Omnifocused 3D Display Using the Nonfrontal Imaging Camera

Andrés Castaño and Narendra Ahuja
Department of Electrical Engineering and Beckman Institute
University of Illinois at Urbana-Champaign
Urbana, IL, 61801

Abstract

The Nonfrontal Imaging Camera (NICAM) is a system that simultaneously provides omnifocused intensity and range panoramics (i.e., images and depth maps). In this paper we discuss the process of acquisition of range, the parameters that affects it and provide experimental results. The camera-centered image and depth map obtained by the NICAM provide the 3-D position and radiance of every object of the scene. This information can be used to reconstruct the 3-D world as it was seen from the point of view of the camera, without incurring in the error-prone and expensive step of view integration required by other systems to create a complete 3-D world from information obtained from different points of view. The 3-D world obtained can be easily integrated into a virtual reality environment.

1 Introduction

A lot of research has been done in both the acquisition of panoramic images and range maps using off-the-shelf cameras. On one hand, a panoramic image provides a large field of view which is desirable in a large number of applications. For example, panoramic images can be of great help in robot navigation since the robot can monitor all its surroundings simultaneously. Unfortunately, the panoramic images used in navigation have deficiencies since they are either obtained through a distortion process (e.g., silvered cones or spheres [3], [8], [9]) or are constructed by a mosaic process that does not provide a seamless image [10]. In either case, the depth of field (DOF) is fixed and limited by the settings of the camera, i.e., the images are subjected to the same focusing limitations present in standard images. These same limitations are present in other panoramic imaging systems used for virtual reality, surveillance, tele-presence and security. On the other hand, range maps have widespread use in applications that vary from scene and object modeling to robot navigation [7]. The most used techniques to obtain range maps using cameras are stereo (based on triangulation) and depth from focus (based on the physical principles of the lens). Depth from focus techniques, although less accurate than the stereo techniques, have the advantage of using only one camera instead of two [2].

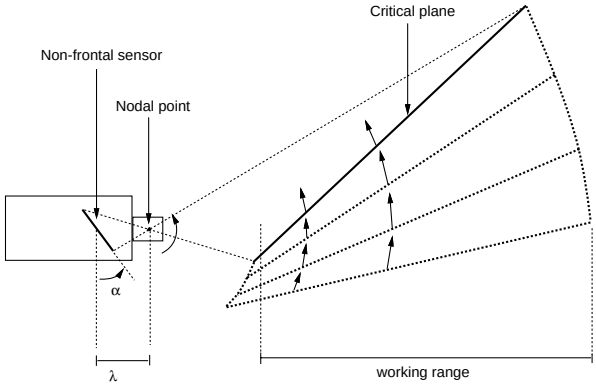

Figure 1: Top view of the NICAM imaging system.

The Nonfrontal Imaging Camera (NICAM) is a dual sensor that outputs both a panoramic image and its corresponding range map accomplishing a task that otherwise would require separate imaging and range sensors [1], [5]. Additionally, the panoramic image produced does not have the limitations on the DOF that other panoramic systems have, i.e., all the scene is in focus. In this camera, the sensor is tilted with respect to the lens in such a way that the objects that lie in the critical plane of focus are, not in a plane parallel to the lens as is the case in a standard camera, but in a plane oblique to the plane of the lens, as shown in Figure 1. When the whole system is rotated about a vertical axis located at the nodal point of the lens, this oblique in-focus plane sweeps the scene bringing all the objects within the range into focus at one moment or another, as shown in Figure 2. The correctly focused pixel is selected using a depth-from-focus algorithm which also provides an estimate of the distance from the lens to the object that irradiated the pixel.

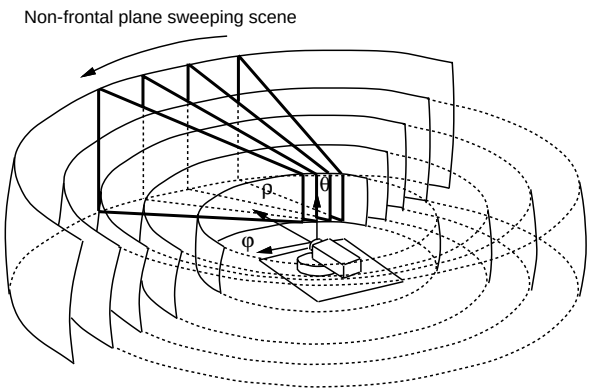

Figure 2: Nonfrontal sensor sweeping a scene.

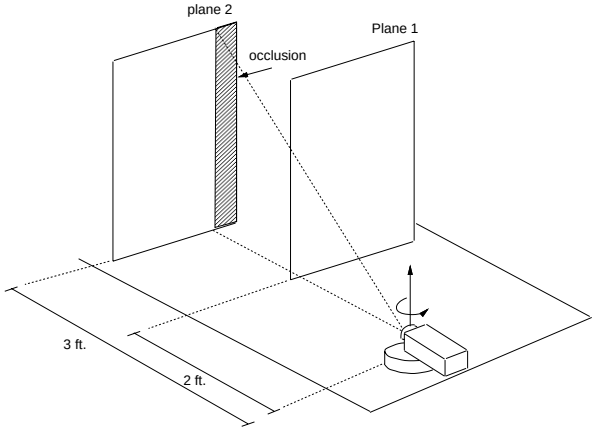

Figure 3: Sample experimental setup.

Finally, these focused pixels are used to build a fully focused panoramic of the scene swept by the camera.

The organization of this paper is as follows. In Section 2 we discuss the depth-from-focus algorithm as it applies to the NICAM system. In Section 3 we discuss how the camera parameters and the panoramic acquisition parameters affect the quality of the range. In Section 4 we discuss the combined use of the intensity and range panoramics to reconstruct the 3-D world as it was seen by the camera. Finally, in Section 5 we present our conclusions.

2 NICAM depth-from-focus

Fully-focused intensity and range panoramics can be easily obtained from the output of a NICAM system. In Section 2.1 we describe how to create range maps of the scene using a depth-from-focus algorithm. In Section 2.2 we describe how to create fully-focused panoramic images using the results obtained from the depth-from-focus algorithm. To illustrate the results, we obtain the intensity and range panoramics of the two-planes setup shown in Figure 3.

2.1 Range estimation

Depth-from-focus is a process used to find the distance from the lens to an object based on the degree of focus of its image. The distance relationship between a 3-D point P and its focused image Q is given by the thin-lens law,

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

where f is the focal length of the lens, v is the distance between P and the lens and u is the distance between the lens and the focused image Q. Thus, we can estimate the distance from the lens to P using

$$v = \frac{uf}{u - f} \tag{2}$$

if we can estimate u. To do this we first need to be able to find when a pixel is in focus.

The 3-D point P is imaged multiple times as the plane of critical focus sweeps the scene (as shown in Fig. 2). The image of P will be in focus only on the picture acquired when the plane of critical focus was intersecting P; all other pictures will contain out-of-focus images of P. Each picture is dewarped carefully from the original planar surface to a spherical surface, to force the different pictures to match perfectly, on top of each other, when overlapped. From all the versions of the images of P, we assign Q, the focused image of P, to the image of P with the largest gradient found using a Sobel operator (similar evaluations could be done with a Laplacian operatior [6]). In other words, we are using the gradient as a high-pass filter and selecting Q to be the image of P with the largest high-frequency component. This is an appropriate assignment because well-focused images are sharp, with a lot of detail (i.e., have high frequency components) while out-of-focus images are blurred and details cannot be distinguished (i.e., do not have high frequency components). Now, we can use information related to Q to estimate u, the distance between Q and the lens.

The normal distance u between the lens and a pixel (I, J) of the sensor of a NICAM system is

$$u(I) = \lambda - S_x(I - I_o)\sin\alpha \tag{3}$$

where λ is the distance between the lens and the sensor along the optical axis, α is the angle of tilt of the sensor, S_x is the size of the horizontal sampling interval, and I_o is the abscissa of the center of the image [1]. Therefore, evaluating Eq. (3) on Q_x and replacing $u(Q_x)$ in Eq. (2) provides an estimate of the distance between the lens and P.

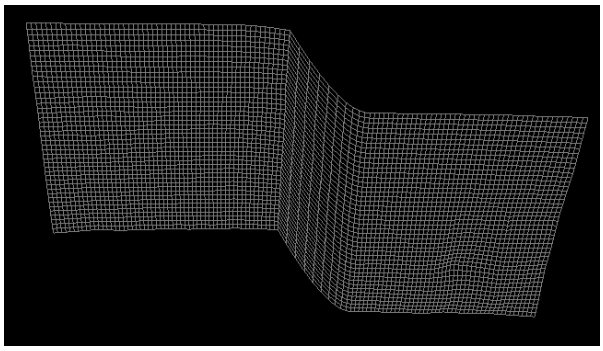

Figure 4: Range estimate from sample.

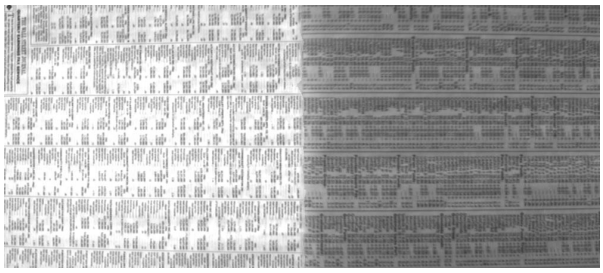

Figure 5: Omni-focused image from sample.

The process of estimating the distance to a 3-D point P based on the information of pixel $Q = (Q_x, Q_y)$ is repeated for every pixel of the panoramic and the result is a depth map of the scene. Figure 4 shows a wireframe version of the depth map obtained for the setup shown in Figure 3.

2.2 Omni-focused images

The process of obtaining fully-focused panoramics uses the results obtained by the depth-from-focus algorithm. A fully focus intensity panoramic is an image in which the intensities of all its pixels correspond to the intensity of the correctly focused pixel found using the depth-from-focus algorithm. Figure 5 shows the intensity panoramic obtained for the experiment shown in Figure 3. The images of both planes are correctly focused.

Intensity panoramics obtained with the NICAM have an extremely large depth of field equal to the distance between the closest and farthest points swept by the inclined plane of critical focus. Hence, the panoramics have the focused images of both near and far objects. This characteristic of the system is particularly useful for imaging of dark scenes, where the depth of field cannot be extended by closing the aperture. Figures 6 and 7 show two images of the same scene. Figure 6 was obtained with a NICAM working as a standard camera (i.e., CCD sensor parallel to the lens). Since the plane of critical focus is located at

Figure 6: Image of scene taken with a standard camera.

Figure 7: Image of scene taken with omni-focus camera.

4 ft. from the camera, the images of objects located closer or farther than 4 ft. appear out of focus. Figure 7 is a portion of a panoramic obtained with a NICAM (i.e., CCD sensor was tilted). In this case, the depth of field is much larger and covers all the objects of the scene and thus, all the scene is in focus.

3 Analysis of errors in range

There are two factors that affect the quality of the range estimates. The first factor that affects the range estimates is the influence of errors of the camera parameters and the selection of the focused pixel [1]. The range estimate is obtained from Eqs. (2) and (3):

$$v = \frac{(\lambda - S_x(I - I_o)\sin\alpha)f}{(\lambda - S_x(I - I_o)\sin\alpha) - f} \tag{4}$$

The error in range as a function of the camera parameters and the selected pixel (I, J) is the summation of the errors due to each parameter:

$$\Delta v = \sum \left|\frac{\partial v}{\partial k}\right| \Delta k, \qquad k = f, \lambda, S_x, I_o, I, \alpha \tag{5}$$

This expression reduces to

$$\Delta v \approx \frac{f^2}{(\lambda - f)^2}(\Delta\lambda + S_x \sin\alpha(\Delta I + \Delta I_o)) \qquad (6)$$

assuming that $\lambda \gg S_x(I - I_o)\sin\alpha$, $\Delta S_x \approx 0$, $\Delta\alpha \approx 0$, and $\Delta f = 0$. This indicates that the larger errors in range are caused by uncertainties in the value of λ. Also, errors in the abscissa of the center of the image and in the selection of the pixel cause minor errors in range, particularly when low resolution sensors (large S_x) and large sensor tilts (large α) are used. Notice however, that regardless of the particular influence of each parameter on the quality of the range estimate, their correct estimation is extremely important to achieve the correct image dewarping and brightness correction that allows different images to overlap perfectly. To explain a final implications of Eq. (6) we now define the working range of the camera.

The working range of the NICAM is defined as the normal distance between the closest and farthest points that can be focused at some point with the tilted sensor, as shown in Figure 1. The working range can be controlled using either α, the angle of tilt of the sensor, or λ, the normal distance between sensor and lens. To analyze Eq. (6), we are particularly interested in λ. If we increase λ the working range will reduce and move closer to the camera. This near working range could be used in object modeling, for example, where we are interested in range measurements of distances to objects that lie between 1 and 2 ft. away from the camera. If we reduce λ, the working range will be enlarged and move away from the camera. This far working range could be used for robot navigation, for example, where we are interested in range measurements of distances between 5 and 30 ft. As expected, Eq. (6) indicates that regardless of the errors in camera parameters and selection of pixel, the error in range is small if we are working with a near working range, i.e., if $\lambda \gg f$. Likewise, the error in range is large when we are using a far working range.

The second factor that affects the quality of the range estimates is the depth of field of the camera. In theory, a point P located at a distance v_0 from the lens with focal length f will have a focused image Q only when the sensor is located at a distance of u_0 from the lens that satisfies the thin-lens law given by Eq. (1). If P moves toward or away from the lens to v_1 or v_2, the image of P will no longer be the focused point Q but a blurred circle, as shown in Figure 8. However, the energy received by the pixel that is imaging P will be always the same as long as the blurred circle remains inside the boundaries of the pixel and thus, the pixel

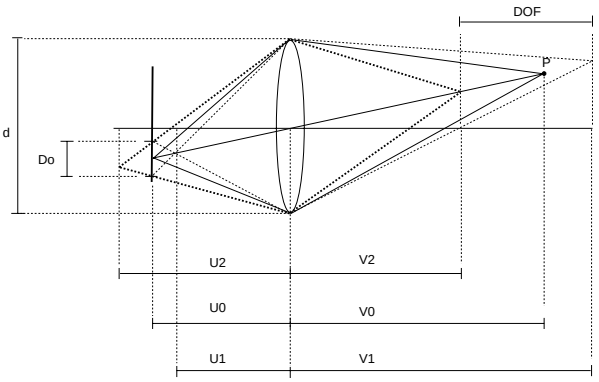

Figure 8: Depth-of-field of a CCD sensor.

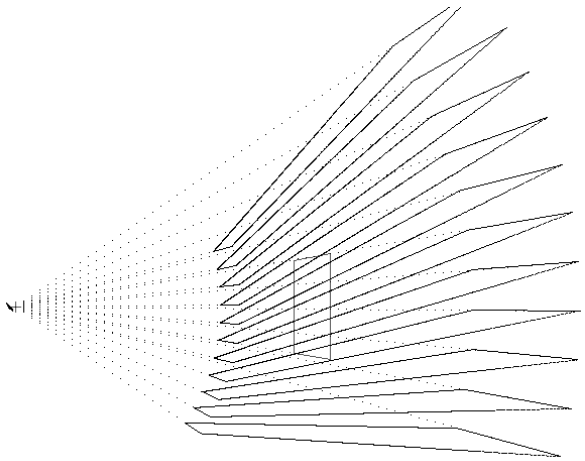

Figure 9: Focused areas of a nonfrontal camera.

will appear to be in focus. The distance between the closest and farthest points where P can move while keeping the blur circle within the boundaries of the pixel is called the depth of field (DOF). In range estimation, it is extremely important to have a narrow DOF since the DOF sets the maximum accuracy at which we can determine the distance to an object.

The focused areas that surround the plane of critical focus of a nonfrontal camera extend from the near field to the far field, as shown in Figure 9. For comparison purposes, Figure 9 also shows the rectangular DOF that the camera would have had if the sensor was not tilted. As the camera spins and takes pictures, different focused areas will be created. If the location of the 3-D point whose distance to the camera we are trying to estimate lies within one of the focused areas, then the maximum accuracy of its range will be given by the DOF which can be approximated to:

$$DOF = \frac{2\ d\ D_o\ \cos\alpha\ f\ v_0\ (v_0 - f)}{f^2\ d^2\ - {D_o}^2\ \cos^2\alpha\ (v_0 - f)^2} \qquad (7)$$

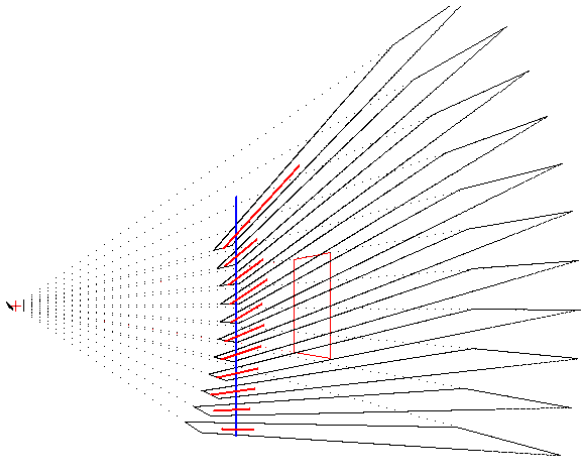

Figure 10: Range error incurred at estimating the distance to a plane.

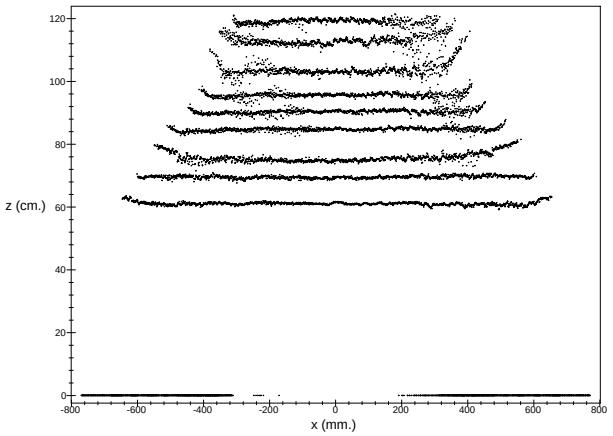

Figure 11: Range error incurred at estimating the distance to a plane.

where D_o is the maximum diameter that the blur circle can have and d is the aperture of the lens, as shown in Figure 8 [1]. This equation indicates that we can obtain narrow DOFs if we favor the use of lenses with large focal lengths f and the use of wide apertures d. Unfortunately, this means that acquiring good depth maps and acquiring panoramics with a large vertical field of view are conflicting goals.

A final source of error associated with the camera functionality appears when the location of the 3-D point whose distance to the camera we are trying to estimate does not lie within one of the focused areas but in between two of them. In this case, the uncertainty on the range will be larger than the DOF itself as shown in Figure 10. In the Figure we see the range estimates of a plane located parallel to the plane of the lens. Due to the presence of gaps between the focused ares, some of these estimates are much larger than the corresponding DOF for that distance. To eliminate these gaps and limit the range error to the DOF we can either reduce α, the tilt of the sensor, or increase the number of pictures per degree that the camera takes as it sweeps the scene. Unfortunately, both solutions have their down-sides: reducing α reduces the working range of the camera and increasing the number of pictures per degree increases the amount of processing required for a given panoramic. Figure 11 shows the range estimates of a plane that was moved from 2 ft. to 4 ft. in steps of 3 in. obtained using $f = 25$ mm, $\alpha = 7.2^o$, f-stop $= 4.0$, and a rotation step size of 0.2^o.

4 Combined use of image and range

The output of the NICAM system can be used to recreate the 3-D environment, as it was seen from the point of view of the camera. Since we have obtained both a focused image of the scene and a range estimate to every point of the scene, we can recreate the structure and appearance of the world by texture-mapping the scene directly onto the range estimate. Figure 12 shows a perspective view of the intensity panoramic shown in Figure 3 texture mapped onto the wireframe of the depth map shown in Figure 5.

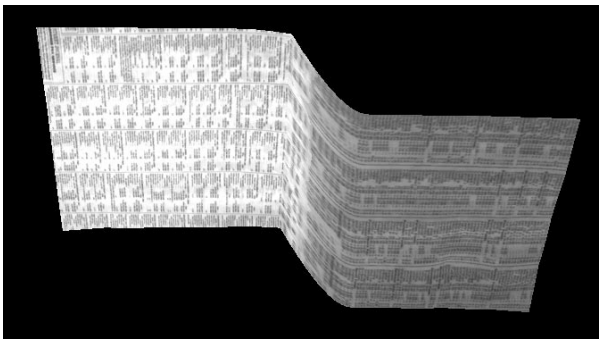

Figure 12: Combined used of image and range from sample.

Figures 13 - 15 show a series of images of a 4 cm. high chess piece. The chess piece was reconstructed using the information obtained from the point of view of the camera (we used a 25 mm. lens and a 5 mm. lens extensor, $\alpha = 7.2^o$). Figure 13 shows the 3-D view that most closely resembles the object as seen from the camera. Figure 14 shows the estimate of a frontal view of the chess piece. The right side of the chess piece is missing because this part of the object was hidden to the camera (self-occlusion). Figure 15 shows the estimate of a profile of the chess piece. Again, parts that were hidden from the camera (e.g., the area under

the chin) cannot be estimated correctly.

5 Conclusions

Experimental results suggest that the Non-frontal imaging camera can successfully acquire range estimates based on depth-from-focus algorithms. The accuracy of the range estimate depends on the particular algorithm chosen, the error of the calibration parameters and the functional parameters used to collect the images of the scene. The fully-focused panoramics and depth maps acquired by the system can be used to reconstruct the 3-D world as it was sensed from the point of view of the camera.

Acknowledgments

We thank Manoj Aggarwal and John Hart for helpful feedback and discussions. This work has been supported in part by the Defense Advanced Research Projects Agency under grant *****.

References

[1] A. Castano. Range from a nonfrontal imaging camera. Ph.D. Thesis, University of Illinois at Urbana-Champaign, 1998.

[2] S. Das and N. Ahuja. Performance Analysis of Stereo, Vergence, and Focus as Depth Cues for Active Vision. *IEEE Trans. Pattern Anal. Machine Intell.*, pp. 1213-1219, Dec 1995.

[3] J. Hong, X. Tan, B. Pinette, R. Weiss and E. M. Riseman. Image-based homing. *Proc. IEEE Int. Conf. Robotics Automat.*, pp. 620-625, 1991.

[4] B. K. P. Horn. Robot vision. MIT press, 1986.

[5] A. Krishnan and N. Ahuja. Obtaining focused images using a non-frontal imaging camera. *Image Understanding Workshop*, pp. 617-620, 1994.

[6] S. Nayar, M. Watanabe and M. Noguchi. Real-time focus range sensor. *IEEE International Conference on Computer Vision*, pp. 995-1001, 1995.

[7] I. R. Nourbakhsh, D. Andre, C. Tomasi and M. Genesereth. Obstacle avoidance via depth from focus. *Image Understanding Workshop*, pp. 1339-1344, 1996.

[8] C. Pégard and E. M. Mouaddib. A mobile robot using a panoramic view. *Proc. IEEE Int. Conf. Robotics Automat.*, pp. 89-94, 1996.

[9] Y. Yagi, S. Kawato and S. Tsuji. Real-time omnidirectional image sensor (COPIS) for vision-guided navigation. *IEEE Journal Robotics Automat.*, pp. 11-22, feb 1994.

[10] J. Y. Zheng and S. Tsuji. Panoramic representation for route recognition by mobile robot. *International Journal of Computer Vision*, pp. 55-76, 1992.

Figure 13: Perspective view of a chess piece.

Figure 14: Estimated frontal view of a chess piece.

Figure 15: Estimated profile view of a chess piece.