of a parametric model of an optical microscope," *Opt. Eng.*, vol. 38, no. 12, pp. 1989–1995, 1999.

[55] M. Ammi, V. Frémont, and A. Ferreira, "Automatic camera-based microscope calibration for a telemicromanipulation system using a virtual pattern," *IEEE Trans. Robot.*, vol. 25, no. 1, pp. 184–191, Feb. 2009.

[56] L. Ren, L. Wang, J. K. Mills, and D. Sun, "Vision-based 2-D automatic micrograsping using coarse-to-fine grasping strategy," *IEEE Trans. Ind. Electron.*, vol. 55, no. 9, pp. 3324–3331, Sep. 2008.

[57] S. Yao, H. Li, K. Wang, and X. Zhang, "High-accuracy calibration of a visual motion measurement system for planar 3-DOF robots using Gaussian process," *IEEE Sensors J.*, vol. 19, no. 17, pp. 7659–7667, Sep. 2019.

[58] J.-A. Conchello and J. W. Lichtman, "Optical sectioning microscopy," *Nature Methods*, vol. 2, no. 12, pp. 920–931, 2005.

[59] C. Wang, J. Caja, E. Gómez, and P. Maresca, "Procedure for calibrating the Z-axis of a confocal microscope: Application for the evaluation of structured surfaces," *Sensors*, vol. 19, no. 3, p. 527, Jan. 2019.

[60] A. Mínguez Martínez and J. de Vicente y Oliva, "Industrial calibration procedure for confocal microscopes," *Materials*, vol. 12, no. 24, p. 4137, Dec. 2019.

[61] N. J. Mudrak, P. S. Rana, and M. A. Model, "Calibrated brightfield-based imaging for measuring intracellular protein concentration," *Cytometry A*, vol. 93, no. 3, pp. 297–304, Mar. 2018.

[62] L. Quan, "Self-calibration of an affine camera from multiple views," *Int. J. Comput. Vis.*, vol. 19, no. 1, pp. 93–105, Jul. 1996.

[63] L. Yao and H. Liu, "A flexible calibration approach for cameras with double-sided telecentric lenses," *Int. J. Adv. Robot. Syst.*, vol. 13, no. 3, p. 82, Jun. 2016.

[64] X. Liu, Z. Li, P. Miraldo, K. Zhong, and Y. Shi, "A framework to calibrate the scanning electron microscope under variational magnifications," *IEEE Photon. Technol. Lett.*, vol. 28, no. 16, pp. 1715–1718, Aug. 15, 2016.

[65] M. Ritter, M. Hemmleb, B. Lich, P. Faber, and H. Hohenberg, "SEM/FIB stage calibration with photogrammetric methods," in *Proc. ISPRS Commission V Symp. Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 36, 2006.

[66] L. Cui and E. Marchand, "Calibration of scanning electron microscope using a multi-image non-linear minimization process," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2014, pp. 5191–5196.

[67] P. Miraldo and H. Araujo, "Calibration of smooth camera models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2091–2103, Sep. 2013.

[68] Y. A. Novikov, "Calibration of a scanning electron microscope from two coordinates," *J. Surf. Investigation: X-ray, Synchrotron Neutron Techn.*, vol. 11, no. 4, pp. 890–896, Jul. 2017.

[69] C. P. Volk, Y. A. Novikov, A. V. Rakov, and P. A. Todua, "Calibrating a scanning electron microscope in two coordinates by the use of one certified dimension," *Meas. Techn.*, vol. 51, no. 6, pp. 605–608, Jun. 2008.

[70] S. Pang, X. Zhang, X. Zhang, and Y. Lu, "A magnification-continuous calibration method for SEM-based nanorobotic manipulation systems," *Rev. Sci. Instrum.*, vol. 90, no. 5, May 2019, Art. no. 053706.

[71] F. Marinello, P. Bariani, E. Savio, A. Horsewell, and L. De Chiffre, "Critical factors in SEM 3D stereo microscopy," *Meas. Sci. Technol.*, vol. 19, no. 6, Jun. 2008, Art. no. 065705.

[72] G. Lu, Y. Zhu, G. Su, Z. Zhang, and P. Yan, "Efficient block matching using improved particle swarm optimization with application to displacement measurement for nano motion systems," *Opt. Lasers Eng.*, vol. 111, pp. 246–254, Dec. 2018.

[73] J. Joglekar and S. S. Gedam, "Area based image matching methods—A survey," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 2, no. 1, pp. 130–136, 2012.

[74] B. Pan, K. Qian, H. Xie, and A. Asundi, "Two-dimensional digital image correlation for in-plane displacement and strain measurement: A review," *Meas. Sci. Technol.*, vol. 20, no. 6, Jun. 2009, Art. no. 062001.

[75] L. Di Stefano, S. Mattoccia, and F. Tombari, "ZNCC-based template matching using bounded partial correlation," *Pattern Recognit. Lett.*, vol. 26, no. 14, pp. 2129–2134, Oct. 2005.

[76] B. Pan, K. Li, and W. Tong, "Fast, robust and accurate digital image correlation calculation without redundant computations," *Experim. Mech.*, vol. 53, no. 7, pp. 1277–1289, Sep. 2013.

[77] H. Wu, X. Zhang, R. Wang, and Z. He, "Displacement measurement of the compliant positioning stage based on a computer micro-vision method," *AIP Adv.*, vol. 6, no. 2, Feb. 2016, Art. no. 025009.

[78] X. Zhang, X. Zhang, H. Wu, J. Gan, and H. Li, "A high accuracy algorithm of displacement measurement for a micro-positioning stage," *AIP Adv.*, vol. 7, no. 5, May 2017, Art. no. 055301.

[79] M. D. Jenkins, P. Barrie, T. Buggy, and G. Morison, "Extended fast compressive tracking with weighted multi-frame template matching for fast motion tracking," *Pattern Recognit. Lett.*, vol. 69, pp. 82–87, Jan. 2016.

[80] F. Zhong and C. Quan, "Digital image correlation in polar coordinate robust to a large rotation," *Opt. Lasers Eng.*, vol. 98, pp. 153–158, Nov. 2017.

[81] H. Li, X. Zhang, B. Zhu, and S. Fatikow, "Online precise motion measurement of 3-DOF nanopositioners based on image correlation," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 3, pp. 782–790, Mar. 2019.

[82] H. Li, B. Zhu, Z. Chen, and X. Zhang, "Realtime in-plane displacements tracking of the precision positioning stage based on computer micro-vision," *Mech. Syst. Signal Process.*, vol. 124, pp. 111–123, Jun. 2019.

[83] H. Li, X. Zhang, S. Yao, B. Zhu, and S. Fatikow, "An improved template-matching-based pose tracking method for planar nanopositioning stages using enhanced correlation coefficient," *IEEE Sensors J.*, vol. 20, no. 12, pp. 6378–6387, Jun. 2020.

[84] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," in *Proc. 10th IEEE Int. Conf. Comput. Vis. (ICCV)*, vols. 1–2, 2005, pp. 1508–1515.

[85] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[86] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, Jun. 2008.

[87] T. Lindeberg, "Image matching using generalized scale-space interest points," *J. Math. Imag. Vis.*, vol. 52, no. 1, pp. 3–36, May 2015.

[88] C. Singh and E. Walia, "Fast and numerically stable methods for the computation of Zernike moments," *Pattern Recognit.*, vol. 43, no. 7, pp. 2497–2506, Jul. 2010.

[89] C. G. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vis. Conf.*, 1988, vol. 15, no. 50, pp. 10–5244.

[90] N. Singhal, Y.-Y. Lee, C.-S. Kim, and S.-U. Lee, "Robust image watermarking using local Zernike moments," *J. Vis. Commun. Image Represent.*, vol. 20, no. 6, pp. 408–419, Aug. 2009.

[91] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571.

[92] A. Alahi, R. Ortiz, and P. Vandergheynst, "FREAK: Fast retina keypoint," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 510–517.

[93] E. Mair, G. D. Hager, D. Burschka, M. Suppa, and G. Hirzinger, "Adaptive and generic corner detection based on the accelerated segment test," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2010, pp. 183–196.

[94] H. P. Moravec, "Obstacle avoidance and navigation in the real world by a seeing robot rover," Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, Tech. Rep. CMU-RI-TR-80-03, 1980.

[95] S. M. Smith and J. M. Brady, "Susan—A new approach to low level image processing," *Int. J. Comput. Vis.*, vol. 23, no. 1, pp. 45–78, 1997.

[96] T. Lindeberg, "Feature detection with automatic scale selection," *Int. J. Comput. Vis.*, vol. 30, no. 2, pp. 79–116, 1998.

[97] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image Vis. Comput.*, vol. 22, no. 10, pp. 761–767, Sep. 2004.

[98] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[99] J. Heikkilä, "Pattern matching with affine moment descriptors," *Pattern Recognit.*, vol. 37, no. 9, pp. 1825–1834, Sep. 2004.

[100] C. Clemente, L. Pallotta, A. De Maio, J. J. Soraghan, and A. Farina, "A novel algorithm for radar classification based on Doppler characteristics exploiting orthogonal pseudo-Zernike polynomials," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 1, pp. 417–430, Jan. 2015.

[101] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2548–2555.

[102] H. Wu, X. Zhang, J. Gan, H. Li, and Z. He, "High-precision displacement measurement method for three degrees of freedom-compliant mechanisms based on computer micro-vision," *Appl. Opt.*, vol. 55, no. 10, pp. 2594–2600, 2016.

[103] J. Peng, S. Peng, and Y. Hu, "Partial least squares and random sample consensus in outlier detection," *Analytica Chim. Acta*, vol. 719, pp. 24–29, Mar. 2012.

[104] X. Yin, G. Jiang, and S. Song, "Research on an automatic tracking strategy based on CCD image sensor in micromanipulation," *IEEE Access*, vol. 6, pp. 76374–76380, 2018.

[105] C. Zhao, C. F. Cheung, and P. Xu, "Optical nanoscale positioning measurement with a feature-based method," *Opt. Lasers Eng.*, vol. 134, Nov. 2020, Art. no. 106225.

[106] P. Sidiropoulos and J.-P. Muller, "Matching of large images through coupled decomposition," *IEEE Trans. Image Process.*, vol. 24, no. 7, pp. 2124–2139, Jul. 2015.

[107] X. Peng, Q. Bai, X. Xia, Z. Huang, K. Saenko, and B. Wang, "Moment matching for multi-source domain adaptation," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2019, pp. 1406–1415.

[108] M. Habrat and M. Młynarczuk, "Evaluation of local matching methods in image analysis for mineral grain tracking in microscope images of rock sections," *Minerals*, vol. 8, no. 5, p. 182, Apr. 2018.

[109] F. Mualla, S. Schöll, B. Sommerfeldt, A. Maier, and J. Hornegger, "Automatic cell detection in bright-field microscope images using SIFT, random forests, and hierarchical clustering," *IEEE Trans. Med. Imag.*, vol. 32, no. 12, pp. 2274–2286, Dec. 2013.

[110] B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE Trans. Image Process.*, vol. 5, no. 8, pp. 1266–1271, Aug. 1996.

[111] J. Z. Galeano *et al.*, "Position-referenced microscopy for live cell culture monitoring," *Biomed. Opt. Exp.*, vol. 2, no. 5, pp. 1307–1318, 2011.

[112] S. Ri, S. Hayashi, S. Ogihara, and H. Tsuda, "Accurate full-field optical displacement measurement technique using a digital camera and repeated patterns," *Opt. Exp.*, vol. 22, no. 8, pp. 9693–9706, 2014.

[113] V. Guelpa, P. Sandoz, M. A. Vergara, C. Clévy, N. Le Fort-Piat, and G. J. Laurent, "2D visual micro-position measurement based on intertwined twin-scale patterns," *Sens. Actuators A, Phys.*, vol. 248, pp. 272–280, Sep. 2016.

[114] V. Argyriou and T. Vlachos, "A study of sub-pixel motion estimation using phase correlation," in *Proc. BMVC*. State College, PA, USA: Citeseer, 2006, pp. 387–396.

[115] H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Trans. Image Process.*, vol. 11, no. 3, pp. 188–200, Mar. 2002.

[116] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Opt. Lett.*, vol. 33, no. 2, pp. 156–158, 2008.

[117] H. S. Stone, M. T. Orchard, E.-C. Chang, and S. A. Martucci, "A fast direct Fourier-based algorithm for subpixel registration of images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 10, pp. 2235–2243, Oct. 2001.

[118] W. S. Hoge, "A subspace identification extension to the phase correlation method," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 277–280, Feb. 2003.

[119] X. Tong *et al.*, "A novel subpixel phase correlation method using singular value decomposition and unified random sample consensus," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4143–4156, Aug. 2015.

[120] V. Argyriou and G. Tzimiropoulos, "Frequency domain subpixel registration using HOG phase correlation," *Comput. Vis. Image Understand.*, vol. 155, pp. 70–82, Feb. 2017.

[121] Z. Jiang, Q. Kemao, H. Miao, J. Yang, and L. Tang, "Path-independent digital image correlation with high accuracy, speed and robustness," *Opt. Lasers Eng.*, vol. 65, pp. 93–102, Feb. 2015.

[122] A. N. Andre, P. Sandoz, B. Mauze, M. Jacquot, and G. J. Laurent, "Sensing one nanometer over ten centimeters: A microencoded target for visual in-plane position measurement," *IEEE/ASME Trans. Mechatronics*, vol. 25, no. 3, pp. 1193–1201, Jun. 2020.

[123] Z.-H. Chen and P. S. Huang, "A vision-based method for planar position measurement," *Meas. Sci. Technol.*, vol. 27, no. 12, Dec. 2016, Art. no. 125018.

[124] D. Sun, S. Roth, and M. J. Black, "Secrets of optical flow estimation and their principles," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2432–2439.

[125] B. K. P. Horn and B. G. Schunck, "Determining optical flow," in *Techniques and Applications of Image Understanding*, vol. 281. Bellingham, WA, USA: International Society for Optics and Photonics, 1981, pp. 319–331.

[126] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. 7th Int. Joint Conf. Artif. Intell. (IJCAI)*, vol. 2. San Francisco, CA, USA: Morgan Kaufmann, 1981, pp. 674–679.

[127] B. Pan, A. Asundi, H. Xie, and J. Gao, "Digital image correlation using iterative least squares and pointwise least squares for displacement field and strain field measurements," *Opt. Lasers Eng.*, vol. 47, nos. 7–8, pp. 865–874, Jul. 2009.

[128] H. Tho Ho and R. Goecke, "Optical flow estimation using Fourier Mellin transform," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[129] J.-Y. Bouguet *et al.*, "Pyramidal implementation of the affine Lucas Kanade feature tracker description of the algorithm," *Intel Corp.*, vol. 5, nos. 1–10, p. 4, 2001.

[130] T. Brox and J. Malik, "Large displacement optical flow: Descriptor matching in variational motion estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 500–513, Mar. 2011.

[131] I. Ishii, T. Taniguchi, K. Yamamoto, and T. Takaki, "High-frame-rate optical flow system," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 1, pp. 105–112, Jan. 2012.

[132] C. Hartmann, J. Wang, D. Opristescu, and W. Volk, "Implementation and evaluation of optical flow methods for two-dimensional deformation measurement in comparison to digital image correlation," *Opt. Lasers Eng.*, vol. 107, pp. 127–141, Aug. 2018.

[133] T. Guo, H. Chang, J. Chen, X. Fu, and X. Hu, "Micro-motion analyzer used for dynamic MEMS characterization," *Opt. Lasers Eng.*, vol. 47, nos. 3–4, pp. 512–517, Mar. 2009.

[134] D. Guo, A. L. van de Ven, and X. Zhou, "Red blood cell tracking using optical flow methods," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 3, pp. 991–998, May 2014.

[135] B. Li, R. Chellappa, Q. Zheng, and S. Z. Der, "Model-based temporal object verification using video," *IEEE Trans. Image Process.*, vol. 10, no. 6, pp. 897–908, Jun. 2001.

[136] T. Drummond and R. Cipolla, "Real-time visual tracking of complex structures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 932–946, Jul. 2002.

[137] K. B. Yesin and B. J. Nelson, "A CAD model based tracking system for visually guided microassembly," *Robotica*, vol. 23, no. 4, pp. 409–418, Jul. 2005.

[138] N. Peterfreund, "Robust tracking of position and velocity with Kalman snakes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 6, pp. 564–569, Jun. 1999.

[139] C. Yang, R. Duraiswami, and L. Davis, "Fast multiple object tracking via a hierarchical particle filter," in *Proc. 10th IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 1, Oct. 2005, pp. 212–219.

[140] F. Sardari and M. Ebrahimi Moghaddam, "A hybrid occlusion free object tracking method using particle filter and modified galaxy based search meta-heuristic algorithm," *Appl. Soft Comput.*, vol. 50, pp. 280–299, Jan. 2017.

[141] W. Yi, M. R. Morelande, L. Kong, and J. Yang, "A computationally efficient particle filter for multitarget tracking using an independence approximation," *IEEE Trans. Signal Process.*, vol. 61, no. 4, pp. 843–856, Feb. 2013.

[142] X. Li, T. Zhang, X. Shen, and J. Sun, "Object tracking using an adaptive Kalman filter combined with mean shift," *Opt. Eng.*, vol. 49, no. 2, Feb. 2010, Art. no. 020503.

[143] H. Wang *et al.*, "Contact assembly of cell-laden hollow microtubes through automated micromanipulator tip locating," *J. Micromech. Microeng.*, vol. 27, no. 1, Jan. 2017, Art. no. 015013.

[144] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[145] W. Hu, X. Zhou, W. Li, W. Luo, X. Zhang, and S. Maybank, "Active contour-based visual tracking by integrating colors, shapes, and motions," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1778–1792, May 2013.

[146] K. Zhang, L. Zhang, K.-M. Lam, and D. Zhang, "A level set approach to image segmentation with intensity inhomogeneity," *IEEE Trans. Cybern.*, vol. 46, no. 2, pp. 546–557, Feb. 2016.

[147] J. Liu *et al.*, "Locating end-effector tips in robotic micromanipulation," *IEEE Trans. Robot.*, vol. 30, no. 1, pp. 125–130, Feb. 2014.

[148] M. Maška, O. Daněk, S. Garasa, A. Rouzaut, A. Muñoz-Barrutia, and C. Ortiz-de-Solorzano, "Segmentation and shape tracking of whole fluorescent cells based on the Chan–Vese model," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 995–1006, Jun. 2013.

[149] X. Liu, Y. Wang, and Y. Sun, "Cell contour tracking and data synchronization for real-time, high-accuracy micropipette aspiration," *IEEE Trans. Autom. Sci. Eng.*, vol. 6, no. 3, pp. 536–543, Jul. 2009.

[150] P. Cheng and C.-H. Menq, "Visual tracking of six-axis motion rendering ultraprecise visual servoing of microscopic objects," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 4, pp. 1564–1572, Aug. 2018.

[151] B. E. Kratochvil, L. Dong, and B. J. Nelson, "Real-time rigid-body visual tracking in a scanning electron microscope," *Int. J. Robot. Res.*, vol. 28, no. 4, pp. 498–511, Apr. 2009.

[152] C. Dahmen and T. Tiemerding, "Fast and robust position determination in the scanning electron microscope," in *Proc. Australas. Conf. Robot. Autom.*, 2013, pp. 2–4.

[153] C. Ru, Y. Zhang, H. Huang, and T. Chen, "An improved visual tracking method in scanning electron microscope," *Microsc. Microanal.*, vol. 18, no. 3, pp. 612–620, Jun. 2012.

[154] Z. Gong, B. K. Chen, J. Liu, and Y. Sun, "Robotic probing of nanostructures inside scanning electron microscopy," *IEEE Trans. Robot.*, vol. 30, no. 3, pp. 758–765, Jun. 2014.

[155] C. van der Wel and D. J. Kraft, "Automated tracking of colloidal clusters with sub-pixel accuracy and precision," *J. Phys., Condens. Matter*, vol. 29, no. 4, Feb. 2017, Art. no. 044001.

[156] Z. Chen, D. Zhou, H. Liao, and X. Zhang, "Precision alignment of optical fibers based on telecentric stereo microvision," *IEEE/ASME Trans. Mechatronics*, vol. 21, no. 4, pp. 1924–1934, Aug. 2016.

[157] K. M. Taute, S. Gude, S. J. Tans, and T. S. Shimizu, "High-throughput 3D tracking of bacteria on a standard phase contrast microscope," *Nature Commun.*, vol. 6, no. 1, pp. 1–9, Dec. 2015.

[158] S. Xiao and Y. Li, "Visual servo feedback control of a novel large working range micro manipulation system for microassembly," *J. Microelectromech. Syst.*, vol. 23, no. 1, pp. 181–190, Feb. 2014.

[159] C. Y. Wong and J. K. Mills, "Cleavage-stage embryo rotation tracking and automated micropipette control: Towards automated single cell manipulation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 2351–2356.

[160] X. Zhang, X. Zhang, H. Wu, H. Li, and J. Gan, "A robust rotation-invariance displacement measurement method for a micro-/nano-positioning system," *Meas. Sci. Technol.*, vol. 29, no. 5, May 2018, Art. no. 055402.

[161] Z. Wu *et al.*, "A novel self-feedback intelligent vision measure for fast and accurate alignment in flip-chip packaging," *IEEE Trans. Ind. Informat.*, vol. 16, no. 3, pp. 1776–1787, Mar. 2020.

[162] Y. Xue *et al.*, "High-accuracy and real-time 3D positioning, tracking system for medical imaging applications based on 3D digital image correlation," *Opt. Lasers Eng.*, vol. 88, pp. 82–90, Jan. 2017.

[163] W. Shang, H. Lu, W. Wan, T. Fukuda, and Y. Shen, "Vision-based nano robotic system for high-throughput non-embedded cell cutting," *Sci. Rep.*, vol. 6, no. 1, pp. 1–14, Mar. 2016.

[164] G. F. Bomarito, J. D. Hochhalter, T. J. Ruggles, and A. H. Cannon, "Increasing accuracy and precision of digital image correlation through pattern optimization," *Opt. Lasers Eng.*, vol. 91, pp. 73–85, Apr. 2017.

[165] W. Huang, C. Ma, and Y. Chen, "Displacement measurement with nanoscale resolution using a coded micro-mark and digital image correlation," *Opt. Eng.*, vol. 53, no. 12, Dec. 2014, Art. no. 124103.

[166] C. Zhao, C. Cheung, and M. Liu, "Integrated polar microstructure and template-matching method for optical position measurement," *Opt. Exp.*, vol. 26, no. 4, pp. 4330–4345, 2018.

[167] H. Li, X. Zhang, B. Zhu, Y. Lu, and H. Wu, "Micro-motion detection of the 3-DOF precision positioning stage based on iterative optimized template matching," *Appl. Opt.*, vol. 56, no. 34, pp. 9435–9443, 2017.

[168] B. Pan, "Bias error reduction of digital image correlation using Gaussian pre-filtering," *Opt. Lasers Eng.*, vol. 51, no. 10, pp. 1161–1167, Oct. 2013.

[169] D. Mas, J. Perez, B. Ferrer, and J. Espinosa, "Realistic limits for subpixel movement detection," *Appl. Opt.*, vol. 55, no. 19, pp. 4974–4979, 2016.

[170] V. Guelpa, G. Laurent, P. Sandoz, J. Zea, and C. Clévy, "Subpixelic measurement of large 1D displacements: Principle, processing algorithms, performances and software," *Sensors*, vol. 14, no. 3, pp. 5056–5073, Mar. 2014.

[171] A. Ya'akobovitz, S. Krylov, and Y. Hanein, "Nanoscale displacement measurement of electrostatically actuated micro-devices using optical microscopy and digital image correlation," *Sens. Actuators A, Phys.*, vol. 162, no. 1, pp. 1–7, Jul. 2010.

[172] B. Tamadazte, E. Marchand, S. Dembélé, and N. Le Fort-Piat, "CAD model-based tracking and 3D visual-based control for MEMS microassembly," *Int. J. Robot. Res.*, vol. 29, no. 11, pp. 1416–1434, Sep. 2010.

[173] M. A. Sutton, J. H. Yan, V. Tiwari, H. W. Schreier, and J. J. Orteu, "The effect of out-of-plane motion on 2D and 3D digital image correlation measurements," *Opt. Lasers Eng.*, vol. 46, no. 10, pp. 746–757, Oct. 2008.

[174] L. Cui, E. Marchand, S. Haliyo, and S. Régnier, "Three-dimensional visual tracking and pose estimation in scanning electron microscopes," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 5210–5215.

[175] G. Behan, E. C. Cosgriff, A. I. Kirkland, and P. D. Nellist, "Three-dimensional imaging by optical sectioning in the aberration-corrected scanning transmission electron microscope," *Phil. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 367, no. 1903, pp. 3825–3844, Sep. 2009.

[176] Y. Shechtman, S. J. Sahl, A. S. Backer, and W. E. Moerner, "Optimal point spread function design for 3D imaging," *Phys. Rev. Lett.*, vol. 113, no. 13, Sep. 2014, Art. no. 133902.

[177] Y. Sun, S. Duthaler, and B. J. Nelson, "Autofocusing in computer microscopy: Selecting the optimal focus algorithm," *Microsc. Res. Techn.*, vol. 65, no. 3, pp. 139–149, 2004.

[178] A. V. Kudryavtsev, S. Dembélé, and N. Piat, "Full 3D rotation estimation in scanning electron microscope," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 1134–1139.

[179] C. Franck, S. Hong, S. Maskarinec, D. Tirrell, and G. Ravichandran, "Three-dimensional full-field measurements of large deformations in soft materials using confocal microscopy and digital Volume correlation," *Experim. Mech.*, vol. 47, no. 3, pp. 427–438, 2007.

[180] M. Jähnisch and S. Fatikow, "3-D vision feedback for nanohandling monitoring in a scanning electron microscope," *Int. J. Optomechtron.*, vol. 1, no. 1, pp. 4–26, Mar. 2007.

[181] H. Yamamoto and T. Sano, "Study of micromanipulation using stereoscopic microscope," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 2, pp. 182–187, Apr. 2002.

[182] H. Wu, X. Zhang, J. Gan, H. Li, and P. Ge, "Displacement measurement system for inverters using computer micro-vision," *Opt. Lasers Eng.*, vol. 81, pp. 113–118, Jun. 2016.

[183] S.-D. Wei and S.-H. Lai, "Fast template matching based on normalized cross correlation with adaptive multilevel winner update," *IEEE Trans. Image Process.*, vol. 17, no. 11, pp. 2227–2235, Nov. 2008.

[184] S. Omachi and M. Omachi, "Fast template matching with polynomials," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2139–2149, Aug. 2007.

[185] C. Zhao, Y. Li, Y. Yao, and D. Deng, "Random residual neural network–based nanoscale positioning measurement," *Opt. Exp.*, vol. 28, no. 9, pp. 13125–13130, 2020.

[186] S. Baker and I. Matthews, "Lucas-kanade 20 years on: A unifying framework," *Int. J. Comput. Vis.*, vol. 56, no. 3, pp. 221–255, Feb. 2004.

[187] X. Shao, X. Dai, and X. He, "Noise robustness and parallel computation of the inverse compositional Gauss–Newton algorithm in digital image correlation," *Opt. Lasers Eng.*, vol. 71, pp. 9–19, Aug. 2015.

[188] V. Balntas, K. Lenc, A. Vedaldi, and K. Mikolajczyk, "HPatches: A benchmark and evaluation of handcrafted and learned local descriptors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 5173–5182.

[189] E. Simo-Serra, E. Trulls, L. Ferraz, I. Kokkinos, P. Fua, and F. Moreno-Noguer, "Discriminative learning of deep convolutional feature point descriptors," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 118–126.

[190] J. Ren, T. Vlachos, Y. Zhang, J. Zheng, and J. Jiang, "Gradient-based subspace phase correlation for fast and effective image alignment," *J. Vis. Commun. Image Represent.*, vol. 25, no. 7, pp. 1558–1565, Oct. 2014.

[191] Y. Deng, P. Coen, M. Sun, and J. W. Shaevitz, "Efficient multiple object tracking using mutually repulsive active membranes," *PLoS ONE*, vol. 8, no. 6, Jun. 2013, Art. no. e65769.

[192] K. Li, E. D. Miller, M. Chen, T. Kanade, L. E. Weiss, and P. G. Campbell, "Cell population tracking and lineage construction with spatiotemporal context," *Med. Image Anal.*, vol. 12, no. 5, pp. 546–566, Oct. 2008.

[193] T. Sievers and S. Fatikow, "Real-time object tracking for the robot-based nanohandling in a scanning electron microscope," *J. Micromechtron.*, vol. 3, nos. 3–4, pp. 267–284, Sep. 2006.

[194] H. Li, J. Lu, G. Shi, and Y. Zhang, "Tracking features in retinal images of adaptive optics confocal scanning laser ophthalmoscope using KLT-SIFT algorithm," *Biomed. Opt. Exp.*, vol. 1, no. 1, pp. 31–40, 2010.

[195] T. T. Ashley, E. L. Gan, J. Pan, and S. B. Andersson, "Tracking single fluorescent particles in three dimensions via extremum seeking," *Biomed. Opt. Exp.*, vol. 7, no. 9, pp. 3355–3376, 2016.

[196] H. Li, X.-M. Zhang, L. Zeng, and Y.-J. Huang, "A monocular vision system for online pose measurement of a 3RRR planar parallel manipulator," *J. Intell. Robotic Syst.*, vol. 92, no. 1, pp. 3–17, Sep. 2018.

[197] F. Timischl, M. Date, and S. Nemoto, "A statistical model of signal-noise in scanning electron microscopy," *Scanning*, vol. 34, no. 3, pp. 137–144, May 2012.

[198] N. Marturi, S. Dembélé, and N. Piat, "Scanning electron microscope image signal-to-noise ratio monitoring for micro-nanomanipulation: SEM image SNR monitoring for micro-nanomanipulation," *Scanning, J. Scanning Microscopies*, vol. 36, no. 4, pp. 419–429, Jul. 2014.

[199] K. S. Sim, V. Teh, and M. E. Nia, "Adaptive noise Wiener filter for scanning electron microscope imaging system," *Scanning*, vol. 38, no. 2, pp. 148–163, Mar. 2016.

[200] D. Hwan Kim, S. Jae Kim, and S. Kyu Oh, "Image improvement with modified scanning waves and noise reduction in a scanning electron microscope," *Nucl. Instrum. Methods Phys. Res. A, Accel. Spectrom. Detect. Assoc. Equip.*, vol. 620, nos. 2–3, pp. 112–120, Aug. 2010.

[201] S. Maraghechi, J. P. M. Hoefnagels, R. H. J. Peerlings, O. Rokoš, and M. G. D. Geers, "Correction of scanning electron microscope imaging artifacts in a novel digital image correlation framework," *Experim. Mech.*, vol. 59, no. 4, pp. 489–516, Apr. 2019.

[202] M. T. Snella, "Drift correction for scanning-electron microscopy," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 2010.

[203] A. C. Malti, S. Dembélé, N. Piat, C. Arnoult, and N. Marturi, "Toward fast calibration of global drift in scanning electron microscopes with respect to time and magnification," *Int. J. Optomechtron.*, vol. 6, no. 1, pp. 1–16, Jan. 2012.
[204] H. Li, F. T. von Kleist-Retzow, O. C. Haenssler, S. Fatikow, and X. Zhang, "Multi-target tracking for automated RF on-wafer probing based on template matching," in *Proc. Int. Conf. Manipulation, Autom. Robot. at Small Scales (MARSS)*, Jul. 2019, pp. 1–6.
[205] W. Chen, X. Zhang, and S. Fatikow, "A novel microgripper hybrid driven by a piezoelectric stack actuator and piezoelectric cantilever actuators," *Rev. Sci. Instrum.*, vol. 87, no. 11, Nov. 2016, Art. no. 115003.
[206] V. Guelpa *et al.*, "3D-printed vision-based micro-force sensor dedicated to *in situ* SEM measurements," in *Proc. IEEE Int. Conf. Adv. Intell. Mechatronics (AIM)*, Jul. 2017, pp. 424–429.
[207] S. Yang and Q. Xu, "A review on actuation and sensing techniques for MEMS-based microgrippers," *J. Micro-Bio Robot.*, vol. 13, nos. 1–4, pp. 1–14, Oct. 2017.
[208] B. Zhu *et al.*, "Design of compliant mechanisms using continuum topology optimization: A review," *Mechanism Mach. Theory*, vol. 143, Jan. 2020, Art. no. 103622.
[209] S. J. Ralis, B. Vikramaditya, and B. J. Nelson, "Micropositioning of a weakly calibrated microassembly system using coarse-to-fine visual servoing strategies," *IEEE Trans. Electron. Packag. Manuf.*, vol. 23, no. 2, pp. 123–131, Apr. 2000.
[210] S. Fahlbusch *et al.*, "Nanomanipulation in a scanning electron microscope," *J. Mater. Process. Technol.*, vol. 167, nos. 2–3, pp. 371–382, Aug. 2005.
[211] X. Ye, Y. Zhang, C. Ru, J. Luo, S. Xie, and Y. Sun, "Automated pick-place of silicon nanowires," *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 3, pp. 554–561, Jul. 2013.
[212] S. Zimmermann, T. Tiemerding, and S. Fatikow, "Automated robotic manipulation of individual colloidal particles using vision-based control," *IEEE/ASME Trans. Mechatronics*, vol. 20, no. 5, pp. 2031–2038, Oct. 2015.
[213] S. He *et al.*, "A flip-chip alignment system with the property of deviation self-correction at the nanoscale," *IEEE Trans. Ind. Electron.*, vol. 68, no. 3, pp. 2345–2355, Mar. 2021.
[214] Y. Sun and B. J. Nelson, "Biological cell injection using an autonomous MicroRobotic system," *Int. J. Robot. Res.*, vol. 21, nos. 10–11, pp. 861–868, Oct. 2002.
[215] L. S. Mattos, E. Grant, R. Thresher, and K. Kluckman, "Blastocyst microinjection automation," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 5, pp. 822–831, Sep. 2009.
[216] X. Liu, Z. Lu, and Y. Sun, "Orientation control of biological cells under inverted microscopy," *IEEE/ASME Trans. Mechatronics*, vol. 16, no. 5, pp. 918–924, Oct. 2011.
[217] Y. Shen *et al.*, "Development of the auto manipulation system towards the single cell automatic analysis inside an environmental SEM," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2012, pp. 4594–4599.
[218] S. Chowdhury, A. Thakur, P. Švec, C. Wang, W. Losert, and S. K. Gupta, "Automated manipulation of biological cells using gripper formations controlled by optical tweezers," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 2, pp. 338–347, Apr. 2014.
[219] B. K. Chen *et al.*, "Nano-dissection and sequencing of DNA at single sub-nuclear structures," *Small*, vol. 10, no. 16, pp. 3267–3274, Aug. 2014.
[220] X. Wang *et al.*, "Intracellular manipulation and measurement with multipole magnetic tweezers," *Sci. Robot.*, vol. 4, no. 28, Mar. 2019, Art. no. eaav6180.
[221] Q. Xu, Y. Li, and N. Xi, "Design, fabrication, and visual servo control of an XY parallel micromanipulator with piezo-actuation," *IEEE Trans. Autom. Sci. Eng.*, vol. 6, no. 4, pp. 710–719, Oct. 2009.
[222] J. Gan, X. Zhang, H. Li, and H. Wu, "Full closed-loop controls of micro/nano positioning system with nonlinear hysteresis using micro-vision system," *Sens. Actuators A, Phys.*, vol. 257, pp. 125–133, Apr. 2017.
[223] H. Li, B. Zhu, X. Zhang, J. Wei, and S. Fatikow, "Pose sensing and servo control of the compliant nanopositioners based on microscopic vision," *IEEE Trans. Ind. Electron.*, vol. 68, no. 4, pp. 3324–3335, Apr. 2021.
[224] C. Ru *et al.*, "Automated four-point probe measurement of nanowires inside a scanning electron microscope," *IEEE Trans. Nanotechnol.*, vol. 10, no. 4, pp. 674–681, Jul. 2011.
[225] M. Xie, J. K. Mills, Y. Wang, M. Mahmoodi, and D. Sun, "Automated translational and rotational control of biological cells with a robot-aided optical tweezers manipulation system," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 2, pp. 543–551, Apr. 2016.
[226] F. Karimirad, B. Shirinzadeh, W. Yan, and S. Fatikow, "A vision-based methodology to dynamically track and describe cell deformation during cell micromanipulation," *Int. J. Optomechtron.*, vol. 7, no. 1, pp. 33–45, Feb. 2013.
[227] S. Zhuang, W. Lin, H. Gao, X. Shang, and L. Li, "Visual servoed zebrafish larva heart microinjection system," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 3727–3736, May 2017.
[228] C. Y. Wong and J. K. Mills, "Cell extraction automation in single cell surgery using the displacement method," *Biomed. Microdevices*, vol. 21, no. 3, p. 52, Sep. 2019.
[229] C. C. Wong *et al.*, "Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage," *Nature Biotechnol.*, vol. 28, no. 10, p. 1115, 2010.
[230] J. Liu *et al.*, "Robotic adherent cell injection for characterizing cell–cell communication," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 1, pp. 119–125, Jan. 2015.
[231] M. A. Greminger and B. J. Nelson, "Vision-based force measurement," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 3, pp. 290–298, Mar. 2004.
[232] Y. H. Anis, J. K. Mills, and W. L. Cleghorn, "Visual measurement of MEMS microassembly forces using template matching," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2006, pp. 275–280.
[233] X. Liu, Y. Sun, W. Wang, and B. M. Lansdorp, "Vision-based cellular force measurement using an elastic microfabricated device," *J. Micromech. Microeng.*, vol. 17, no. 7, p. 1281, 2007.
[234] F. Karimirad, S. Chauhan, and B. Shirinzadeh, "Vision-based force measurement using neural networks for biological cell microinjection," *J. Biomech.*, vol. 47, no. 5, pp. 1157–1163, Mar. 2014.
[235] D. J. Cappelleri, G. Piazza, and V. Kumar, "A two dimensional vision-based force sensor for microrobotic applications," *Sens. Actuators A, Phys.*, vol. 171, no. 2, pp. 340–351, Nov. 2011.
[236] Z. Lu, P. C. Y. Chen, and W. Lin, "Force sensing and control in micromanipulation," *IEEE Trans. Syst., Man Cybern. C, Appl. Rev.*, vol. 36, no. 6, pp. 713–724, Nov. 2006.
[237] A. Ferreira, C. Cassier, and S. Hirai, "Automatic microassembly system assisted by vision servoing and virtual reality," *IEEE/ASME Trans. Mechatronics*, vol. 9, no. 2, pp. 321–333, Jun. 2004.
[238] X. Liu, K. Kim, Y. Zhang, and Y. Sun, "Nanonewton force sensing and control in microrobotic cell manipulation," *Int. J. Robot. Res.*, vol. 28, no. 8, pp. 1065–1076, Aug. 2009.
[239] Y. Xie, D. Sun, C. Liu, H. Y. Tse, and S. H. Cheng, "A force control approach to a robot-assisted cell microinjection system," *Int. J. Robot. Res.*, vol. 29, no. 9, pp. 1222–1232, Aug. 2010.
[240] Q. Zhao, M. Sun, M. Cui, J. Yu, Y. Qin, and X. Zhao, "Robotic cell rotation based on the minimum rotation force," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 4, pp. 1504–1515, Oct. 2015.
[241] M. A. Sutton *et al.*, "Scanning electron microscopy for quantitative small and large deformation measurements part II: Experimental validation for magnifications from 200 to 10,000," *Experim. Mech.*, vol. 47, no. 6, pp. 789–804, Dec. 2007.
[242] R. Wu, H. Qian, and D. Zhang, "Robust full-field measurement considering rotation using digital image correlation," *Meas. Sci. Technol.*, vol. 27, no. 10, Oct. 2016, Art. no. 105002.
[243] B. Dong, C. Li, and B. Pan, "Ultrasensitive video extensometer using single-camera dual field-of-view telecentric imaging system," *Opt. Lett.*, vol. 44, no. 18, pp. 4499–4502, 2019.
[244] Y. Gao, T. Cheng, Y. Su, X. Xu, Y. Zhang, and Q. Zhang, "High-efficiency and high-accuracy digital image correlation for three-dimensional measurement," *Opt. Lasers Eng.*, vol. 65, pp. 73–80, Feb. 2015.
[245] R. Zhang and L. He, "Measurement of mixed-mode stress intensity factors using digital image correlation method," *Opt. Lasers Eng.*, vol. 50, no. 7, pp. 1001–1007, Jul. 2012.

Sheng Yao received the B.Eng. degree in mechatronics engineering from the South China University of Technology, Guangzhou, China, in 2016, where he is currently pursuing doctoral studies with the Guangdong Key Laboratory of Precision Equipment and Manufacturing Technology.

From 2018 to 2019, he studied in the Nonlinear Systems Control Laboratory, University of Toronto, Toronto, ON, Canada, as a joint Ph.D. student. His research interests include high-precision vision systems and vision-based micromanipulation, with applications to MEMS and cell surgery.

Hai Li received the Ph.D. degree in mechanical engineering from the South China University of Technology, Guangzhou, China, in 2018.

He spent one year as a Post-Doctoral Fellow with the Division for Microrobotics and Control Engineering (AMiR), University of Oldenburg, Oldenburg, Germany. He is currently an Assistant Professor at the South China University of Technology. His research interests include vision-based precision measurement and servo control, micropositioning/nanopositioning, and manipulation.

Shuiquan Pang received the B.Eng. degree from the South China University of Technology, Guangzhou, China, in 2017, where he is currently pursuing the Ph.D. degree with the Guangdong Key Laboratory of Precision Equipment and Manufacturing Technology.

His current research interests include vision-based microscale/nanoscale measurement, SEM image processing, and micromanipulation/nanomanipulation.

Benliang Zhu received the Ph.D. degree in mechanical engineering from the South China University of Technology, Guangzhou, China, in 2014.

He was a Guest Scholar at Kyoto University, Kyoto, Japan, supported by the JSPS Fellowship. He is currently an Associate Professor with the School of Mechanical and Automotive Engineering, South China University of Technology. His research mainly focuses on compliant mechanisms/tools and their applications in the micromanipulation/nanomanipulation systems, and other emerging fields.

Xianmin Zhang received the Ph.D. degree from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1993.

He has been the Director of the Guangdong Key Laboratory of Precision Equipment and Manufacturing Technology since 2010, and he has been the Dean and the Chair Professor of the School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou, China, since 2013.

Dr. Zhang has served as the Chair of the China Committee of the International Federation for the Promotion of Mechanism and Machine Science (IFToMM) from 2016 to 2020. He is a member of the IFToMM Executive Committee.

Sergej Fatikow received the Ph.D. degree in electrical engineering and computer science from Ufa State Aviation Technical University, Ufa, Russia, in 1988.

In 1990, he moved to the University of Karlsruhe, Karlsruhe, Germany, where he initiated the new research field of microrobotics. Since 2001, he has been a Full Professor with the Department of Computing Science and the Head of the Division for Microrobotics and Control Engineering (AMiR), University of Oldenburg, Oldenburg, Germany. His research interests include microrobotics and nanorobotics automation at the nanoscale, sensor feedback at the nanoscale, and neurofuzzy robot control.