Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

An adaptive depth of field imaging system for micromanipulation

Xiaodong Tao, Deokhwa Hong and Hyungsuck Cho*

Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Korea
(e-mail: hscho@lca.kaist.ac.kr).

Abstract: Vision-based techniques used in automatic micromanipulation are limited by small depth of field (DOF) because of high magnification and big numerical aperture of optical system. This paper proposed a new adaptive DOF image system which can adaptively adjust DOF during micromanipulation while maintaining high resolution of the images. The proposed system uses a spatial light modulator instead of fixed phase mask to make flexible wave-front coding, which can change the DOF during micro manipulation. An DOF-extension scheme is proposed based on the depth information of micro objects, which is estimated by the stereo vision method. The large DOF in the coarse assembly and high resolution in the fine assembly are achieved. The image qualities of multi micro objects in different depths are improved during the micromanipulation process. Experiment results show the feasibility of the proposed system in the micromanipulation application.

1. INTRODUCTION

Vision-based techniques used in automatic micromanipulation are limited by small depth of field (DOF), occlusion, and illumination. In micromanipulation, multiple micro parts in different depth need to be detected by the imaging system with high resolution. Shallow DOF of the imaging system can cause the lack of visual information, which may lead to failure of visual servoing. Especially when there are multiple objects to be viewed simultaneously, and they are not located in the same depth, some of them would be blurred due to shallow DOF problem. [Morris *et al*., 1990a, b]

There have been several methods to extend DOF. One solution is to construct extended focus image from multiple images obtained by scanning the optical system along the optical axis [Hausler *et al.*, 1972, Pieper *et al.*, 1983]. In micromanipulation, this method has been applied to perform object recognition and three-dimensional visualization [Fatikow *et al.*, 2000, Greminger *et al.*, 2005]. Because of the time for mechanical scan, image sequence capture and image processing, it is not suitable for real time visual guided micromanipulation. Another method is the wave-front coding method proposed in 1995 [Dowski *et al.*, 1995, Cathey *et al.*, 2002], followed by further modifications applied to the phase mask [Ben-Eliezer *et al.*, 2003, 2005]. In their solution, a special designed phase mask was put in optical path of the microscope that allows an extension of the depth of focus of the images observable by a microscope. Digital holography is also been researched for extended DOF [Ferraro *et al.*, 2005]. But the computation time is too long for the real time application. And it is not suitable for the objects with diffuse surface and complicated surface.

Among above solutions for extended DOF, wave front coding method is a promising method for the real time application because of time efficiency without mechanical scanning. But the resolution reduction in the conventional method hinders its real time application. The resolution of the image will decrease when the extended DOF increases. It can not meet the requirement for micromanipulation to detect micro parts in a large DOF at the initial state yet to maintain high resolution for the final state. In order to solve this problem, this paper proposed a new adaptive DOF imaging system which can adaptively adjust DOF in micromanipulation while maintaining good resolution of the images for multi micro objects. Different from the conventional wave-front coding method, the proposed system use a liquid crystal spatial light modulator (LC-SLM) to make the flexible wave-front coding. A DOF-extension scheme is proposed to achieve high resolution images for all the micro objects in different depth. The phase mask on the LC-SLM is adjusted based on the relationship between the DOF and phase mask strength. The depth information is estimated by the stereo vision system. The real experiments are demonstrated to show the feasibility of the proposed system.

This paper is organized as follows: In Section 2, we introduce the concept of flexible wavefront-coding method using a LC-SLM. Then Section 3 presents the depth estimation from stereo vision. Visual guided micromanipulation with adaptive DOF was presented in Section 4. Section 5 shows the experimental results.

2. FLEXIBLE WAVEFRONT-CODING METHOD USING SPATIAL LIGHT MODULATOR

2.1 Concept of Wavefront-Coding Method

In the conventional wavefront coding system, a special designed phase mask is placed at the system aperture plane to make a phase modulation as shown in Fig. 1. The cubic phase modulation in the pupil plane is expressed by generalized pupil function [Goodman *et al.*, 2005] as in (1).

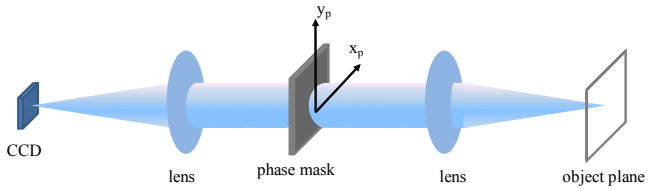


Fig. 1. General composition of wavefront coding systems

$$p(x,y)=\begin{cases}\exp\left[\mathrm{i}\alpha\left(x^3+y^3\right)\right] & \text{for } |x|\le 1, |y|\le 1\\ 0 & \text{otherwise}\end{cases} \quad (1)$$

where $p\ (x,y)$ represents generalized pupil function, x and y denote normalized pupil coordinates and α denotes the strength of the mask. With larger α, incoming light is more strongly modulated. The effect of the phase mask can be well explained by its modulation transfer function (MTF). For 1-D case, the corresponding MTF is calculated as follows [Dowski *et al.*, 1995]

$$M(u,\psi)=\begin{cases}\left(\dfrac{\pi}{12|\alpha u|}\right)^{1/2}\exp\left(\mathrm{i}\dfrac{\alpha u^3}{4}\right) & |\alpha|\gg 20 \quad u\neq 0\\ 1 & u=0\end{cases} \quad (2)$$

where u and ψ represent spatial frequency and defocus, respectively. As can be seen in (2), if α is large enough, the MTF is independent of the amount of defocus ψ. This means that the point spread function (PSF) of the imaging system with the cubic phase mask at the pupil plane does not change. Therefore, the image is equally blurred regardless of the location of the object, and this fact enables applying simple restoration algorithm for post processing to get a clear image. This result can be simply extended to 2D case due to separability of the phase mask [Cathey *et al.*, 2002].

Once the image is wavefront-coded, the image captured by the camera is blurry because the PSF with cubic phase modulation is not as sharp as the ideal one. It can be restored to a clear image by applying well-known Wiener filtering algorithm [Gonzalez *et al.*, 2002]. The brief explanation of Wiener filtering is as follows: Knowing the Fourier transform of the wavefront-coded image and the PSF denoted as $G(\mu, \nu)$ and $H(\mu, \nu)$, and noise to signal ratio K, the restored image in the frequency domain $F(\mu, \nu)$ is calculated by the following equation.

$$F(\mu,\nu)=\left[\frac{1}{H(\mu,\nu)}\frac{|H(\mu,\nu)|^2}{|H(\mu,\nu)|^2+K}\right]G(\mu,\nu) \quad (3)$$

$$\begin{cases}H(\mu,\nu)=\mathcal{F}\{h(m,n)\}\\ G(\mu,\nu)=\mathcal{F}\{g(m,n)\}\\ F(\mu,\nu)=\mathcal{F}\{f(m,n)\}\end{cases}$$

where $h(m,n)$ is the PSF of the imaging system, $g(m,n)$ is the intermediate image and $f(m,n)$ is the restored image. $\mathcal{F}\{\bullet\}$ represents Fourier transform. Thus, the final image of the system is obtained by calculating inverse-Fourier transform of $F(\mu,\nu)$.

2.2 Flexible wavefront coding using LC-SLM

One property of the wavefront coding method is that the image resolution is related to the extended DOF. The large extended DOF will sacrifice the image resolution. With the fixed phase mask, wavefront coding method is only suitable to specific application because trade off between extended DOF and image resolution. For instance, in microassembly, at the initial state, multiple micro parts are often located at different depth. The depth difference between these micro parts is often ranging from hundreds of micrometers to several millimetres. But at final state of assembly, the micro parts will be assembled together in a small DOF with high accuracy. Therefore, conventional fixed phase mask can not achieve optimal trade off between depth of field and image resolution.

In order to achieve flexible wavefront coding, we introduce a LC-SLM in the pupil plane. Combined with a polarizer and an analyzer, the LC-SLM can be performed as a phase modulator. Therefore the parameter α can be adjusted during operation. The systems layout is shown in Fig. 2. After the calibration of the system, the relationship between mask strength α and extended depth of field can be achieved from experiments [Hong *et al.*, 2007]. For different resolution, the extended DOF is different. Fig. 3 shows one curve with 20lp/mm resolution. This relationship can be expressed as a function after doing polynomial fitting as follows.

$$\alpha=f(D) \quad (4)$$

where D is the depth field. The resolution of target in the focus plane with different α is shown in Fig. 4. As can be seen, the resolution decreases from 53lp/mm to 35lp/mm as α increase from 0 to 100. Therefore with the high mask strength α, the detailed features of the target on the focus plane can not be observed.

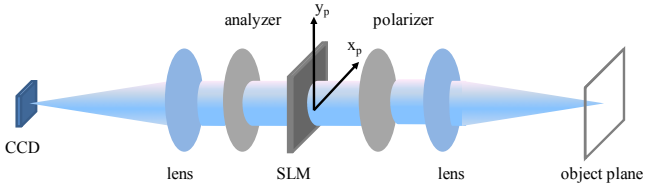


Fig. 2. System layout for flexible wavefront coding

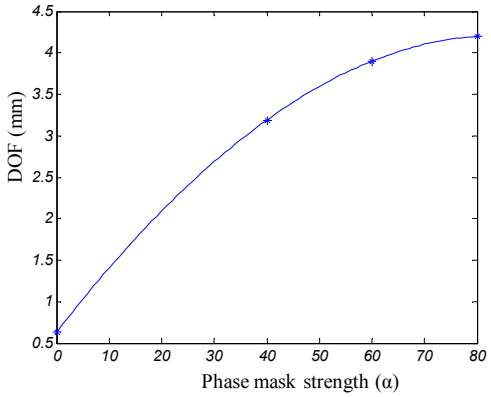


Fig. 3. DOF vs. Phase mask strength for 20lp/mm resolution

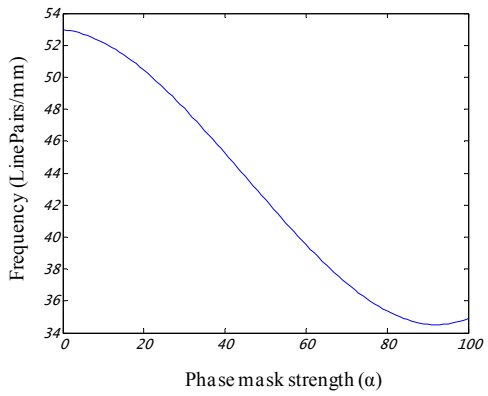


Fig. 4. Resolution vs. Phase makes strength for target on focus plane

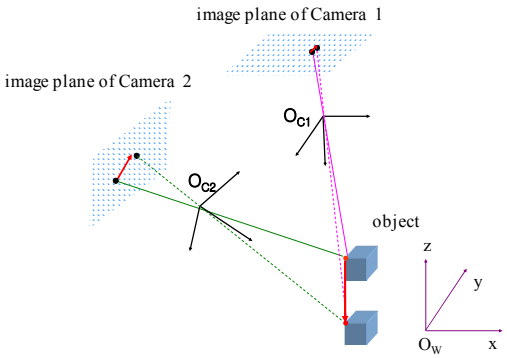


Fig. 5. Configuration of stereo vision

3. ADAPTIVE DOF USING STEREO VISION

In order to extend DOF and keep the resolution high enough for all the targets on the same time, the phase mask strength is adjusted according to its relationship with DOF as introduced in Section 2.2. Therefore depth information of the target is critical for the proposed system. Although the depth from focus method can be applied by using single camera [Nayar *et al.*, 1990], it is not suitable for extended DOF system because of large DOF. And it can not operate in real time. Therefore in this paper, stereo vision system is used to achieve the depth information. The configuration of the stereo vision system is shown in Fig. 5.

The depth information can be observed by introducing another camera with a tile angle to the optical axis. In the micromanipulation application, cameras have high magnification and the size of an object is small compared with the distance from the lens. Therefore, a scaled orthographic projection model is used for camera calibration [Faugeras *et al.*, 1993]. The camera model is shown in (5).

$$\begin{bmatrix} {}^{I}X \\ {}^{I}Y \\ 1 \end{bmatrix} = \begin{bmatrix} m & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} {}^{C}X \\ {}^{C}Y \\ {}^{C}Z \\ 1 \end{bmatrix} \tag{5}$$

where ${}^{I}X$, ${}^{I}Y$ represent the image coordinate, ${}^{C}X$, ${}^{C}Y$, ${}^{C}Z$ represent the camera coordinate. m is the magnification factor. Then the transformation from the world frame to the pixel frame can be expressed as follows,

$$\begin{aligned} \begin{bmatrix} {}^{p}u \\ {}^{p}v \\ 1 \end{bmatrix} &= \begin{bmatrix} m(k_u r_1 + f_u r_2) & m(k_u t_x + f_u t_y) + u_0 \\ m f_v r_2 & m f_v t_y + v_0 \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} {}^{W}X \\ {}^{W}Y \\ {}^{W}Z \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} q_1^T & q_{14} \\ q_2^T & q_{24} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} {}^{W}M \\ 1 \end{bmatrix} \end{aligned} \tag{6}$$

where $f_u = -k_u / \tan\theta$, $f_v = k_v / \sin\theta$; k_u and k_v are scale factors in two directions. θ is the skew angle of the pixel plane; u_0 and v_0 are the pixel coordinate of camera center; ${}^{W}X$, ${}^{W}Y$ and ${}^{W}Z$ are the world coordinate of calibration pattern; ${}^{P}u$ and ${}^{P}v$ are the pixel coordinate of calibration pattern. r and t are extrinsic parameters of the camera. (6) can be modified for N calibration patterns as follows,

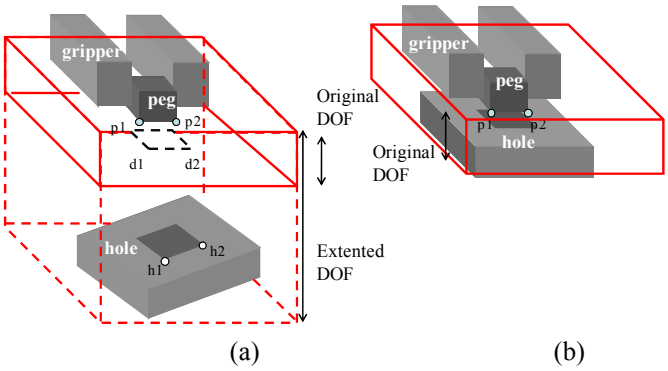
$$\begin{bmatrix} {}^{W}M_1^T & 1 & 0 & 0 & 0^T & 0 \\ 0 & 0 & 0 & 0 & {}^{W}M_1^T & 1 \\ & & \vdots & & & \\ {}^{W}M_N^T & 1 & 0 & 0 & 0 & 0 \\ 0^T & 0 & 0 & 0 & {}^{W}M_N^T & 1 \end{bmatrix} \begin{bmatrix} q_1 \\ q_{14} \\ q_2 \\ q_{24} \end{bmatrix} = \begin{bmatrix} {}^{p}u_1 \\ {}^{p}v_1 \\ \vdots \\ -{}^{p}u_N \\ -{}^{p}v_N \end{bmatrix} \tag{7}$$

where ${}^{W}M_i^T$ is world coordinate of the ith pattern, and (${}^{p}u_i$, ${}^{p}v_i$) is its pixel coordinate. Least-squares solution of (7) can be solved by pseudo-inverse method [Faugeras *et al.*, 1993].

After calibrating camera 1 and camera 2 with the method introduced above, world coordinate of target can be obtained from the image coordinate of features.

$$\begin{bmatrix} {}^{W}X \\ {}^{W}Y \\ {}^{W}Z \end{bmatrix} = \begin{bmatrix} {}^{C1}q_1^T \\ {}^{C1}q_2^T \\ {}^{C2}q_1^T \\ {}^{C2}q_2^T \end{bmatrix}^{+} \begin{bmatrix} {}^{C1}u - {}^{C1}q_{14} \\ {}^{C1}v - {}^{C1}q_{24} \\ {}^{C2}u - {}^{C2}q_{14} \\ {}^{C2}v - {}^{C2}q_{24} \end{bmatrix} \tag{8}$$

where ${}^{C1}q_i$ and ${}^{C2}q_i$ are calibrated parameters for camera 1 and camera 2 respectively. $\left({}^{c1}u,\ {}^{c1}v\right)$ and $\left({}^{c2}u,\ {}^{c2}v\right)$ are image coordinates of the target in camera 1 and camera 2 respectively.



(a) (b)

Fig. 6. Micorassembly scheme using adaptive DOF, (a) coarse assembly and (b) fine assembly

4. VISUAL GUIDED MICROMANIPULATION WITH ADAPTIVE DOF

4.1 Task Description

In this paper, the micromanipulation task is to assemble a 450 μm × 400 μm × 300 μm micro peg, held by a micro gripper, to a 650 μm × 500 μm × 350 μm hole. The peg is located near the optical axis on the focus plane by autofocusing. The initial distance between the peg and hole is 4 mm in the Z direction.

The microassembly process includes two steps, as shown in Fig. 6, at the initial state of the microassembly, the two parts are far from each other. The proposed system will extend DOF to observe the features on both of the hole and peg. During the coarse assembly, the hole will move near to the peg as shown in the dashed line of Fig. 6 (a). The depth information is estimated by stereo vision. From (4), the desired mask strength is obtained to display phase mask on the LC-SLM. The extended DOF keep decreasing according to the depth of the hole. Therefore the resolution of the peg will increase during coarse assembly. In fine assembly, both peg and hole are located in original DOF as shown in Fig. 6 (b), which will provide high accuracy for visual guided assembly.

4.2 Visual servoing for micromanipulation

In coarse assembly, image based visual servoing is performed in translations along x and y and the rotation along z. Position based visual servoing is performed along translation along z. The corner features p1, p2, h1 and h2 on the peg and hole are detected before visual servoing. The 20×20 pixel templates around the features p1, p2, h1 and h2 are obtained when these features are in focus by autofocusing process. During visual servoing, the positions of features h1 and h2 are tracked by sum-of-squared difference (SSD) tracking method [Nickels *et al.*, 2002]. The velocities of stages are derived as follows:

$$\begin{bmatrix} V_X \\ V_Y \\ \omega_Z \end{bmatrix} = K_f J_{vf}^{+} \begin{bmatrix} u_{d1} - u_{h1} \\ v_{d1} - v_{h1} \\ u_{d2} - u_{h2} \\ v_{d2} - v_{h2} \end{bmatrix} \tag{9}$$

where $J_{vf} = \begin{bmatrix} k_u m & f_u m & -v_{h1}' \\ 0 & f_v m & u_{h1}' \\ k_u m & f_u m & -v_{h2}' \\ 0 & f_v m & u_{h2}' \end{bmatrix}$

(u_{h1}, v_{h1}), (u_{h2}, v_{h2}), (u_{d1}, v_{d1}) and (u_{d2}, v_{d2}) are the image coordinates of features h1,h2, and desired position d1 and d2 as shown in Fig. 6 (a). (u_{h1}', v_{h1}') and (u_{h2}', v_{h2}') are the feature position vector from the rotation center. J_{vf}^{+} is the pseudo inverse of image Jacobian J_{vf}.

Position based visual servoing is applied along Z direction, where the velocity of the stage in the Z direction is given as follows:

$$V_z = k_Z [{}^{w}Z_d - {}^{w}Z_h] \tag{10}$$

where ${}^{w}Z_d$ is the desired depth of the hole. ${}^{w}Z_h$ is the real depth of the hole calculated by (8). k_Z is the gain. In fine assembly, peg moves towards to hole with high accuracy. Only translations are considered in this state. The visual servoing strategy is similar to coarse assembly.

5. EXPEIRMENTAL RESULTS

5.1 Experimental setup

Experimetnal setup include two parts, the adaptive DOF system and the micromanipulation system as shown in Fig 7. The adaptive DOF system is shown in Fig. 8. A LC-SLM (HOLOEYE®) is located at the system pupil plane sandwiched by two polarizers to perfrom phase modulation. Other optical components are suppplied by Edmund®. In the micromanipulation system, the hole is put on the coarse X and Y stages, whose resolutions are 3um. The resolution of rotation stage is 32urad, wich supplied by Physik Instrumente. The reslution of Z stage is 0.03 um. The gripper is installed on the fine X and Y stage (Physik Instrumente®), whose resolutions are 0.03um. The specificaion of the imaging system is shown in Table 1.

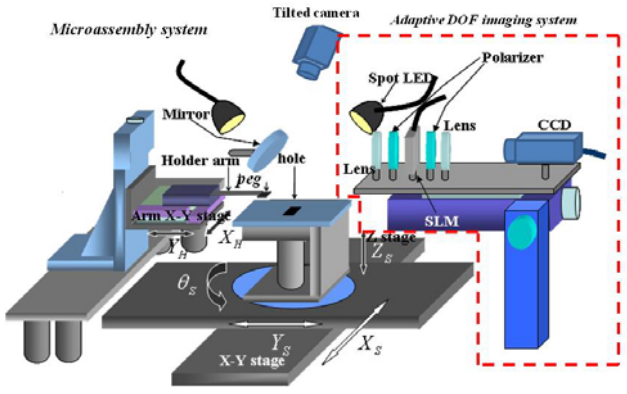


Fig. 7. Experiment setup of the whole system.

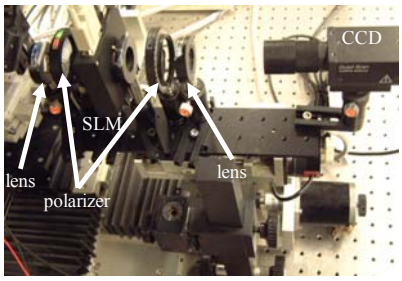


Fig. 8. Experiment setup for the adaptive DOF imagin system

Table 1. Specification of the imaging system

Magnification	1×
CCD size	659×494 pixels
Frame rate	60fps
Pixel size	7.4 μ m×7.4 μ m
Aperture Size	10mm
processing time:	100ms/iteration
Computer:	CPU: 1.7GHz; RAM:128M
Compiler	Microsoft Visual C++

5.2 Experiment result for micromanipulation

The micromanipulation task has described in Section 4.1. Here only coarse assembly is shown to evaluate the ability of the system to adaptively change the DOF. The trajectory of the hole is shown in Fig. 9. At the initial state, the distance between peg and hole is 4mm in Z direction. Fig. 10 shows the image without extended DOF. As can been seen, the peg is blurry because of the shallow DOF. The vision information is not reliable for the visual guided micromanipulation. In order to reduce computation time, sub area (512×400 pixels) is analysed instead of whole image as shown in the dash frame in Fig. 10. Fig. 11 shows the images during assembly with the adaptive DOF. The phase mask is adjusted according to (4) automatically. We also captured images without adaptive DOF during assembly as shown in Fig. 12. The result shows that the proposed system can achieve high quality images for both the peg and the hole. Because the SSD value can measure the similarity between template image and candidate image, Table.2 shows the minimum SSD values for features h2 and h1 on the hole with and without the adaptive DOF. The smaller values are achieved for the adaptive DOF system, which means the high quality of the images is obtained. For the fixed extended DOF method, the captured image with α=80 are shown in Fig. 13. Because the gripper is fixed at the focus plane, the image of the gripper does not change with the fixed extended DOF method. In order to compare the image quality of the gripper, only one image is shown when the hole is at E in Fig. 13. Table 3 shows the minimum SSD value for features p1 and p2 on the peg. It keeps decreasing when the adaptive DOF are used, whereas the higher values are achieved for the fixed extended DOF. As can been, the image of micro peg is degenerate compared with the images with the adaptive DOF in Fig. 11, which satisfies the discussion in the Section 2.2.

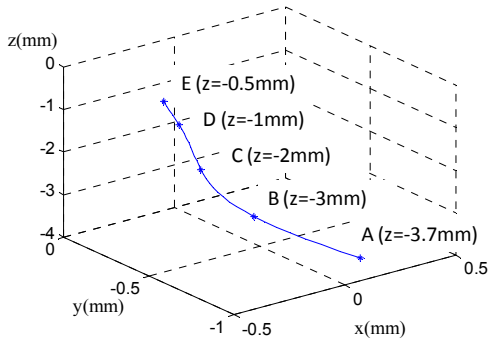


Fig. 9. The trajectory of the hoel

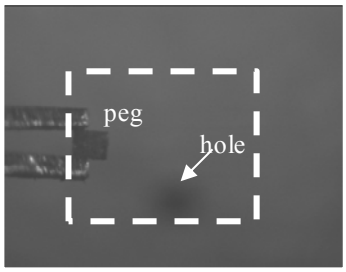


Fig. 10. Initial state without extended DOF

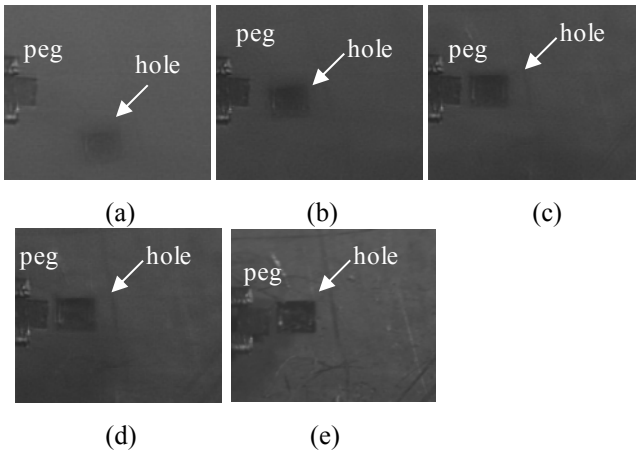


Fig. 11. Captured images during assembly with adaptive DOF, where the hole locates at position (a) A (α=53.5), (b) B (α=35.44), (c) C (α=18.9), (d) D (α=3.96) and (e) E (α=0.97).

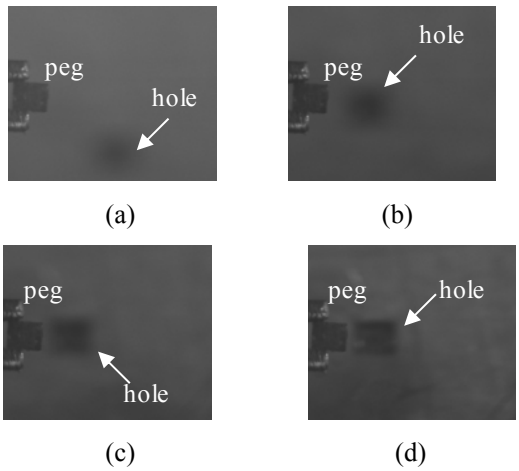


Fig. 12. Captured Images without adaptive DOF, where the hole locates at position (a) A, (b) B, (c) C and (d) D

Table 2. Minimum SSD values for features on the hole

Position	Minimum SSD values for features on the hole	
	With adaptive DOF	Without adaptive DOF
A	2.45	3.13
B	2.40	2.57
C	2.2	2.41
D	1.81	2.02
E	1.21	

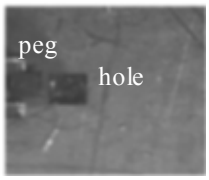


Fig.13. Images captured during assembly with fixed extended DOF (α=80), where the hole locates at E.

Table 3. Minimum SSD values for features on the peg

Position	Minimum SSD values for features on peg	
	With adaptive DOF	With fixed extended DOF (α=80)
A	2.21	2.30
B	2.10	2.32
C	2.03	2.31
D	1.68	2.28
E	0.48	2.29

6. CONCLUSIONS

We have proposed a flexible adaptive depth of field imaging system for micromanipulation. Different from conventional wave-front coding method with fixed phase mask, which suffer the contradictory between resolution and DOF, the proposed system use LC-SLM to make flexible wave-front coding to adjust the DOF according to the environment to avoid this problem. Especially in microassembly application, the proposed system utilized the properties of relationship between resolution and DOF to supply high DOF at the coarse microassembly and high resolution at the fine microassembly. The experiment shows the feasibility of the proposed system. In the future, more efficient reconstruction algorithm will be studied to improve the real time performance.

REFERENCES

Ben-Eliezer, E., Z. Zalevsky, E. Marom and N. Konforti (2003). All-optical extended depth of field imaging system, *J.Opt.A: Pure Appl. Opt.*, **5**, 164-169.

Ben-Eliezer, E., E. Marom, N. Konforti and Z. Zalevsky (2005). Experimental realization of an imaging system with an extended depth of field, *Applied Optics*, **44(14)**, 2792-2798.

Cathey, W. and E. Dowski (2002). New paradigm for imaging systems, *Appl. Opt.*, **41(29)**, 6080-6092.

Dowski, E. and W. Cathey (1995). Extended depth of field through wave-front coding, *Appl. Opt.*, **34(11)**, 1859-1866 .

Fatikow, S. and J. Seyfried (2000). A Flexible microrobot-based microassembly station, *Intelligent and Robotic Systems:Theory and Applications*, **27(1-2)**, 135-169.

Ferraro P, S. Grilli, D. Alfieri, S. De Nicola, A. Finizio, G. Pierattini, B. Javidi, G. Coppola, V. Striano (2005). Extended focused image in microscopy by digital holography. *Opt Exp.,* **13**, 6738–6749.

Faugeras O. (1993). *Three-Dimensional computer vision*, MIT Press, Cambridge, Massachusetts, London, England.

Gonzalez, R. and R. Woods (2002). *Digital Image Processing*, Prentice Hall.

Goodman, J. (2005). *Introduction to Fourier Optics,* Roberts & Company, Colorado

Greminger, M.A. and J.F. Jones (2005). Real-time three-dimensional visualization of standard light microscope image sequences for microassembly, The 6th IEEE International Symposium on Assembly and Task Planning,194-199

Hong, D, K. Park, H. Cho, and M. Kim (2007). Flexible depth-of-field imaging system using a spatial light modulator, Appl. Opt., **46**, 8591-8599.

Hausler, G. (1972). A method to increase the depth of focus by two step image processing, Opt. Commun., **6**, 38.

Nayar, S.K., and Y. Nakagawa (1990). Shape from focus: an Effective approach for rough surfaces, *Proc. IEEE Int. Conf. On Robotics and Automation*, 218-225.

Nickels, K. and S. Hutchinson (2002). Estimating uncertainty in SSD-based feature tracking, Image and vision computing, **20(1)**, 47-58.

Pieper. R.J. and A. Korpel (1983). Image processing for extended depth of field, Appl. Opt., **22**, 1449-1453.

Yamauchi, M. and T. Eiju (1995). Optimization of twisted nematic liquid crystal panels for spatial light phase modulation, *Opt. Commun*., **115**, 19-25.