Parallel Depth Recovery by Changing Camera Parameters

Murali Subbarao
Department of Electrical Engineering, State University of New York at Stony Brook
Stony Brook, NY 11794-2350, USA.

Abstract

A new method is described for recovering the distance of objects in a scene from images formed by lenses. The recovery is based on measuring the change in the scene's image due to a known change in the three intrinsic camera parameters: (i) distance between the lens and the image detector, (ii) focal length of the lens, and (iii) diameter of the lens aperture. The method is *parallel* involving simple *local computations*. In comparison with stereo vision and structure-from-motion methods, the *correspondence* problem does not arise. This method for depth-map recovery may also be used for (i) obtaining focused images (i.e. images having large depth of field) from two images having finite depth of field, and (ii) rapid autofocusing of computer controlled video cameras.

1. Introduction

Here we describe a new passive ranging method which in principle is fast and involves relatively weak assumptions that are generally valid. The method is basically a generalized version of the 'depth-from-focusing' [4,6,14,7] method. It is general in the sense that it is not required to focus an object in order to find its distance. Instead, two images of the object (which may or may not be focused) acquired with different camera parameter settings are processed to determine the distance. In particular, this method requires only two images and is *parallel* (hence fast) in contrast to the depth-from-focusing method which requires a large number of images and is *sequential* (hence slow).

2. Previous work

Pentland [10,11,12] is perhaps the first to investigate monocular depth recovery in parallel from *images formed by a lens*. Pentland proposes two methods. The first method is based on measuring the "blur" of edges which are step discontinuities in the focused image. Recently Grossman [3] has reported results of some experiments based on this same principle. A related method and experiments are described in [20].

Pentland's second method is based on comparing two images, one image formed with a very small (pin-hole) aperture and the other image formed with a normal aperture. Application of this method poses some practical difficulties. A very small aperture increases the effects of diffraction which distorts the image. Further, a smaller aperture gathers lesser light and therefore increases film exposure period. This could slow down the method.

The method to be presented here originates from the new methods for depth recovery proposed in [16]. The original methods are based on measuring the change in an image due to a very small or *infinitesimal* change in one of the camera parameters. These methods are sensitive to noise as the change in an image is small if the change in a camera parameter is small. In the new method the change in the camera parameters can be very large, thus permitting large changes in the observed image. This makes the method realizable in practice and more robust than the earlier methods. Further, it is more general than the original methods. In the original methods only one out of three camera parameters are changed at a time. In the method described here, any one, any two, or all three camera parameters may be changed simultaneously.

The main ideas behind the new method were developed independently by us although Pentland's earlier work is related. Pentland's method can be derived as a special case of the new method.

3. Camera configuration

The first camera parameter, the distance between the lens system and the image detector, is changed by *moving the image detector* back and forth along the optical axis. (Note: In most commercially available cameras, the lens is moved instead of the image detector or film. However lens motion introduces the same *correspondence problem* encountered in structure-from-motion where the camera is moved along the optical axis. Apparently, the fact that –moving the image detector instead of the lens avoids correspondence problem– has not been noted in previous work on depth-from-focusing.) The second parameter, the focal length of the lens system, is changed by moving back and forth the lens nearer to the image detector (L2 in Figure 1). The other lens viewing the scene directly is not moved. (Note: it is the motion of the front lens L1 or *object piece* that introduces the correspondence problem; hence it should not be moved.) The effective focal length f is determined from the well known formula

$$\frac{1}{f} = \frac{1}{f_a} + \frac{1}{f_b} - \frac{l}{f_a f_b} \qquad (1)$$

where f_a, f_b are the focal lengths of the two lenses and l is the distance between the two lenses. The positions of principal planes and principal points are determined using well known methods in geometric optics [9]. The third camera parameter, the diameter of the camera's aperture, is controlled separately by changing the diameter.

CH2664-1/88/0000/0149$01.00 © 1988 IEEE

4. Theoretical basis of the method

4.1 Spatial and gray-level scaling

The magnification of an observed image is proportional to the distance s of the image detector from the second principal plane (see Figure 1). In order to simplify locating corresponding points (or regions) in images acquired with different s values, we shall scale all images to have the same magnification. In the following discussion we fix this magnification to be that corresponding to $s=1$. Analogous to this correction for spatial rescaling is the gray-level rescaling. Changing the camera parameters could change the irradiance at the image detector. For example, reducing the diameter of the aperture reduces the image irradiance. This change in image brightness is compensated as follows. After spatial rescaling of the images, normalize (or rescale) the gray-levels of the images so that the mean gray value is the same for all images. This is done by multiplying the gray values by a constant factor. (This gray level normalization should be applied after correcting for the vignetting effect.)

4.2 Point spread function of a lens system

Let P be a point on a visible surface in the scene and p be its focused image (see Figure 1). The relation between the positions of P and p is given by the lens formula

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \tag{2}$$

where f is the effective focal length, u is the object distance from the first principal plane, and v is the image distance from the second principal plane. (The lens formula is valid exactly only for an aberration free optical system and for points near the optical axis. Later we will comment on how to use our method in the other cases excluded here.) If P is not in focus then it gives rise to a circular image called the *blur circle* on the image detector. From simple plane geometry (see Figure 1) and the lens formula (2) we can show that the diameter of the blur circle is given by

$$D\ s\left(\frac{1}{f} - \frac{1}{u} - \frac{1}{s}\right) \tag{3}$$

where D is the diameter of the lens. Normalizing the spatial magnification of the blur circle (by dividing its diameter by s) we obtain its new diameter d to be

$$d = D\left(\frac{1}{f} - \frac{1}{u} - \frac{1}{s}\right). \tag{4}$$

Note that d can be either positive or negative depending on whether $s \geq v$ or $s < v$. In the former case the image detector is *behind* the focused image of P and in the latter case it is *in front* of the focused image of P. According to geometric optics, the intensity within the blur circle is approximately constant, but due to aberrations, diffraction, and other effects [2,5,15,12,16,17,18] the following two-dimensional Gaussian has been suggested as an alternative model:

$$h(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2}\frac{x^2+y^2}{\sigma^2}} \tag{5}$$

where σ is the spread parameter such that

$$\sigma = k\ d \quad \text{for} \quad k>0. \tag{6}$$

(k is a constant of proportionality characteristic of a given camera; it is determined initially by an appropriate calibration procedure.) For the purpose of illustration, we proceed with the above model, but later we indicate how the method is applied for any circularly symmetric distribution. Now the blurred image of point P is actually the *point spread function* of the camera system. Therefore, for a linear shift-invariant camera system (cf. [13]), an observed image is the result of convolving the focused image with the camera's point spread function.

The *focused image* $f(x,y)$ of a scene for a given setting of the camera parameters is defined as follows. For any point p (see Figure 1) at position (x,y) on the image detector, consider a line through that point and the second principal point (Q2). Let P be the point on a visible surface in the scene whose focused image lies on this line at position p'. Then $f(x,y)$ is defined as the image irradiance at p' due to the light from P.

From equations (4) and (6) we have

$$\sigma = kD\left(\frac{1}{f} - \frac{1}{u} - \frac{1}{s}\right). \tag{7}$$

The equation above differs from the corresponding equations in [16,17,18] by a factor of s. The reason for this discrepancy is that the image magnification which depends on s was not normalized. Therefore, one of the depth recovery method based on changing s there needs minor modification. This magnification correction appears to have also been overlooked in the implementation of all depth-from-focusing methods [8,14,7] known to the author with the exception of Horn's [4]. We believe that applying this magnification correction will improve the reported experimental results and also alleviate some of the problems associated with local extrema and region correspondence.

$h(x,y)$ in equation (5) is defined in terms of σ and therefore is different for points at different distances from the lens plane. The "volume" of $h(x,y)$ can be shown to be unity. (The volume of the point spread function of a non-light-absorbing lens is unity irrespective of the form of the point spread function.) The Fourier transform of $h(x,y)$ is

$$H(\omega,\nu) = e^{-\frac{1}{2}(\omega^2+\nu^2)\sigma^2} \tag{8}$$

where ω, ν are spatial frequencies in radians per unit distance. In the above equation, the blur parameter σ is different for objects in the scene at different distances from the camera. In this paper we assume that an observed image has been subdivided into smaller subimages within which the distance of objects are nearly constant. The subimages are then processed in parallel indivi-

dually. If this assumption is not valid inside a subimage, then our method gives an "average" distance of objects in that subimage which is still a useful piece of information.

Dividing an image into subimages introduces some errors due to border effects. An image region cannot be analyzed in isolation because, due to blurring (caused by the finite spread of the point-spread-function), the intensity at the border of the region is affected by the intensity immediately outside the region. We call this the *image overlap* problem because the intensity distribution produced by adjacent patches of visible surfaces in the scene overlap on the image detector. In indoor scenes such as the environments of industrial vision systems, the image overlap problem can be completely avoided through selective illumination of the scene. For example, the scene can be illuminated by square bright patches separated by wide dark bands with no illumination. In this case the boundaries of the subimages can be chosen to be in the middle of the dark bands. Border effects are then avoided because the image intensity is zero at and near the borders. In situations where the illumination cannot be controlled (e.g. outdoor scenes), the image overlap problem may be reduced as follows. The image intensity is first multiplied by a suitable center weighted (e.g. a Gaussian) mask centered at the region of interest. The resulting weighted image is then used for depth recovery. Because the weights are higher at the center than at the periphery, this scheme gives a depth estimate which is approximately the depth along the center of the field of view.

4.3 Power spectral density

Let $g(x,y)$ be the observed image of an object on the image detector, and $f(x,y)$ be the corresponding focused image. Also, let $G(\omega,\nu)$ and $F(\omega,\nu)$ be the corresponding Fourier transforms. We have

$$g(x,y) = h(x,y)*f(x,y) \tag{9}$$

where * denotes the convolution operation. Therefore, the power spectral density for a Gaussian point spread function is

$$P(\omega,\nu) = e^{-(\omega^2+\nu^2)\sigma^2} F F^*. \tag{13}$$

4.4 Depth recovery by changing camera parameters by large values

The blur parameter σ for an image region can be changed by changing one or more of the camera parameters: s, f, and D (see equation 7). Let $P_1(\omega,\nu)$ and $P_2(\omega,\nu)$ denote powers for two different camera parameter settings: s_1, f_1, D_1 and s_2, f_2, D_2. Let σ_1 and σ_2 be the corresponding blur parameters. Then, from equation (13) we have

$$\frac{P_1(\omega,\nu)}{P_2(\omega,\nu)} = e^{-(\omega^2+\nu^2)(\sigma_1^2-\sigma_2^2)} \tag{14}$$

Taking logarithm on either side of equation (14) and rearranging terms, we get

$$\sigma_1^2-\sigma_2^2 = \frac{-1}{\omega^2+\nu^2} \ln\left(\frac{P_1(\omega,\nu)}{P_2(\omega,\nu)}\right). \tag{15}$$

For some (ω,ν), the right hand side of equation (15) can be computed from the given image pair. Therefore equation (15) can be used to estimate $\sigma_1^2-\sigma_2^2$ from the observed images. In principle, measuring the power spectral density at a single point (ω,ν) is sufficient to obtain the value of $\sigma_1^2-\sigma_2^2$, but a more robust estimate can be obtained by taking the average over some domain in the frequency space. Let the estimated average be C given by

$$C = \frac{1}{A}\iint_R \frac{-1}{\omega^2+\nu^2} \ln\left(\frac{P_1(\omega,\nu)}{P_2(\omega,\nu)}\right) d\omega d\nu \tag{16}$$

where R is a region in the (ω,ν) space not containing points where $P_1(\omega,\nu) = P_2(\omega,\nu)$, and A is the area of R. Therefore, from the observed images we get the following constraint between σ_1 and σ_2:

$$\sigma_1^2-\sigma_2^2 = C\ . \tag{17}$$

Above we have one equation in two unknowns. We obtain an additional equation from the actual camera parameter settings as follows.

From equation (7) we have

$$\sigma_1 = k_1 D_1 \left(\frac{1}{f_1} - \frac{1}{u} - \frac{1}{s_1}\right), \text{ and} \tag{18a}$$

$$\sigma_2 = k_2 D_2 \left(\frac{1}{f_2} - \frac{1}{u} - \frac{1}{s_2}\right). \tag{18b}$$

Eliminating u from the above two equations we get

$$\sigma_1 = \alpha\sigma_2+\beta \tag{19a}$$

where

$$\alpha = \frac{k_1 D_1}{k_2 D_2}\ ,\ \beta = k_1 D_1\left(\frac{1}{f_1} - \frac{1}{f_2} + \frac{1}{s_2} - \frac{1}{s_1}\right). \tag{19b}$$

Equations (17,19a) together constitute two equations in two unknowns. From these equations we get

$$(\alpha^2-1)\sigma_2^2 + 2\alpha\beta\sigma_2 + \beta^2 = C. \tag{20}$$

Above we have a quadratic equation in σ_2 which is easily solved. In general there will be two solutions. However a unique solution is obtained if $D_1=D_2$. We can also derive other special cases where a unique solution is obtained (e.g.: $D_1\neq D_2$, $s_1=s_2=f_1=f_2$; in this case only the negative solution of σ is acceptable which is unique.) Having solved for σ_2 we obtain the distance u from equation (18b) as

$$u = \frac{-k_2 D_2 s_2 f_2}{s_2 f_2 \sigma_2 + k_2 D_2 (f_2 - s_2)}.$$

4.5 Depth recovery by changing camera parameters by small values

Next we describe a method for depth-map recovery by changing camera parameters by small values. This method is a generalization of the methods proposed in [16]. At present it remains to be investigated experimentally whether this method is better than the previous method when the change in the camera parameters is very

small.

Let two images be acquired with camera settings s, f, D and $s+ds, f+df, D+dD$ where ds, df, dD are small changes in the respective parameters. From (7) we have

$$d\sigma = \frac{\partial\sigma}{\partial s}ds + \frac{\partial\sigma}{\partial f}df + \frac{\partial\sigma}{\partial D}dD. \tag{21a}$$

$$= \frac{kD}{s^2}ds - \frac{kD}{f^2}df + \frac{\sigma}{D}dD. \tag{21b}$$

From (13) we get

$$dP = -2(\omega^2+\nu^2)\, P\, \sigma\, d\sigma. \tag{22}$$

Therefore

$$\sigma\, d\sigma = -\frac{1}{2}\frac{1}{\omega^2+\nu^2}\frac{dP}{P}. \tag{23}$$

Let

$$C = -\frac{1}{2}\frac{1}{\omega^2+\nu^2}\frac{dP}{P}. \tag{24}$$

In principle, measuring P and dP at a single point (ω,ν) is sufficient to compute C, but a more robust estimate of C can be computed from the observed images as

$$C = -\frac{1}{2A}\int\int_R \frac{1}{\omega^2+\nu^2}\frac{dP(\omega,\nu)}{P(\omega,\nu)}\, d\omega\, d\nu \tag{25}$$

where, as before, R is a region in the (ω,ν) space not containing points where $dP(\omega,\nu)=0$, and A is the area of R. From (21b,23,24) we obtain

$$\sigma\left(\frac{kD}{s^2}ds - \frac{kD}{f^2}df + \frac{\sigma}{D}dD\right) = C. \tag{26}$$

σ is obtained by solving the above quadratic equation. In general two solutions are obtained. However a unique solution is obtained if $dD=0$.

4.6 Notes:

Enhancing depth-of-field: A focused image may be obtained in principle from two observed images as follows. As in the depth recovery method, first the spread of the Gaussian point spread function is estimated for one of the observed images. The observed image is then deconvolved with the corresponding point spread function. The resulting image is the required focused image. However, in practice, *deconvolution poses many serious difficulties* (especially in the presence of noise).

Multiple images: Although our method requires only two images, the estimate of depth can be made more robust if more images are used. If n images are available for different camera parameter settings, then $n-1$ independent estimates of depth can be made and the mean of these gives a robust estimate of the actual depth. Alternative schemes are also possible for using multiple images.

Domain of analysis: Our method is based on a Fourier domain analysis of the images. It is possible to do a corresponding analysis in the spatial or other suitable domain. We have chosen the Fourier domain for its simplicity and in particular analyze the power spectrum as it can be computed very fast by optical methods.

5. Practical Considerations

Lens formula: We said earlier that the lens formula in equation (2) is valid exactly only for an aberration-free lens and for points near the optical axis. For an actual camera, one may consider v to be that distance of the image detector from the lens for which the image of a point at distance u is "sharpest" (we will give a precise definition of sharpness shortly). Therefore, for a given camera, one can experimentally determine v as a function of u, f and the direction (or angular position) of a point in the scene. Having determined this function, one can use it in place of the lens formula and derive the corresponding equations. Even if there is no satisfactory piecewise parametric representation of the function, a table-look-up method can be used. Only the computational steps become clumsy.

Arbitrary point spread function: For the purpose of illustration we have taken the point spread function to be a Gaussian. For this case, as we have seen above, neat closed form solution exists for depth recovery. However, for an actual camera, the form of the point spread function could be significantly different from the Gaussian form. (This is the case with the camera used in our experiments.) Simple arguments based on geometric optics suggests a cylindrical or "pill-box" point spread function. For this case too, closed form solution exists (in the form of the roots of an infinite series derived from the ratio of two first order Bessel functions) although the expression is cumbersome. However, due to diffraction effects, lens aberrations, and, quantization and digitization, the effective point spread function of a practical camera becomes complicated. In this case we characterize the point spread function by a single parameter for the purpose of depth recovery. We call it the *spread parameter* of the point spread function, defined as *the square root of the second central moment of the point spread distribution.* (This corresponds to the "standard deviation" of the distribution of the point spread function. It can also be interpreted as the "radius of gyration" about the "center of mass".) Therefore, if $h(x,y)$ is the point spread function and σ is its spread parameter, then we have

$$\sigma^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[(x-\bar{x})^2+(y-\bar{y})^2]\, h(x,y)\, dxdy \tag{27}$$

where $(\bar{x},\bar{y})$ is the location of the "center of mass" of the distribution defined by

$$\bar{x} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, h(x,y)\, dx\, dy\ , \tag{28a}$$

$$\bar{y} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, h(x,y)\, dx\, dy. \tag{28b}$$

(Note that the "total mass" of the distribution is unity,

i.e. $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\, dx\, dy = 1\ . \qquad) \tag{29}$$

A measure of "sharpness" of an image is given by the spread parameter σ of the point spread function. An image is sharpest when the corresponding σ is a global minimum.

Irrespective of the form of the point spread function, given two images of a scene recorded with two different camera parameter settings, we obtain two constraints on the spread parameters of the point spread functions corresponding to the two images. The first one is derived from the two observed images (which corresponds to equation (17) for a Gaussian point spread function), and the second one is derived from the actual values of the camera parameters (corresponding to equation (19a) for an aberration-free optical system for points near the optical axis). Solving these two constraints simultaneously enables the determination of distance of objects. However, in practice, the camera has to be calibrated (only once at the begining) and the actual computational method may be complicated. The details depend on the characteristics (e.g. the form of the point spread function) of the camera system.

Error sensitivity: The effective range of the system depends on many factors such as the values of the camera parameters, illumination condition of the scene, aberrations of the optical system, image quality (i.e. spatial and gray level resolution), etc. The method will fail for polished (mirror-like) objects. A general and complete analysis remains to be done, but one calculation under simplified assumptions leads to the following conclusions: (i) the method is more accurate for nearby objects than for far away objects, (ii) the effective range of the method in practical applications is about one hundred times the focal length of the camera system, and (iii) for far away objects, even if the accuracy of the quantitative estimate of depth may be unsatisfactory, the method can often provide useful qualitative information such as, for example, "object A is nearer than object B", "there are no obstacles within distance X", etc.

Relevance to human vision: In the human visual system, focusing occurs by changing the focal length of the lens. In our method for depth recovery, the two images may be obtained by changing only the focal length. This suggests that humans could, in principle, perceive the depth of all objects in the field of view even if the objects are not in focus. There is evidence in support of the fact that the human eye deliberately exhibits small fluctuations in the focal length of the lens to obtain two images. The following paragraph is quoted from [23] (page 18):

> "... the state of accommodation of the un stimulated eye is not stationary, but exhibits micro fluctuations with an amplitude of approximately 0.1 D (diopter: a unit of lens power given by the reciprocal of focal length expressed in meters) and a temporal frequency of 0.5 cycles/second. He (Cambell, [1]) demonstrated convincingly that these were not a manifestation of instrumental noise, since they occurred synchronously in both eyes. It follows that their origin is central."

Our method implies that such fluctuations could be used to perceive depth in the entire scene simultaneously.

Plain objects: Objects like machine parts, wall, door, road, etc. are often "plain" or "textureless", i.e. their surfaces are smooth and have no reflectance variation. Therefore they appear as objects with constant brightness under uniform illumination. Our method fails for such objects due to the lack of spatial frequency content. However, if one has control over the illumination of the scene (as in indoor scenes), one can introduce "texture" by projecting an arbitrary light pattern (e.g. a random dot pattern) onto the surface of objects. Then our method becomes applicable.

6. Preliminary experiments

At present it appears that camera systems calibrated with respect to s, f, D are not available. Therefore the method proposed here could not be tested in its complete generality. Further, most existing camera systems have small apertures. This appears to be a deliberate design decision to maximize the depth-of-field. For such cameras objects at all distances are nearly focused and consequently have low depth discrimination. Therefore a camera system designed specifically for depth recovery (having small depth-of-field) should perform significantly better than the camera used by us.

The Panasonic WV-CD 50, a CCD camera, focal length 16 mm, and aperture diameter 11.4 mm, was used in our experiments. The goal of our experiments was two fold: (i) to verify the applicability of our mathematical model to practical camera systems, and (ii) to test the usefulness of the method in practical applications. Two sets of experiments were conducted.

The first set of experiments are described in detail in [20]. The pictures of a step edge were used to compute the spread σ as a function of object distance. Two of the edge pictures and a typical plot of σ against inverse distance are shown in figures 2 and 3. The plot verifies that the spread parameter σ defined by equation (27) for the point spread function of our camera (which was neither Gaussian nor pill-box or cylindrical form) is a meaningful and consistent measure. They show that, for a given setting of camera parameters, *the spread parameter is linearly related to inverse distance* as implied by equation (7). Indeed we find that (see Figure 3) this linear relation predicted by our mathematical model holds remarkably well for our camera system.

The second set of experiments is related to depth recovery from two images. Due to the lack of a suitable camera system whose parameters can be changed and measured, only the following limited experiment could be conducted. A painted cloth hung normal to the optical axis at about 1 meter from the camera was used as the object. A sequence of pictures A1-A7 of the object (some shown in Figure 4) were taken by continuously decreasing *the distance between the lens and the image detector.* We could see that the blur in the pictures gradually decreased, reached a minimum for A4 (which was visually judged to be focused) and then increased again. The size of each

picture was 480×480. Four subpictures of size 32×32 were taken from near the center of each of these pictures and their power spectrums were computed. Then the quantities $\sigma_{A1}^2 - \sigma_{A4}^2$ for $i=1,2,...,7$ were computed using equation (16) (see equation 17). These quantities have been plotted against the picture sequence in Figure 5. In the graph the symbols: x,*,+, and o correspond to the four subpictures. The horizontal axis crudely corresponds to s, the distance between the lens and the image detector. The actual distances could not be measured in the camera system available to us. For ideal data and an exact mathematical model, the positions of x,*,+ and o would all be the same for each picture. We see that this is approximately true of Figure 5. The computed blur is (i) nearly the same for different parts of the pictures, and (ii) a strictly monotonic function in two intervals. These observations imply that the computed measure depends only on the camera parameters and the *distance* of objects in the scene, but not on the *appearance* of objects. We attribute most of the deviation in the plot to the magnification correction that could not be applied due to the unavailability of the correction factor. Further, the image overlap problem and noise (quantization, digitization, etc) contribute to the deviation from the predicted behavior.

Experiments were conducted for two more sets of pictures. The results were similar to the one above. We conclude from these experiments that (i) our mathematical formulation is a reasonable model of actual camera systems, and (ii) approximate depth information can be obtained in practical applications. Clearly more experiments are needed. Design and fabrication of a custom designed camera system is essential for future experiments.

7. Related Work

In this paper we have concentrated on the *method* for depth recovery. A formal development of the underlying *theory* from first principles is presented in [21]. Detailed computational steps, and camera design are described in [19]. A scheme for incorporating the depth recovery method into a binocular system to help establish correspondence and a method for motion recovery are described in [18].

8. References

[1] Cambell, F. W. 1960. *Jour. of Optical Soc. of America,* 50, 738.

[2] Goodman, J.W. 1968. *Introduction to Fourier Optics.* McGraw-Hill, Inc.

[3] Grossman, P. 1987 (Jan.). Depth from focus. *Pattern Recognition Letters 5.* pp. 63-69.

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

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

[6] Jarvis, R. A. 1983. *IEEE Trans. on Patt. Anal. and Mach. Intel.,* PAMI-5, No. 2, pp. 122-139.

[7] Krotkov, E. 1987. Ph.D. thesis. MS-CIS-87-29. Grasp lab 101. Dept. of Comp. Scn., Univ. of Pennsylvania.

[8] Ligthart, G. and F. C. A. Groen. 1982. *Proc. of the Inter. Conf. on Pattern Recognition.*

[9] Ogle, K.N. 1968. *Optics.* Chapter VII, Section 2. Charles C Thomas . Publisher. Springfield, Illinois.

[10] Pentland, A.P. 1982. *Proc. of DARPA Image Understanding Workshop.* Palo Alto, pp. 253-259.

[11] Pentland, A. P. 1985. *Proc. of Inter. Joint Conf. on Artif. Intel.,* pp. 988-994.

[12] Pentland, A.P. 1987. A new sense for depth of field. *IEEE Trans. on Patt. Anal. and Mach. Intel.,* Vol. PAMI-9, No. 4, pp. 523-531.

[13] Rosenfeld, A., and A.C. Kak. 1982. *Digital Picture Processing,* Vol. I. Academic Press, Inc.

[14] Schlag, J.F., A.C. Sanderson, C.P. Neuman, and F.C. Wimberly. 1983. CMU-RI-TR-83-14. Robotics Institute. Carnegie-Mellon University.

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

[16] Subbarao, M. 1987a. *Proc. of IEEE computer soc. workshop on computer vision, Miami Beach, Dec. 1987, pp. 58-65.* (Also in Tech. Rep. 87-02, Image Analysis and Graphics Lab, SUNY at Stony Brook.)

[17] Subbarao, M. 1987b. Tech. rep. 87-03. Image Analysis and Graphics Lab, SUNY at Stony Brook.

[18] Subbarao, M. 1987c. Tech. rep. 87-04. Image Analysis and Graphics Lab, SUNY at Stony Brook.

[19] Subbarao, M. 1987d, Nov. U.S. patent application no. 126407 (pending).

[20] Subbarao, M. and G. Natarajan. 1988. *Proc. of the IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition,* Ann Arbor, Michigan, pp. 498-503.

[21] Subbarao, M. 1988 (to appear). Efficient depth recovery through inverse optics. *Machine Vision for Inspection and Measurement,* Editor: H. Freeman, Academic Press.

[22] Tenenbaum, J. M. 1970. Accommodation in Computer Vision, Ph.D. Thesis, Stanford Univ.

[23] Weale, R.A. 1982. *Focus on Vision.* Harvard University Press, Cambridge, Massachusetts.

Acknowledgement: I am thankful to Mr. G. Natarajan for doing some of the experiments and to anonymous referees and Prof. A. Pentland for helpful comments at various stages of this work. Dr. Bill Sakoda helped in acquiring data for the experiments. The EE department of SUNY at Stony Brook is expected to pay the page charges and my expenses to attend the conference.

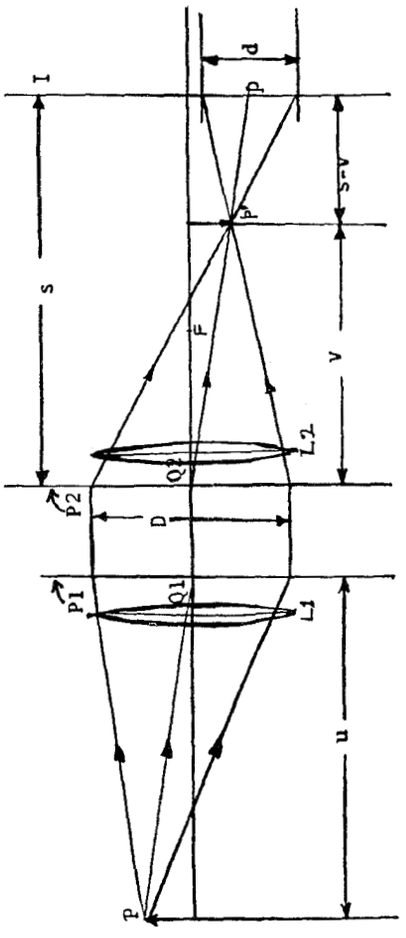

Figure 1. The camera geometry and camera parameters.
P1: first principal plane P2: second principal plane
P : object point p: image point I: image detector plane
s,f,D: camera parameters u: object distance v: image distance
d: blur circle diameter Q1,Q2: first & second principal points

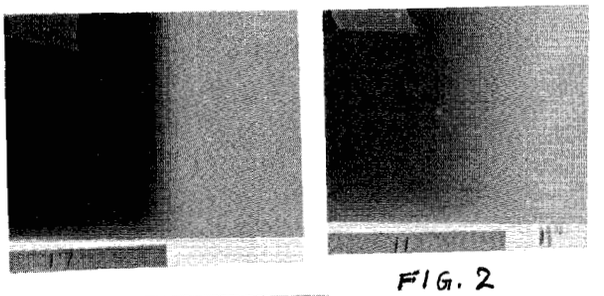

FIG. 2

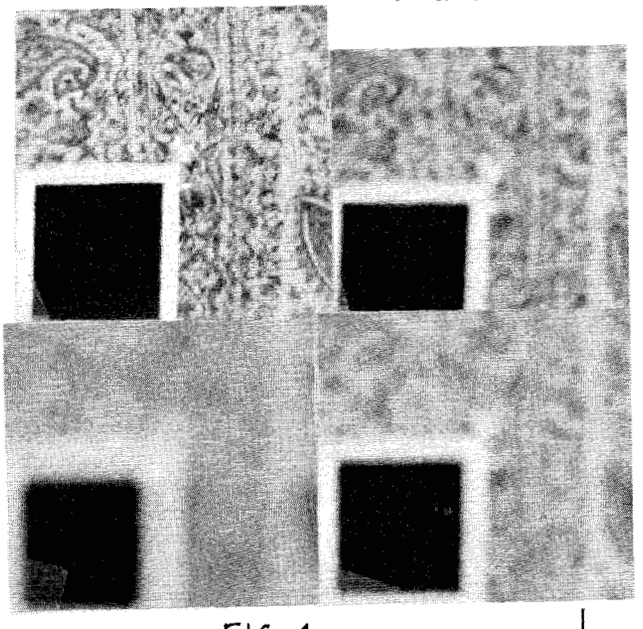

FIG. 4

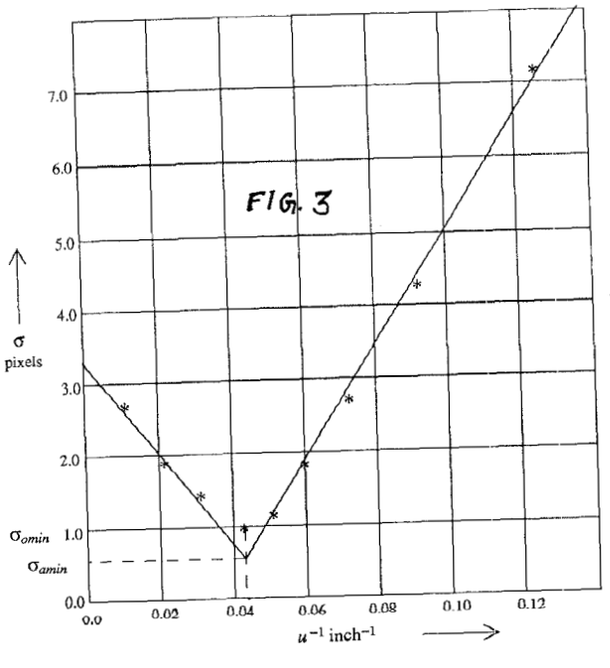

FIG. 3

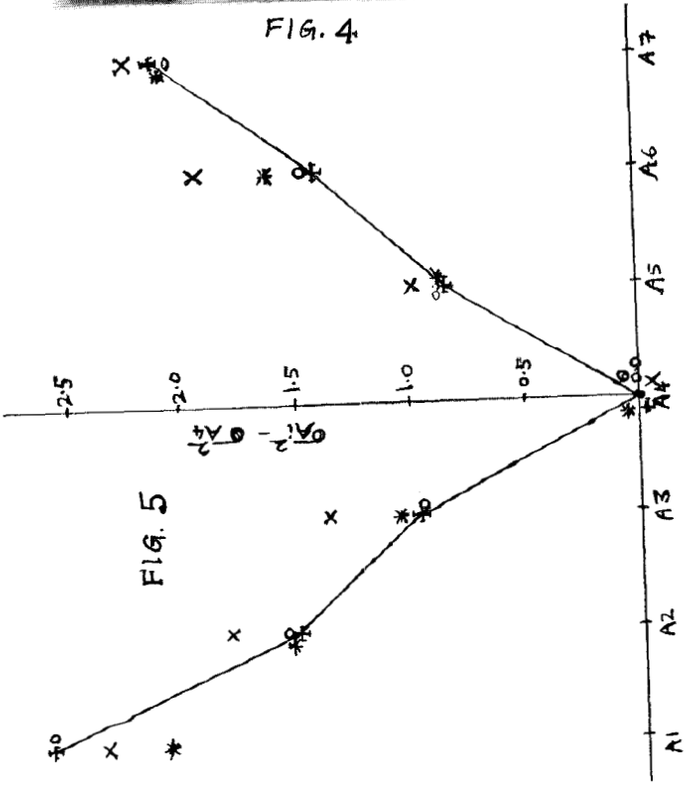

FIG. 5