Continuous Focusing of Moving Objects Using DFD1F

Tse-Chung Wei and Murali Subbarao
Department of Electrical Engineering
State University of New York at Stony Brook
Stony Brook, New York 11794-2350

Abstract

A method is presented for continuous focusing of moving objects. It is based on DFD1F [20] which is a method for determining distance of stationary objects using image defocus information. The proposed method requires recording of two images of a moving object with different degrees of blur. The change in blur is caused by varying camera parameters such as lens position, focal length, and aperture diameter of the recording camera. The two images must be recorded simultaneously in a short time period. A new camera structure is proposed for such recording of the images. In the proposed method, the requirement of a large memory space has been avoided for storing the MTF data of the camera's optical system. This is achieved by using a parameterization scheme for the MTF data. The method has been implemented on an actual camera system. Experimental results on this system indicate that the method yields an RMS error in focusing of about 4.3% in lens position. The image blur caused by a focusing error of this magnitude is barely noticeable by humans. Therefore, in addition to machine vision, the method has practical applications in video cameras such as camcorders.

1 Introduction

In an imaging system such as a convex lens, the image of an object in front of the lens will be usually blurred. The degree of blur depends on the focal length f of the lens, the distance u of the object from the camera and the distance s between image detector and the lens (see Fig. 1). A well-known relation is the lens formula

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \tag{1}$$

Following this formula, a focused image is obtained under the condition that s equals v. The degree of blur increases as the difference between s and v increases. If f and v are known, then the distance (depth) u of the object can be found using this formula.

In *Depth-from-Focus* (DFF) approach [1, 2, 3, 4, 12], a search is made for the lens position s or/and the focal length f which brings a stationary object to focus. This involves acquiring about 10 images with different s or/and f and finding the image which is in best focus. This approach is slow due to the mechanical motion of camera parts to change s or/and f to record the required 10 or so images.

In *Depth-from-Defocus* (DFD) approach [5, 13, 14, 18, 7, 15, 20] there is no need to search for s or/and f values which correspond to focusing the object. The level of defocus is used in determining distance. This approach involves processing only a few images (about two or three) as compared to a large number (about 10) of images in the DFF methods. In addition, only a few images are sufficient to determine the distance of all objects in a scene irrespective of whether the objects are focused or not. Therefore this method is much faster than DFF due to the reduction in the mechanical motion of camera parts.

The DFD approach has been successfully applied to high contrast step edges [5, 6, 8, 16, 17]. Application of DFD to arbitrary objects has been investigated by many researchers [5, 13, 14, 18, 7, 15, 20].

The problem of continuous focusing of a moving object using DFD approach has not been investigated in the previous literature. In this paper we propose a

camera structure and a method for this problem. The method is based on a DFD method called DFD1F for stationary objects proposed by us [20].

The major distinguishing features of DFD1F in comparison with prior DFD approaches are- (i) it requires the computation of only a few (about 6) and that too only one-dimensional Fourier coefficients (hence the suffix 1F in DFD1F), (ii) it is general in that it is not restricted to any particular model of the point spread function of the camera system, (iii) only a few (two or three) images acquired with different camera parameter settings are needed, (iv) there is no restriction on the camera parameter settings such as pin-hole aperture, etc., and (v) the method has been demonstrated on a very large database of natural images.

Finding the distance of a moving object at a given time instant using a DFD method requires the simultaneous recording of two images of the object at the given time instant. The images have to be recorded with different camera parameter settings. A camera system is proposed to accomplish this task.

A straightforward adaptation of DFD1F for moving objects would require a large amount of memory space. The memory space is mainly needed to store a large look-up table representing the low-frequency MTF data of the optical system of the camera. A parameterization scheme is proposed for reducing the memory requirement.

The method proposed here has been implemented on an actual camera system. The results of experiments on this system are reported here. The results indicate that the method has a root-mean-square (RMS) error of about 4.3% in the lens position for focusing. The blurring due to this focusing error is barely noticeable by the human vision system. Therefore the method has practical applications in both machine vision and consumer video cameras. We believe that the accuracy can be improved by using more accurate camera MTF and parameter data than the one used in our experiments. The proposed method has also been tested on simulated image data. Very accurate results were obtained in this case.

2 DFD1F

In this section we briefly describe DFD1F and present the background material for the proposed method for continuous focusing of moving objects. Detailed description of DFD1F can be found in [20].

2.1 Camera Model

A model for a camera with variable parameters is shown in Fig. 2. It consists of an optical system with two lenses L1 and L2. The effective focal length f is varied by moving one lens with respect to the other. O.A. is the optical axis, P1 and P2 are the principal planes, Q1 and Q2 are the principal points, ID is the image detector, D is the aperture diameter, s is the distance between the second principal plane and the image detector, u is the distance of the object from the first principal plane, and v is the distance of the focused image from the second principal plane.

Object distance u is the only unknown parameter. The other parameters s, f, and D are adjustable in the camera. We will refer these adjustable parameters together as *camera parameters* and denote them by $\mathbf{e}$, i.e.

$$\mathbf{e} \equiv (s, f, D). \tag{2}$$

We take the optical system to be circularly symmetric around the optical axis. In order to illustrate the theoretical basis of our method, we use a paraxial geometric optics model [9] for image formation. This is a good approximation in practice to the actual image formation process.

2.2 PSF and MTF

According to paraxial geometric optics, a point light source at point P in Fig. 2 will form a *blur circle* on the image detector for a circular aperture. Let R be the radius of the blur circle. Using the lens formula (1) and Fig. 1 we obtain

$$R = s\frac{D}{2}\left[\frac{1}{f} - \frac{1}{u} - \frac{1}{s}\right] \tag{3}$$

This radius R is a function of object distance u and the camera parameters $\mathbf{e}$ defined earlier in Eq. (2). Therefore it can be denoted by $R(\mathbf{e}, u)$. The Point Spread Function (PSF) corresponding to the blur circle is a cylindrical function of the form:

$$h_a(x, y; \mathbf{e}, u) = \begin{cases} \frac{1}{\pi R^2(\mathbf{e},u)} & \text{if } x^2 + y^2 \leq R^2(\mathbf{e}, u) \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

The Fourier transform of the above PSF is the Optical Transfer Function (OTF):

$$H_a(\omega, \nu; \mathbf{e}, u) = 2\frac{J_1(R(\mathbf{e}, u)\ \rho(\omega, \nu))}{R(\mathbf{e}, u)\ \rho(\omega, \nu)} \tag{5}$$

where ω, ν, and ρ are spatial frequencies specified in radians per unit distance, with $\rho(\omega, \nu) = \sqrt{\omega^2 + \nu^2}$, and J_1 is the first order Bessel function. The magnitude of the OTF is the Modulation Transfer Function (MTF).

As an alternative to the PSF derived above using geometric optics, a two-dimensional Gaussian has been proposed [5, 13, 10, 11]. The spread parameter σ of this Gaussian is a function of object distance u and camera parameters $\mathbf{e}$. Except when the blur circle radius R is very small, a good approximation is the relation $\sigma = R/\sqrt{2}$ [17]. Therefore we can write

$$\sigma(\mathbf{e}, u) = s\frac{D}{2\sqrt{2}}\left[\frac{1}{f} - \frac{1}{u} - \frac{1}{s}\right] \tag{6}$$

The Gaussian PSF is

$$h_b(x, y; \mathbf{e}, u) = \frac{1}{2\pi\sigma^2(\mathbf{e}, u)} \exp\left(-\frac{x^2 + y^2}{2\sigma^2(\mathbf{e}, u)}\right) \tag{7}$$

and the corresponding OTF is

$$H_b(\omega, \nu; \mathbf{e}, u) = \exp\left(-\frac{1}{2}(\omega^2 + \nu^2)\ \sigma^2(\mathbf{e}, u)\right) \tag{8}$$

Fig. 4 shows a plot of the MTFs according to geometric optics, corresponding to a Gaussian PSF, and the actual MTF of the lens used in our experiments (this MTF data was obtained by the lens manufacturer through computer simulation).

2.3 Finding Distance

The image formation in a camera system can be thought of as the superposition of the images of an infinite number of point light sources lying on visible surfaces

in the 3D scene. In general, the blur circle radius changes from point to point due to variation in distance. For a flat region perpendicular to the optical axis, the distance u is a constant inside the region. Therefore, PSF for all points in this region will remain the same. Consequently, the image formed on the image detector can be modeled by a two-dimensional convolution of the PSF and the brightness distribution or focused image in that region. Let $f(x, y)$ be the focused image for that flat region with constant distance u. In DFD1F, the camera parameter setting is first set to $\mathbf{e_1}$, and a blurred image $g_1(x, y)$ is obtained. The PSF $h_1(x, y)$ for this blurring process may be modeled by Eq. (4) or Eq. (7). Similarly, a second image $g_2(x, y)$ with a different degree of blur is obtained by another camera setting $\mathbf{e_2}$. In the frequency domain we have

$$\begin{aligned} G_1(\omega, \nu) &= H(\omega, \nu; \mathbf{e_1}, u)\, F(\omega, \nu) \quad \text{and} \\ G_2(\omega, \nu) &= H(\omega, \nu; \mathbf{e_2}, u)\, F(\omega, \nu) \end{aligned} \tag{9}$$

The effect of the focused image F can be cancelled by taking the ratio of G_1 and G_2:

$$\frac{G_1(\omega, \nu)}{G_2(\omega, \nu)} = \frac{H(\omega, \nu; \mathbf{e_1}, u)}{H(\omega, \nu; \mathbf{e_2}, u)} \tag{10}$$

The left hand side above can be obtained by dividing the Fourier transforms of the two blurred images. The right hand side can be pre-computed and stored in a table as a function of the index variables ω, ν and distance u (the camera settings $\mathbf{e_1}$ and $\mathbf{e_2}$ remain the same). The distance u can be solved by searching for the value of u in the precomputed table where the left hand side and right hand side are almost equal.

Theoretically, only one data point (specified by ω and ν) may be sufficient to solve for u. In practice, using several data points will be needed for robustness against noise. If we use only the points along ν axis then $\omega = 0$ in Eq. (10). The Fourier coefficients along ν axis can be computed by first summing up each row of the image to obtain a one-dimensional signal and then performing a one-dimensional Fourier transform on this signal. The processing is then reduced to one-dimensional for a two-dimensional image.

Instead of solving Eq. (10) directly, in DFD1F both sides of the equation are transformed by first taking logarithm and then dividing by $\omega^2 + \nu^2$. This has been

called the log-by-rho-squared ($\log/\rho^2$) transform [20]. The resulting expression is then solved by a table look up procedure. This has some practical advantages. Also, only low spatial frequency data points are used to reduce the effects of noise and zero-crossings of the OTF.

3 Continuous Focusing of Moving Objects

In DFD1F, the camera parameters have to be changed after recording the first image g_1 but before recording the second image g_2. In most camera systems, this takes a few seconds of time since some mechanical parts (e.g. lens, aperture, etc.) have to be moved. In the case of moving objects this time delay is unacceptable because the object would have changed its position during the delay period. The images g_1 and g_2 must correspond to the same position of the object. Therefore the two images have to be recorded simultaneously in a short period of time. Fig. 3 shows camera structures for accomplishing this.

In Fig. 3(a), two identical cameras and a beam splitter with a mirror are used. The cameras are identical in all respects except that their camera settings are different.

In Fig. 3(b), a single lens with two image detectors (ID1 & ID2) and a beam splitter are used. There is a displacement of $2\triangle S$ between the two image detectors. This will serve as the change in lens position for the camera settings $\mathbf{e}_1$ and $\mathbf{e}_2$. Since the purpose of ID2 is to obtain one of the two images used by DFD1F, it can be a smaller CCD array as long as it is large enough to provide the image block used in focusing. In our experimental setting, a 128×128 CCD will be sufficient. A linear CCD with elongated sensor elements (fig. 3(f)) can also be used. The wider pixel size will automatically sum up the image along one direction to obtain the one-dimensional signal we need. The other image needed for DFD1F will be a subimage of the larger CCD. In this case, it will be a subimage from ID1.

Figure 3(c) shows a minor variation of Fig. 3(b) where two beam splitters and an additional image detector are used. The image detectors marked ID1 and ID2 can be smaller, for their purpose is obtain the two images used by DFD1F. In this case, image summation along one direction can be done by using a regular linear CCD (with square sensor elements) by rotating the half silvered mirror hsm1 (about the

point Q) by a small angle.

In Fig. 3(d), a beam splitter and a lens are used for varying the focal length. The two images used in focusing will be the image on ID2 and a subimage from ID1. Therefore, ID2 can be made smaller.

In Fig. 3(e), a beam splitter and an aperture are used for changing the aperture. This setup is similar to Fig. 3(d), except that the variation is aperture diameter instead of focal length.

It is possible to combine the features of Fig. 3(b, d & e) in a single camera where $\mathbf{e}_1$ and $\mathbf{e}_2$ are different in more than one camera parameter.

3.1 Parameterization of MTF

In this paper we consider the case of changing lens position s for continuous focusing. An alternative to this (which is employed by the human vision system) is to change the focal length f. In order to continuously change s for continuous focusing, it is necessary to store the MTF data for every possible value of s for each possible object distance. This would require a large memory space. For example, in our camera system, this would require storing about 6x100x100 floating point numbers. Here we propose a scheme for parameterizing the MTF data so that the memory requirement is drastically reduced (to about 100 floating point numbers for our camera).

The problem of storing MTF data arises because, the two PSF models based on paraxial geometric optics (Eq. 4) and and a Gaussian (Eq. 7) are not sufficiently accurate for actual camera systems. The two PSF models are good approximations but not adequate. This is particularly true for small F-number (less than 8) cameras.

The motivation for our parameterization scheme is based on the observation that the MTF of our camera is roughly if not exactly a Gaussian for low frequencies. This is clear from Figure 4. It shows a plot of a typical MTF (marked Lens Data) for the lens used in our experiments. Figure 5 shows a plot of σ obtained by applying the $\log/\rho^2$ transform to the lens MTF and then taking square root:

$$\sigma(\rho, \mathbf{e}, u) \;=\; \sqrt{\frac{-2}{\rho^2} \ln H_c(\rho; \mathbf{e}, u)} \tag{11}$$

where H_c represents the lens MTF. In this plot we see that σ is almost a constant with respect to ρ for low frequencies. The average value of σ in the range $\rho = 1$ to $\rho = 3.6$ is taken as a measure of blur in one version of DFD1F. It can be used as a spread parameter of the PSF. The plot of a Gaussian MTF with this average σ is shown in Figure 4 for comparison. We see that the Gaussian is close to the lens MTF for low frequencies (up to $\rho = 3.6$). As expected, applying the $\log/\rho^2$ transform to the Gaussian MTF and taking the square root gives a function which is constant with respect to ρ for all frequencies (see Fig. 5). Figures 4 and 5 also show the plots of the MTF as predicted by paraxial geometric optics model and the square root of the $\log/\rho^2$ transform of the MTF. In this case the radius of the blur circle is taken to be $\sqrt{2}$ times the σ of the Gaussian.

From the above discussion we conclude that in practical applications a blur parameter σ can be defined as in Eq. (11) which is almost a constant with respect to low frequencies. This can be used to characterize the MTF data of a lens system.

If the lens MTF is exactly a Gaussian, then from Eq. 6 we have:

$$\sigma = \frac{Ds}{2\sqrt{2}} \left(\frac{1}{f} - \frac{1}{u} - \frac{1}{s} \right) \tag{12}$$

The minimum value of s is equal to f because objects at infinity come to focus at $v = f$ (see lens formula) and all other objects come to focus at $v > f$. Normalizing the magnitude of σ corresponding to $s = s_0$ [20] and using the lens formula, we obtain

$$\sigma = \frac{Ds_0}{2\sqrt{2}} \left(\frac{1}{v} - \frac{1}{s} \right) \tag{13}$$

Assuming $s_0 = f$ and using the approximation $vs \approx f^2$ which holds for most cameras, we obtain

$$\sigma = \frac{s - v}{2\sqrt{2}F\#} \tag{14}$$

where $F\#$ denotes the F-number (equal to f/D) of the camera. From the above expression we see that σ is directly proportional to the difference $s - v$ and inversely proportional to the F-number of the camera.

In the camera used in our experiments, the parameter s is changed by moving the lens with respect to the image detector. The lens motion is effected by a stepper motor with 97 steps numbered 0 to 96. When the lens is at step 0, except for

assembly error, $s = f$ and therefore an object at infinity will be focused on the image detector. When the lens is moved to the other end (lens step 96), an object at a distance of about 0.55 meter will be focused. Each lens step corresponds to a relative displacement of the lens with respect to the image detector of about 0.025mm. For each lens position there corresponds a unique object distance for which the object will be in best focus on the image detector. For example, except for a constant due to assembly error between the lens and the camera, Table 1 shows the relation between the lens position and the distance of an object in best focus. It is therefore convenient to specify object distances in terms of lens position. For example, according to Table 1, if the distance of an object is said to be step 30, it means that the object is at distance 1.320 meters from the camera.

The relation between the lens position s (in mm) and the corresponding step number i is given by $s = f + i\ \delta s$ where δs is the lens displacement in mm per step of the motor. Similarly, if j specifies object distance in lens step, then we have $v = f + j\ \delta s$. Using these relations and Eq. 14 we can write

$$\sigma = \frac{\delta s}{2\sqrt{2}F\#}(i - j) \tag{15}$$

The above relation implies that the blur parameter σ which characterizes the lens MTF varies linearly with respect to both lens position i and object distance j with the same proportionality constant or slope. It is found that, after a minor modification, this model holds well for the actual lens MTF. The modification that is needed is the addition of a constant σ_{min}. This can be justified as follows. When an object is in best focus we have $i = j$ in the above expression. Therefore σ is zero according to this model. However, in practice, even when a point light source is in best focus, its' image is not a point which is dimensionless but the wellknown Airy pattern (bright and dark rings due to diffraction). Further, when the aperture is large (F-number of less than 8), the paraxial assumption (i.e. the light rays incident on the lens are almost parallel to the optical axis) does not hold. Therefore, instead of a single point where all rays from a point source are focused, there is only what is called a *circle of least confusion* which has finite dimensions. For these reasons, we propose the following model for the blur parameter σ

$$\sigma = \sigma_{min} + K\,|i - j| \tag{16}$$

(Another alternative model is $\sigma^2 = \sigma_{min}^2 + K^2(i-j)^2$.) The parameters σ_{min} and K are both approximately inversely proportional to the F-number of the camera system.

We next show some plots for the MTF data of our camera which indicate that the above model for σ is a good approximation to our camera.

Figure 6 shows a plot of the MTF data for the lens system used in our experiments. In this plot, one axis corresponds to spatial frequency ρ and the other axis corresponds to distance of the object being imaged. (The distance of the object is specified in terms of the lens position or lens step number.)

Figure 7 shows a plot of the MTF data in Figure 6 after the $\log/\rho^2$ transform and taking square root. Here we see that, as before, the value of σ is almost a constant with respect to low frequencies for all object distances (lens steps 0 to 90). Therefore, σ has been averaged with respect to low frequencies and the resulting plot is shown in Figure 8. Here we see that σ is almost linear with respect to object distance specified in lens step. σ has a minimum value σ_{min} at step 40 corresponding to the distance of the best focused object. On either side of the minimum, the slope is almost the same.

Figure 8 corresponds to an object distance of step 40. Similar plots have been obtained for object distances 0,10,20,...,90, and are shown as a 3D plot in Figure 9. We see that σ is a minimum along the diagonal and varies linearly on either side of the diagonal. The minimum value σ_{min} along the diagonal is almost a constant. The axes in this plot are lens position and object distance specified in step numbers. The slope on either side of the diagonal are almost the same. This implies that σ depends only on the difference between lens step i and object step j. These plots indicate that our proposed model (Eq. 16) can be used for practical camera systems. For the plot data in Figure 9, $\sigma_{min} = 8.904 \times 10^{-4}$ and $K = 1.343 \times 10^{-3}\mathrm{mm}$.

4 Computational Steps

A flow chart of the algorithm for continuous focusing is shown in Fig. 10. Initially the lens is moved to step 15 which corresponds to focusing an object at about 3 meters distance. The variable Lens_Step in the flow chart corresponds to the

position of the lens at any given instant. The stored table $\overline{T}_s[i]$ is computed next for two camera settings $\mathbf{e}_1$ and $\mathbf{e}_2$. In our experiments, the only camera parameter that was different for the two camera settings was the lens position. The first one was Lens_Step-15 and the second one was Lens_Step+15. In principle, other parameters such as focal length and aperture diameter, could also be varied. The parameters could be varied either one at a time or, two or more simultaneously. In our experiments, the stored table was computed as

$$\overline{T}_s[i] \;=\; [\sigma_{min} + K\ |i - (Lens_Step - 15)|]^2 - [\sigma_{min} + K\ |i - (Lens_Step + 15)|]^2 \quad (17)$$

Two images g_1 and g_2 are recorded corresponding to the camera settings $\mathbf{e}_1$ and $\mathbf{e}_2$. The image size in our experiments was 128×128. Both were summed along rows to obtain one-dimensional signals.

The two images g_1 and g_2 are normalized as in DFD1F with respect to mean brightness. In our implementation, normalization with respect to magnification was not done as the change in magnification was small (about 2-3%).

A few low frequency Fourier coefficients of g_1 and g_2 are computed. In the experiments, the first 6 coefficients were computed. The table $T_c[j]$ is then computed using the $\log/\rho^2$ transform as

$$T_c[j] = \frac{-2}{\rho_j^2} \ln \left| \frac{G_1(\rho_j)}{G_2(\rho_j)} \right| \quad (18)$$

Next the mean value of $\overline{T}_c$ is computed as

$$\overline{T}_c = \frac{1}{k} \sum_{j=1}^{k} T_c[j] \quad (19)$$

The mean $\overline{T}_c$ is compared with the stored table values $\overline{T}_s[i]$ and the index i for which the two values are closest is found. This index gives the lens step position for focusing the object. The index is also used to find the actual distance of the object through another table lookup. The lens is moved to the focusing step position and the variable Lens_Step is set to this new position.

Next the above algorithm is repeated beginning from the computation of $\overline{T}_s[i]$

The algorithm terminates when the camera power is turned off.

5 Experiments

Three poster pictures–FACE, NAVY, and OPTCON– shown in Figures 11, 12, and 13 respectively, were used as test objects. The reason for using planar objects is that it simplifies error analysis in the image window being processed. The estimated distance of the object can be compared with the actual distance to compute RMS errors. For 3D objects with depth variation in the image window of interest, the estimated distance will be some kind of "average" of different points in the window.

Experiments were done under the following camera settings: focal length: 35 mm, F#: 4, White balance: off, Gamma compensation: off, camera gain control: +6dB, illumination: 300 Lux. Each picture was placed at 24 different positions in sequence at time instants 1 to 24. The initial position was about 1 meter from the camera. The object was then moved gradually closer to the camera to a distance of about 0.6 meter. Next the object was moved gradually away up to a distance of about 5 meters. Then the object was moved back in steps to about 1 meter from the camera.

In Figure 14, the plot labeled 'Actual' shows the actual distance (in lens step number) of the "moving" object at different time instants. The estimated distance at each time instant for the three objects are plotted in Figure 14. We see that at the beginning there is a kind of "warm-up" period when the errors are relatively large. This is because, at the beginning the lens position (at step 15) was very far from the focused lens position (around step 55). Therefore the recorded images were highly blurred resulting in more error. After a few time instants, the camera "locks" onto the "moving" object and continuously focuses onto the object. During this "locked" period, focusing error is small because the lens position is not too far from the focused position and therefore the recorded images are less blurred.

In the beginning, no matter where the object is, the initial lens position will be at step 15. At each time instant, the camera records two images, one at 15 steps behind and another at 15 steps ahead of the current lens position. Using these images, it estimates the distance of the object and moves the lens to focus it. After moving the lens, it again records two more images and repeats the process. There are $24 \times 3 = 72$ data points in Fig. 14. The RMS error based on these 72 focusing results

is about 4.2 lens steps out of 97 steps, or about 4.3%. The image blur due to a lens position error of this magnitude is small and is not easily noticeable by humans. Therefore, in addition to machine vision, the method is useful in camcorders.

Figure 15 shows the results of experiments on simulated image data. Paraxial geometric optics model of image formation was used to compute the blurred images corresponding to the three images in Figures 11, 12, and 13. We see that the focusing results are very good as expected.

6 Conclusion

We have described a camera structure and method for continuous focusing of moving objects. The method is based on the DFD1F method which uses image defocus information. Experimental results show that the method is useful in both machine vision and consumer video cameras.

References

[1] B. K. P. Horn, "Focusing", Artificial Intelligence Memo No. 160, MIT, 1968.

[2] R. A. Jarvis, "A perspective on range finding techniques for computer vision", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-5, No. 2, pp. 122–139, March 1983.

[3] E. Krotkov, "Focusing", *International Journal of Computer Vision*, 1, 223-237, 1987.

[4] S. Nayar, "Shape from Focus System for Rough Surfaces", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Champaign, Illinois, pp. 302-308, June 1992.

[5] A. P. Pentland, "A new sense for depth of field", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-9, No. 4, pp. 523–531, 1987.

[6] P. Grossman, "Depth from focus", *Pattern Recognition Letters* 5, pp. 63–69, Jan. 1987.

[7] J. Ens and P. Lawrence, "An Investigation of Methods for Determining Depth from Focus", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-15, No. 2, pp. 97–108, February 1993.

[8] S. H. Lai, C. W. Fu and S. Chang, "A Generalized Depth Estimation Algorithm with a Single Image" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-14, No. 4, pp. 405–411, April 1992.

[9] J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics,* John Wiley & Sons, New York, 1978.

[10] B. K. P. Horn, *Robot Vision,* McGraw-Hill Book Company, 1986.

[11] W. F. Schreiber, *Fundamentals of Electronic Imaging Systems,* Springer-Verlag, Section 2.5.2., 1986.

[12] M. Subbarao and T. Choi, "Focusing Techniques", *Proceedings SPIE,* Boston, Massachusetts, Vol 1823, pp. 163–174, November 1992.

[13] M. Subbarao, "Parallel depth recovery by changing camera parameters", *Second International Conference on Computer Vision,* Florida, USA, pp. 149-155, December 1988.

[14] M. Subbarao, "Efficient depth recovery through inverse optics", Editor: H. Freeman, *Machine Vision for Inspection and Measurement,* Academic press, Boston, pp. 101-126, 1989.

[15] M. Subbarao and G. Surya, "Application of Spatial-Domain Convolution/Deconvolution Transform for Determining Distance from Image Defocus", *Proceedings SPIE,* Boston, Massachusetts, Vol 1822, pp. 159–167, November 1992.

[16] M. Subbarao, and G. Natarajan, "Depth recovery from blurred edges", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Ann Arbor, Michigan, pp. 498-503, June 1988.

[17] M. Subbarao, "Determining distance from defocused images of simple objects", Tech. Report No. 89.07.20, Computer Vision Laboratory, Dept. of Electrical Engineering, State University of New York, Stony Brook, NY 11794-2350.

[18] M. Subbarao, "Computational methods and electronic camera apparatus for determining distance of objects, rapid autofocusing, and obtaining improved focus images", U.S. patent application serial number 07/373,996, June 1989 (pending).

[19] M. Subbarao, and M. C. Lu, "Computer Modeling and Simulation of Camera Defocus", *Proceedings SPIE*, Boston, Massachusetts, Vol 1822, pp. 110–120, November 1992.

[20] M. Subbarao, and T. Wei, "Depth from Defocus and Rapid Autofocusing: A Practical Approach", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Champaign, Illinois, pp. 773-776, June 1992. (Detailed version of the paper is submitted to IEEE–PAMI)

Lens Step	0	5	10	15	20	25	30	35	40	45
Distance(m)	∞	5.300	3.750	2.850	2.500	1.930	1.720	1.465	1.320	1.170
Lens Step	50	55	60	65	70	75	80	85	90	95
Distance(m)	1.080	0.965	0.900	0.822	0.770	0.715	0.670	0.628	0.595	0.560

Table 1. Lens Step vs Best Focused Distance

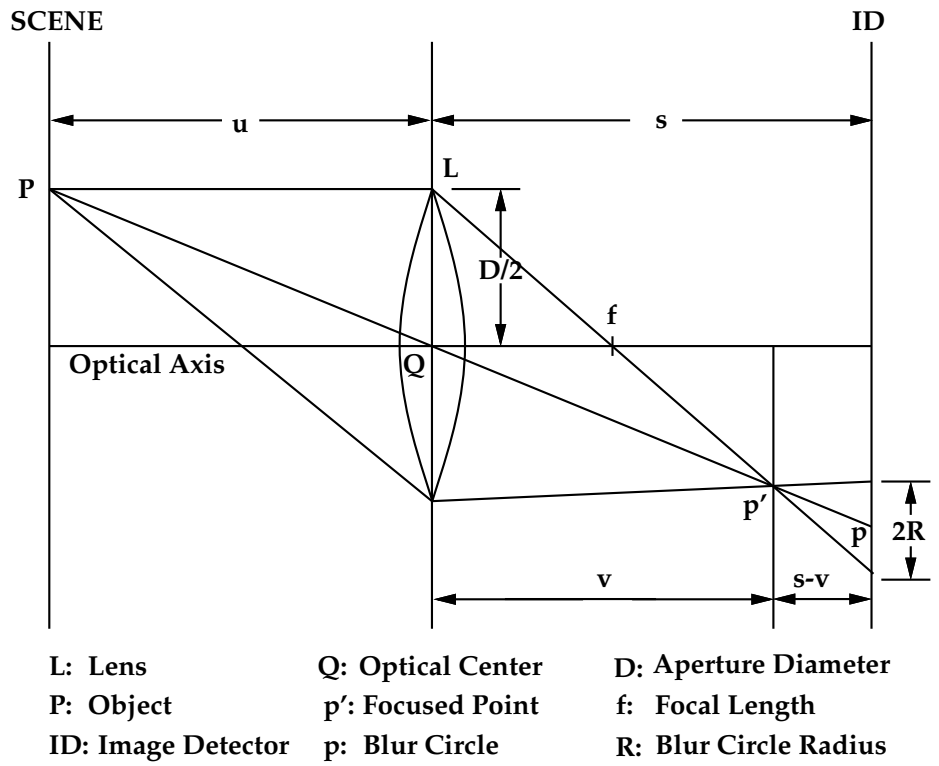


Fig. 1 Image Formation in a Convex Lens

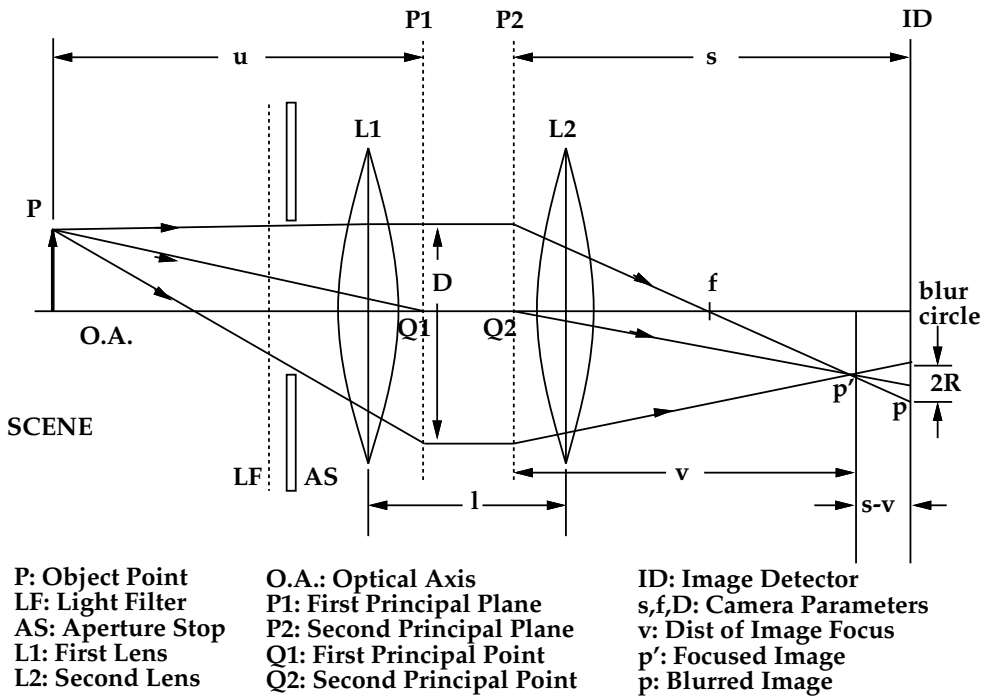
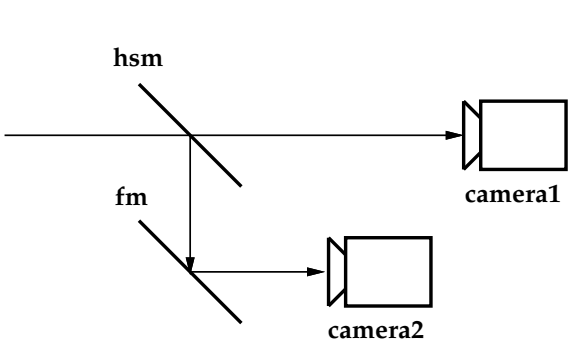
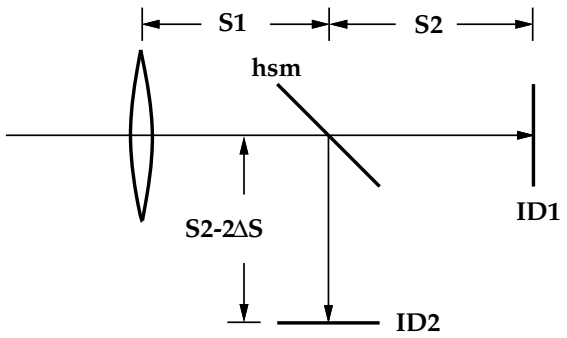


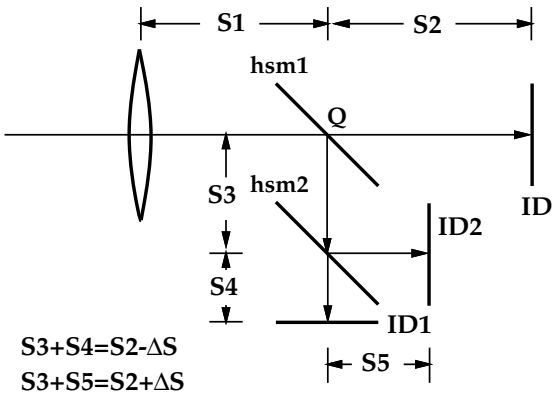
Fig. 2 Camera Model and Camera Parameters



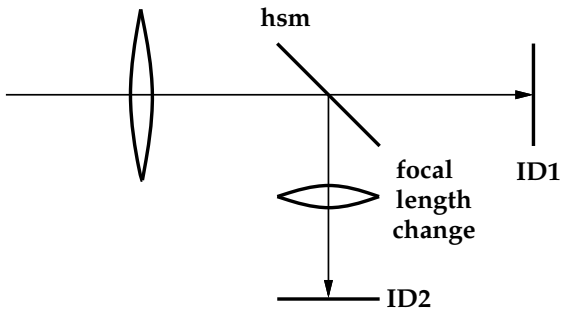
(a) two cameras



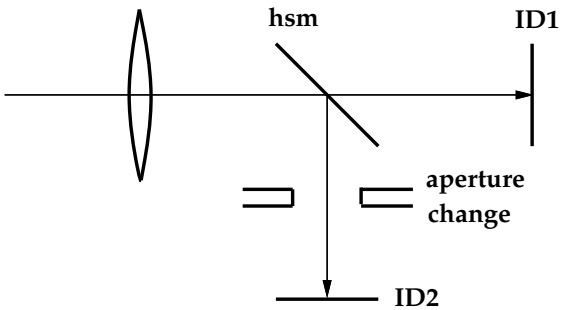
(b) two CCDs



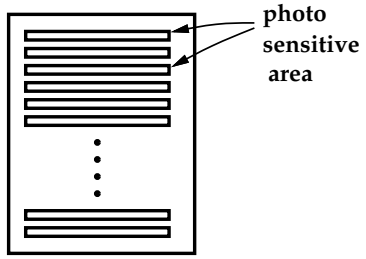
(c) three CCDs



(d) change focal length



(e) change aperture



(f) Linear CCD with elongated sensor elements

Fig. 3 Camera Structures for Focusing on Moving Object

hsm: half silvered mirror

fm: full mirror

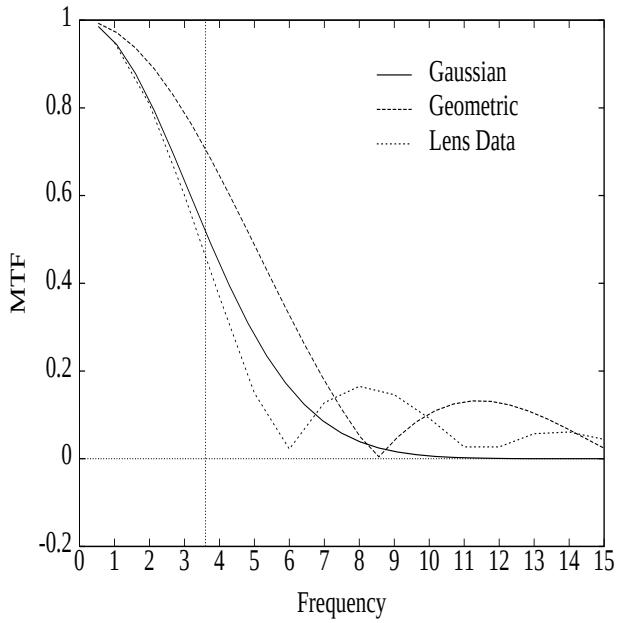


Fig. 4 Gaussian, Geometric and Lens MTF

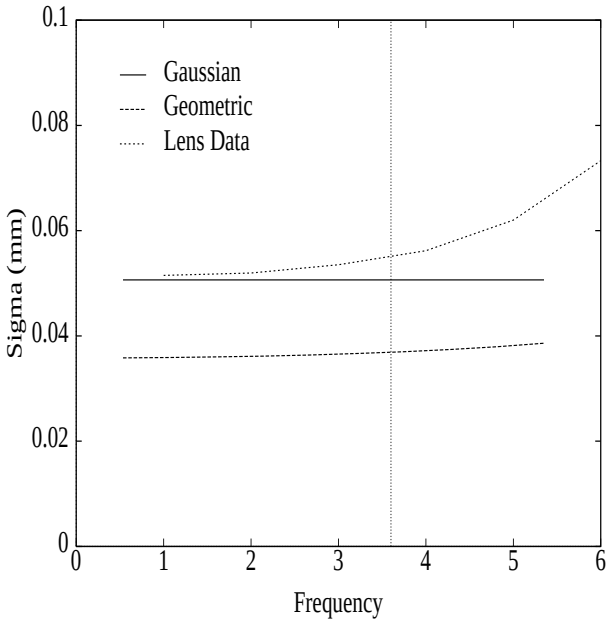


Fig. 5 Estimated σ from MTF

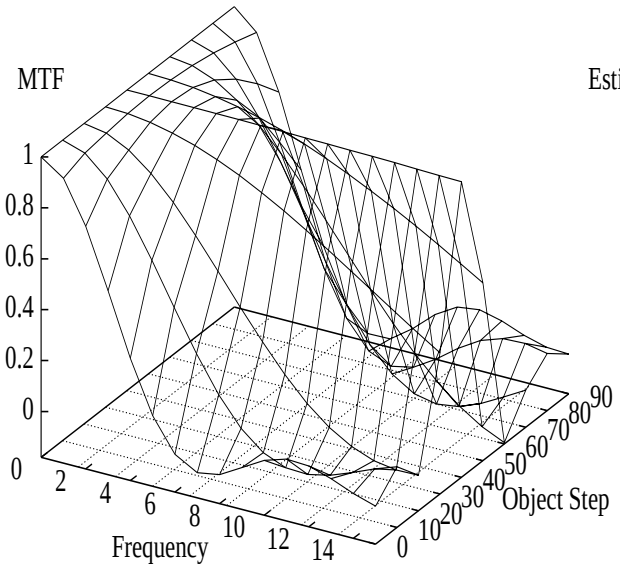


Fig. 6 Lens MTF for Lens Step 40

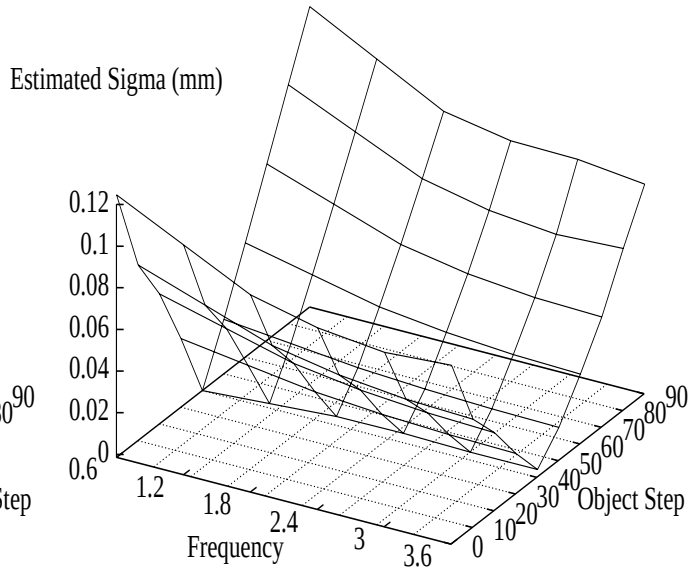


Fig. 7 Estimated σ from Fig. 6

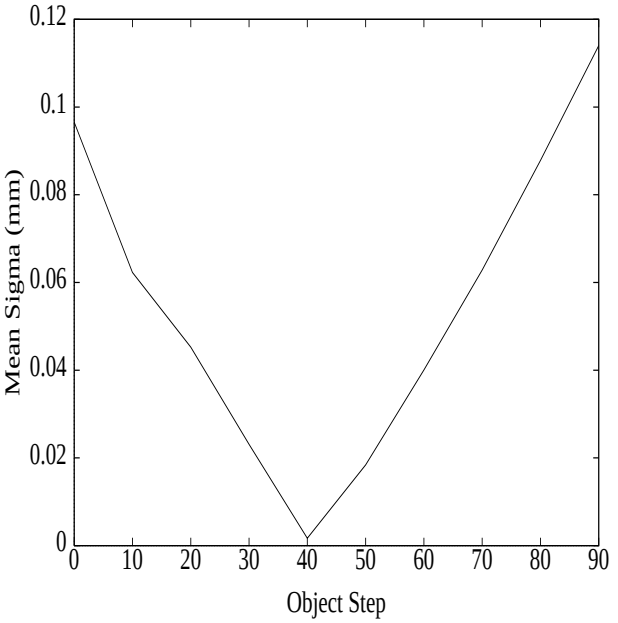


Fig. 8 Mean Sigma from Fig. 7

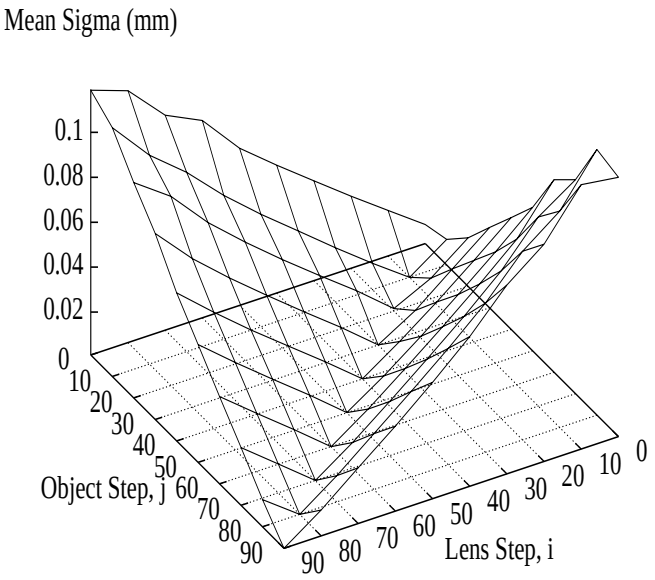


Fig. 9 Estimated σ Lens MTF

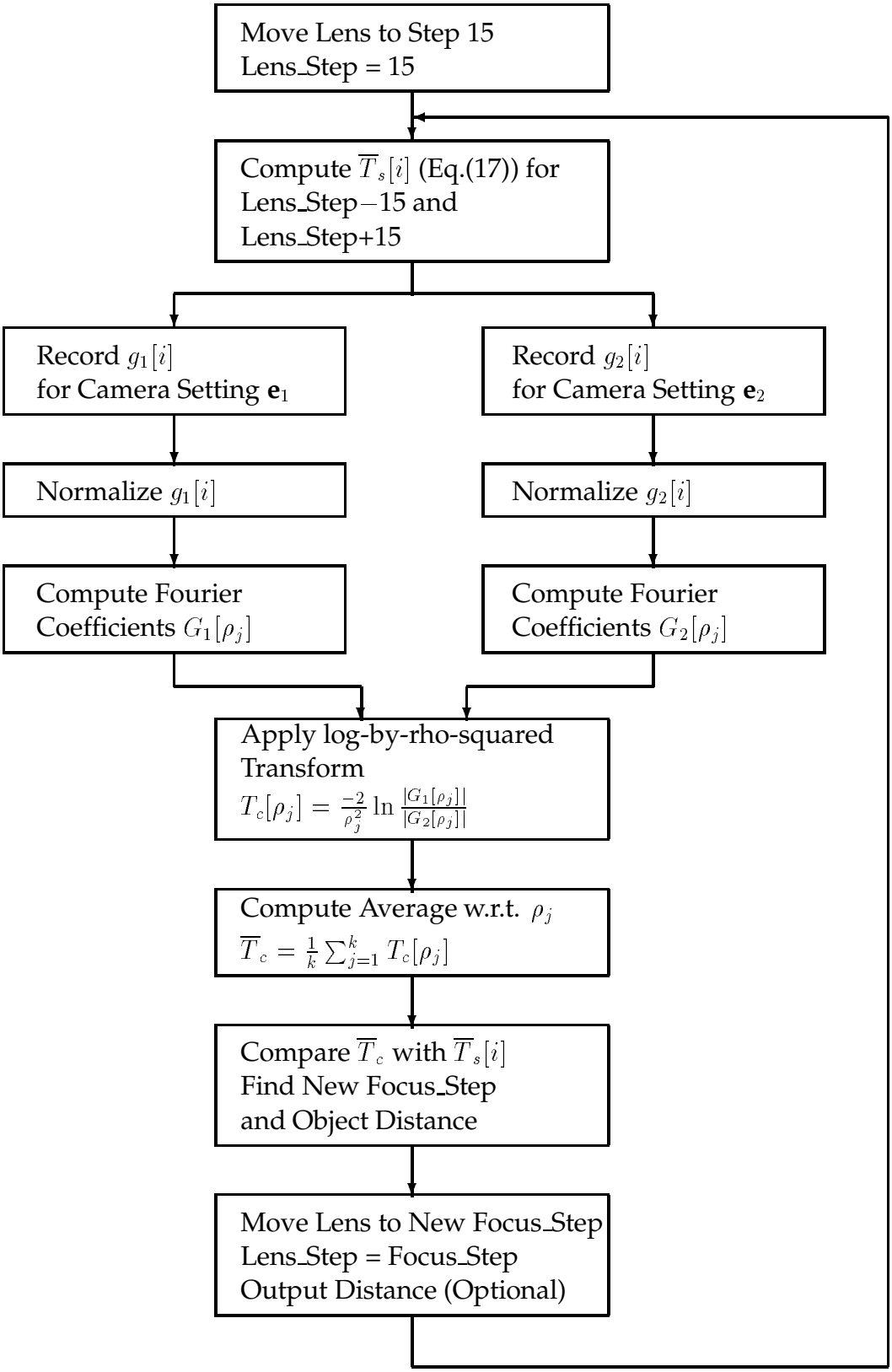


Fig. 10 Flow Chart for Focusing on Moving Object

Fig. 11 Test Image, FACE

Fig. 12 Test Image, NAVY

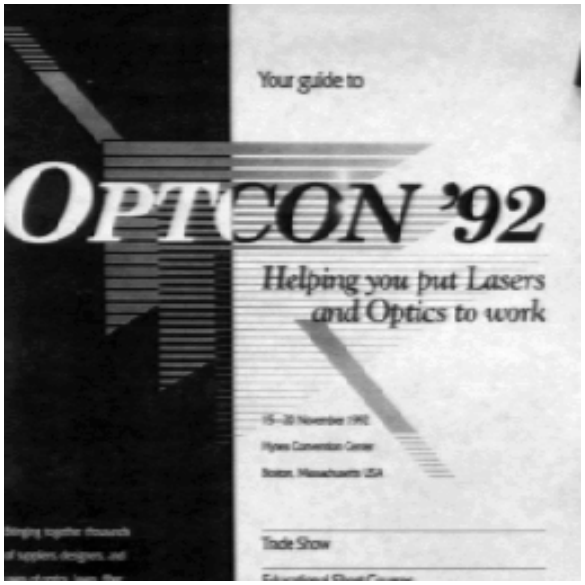


Fig. 13 Test Image, OPTCON

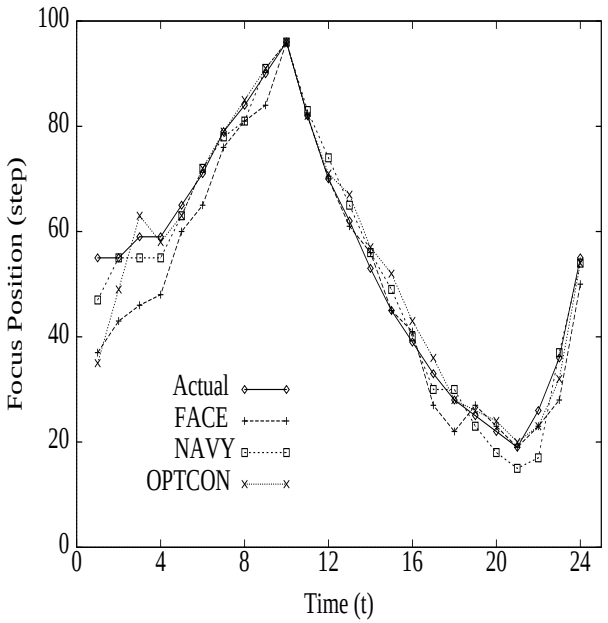


Fig. 14 Experiment Results

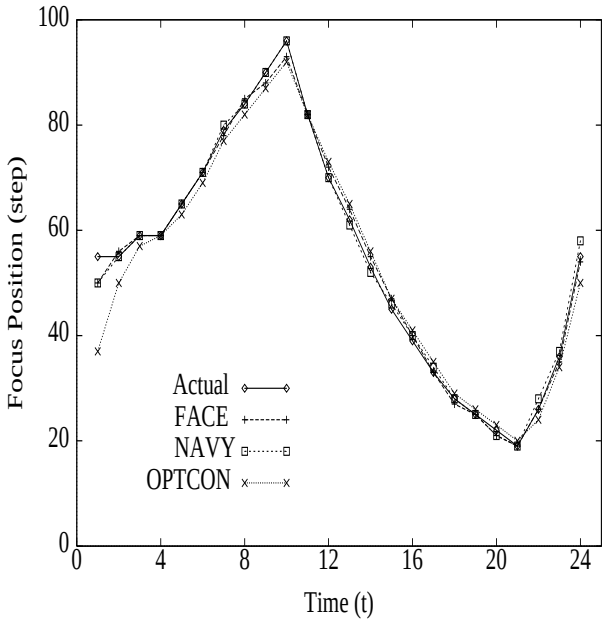


Fig. 15 Simulation Results