Optical Engineering 58(5), 053102 (May 2019)

Evaluation method of ensquared energy based on dynamic factors in gene sequencing system

Zhiyuan Sun[a] and Wanyu Li[b,*]
[a]Chinese Academy of Sciences, Changchun Institute of Optics, Fine Mechanics, and Physics, Changchun, China
[b]First Hospital of Jilin University, Hepatology Department, Changchun, Jilin Province, China

Abstract. In a gene sequencing system, ensquared energy is an important parameter for base-calling software. Nevertheless, the displacement between wafer and camera, which is caused by vibration and position error of the wafer stage while photographing, will decrease the image quality and further influence the ensquared energy. Our study established the relationship model between the dynamic factors and ensquared energy. Several experiments were carried out on the sequencing platform. The results revealed that the ensquared energy has a linear relationship with the location standard deviation of dynamic errors. In the sequencing platform described, if the ensquared energy value required is larger than 65%, the location standard deviation should be controlled below 140 nm. Furthermore, the ensquared energy calculated by a DNA nanoballs image is accordant with the conclusion in the meantime. The performance parameters of the wafer stage and vibration suppression could be drawn up reasonably. [DOI: 10.1117/1.OE.58.5.053102]

Keywords: gene sequencing; fluorescence microscopy; ensquared energy; dynamic factors.

Paper 190055 received Jan. 14, 2019; accepted for publication Apr. 30, 2019; published online May 20, 2019.

1 Introduction

The most common technologies of second generation sequencing rely on optical detection of fluorescent substance.[1,2] The fluorescent dyes attached on DNA nanoballs (DNBs) are activated by lasers, and then the excited fluorescence illuminates the detector to form the DNB images. A base with a quality score is called for a given DNB spot by special base-calling software. Ensquared energy[3,4] is an important factor that influences the quality score correlated with base accuracy. If the ensquared energy is small, the energy distribution is dispersive, crosstalk and background intensity will decrease the quality score. Generally speaking, the ensquared energy should be larger than 65% for precise base calling. Therefore, it should be focused on the factors that influence the ensquared energy.

Mathematical properties of the encircled and ensquared energy functions for the diffraction-limited point-spread function (PSF) were presented by Torben.[5] In recent years, requirements for ensquared energy have been an important part of the system requirements for various imaging instruments.[6–8] However, the research mentioned above is all related to static imaging. Few studies focus on the relation between ensquared energy and dynamic parameters.

In the sequencing field, in order to acquire high-sequencing throughput, the core is to increase scanning speed of the wafer stage. The two main operating modes of the wafer stage are uniform motion mode and stepping motion mode. In uniform motion mode, the vibration caused by the wafer stage is small while the galvanometer should be added into system to compensate the uniform motion of the wafer stage,[9] which greatly increases the system complexity and assembly difficulty. As a consequence, more systems prefer stepping motion mode. In stepping motion mode, two dynamic parameters should be focused on.

First, one dynamic factor is the vibration caused by the wafer stage. If the wafer stage moves to the next field after it finishes capturing the previous position, it will experience acceleration and deceleration. As vibrations have certain attenuation period, residual vibration will exist in photographing period. These leftover vibrations will bring out image shift and thus cut down ensquared energy.

Second, position error of the wafer stage also should be carefully considered. To gain optimal sequencing throughput, the stabilizing process of the wafer stage should not continue to the status that no position error exit. Therefore, in the photographing period, the image shift caused by position error of the wafer stage will reduce ensquared energy as well.

In this study, we presented the influence of dynamic parameters on the ensquared energy. In Sec. 2, we described the sequencing technology centered on a hybridization and ligation method. In Sec. 3, the model between dynamic parameters and ensquared energy was established. Meanwhile, the composition of sequencing platform and the displacement measurement interferometer (DMI) system was also described in this section. In Sec. 4, the DNB imaging experiments were carried through, and conclusions are drawn in Sec. 5.

2 Background

The sequencing technology described in this study is centered on a hybridization and ligation method and relies on optical detection of fluorescent events.[10] The schematic flow diagram[11] is shown in Fig. 1.

First, sequencing fragments are prepared by sonication of genomic DNA followed by a series of repeated adapter site insertions, template circularization, restriction enzyme scission, and polymerase amplification. In the end, DNBs are generated. Second, DNBs are selectively attached to a coated

*Address all correspondence to Wanyu Li, E-mail: liwanyu2006@163.com

0091-3286/2019/$25.00 © 2019 SPIE

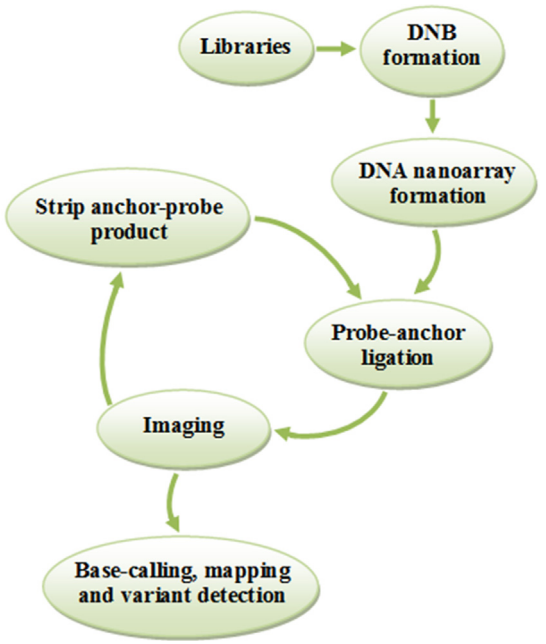


Fig. 1 Schematic flow diagram of the sequencing process.

silicon chip that is patterned with active sites to generate the DNB array chip.[12] Third, a library of common probes is used in combination with standard anchors and extended anchors to perform an unchained hybridization and ligation assay. Each hybridization and ligation cycle is followed by fluorescent imaging of the DNB spotted chip and subsequently regeneration of the DNBs. This cycle is repeated until the entire combinatorial library of probes and anchors is examined. Finally, base-calling, mapping, assembly, and analysis software are implemented to rapidly reconstruct genomes from billions of paired-end reads.

The optical platform realizes the rapid reading of submicron DNA nanoarrays. The imaging process based on the optical platform is shown in Fig. 2. The light emitted by laser is reflected by beam-splitters and enters into objective, followed by the light focused on the wafer. The fluorescent dyes attached on DNBs are motivated by lasers to emit fluorescence. The excited fluorescence transfers to objective, and then passes through filter, tube and illuminates the detector of the camera. In order to achieve high-throughput sequencing, the wafer, which is loaded with high-density DNB array, is scanned by a three-dimensional (3-D) motion platform. Among the three axes, two axes are employed to realize horizontal sweep, and the other axis cooperate with the autofocus module to realize the optimal focal plane detecting during scanning process.

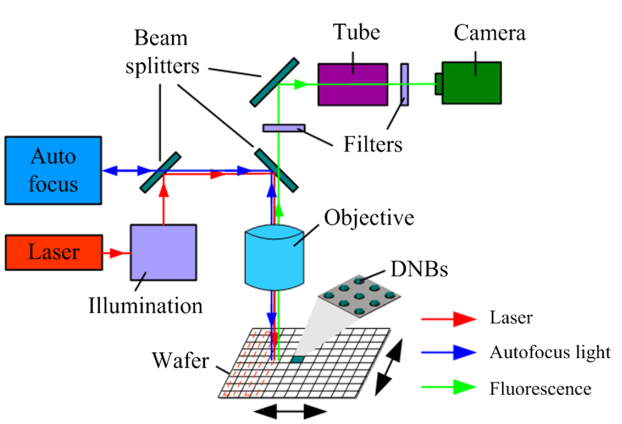


Fig. 2 The imaging process based on optical platform.

3 Description of the Method

3.1 *Process of Image Degradation*

The image quality of the sequencing system is influenced by the aberration of the optical system, laser thermal influence, the image shift aroused by stepping motion, and so on. All these factors will cause image blur and deformation, which is image degradation.

It is assumed that degenerative system $h(x, y)$ is the space invariant linear system, and then the degenerated model of continuous function could be expressed as

$$g = S(f * h) + n, \tag{1}$$

where $f(x, y)$ is the target image, $h(x, y)$ is the degraded model, * is convolution process, S is the sampling course, $n(x, y)$ is the noise function, and $g(x, y)$ is the degraded image. The degradation model is shown in Fig. 3.

In statics imaging condition, the degraded model $h(x, y)$ mainly refers to PSF of the optical system. Nevertheless, when in dynamic imaging condition, image shift caused by moving components will degrade image quality together with PSF of the optical system.

3.2 *Modeling Establishment*

In this section, it is set forth the physical process that how the dynamic parameters influence the ensquared energy.

Proprietary techniques could be used to control DNBs size, density, and binding affinity to wafer. Therefore, DNBs on wafer have approximately identical size and luminance. One DNB is taken out to implement the analysis. Without considering all factors that could lead to degradation, but considering the magnification of the optical system, the theoretical DNB (energy normalized to 1) in the image space is established.

The degenerate function includes the static factors and dynamic factors. The statics factors refer to the aberration of the optical system and heat effect of laser and could be evaluated by PSF. On the other hand, vibration and position error of the wafer stage form dynamic errors, which would cause the image shift in the photographing cycle. First, the statics influence is considered in the analysis, the convolution of the theoretical DNB model and PSF of the optical system will form a blurry spot on the camera. Second, owing to dynamic factors, image shift exists in a sampling cycle. It is supported that the photographing time is T and

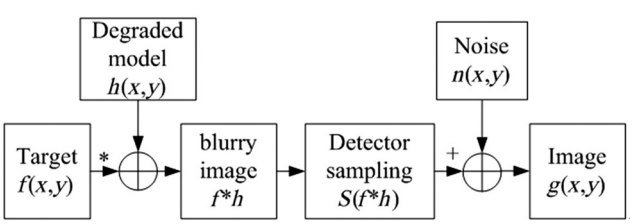


Fig. 3 Model of image degeneration.

the sample rate of the wafer stage is f_{stage}, then the frames numbers captured in a sampling cycle is $n = (T \times f_{\text{stage}})$. Therefore, in a photographing cycle, the blurry DNB image is the superimposed of these N blurry spots. The dispersed DNB image matrix could be expressed as

$$R_{\text{image}} = \sum_{i=1}^{n} f(l_{xi}, l_{yi}) * \text{PSF}, \tag{2}$$

where $f()$ represents the theoretical DNB model, PSF is point spread function of the optical system, $L_x = (l_{x1}, l_{x2}, \cdots, l_{xn})$ and $L_Y = (l_{y1}, l_{y2}, \cdots, l_{yn})$ are the offset matrices between DNB and camera in X and Y scanning direction, respectively. L_x and L_Y could be obtained by DMI system or linear encoders.

After limited sampling of detector, the energy of continuous blurry spots gathered by the discrete pixels in the nine block box. Therefore, the ensquared energy could be calculated by the grayscale of nine pixels in nine block box. It is important to note that the calculation of ensquared energy should take account the offset between the center of nine block box and the centroid of DNB image.

In a photographing cycle, the locations of DNB are discrete. As described before, the distribution of blurry spot in the image space is related to the locations of N frames in one sampling cycle. Therefore, the location standard deviation (STD for short) of N frames captured in one sampling cycle should be figured out to manifest the dispersion degree, which is shown as

$$\text{STD} = \sqrt{\frac{\sum_{i=1}^{N} (l_{xi} - \overline{l_x})^2 + \sum_{i=1}^{N} (l_{yi} - \overline{l_y})^2}{N-1}}, \tag{3}$$

where $\overline{l_x} = \frac{\sum_{i=1}^{N} l_{xi}}{N}$ and $\overline{l_y} = \frac{\sum_{i=1}^{N} l_{yi}}{N}$.

The entire algorithm of flowchart is shown in Fig. 4. The relationship model between offset in one sampling cycle and ensquared energy is built up. In other words, the model between dynamic parameters and ensquared energy has been established.

3.3 *Sequencing Platform*

In order to verify the influence of dynamic parameters on ensquared energy, several experiments were carried out on sequencing platform, shown in Fig. 5. The sequencing platform consists of marble, motion stage for 3-D translation, objective, illumination module, beam splitter module, tube module, autofocus module, camera, and so on. Horizontal motion stage employs Aerotech air-bearing linear stage. Three beam splitter modules split four-ways laser beams, autofocus beam, and fluorescence beams, and several filters are installed in the platform to reduce the laser influence on fluorescence. Meanwhile, four complementary metal oxide semiconductor (CMOS) cameras are used to capture fluorescence signals in different wave bands. In order to reduce the external vibration influence, the instrument is placed on a site with VC-D vibration criterion.[13]

3.4 *Dynamic Parameters Measurement Based on DMI System*

The dynamic parameters happened in stepper motion progress are measured by means of DMI system, shown in Fig. 6.

The laser is split into two paths by a 50% beam splitter. Two plane mirrors are placed on the upper platform of marble and the Z wafer stage, respectively. The displacement of the two plane mirrors is measured by high-stability plane mirror interferometer in the respective light path while the wafer stage acts as step motion. After recombining and data processing of the two mirrors displacement signals, the relative displacement between the two mirrors is obtained. Since the optical modules are rigidly connected on the upper platform, the result of DMI measurement is actually the relative displacement between camera and wafer. The measurement results include the vibration of the wafer stage and the position error of the wafer stage.

4 Results

4.1 *Static Images*

First, the sequencing wafer is imaged at static condition. The aberration of imaging optical system and error caused by laser thermal influence will affect the image quality of static

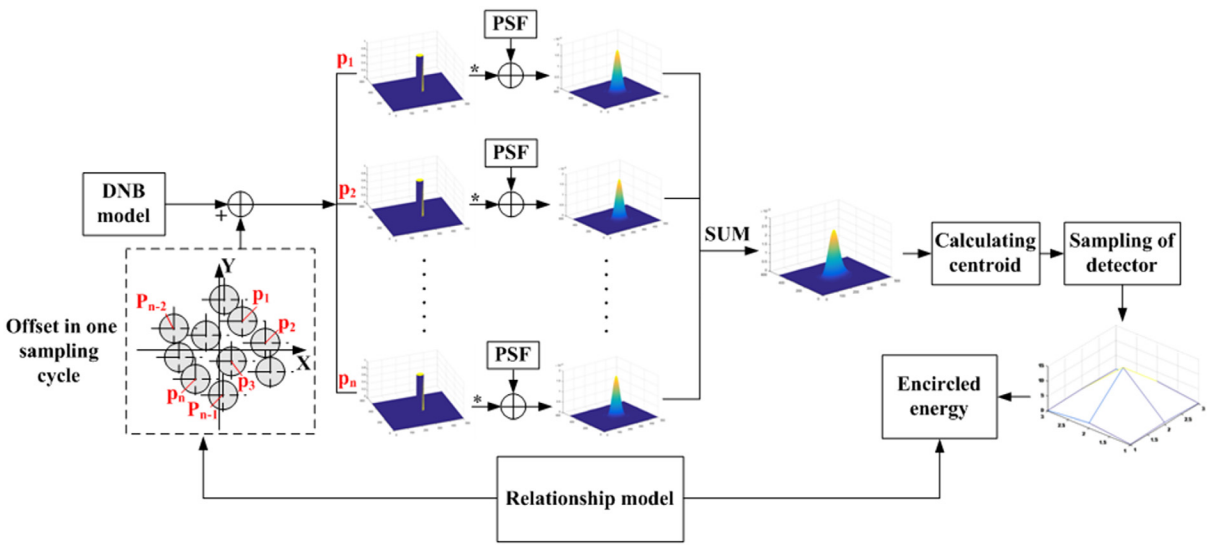


Fig. 4 Algorithm of flowchart.

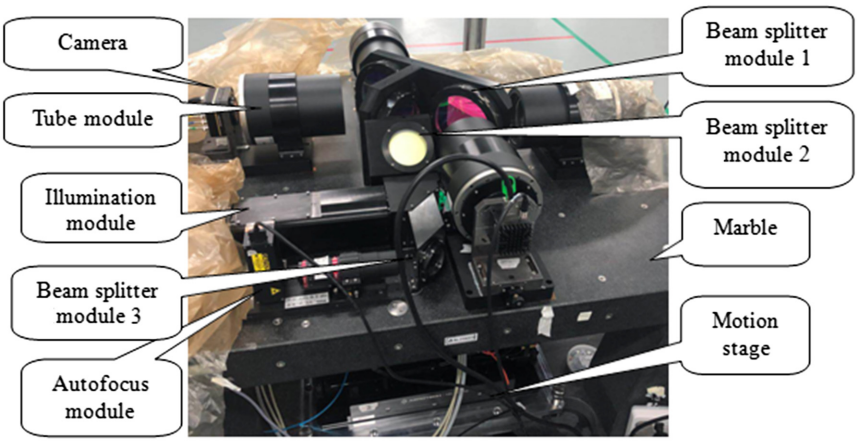


Fig. 5 Photograph of sequencing platform.

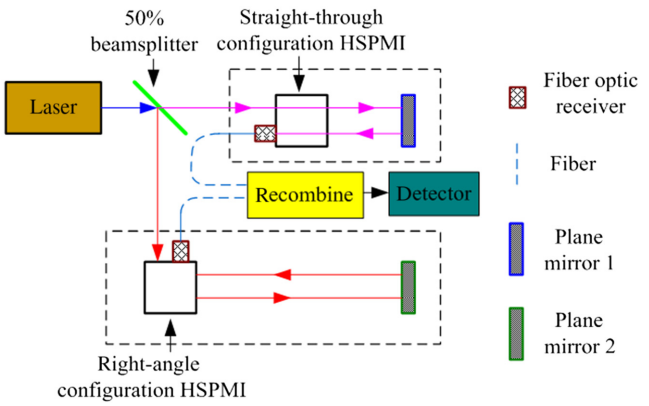
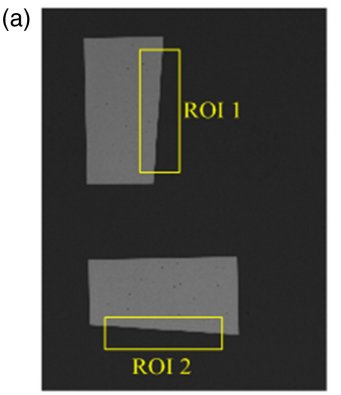


Fig. 6 Dynamic errors measurement schematic based on DMI system.



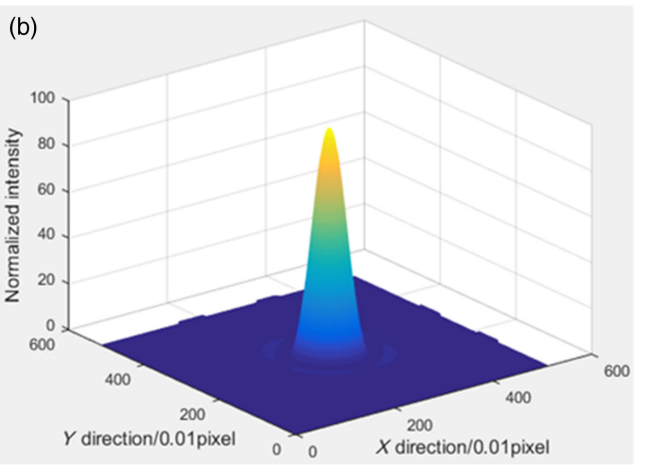


Fig. 7 PSF measurement of sequencing system: (a) image of processing chip and (b) PSF of the sequencing system.

imaging. Among these, the aberrations of imaging optical system mainly refer to the principle errors of the optical system, the optical design errors, and various errors in manufacturing. The factors described above all belong to system error, which could be expressed by PSF of the sequencing system.

A processing chip used for evaluating PSF is placed on the chuck above the wafer stage. The processing chip is focused on the optimal focal plane with the help of autofocus module. Partial image of the processing chip is shown in Fig. 7(a). Two regions of interest are selected to calculate the PSF in orthogonal direction. By means of knife-edge method[14] and Fermi function, the PSF result is shown in Fig. 7(b).

In this system, the image device is CMV50000 CMOS sensor, with the pixel size 4.6×4.6 μm. The ensquared energy is calculated in the region of 3×3 pixels, and the area is 13.8×13.8 μm accordingly.

Afterward, the biological wafer is absorbed on the chuck for DNB imaging. The images of DNB are photographed at the optimal focal plane, which is shown in Fig. 8. It has been calculated that the ensquared energy is about 73% at static status.

4.2 *Dynamic Images*

Second, the wafer stage moves in stepping mode in dynamic imaging condition. The overall period is 75 ms. The photographing period is 15 ms, and residual time consists of motion period and setting time. The sample frequency of the wafer stage is 2000 Hz, so the frames numbers captured in a sampling cycle is 30.

In addition to the factors described before in static image, the vibration that is caused by the wafer stage stepping motion and the position error that happened at the photographing cycle will bring out image shift, which will influence the dynamic images quality.

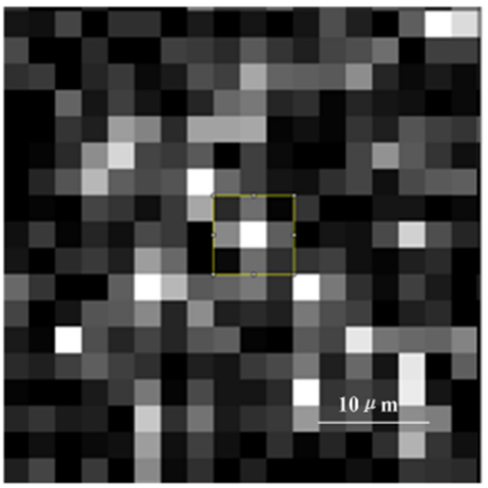


Fig. 8 The DNB image captured at static condition.

In order to acquire the statistical property, about 800 fields were scanned by the wafer stage. When the wafer stage is carried through stepping motion, the displacement between the camera and wafer stage in two scanning direction were measured by the DMI systems.

The location STD distribution histogram of all fields is shown in Fig. 9. Meanwhile, the proportions that different ranges account for are listed in Table 1.

From Fig. 9 and Table 1, 87% images have the location STD between 60 and 160. According to the model described in Sec. 3, the corresponding ensquared energies at different ranges are calculated, the result is shown in Fig. 10.

From Fig. 10, it is indicated that the ensquared energy has approximately linear relationship with the location STD of X and Y offsets. As the location STD increases, the ensquared energy decreases. In this sequencing system, if the ensquared energy value required is larger than 65%, the location STD should be controlled below 140 nm.

One image of which the location STD is between 120 and 140 nm is selected to calculate the ensquared energy. The corresponding DNB figure is shown in Fig. 11. After calculating, the ensquared energy is 66.44%. The ensquared energy value conforms to the statistics result.

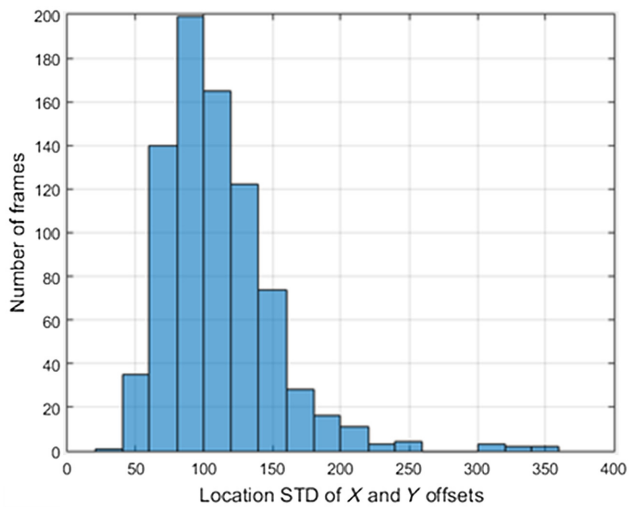


Fig. 9 The distribution of location STD.

Table 1 Proportion of different ranges.

Range of location STD (nm)	Numbers relative to all frames (%)
<60	4.47
60 to 80	17.39
80 to 100	24.72
100 to 120	20.5
120 to 140	15.16
140 to 160	9.19
>160	8.57

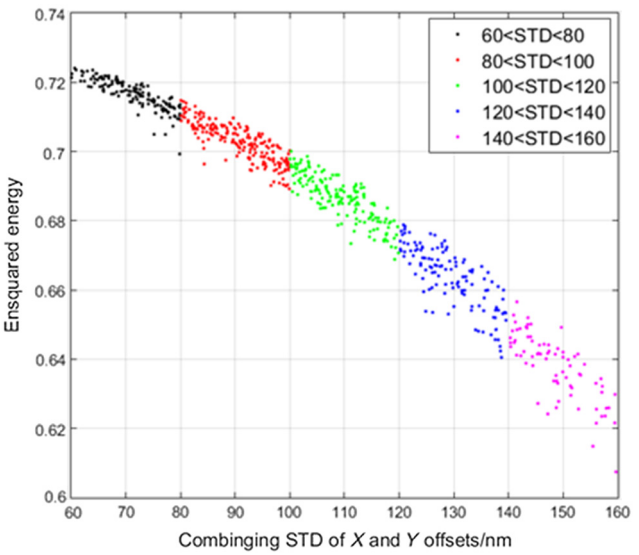


Fig. 10 The ensquared energy results.

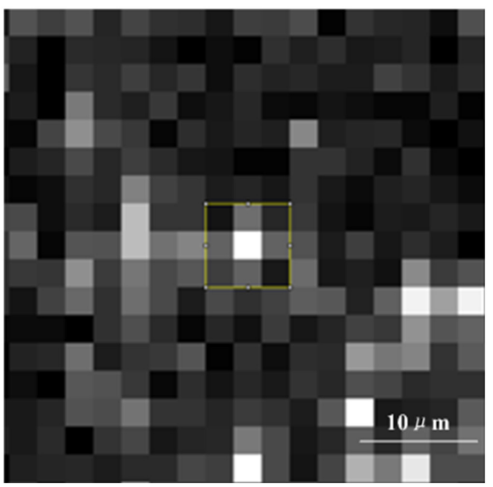


Fig. 11 The DNB image captured at dynamic status.

5 Discussion

In a gene sequencing system, ensquared energy is a crucial parameter for base-calling software to recognize the DNA bases. A model between dynamic parameters and ensquared

energy is built up in this study. Several imaging experiments are carried through in the sequencing platform. First, the static elements that mainly include optical system aberration and laser thermal influence are calibrated by the processing chip. The PSF of the sequencing system is calculated by means of the knife-edge feature on the processing chip. Second, the DMI systems are set up to measure the displacement between the wafer stage and camera at about 800 times motions. Therefore, the dynamic factors that primarily contain the vibration caused by the wafer stage stepping motion and the position error happened at the photographing cycle are measured. Finally, the PSF and the dynamic errors data are used for calculating the ensquared energy at different dynamic performances. The statistics results indicated that the ensquared energy has approximately linear relationship with the location STD. In this sequencing system, if the ensquared energy value required is larger than 65%, the location STD should be controlled below 140 nm. The ensquared energy calculated by DNB image is accordant with the conclusion in the meantime.

This work established the model between dynamic parameters and ensquared energy. Applying the technique, we can conclude how dynamic parameters will meet the ensquared energy demand. The dynamic performance parameters of the wafer stage and vibration suppression grade could be drawn up reasonably by this study. Meanwhile, error allocation in the sequencing system is facilitated.

Acknowledgments

This study was supported by grants from the First Hospital of Jilin University (No. JDYYJC006), Tianqing Liver Disease Research Fund (No. TQGB20180160), Department of Science and Technology of Jilin Province (No. 20180520116JH). The authors declare that they have no conflict of interest.

References

1. M. L. Metzker, "Sequencing technologies—the next generation," *Nat. Rev. Genet.* **11**(1), 31–46 (2010).
2. R. H. Deurenberg et al., "Application of next generation sequencing in clinical microbiology and infection prevention," *J. Biotechnol.* **243**, 16–24 (2017).
3. J. M. Nichols and C. Miller, "Analytical expression for the average ensquared energy," *J. Opt. Soc. Am. A* **32**(4), 654–659 (2015).
4. W. J. Smith, *Modern Optical Engineering*, McGraw-Hill, New York (2000).
5. B. A. Torben, "Accurate calculation of diffraction-limited encircled and ensquared energy," *Appl. Opt.* **54**(25), 7525–7533 (2015).
6. R. A. David et al., "Performance modeling for the RAVEN multi-object adaptive optics demonstrator," *Publ. Astron. Soc. Pac.* **124**, 469–484 (2012).
7. A. G. Basden, C. J. Evans, and T. J. Morris, "Wide-field adaptive optics performance in cosmological deep fields for multi-object spectroscopy with the European Extremely Large Telescope," *Mon. Not. R. Astron. Soc.* **445**(4), 4008–4014 (2014).
8. P. Zhang et al., "Impact of point spread function on infrared radiances from geostationary satellites," *IEEE Trans. Geosci. Remote Sens.* **44**, 2176–2183 (2006).
9. L. A. Shaw et al., "Scanning two-photon continuous flow lithography for synthesis of high-resolution 3-D microparticles," *Opt. Express* **26**(10), 13543–13548 (2018).
10. R. Drmanac et al., "Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays," *Science* **327**(5961), 78–81 (2010).
11. T. P. Niedringhaus et al., "Landscape of next-generation sequencing technologies," *Anal. Chem.* **83**(12), 4327–4341 (2011).
12. J. H. Han et al., "Electrophoretic build-up of multi nanoparticle array for a highly sensitive immunoassay," *Biosens. Bioelectron.* **41**(1), 302–308 (2013).
13. K. A. Salyards and L. M. Hanagan, "Evaluation of vibration assessment criteria and their application to stadium serviceability," *J. Perform. Constr. Facil.* **24**, 100–107 (2010).
14. H. Li, C. Yan, and J. Shao, "Measurement of the modulation transfer function of infrared imaging system by modified slant edge method," *J. Opt. Soc. Korea* **20**(3), 381–388 (2016).

Zhiyuan Sun received his BS degree in mechanical manufacture and automation from the Northeastern University in 2003 and his PhD in optical engineering from the Changchun Institute of Optics, Fine Mechanics, and Physics (CIOMP) in 2008. He is an associated professor at CIOMP. He is the author of more than 10 journal papers and has written 4 patents. His current research interests include fluorescent microscopy and gene sequencing system.

Wanyu Li received her BS degree in clinical medicine in 2005 and her PhD in hepatology from Jilin University in 2011. She is an associated professor at the first hospital of Jilin University. She is the author of more than 20 journal papers and has written 5 book chapters. Her current research interests are liver immunology and genetics.