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 Visual Servoing

Deokhwa Hong*, Ferrokh Janabi-Sharifi**
Hyungsuck Cho*

** Department of Mechanical Engineering, KAIST, Daejeon Korea, (Tel: +82-42+869+3253; e-mail: hscho@ lca.kaist.ac.kr).*
*** Department of Mechanical and Industrial Engineering Ryerson University, Canada, (e-mail: fsharifi@ryerson.ca)*

Abstract: Limitations related to depth of field (DOF) is one of the most challenging issue for visually-servoed micro-operations. Research on extending DOF had limited success for direct application to visual servoing.. Recently, an all-optical system has also been proposed to extend DOF in imaging systems by spatial multiplexing of several Fresnel lenses. Before its application to visual servoing, high resolution of visual feedback must be warranted. This is particularly necessary for fine vision-based positioning tasks. Thus, providing an adaptive scheme for DOF according to objects' distribution will enable the system to obtain the best resolved image for a given distribution. The main contribution of this paper is a proposal for an adaptive DOF-extension scheme in micro-visual servoing while maintaining good resolution and SNR of the images related to target objects. Simulation results are provided to support the proposed method.

1. INTRODUCTION

Using a light microscope in micro robotics area has mainly two problems: a small filed-of-view (FOV) and a small depth-of-filed (DOF). Small FOV can be resolved to some extent by coarse servoing using global vision system, but the DOF problem cannot be resolved in this way because the viewing volume is almost flat. The simplest way of extending DOF is to reduce the aperture of the imaging system. But in this case, the image will suffer from low signl-to-noise ratio (SNR) due to decreased incident light power and the optical resolution.

To overcome the small DOF problem in micro robotics area, Fatikow and Woern used multi-focusing technique (Fatikow et al., 2000; Woern, et al, 2000). By capturing several images at consecutive focus levels and combing the sharp areas of each image, a focused image can be generated. Using this technique one can obtain an approximate depth information and all-focused image, and this synthesized image can be used for object recognition. This method, however, cannot be utilised in real time because it requires several images at different height and the synthesizing step is also time-consuming. Active zooming control was also proposed to resolve small FOV and small DOF problems simultaneously (Tao et al., 2005). Tao and Cho took advantage of the fact that both FOV and DOF are related to magnification of the imaging system: as one increases the magnification, FOV and DOF will decrease. However, this method has two problems. First, FOV and DOF are coupled by magnification, Sometimes it is desired to extend DOF yet keeping FOV as large as possible. Second, extending DOF by changing magnification is limited and this method may not work in a higher numerical aperture system which has smaller DOF.

Extending DOF has interested many researchers for decades beyond micro robotics area. The early stage of this research was mainly about placing an optical power absorbing apodizer in the aperture (Ojeda-Castaneda et al., 1988; Ojeda-Castaneda et al., 1989). However, this kind of approach is basically similar to closing the aperture originating from blocking the incident optical power. Therefore, such approaches suffer from similar problems with decreasing the aperture: low SNR and low optical resolution. In 1995, Dowski and Cathey proposed a method of wave-front coding to extend DOF (Dowski and Cathey, 1995). Further details were provided in (Cathey and Dowski, 2002). The basic concept of the method is to use phase mask in the pupil plane to make the point spread function (PSF) constant even if the object moves back and forth from the focus, and to use deblurring technique to make the image clear. The advantage of the method comes from the fact that PSF remains almost constant along the entire image, and hence, a simple restoration algorithm such as Wiener filtering can be applied.

The main problem in using the wave-front coding method in visually-servoed micro-robotics is its computational time required for the post-processing stage. Image restoration is difficult to achieve in real-time, especially, when the image is large. Ben-Eliezer et al. proposed an all-optical extended depth of field imaging system (Ben-Eliezer et al., 2003; Ben-Eliezer et al., 2005; Ben-Eliezer et al., 2006). In the first two papers, they proposed a composite phase mask consisting of 16 Fresnel lenses (FL), and in the third paper, they proposed a radial mask to extend DOF. The merit of these methods is that they do not need post processing. Thus, one can obtain DOF-extended image in real-time. However, a trade-off exists between DOF and contrast. Using the phase masks proposed in these papers, one should sacrifice the image

contrast to extend DOF. Besides, this method cannot be applied directly to a visual servoing system, because image quality is as important as large DOF. That is, the image should be good enough to localize features for final positioning.

Therefore, providing an adaptive scheme for DOF according to objects' distribution will enable obtaining the best resolved image for a given distribution. When the depth range of objects' distribution is large, the images with large DOF are desirable. When the depth range is small, images with small DOF become preferable. The main contribution of this paper is a proposal for an adaptive DOF-extension scheme in micro-visual servoing while maintaining good resolution and SNR of the images related to target objects. In this paper, basic concept of DOF adaptation is explained and some simulation results are presented to support the proposed method.

2. METHODOLOGY

2.1 Basic Concept

The objective is to visually servo two objects (at different depths) to the desired positions. Focus measure calculated in windows defined around the objects is used as a feedback for DOF adaptation. The overall block diagram is drawn in Fig. 1.

Defocus of an object can be expressed by a dimensionless misfocus parameter as follows (Goodman, 1996);

$$\delta = \frac{\pi D^2}{4\lambda}\left(\frac{1}{d_{obj}}+\frac{1}{d_{image}}-\frac{1}{f}\right) \tag{1}$$

where λ, D and f represent the wavelength of light, the diameter of the aperture, and the focal length of the system, respectively. Also d_{obj} and d_{image} are the distances of the object and the image planes from the principal planes. If the object is well focused, from the lens equation (Goodman, 1996) the value inside the parenthesis in (1) becomes zero. But when it is deviated from the focus, absolute value of δ increases.

In (Ben-Eliezer et al., 2003), a composition of sparsely distributed FL's, i.e., a composite phase mask (CPM), was used for DOF extension. The CPM is placed at the pupil plane of the imaging system, and hence, the resulting pupil function is as follows:

$$P(u,v) = \sum_k P_k(u,v) \tag{2}$$

where (u,v) are normalized coordinates with respect to the full pupil size, $P_k(u,v)$ is the pupil function associated with the k th FL, and $P(u,v)$ is the overall pupil function which is composition of $P_k(u,v)$'s. $P_k(u,v)$ is defined as follows

$$P_k(u,v) = \begin{cases} M_k(u,v)\exp[-\mathrm{j}\psi_k(u^2+v^2)] & |u|\le 1;|v|\le 1 \\ 0 & otherwise \end{cases} \tag{3}$$

where ψ_k is the dimensionless misfocus parameter defined by Goodman (1996), and $M_k(u,v)$ is defined as

$$M_k(u,v) = \begin{cases} 1, & (u,v)\in k \text{ th FL} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Where $M_k(u,v)$ represents the points belonging to the k th FL. Each point belongs to only one FL. Each FL occupies same area through different k's, and $M_k(u,v)$ is randomly generated.

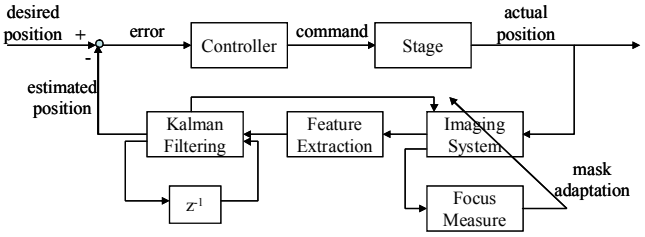

Fig. 1 Overall block diagram

Each FL changes focal length of the original imaging system. Therefore, composition of FL's into CPM can be interpreted as making the imaging system have multiple focal lengths, contributing to DOF increase. Change of the focal length by introducing the k th FL is as follows

$$\frac{1}{f_k} = \frac{1}{f} + \psi_k \frac{4\lambda}{\pi D^2} \tag{5}$$

where f_k is the changed focal length.

Sixteen different ψ_k's were used with fixed interval of two in (Ben-Eliezer et al., 2003). In fact, there is a risk of degrading the obtained image before achieving desired DOF. , On the other hand, if DOF is not extended enough, objects out of DOF might not be well-observed. Complication arises since an object's position is unknown in advance. Therefore in this project, a proposal is made to change the interval of ψ_k's to enable adaptation of the phase mask, while fixing the number of different ψ_k's as follows

$$\psi_k = s_m(-8.5+k), \quad k = 0,\cdots,15 \tag{6}$$

where s_m is the strength of the mask. Adaptive s_m scheme will enable higher quality images and hence improved tracking scheme.

2.2 Focus Measure

Using the fast wavelet transform algorithm (Gonzalez and Woods, 2002), an image is decomposed into the approximation (C_{a1}), the horizontal detail (C_{h1}), the vertical detail (C_{v1}), and the diagonal detail (C_{d1}) (Gonzalez and Woods, 2002). The approximation, C_{a1} , is again decomposed into another approximation (C_{a2}) and details of lower resolution (C_{h2} , C_{v2} , C_{d2}). This is illustrated in Fig. 2.

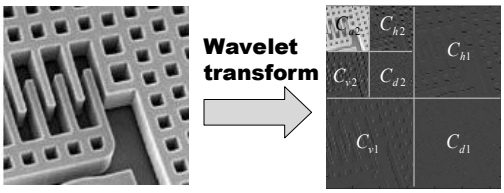

Fig. 2 Illustratioin of wavelet transform

The first details (C_{h1} , C_{v1} , C_{d1}) and the second details (C_{h2} , C_{v2} , C_{d2}) represent the highest and second highest frequency components of the image in each direction. Due to noise filtering, the first details are rejected from focus measure calculation. Instead, the second details are used. The focus measure is calculated by taking Laplacian of C_{h2} and C_{v2} and integrating the squared values as follows:

$$F = \frac{1}{H_2 \bullet W_2}\left[\sum_{H_2}\sum_{W_2}\{L * C_{v2}(x,y)\}^2 + \sum_{H_2}\sum_{W_2}\{L * C_{h2}(x,y)\}^2\right] \tag{7}$$

where H_2 and W_2 denote the height and width of the second highest detail images and L is the Laplacian operator defined by

$$L = \begin{bmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{bmatrix}. \tag{8}$$

Computational time required to calculate focus measure using this method was in average 27 ms for a 128×128 image, when it was implemented in Matlab code (within a Pantium IV of 2.8 GHz, 256 MB RAM).

2.3 Mask Adaptation

Mask adaptation means adjusting the strength of the mask automatically so that the best resolved image can be obtained for a given defocus of the object. Mask adaptation, in other word, is optimizing the strength of the mask to maximize the focus measure. Assuming that δ is constant, this can be formulated as an optimization problem as follows:

$$\hat{s}_m = \arg\min_{s_m}\{F(s_m,\delta)\}$$
$$\text{subject to : } s_m \in [0,\infty) \text{ and } \delta = \text{constant} \tag{9}$$

Constant δ assumption is valid when adaptation is faster enough than δ changes.

To solve the problem, adopting a proper optimization method is important. Because adaptation should be fast enough to be implemented as a subroutine of tracking, the optimization method should be simple and fast. Interesting enough, our investigation revealed that that almost always the plot of F versus s_m is unimodal. Therefore, a simple gradient-based optimization algorithm can be used for mask adaptation. The iterative adaptation rule is as applied until s_m converges. That is

$$s_m(n+1) = s_m(n) + \alpha \left.\frac{\partial F}{\partial s_m}\right|_{s_m(n),\delta} \tag{10}$$

where α is the learning rate. To provide escape from local maxima, it is helpful to start adaptation with large steps.

2.4 Tracking Using Kalman Filter

It is strongly expected that mask adaptation will enhance tracking performance in the case of objects moving out of DOF. To investigate how mask adaptation can improve tracking performance, window based tracking was implemented and tracking results with and without mask adaptation were compared. For tracking, a Kalman filtering approach was adopted. The general state and measurement equation is as follows:

$$\mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n + \boldsymbol{\eta}, \ \mathbf{y}_n = \mathbf{H}\mathbf{x}_n + \boldsymbol{\xi} \tag{11}$$

where $\mathbf{x}$ and $\mathbf{y}$ are the state and measurement vectors, respectively. Also $\mathbf{A}$ and $\mathbf{H}$ are state transition and

measurement matrices, respectively. Furthermore, $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ represent process and measurement noises, respectively. These noises are assumed to be independent and white, with normal probability distribution of

$$p(\boldsymbol{\eta}) \sim N(\mathbf{0}, \mathbf{Q}), \; p(\boldsymbol{\xi}) \sim N(\mathbf{0}, \mathbf{R}) \tag{12}$$

where $\mathbf{Q}$ and $\mathbf{R}$ are process and measurement noise covariance matrices, respectively. Because no information about the motion of the object is provided, constant velocity model was used for the state equation. In this case, $\mathbf{x}$ and $\mathbf{y}$ become

$$\mathbf{x} = \begin{bmatrix} x & \dot{x} & y & \dot{y} \end{bmatrix}^{\mathrm{T}}, \; \mathbf{y} = \begin{bmatrix} x & y \end{bmatrix}^{\mathrm{T}} \tag{13}$$

and the state transition and measurement matrices become

$$\mathbf{A} = \begin{bmatrix} 1 & dt & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & dt \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{14}$$

Tracking is performed in two steps: 'predict' and 'correct'. Fig. 3 shows how Kalman filter operates.

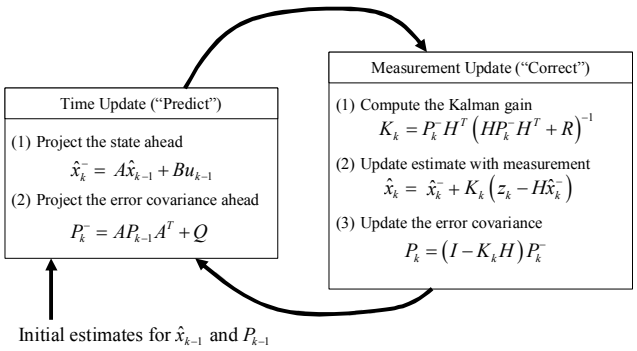

Fig. 3 Operation of Kalman filter .

3. SIMULATION RESULTS

3.1 Mask Adaptation

The gradient based mask adaptation algorithm shown in Eq. (10) was implemented. The learning rate α was set to 0.1 and the stopping condition was $\partial F / \partial s_m < 0.02$. A MEMS image was used in the simulations. Fig. 4 to 6 show an example of mask adaptation. In this simulation, it was observed that adaptation is completed within 5 iterations in many cases and mostly within 10 iterations. Since twice measurement of focus is needed for one iteration of adaptation to obtain the gradient, and about 27 ms is required for calculating focus measure of 128×128 image, one can say that adaptation can be performed faster than 2Hz.

3.2 Tracking

Tracking was performed by Kalman filtering. Process and measurement noise covariance matrices were set to $\mathbf{Q} = \mathbf{I}_{4\times4}$ and $\mathbf{R} = \mathbf{I}_{2\times2}$, respectively. Initial position of the feature was approximately given by the user and initial velocity was set to zero. The initial error covariance matrix was set to $\mathbf{P}_1 = 3\mathbf{I}_{4\times4}$. Sum of Squared Differences (SSD) based template matching method was used for feature extraction. The location of the searching window was determined by prediction of $\mathbf{x}$, and its size changed according to norm of $\mathbf{P}_k$.

Fig. 7 shows the moving trajectory of a MEMS object, and template for feature extraction. The tracking result without mask adaptation is shown in Fig. 8 and its error histories are shown in Fig. 9. From the result, it can be seen that large feature extraction error due to insufficient visual information also leads to large tracking error.

The tracking result with mask adaptation is shown in Fig. 10 and its error histories are shown in Fig. 11. Interestingly both feature extraction and tracking errors are reduced.

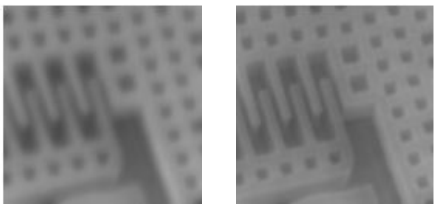

Fig. 4 Mask adaptation result.

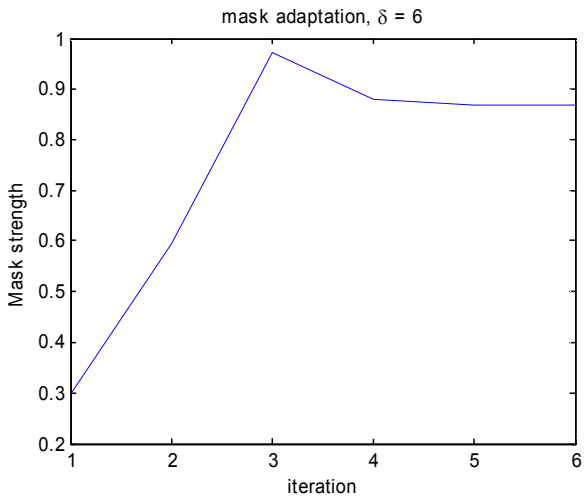

Fig. 5 Change of s_m .

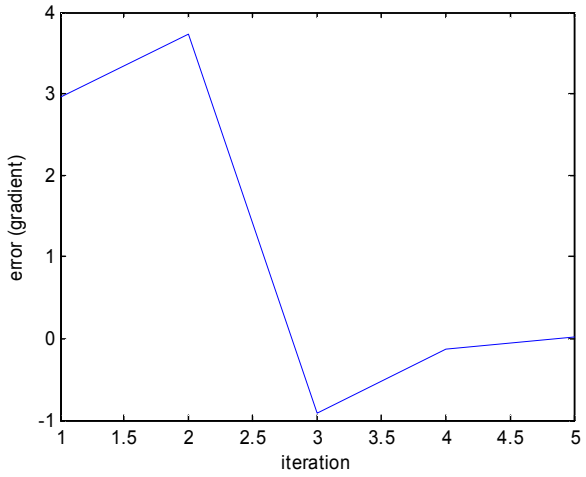

Fig. 6 Error plot of mask adaptation.

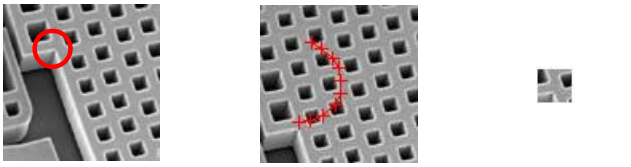

Fig. 7 Object trajectory and template.

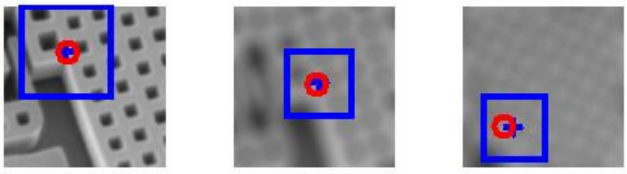

Fig. 8 Tracking result without mask adaptation.

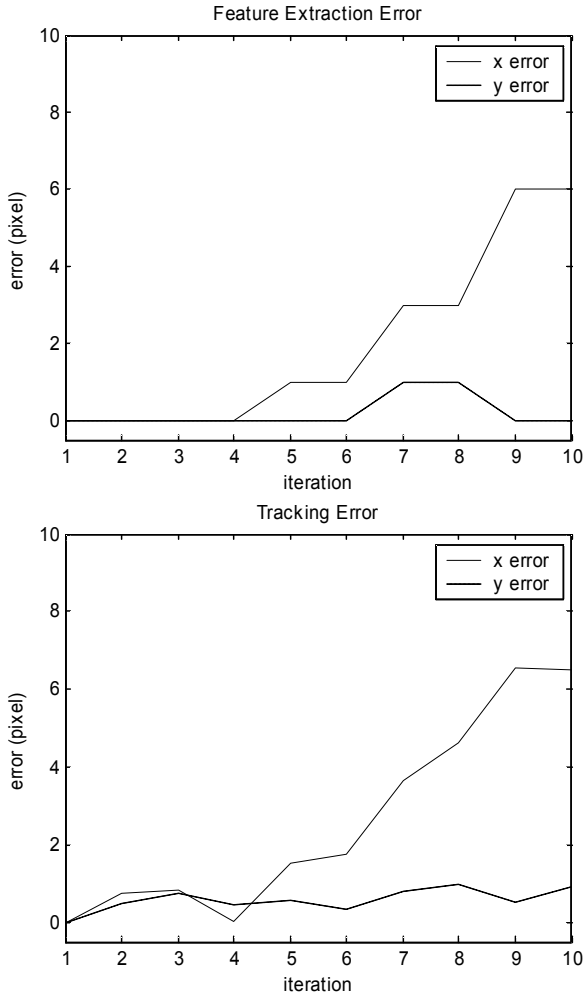

Fig. 9 Feature extraction and tracking error without mask adaptation.

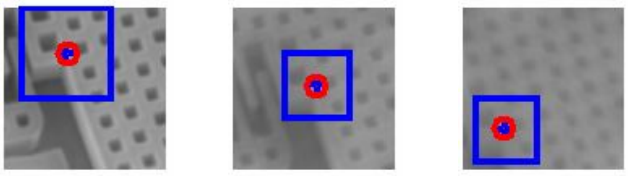

Fig. 10 Tracking result with mask adaptation.

4. CONCLUSION

To solve small DOF problem in MEMS visual servoing, DOF adaptation method was proposed.. DOF adaptation was done by adjusting the strength of the phase mask which was placed at the pupil plane. For the adaptation, a focus measure based on wavelet transform was proposed in this work. Based on the developed focus measure, gradient-based mask adaptation was simulated and an adaptation speed of 2Hz was achieved. The speed can be increased further by reducing the size of the area where focus should be measured and also by using more powerful processor. Because focus measure vs. mask strength curve is almost unimodal, local maxima problem was not

encountered in the simulations. Simulation results of tracking supported the reduction in the feature extraction errors enhanced tracking performance. For the future works, this concept should be examined by experiments. Work will be conducted to further improve the computational complexity of the focus measure.

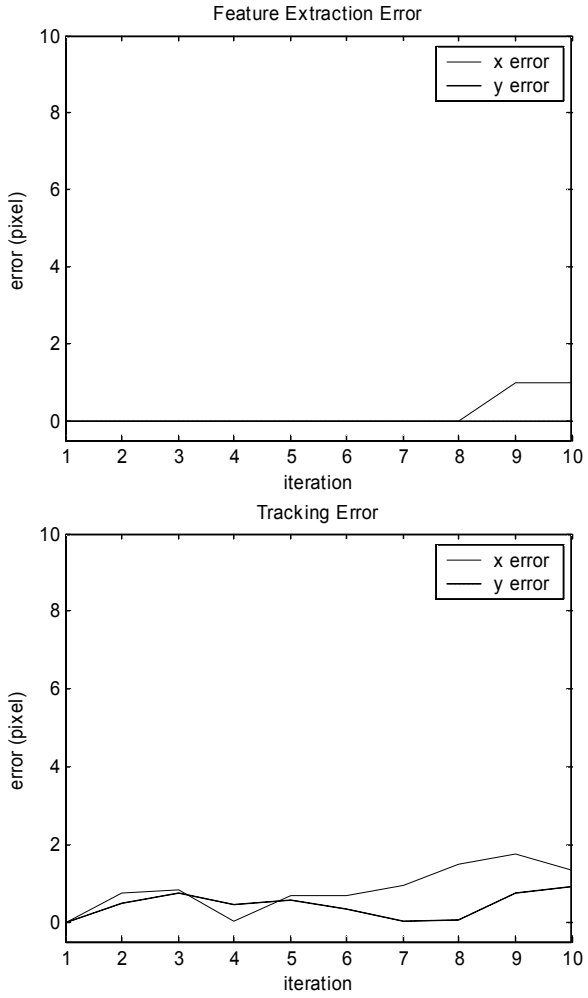

Fig. 11 Feature extraction and tracking error with mask adaptation.

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**, pp. 164-169.

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

Ben-Eliezer, E., E. Marom, N. Konforti, Z. Zalevsky (2006). Radial mask for imaging systems that exhibit high resolution and extended depth of field, *Applied Optics*, **45**(9), pp. 2001-2013.

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

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

Fatikow, S., J. Seyfried, St. Fahlbusch, A. Buerkle, F. Schmoeckel and H. Woern (2000). Intelligent Microrobotic System for Microassembly Tasks, *1st Int. Conference on Mechatronics and Robotics*, St.-Petersburg, Russia, May 29-June 2, pp. 22-26.

Gonzalez, R., R. Woods (2002). *Digital Image Processing*, 2nd, chap.7, pp. 349-408, Prentice-Hall, New Jersey

Goodman, J., (1996). *Introduction to Fourier Optics*, pp. 126–51, McGraw-Hill, New York.

Ojeda-Castaneda, J., R. Ramos, and A. Noyola-Isgleas (1988). High focal depth by apodization and digital restoration, *Applied Optics*, **27**, pp. 2583-2586.

Ojeda-Castaneda, J., E. Tepichin, and A. Diaz (1989). Arbitrary high focal depth with a quasioptimum real and positive transmittance apodizer, *Applied Optics*, **28**, pp. 2666-2670.

Tao, X., H. Cho, Y. Cho (2005). Microassembly of peg and hole using active zooming, *Proceedings of SPIE*, **6052**, pp. 605204-1 - 605204-12.

Woern, H., J. Seyfried, St. Fahlbusch, A. Buerkle, F. Schmoeckel (2000). Flexible microrobots for micro assembly tasks, *Micromechatronics and Human Science, 2000, Proceedings of 2000 International Symposium on*, Karlsruhe Univ, pp. 135-143.