Use of a non-frontal camera for extended depth of field in wide scenes

Arun Krishnan Narendra Ahuja

e-mail: arunki@vision.csl.uiuc.edu
e-mail: ahuja@vision.csl.uiuc.edu
Beckman Institute,
University Of Illinois
405 North Mathews Ave.,
Urbana IL 61801, U.S.A

ABSTRACT

This paper describes a method for obtaining a composite focused image from a monocular image sequence. The image sequence is obtained using a novel non-frontal camera that has sensor elements at different distances from the lens. This paper first describes the motivation behind the non-frontal camera, followed by the description of an algorithm to obtain a focused image of a large scene. Large scenes are scenes that are deep and wide (panoramic). Consequently, the camera has to be panned in order to image all objects/surfaces of interest. The described algorithm integrates panning and generation of focused images. Results of experiments to generate extended depth of field images of wide scenes are also shown.

1. INTRODUCTION

This paper is concerned with creating a composite image of a panaromic scene that has objects at different depths and which cannot all be imaged in one image frame. A camera therefore has to pan across the scene to image all points of interest.

Obtaining focused images of even one frame is difficult when the visible object points lie at different depths from the camera. For a given position of the sensor plane from the lens, only parts of the scene that lie within the depth of field will appear in sharp focus. One way to compose a focused image of a scene with objects at different depths is to first obtain a series of images, each taken with different sensor plane positions (different values of v). Each image can be analyzed to determine sharply focused regions, and then the focused regions from each image can be put together to obtain a final focused image of the scene [1, 2]. For scenes that are wider than the field-of-view of a camera, image sequences will have to be taken with varying values of v, for multiple pan angles until all parts of the scene are covered. As will be shown later in this paper, a non-frontal camera simplifies this process considerably by integrating the two mechanical motions of panning and focusing into one.

Since in common cameras and lens assemblies change of focus (v) is usually done by either a manual control

0-8194-1321-6/93/$6.00

or a mechanical control of the focus ring in the lens system, integration of panning with focusing leads to increased operational speed. Further, when the sensor plane is displaced within the lens system, some elements may undergo translational motion or a screw type rotational motion with the final effect being that the distance of the sensor plane (fixed to the camera) from the lens center may change. Thus focusing action can sometimes also have the side effect of changing the distance of the viewed object from the lens center which is avoided by the above integration.

Another common method of obtaining a focused image of a scene with objects at different depths is to adjust the camera parameters to cause a large enough depth of field, for example by reducing the aperture down to a pinhole. This has disadvantages in that it is not always possible to increase the depth of field (for example in microscopes) and in situations with low ambient lighting it is not possible to reduce the aperture down to a pinhole without affecting the quality of the image. In any case, the camera will have to pan in order to image all parts of a wide scene.

2. A NON-FRONTAL CAMERA

2.1. Motivation

The following observations underlie the proposed non-frontal camera [3]. In a normal camera, all points on the sensor plane lie at a fixed distance (v) from the lens. So all scene points are always imaged with a fixed value of v, regardless of where on the sensor plane they are imaged, i.e., regardless of the camera pan angle. If we instead have a sensor surface such that the different sensor surface elements are at different distances from the lens, then depending upon where on the sensor surface the image of a scene point is formed (i.e., depending on the camera pan angle and the location of the object), the imaging parameter v will assume different values. So for fixed scene points, the v value for that scene point will vary as a function of the pan angle. This means that by controlling just the pan angle, we could achieve both goals of the traditional mechanical movements, namely, that of changing v values as well as that of scanning the visual field, in an integrated way. Another use of a non-frontal camera is in finding the range of points in a scene [4, 5].

2.2. Tilted sensor plane

Consider a non-frontal camera with a sensor plane that is not perpendicular to the optical axis as shown in Figure 1. The tilt of the sensor plane is controllable though for the experiments described in this paper the tilt will be kept constant. The entire camera pans about a vertical axis that goes through the optical center of the lens assembly.

The tilt of the sensor plane causes different sensor pixels to be at different distances from the lens. For a fixed tilt, consider the set of object points that will image with sharpest focus on the sensor surface (Sharp

Focus, or SF surface). Derivation of the lens law for the above optical arrangement (which is not included due to page limitations) shows that the SF surface will be as shown in Figure 2. The significance of the SF surface is that when the camera undergoes a pan rotation, the SF surface sweeps out a region in object space. All object points that lie within the swept volume will be imaged with sharp focus in an image among the sequence captured by the camera during it's motion. A regular frontal camera on the other hand has a SF surface that is roughly planar and perpendicular to the optical axis so any pan rotation causes only a small swept volume.

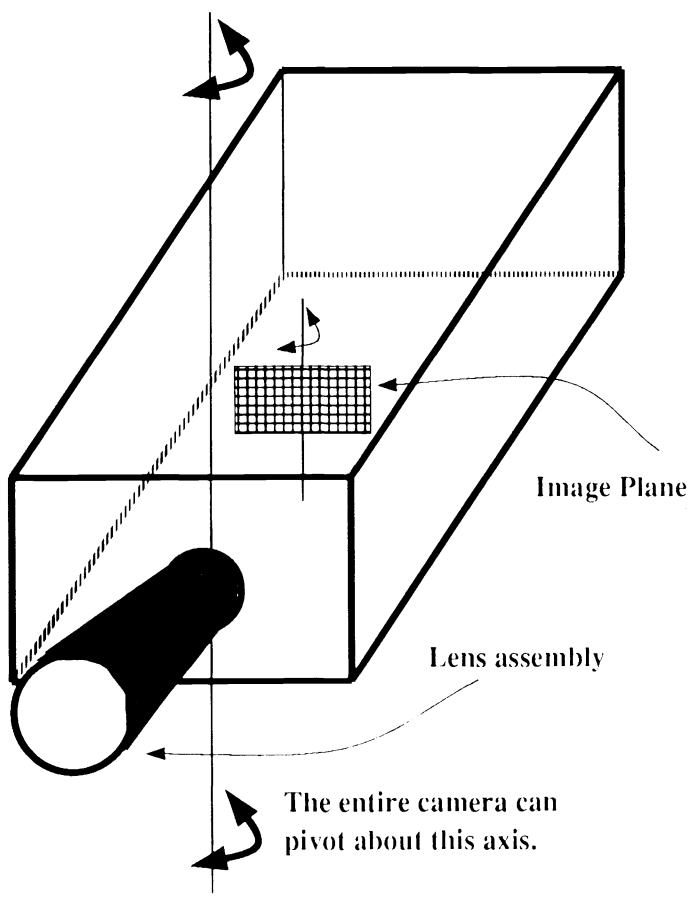


Figure 1: Schematic of the modified camera

3. OBTAINING A FOCUSED IMAGE

The non-frontal camera will image with sharp focus a volume of object space that is also tilted with respect to the optical axis. So in any one image frame, there will be scene points in sharp focus that are at different depths from the camera. If the entire camera is now rotated in sequential small steps about a vertical axis through the lens center, then the entire scene will be sequentially imaged by a plurality of frames. By

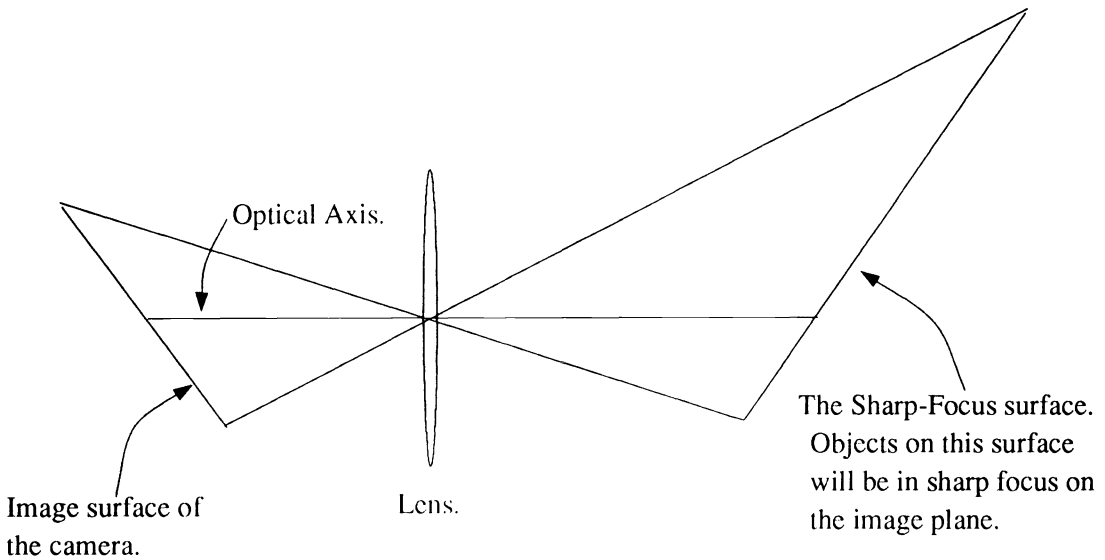


Figure 2: The SF surface for the proposed camera with a tilted image plane. The SF surface is not parallel to the lens and the optical axis is not perpendicular to the SF surface.

choosing the angle of rotation corresponding to each frame to be sufficiently small, each scene point will be imaged in multiple frames, but by different sensor elements in different frames. Since the sensor elements are located at different distances from the lens center, the scene point will therefore be imaged with different levels of blur, the sharpest focus image occurring for some camera rotation, and progressively blurred images as the camera rotates further away. The non-frontal camera thus integrates paning and focusing into just panning alone.

To determine when a scene is imaged with the least blur, several autofocus methods have been proposed in the past. Horn [6] describes a Fourier-transform method, in which the normalized high-frequency energy from a one-dimensional FFT is used as the criterion function to maximize. Sperling suggests the squared Laplacian as a measure of image blur which must be minimized [7]. Tennenbaum uses a thresholded gradient magnitude method, in which Sobel operators are used to estimate the gradient [8] . The criterion function used is the sum, over some image window, of the squared gradient magnitudes exceeding a certain threshold. This has also been used by Krotkov [9, 10]. Jarvis suggests sharpness measures based on entropy, variance, and gradient [11]. A survey and comparison of several criterion values is presented in [12]. The criterion functions described there make use of such measures as signal power, gray level standard deviation, thresholded pixel counts, and summation of squared gradient in one dimension. Schlag et al. also implement and compare

several auto-focusing methods, including the gradient, Laplacian, and entropy [13]. Darrell and Wohn describe a depth from focus method that varies the focus distance and uses Laplacian and Gaussian pyramids to obtain the range [14]. Nayar et. al. propose a sum-modified-Laplacian where they add the magnitudes of the second derivatives along individual axes [15]. This was done to avoid unstable behavior for textured images whose second derivatives have opposite signs and hence cancel out. For the results reported in this paper, we use the Tennegrad criterion function.

Algorithm Let the image plane have N x N pixels and let the focus map be a large array of size **N x sN**, where **s >= 1** is a number that depends on how wide a scene is to be imaged. The kth image frame is represented by I_k and the desired, cumulative, focused image is represented by R. Every element in the focus array is a structure that contains the focus criterion values for different image indices, i.e., for different pan angles. When the stored criterion value shows a maximum, then the index corresponding to the maximum can be used to determine the focused image for that scene point. Let the camera start from one side of the scene and pan to the other side.

Set $j = 0$, $\phi = 0$. Initialize all the arrays and then execute the following steps.

- Capture the jth image I_j.
- Pass the image through a focus criterion filter to yield an array C_j of criterion values.
- For the angle ϕ (which is the angle that the camera has turned from its starting position), calculate the offset into the focus map required to align image I_j with the previous images. For example, Pixel $I_j[50][75]$ might correspond to the same object as pixels $I_{j+1}[50][125]$ and $I_{j+2}[50][175]$.
- Check to see if the criterion function for any scene point has crossed the maximum. If so, compute the intensity of the focused image for that scene point using the pan angle (and hence v value) for the image with maximum criterion value.
- Rotate the camera by a small amount. Update ϕ and j.
- Repeat the above steps until the entire scene is imaged.

Figure 3 illustrates the geometrical relationships between successive pan angles, pixels of the images obtained, and the focus array elements.

4. IMPLEMENTATION AND RESULTS

A non-frontal camera with a controllable sensor plane tilt was built and the following experiments were conducted.

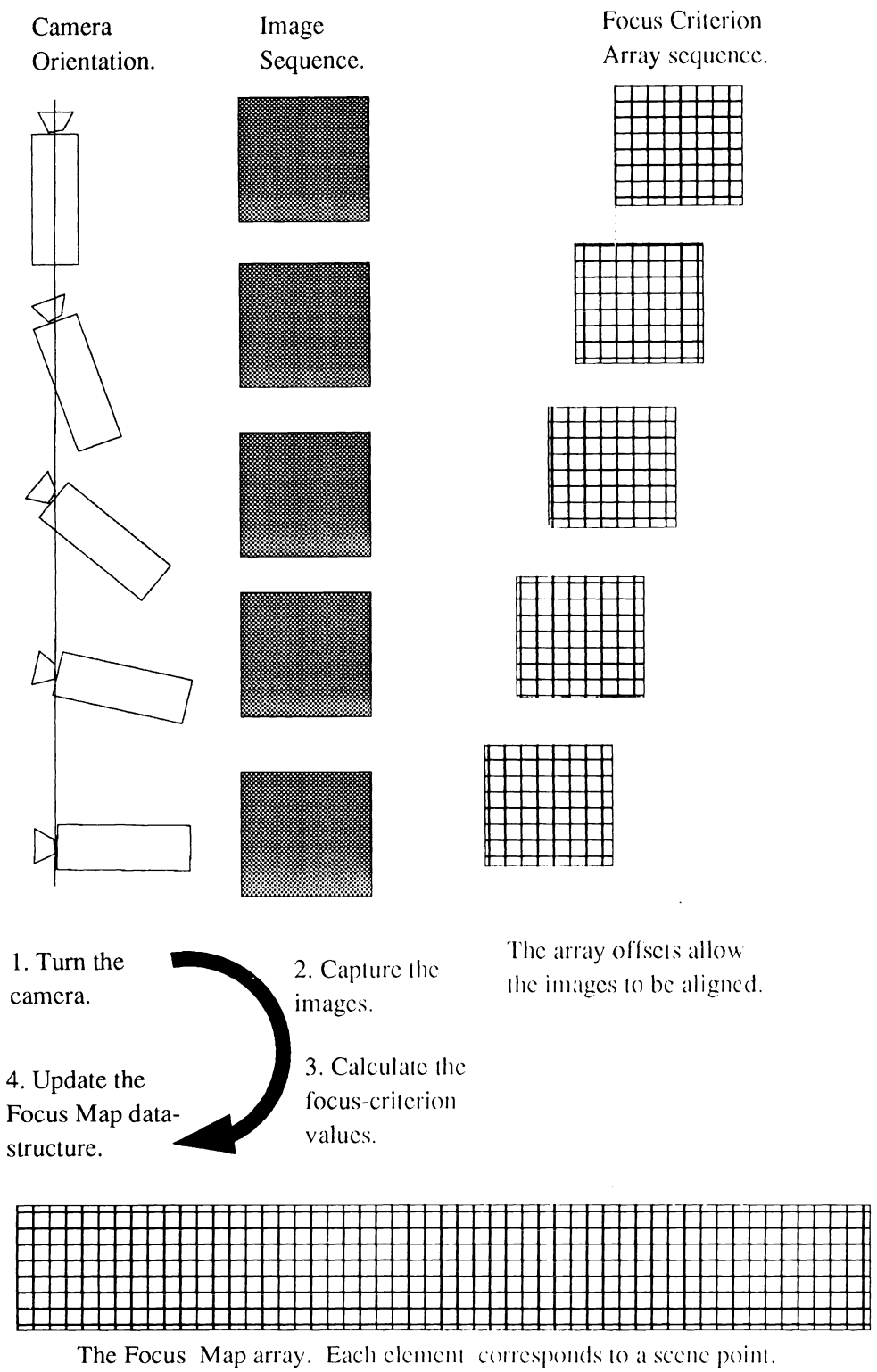


Figure 3: Panning camera, focus array, and the images obtained at successive pan angles. Each focus array element corresponds to one radial direction from the lens center and is associated with multiple criterion function values which are computed from different overlapping views. The maximum of the values in any radial direction is the one finally selected for the corresponding focus array element, to compute the image intensity value in that direction.

4.1. Illustration of SF surface

An example of focusing on two points that are far apart in range is shown in Figure 4. This shows that as expected, the SF surface is not perpendicular to the optical axis.

4.2. Experiments 1 and 2

These experiments obtain focused images of wide scenes. The scene in experiment 1 consists of, from left to right, a planar surface (range = 73 in), part of the background curtain (range = 132 in), a planar surface (range = 54in) and a planar surface (range = 38 in). The scene in experiment 2 consists of, from left to right, a planar surface (range = 70 in), a planar surface (range = 50 in), and a face of a box (at a depth of 35 in).

The camera is turned in small steps of 50 units (of the stepper motor), that roughly corresponds to a shift of 15 pixels (in pixel columns) between images. A scene point will thus be present in a maximum of thirty four (roughly $\frac{512}{15}$) images. Each frame was processed in sequence with the focus criterion values calculated for all pixels. The focus map data structure was updated with the criterion values and then the next frame was processed. After all the images for a particular scene point are obtained (a maximum of 34 frames), the column number where the focus criterion value peaked was found out. This process of determining the peaks is carried out in conjunction with processing of new incoming frames and updating the focus map data structure.

Figures 5(a) and 5(b) show the composite focused images for Experiments 1 & 2.

4.3. Results

For any scene point, the image of sharp focus is found by calculating the focus criterion function and determining it's maximum. The focus criterion function is a measure of the spatial high frequency content of a window around the image point. This has two effects on the results.

- Window effects: The bigger the window, the more accurate is the high frequency content. But the window might go across occlusion boundaries thereby giving false peaks in the criterion function.
- Feature less regions: For scene points that are inherently featureless or have only low spatial frequency content (uniform brightness patches), the focus criterion function will give many false peaks and will not be reliable.

The composite images shown in Figure 5 show errors like blurred lettering and edge bleeding around occlusion boundaries and uniform brightness surface patches.

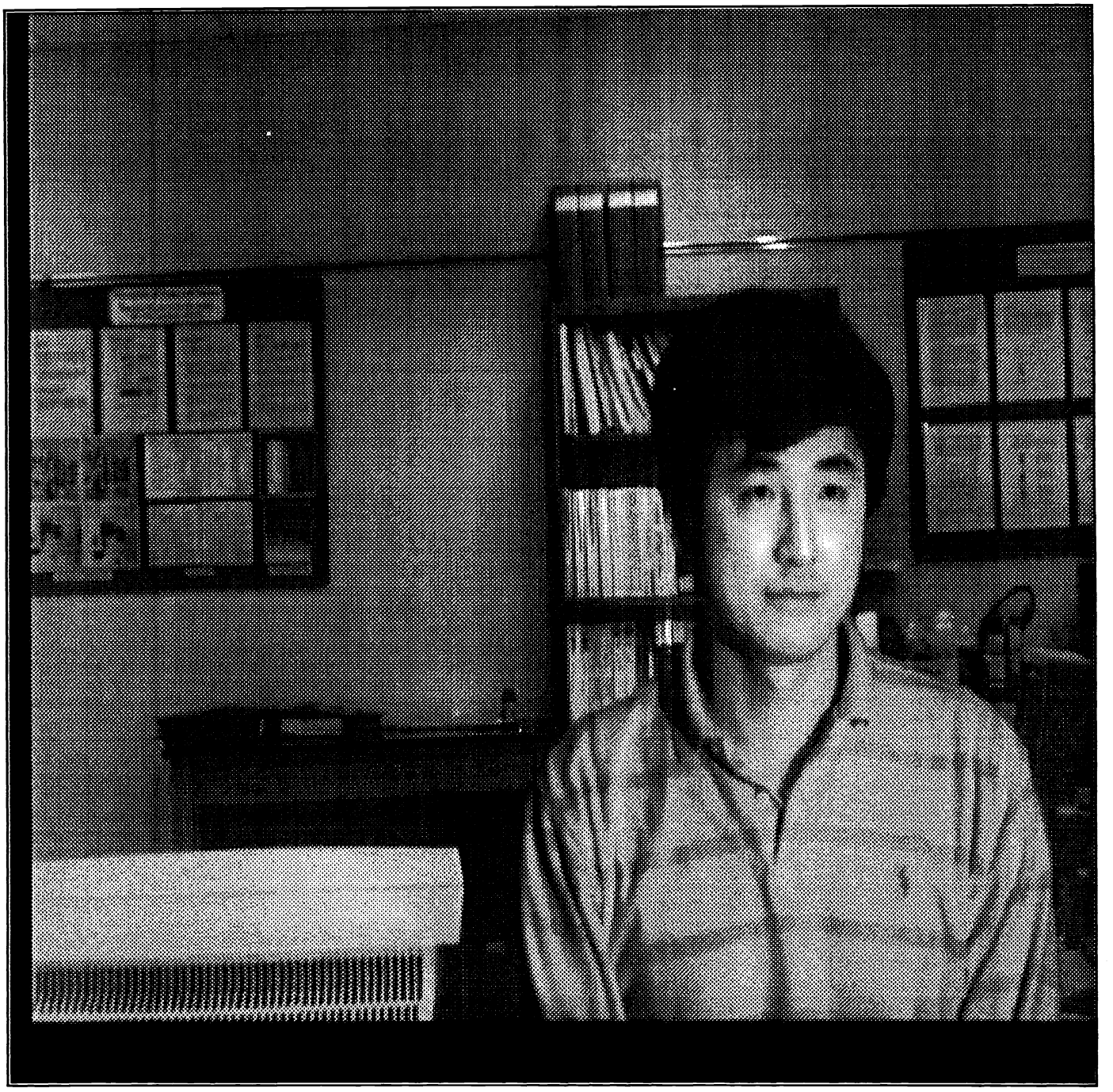
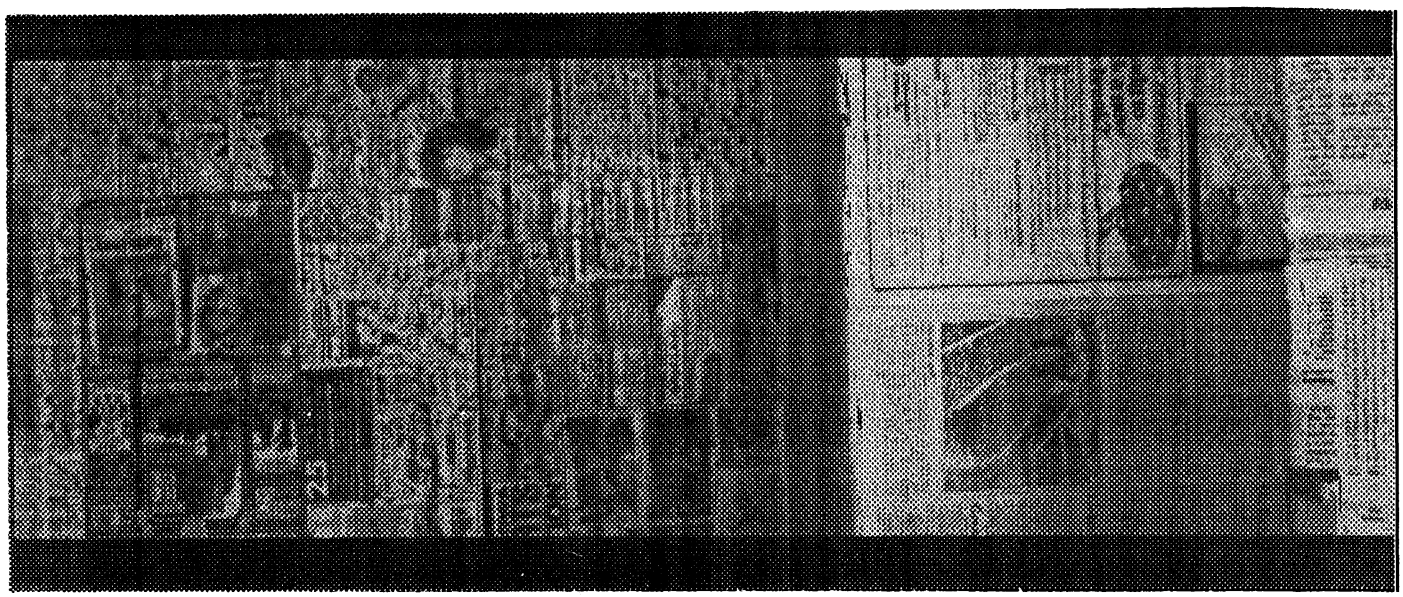
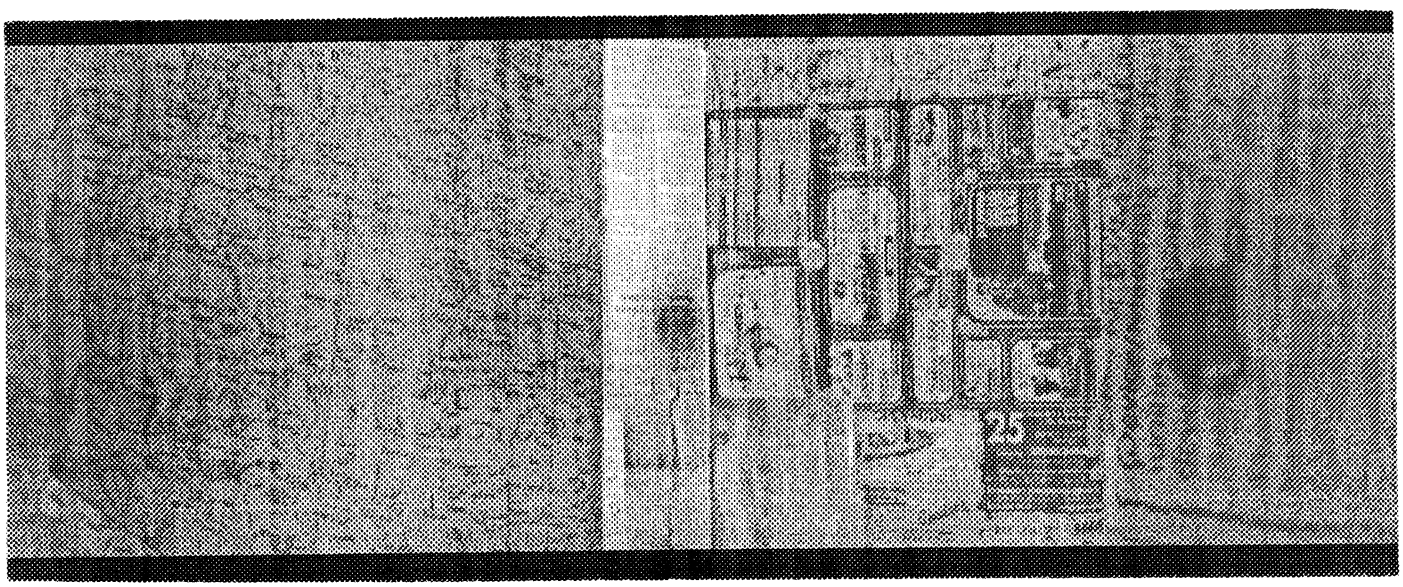


Figure 4: Image focused on two points that are far apart. One point in focus is the person in the foreground. The other point is the chart on the far left side of the wall.



(a) Focused Image for Experiment 1



(b) Focused Image for Experiment 2

Figure 5: Composite focused images.

5. SUMMARY AND CONCLUSIONS

In this paper we have shown that using a camera whose image plane is not perpendicular to the optical axis, allows us to determine a composite focused image of a wide scene. We showed that the SF surface, object points which appear in sharp focus, for the non-frontal imaging camera is approximately an inclined plane. When the camera's pan angle direction changes, by turning about the lens center, an SF volume is swept out by the SF surface. The scene points within this volume comprise those for which sharp focused images can be obtained. We have described an algorithm that determines the composite focused image. We have also described the results of experiments that were conducted to prove the feasibility of our method and point out the shortcomings of the algorithm.

6. ACKNOWLEDGEMENTS

The support of the National Science Foundation and Defence Advanced Research Projects Agency under grant IRI-89-02728 and U.S. Army Advance Construction Technology Center under grant DAAL 03-87-K-0006 is gratefully acknowledged.

7. REFERENCES

[1] R. J. Pieper and A. Kopel, "Image processing for extended depth of field," *Applied Optics*, vol. 22, pp. 1449–1453, May 1983.

[2] K. Kaneda, S. Ishida, A. Ishida, and E. Nakamae, "Image processing and synthesis for extended depth of field of optical microscopes," *Visual Computer*, vol. 8, pp. 351–360, June 1992.

[3] A. Krishnan and N. Ahuja, "Range estimation from focus using a non-frontal imaging camera," in *Proceedings of the DARPA Image Understanding Workshop*, pp. 959–965, Washington D.C., April 1993.

[4] A. Krishnan and N. Ahuja, "Range estimation from focus using a non-frontal imaging camera," in *[To appear in] the Proceedings of AAAI*, Washington D.C., July 1993.

[5] B. Wilcox, "Personal communication on tilted image plane cameras," May 1993.

[6] B. K. P. Horn, "Focusing," Tech. Rep. 160, MIT Artificial Intelligence Lab, Cambridge, Mass., 1968.

[7] G. Sperling, "Binocular vision: A physical and a neural theory," *Amer. J. Psychology*, vol. 83, pp. 461–534, 1970.

[8] J. M. Tenebaum, *Accomodation in Computer Vision*, PhD thesis, Stanford University, Palo Alto, Calif., 1971.

[9] E. P. Krotkov, "Focusing," Tech. Rep. MS-CIS-86-22, GRASP Laboratory, University of Pennsylvania, April 1986.

[10] E. P. Krotkov, J. Summers, and F. Fuma, "Computing range with an active camera system," in *Eighth International Conference on Pattern Recognition*, pp. 1156–1158, October 1986.

[11] R. A. Jarvis, "Focus optimisation criteria for computer image processing," *Microscope*, vol. 24, pp. 163–180, 1976.

[12] G. Ligthart and F. C. A. Groen, "A comparison of different autofocus algorithms," in *Proc. Sixth Intl. Conf. Pattern Recognition*, pp. 597–600, Oct. 1982.

[13] J. F. Schlag, A. C. Sanderson, C. P. Neuman, and F. C. Wimberly, "Implementation of automatic focusing algorithms for a computer vision system with camera control," Tech. Rep. CMU-RI-TR-83-14, Carnegie-Mellon University, August 1985.

[14] T. Darrell and K. Wohn, "Pyramid based depth from focus," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 504–509, 1988.

[15] S. K. Nayar and Y. Nakagawa, "Shape from focus: An effective approach for rough surfaces," in *Proc. IEEE Intl. Conf. Robotics and Automation*, pp. 218–225, Cincinnati, Ohio, May 1990.