2022 IEEE 22nd International Conference on Nanotechnology (NANO)
July 4-8, 2022
Palma de Mallorca, Spain

A real-time target tracking method for SEM images based on template matching

Liang Fang, Zhi Qu, Lue Zhang*, and Zhan Yang, Senior Member IEEE

***Abstract*— Micro-nano materials and their manufacture have become increasingly commonly used in the field of chip manufacturing as nanotechnology has advanced. In this area, the use of carbon nanotubes (CNTs) to build high-performance three-dimensional micro-nano devices has gotten a lot of interest. The micro-nano manipulation robot system built in the scanning electron microscope (SEM) is currently used to assemble carbon nanotube devices. Secondary electron detectors in SEMs are used to record secondary electron scanning pictures as a perceptual tool for automated assembly visual guidance. This paper describes how SEM images are used to provide visual guidance for the assembly of CNTs. SEM image noise sources and components were investigated. A carbon nanotube image tracking system based on template matching with scale invariance and rotation invariance was suggested, based on noise components and denoising. Its accuracy error is not more than 6 pixels and the frame rate reaches 26.39 FPS which meets the real-time requirements.**

I. Introduction

Carbon-based chips are predicted to become the favored option for the development of next-generation circuits as Moore's Law approaches its limits. [1] Due to their superior electrical qualities, quasi-one-dimensional scale, and stable structure, carbon-based chips constructed from carbon nanotubes (CNTs) are a preferable candidate for alternative silicon-based circuits. [2-3] The process of transferring and installing carbon nanotubes on electrodes is referred to as carbon nanochip assembly, and the interaction between CNTs and metal electrode surfaces induced by different assembly processes is also distinct. In general, the contact resistance is inversely proportional to the length of contact between CNTs and electrodes, and increased contact resistance reduces chip sensitivity. [4] As a result, the assembly process chosen will have a direct impact on the performance of carbon nanochips.

CNTs can be assembled using two different processes. Chemical vapor deposition, random dispersion, and dielectric electrophoresis are examples of classic assembly methods. The other is a robotic assembly method, which mainly relies on the micro-nano manipulation robot to achieve the assembly of CNTs.[5-6] The assembling and testing process will be affected by scale effect, channel effect, and other variables due to the size limitation of the equipment itself, the needs of correct assembly, and testing functions in the nano environment. The size of CNTs is at the nanoscale and there are clusters so the traditional method is difficult to meet the precise placement requirements of CNTs and the micro-nano manipulation robot system based on SEM can complete the manipulable tasks such as picking, handling, assembling, and measuring nanodevices, showing great application potential.

To a large extent, the micro-nano manipulation robot system built in the scanning electron microscope (SEM) at this stage requires the experimenter to observe the posture and position of the three-dimensional micro-nano device through the SEM image. Compared with automatic manipulation, manual manipulation not only consumes more time but also the probability of success depends on the proficiency of the experimenter. The manipulation without visual feedback and force feedback has a high probability of damaging micro-nano devices.[7-8] In addition, various interferences such as image noise and drift distortion are caused by the construction of the micro-nano manipulation robot system in the SEM. The image quality and magnification offered are also greatly reduced. These conditions are incompatible with the assembly of micro-nano devices. [9]

A lot of research has been done on the automation of nano-manipulation by many researchers. Sun Yu's team has developed a nanorobot system that realizes the automatic picking of individual nanowires. The system selects a suitable nanowire from a growth substrate, picks up a nanowire, and places it on a MEMS device.[10] Toshio Fukuda's team designed an automated CNT picking manipulation system that automatically binarized images by histogram thresholding to segment CNTs and atomic force microscopy (AFM) cantilevers and controlled the cantilever to automatically pick up CNTs.[11] Fatikow's team proposed a multi-target detection method based on template matching to detect RF wafers. This method can accurately measure the pose of multiple targets, even if the targets overlapped.[12] Different from the above methods, the real-time tracking of the target and online processing of micro-nano manipulation based on the SEM environment were focused on in our paper. First of all, according to the characteristics of SEM imaging, the edge noise reduction of the SEM image was processed. Secondly, a similarity matrix was added to the template matching based on the gray-scale similarity measure to identify the rotation and transformation of the target tracking at different magnifications. Finally, the image pyramid sampling method was applied to realize the online manipulation of the above method and enhance the frame rate. The commonly used setups and the main process are shown in Figure.1. The electrode materials of the chip are Cr and Au, the substrate material is silica, and CNTs are used instead of silicon-based materials as the conductive trenches.

* Research supported by National Key Research and Development Program of China, grant number 2018YFB1304901 and National Natural Science Funds of China (U1813211).

Liang Fang, Zhan Yang and Lue Zhang* are with School of Mechanical and Electric Engineering, Soochow University, Suzhou, China.

Zhi Qu is with School of Biomedical Engineering, City University of Hong Kong, Hongkong, China.

2022 IEEE 22nd International Conference on Nanotechnology (NANO) | 978-1-6654-5225-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/NANO54668.2022.9928655

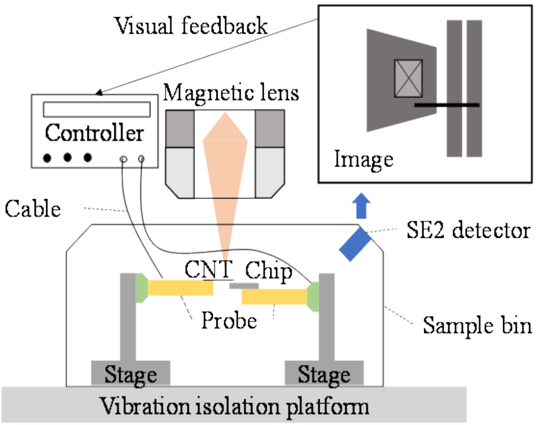

Figure.1 The commonly used setups and the main process for CNTs assembling. Setups: the SEM, the micro-nano manipulation robot system, the CNTs and the chips.

II. SEM Image Noise Cause And Component Analysis

The image of the micro-nano sample is collected using a SEM (Carl Zeiss MERLIN Compact SEM) in this paper. The electron beam scans the sample point by point in the SEM. The external environment interferes with the imaging process. The SEM in this paper incorporates a symmetrically distributed micro-nano manipulation robot system, which comprises of four manipulation units, each of which is comprised of a SmarAct three-axis micro-movement platform (SmarAct, SLC-1720-s) and a piezoelectric actuator Picmotor (New Focus Inc., 8301-UHV). The addition of a micro-nano manipulation robot system introduces new thermal, electromagnetic, and vibration interference, which is reflected in the image as a superposition of several noises.

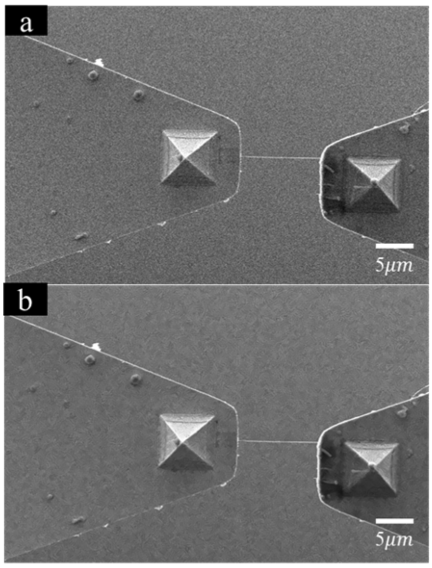

Figure. 2 (a) Original image before filtering;

(b) Image after filtering;

Noise has been analyzed in our previous work. [13] For diverse micro-nano experimental material samples, Zeiss MERLIN Compact SEM and FEI Helios G4 SEM were utilized to acquire experimental image samples at varying working distances, acceleration voltages, magnifications, and electronic detectors, respectively. Thermal noise and electromagnetic interference from electronic components, as well as environmental vibration, are the principal sources of image noise in scanning electron microscopes. In the image, thermal noise produces Gaussian noise, while electromagnetic interference and ambient vibration provide Poisson noise. A convolutional neural network was used to filter SEM pictures for different noises, and it was able to filter out Gaussian and Poisson noise directionally.

III. Target Tracking Based On Template Matching

The template matching method is widely used. The target search can be regarded as the process of locating the template in the image to be matched, and the two are judged according to different features, such as gray level, feature point, gradient, and other measurement methods. The SEM picture is a single-channel grayscale image with a small and fast grayscale-based feature computation.

A. Similarity Calculation

To realize the tracking of the target in the SEM image, the normalized cross-correlation-based (NCC) template matching method has been used, which calculates the similarity between the template image and the matched image by comparing the similarity of the gray values of the two images.[14] NCC is a normalized similarity measurement method, which can effectively reduce the influence of light intensity changes on the measurement results. In other words, it avoids matching errors due to changes in image pixels. The candidate window refers to the coverage region in the image to be matched that is equivalent to the size of the template. A candidate window is created each time the template moves a position in the matched image. The formula NCC is as follows:

$$NCC(x_w^j, x_t) = \frac{\sum_{x=0}^{W_{t-1}} \sum_{y=0}^{H_{t-1}} x_w^j(u_j+x\ v_j+y) \bullet x_t(x,y)}{\left[\sum_{x=0}^{W_{t-1}} \sum_{y=0}^{H_{t-1}} (x_w^j(u_j+x\ v_j+y))^2 \bullet \sum_{x=0}^{W_{t-1}} \sum_{y=0}^{H_{t-1}} (x_t(x,y))^2 \right]^{0.5}} \tag{1}$$

where X_w^j represents the jth candidate window. Correspondingly, the candidate window with the best matching similarity is called the best matching window. For the matched image P with the size of $Wp \times Hp$, and the template Xt with the size of $Wt \times Ht$, usually $Wp > Wt$, $Hp > Ht$, where the upper left corner of the candidate window X_w^j is located in the (uj, vj) of the matched image at the place. The formula NCC represents the pixel matching degree between the candidate window and the template, and its value range is "-1, 1". When applied, take the absolute value, 1 means a high degree of similarity, and 0 means a low degree of similarity.

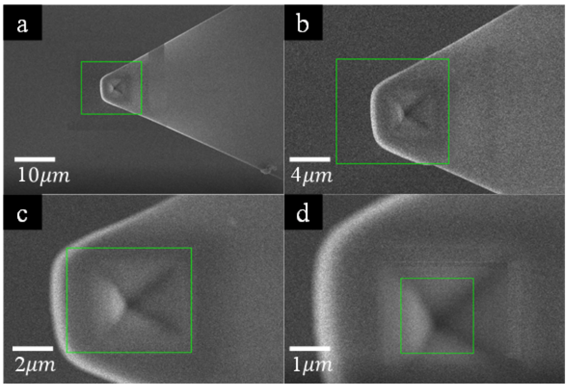

Figure 3 Selection of ROIs at different magnifications in (a) 1.0k (b) 2.5k (c) 5k (d) 10k.

To verify the stability of target tracking, images with a resolution size of 1024 × 768 at 1thousand, 2.5thousand, 5 thousand, and 10 thousand magnifications. were selected respectively, the scan speed is 4, and region of interest(ROI) was selected as the probe's pyramid area, as shown in Figure 3.

The maximum pixel error in the X-axis direction is 2 pixels, and the maximum pixel error in the y-axis is 6 pixels. Since the actual distance represented by a pixel is related to the magnification and image resolution, increasing the resolution of the image can effectively reduce errors while keeping the magnification unchanged.

B. Handling rotation and scale

In the micro-nano manipulation, the translational motion of three degrees of freedom, and rotational motion are also involved and the magnification is adjusted to obtain the optimal size of the target in the field of view. Traditional template matching can only match templates with the same angle and size. If the pose of the template and the target change, it is difficult to match correctly. As a result, rotation and scale invariance are required to follow the position of the object in the field of vision in real time. The similarity transformation matrix based on NCC is introduced, which can be expressed as the following formula:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} s\cos\theta & s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \qquad (2)$$

The similarity transformation is composed of translation transformation (t), rotation transformation (R), scale scaling (S), etc., with 4 degrees of freedom, which can meet the translational motion of the X-Y plane and the rotational motion of the Z-axis of the micro-nano manipulation robot, and the SEM perspective scale scaling requirements.

C. Search Strategy: Pyramid Sampling

To speed up the algorithm, the search strategy of template matching is needed. The traditional search method is to calculate all the candidate windows one by one and select the window with the highest similarity as the best matching window. This method is only suitable for images with smaller sizes. When the input image size is larger, the search time increases significantly. To this end, the image pyramid search method is used. First, pyramid downsampling is performed on the input matched image and the template image. Make the original image the bottommost image. The Gaussian kernel function to convolve it, and then downsample the convolved image to obtain the image of the previous layer, repeat the above convolution and downsampling, and obtain a pyramid-shaped multi-layer image. The number of pixels in each layer from the bottom to the top keeps decreasing and becomes rougher. The Gaussian smoothing filter is performed so that the down-sampled pixels can better represent the pixels of the original image. Starting from a set of images in the top layer, a complete matching is performed to find the best position for the current layer. Then map the result of the top-level matching to the next layer of the pyramid. The search range of 5x5 at the matching points of the map is expended and then a local matching calculation is performed. At this time, only 25 small matching calculations are required. After the calculation is completed, the extreme value of this small area is extracted as the best matching point of this layer and this process is repeated until the bottom layer of the pyramid (the original search image).

The number of layers of the pyramid has a direct impact on the efficiency of the entire algorithm. The maximum frame rate of our SEM output video is 10, to achieve real-time effects, the time to process a frame of an image should not exceed 100 ms and the time of User Datagram Protocol(UDP) transmits coordinates is negligible. The image resolution and algorithm affect the speed of image processing. Aside from the image resolution, the relationship between the number of layers of the pyramid and the calculation time was shown in the following table 1:

Table.1 The relationship between the running time and the number of pyramid levels

Pyramid levels	0	1	2	3
NCC	667.22	153.58	37.89	8.98
SIFT+FLANN	198.57	61.93	25.93	Not enough feature points to match

The SIFT+FLANN algorithm is selected to compare with our method because it ensures the accuracy and at the same time the speed can be guaranteed. When the pyramid selects more than one layer, it can meet the real-time requirements. As the number of pyramid layers increases, the matching points gradually decrease so that they cannot be matched. The NCC algorithm requires at least two layers of pyramids to accelerate. When two layers are selected, the running time is 37.89 ms. It can meet real-time requirements. The running time is reduced to 8.98 ms when a 3-layer pyramid is used. In our experiments, although the running time of the 3-layer pyramid is further reduced, its stability is not as high as the 2-layer pyramid match. In the case of ensuring matching accuracy, a 2-level pyramid has been chosen for acceleration.

IV. Experiments And Results

To verify the effectiveness and performance of the method, the experiment of placing the CNTs on the chip by

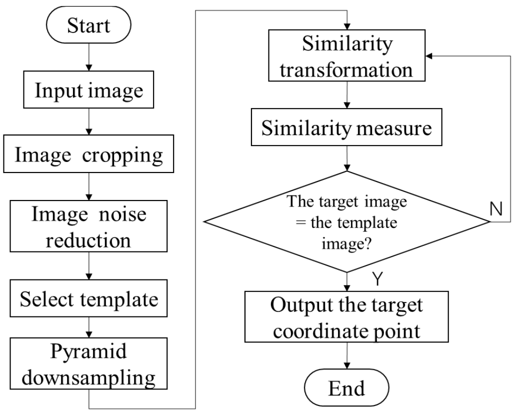

Figure 4. A target tracking framework based on template matching

using an AFM probe is taken. Transmitting the real-time images collected by the SEM through UDP and feeding back the real-time position coordinates of the probe to the Simulink. The software used in this experiment is visual studio 2019. The local language is C++. The whole method flow is shown in figure 4.

The video of this experiment was obtained by the smartSEM software. The video collected is composed of a group of consecutive SEM image frames. The resolution is 1024 × 768 and the FPS is 10. The scanning speed is 4, the acceleration voltage is 5 kV, and the SE2 detector is used to capture images.

The process of placing CNT on the chip was shown. Since the CNT fits the probe under the van der Waals force, the probe with more obvious features was selected as the ROI as shown in Fig. 5(a). The probe was moved upward close to the chip and then dropped in depth until the CNT made contact with the electrode. The CNT was moved in the horizontal direction and when the CNT is deformed by the van der Waals force, the contact is successful. The trajectories for the obtained pixel coordinates were shown in Figure.5 (b).

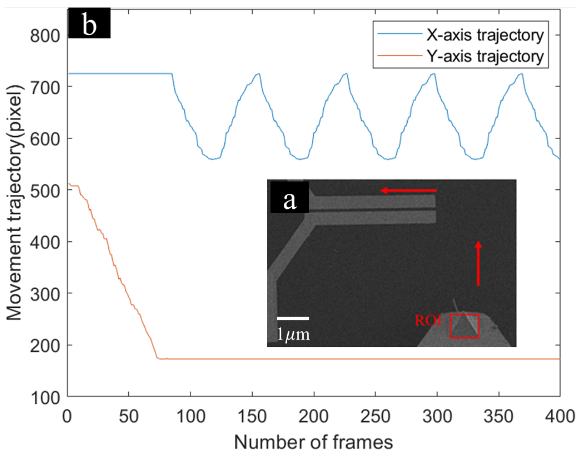

Figure 5. (a) CNT placement diagram

(b) traced trajectory

V. Conclusion

Based on the structure and imaging principle of SEM, the causes of noise in secondary electron images and backscattered electron images were discussed. An appropriate filter for different noises was selected to filter them. For target tracking, the NCC-based template matching method was employed, with similarity transformation added to account for rotation invariance and scale scaling. Pyramid downsampling accelerates the method to attain real-time processing. The target coordinates will be created to the Dynamic Movement Primitives (DMP) motion trajectory and paired with Simulink to build a servo feedback controller to achieve the automatic construction, cutting, picking, handling, and other manipulation of CNTs in future study.

Acknowledgment

* Research supported by National Key Research and Development Program of China, grant number 2018YFB1304901, and National Natural Science Funds of China (U1813211).

References

[1] W. Zhou and Z. L. Wang, Scanning microscopy for nanotechnology: techniques and applications. Springer science & business media, 2007..

[2] M. M. Waldrop, "More than moore," Nature, vol. 530, no. 7589, pp. 144-148, 2016.

[3] M. Hartmann et al., "CNTFET technology for RF applications: Review and future perspective," IEEE Journal of Microwaves, vol. 1, no. 1, pp. 275-287, 2021.

[4] Z. Yang et al., "Mechatronic development and vision feedback control of a nanorobotics manipulation system inside SEM for nanodevice assembly," Sensors, vol. 16, no. 9, p. 1479, 2016.

[5] T. Fukuda, F. Arai, and L. Dong, "Assembly of nanodevices with carbon nanotubes through nanorobotic manipulations," Proceedings of the IEEE, vol. 91, no. 11, pp. 1803-1818, 2003.

[6] C. Shi et al., "Recent advances in nanorobotic manipulation inside scanning electron microscopes," Microsystems & Nanoengineering, vol. 2, no. 1, pp. 1-16, 2016.

[7] S. Frégonèse, C. Maneux, and T. Zimmer, "A compact model for dual-gate one-dimensional FET: Application to carbon-nanotube FETs," IEEE transactions on electron devices, vol. 58, no. 1, pp. 206-215, 2010.

[8] A. D. Franklin et al., "Carbon nanotube complementary wrap-gate transistors," Nano letters, vol. 13, no. 6, pp. 2490-2495, 2013.

[9] Y. Ouyang and J. Guo, "Assessment of carbon nanotube array transistors: A three-dimensional quantum simulation," Solid-state electronics, vol. 61, no. 1, pp. 18-22, 2011.

[10] X. Ye, Y. Zhang, C. Ru, J. Luo, S. Xie, and Y. Sun, "Automated pick-place of silicon nanowires," IEEE Transactions on Automation Science and Engineering, vol. 10, no. 3, pp. 554-561, 2013.

[11] L. Dong, T. Fukuda, and F. Arai, "Nanorobotic systems," International Journal of Advanced Robotic Systems, vol. 2, no. 3, pp. 264-275, 2008.

[12] 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 2019 International Conference on Manipulation, Automation and Robotics at Small Scales (MARSS), 2019, pp. 1-6: IEEE.

[13] Z. Qu, Z. Yang, and C. Ru, "Edges Detection of Nanowires and Adaptively Denoising with Deep Convolutional Neural Network from SEM Images," in 2020 IEEE 20th International Conference on Nanotechnology (IEEE-NANO), 2020, pp. 146-149: IEEE.

[14] J. N. Sarvaiya, S. Patnaik, and S. Bombaywala, "Image Registration by Template Matching Using Normalized Cross-Correlation," presented at the 2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies, 2009.