23

Applications of Computer Vision in Micro/Nano Observation

Yangjie Wei[1], Chengdong Wu[2] and Zaili Dong[3]
[1]*Graduate School of Chinese Academy of Sciences & The State Key Laboratory of Robotics,Institute of Automation, Chinese Academy of Sciences*
[2]*School of Information Science & Engineering, Northeast University*
[3]*State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences,*
China

1. Introduction

Nowadays, micro/nano science and technology has been one of the most attractive research fields. However, real time and accurate observation in micro/nano manipulation is a top important enabling technique. Most recently, with the great development of microscopes and computer vision techniques, real time visualization, including 2D motion measurement and 3D reconstruction, on micro/nano scale is becoming possible.

As for 2D motion measurement, visual motion measurement on micro/nano scale is still an open problem. Many researchers have designed different algorithms, and most of them are based on a block matching algorithm, which locates matching blocks in a researched digital image for the purposes of distance or similarity estimation. Usually, block matching based methods can achieve better performances when the texture is not relevant, or the aliasing problem in the derivative estimation, which is caused by the large inter-frame displacements (Giachetti & Torre, 1996). Images, however, are typically processed assuming a uniform grid of pixels. While straightforward, the uniform grid representation does not scale well in a multi-scale setting, because it requires an excessive amount of refinement to capture small details in a image, including sub-pixel resolution. The motion to be estimated is, on most situations in micro/nano manipulation, small and not integer. Therefore, it is necessary to improve the existing algorithms and obtain higher precision not limited by the pixel dimension, i.e., sub-pixel motion estimation. In 1989, Anandan reaches a sub-pixel precision by locally approximating the difference function with a quadratic surface and published (Horn, 1986; Horn & Schunck,1981; Singh, 1990), however, the sub-pixel estimation resolution usually introduces more computational burden.

As far as 3D reconstruction is concerned, depth measurement, i.e., methods to attain 3D information from 2D images, is an important research field in computer vision, and now it has been one of the key techniques in many fields, such as medicine, robotics, remote sensing and micro/nano manipulation. In recent years, there are various 3D reconstruction methods, including volumetric methods, depth from stereo (DFS), depth from focus (DFF) and depth from defocus (DFD)(Yin,1999), researched and used in real applications.

Volumetric methods usually reconstruct 3D models of external anatomical structures from 2D images. They represent the final volume using a finite set of 3D geometric primitives. Then, from an image sequence acquired around the object to reconstruct, the images are calibrated and the 3D models of the referred object are built using different approaches of volumetric methods. These methods work in the object volumetric space and do not require a matching process between the images used. Thus, typically, the 3D models are built from a sequence of images, acquired using a turntable device and an off-the shelf camera (Teresa et al,in press, 2008). However, in some real applications, we do not need to reconstruct the 3D model of objects, because depth is enough to understand the 3D relationship of scenes.

DFS estimates depth from two images of the same scene captured by cameras at different positions and with different postures (Wu, 1999). Because it needs to extract and match feature points in these images, the computational task is so huge. As for DFF, it uses a mapping relation between focus and depth to estimate depth. It obtains a sequence of images with different depth, measures the focus degree using a measurement operator (Bove 1993; Nayar,1992), and attains the desired depth when the measurement value is maximal or minimal. Compared to DFS, DFF is simple in principle, but its estimation accuracy is highly related to the number of images.

DFD is first introduced by Pentland in 1987(Pentland,1987). It has been proved to be an effective depth reconstruction method by using the concept of blurring degree of region images with limit depth of field (Girod & Scherock, 1989; Pentland et al, 1994; Navar et al, 1996). Usually, DFD algorithm captures two images obtained with different camera parameters, measures blurring degree of every point, and estimates depth using the point spread function. During the past years, DFD has become attractive because 1) it requires only two images; 2) it avoids matching and masking problems; 3) it is effective both in the frequency domain and in the spatial domain (Gokstorp,1994; Subbarao & Surya,1994). However, since all above DFD methods need to capture two defocused images with changed camera parameters, they can not be used in applications with high level magnification microscopes, such as micro/nano manipulation, because on these situations, it is destructive to change camera parameters. This is the main reason why DFD has not been used in micro/nano manipulation until now.

2. 2D motion measurement

Computer vision is one of the most important techniques used in motion measurements, especially 2D motion measurement, because the instruments used in computer vision are comparatively cheaper, the measurement process is simple and the result is direct. In recent years, with the development of revolution and sensitivity on visual sensors, the measurement scale of computer vision has reached micro/nano scale.

Block matching method (BMA) is one of the most widely applied methods to compute the visual 2D motion from images, i.e. to estimate the 2D motion projected on the image plane by the objects moving in the 3D scene, as it is less susceptible to a random error source than edge based or image moment methods.

2.1 BMA method

The foundational principle of BAM is to find a matching block from an image X in some other image Y, which may appear before or after X, and through comparing them to

measure difference, such as distance or similarity, between two images. Therefore to select a criteria to determine whether a given block in image Y matches the search block in image X is top important, i.e., the object function.

BMA based techniques usually can be divided into two classes according to the measurement criterion, including the minimal difference and the maximal similarity. The widely used object functions based on difference measurements include Sum-of-Squared-Differences (SSD), Sum-of-Absolute-Differences (SAD) which can transferred into Local-SAD (LSAD) when its intensity is locally scaled and Zero-SAD (ZSAD) with setting the average gray level difference equal to zero. If the difference minimum is replaced by the maximum of a correlation measurement, some object functions can be got, such as Normalized-Cross-Correlation (NCC)(Qi & Michale,1987), Approximate-Maximum-Direct-Correlation (AMDC)(Kim & Meng,2007), or some other variations those are all approximate maximum likelihood estimators(Robinson & Milanfar,2004).

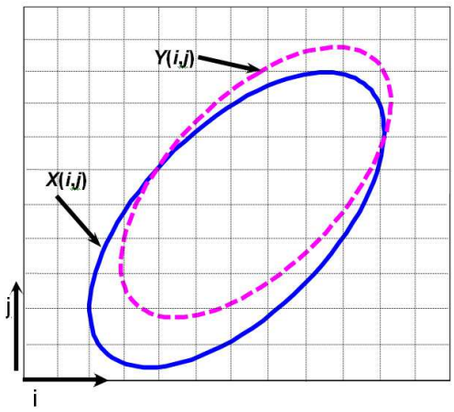


Fig. 1. Motion of the continuous image $F(i,j)$ with respect to the pixel grid

Here, SSD, LSAD, ZSAD and NCC are all adopted to estimate the motion between two neighbor images in a same image sequence. The theory is shown in Fig. 1. Generally, $X(i,j)$ and $Y(i,j)$ are referred to as the model image and the target image respectively. $F(i,j)$ is the continuous image function, $\varepsilon_{x,y}$ represents additive noise, $s=(s_x, s_y)$ is the shift between the model image and the target image.

$$X_{i,j} = F(i,j) + \varepsilon_x \tag{1}$$

$$Y_{i,j} = F(i - s_x, j - s_y) + \varepsilon_y \tag{2}$$

The object functions for SSD, LSAD, ZSAD and NCC are respectively defined as follow,

$$R(u,v)_{SSD} = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [x_{i,j} - y_{i+u,j+v}]^2 \tag{3}$$

$$R(u,v)_{LSAD} = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left| x_{i,j} - \frac{\overline{X_{i,j}}}{\overline{Y_{i+u,j+v}}} y_{i+u,j+v} \right| \tag{4}$$

$$R(u,v)_{ZSAD} = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [x_{i,j} - \overline{X_{i,j}} - y_{i+u,j+v} + \overline{Y_{i+u,j+v}}] \tag{5}$$

$$R(u,v)_{NCC} = \frac{\frac{1}{n \times m} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} X_{i,j} Y_{i+u,j+v}}{\sqrt{\frac{1}{n \times m} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} X_{i,j}^2} \sqrt{\frac{1}{n \times m} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} Y_{i+u,j+v}^2}} \tag{6}$$

$$X_{i,j} = x_{i,j} - \frac{1}{n \times m} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} x_{i,j} = x_{i,j} - \overline{X_{i,j}} \tag{7}$$

$$\begin{aligned} Y_{i+u,j+v} &= y_{i+u,j+v} - \frac{1}{n \times m} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} y_{i+u,j+v} \\ &= y_{i+u,j+v} - \overline{Y_{i+u.j+v}} \end{aligned} \tag{8}$$

where $x_{i,j}$, $y_{i,j}$ are the original gray intensity of each point in the model and target image respectively, $\overline{X_{i,j}}, \overline{Y_{i,j}}$ are the mean of each image inside their respective "block", u, v are coordinates of the model image block, and $R(u ,v)$ is the object function between the model block and the target block. The shift $s = (s_x, s_y)$ is estimated by finding the peak of the objective function.

The shift between the model image and the target image can be denoted as,

$$s = s_n + s_\Delta \tag{9}$$

where $s_n = (n_x, n_y)$ is the integer shift and $s_\Delta = (\varepsilon_x, \varepsilon_y)$ is the sub-pixel shift. If the evaluation step is one pixel, s_n can be obtained from,

$$R(n_x, n_y) = \max\{R(u,v)\} \tag{10}$$

However, the peak position above can only be solved with pixel-level accuracy, and it is not enough in many applications, especially in micro/nano manipulation. In order to attain the sub-pixel resolution, a quadratic curve fitting around the peak s_n is usually used to estimate the sub-pixel shift s_Δ as follows,

$$f(x) = a_x x^2 + b_x x + c_x \tag{11}$$

$$f(y) = a_y y^2 + b_y y + c_y \tag{12}$$

where a_x, b_x, c_x, a_y, b_y, c_y are coefficients of the quadratic curves along x and y axis, which are the parameters to be estimated. The shift $s_\Delta = (\varepsilon_x, \varepsilon_y)$ is estimated by finding the peak of the following $f(x)$ and $f(y)$.

$$\widetilde{\varepsilon_x} = \max(f(x)) = 2a_x x + b_x \tag{13}$$

$$\widetilde{\varepsilon_y} = \max(f(y)) = 2a_y y + b_y \tag{14}$$

If three points are used, the estimation results of the shifts can be denoted,

$$\widetilde{\varepsilon_x} = \frac{R'(-1,0) - R'(1,0)}{2[R'(-1,0) + R'(1,0) - 2R'(0,0)]} \tag{15}$$

$$\widetilde{\varepsilon_x} = \frac{R'(0,-1) - R'(0,1)}{2[R'(0,-1) + R'(0,1) - 2R'(0,0)]} \tag{16}$$

where $R'(u - n_x, v - n_y)=R(u, v)$.

However, the precision of the method mentioned above is usually low as it has no additional points to be used. Thus, one simple and effective way to improve the estimation precision is to properly increase the interpolated points and the order of the fitting polynomials.

2.2 Improved block matching algorithm

2.2.1 The searching region

As is known, the searching region is the main factor which influences the computational cost and affects the performance of the BMA. Thus, in this section, with respect to an image sequence, an improved method is proposed to reduce the searching region effectively.

Since our aim here is to estimate the shift in the image sequence, and generally the motion is small between two neighbor images, it is not necessary to calculate $R(u, v)$ with blocks throughout the whole image. Furthermore, the computation procedure not only increases the computational burden but also adds some opportunities of wrong matching when the texture or gray level of the target image is very similar.

Assuming that the largest shift is known, our proposed new BMA can be denoted as following steps,

- First, define the initial position**.** Since the shift between every two images is very small, it is reasonable to take the position of the block in the model image as the initial position in the target image.
- Second, define the maximal distance that the block can move in target image as $step_{max\_x}$ and $step_{max\_y}$, where $step_{max\_x}$ and $step_{max\_y}$ can be determined through experience.
- Third, move step by step along x and y axis and calculate $R(u, v)$ at each step. During this process, the total distance along x axis should be always equal or lesser than $step_{max\_x}$, the total distance along y axis should be always equal or lesser than $step_{max\_y}$.
- Finally, find out the minimal or maximal value of the object function and the corresponding moving steps.

It is clear that using the improved algorithm the computational burden can be greatly reduced because of the reduced searching region. Besides, if the whole image is similar in texture and gray level, or the block is very small, the improved method can also effectively decrease casual matching errors.

2.2.2 The block size

In BMA, the size of a block is another important factor to influence the matching precision and the complexity. If the block is too small, the matching information is limited and it would influence the matching precision greatly. While if the block is too large, the deformation during movement cannot be omitted. Besides, the larger block can result in a large computational burden. Therefore, to research the relationship between the block size and the object function is meaningful.

A standard grid used as the target object is shown in Fig.2, the magnification of the microscope's objective lens is $60\times$, which means that the underlying image is very smooth. The black squares designate the "blocks". First, we tested the matching error of different object functions on different block sizes. The results are shown in Fig.3, where the vertical axis denotes the estimation errors, with unit of pixel; the horizontal axis denotes the block size, with unit of pixels.

1. The estimation error of NCC is sensitive to the block size and has no obvious rule. Therefore if one wants to estimate the sub-pixel shift between two images by using NCC, the block size should not be the main regulating parameter.
2. The estimation error of LSAD is sensitive to the block size too, the larger the block size is selected, the smaller the estimation errors. Thus, with respect to the LSAD sub-pixel estimation, it is often beneficial to increase the block size.
3. As far as other functions are concerned, including SSD and ZSAD, they have the similar estimation results which can achieve much smaller errors comparatively, and they are not sensitive to block size as NCC and LSAD.
4. As far as the standard grid block, the most appropriate size is 50×50 pixels, which includes an integral corner of a white grid and a little black background. It proves this kind of block both can express the internal difference of the block approximately and can eliminate the coupling between x and y direction.

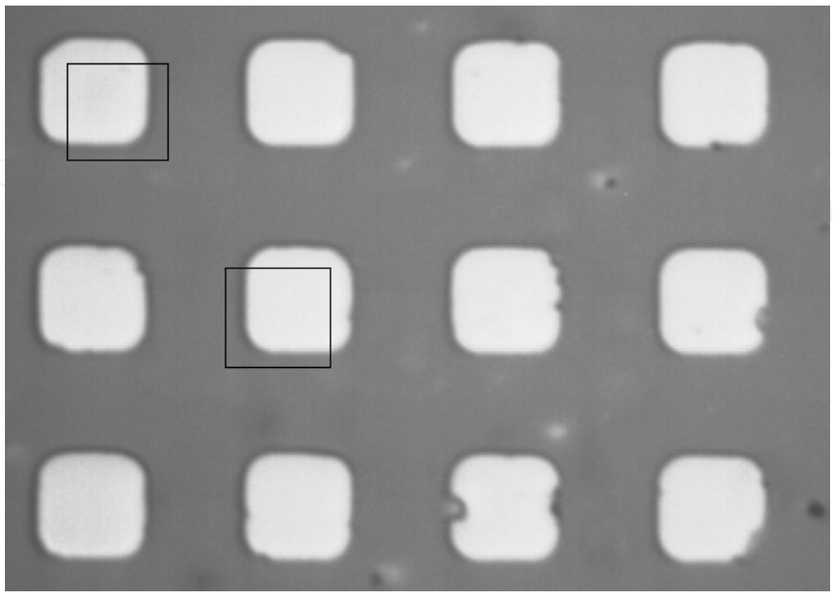


Fig. 2. The standard grid micrograph

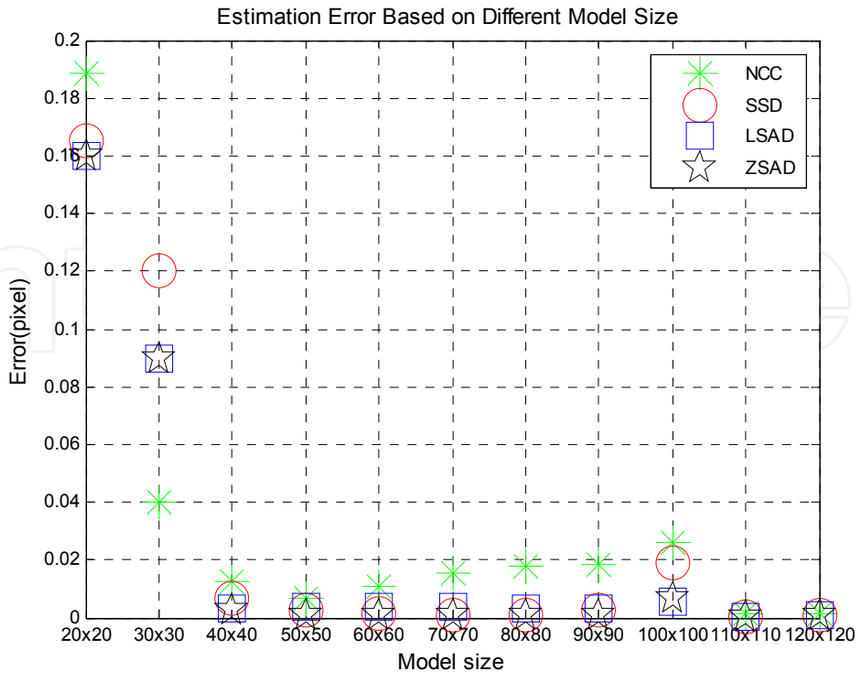


Fig. 3. Estimation errors with different block size

2.2.3 The sub-pixel fitting precision

Third, it is well known that estimation error of the curve fitting may become smaller when the order of a fitting polynomial increases. In this experiment, the comparison of the sub-pixel shift estimation error among different fitting polynomials including quadratic curve, cubic curve, quartic curve and spline curve, was conducted. Here, we used the same five fitting points, where the middle point is the peak and the other four points are on its both sides evenly. The main results are shown as Fig.5 to Fig.8, from which the following results can be obtained,

1. No matter what objective functions are used, the estimation error of the quadratic curve is always larger than the original method because of redundancy. While due to high smoothness, the spline curve can achieve an optimal result for all objective functions.
2. For NCC and SSD, the sub-pixel estimation is better enough when the fitting function is the cubic curve. That means, if the higher order fitting function is selected, the improvement in the estimation precision is unclear compared with the introduced computational burden. As far as the LSAD and ZSAD are concerned, the estimation precision improvement of the quartic fitting function is obvious. Thus, it can be concluded that the objective function is the main factor to decide the order of fitting equation.
3. Since the computational burden has been greatly reduced by using the new BMA proposed in section 2.2.1, much higher precision can be achieved by using higher order fitting functions and larger number of fitting points. That can be properly selected based on the preceding conclusions in real applications.

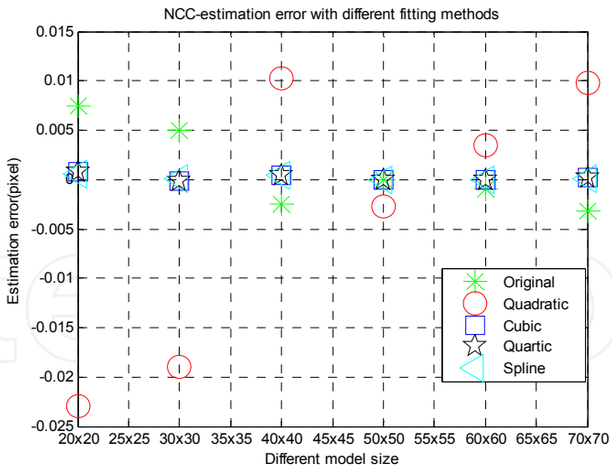


Fig. 4. Estimation errors with NCC

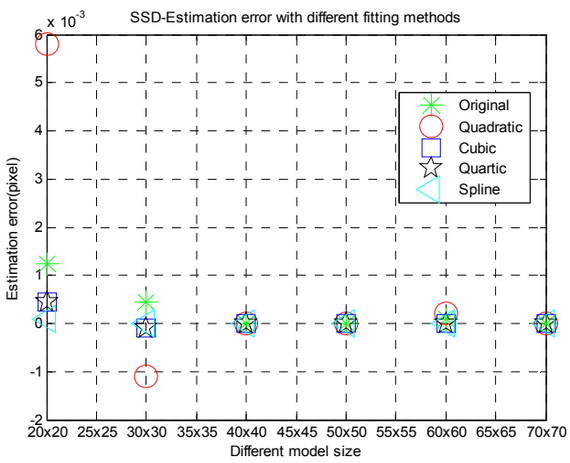


Fig. 5. Estimation error with SSD

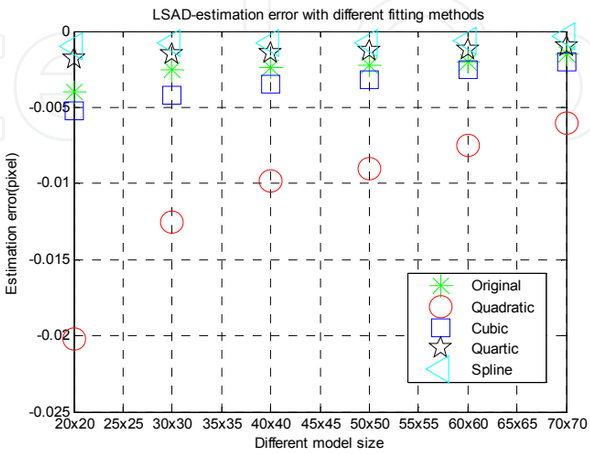


Fig. 6. Estimation error with LSAD

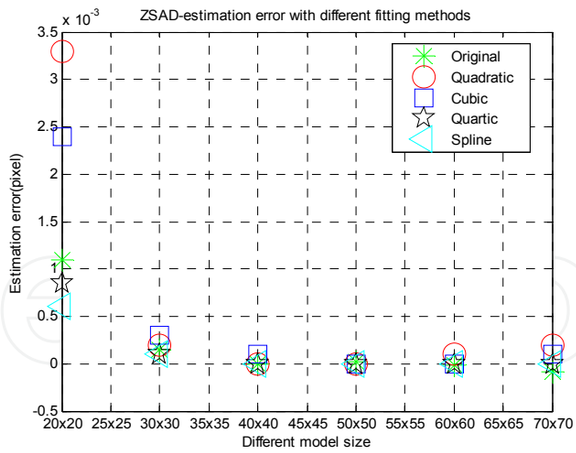


Fig. 7. Estimation error with ZSAD

Based on the theory and experiment results in 2.2, if we want to attain a matching result with high precision and low computational task using the grid block, the following parameters should be chosen: the reduced searching region, ZSAD object function, 50×50 pixels block size and the quartic sub-pixel fitting function.

2.3 The driving characteristic of a piezoelectric actuator

In order to validate the precision of our improved BAM practically, we used the PI nano platform to control the motion of the standard grid and calculated it with our method. Because the platform can output the nano motion, the shift of the grid is known. The KH-7700 microscope was used to capture the images of the grid, and the shift of each step was 50nm. Since the horizontal pixel is 57.47nm, the practical calculation result should be 0.89 pixels. Fig.9 is the result of the integral shift and Fig.10 is the shift of sub-pixel where the vertical axis denotes the motion, with unit of pixel, and the horizontal axis denotes the shift steps; the line with "*" is the true movement and the line with "o" is the calculation movement.

From Fig.9-10, we can see that the precision of our improved sub-pixel motion measurement method is very high, the integral pixel measurement is exactly equal to the true value and the sub-pixel measurement result is close to the true value. Therefore, the method can be used to measure the practical shifts in micro/nano manipulation.

Then, the sub-pixel block matching method of displacement measurement based on computer vision was used to measure the driving characteristic curve of a piezoelectric actuator practically. The result was shown in Fig.11, where the vertical axis denotes the motion, with unit of nm, and the horizontal axis denotes the driving voltage, with unit of V. Fig.11(a) is the driving curve when the voltage increases to 200V and then decreases to 0V smoothly; Fig.11(b) is the driving curve when the voltage increases to 150V and then decreases to 0V; Fig.11(c) is the driving curve when the changing routine of the voltage is 0V-200V-0V, 0V-150V-0V, 0V-100V-0V, 0V-50V-0V; Fig.11(d) is the driving curve when the changing routine of the voltage is 0V-200V-0V, 0V-160V-0V, 0V-120V-0V, 0V-80V-0V, 0V-40V-0V.

The measurement results of the piezoelectric actuator driving characteristic are consistent with the physics analysis. Furthermore, the proposed method, which is simple in manipulation and credible in measurement results, satisfies the requirement of the micro/nano measurement with high precision.

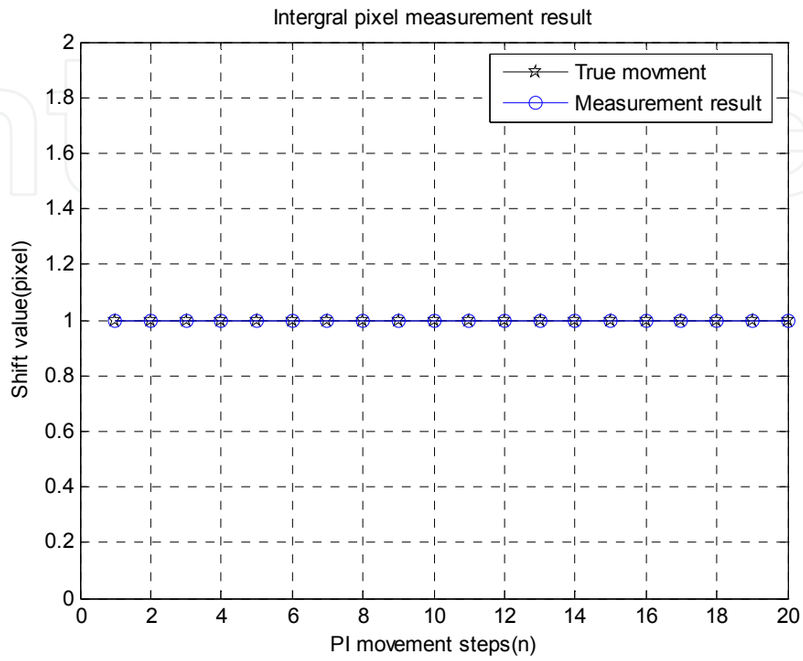


Fig. 8. The integral pixel measurement result

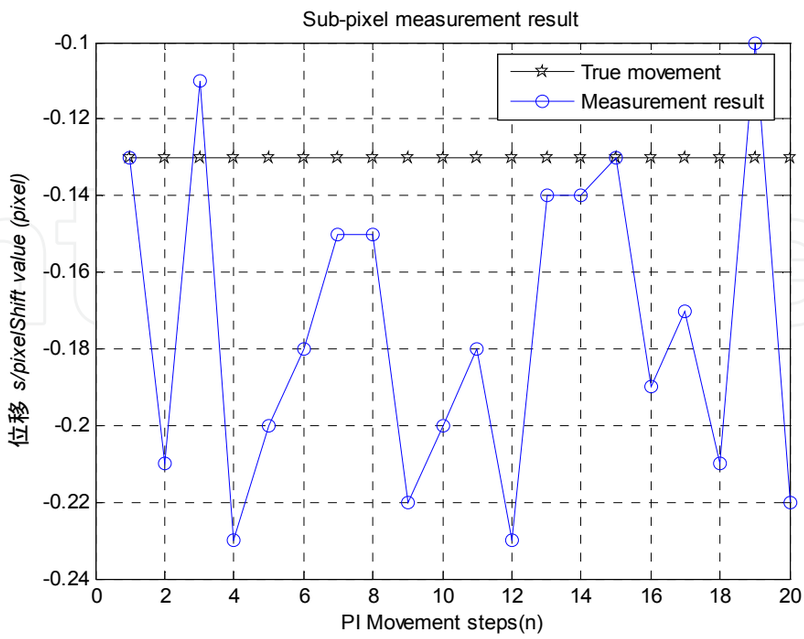
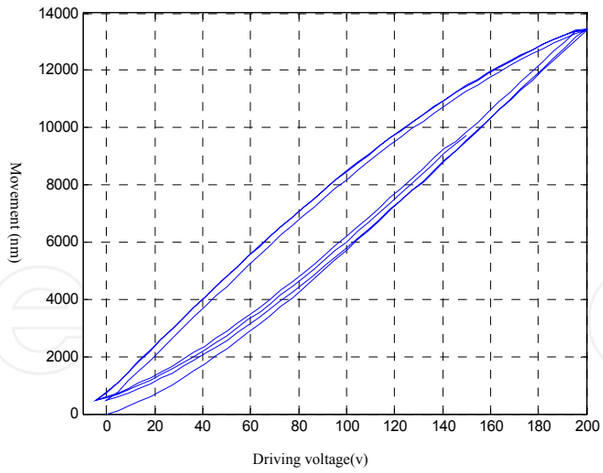
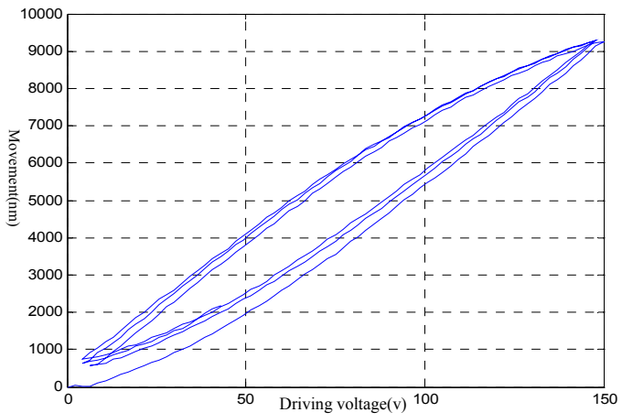


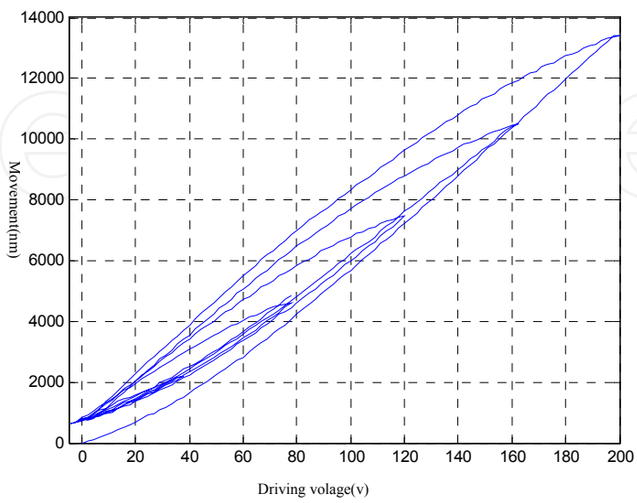
Fig. 9. The sub- pixel measurement result



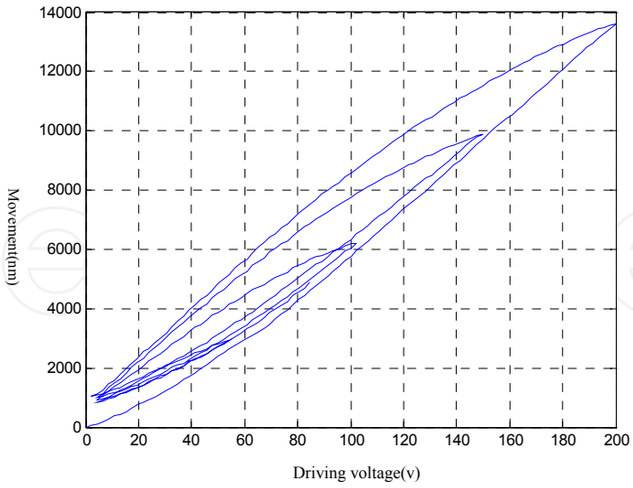
(a) 0V-200V-0V



(b) 0V-150V-0V



(c) 0V-200V-0V, 0V-150V-0V, 0V-100V-0V, 0V-50V-0V



(d) 0V-200V-0V, 0V-160V-0V, 0V-120V-0V, 0V-80V-0V, 0V-40V-0V

Fig. 10. The driving characteristic of a piezoelectric actuator

3. 3D reconstruction

Shape, or depth profile, reconstruction, is based on measurement depth information from 2D images, and now it has been widely used in many fields, such as medicine, robotics, and remote sensing.

All existing DFD algorithms can be divided into two kinds, local DFD algorithm and global DFD algorithm. In local DFD, a window around every pixel point is predefined, and the point's blurring is defined as that of the window (Pentland,1987; Vinay & Subhasis,2007). However, the difficulty of selecting proper size of window is a well known disadvantage of DFD algorithm, because there is a trade-off between having a window that is as large as possible to average out noise, but as small as possible to guarantee that within it (Ens &Lawrence ,1993; Nair & Stewart,1992). As far as global DFD is concerned, its main idea is completely different with local DFD algorithm since it works on the entire image without information of its radiance, or the appearance of the surfaces, and depth. Therefore, it is necessary to construct the depth model and the radiance model simultaneously (Favaro et al 2008, 2003 2002). This, however, will bring the problem of huge computation cost. A general method to solve this problem is to simplify the imaging model, for example, assuming the scene contains "sharp edges", that is, there are discontinuities in the scene (Asada et al 1998). Another way is to use a cubic function or structure light to approach the radiance (Nayar et al 1996; Lagnado & Osher,1997). Unfortunately, both local DFD and global DFD are on the basis of attaining two defocused images with different camera parameters which may destroy the camera drastically if the camera's amplification level is high.

In this section a novel DFD method with single fixed optical microscope is proposed to reconstruct the shape of samples on micro/nano scale. In the method, the blurring image model is constructed with the relative blurring and the diffusion equation, and the relation between depth and blurring is discussed from four aspects. The method proposed needs

only one microscope with unchanged camera parameters, so the reconstruction process is very simple. The experiments and error analysis results show that it can reconstruct shape on micr/nano scale.

3.1 The imaging model for defocus

In the defocus imaging model, a defocused image can be theoretically considered as the summation of some defocused points, and this process can be denoted by the following convolution function normally:

$$E(x,y) = I(x,y) * h(x,y) \tag{17}$$

where $E(x, y)$ and $I(x, y)$ are the defocused image and the focused image, respectively, $h(x, y)$ is the point spread function.

When the point spread function is approximated by a shift-invariant Gauss function, the imaging model in Eq.(17) can be formulated in terms of the isotropic heat equation:

$$\begin{cases} \dot{u}(x,y,t) = a\Delta u(x,y,t)\ a \in [0,\infty)\ t \in (0,\infty) \\ u(x,y,0) = E(x,y) \end{cases} \tag{18}$$

where a is the diffusion coefficient, $\dot{u} \doteq \frac{\partial u}{\partial t}$, "$\Delta$"denotes the Laplacian operator, $\Delta u = \frac{\partial^2 u(x,y,t)}{\partial x^2} + \frac{\partial^2 u(x,y,t)}{\partial y^2}$.

If the depth map is an equifocal plane, a is constant. Otherwise, a is shift-variant, and the diffusion equation becomes:

$$\begin{cases} \dot{u}(x,y,t) = \nabla \cdot (a(x,y)\nabla u(x,y,t))\ t \in (0,\infty) \\ u(x,y,0) = r(x,y) \end{cases} \tag{19}$$

where "∇"denotes the gradient operator and "$\nabla\cdot$"is the divergence operator, $\nabla = \left[\frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\right]^T, \nabla\cdot = \frac{\partial}{\partial x} + \frac{\partial}{\partial y}$.

It is also easy to verify that the variance σ is related to the diffusion coefficient a via:

$$\sigma^2 = 2ta \tag{20}$$

Suppose there are two images $E_1(x, y)$ and $E_2(x, y)$ for two different focus setting, also, $\sigma_1 < \sigma_2$ (that is, $E_1(x, y)$ is more defocused than $E_2(x, y)$), then $E_2(x, y)$ can be written as:

$$\begin{aligned} E_2(x,y) &= \iint \frac{1}{2\pi\sigma_2^2} \exp\left(-\frac{(x-u)^2 + (y-v)^2}{2\sigma_2^2}\right) r(u,v)\,du\,dv \\ &= \iint \frac{1}{2\pi\Delta\sigma^2} \exp\left(-\frac{(x-u)^2 + (y-v)^2}{2\Delta\sigma^2}\right) E_1(u,v)\,du\,dv \end{aligned} \tag{21}$$

where $\Delta\sigma^2 \triangleq \sigma_2^2 - \sigma_1^2$ is called the relative blurring(Favaro et al, 2008). So Eq.(18) can be written as:

$$\begin{cases} \dot{u}(x,y,t) = a\Delta u(x,y,t)\ a \in [0,\infty)\ t \in (0,\infty) \\ u(x,y,0) = E_1(x,y) \end{cases} \tag{22}$$

Eq.(19) becomes:

$$\begin{cases} \dot{u}(x,y,t) = \nabla \cdot (a(x,y)\nabla u(x,y,t))\ t \in (0,\infty) \\ u(x,y,0) = E_1(x,y) \end{cases} \tag{23}$$

When the time-shifted is Δt ,the solution of the diffusion equation is $u(x,y,\Delta t) = E_2(x,y)$,and Δt can be defined as,

$$\Delta\sigma^2 = 2(t_2 - t_1)a \doteq 2\Delta t a \tag{24}$$

Thus, the relation between the relative blurring and the depth map can be denoted as:

$$\Delta\sigma^2 = \gamma^2(b_2^2 - b_1^2) \tag{25}$$

where γ is a constant between the blurring radius and the blurring degree, b_i (i=1,2) is the radius of the burring round :

$$b = \frac{Dv}{2}\left|\frac{1}{f} - \frac{1}{v} - \frac{1}{s}\right| \tag{26}$$

where s denotes depth of the blurring point and D denotes the radius of the lens.

3.2 The new shape reconstruction method

Suppose $E_1(x, y)$, whose depth map is $s_1(x, y)$, is the defocused image attained before depth variation, and $E_2(x, y)$ is another defocused image attained after depth variation, in this section, we will propose a new shape from defocus method in which the depth map $s_2(x, y)$ is attained through a depth change Δs, and the main theory is shown in Fig.12.

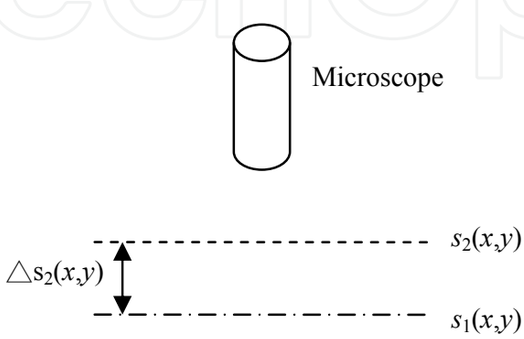


Fig. 11. The main theory of our method

Suppose s_0 is the focus depth, and $s_1(x,y) - s_2(x,y) = \Delta s(x,y)$. Based on the diffusion equations in section 2, the following functions can be given:

$$\begin{cases} \dot{u}(x,y,t) = \nabla \cdot (a(x,y)\nabla u(x,y,t)) \; t \in (0,\infty) \\ u(x,y,0) = E_1(x,y) \\ u(x,y,\Delta t) = E_2(x,y) \end{cases} \tag{27}$$

where the relative blurring can be denoted as:

$$\begin{aligned} \Delta\sigma^2(x,y) &= \gamma^2(b_2{}^2(x,y) - b_1^2(x,y)) \\ &= \frac{\gamma^2 D^2 v^2}{4}\left[\left(\frac{1}{f} - \frac{1}{v} - \frac{1}{s_2(x,y)}\right)^2 - \left(\frac{1}{f} - \frac{1}{v} - \frac{1}{s_1(x,y)}\right)^2\right] \\ &= \frac{\gamma^2 D^2 v^2}{4}\left[\left(\frac{1}{s_0} - \frac{1}{s_2(x,y)}\right)^2 - \left(\frac{1}{s_0} - \frac{1}{s_1(x,y)}\right)^2\right] \end{aligned} \tag{28}$$

Define: $k = \frac{4\Delta\sigma^2}{\gamma^2 D^2 v^2} + \left(\frac{1}{s_0} - \frac{1}{s_1(x,y)}\right)^2$, thus the desired depth map is:

$$s_2(x,y) = 1/(\frac{1}{s_0} \pm \sqrt{k}) \tag{29}$$

In real applications, it is reasonable to discuss the following four cases when the distance between the sample and the microscope is becoming shorter and shorter.

a. $s_1 > s_2 > s_0$

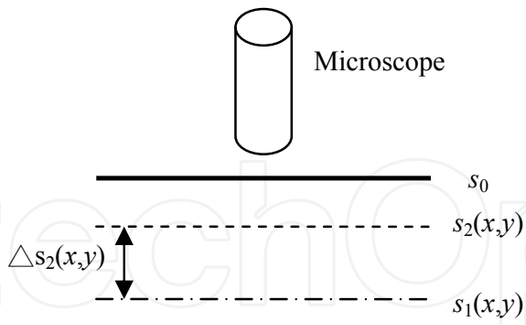


Fig. 12. The theory of case A

In this case, $E_1(x, y)$ and $E_2(x, y)$ are on the large side of s_0, and $E_1(x, y)$ is more defocused than $E_2(x, y)$, so it is backward diffusion process from $E_1(x, y)$ to $E_2(x, y)$, that is, the diffusion efficient a is negative. The theory is shown as Fig. 13 and the finial depth can be denoted as:

$$s_2(x,y) = 1/(\frac{1}{s_0} - \sqrt{k}) \tag{30}$$

b. $s_0 > s_1 > s_2$

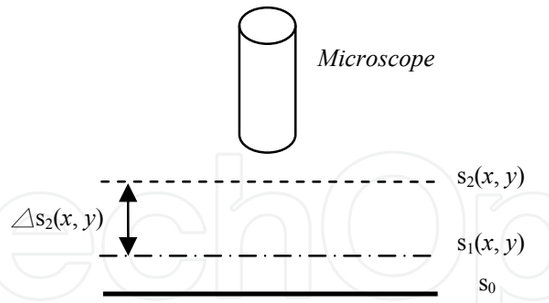


Fig. 13. The theory of case B

As is shown in Fig.14, here, $E_1(x, y)$ and $E_2(x, y)$ are on the small side of s_0, $E_1(x, y)$ is less defocused than $E_2(x, y)$, so it is afterward diffusion from $E_1(x, y)$ to $E_2(x, y)$, and the diffusion efficient a is positive. The finial depth can be denoted as:

$$s_2(x,y) = 1/(\frac{1}{s_0} + \sqrt{k}) \tag{31}$$

c. $s_1 > s_0$, $s_2 < s_0$, $(s_0 - s_2) < (s_1 - s_0)$

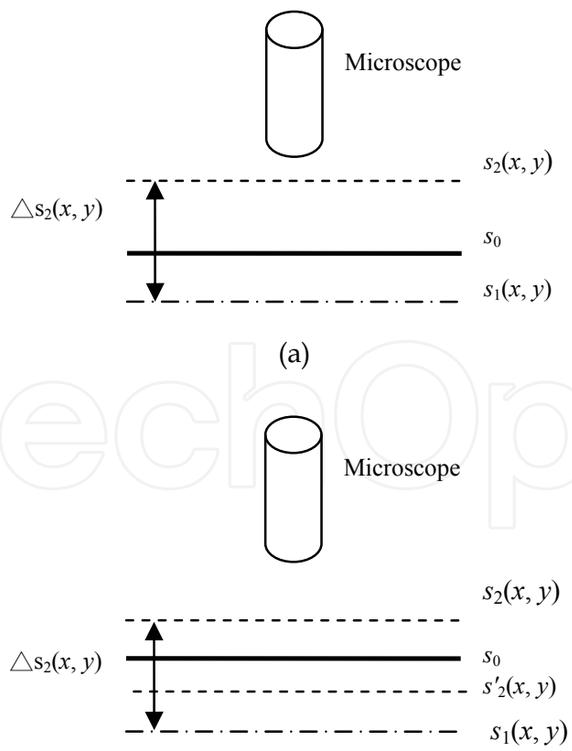


Fig. 14. The theory of case *C*

This case is a little more complicated than the first two scenarios. $E_1(x, y)$ is more defocused than $E_2(x, y)$, but they are not on the same side of s_0. Suppose $s'_2(x, y)$ is the symmetrical depth of $s_2(x, y)$ about s_0, the process can be transferred from Fig.15(a) to Fig.15(b), and the finial depth can be denoted as:

$$s'_2(x,y) = 1/(\frac{1}{s_0} - \sqrt{k}) \tag{32}$$

$$s_2(x,y) = s_0 - (s'_2(x,y) - s_0) = 2s_0 - s'_2(x,y) \tag{33}$$

d. $s_1 > s_0$, $s_2 < s_0$, $(s_0 - s_2) > (s_1 - s_0)$

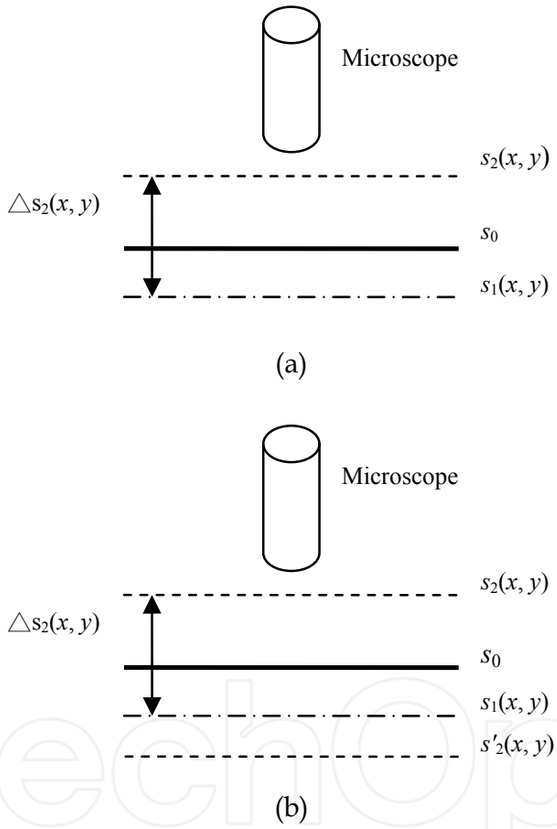


Fig. 15. The theory of case *D*

Here, $E_1(x, y)$ is less defocused than $E_2(x, y)$, and they are not on the same side of s_0. Suppose $s'_2(x, y)$ is the symmetrical depth of $s_2(x, y)$ about s_0, the process can be transferred to Fig.16(b), and the finial depth can be denoted as:

$$s'_2(x,y) = 1/(\frac{1}{s_0} + \sqrt{k}) \tag{34}$$

$$s_2(x,y) = s_0 - (s'_2(x,y) - s_0) = 2s_0 - s'_2(x,y) \tag{35}$$

As a global algorithm, we construct the following optimization problem to calculate the solutions of the diffusion equations.

$$\tilde{s} = \arg\min_{s_2(x,y)} \iint \left(u(x,y,\Delta t) - E_2(x,y)\right)^2 dxdy \tag{36}$$

However, the optimization process above is ill posed(Favaro et al 2008), that is, the minimum may not exist, and even if it exists, it may not be stable with respect to data noise. A common way to regularize the problem is to add a Tikhonov Penalty:

$$\begin{aligned} \tilde{s} = \arg\min_{s_2(x,y)} & \iint \left(u(x,y,\Delta t) - E_2(x,y)\right)^2 dxdy \\ & + \alpha \left\|\nabla s_2(x,y)\right\|^2 + \alpha k \left\|s_2(x,y)\right\|^2 \end{aligned} \tag{37}$$

where the additional term imposes a smoothness constraint on the depth map. In practice, we use $\alpha > 0$, $k > 0$ which are all very small, because this term has no practical influence on the cost energy denoted as:

$$\begin{aligned} F(s) = & \iint \left(u(x,y,\Delta t) - E_2(x,y)\right)^2 dxdy \\ & + \alpha \left\|\nabla s\right\|^2 + \alpha k \left\|s\right\|^2 \end{aligned} \tag{38}$$

Thus the solution process is equal to the following:

$$\begin{aligned} & \tilde{s} = \arg\min_{s} F(s) \\ & s.t. \quad Eq.(34), Eq.(37) \end{aligned} \tag{39}$$

Eq. (39) is a dynamic optimization which can be solved by the gradient flow, the algorithm can be divided into the following steps (the detailed process can be seen in literature (Favaro et al 2008)):

1. Give camera parameters f, D, γ , v,s_0;two defocus images E_1,E_2；a threshold ε ； α and optimization step β ;
2. Initialize the depth map with a plan s, to be simple, we can suppose that the initial plane is an equifocal plane;
3. Compute Eq.(28)and attain the relative blurring;
4. Compute Eq.(27) and attain the solution $u(x,y,\Delta t)$ of diffusion equations ;
5. Compute Eq.(38) with the solution of step(4). If the cost energy is below ε , stop; or compute the following equation with step β ,
6. $\dfrac{\partial s}{\partial t} = -F'(s)$ (40)
7. Compute Eq.(26), update the depth map, and return to step(3).

So if the initial depth is known, maybe it is just a general value, the dynamic depth, as well as the expected shape, can be reconstructed.

3.3 Experiment results

In order to validate the new algorithm, we used it to reconstruct the shapes of a nano standard grid which is 500nm high, two AFM cantilevers. We used the microscope of HIROX -7700 shown as Fig.17, and magnify the grid into 7000 times. The rest parameters are as the following: f=0.357mm, s_0 =3.4mm, F-number =2, D= f/2.

Fig. 17. HIROX-7700

In order to investigate the influence of different region size on the algorithm, we tested the grid with three kinds of region size and two kinds of AFM cantilever. As for the grid, through comparing to the true grid, the error maps in each experiment are constructed and the mean square error of the proposed method was calculated to test the precision. When testing the AFM cantilevers, we used PI nano platform to test the reconstruction precision.

3.3.1 Shape reconstruction of the nano grid

Firstly, the experiment using 120×110 pixel grid region was conducted. The results are shown in Fig.18 to Fig.21. Fig.18 is two defocused images in which the left is the image before variation and the right is that after variation; Fig.19 is the constructed 3D shape of the nano grid. In order to investigate the precision of the new algorism, we constructed the error map Φ between the true shape s and the estimated shape $\tilde{s}$, and computed the mean square error φ of the whole image. The compute formulas are shown in Eq.(41) and Eq.(42). Fig. 20 is the true shape of the grid and Fig.21 is the error map.

$$\phi = \tilde{s} / s - 1 \tag{41}$$

$$\varphi = \sqrt{\mathrm{E}\left[(\tilde{s} / s - 1)^2 \right]} \tag{42}$$

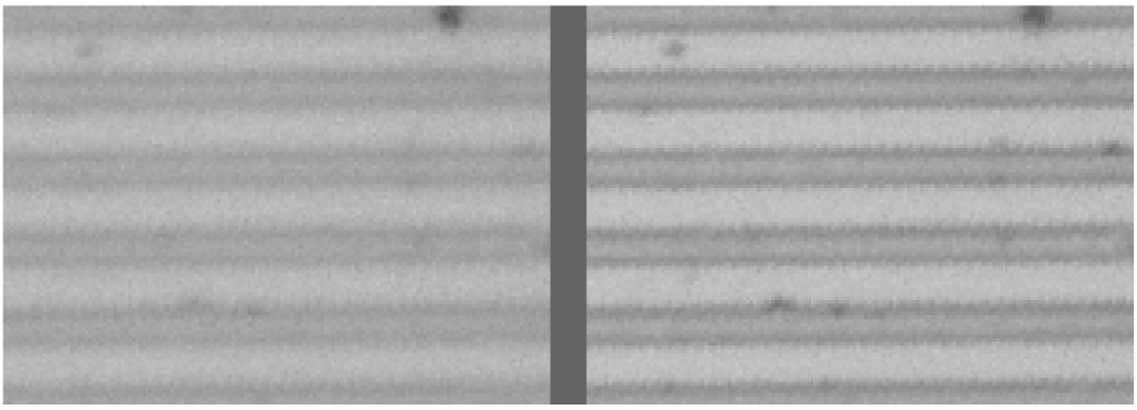


Fig. 18. The defocused images

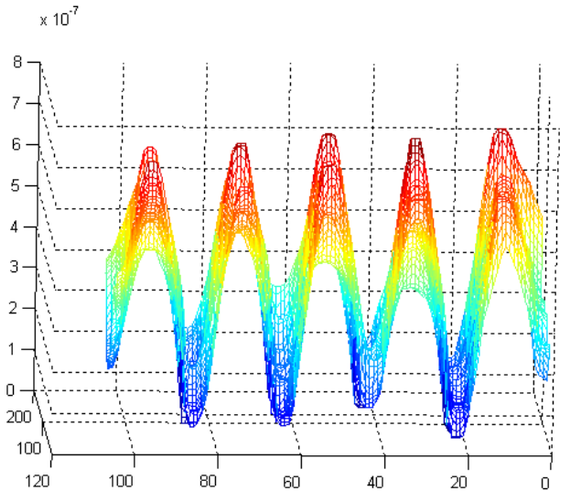


Fig. 19. The constructed 3D shape

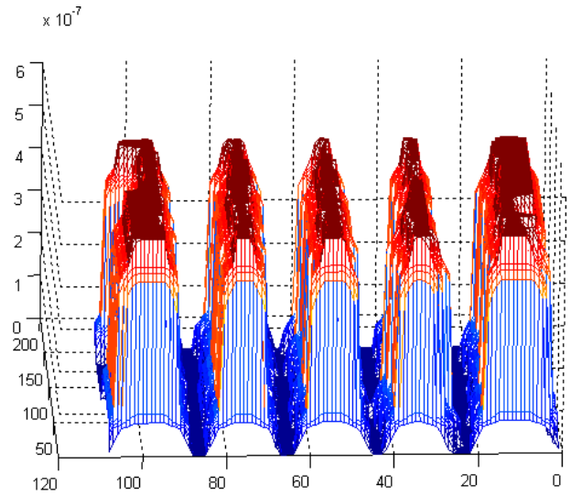


Fig. 20. The true 3D shape

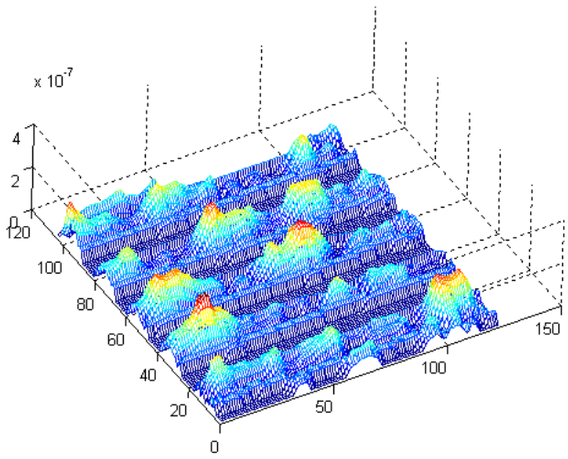


Fig. 21. The error map

From Fig. 10-21, we can see that the new algorithm can attain good results in constructing nano grid shape, and the precision of the proposed method is very high. The mean square error of the whole image is equal to -0.048, and the average error is -9.26 nm.

Secondly, we tested our algorithm on the grid region of 120×50 pixels, the results are shown as Fig.22- Fig.25. The mean square error is equal to -0.038 and the average error is 10.39 nm. From the figures, we can see that the error of our construction algorithm is slightly larger at the edge of the image and smaller at other region, this results from the optimization method. However, the average error is only about 2.3% and it can certainly satisfy the demand of micro/nano magnification.

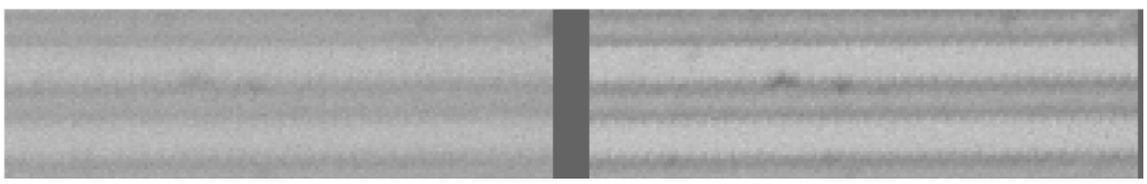


Fig. 22. The defocused images

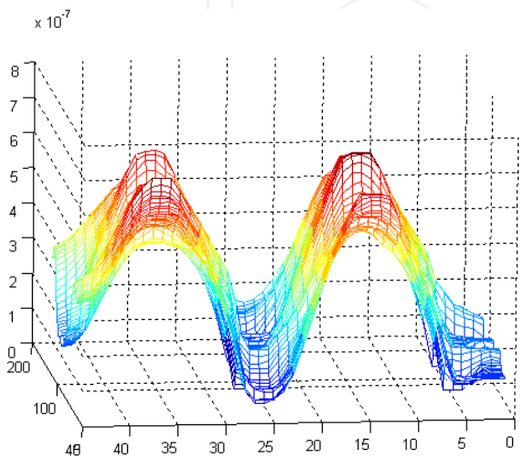


Fig. 23. The constructed 3D shape

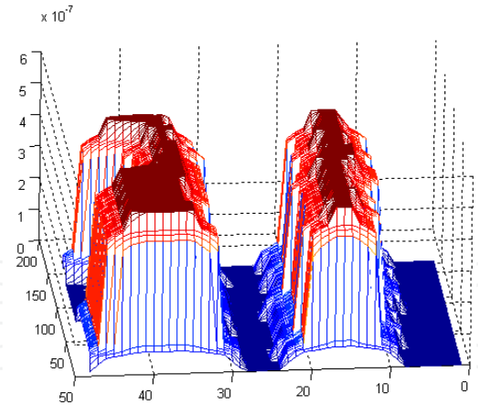


Fig. 24. The true 3D shape

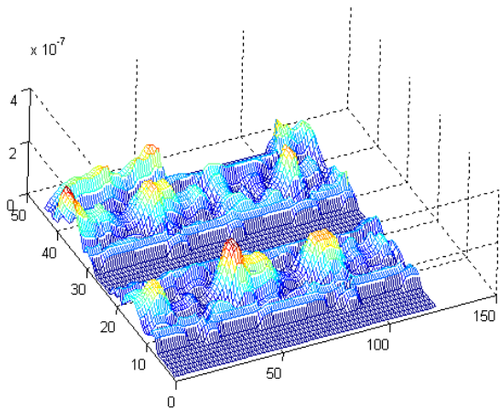


Fig. 25. The error map

3.3.2 Shape reconstruction of the AFM cantilever

The raise height of the cantilevers was controlled by the Iphysik Instrumente(PI) nano platform. Furthermore, we provided the performance of the algorithm on three kinds of raise height: 500nm, 300nm and 100nm. The theory of the experiment is shown as Fig.26.

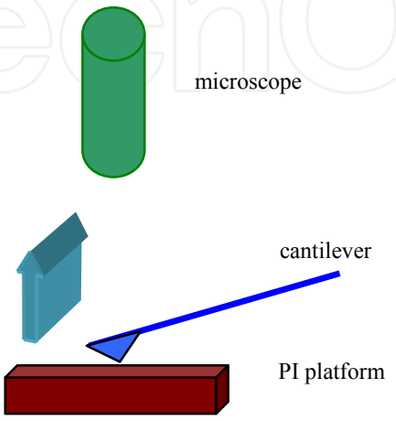


Fig. 26. The experiment theory of reconstruction AFM cantilever

Firstly, the experiment using the conductive AFM cantilever was conducted. Fig.27 is two defocused images, in which the left is the image before variation and the right is that after variation; Fig. 28(a)- Fig.28(c) are the constructed 3D shapes of the bended cantilever when the PI platform rises 500nm, 300nm and 100nm.

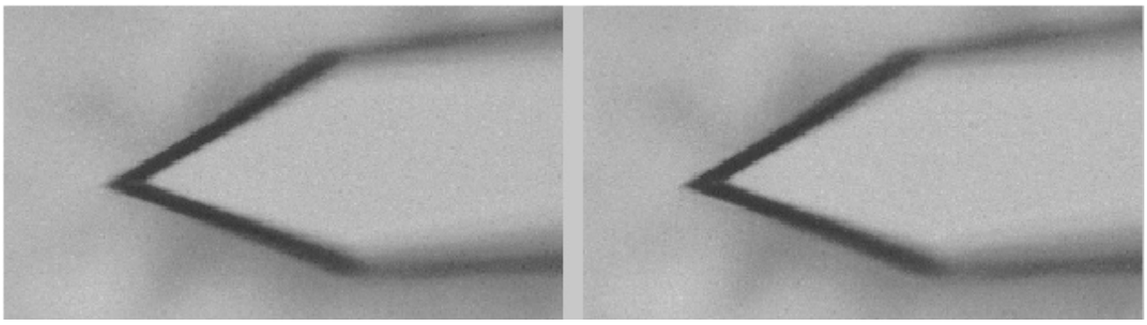


Fig. 27. The defocus images of the conductive cantilever

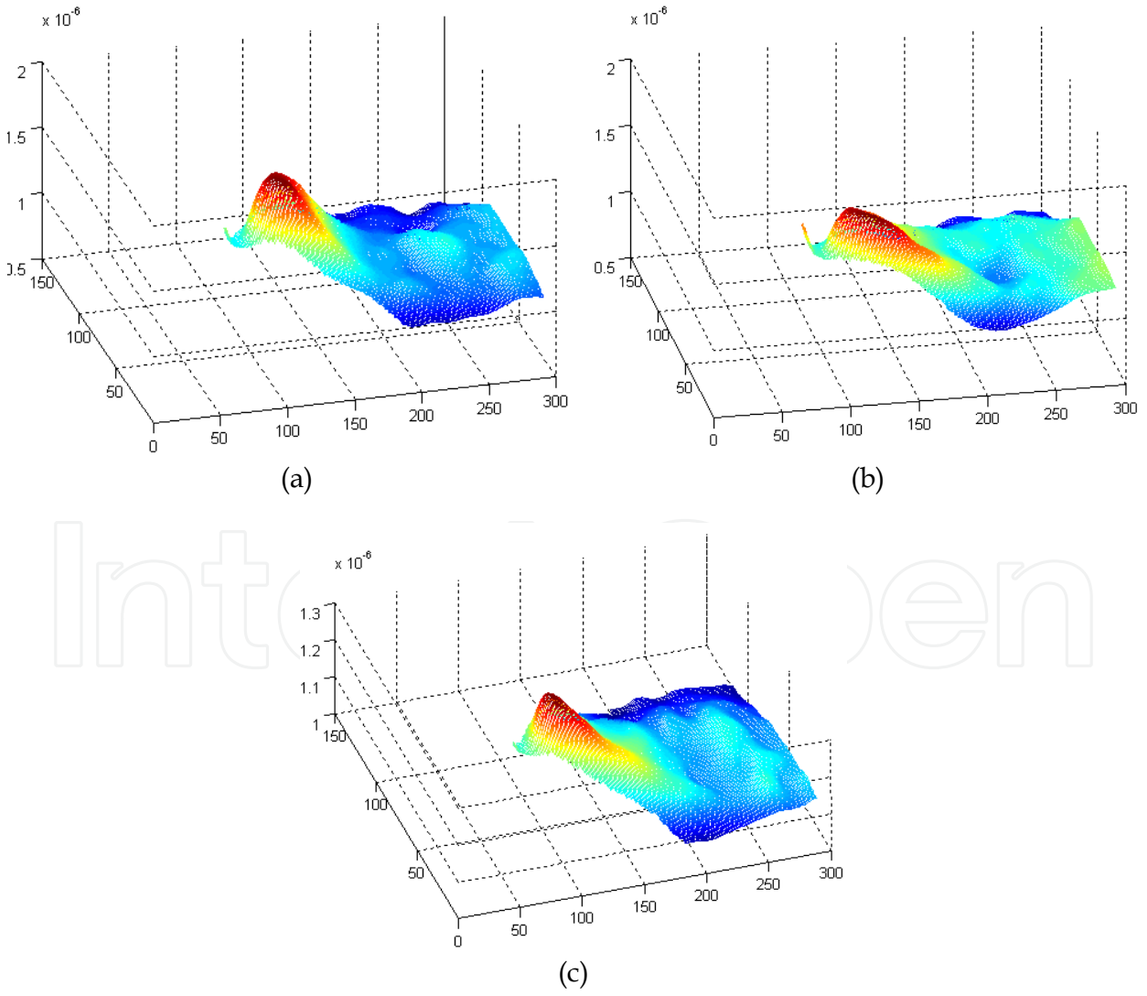


Fig. 28. The constructed 3D shape for 500nm, 300nm and 100nm

From Fig. 28., we can see that when the PI platform rises, the top end of the conductive cantilever bends obviously, and the deflection decreases gradually from the top end to the trailing end until close to a steady value; the bended degree is a monotonic function with the raise height. In order to contrast the bended precision, we choose the section image of them on the same position, and show them in Fig.29. From it, we can see that the deflection height is proportionately increases when the platform rises, and the height difference between the top end and the steady value is exactly equal to the raise height of the PI platform.

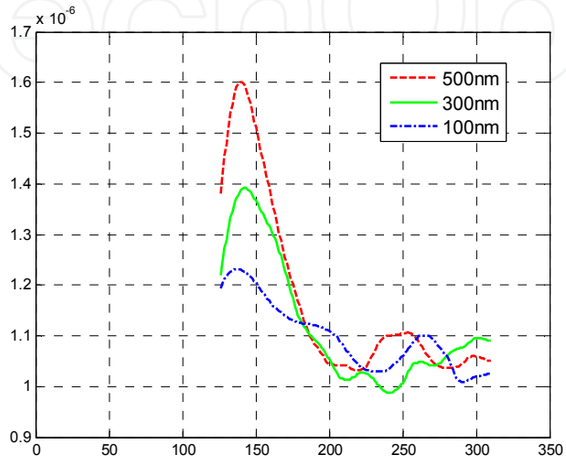


Fig. 29. The contrast of the conductive cantilever

Secondly, the experiment using the triangle cantilever was conducted. Fig.30 is two defocus images, in which the left is the image before variation and the right is that after variation; Fig.31(a)- Fig.31(c) are the constructed 3D shapes of bended cantilever when different raise height is 500nm, 300nm and 100nm; Fig.31 is the contrast image of these three sections. From them we can get the same conclusions as the last experiment, but the sensitivity of the triangle cantilever is lower because the reconstructed shapes are a little rough.

From these experiments, we can see that, regardless of the cantilever shape, our algorithm all can reconstruct the global bended shape exactly with only two defocused images. The following conclusion can be given:

1. The most obvious bend of the cantilever concentrates on the region near to the tip, it is reasonable because when the PI platform works up, the stress all concentrates on the tip due to our experiment theory.
2. The cantilever's original shape, material and illumination can influence the reconstruction result to some extent. For example, the conductive cantilever is thinner than the triangle cantilever, and the shape reconstruction of it is smoother due to its higher sensitivity; the black edge of the cantilever results in a little error in the result.
3. The raise height is larger, the calculation result is exacter, and the reconstruction image is smoother
4. No matter how much the raise height is, the reconstruction height tends to be steady finally. Furthermore, the height difference between the maximum value and the original value is equal to the raise height of the PI platform.

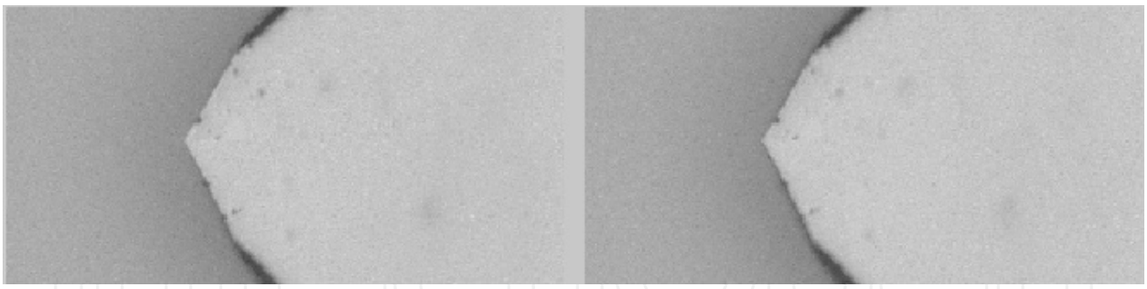
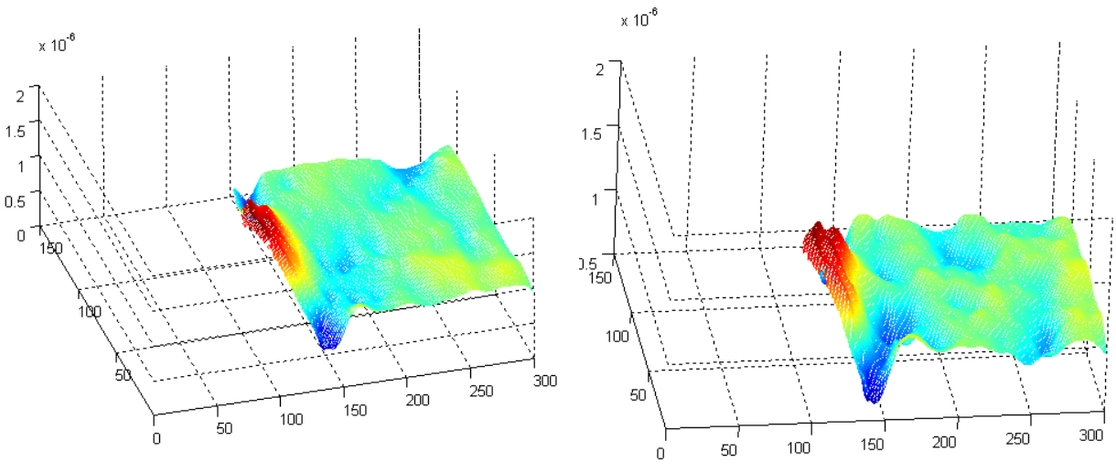
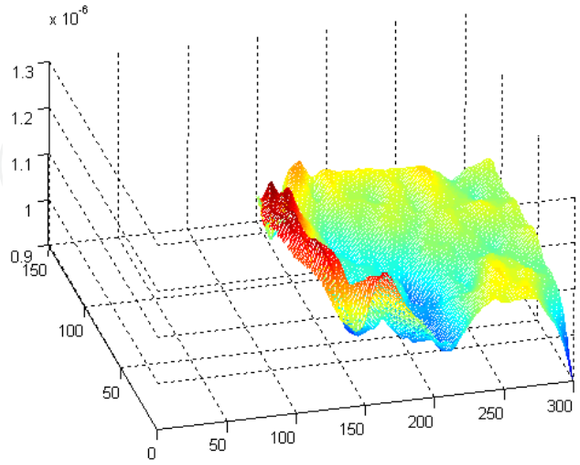


Fig. 30. The defocus images of the triangle cantilever



(a) (b)



(c)

Fig. 31. The constructed 3D shape for 500nm, 300nm and 100nm

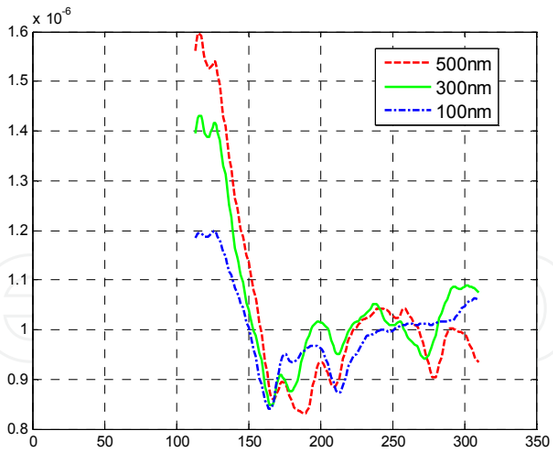


Fig. 32. The contrast of the triangle cantilever

4. Conclusion

In this chapter, two typical micro vision algorithms were researched to model the process in micro/nano size: sub-pixel 2D motion measurement and DFD 3D reconstruction.

As for 2D motion measurement, this chapter mainly researched the problem of motion measurement based on the sub-pixel estimation for image sequences. Firstly, three important factors, including the searching region, the model size, and the fitting precision of sub-pixel, are analyzed and researched in detail, and the most appropriated parameters are chosen with respect to both the experiment results and the measurement characteristic of micro/nano image sequences. Then, the nano platform with high precision, the microscope with high magnification and the standard grid are used together to validate the measurement precision of this method. Finally, the proposed method is used to measure the driving characteristic curve of a piezoelectric actuator practically. The experimental results of the piezoelectric actuator driving characteristic measurement are consistent with the physics analysis. Also, the proposed method, which is simple in manipulation and credible in measurement results, satisfies the requirement of the micro/nano measurement with high precision.

On the other hand, a global shape reconstruction of the standard nano grid using single optical microscope was researched based on a new DFD method. Our primary contribution is to suppose a new global DFD algorithm. Therefore, it can be used to attain 3D information in one-eye vision, hand-eye system, especially in micro/nano manipulation. The second contribution is proposing a series of experiments to validate the method on micro/nano scale. The results below are significant: the computer vision can be used to reconstruct the global shape of the samples in micro/nano manipulation using defocused images without changing camera parameters.

5. References

Asada N.; Fujiwara H. & Matsuyama T.(1998). *Edges and depth from focus*. International Journal of Computer Vision, Vol.26, No. 2, pp. 153-163.

Bove V. M.(1993). *Entropy-based depth from focus*. Journal of Optical Society of America – A, Vol.10, No.4, pp.561-566.

Ens J.; Lawrence P.(1993). An investigation of methods for determining depth from focus, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 15, No.2, pp. 97-108.

Favaro P.; Burger M. & Osher S. J.(2008). *Shape from defocus via diffusion*, IEEE Transaction of Pattern Recognition and Machine Intelligence, Vol.30, No.3, pp. 518-531.

Favaro P.; Mennucci A. & Soatto S.(2003). *Observing shape from defocused images*, International Journal of Computer Vision, Vol. 52, No.1, pp. 25-43.

Favaro P.; Mennucci A.(2002). *Learning shape from defocus*, pp. 735~745, Proceedings of European Conference on Computer Vision, Copenhagen, Denmark, May 27-June 2, 2002.

Giachetti A.; Torre V.(1996). *Refinement of optical flow estimation and detection of motion edges*, Proceeding of European Conference on Computer Vision, Cambridge, UK, April 13-14,1996.

Girod B.; Scherock S.(1989). *Depth from defocus of structured light*, pp. 209-215, Proceedings of Optics, Illumination, and Image Sensing for Machine Vision IV, Philadelphia, USA, October 14-15,1989.

Gokstorp M.(1994). *Computing depth from out-of-focus blur using a local frequency representation*, pp.153-158 Proceedings of International Conference on Pattern Recognition, Jerusalem, Israel, October 9-10, 1994.

Horn B. K. P.; Schunck, B.G..(1981). *Determining optical flow*, Artificial Intelligence, Vol. 17,pp. 185–203 .

Horn B.(1986). *Robot Vision*. Cambridge, MA: MIT Press, 1986.

Kim J. H.; Meng C. H.(2007). *Visually servoes 3-D alignment of multiple objects with sub-nanometer precision*. IEEE Transactions on Nanotechnology, Vol.7, No.3, pp.321-330.

Lagnado R.; Osher S.(1997). *A technique for calibrating derivative security pricing models: numerical solution of an inverse problem*, Journal of Computational Finance, Vol. 1, No.1, pp. 13-26.

Nair N., and Stewart C.(1992). *Robust focus ranging*, pp. 309-314, Proceedings of IEEE Computer Vision and Pattern Recognition, Champaign, USA, June 16-18, 1992.

Navar S. K.;Watanabe M. & Noguchi M.(1996). *Real-time focus range sensor*, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.18, No.12, pp. 1186-1198.

Nayar S. K.(1992). *Shape from focus system*, pp. 302-308, Proceedings of IEEE Computer Vision and Pattern Recognition, Champaign, IL, USA, October 14-15,1992.

Nayar S. K.; Watanabe S. & Noguchi M.(1996). *Real-time focus range sensor*, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 18, No.2, pp. 1186-1198.

Pentland A. P.(1987). *A new sense for depth of field*. IEEE Transaction on Pattern and Machine Intelligence, Vol.9,No. 4, pp.523-531.

Pentland A. P.; Scheroch S. & Darrell T.(1994). *Simple range cameras based on focus error*, Journal of the Optical Society of America - A, 1994, Vol.11, No.11, pp. 2925-2934.

Qi T.; Michale N. H.(1986). *Algorithms for sub-pixel registration*, Computer Vision, Graphics, and Image Processing,Vol. 35,pp.220-233.

Robinson D.; Milanfar P.(2004). *Fundamental performance limits in image registration*. IEEE Transactions on Image Process, Vol.13, No.9, pp. 1185-1199.

Singh, A.(1990). *An estimation-theoretic framework for image flow computation*, pp. 168–177, Proceedings of 3rd International Conference on Computer Vision, Osaka, Japan., December 4-7,1990.

Subbarao M.; Surya G.(1994). *Depth from defocus: A spatial domain approach*, International Journal of Computer Vision, Vol.13, No.3, pp. 271-294.

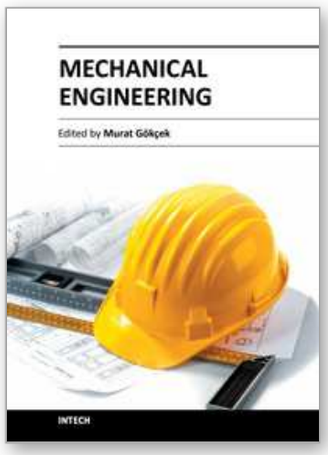
Teresa C. S. A; João M. R. S. T & Mário A. P. V. *Three-dimensional reconstruction and characterization of human external shapes from two-dimensional images using volumetric methods*, Computer Methods in Biomechanics and Biomedical Engineering, DOI: 10.1080/10255840903251288 (in press).

Teresa C. S. A; João Manuel R. S. T & Mário A. P. V.(2008). *3D Object Reconstruction from Uncalibrated Images using an Off-the-Shelf Camera*, Advances in Computational Vision and Medical Image Processing: Methods and Applications, pp. 117-136.

Vinay, P. N.; Subhasis C.(2007). *On defocus, diffusion and depth estimation*, Pattern Recognition Letters, Vol. 28, No.3, pp. 311-319.

Wu L. D. (1993).*Computer Vision*. Fudan University Press, 1993.

Yin C. Y.(1999). *Determining residual nonlinearity of a high-precision heterodyne interferometer*, Optical Engineering, Vol.38, No. 8, pp. 1361-1365.



Mechanical Engineering

Edited by Dr. Murat Gokcek

ISBN 978-953-51-0505-3
Hard cover, 670 pages
Publisher InTech
Published online 11, April, 2012
Published in print edition April, 2012

The book substantially offers the latest progresses about the important topics of the "Mechanical Engineering" to readers. It includes twenty-eight excellent studies prepared using state-of-art methodologies by professional researchers from different countries. The sections in the book comprise of the following titles: power transmission system, manufacturing processes and system analysis, thermo-fluid systems, simulations and computer applications, and new approaches in mechanical engineering education and organization systems.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Yangjie Wei, Chengdong Wu and Zaili Dong (2012). Applications of Computer Vision in Micro/Nano Observation, Mechanical Engineering, Dr. Murat Gokcek (Ed.), ISBN: 978-953-51-0505-3, InTech, Available from: http://www.intechopen.com/books/mechanical-engineering/applications-of-computer-vision-in-micro-nano-observation

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen