Integration of Defocus and Focus Analysis with Stereo for 3D Shape Recovery

Murali Subbarao Ta Yuan Jenn-Kwei Tyan
Department of Electrical Engineering,
State University of New York
Stony Brook, NY 11794-2350
email: murali@sbee.sunysb.edu , tyuan@sbee.sunysb.edu , jktyan@sbee.sunysb.edu

ABSTRACT

Image Defocus Analysis (IDA), Image Focus Analysis (IFA) and Stereo Image Analysis (SIA), are integrated for recovering the three-dimensional (3D) shape of objects. Integrating IDA, IFA, and SIA, has important advantages because IDA and IFA are less accurate than stereo but they do not suffer from the *correspondence problem* associated with stereo. Therefore, a rough 3D shape is first obtained using IDA and IFA without encountering the correspondence problem. The amount of computation and accuracy at this stage is optimized by using IDA first and then the IFA. Accuracy is further improved by projecting a high contrast pattern on to the object of interest. The rough shape thus obtained is used in a stereo matching algorithm to solve the correspondence problem efficiently. The amount of computation in matching is reduced since the search for correspondence is done in a narrow image region determined by the approximate shape. Also, the knowledge of approximate shape improves the matching accuracy by minimizing false matches due to occlusion. The method for integrating IDA, IFA, and SIA is presented in detail. The method has been implemented on a vision system named SVIS. Experimental results of the method are presented.

Keywords: Image Focus Analysis, Image Defocus Analysis, Stereo Image Analysis, Ranging, Depth-Map.

1 Introduction

In machine vision, three of the important techniques for ranging and three-dimensional (3D) shape recovery are– *image defocus analysis* (IDA)[7,9,3,8] *image focus analysis* (IFA) (e.g.[2,4–6]) and *stereo image analysis* (SIA).[1]

In IFA, a large sequence of image frames of a 3D scene is recorded with different camera parameters (e.g. focal length or/and lens to image detector distance). In each image frame, different objects in the scene will be blurred by different degrees depending on their distance from the camera lens. Each object will be in best focus in only one image frame in the image sequence. The entire image sequence is processed to find the best focused image of each object in the 3D scene. The distance of each object in the scene is then found from the camera parameters that correspond to the image frame that contains the best focused image of the object.

In IDA only a few (2 or 3) image frames of a scene are recorded with different camera

parameters. The degree of defocus of each object in these few image frames along with the corresponding camera parameters are analyzed to find the focused image and distance of every object in the scene. In comparison with IFA, IDA requires (i) less number of images, (ii) less computation, but (iii) more information about the defocusing charecteristics of the camera. In addition, IDA is less accurate than IFA. In this paper, we integrate the two techniques to obtain accuracy equivalent to IFA but with less number of images and computation than IFA. This is accomplished by first using IDA to obtain a rough estimate of depth-map, and improving the accuracy of the estimate using IFA in a narrow range around the estimated depth-map.

In SIA two or more images are recorded from two or more spatial locations by displacing the camera. Then distance of objects in the scene are found through "triangulation". Depending on camera parameter values, in some vision systems (e.g. the human vision system), IDA and IFA methods are less accurate than stereo vision in providing the depth-map of a scene. However, unlike stereo vision, IDA and IFA do not suffer from the *correspondence* and *occlusion* problems. In this paper, we describe a technique for integrating IFA and IDA with stereo vision. This technique has the potential to result in a fast, reliable, and accurate method for ranging and 3D shape recovery. The rough depth-map provided by IDA and IFA is used to smplify the stereo correspondence and occlusion detection problems. The rough depth-map essentially reduces the range of stereo disparity which is searched for stereo matching. In addition, false matches due to occlusion are also reduced. Therefore, stereopsis yields a more accurate 3D shape of objects. If the object whose 3D shape is to be measured does not have sufficient contrast information, then contrast is introduced by projecting a light pattern onto the object. This facilitates the application of IFA, IDA, and SIA.

The technique described in this paper is implemented on a camera system named Stonybrook VIsion System (SVIS). Methods for calibrating the camera system and results of experiments on SVIS are presented.

2 Image Focus and Defocus Analysis

The *depth-map* and the *focused image* of a scene are respectively the *geometric* and *photometric* information that are of interest in machine vision. IFA and IDA are useful in recovering both these types of information. IFA methods (see Fig. 1) are based on the fact that for an aberration-free convex lens, (i) the radiance at a point in the scene is proportional to the irradiance at its *focused image*[1] (photometric constraint), and (ii) the position of the point in the scene and the position of its focused image are related by the *lens formula* (geometric constraint)

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

where f is the focal length, u is the distance of the object from the lens plane, and v is the distance of the focused image from the lens plane (see Fig. 1). Given the irradiance and the position of the focused image of a point, its radiance and position in the scene are uniquely determined. In a sense, the positions of a point-object and its image are *interchangeable,* i.e. the image of the image is the object itself. Now, if we think of an object surface in front of the lens to be

comprised of a set of points, then the focused images of these points define another surface behind the lens (see Fig. 1). This surface is defined to be the *Focused Image Surface* (FIS) and the image irradiance on this surface to be the *focused image.* There is a *one to one correspondence* between FIS and the object surface. The geometry (i.e. the 3D shape information) and the radiance distribution (i.e. the photometric information) of the object surface are uniquely determined by the FIS and the focused image.

In traditional IFA methods (e.g.[2,4–6]) a sequence of images are obtained by continuously varying the distance s between the lens and the image detector or/and the focal length f (see Fig. 1). For each image in the sequence, a focus measure is computed at each pixel (i.e. each direction of view) in a small (about 15×15) image neighborhood around the pixel. At each pixel, that image frame among the image sequence which has the maximum focus measure is found by a search procedure. The grey level (which is proportional to image irradiance) of the pixel in the image frame thus found gives the grey level of the focused image for that pixel. The values of s and f for this image frame are used to compute the distance of the object point corresponding to the pixel. An example of a focus measure is the grey level variance. IFA methods involve a search for the values of s or/and f that results in a maximum focus measure and these methods require the acquisition and processing of a large number of images.

IDA[7,9,3,8] methods do not require focusing the object. They take the level of defocus of the object into account in determining distance and focused image. IDA methods do not involve searching for f and s values which correspond to focusing the object. Therefore these methods require processing only a few images (about 2-3) as compared to a large number of images in the IFA methods. In addition, only a few images are sufficient to determine the distance of all objects in a scene using the IDA methods, irrespective of whether the objects are focused or not. The two main disadvantages of the IDA methods are (i) they require accurate camera calibration for the camera charactristics (a blur parameter as a function of camera parameters), and (ii) they are less accurate than IFA methods.

An IDA method can be combined with an IFA method to reduce the number of images acquired and processed but attain the same accuracy as IFA. First IDA method is used to obtain a rough depth-map. This requires acquiring and processing only 2 or 3 images. Then IFA is applied to a short sequence of images which are acquired with camera parameters so that only objects in the rough depth-map range estimated by IDA are focused. Acquiring and processing of image frames that correspond to focusing objects that are far away from the depth-map estimated by IDA is avoided. This saves image acquisition and processing time of unnecessary image frames. The accuracy of the depth-map obtained will be the same as that of IFA. In comparison with a bare IFA, the combined IDA-IFA may save much time in the best case when all objects in a scene are at the same distance, but in the worst case when objects in scene are at all possible distances, the combined IDA-IFA will take a bit more time than the bare IFA. In a typical application, the combined method can be expected to save some modest time.

In SIA,[1] depth-map is recovered by comparing two images (a right image and a left image) acquired from different camera positions (a right position and a left position). The line connecting the two camera positions is called the *baseline.* In this paper we consider only the simplest case

where the baseline is perpendicular to the optical axis of the camera. In this case, the image displacement d (also called *disparity*) of a point in the left image with respect to its position in the right image is $d = bs/z$ where b is the baseline distance, s is the ("focal length") distance between the optical center and the image detector, and z is the distance of the point from the camera along the optical axis. The image displacement of points in the right image is obtained by *matching* it to the *corresponding* point in the left image. The matching process is computationally intensive since all permissible values of d need to be searched. In addition, sometimes, due to occlusion, no matching point will be present in the left image. However, if we know the approximate distance z' of a point at actual distance z, then the search for the matching point can be limited to a small range around the expected displacement of $d' = bs/z'$. An estimate of d' can be obtained by the combined IDA-IFA method for depth-map recovery. Thus, integrating IDA, IFA, and SIA, the search interval for the *matching* points can be reduced dramatically and computational time can be saved. Also, the depth-map from IDA-IFA can be used to detect *occlusion* and avoid false matches. The accuracy of the depth-map obtained by the integrated method will be the same as that of SIA.

3 Camera System

The integration of IDA, IFA and SIA was implemented on a camera system named Stonybrook VIsion System or SVIS (see Fig. 7). SVIS is a vision system built over the last 1 year in the Computer Vision Laboratory, State University of New York at Stony Brook. SVIS consists of a digital still camera (DELTIS VC 1000 of Olympus Co.). S-Video signal from the camera is digitized by a frame grabber board (Matrox Meteor Standard board). All processing is done on a PC (Intel Pentium, 200 MHz.). The camera is mounted on a linear motion stage driven by a stepper motor (X-9 stage and MD-2 stepper motor system of Arrick Robotics Inc.). The camera is mounted such that its optical axis is perpendicular to linear motion of the stage. The right and left images for stereo disparity analysis are obtained by moving the camera to different positions and recording images. The stepper motor that moves the camera is controlled through a parallel printer port on the PC. Focusing of the camera is done by a motorized lens system inside the camera. The lens motor is controlled from a serial (RS-232) communication port on the PC. The Matrox frame grabber installed in the PC is used to record 480×640 size monochrome images with 8 bits/pixel. A user friendly windows software interface has been developed to control the whole system under MS Windows 95 OS. It includes convenient controls for manipulating (i) the lens system, (ii) digitizer board, (iii) linear motion stage on which the camera is mounted, and (iv) all the application programs. In addition, an overhead projector is used to project a high contrast pattern onto 3D objects that have low contrast.

The camera lens system has separate controls for zooming and focusing. Zooming can be varied from a focal length of 10.2 mm (WIDE mode) to 19.6 mm. (TELE mode). The experimental results reported in this paper were carried out with a zoom focal length of 19.6 mm (TELE mode). Focusing is done by driving a stepper motor that controls lens position with respect to the image sensing CCD in the camera. The stepper motor has step positions ranging from 70 to 170. Objects at infinity are focused when the lens stepper motor is at step number 70, and objects close by (about 25 cm.) are focused when the lens position is at step 170. Since

each lens step number corresponds to focusing objects at some unique distance, we often use this corresponding step number to specify the distance of objects. If an object is said to be at a distance of step X, it means that the distance of the object is such that the object would be in best focus if the lens is moved to step number X. Specifying distance of objects in terms of lens step numbers is particularly convenient in IDA and IFA.

4 Camera Calibration

The internal parameters of SVIS such as focal length, lens to CCD distance, aperture diameter, baseline distance, etc. were not known accurately. Therefore SVIS had to be calibrated with respect to four important factors. The first was change in image magnification as the lens moved from step number 70 to 170 for autofocusing. This is important because correspondence between different image regions is needed between image frames recorded with different lens positions. This facilitates comparison of focus measures computed in different image frames to find the image frame in which a given image region is in best focus. The second factor needing calibration was the relation between the distance of objects from the camera and the corresponding lens step number at which the objects would be in best focus. The third factor needing calibration was a blur parameter needed for image defocus analysis. Finally, calibration was needed to establish a relation between object distance and the corresponding stereo disparity. Calibrations with respect to each of these four factors were carried out and four tables were created to represent the calibration data. The calibration procedures are described below.

4.1 Magnification Factor

The magnification at lens position of step 70 was taken to be 1 unit and the magnification at other lens positions were estimated as follows. A planar pattern (see Fig. 2) was placed normal to the optical axis at about 600 mm from the camera. The pattern consisted of two rectangles of sizes– 200mm width X 150 mm height, and 120 mm width X 90 mm height– with a common center. The image pixel coordinates of the corner points of the rectangles, the center, and the width and height (in pixels) of the rectangles were recorded as a function of lens position in step number. This was made possible by the interactive Windows interface. The image pixel coordinates of the corner points were recorded by visually pointing the mouse pointer to the percieved image point. When the images were blurred, the corner points spread out into large circular patches. In these cases, the location of the image points were found by visually estimating the center of the circular patches on the computer monitor and pointing the mouse pointer there. The pixel coordinates noted by us had an error of about ± 1 pixel. The length of the diagonals of the rectangles were computed from the pixel coordinates of the corner points. The length of the diagonal was a maximum at step 70 and decreased monotonically. The magnification factor at a given step number was obtained by dividing the length of a diagonal at step 70 and the length of the same diagonal at the given step number. These ratios obtained for different diagonals were averaged to obtain the final estimate of the magnification factor. This mean ratio was recorded at intervals of roughly 10 steps each. Between these intervals, the magnification factor was estimated by linear interpolation. A plot of the magnification factor as a function of

step number is shown in Fig. 3. Any image recorded at a given step s was magnified by the magnification factor corresponding to that step. Thus, images recorded at different lens step positions had the same magnification as that for step number 70.

4.2 Lens step vs focusing distance

The lens position in step number and the corresponding focused distance was obtained using an autofocusing algorithm as follows. A large planar high contrast object was placed normal to the camera's optical axis at a known distance from the camera. The camera was then autofocused by maximizing a focus measure. The lens position in step number that resulted in a maximum focus measure was found by a binary search type of algorithm. The image used was the central 128X128. The focus measure was the sum of square of Laplacian of image grey-level. Then the focused step position and the distance of the object from the camera was recorded. This procedure was repeated for many different distances of the object corresponding to roughly 5 step intervals for the focused lens position. Then the gaps were filled by linear interpolation with respect to step number and reciprocal of object distance. This calibration data was used in finding the distance of object points given the focused lens step number for the object points (see Fig. 4).

4.3 Blur parameter vs focused lens step

The IDA used in our implementation is based on a spatial domain approach proposed in.[7] In IDA, only one camera parameter, the lens position (step number) was varied in acquiring the two needed images. All other parameters (focal length and aperture diameter) were nearly constant. In this case we find that a blur parameter σ_2 (which is proportional to the diameter of blur circle) is related to a quantity G' that can be computed from the two recorded images by:

$$\sigma_2 = \frac{G' - \beta^2}{2\beta} \tag{2}$$

The camera constant β in the above equation is a function of the two camera parameter settings at which the two images are recorded. It can be computed if the camera parameters are known. Since they were not known, it was determined experimentally as follows.

The IDA[7] was implemented with two images recorded at lens positions of step 120 and 155. An object was placed at such a distance that it was focused when the lens was at step 120. In this case the blur parameter for the image recorded at step 120 is zero, but the blur parameter σ_2 for the image recorded at step 155 is $-\beta$. Therefore β is obtained directly by computing the square root of $-G'$ (a quantity that can be computed from the two observed images) and the sign of β is negative (Note that in this case the sign of G' will be negative). This method yielded $\beta = -2.0$ for focal length at WIDE mode end and $\beta = -4.0$ at TELE mode end. β can also be computed by placing an object at a distance corresponding to the focused lens position 155. In this case, the blur parameter for the second image is zero, but that for the first image is equal to β. Therefore β can be estimated as the the negative value of the square root of G'.

Another calibration table is needed that relates the blur parameter of an object in the second image (recorded at step 155) to the focused lens step number of the object. This table is created as follows. First an object is placed at a given distance and the IFA method is used to autofocus the object by a binary search for the maxima of a focus measure. The lens step number that autofocuses the object is recorded. Then the IDA method is applied and the blur parameter σ_2 is calculated. This procedure was repeated for several different objects at the same distance and the average σ_2 and the average focus step number were recorded. This gives one entry of the calibration table. This procedure was repeated for several different object distances at roughly regular intervals in terms of focused lens step number (about 5). The gaps in the table were filled by linear interpolation with respect to lens step number and blur parameter. The resulting data is shown in Fig. 5.

4.4 Disparity vs Distance

First the optical axis of the camera was set perpendicular to the stereo baseline by trial and error as follows. The length of the baseline was known to be 50.0 ± 0.2mm. A line segment of the same length as the baseline (50.0mm) was placed at about 500 mm from the camera such that the line segment was parallel to image rows (i.e. the end points had the same row index coordinates but different column index coordinates). This ensured that the line segment was roughly parallel to the camera baseline. If the coordinates of the right end of the line segment in the right image is the same as the coordinates of the left end of the line segment in the left image, then the optical axis of the camera is perpendicular to the baseline. This fact was used to adjust the optical axis to be perpendicular to the baseline. We believe that an accuracy of ± 2 degree was achieved.

The relation between disparity and distance was obtained as follows. A planar object was placed normal to the optical axis at different distances from the camera. At each distance, the object was autofocused using the IFA method. Then the disparity of a set of points on the planar object was recorded. The points were the corners in a checker pattern. The coordinates of corresponding points in the left and right images were found visually by pointing the mouse to their images on the computer monitor. The average disparity for the points was computed. This disparity was then normalized corresponding to the magnification of the camera for the autofocused lens step position. This is necessary since stereo matching is done on focused images that are reconstructed from magnification normalized images. The resulting disparity was recorded for many different object distances (in the range 250mm to 1000 mm) at roughly regular intervals of disparity. The gaps in these intervals were filled by linearly interpolating the disparity with respect to reciprocal of distance. The resulting calibration data is depicted in Fig. 6.

The calibration data relating the lens step number for autofocusing with object distance, and the data relating disparity with object distance were combined to obtain a table relating lens step number for autofocusing with disparity (see Fig. 8). The two calibration data were combined using the following method. For each lens step position, the corresponding autofocused object distance was found. Then the reciprocal of this distance was compared to the reciprocal of distance in the disparity vs distance data for each possible disparity. The disparity corresponding

to the minimum absolute difference in reciprocal distances was taken for creating the lens step vs disparity table. This table is useful in combining IFA and SIA. If the focused step number of a point is determined using IFA, then rough disparity of the point can be obtained by looking up this table. This rough disparity is then improved through a stereo matching method.

5 Integration

5.1 Combining IDA and IFA

First a rough depth-map is obtained using IDA.[7] One estimate of depth is obtained in each image region of size 48×48. The two needed images were recorded at lens steps 120 and 155. Four image frames were time-averaged to reduce noise. The magnification of the images were normalized using the magnification table. The IDA[7] was applied to images of size 432×432. Thus a coarse depth-map array of size 9×9 was obtained. At this stage, the depths were expressed in terms of the lens step number that focuses objects at that depth. In each image region, the actual depth-map at higher resolutions was assumed to be within ± 10 steps of the estimated depth-map step number. Using this initial depth-map, a higher resolution depth-map of size 27×27 (one estimate in 16×16 image region) was obtained using IFA. The lens step numbers for which image frames needed to be recorded and processed in IFA were determined using the following algorithm. The purpose of this algorithm is to acquire and process only those image frames near the estimated depth-map values. This avoids processing unnecessary image frames in which all image regions are highly blurred.

The estimated depth values (in step numbers) is first quantized to multiples of DEL (=3) steps. Then, for each quantized value s that occurs in the depth-map, $2numdel+1$ ($numdel = 3$) lens step positions at $s + i * DEL$ for $i = 0, \pm 1, \pm 2, \cdots, \pm numdel$ are marked. If any of these steps are outside the range of minimum and maximum step positions, they are discarded. Then image frames are recorded at each marked lens position. All images are normalized with respect to the magnification corresponding to the step position at which they are recorded. Then, in the resulting image sequence, focus measures are computed in image regions of size 16×16. The step number where the focus measure is a maximum in each image region is determined. These image regions with maximum focus measures are synthesized to obtain a focused image of the entire scene. Further, the maximum focus measure and the two focus measures in the preceding image frame (DEL steps below) and the succeeding image frame (DEL steps above) are taken. A local quadratic curve is fitted to the three focus measures (the center one being the maximum) and the position of the maximum of the curve is computed. This position is taken as an improved estimate of the depth-map. If the focus measure for the preceding or succeeding image frame is not available, then this last step is not performed.

5.2 Integration of IDA and IFA with Stereo

The depth-map obtained above is improved using SIA as follows. The combined IDA/IFA is applied at the right camera position, and a focused right image and right depth-map are obtained. The camera is moved by 50 mm to the left position. Then the combined IDA/IFA is applied at

the left camera position, and a focused left image and left depth-map are obtained. Then SIA is applied to the right and left focused images. For each image block in the right focused image, the expected disparity is obtained using the step number vs disparity table. The range of disparity to be searched for establishing correspondence is obtained by finding the disparity for DELST (=10) steps below and DELST steps above the focused step number. For each image block of size 16×16 in the right focused image, the best matching position in the left focused image is found using Sum-of-Squared-Differences (SSD).[10] SSD is computed by finding the grey-level difference between corresponding pixels of the right image block and the left image block, squaring the differences, and summing them. The depth-map obtained by matching 16×16 blocks can be refined by more intelligent matching techniques. This refinement step will be implemented in the future on SVIS. One simple refinement is to match 8×8 blocks in a range of ± 8 pixels from the current estimate of disparity.

6 Experiments

The algorithm for integrating IDA, IFA, and SIA was named DEFOST (DEfocus FOcus and STereo). The implementation of DEFOST was tested on SVIS. The algorithm parameters were set to operate well for objects in the depth range of 30cm. to 80 cm. from the camera. The results on three of the objects– prism, cone and face– are presented here. The objects were placed about 60 cm distance from the camera. In general, the magnification correction data obtained by calibration was sufficiently accurate for IFA and SIA, but not for IDA. The results of IDA was satisfactory at the center of the image where the effect of magnification change was minimum, but the results were not satisfactory away from the center. This was compensated by allowing for error in IDA and applying IFA in a larger range than necessary, thus increasing the number of images and computing used. It should be noted that shape of objects is obtained by looking for changes in scene depth-map. Therefore the percentage error in depth-map will be much less than that in shape. The focused image and 3D shape of a prism recovered from IDA/IFA are shown in Fig. 12 and 13 respectively. The final 3D shape recovered from DEFOST after stereo disparity analysis is shown in Fig. 14 (In all plots of 3D shape in the Figures, X-Y plane (Z= 0) is at 1000 mm from the camera, and Z-axis points towards the camera). These results were obtained by using 2 image frames for IDA and about 10 image frames for IFA for each of the right and left camera positions. The 3D plots of shape are somewhat distorted by the plotting software used by us because the depth-map data obtained by SIA was not corrected for perspective distortion in imaging. The perspective distortion makes a constant size surface patch appear smaller (i.e. projects onto smaller image area on the CCD array) as it moves farther. The DEFOST software implementation was not optimized to minimize computation since the goal was to demonstrate the conceptual feasibility of integrating IDA, IFA, and SIA. Therefore, for each object, shape recovery takes about 3 minutes on the Pentium PC. We believe that this time can be reduced by a factor of 3 or more by optimizing the software design. Results for a cone object are shown in Figs. 15-17, and for a Face object are shown in Figs. 9-11. In general the results are found to be satisfactory. A precise quantitative analysis of the accuracy of shape is yet to be carried out. Significant further improvements in accuracy can be obtained by using more intelligent stereo matching techniques than the simple SSD technique used by us.

7 Conclusion

We have integrated defocus and focus analysis with stereo image analysis for recovering the 3D shape of objects. The integrated method has been implemented on a vision system– SVIS– and the effectiveness of the method has been demonstrated. The main advantage of the method is the reduction in computation and errors in shape caused by stereo correspondence matching. A rough depth-map provided by defocus and focus analysis decreases the search space for correspondence matching and avoids false matches due to occlusion. It is found that careful calibration of the camera system with respect to several factors are very important in building a successful vision system that integrates defocus, focus, and stereo. Details on these calibration methods are presented. Further improvements to our implementation are possible in several respects. These will be explored as part of our future research.

Acknowledgement: The support of this research in part by Olympus Optical Co., Japan, is gratefully acknowledged.

8 REFERENCES

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

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

[3] J. Enns and P. Lawrence, "A Matrix Based Method for Determining Depth from Focus", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 1991.

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

[5] M. Subbarao, T. Choi, and A. Nikzad, "Focusing Techniques", *Journal of Optical Engineering*, Vol. 32 No. 11, pp. 2824-2836, November 1993.

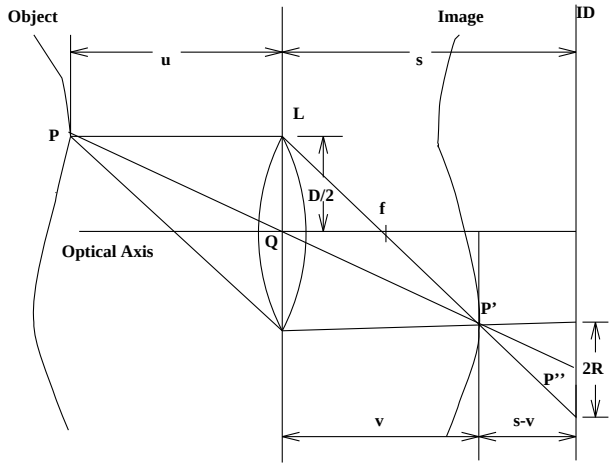
[6] M. Subbarao, and J.K. Tyan, "The Optimal Focus Measure for Passive Autofocusing and Depth-from-Focus", *Proceedings of SPIE conference on Videometrics IV* , Philadelphia, Oct 1995.

[7] M. Subbarao and G. Surya, "Depth from Defocus: A Spatial Domain Approach", *International Journal of Computer Vision*, 13, 3, pp. 271-294 (1994).

[8] G. Surya, "Three-dimensional scene recovery from image defocus", Ph.D. Thesis, Dept. of Electrical Engg., SUNY at Stony Brook, 1994.

[9] T. Wei, "Three-dimensional machine vision using image defocus", Ph.D. Thesis, Dept. of Electrical Engg., SUNY at Stony Brook, 1994.

[10] M. Okutomi and T. Kanade, "A Multiple-Baseline Stereo", *IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition*, (1991).



L Lens | p' Focused Point | D Aperture Diameter
P Object | Q Optical Center | f Focal Length
ID Image Detector | p'' Blur Circle | R Blur Circle Radius

Figure 1: Image Formation in a Convex Lens

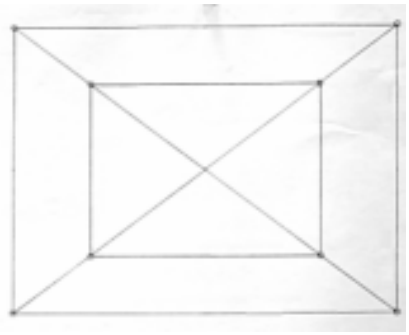


Figure 2: Planar Pattern for Computing Magnification Factor

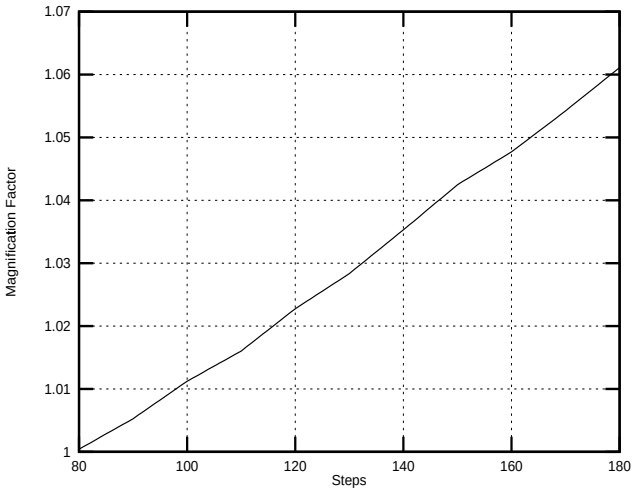


Figure 3: Magnification vs Lens Step

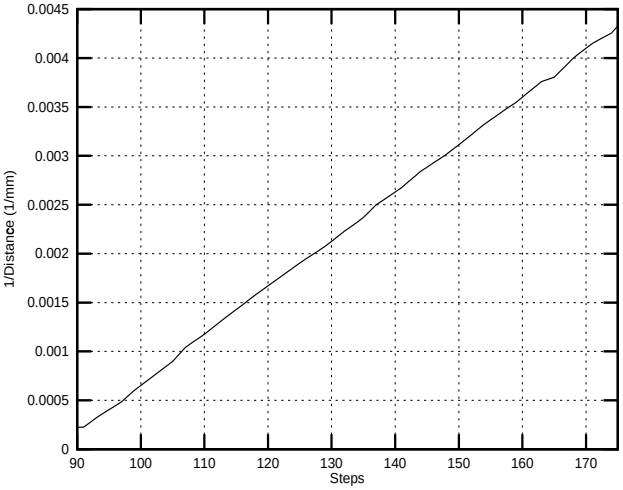


Figure 4: Lens Step vs 1/Focusing Distance

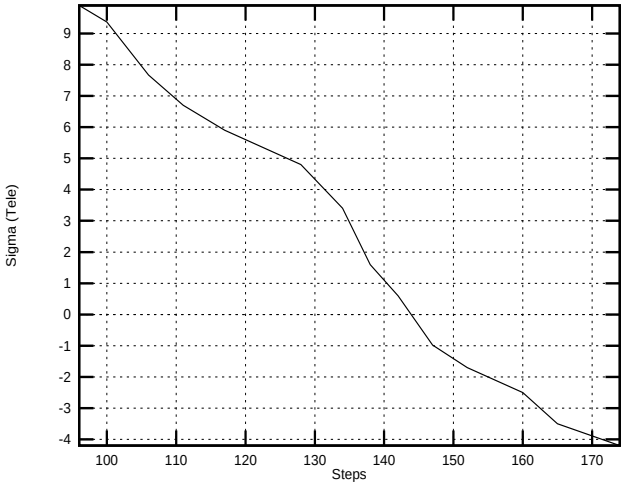


Figure 5: Blur Parameter vs Focused Lens Step

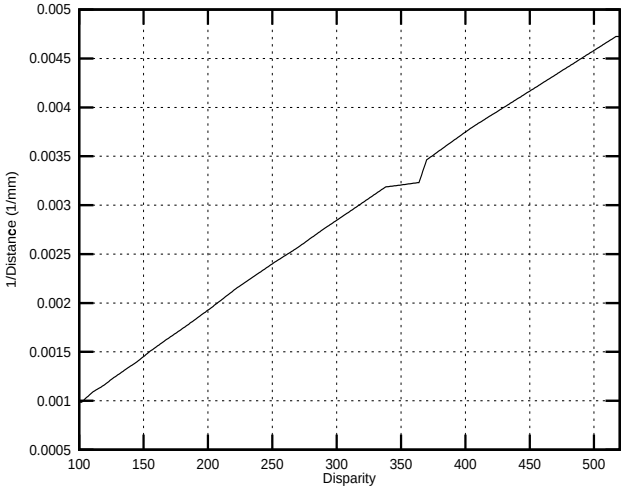


Figure 6: Disparity vs 1/Distance

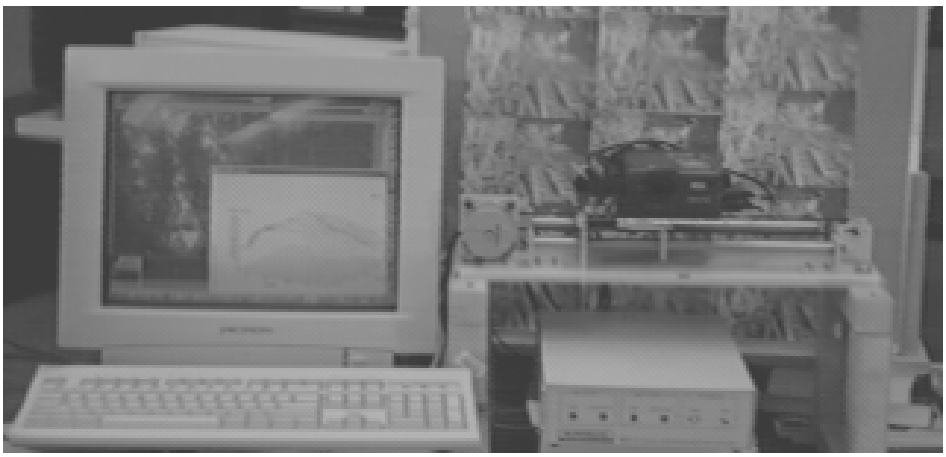


Figure 7: Stonybrook VIsion System (SVIS)

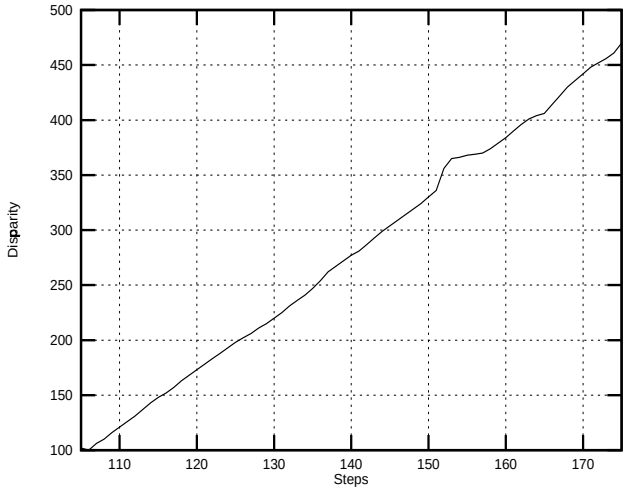


Figure 8: Disparity vs Lens Step

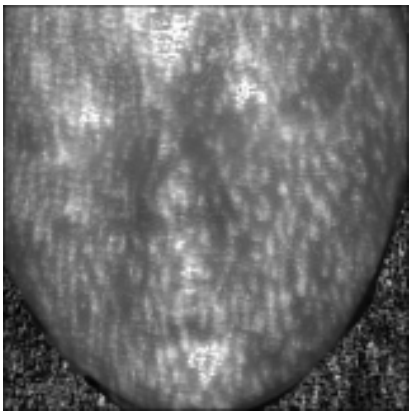


Figure 10: Reconstructed Focused Image of the Face Object (with projected random pattern)

Figure 9: Face Object

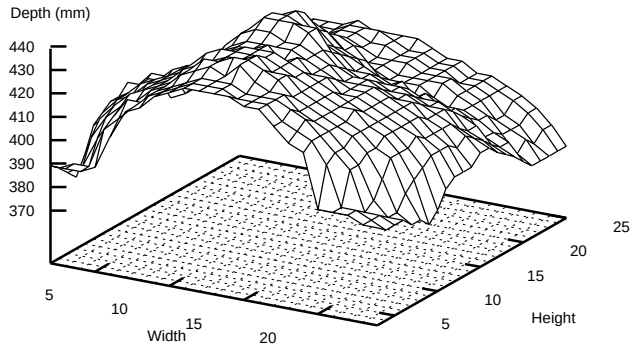


Figure 11: Final Depth Map of the Face Object

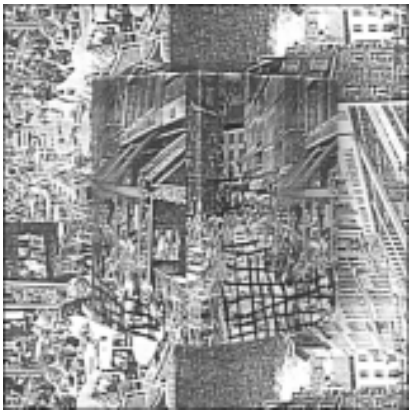


Figure 12: Reconstructed Focused Image of a Prism Object

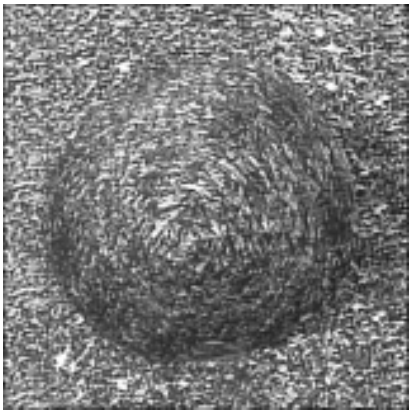


Figure 15: Reconstructed Focused Image of a Cone Object

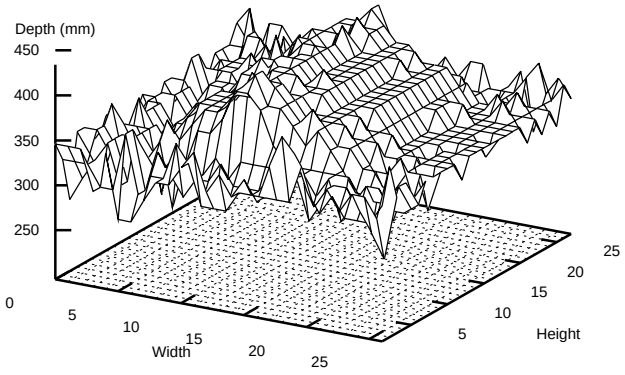


Figure 13: Intermediate Depth Map of the Prism Object

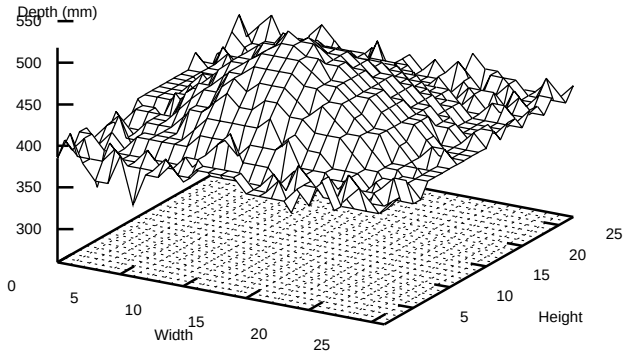


Figure 16: Intermediate Depth Map of the Cone Object

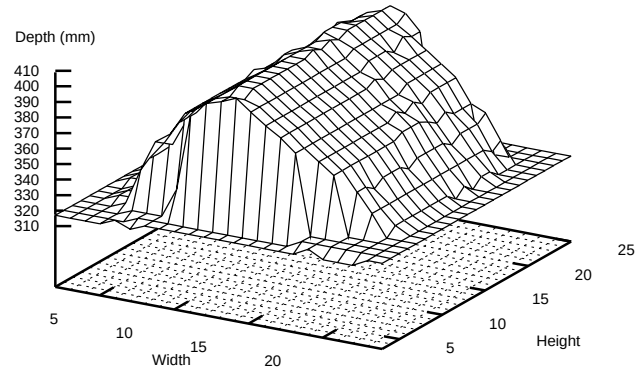


Figure 14: Final Depth Map of the Prism Object

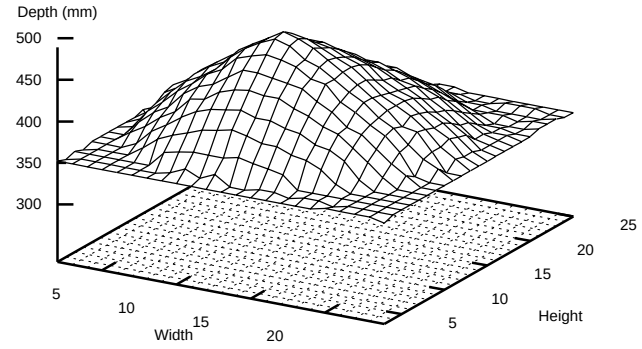


Figure 17: Final Depth Map of the Cone Object