PANORAMIC IMAGE ACQUISITION*

Arun Krishnan†

Imaging and Visualization Department
Siemens Corporate Research
755 College Road East
Princeton, NJ 08540, U.S.A
e-mail: arunki@scr.siemens.com

Narendra Ahuja

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

Abstract

This paper is concerned with acquiring panoramic focused images using a small field of view video camera. When scene points are distributed over a range of distances from the sensor, obtaining a focused composite image involves focus computations and mechanically changing some sensor parameters (translation of sensor plane, panning of camera etc.) which can be time intensive. In this paper we present methods to optimize the image acquisition strategy in order to reduce redundancy. We show that panning a camera about a point f (focal length) in front of the camera eliminates redundancy. The Non-frontal imaging camera (NICAM) with tilted sensor plane has been previously introduced [5] as a sensor that can acquire focused panoramic images. In this paper we also describe strategies for optimal selection of panning angle increments and sensor plane tilt for NICAM. Experimental results are presented for panoramic image acquisition using a regular camera as well as using NICAM.

1 Previous work and introduction

Panoramic images, especially focused images, have many applications in surveillance, robot navigation, art, etc. There have been different panoramic image acquisition methods reported in the past. For satellite imagery, frames acquired with some overlap are registered using the overlapped regions [2, 8]. For short-range imagery, where small viewpoint changes cause changes in the projection parameters, scene points from individual frames are projected on to a common coordinate system and then re-sampled with interpolation to create a regularly sampled panoramic image [4, 9]. Instead of acquiring multiple images and then combining them, Tsuji et. al. obtain a panoramic image by using a panning slit camera [3, 11]. A vertical slit camera is moved in steps of 0.4 degrees and the panoramic image is created by pasting the slit views together. Yagi et. al. use a conical mirror to compress a 360 degree view of the scene into the view of a regular camera [10]. The authors state that the images obtained by the camera have low resolution which makes it impractical for detailed analysis.

A method to handle wide scenes using a regular video camera is to process the scene a part at a time by changing viewpoints and directions. In a single configuration, the camera can acquire a focused view over a given visual angle and over a given depth range (referred to as the SF surface henceforth). Other scene parts that are within the field of view will appear defocused. As the sensor parameters change, the SF surface will move and sweep out a volume in 3D space (henceforth called the SF cone) as shown in Figure 1. Thus to process the entire breath, height, and depth of the scene, the camera orientation and focus settings must be changed. This approach has the advantages that it gives up to 360 degree field of view with uniform resolution, and reduces aberrations. But the disadvantages are an increase in acquisition time and post-processing complexity.

The objective of the work presented in this paper is to improve the performance of this approach with respect to acquisition time and to some extent processing complexity. Methods are described that require minimal number of parameter changes to process the panoramic scene. For frontal cameras with sensor plane perpendicular to the optic axis, these parameters are the location of the pan axis and changes in the focus parameter (distance between sensor plane and lens center). For non-frontal cameras, with sensor planes at non-perpendicular angles to the optic axis, these parameters are the sensor plane tilt angle and location, and the pan angle increments. Sections 2 and 3 deal with frontal and non-frontal cameras respectively. Sections 4 and 5 present the results and conclusions respectively.

2 Focused panoramic image acquisition using frontal cameras

In this section we shall consider using regular frontal cameras (standard cameras with sensor plane perpendicular to the optic axis) to create a composite

*This work was supported by Advanced Reasearch Projects Agency grant N00014-93-1-1167 administered by the Office of Naval Research.

†This paper reports research done while the author was at the University of Illinois

focused image of the scene. The scenario is to pan the camera with a fixed pan axis to image a stripe of the scene. The camera will be panned in regular angle increments. At each pan position, the sensor plane will be translated to obtain a sequence of images with different focus settings. A fully focused composite image will be created for the visible part of the scene. After the camera finishes a 360 degree revolution, the camera will be tilted and the entire process of panning will repeat to image an adjacent stripe. Individual fully focused stripe images will finally be merged to create the composite fully focused image for the entire scene.

2.1 Panning about lens center

Since the camera must be panned to view different directions, a pan axis must be chosen. The orientation of the pan axis must be varied to scan the entire scene (4Π solid angle about a point). However the location of the pan axis must be chosen carefully. In this section we consider the most obvious choice of the axis location, namely, the axis passing through the front nodal point for the lens.

Consider the SF cone of a frontal camera which has a sensor plane of length $2l$ units. Let the sensor surface translate from a distance of v_A from the lens center, to v_B thereby causing the SF surface to move from u_A to u_B. If the focal length is f, then we have

$$v_A = \frac{fu_A}{u_A - f} \quad \& \quad v_B = \frac{fu_B}{u_B - f}$$

If the swept volume is approximated by a trapezium as shown in Figure 1, then its dimensions are: height $= (u_B - u_A)$; smaller side $= \frac{2lu_A}{v_A}$; and larger side $= \frac{2lu_B}{v_B}$

When the sensor surface is at position v_A, the angle subtended by the SF surface at the lens center is $\theta_1 = 2\arctan\left[\frac{l}{v_A}\right]$ and at position v_B is $\theta_2 = 2\arctan\left[\frac{l}{v_B}\right]$

Since $\theta_2 \geq \theta_1$, to cover the entire angle of 2π we will need to take image sets at $\left\lceil \frac{2\pi}{2\arctan\left[\frac{l}{v_A}\right]} \right\rceil$ pan angles. The SF cones for adjoining camera pan values will have overlapping regions, *i.e.* regions which are processed twice.

2.2 Optimal pan axis location

The two inclined sides of the trapezium shown in Figure 1 meet at a point that is at a distance of f in front of the lens center, on the optical axis. At this point, the two parallel sides of the SF cone subtend the same angle of $2 \times \arctan[l/f]$ radians. If the camera is panned about this point as shown in Figure 2, then neighboring SF cones (for successive pan values) would be adjacent, without overlap. So the number of different pan angles required to completely capture a view of 2π radians would be $\left\lceil \frac{\pi}{\arctan[l/f]} \right\rceil$. This would give an optimal packing of the scene with SF cones. Further, it would allow the use of the full extent of the sensor plane for acquiring each image. The following procedure employs such panning and is optimal in the number of pan angles and in the usage of the entire sensor plane.

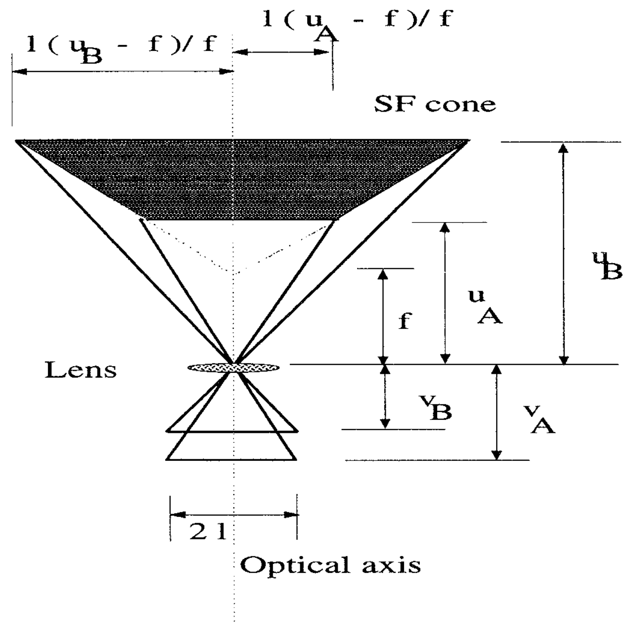

Figure 1: A cross-section of the cone swept by the SF surface as the value of v is changed. Only those points that lie inside the SF cone can be imaged sharply.

Optimal procedure

Step 1 Change v from v_A to v_B and obtain a fully focused image using the procedure described in Section 2.3. For each v, all pixels are used in determining the fully focused image.

Step 2 Pan the camera by an angle of $2 \times \arctan[l/f]$ radians about a point f in front of the lens center and return to Step 1 until the entire scene of interest has been imaged.

2.3 Optimal focus setting variation

For each pan position, the sensor plane needs to translate between two extremes that depend on the distances to the closest scene point and farthest scene point. We shall use the following three criteria to enable optimal movement of the sensor surface

- No scene point is ever outside the DOF of a SF surface.
- DOF at each position of the SF surface should be as large as possible.
- Neighboring DOFs do not overlap.

2.3.1 Fully focused image generation

The exact relationship between the DOF and other variables is described by Equation (3).

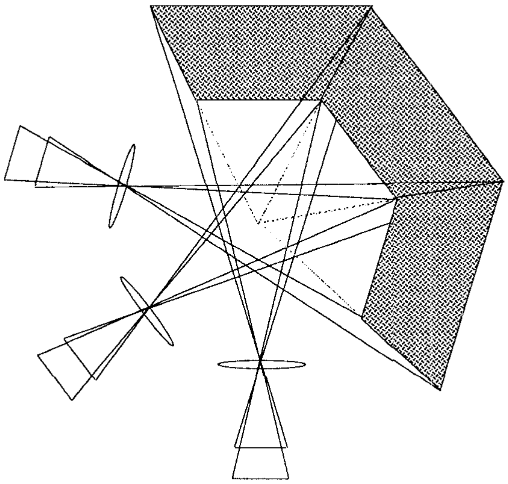

Figure 2: Panning a camera about a point f units along the optic axis from the lens center. For each pan angle, the sensor plane is translated from v_A to v_B to create a SF cone. The SF cones are optimally packed for this choice of pan axis location.

$$u1 = \frac{u0Af}{Af + C(u0 - f)} \quad (1)$$

$$u2 = \frac{u0Af}{Af - C(u0 - f)} \quad (2)$$

$$DOF = u2 - u1 = \frac{2ACu0(u0 - f)}{A^2 f^2 - C^2(u0 - f)^2} \quad (3)$$

where, $u0$ is the object distance about which the DOF is located, $u1$ is the near extreme of the DOF, $u2$ is the far extreme of the DOF, f is the focal length of the lens, A is the radius of the lens aperture, and C is the radius of the circle of confusion.

From Equation (3) we see that for a given value of $u0$, the DOF is a function of f and A. Assuming that C and f remain constant, then the smallest value of A maximizes the DOF[1]. Changing the value of A changes the brightness of the image and so the chosen value of A may not be the mechanical minimum.

Let u_{near} and u_{far} be the desired near and far ends of the scene. The sensor plane distance is changed in every iteration of the algorithm such that every scene point is within the depth of field region around only one SF surface.

Algorithm A

1. Let k = 1, $u1_1 = u_{near}$ which implies that

$$u0_1 = \frac{fu_{near}(C - A)}{Cu_{near} - Af}$$

2. Acquire and analyze the image. Update the focus map for scene points that have a peak in the focus criterion function[2].

3. Determine $u2_k$ using Equation (2) and also calculate the new value of $u0_{k+1}$ using the following formula $u1_{k+1} = u2_k$ which yields

$$u0_{k+1} = \frac{fu0_k(C - A)}{2Cu0_k - f(C + A)}$$

4. If $u2_k > u_{far}$, exit (all points have been imaged in sharp focus), otherwise move sensor plane such that $u = u0_{k+1}$, set k = k + 1 and continue from Step 2

The above procedure is used to translate the sensor plane to view the scene points from near to far (by moving the sensor plane away from the lens). The camera is then panned and the procedure repeated, but with the scene points in focus from far to near (by moving the sensor plane towards the lens).

3 Focused panoramic image acquisition using non-frontal cameras

Panoramic focused image acquisition using a non-frontal imaging camera (NICAM) was first introduced by Krishnan et. al in [6]. To summarize, the sensor plane of a non-frontal camera is at an non-perpendicular angle to the optical axis. The SF surface of the NICAM will also be non-perpendicular to the optical axis as given by *Scheimpflug's condition* [1]. Consider the image of a scene point as the camera pans. Initially the point's image will appear defocused on the sensor plane. As the camera pans, the distance between the lens center and the point's image will change and there will be one particular camera pan angle at which the scene point will image in sharp focus. As the camera pan angle further increases, the scene point will go out of focus again. Thus panning the NICAM once is all that is required to obtain a sharp focused panoramic image of the scene.

The image acquisition protocol using the NICAM requires three parameters: the sensor plane tilt, the sensor plane location, and the pan angle increment. Determining these parameters requires expression for the depth of field which is given in Section 3.1. Section 3.2 gives constraints for the optimal selection for the parameters.

3.1 Depth of field for NICAM

The depth of field for the NICAM varies as a function of the sensor plane tilt (α), the distance between the sensor plane and the lens as measured along the optical axis (d), the position of the image point on the sensor plane (x, y), the aperture of the lens (A) and the circle of confusion radius (C).

[1] Changing the focal length can also change the DOF, but in most lens systems, changes in f automatically changes the focus distance too.

[2] The choice of criterion function does not affect the algorithms presented in this paper.

A simplified expression[3] for the depth of field [7] is given by,

$$\begin{aligned} v0(d,\alpha,x) &= d - x\sin(\alpha) \\ \beta_{max}(d,\alpha,x) &= \frac{A\,[v0 - c\,\sin(\alpha)]}{A\,v0 - c\,d\,\cos(\alpha)} \\ \beta_{min}(v0,\alpha,x) &= \frac{A\,[v0 + c\,\sin(\alpha)]}{A\,v0 + c\,d\,\cos(\alpha)} \end{aligned}$$

$$v1(d,\alpha,x) = v0\,\beta_{max} \quad \text{and} \quad u1(d,\alpha,x) = \frac{f\ v1}{v1 - f}$$

$$v2(d,\alpha,x) = v0\,\beta_{min} \quad \text{and} \quad u2(d,\alpha,x) = \frac{f\ v2}{v2 - f}$$

$$\text{Depth of field } DOF(d,\alpha,x) = f^2 \frac{[\beta_{max} - \beta_{min}]\ v0}{(\beta_{max}\,v0 - f)\,(\beta_{min}\,v0 - f)}$$

Using the same terminology used in Section 2.3.1, $u1$ is the near extent of the DOF and $u2$ is the far extent of the DOF. Note that u and v are measured parallel to the optical axis. The radial distance from the lens center to a scene point whose image appears at location coordinate x on the sensor plane is given by

$$r = u\frac{\sqrt{d^2 + x^2 - 2\,d\,x\,\sin(\alpha)}}{d - x\,\sin(\alpha)} \qquad (4)$$

3.2 Sensor plane tilt and pan increment

Objects that lie within the SF cone of the panning NICAM will be imaged in at least one pan position with maximum focus criterion value. In the ideal case, we want the SF cone to completely sweep out every scene point between r_{min}, the nearest scene point, and r_{max}, the farthest point as shown in Figure 3. Making very small pan angle increments will do that, but at the expense of increased and redundant processing. We can use the fact that the SF surface is actually surrounded by the DOF and can therefore increase the pan angle increment. The larger the pan angle increment, the fewer image acquisitions and computations need to be done. Thus for large pan increments, DOF needs to be large.

One of the variables that affects DOF is the sensor plane tilt. Increasing the sensor plane tilt increases the radial extent of the SF surface. This increases the SF cone swept volume, thus allowing more objects to image in sharp focus during camera pan. But increasing the sensor plane tilt decreases both the field of view of the camera and the DOF. Intuitively, the resolution of the sensor becomes finer as the tilt increases because there are more pixel elements per unit view angle and this decreases the depth of field.

Let the field of view[4] of the sensor be ρ. Let θ be an angular variable that goes from 0 to ρ as x goes from $-L$ to L. x can be written as a function of θ, α and d. Let the pan angle increment be δ.

The following constraints should be satisfied

1. The SF surface including the DOF should span the range from r_{min} to r_{max}.

$$r1(d,\alpha,\theta=0) \le r_{min} < r_{max} \le r2(d,\alpha,\theta=\rho)$$

where $r1$ and $r2$ are the radial distances that correspond to $u1$ and $u2$ respectively using Equation 4.

2. Neighboring SF surfaces including DOF should not have gaps between them. And all scene points between r_{min} and r_{max} should be in at least one SF surface region as illustrated in Figure 4. That is,

$$r2(d,\alpha,theta-\delta) \ge r1(d,\alpha,\theta)$$

for all $\delta \le \theta \le \rho$ and,

$$r2(d,\alpha,\rho-\delta) \ge r_{max}$$

and

$$r_{min} \ge r1(d,\alpha,\delta)$$

The above constraints are not easily amenable to symbolic solutions. The following optimization procedure can be used:

- Step1: Constraint 1 can be exactly solved for α and d given values for $r1(d,\alpha,x=-L)$ and $r2(d,\alpha,x=L)$, say r_A and r_B. Of course, $r_A \le r_{min}$ and $r_{max} \le r_B$
- Step2: Determine the maximum δ value that satisfies Constraint 2 using α and d solutions from Step 1.
- Step 3: Repeat Steps 1 and 2 for different choices of r_A and r_B and use the set of parameters that gives the global maximum for δ.

4 Results

Usual focus control mechanisms in cameras are attached to the lens and work by shifting the lens system. This causes the view point to shift as the camera focuses. The algorithms described in Section 2 require the movement of the sensor plane without moving the lens center. As in the experiments performed, we did not have controllable sensor plane translation, we are unable to present experimental verification of the presented algorithm for fully focused panoramic image acquisition. Instead, we present results for panoramic images with the sensor plane fixed at an intermediate focus setting. This causes parts of the scene to appear defocused.

Figures 5 (a)-(e) and (f)-(i) show 8 consecutive images obtained by panning a camera about a point f in front of the lens center. This covers an angle of approximately 130 degrees.

Figures 5 (e) and (j) show the same scene imaged by NICAM. The aperture and scene brightness were kept the same. All parts of the scene are in focus.

[3] The variation of DOF due to y has been ignored as that involves solving equations of degree 4

[4] ρ is a function of d, the sensor plane tilt α, and the extent of the sensor plane $2l$ and is easily calculated.

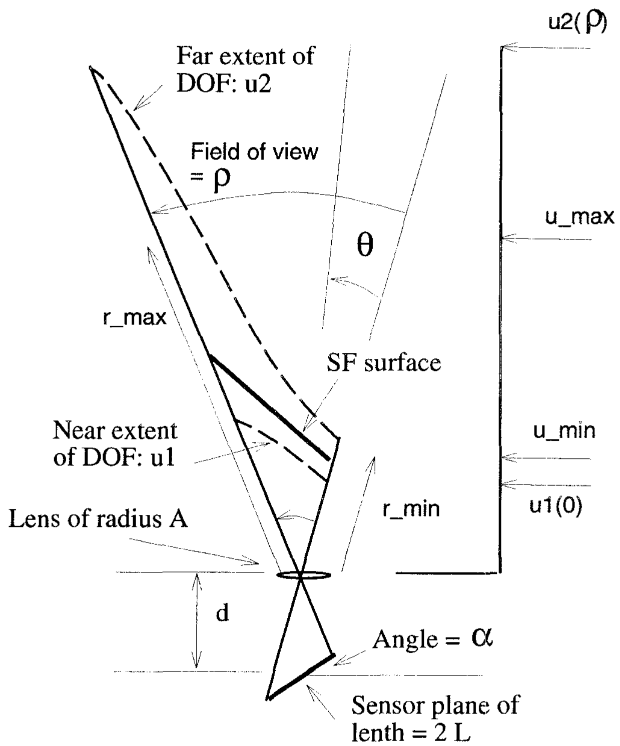

Figure 3: NICAM SF surface with DOF. All scene points between the near extent of DOF (u1) and the far extent of DOF (u2) will appear focused on the sensor plane. The SF surface is shown as a thick line. The field of view is ρ and θ is a variable that goes from 0 to ρ.

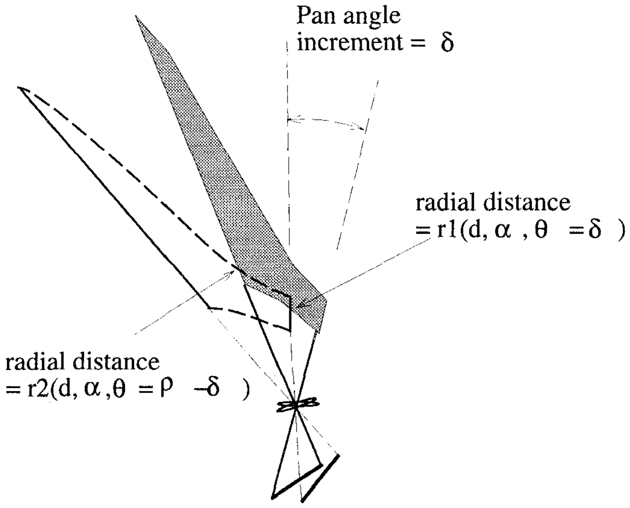

Figure 4: NICAM SF surfaces augmented by the DOF for two consecutive pan positions. The pan angle increment is δ.

5 Conclusions

Optimal control of camera parameters to acquire and process images of a large scene has been discussed. Panning the camera about a point f in front of the lens center, prevents overlap between successive pan positions and thus prevents redundant computation. To create fully focused panoramic images requires individual fully focused images at every pan angle. An optimal focus varying method has been presented that minimizes the number of focus settings used to obtain fully focused images for static scenes.

Methods to determine the optimal parameters for the NICAM were described. Finally, results for panoramic scene acquisition using a frontal camera (without sensor plane adjustment) and results for panoramic scene acquisition using a non-frontal imaging camera were given.

For scenes with bright lighting conditions, the lens aperture can be made small enough to increase the DOF to near infinity. A frontal camera that pans about its lens center would be the fastest and easiest way. If infinite DOF is not possible, then panning frontal camera about a point f in front of the lens center, or a panning NICAM are the choices to obtain a focused panoramic image. Of the two, a panning NICAM would be the prefered choice as it needs only one mechanical motion (the panning motion).

Acknowledgments

We would like to acknowledge the help of Andres Castano in obtaining the experimental results.

References

[1] Michael Bass, editor. *Handbook of optics*, volume vol I. McGraw-Hill, 1995.

[2] P. J. Burt and E. H. Adelson. A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics*, 2(4):217–236, October 1983.

[3] H. Ishiguro, M. Yamamoto, and S. Tsuji. Omnidirectional stereo. *IEEE Trans. Patt. Anal. Mach. Intell.*, 14:2, 1992.

[4] P. Jaillon and A. Montanvert. Image mosaicking applied to three-dimensional surfaces. In *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, pages A253–257, October 1994.

[5] Arun Krishnan and Narendra Ahuja. Range estimation from focus using a non-frontal imaging camera. In *Proceedings of the DARPA Image Understanding Workshop*, pages 959–965, Washington D.C., April 1993.

[6] Arun Krishnan and Narendra Ahuja. Use of a non-frontal camera for extended depth of field in wide scenes. In *Proceedings of the SPIE Conference on Intelligent Robots and Computer Vision XII: Active Vision and 3D Methods*, pages 62–72, Boston MA, September 1993.

[7] Arun Krishnan and Narendra Ahuja. Depth of field for tilted sensor plane. Technical Report UIUC-BI-AI-RCV-94-08, Beckman Institute, University of Illinois, 1994.

[8] D. L. Milgram. Adaptive techniques for photo-mosaicking. *IEEE Transactions on computers*, C-26:1175–1180, 1977.

[9] R. Szeliski. Image mosaicing for tele-reality applications. Technical Report CRL 94/2, Digital Equipment Corporation CRL, 1994.

[10] Y. Yagi, S. Kawato, and S. Tsuji. Real-time omnidirectional image sensor (copis) for vision-guided navigation. *IEEE Transactions on Robotics and Automation*, 10:11–22, 1994.

[11] J. Y. Zheng and S. Tsuji. Panoramic representation for route recognition by a mobile robot. *Int. J. of Comp. Vision*, 9:1:55–76, 1992.

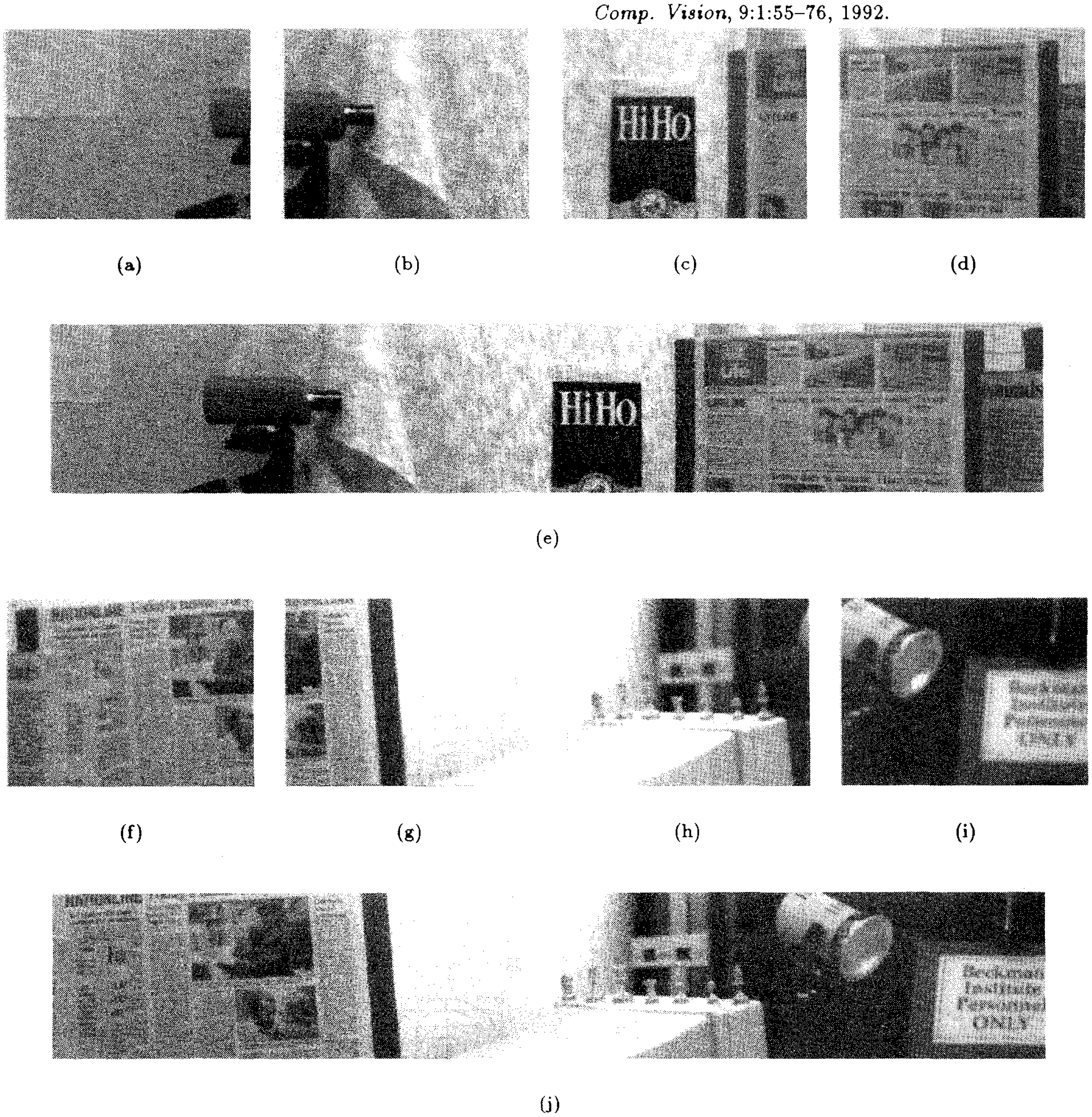

Figure 5: (a)-(d) and (f)-(i) Images taken by panning a frontal camera about a point f in front of the lens center. (e) and (j) are images obtained from NICAM