Nonmetric Calibration of Wide-Angle Lenses and Polycameras

Rahul Swaminathan and
Shree K. Nayar, *Senior Member*, *IEEE*

Abstract—Images taken with wide-angle cameras tend to have severe distortions which pull points towards the optical center. This paper proposes a simple method for recovering the distortion parameters without the use of any calibration objects. Since distortions cause straight lines in the scene to appear as curves in the image, our algorithm seeks to find the distortion parameters that map the image curves to straight lines. The user selects a small set of points along the image curves. Recovery of the distortion parameters is formulated as the minimization of an objective function which is designed to explicitly account for noise in the selected image points. Experimental results are presented for synthetic data as well as real images. We also present the idea of a polycamera which is defined as a tightly packed camera cluster. Possible configurations are proposed to capture very large fields of view. Such camera clusters tend to have a nonsingle viewpoint. We therefore provide analysis of what we call the minimum working distance for such clusters. Finally, we present results for a polycamera consisting of four wide-angle sensors having a minimum working distance of about 4m. On undistorting the acquired images using our proposed technique, we create real-time high resolution panoramas.

Index Terms—Camera calibration, wide-angle lens, radial distortion, decentering distortion, camera clusters, polycamera, nonsingle viewpoint, minimum working distance, real-time panoramic sensor.

1 Introduction

In many vision applications, such as surveillance, it is desirable to capture the entire region of interest with as few cameras as possible. Wide-angle cameras help in this regard, but at the cost of severe image distortions. Fig. 1 clearly illustrates the effects of such image distortions. Wide-angle lenses that adhere to perspective projection would necessitate the use of prohibitively large image detectors. To work around this problem, the lenses are designed to severely bend rays of light around the periphery of the field of view, thus permitting the use of a small image detector. Severe bending of light rays typically leads to a nonsingular entrance pupil. The resulting locus of pupils in three dimensions is called a *diacaustic* [2]. Hence, for a wide-angle lens, complete removal of distortions cannot in general be achieved. For our purposes, we will assume a small pupil locus that can be approximated by a single point.

If the optics of a wide-angle camera system are known a priori (i.e., the distortion parameters), then distortion correction can be easily applied. Unfortunately, such information is seldom available. Therefore, a simple calibration method to estimate the distortion parameters is desirable.

Several calibration techniques have been suggested for recovering lens distortion parameters. Tsai [17] used known points in 3D space to recover some of the distortion parameters. Goshtasby [6] utilized Bezier patches to model the distortions and used a uniform grid placed in front of the camera as a calibration object. Weng et al. [19] also used calibration objects to extract the distortion parameters. All these methods fall in the category of "stellar" calibration, where calibration objects of known dimensions need to be used.

In contrast, Brown [4] proposed a "nonmetric" approach that does not rely on known scene points. Instead, it relies on the fact that straight scene lines must perspectively project to straight lines in the image. Brown's algorithm uses essentially noiseless image data, obtained by imaging plumb-lines suspended against a black background onto a photographic plate. More recently, Kang [7] suggested using radial distortion snakes to estimate the radial distortion parameters. Becker and Bove [1] used three mutually orthogonal sets of parallel lines and a vanishing point constraint to recover distortion parameters. Stein in [15], [13], and [14] uses point correspondences from multiple views to estimate radial distortions as well as the extrinsic parameters. This additional parameter search space makes the problem unstable in the presence of noise. Sawhney and Kumar [11], [12] suggested a novel approach of image-based distortion parameter estimation. This is a direct method and relies solely on multiimage alignment.

Previous work suffers from one or more of the following restrictions: Calibration objects need to be used, not all the distortion parameters are recovered, or the algorithm is highly sensitive to noise. One exception is the work of Becker and Bove [1]. However, Becker and Bove's constraint (triplets of orthogonal lines) is less abundant in urban settings than the randomly oriented straight lines we use. We formulate the estimation of distortion parameters as the minimization of a noise insensitive objective function via efficient search. Experimental results with simulated as well as real data are presented.

In addition, we present the notion of a polycamera (tight cluster of cameras) to capture very large fully connected fields of view. Wide-angle cameras are useful in this context as they minimize the number of cameras needed to cover the desired field of view. Fewer cameras also facilitate tightly packed clusters, which aid in reducing the effects of a nonsingular viewpoint such as parallax. In general, it is not possible to maintain a single viewpoint when using camera clusters, one exception being the system developed by Nalwa (see [9]), which uses mirrors along with image sensors to obtain a single viewpoint. Nonsingular viewpoints lead to parallax between views which, although useful in stereo applications, creates unwanted visual artifacts when merging multiple views together. We propose the idea of a *minimum working distance* beyond which the parallax effects are negligible. Analysis of the minimum working distance for the generic case is provided. Finally, we present a polycamera system we developed based on our designs. For this system, we present its minimum working distance map as well as the high resolution panoramic video computed in real-time.

2 Distortion Model

Distortions in lenses can be decomposed into three components: 1) shift of the optical center, 2) radial distortion, and 3) decentering distortion. In the discussion to follow, we assume the perspective projection of a scene point in the image plane is point $\mathbf{q}'$. Due to distortions in the lens, $\mathbf{q}'$ gets mapped to $\mathbf{q}$ (see Fig. 2). Let (x, y) be the Cartesian and (r, ϕ) be the polar coordinates of $\mathbf{q}$. Similarly, let (x', y') be the Cartesian and (r', ϕ') be the polar coordinates of $\mathbf{q}'$. Also, let the optical center C be located at (x_p, y_p).

2.1 Shift of Optical Center

A shift of the optical center corresponds to a shift of the image detector in the imaging plane. Estimating this distortion component amounts to estimating $C = (x_p, y_p)$, the optical center. The coordinates of an image point upon correction is given by:

$$\bar{x} = x - x_p \ , \ \ \bar{y} = y - y_p.$$

2.2 Radial Distortions

Radial distortions distort image points only along the radial direction. Most wide-angle cameras tend to pull points radially toward the optical center. This is referred to as *barrel distortion* [2].

- *The authors are with the Department of Computer Science, Columbia University, New York, NY 10027.*
E-mail: {srahul, nayar}@cs.columbia.edu.

Manuscript received 21 Apr. 1999; revised 12 June 2000; accepted 25 July 2000.
Recommended for acceptance by R. Collins.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 109659.

0162-8828/00/$10.00 © 2000 IEEE

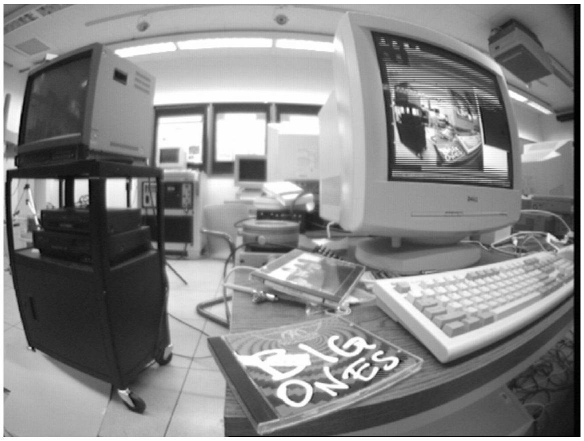


Fig. 1. Images captured with wide-angle cameras have severe distortions that alter the appearances of objects in the scene.

Another type of radial distortion tends to push points away from the optical center along the radial direction and is called *pincushion distortion*. The radial distortion at the point $\mathbf{q}$ is modeled as:

$$\Delta r(\mathbf{q}) = \sum_{i=1}^{\infty} C_{2i+1} r^{2i+1} \text{ where } r = \sqrt{\bar{x}^2 + \bar{y}^2},\ \tan(\phi) = \frac{\bar{y}}{\bar{x}}. \quad (1)$$

C_{2i+1} are the distortion parameters. We ignore terms higher than the fifth-order as their contribution to the distortion is generally negligible [3]. Hence, we have:

$$\Delta r(\mathbf{q}) \quad \approx \quad C_3 r^3 + C_5 r^5. \quad (2)$$

2.3 Decentering Distortions

Decentering distortions are caused by nonorthogonality of the lens components and the image detector with respect to the optical axis. All imaging systems have some degree of decentering distortions. Unlike radial distortions, this component acts tangential to the radial direction. We use Conrady's model [5] for decentering distortion:

$$\begin{aligned}
&\Delta T_x(\mathbf{q}) = \\
&\left[P_1 r^2\left(1 + 2\cos^2(\phi)\right) + 2P_2 r^2 \sin(\phi)\cos(\phi)\right] \cdot [1 + \sum_{i=1}^{\infty} P_{i+2} r^{2i}] \\
&\Delta T_y(\mathbf{q}) = \\
&\left[P_2 r^2\left(1 + 2\sin^2(\phi)\right) + 2P_1 r^2 \sin(\phi)\cos(\phi)\right] \cdot [1 + \sum_{i=1}^{\infty} P_{i+2} r^{2i}],
\end{aligned} \quad (3)$$

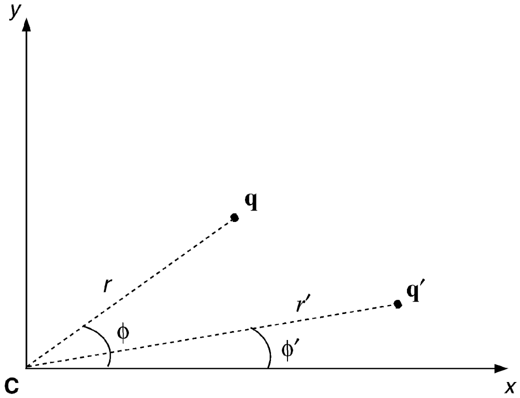


Fig. 2. $\mathbf{q}'$ is the perspective projection of a scene point onto the image plane $x - y$. Due to radial and decentering distortions $\mathbf{q}'$ gets mapped to the point $\mathbf{q}$.

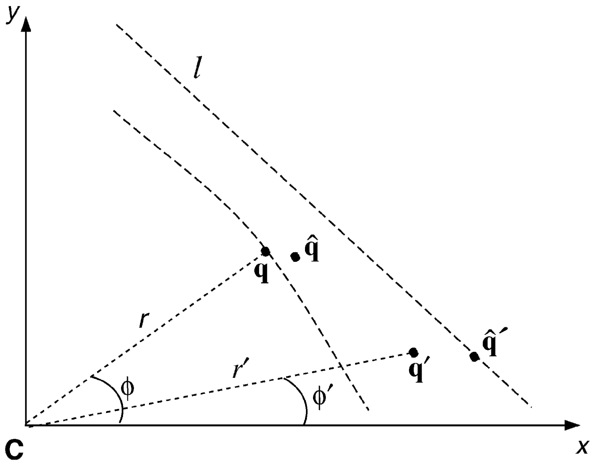


Fig. 3. $\mathbf{q}$ is a point selected by the user and $\mathbf{q}'$ is its undistorted location on applying the (hypothesized) distortion parameters $\mathcal{S}$. l is the "best-fit" line estimated for all $\mathbf{q}'$ which lie on the same scene line. $\hat{\mathbf{q}}$ is a point close to $\mathbf{q}$ such that its undistorted location $\hat{\mathbf{q}}'$ (obtained by applying $\mathcal{S}$ to $\hat{\mathbf{q}}$) lies on l. We wish to minimize the distance between $\mathbf{q}$ and $\hat{\mathbf{q}}$.

where, P_1, P_2, P_{i+2} are the distortion parameters and $\Delta T_x, \Delta T_y$ are the distortions along the x and y directions, respectively. The higher-order terms are again relatively insignificant [3] and, in practice, the model may approximated as:

$$\begin{aligned}
\Delta T_x(\mathbf{q}) &\approx \left[P_1 r^2\left(1 + 2\cos^2(\phi)\right) + 2P_2 r^2 \sin(\phi)\cos(\phi)\right] \\
\Delta T_y(\mathbf{q}) &\approx \left[P_2 r^2\left(1 + 2\sin^2(\phi)\right) + 2P_1 r^2 \sin(\phi)\cos(\phi)\right].
\end{aligned} \quad (4)$$

2.4 Complete Distortion Model

In order to correct for distortions, we need to recover the parameters: $\mathcal{S} = \{C_3, C_5, P_1, P_2, x_p, y_p\}$. The total distortion is modeled as:

$$\begin{aligned}
\Delta x(\mathbf{q}) &\approx \cos(\phi)(\Delta r(\mathbf{q})) + \Delta T_x(\mathbf{q}) \\
\Delta y(\mathbf{q}) &\approx \sin(\phi)(\Delta r(\mathbf{q})) + \Delta T_y(\mathbf{q}).
\end{aligned} \quad (5)$$

3 Objective Function Formulation

Under perspective projection, straight lines in the scene must map to straight lines in the image. Lens distortions cause points along straight scene lines to be mapped to curves in the image. However, only nonradial lines get distorted if decentering distortions are not considered. In this setting, an objective function can be defined which, when minimized, yields the parameters that undistort the curve points to lie on straight lines. The user selects points in the image along the distorted curves which correspond to straight scene lines. These points are used to estimate the distortion parameters.

We present three objective functions, namely, sum of squared distances (from straight lines), normalized sum of squared distances, and one that explicitly estimates noise in the chosen image points. The first two are presented mainly to demonstrate that simple objective functions (similar to ones proposed previously) are highly noise sensitive. In contrast, the third function is designed to explicitly account for noise in the selected points. All our objective functions are minimized using efficient search algorithms. In what follows, our goal will be to recover only the radial and decentering distortion parameters. The shift of the optical center is recovered separately in an iterative fashion.

3.1 Sum of Squared Distances (ξ_1)

Let $\{\mathbf{q} = (x, y)\}$ denote the set of points selected by the user. During search, a set of (hypothesized) distortion parameters $\mathcal{S} = \{C_3, C_5, P_1, P_2\}$ are applied to these selected image points. This

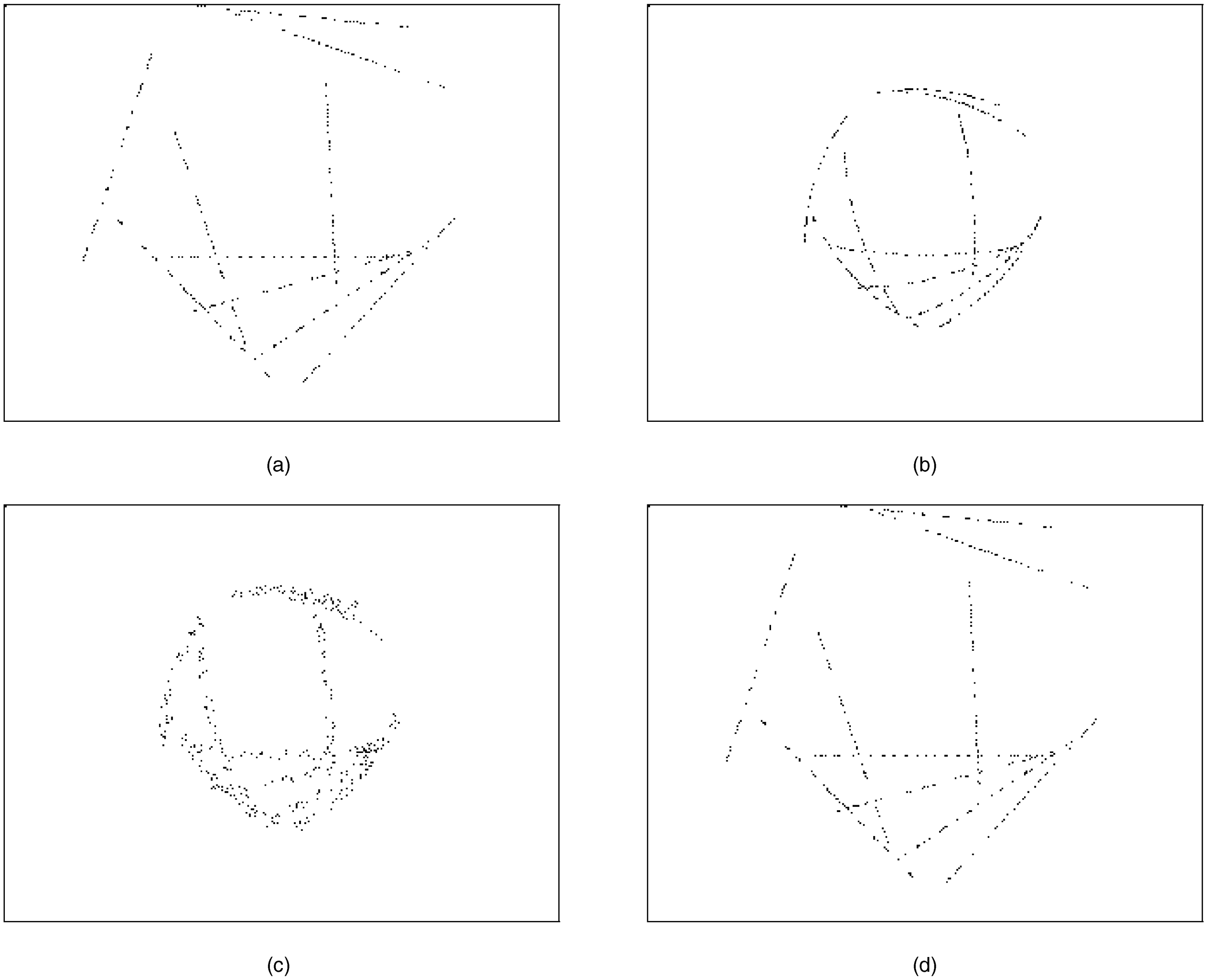


Fig. 4. (a) Points randomly sampled from synthetically generated lines. (b) Known distortions are applied to the points in (a). (c) Uniformly distributed random noise in the interval (-5 pixels, 5 pixels) is added to the distorted points in (b). (d) The distortion parameters are recovered from these noisy image points using the algorithm based on objective function ξ_3. These parameters are used to undo the distortions present in (b). Despite the large amount of noise, the recovery of undistorted image points is found to be accurate and robust.

maps them to supposed undistorted points which we hope will be collinear. Lines are fit to these resulting sets of points $\{\mathbf{q}' = (x', y')\}$ using a least-squares approach. The objective function is then defined as the sum of the squared distances of the points from their corresponding "best-fit" lines. Let the best-fit line for a set of points $\{\mathbf{q}'\}$ be parameterized by (θ, ρ), where θ is the angle the line makes with the horizontal axis and ρ is the distance of the line from the optical center. The error due to a single point $\mathbf{q}$ is then defined as:

$$\begin{aligned} e_1 &= \big(x' \sin(\theta) - y' \cos(\theta) + \rho\big)^2, \text{where :} \\ x' &= x + \Delta x(\mathbf{q}), y' = y + \Delta y(\mathbf{q}). \end{aligned} \tag{6}$$

Let the number of curves selected by the user be L and the number of points on each line l be P_l. Then, the objective function is given by:

$$\xi_1 = \sum_{l=1}^{L} \sum_{p=1}^{P_l} \Big(x'_{p,l} \sin(\theta_l) - y'_{p,l} \cos(\theta_l) + \rho_l \Big)^2, \tag{7}$$

where θ_l and ρ_l are the best-fit line parameters corresponding to image curve l and $(x_{p,l}, y_{p,l})$ is the pth point on line l.

3.2 Normalized Sum of Squares (ξ_2)

Although simple, the above function is very sensitive to noise. This is because noise is magnified by the higher order terms in the model. Thus, points closer to the optical center contribute less to the error metric than points farther away. This effect is partially remedied by normalizing the error e_1 in (6) by the square of the distance ρ_l of the corresponding line l from the optical center. The modified objective function then is:

$$\xi_2 = \sum_{l=1}^{L} \sum_{p=1}^{P_l} \left(\frac{x'_{p,l} \sin(\theta_l) - y'_{p,l} \cos(\theta_l) + \rho_l}{\rho_l} \right)^2. \tag{8}$$

3.3 Explicit Noise Estimation (ξ_3)

The objective functions ξ_1 and ξ_2 are defined in the space of the undistorted points (i.e., after applying $\mathcal{S}$). In this space, the distortion model nonlinearly magnifies noise. It is therefore more appropriate to formulate an objective function in the space of distorted image points. As shown in Fig. 3, let q be the distorted point under consideration and $\mathbf{q}'$ be the "undistorted" point obtained by applying the set of distortion parameters $\mathcal{S}$. Again, as before, l is the best-fit line for the points $\{\mathbf{q}'\}$ which are believed to lie on the same scene line. We now determine (via search) the point $\hat{\mathbf{q}}$ close to q which, when undistorted using $\mathcal{S}$, would lie on l at $\hat{\mathbf{q}}'$. The new error function then is:

$$e_3 = \|\mathbf{q} - \hat{\mathbf{q}}\|^2. \tag{9}$$

Since $\hat{\mathbf{q}}'(\hat{x}', \hat{y}')$ must lie on l, it must satisfy the constraint:

$$\hat{x}' \sin(\theta) - \hat{y}' \cos(\theta) + \rho = 0, \text{where} : \hat{x}' = \hat{x} + \Delta x(\hat{\mathbf{q}}), \hat{y}' = \hat{y} + \Delta y(\hat{\mathbf{q}}). \tag{10}$$

Considering all data points, the objective function is given by:

$$\xi_3 = \sum_{l=1}^{L} \sum_{p=1}^{P_l} \|\mathbf{q}_{p,l} - \hat{\mathbf{q}}_{p,l}\|^2. \tag{11}$$

This objective function is much more resilient to noise. Experimental results included in the following sections illustrate this fact.

4 Minimization of ξ_1, ξ_2, and ξ_3

We use a constrained minimization algorithm (modified simplex) outlined in [10], implemented in the IMSL [18] library to minimize the three objective functions. In general, any search algorithm can be used in its place. The following bounds were used for the search parameters: $C_3 \in (-10^{-5}, 10^{-5})$, $C_5 \in (-10^{-9}, 10^{-9})$, $P_1 \in (-10^{-5}, 10^{-5})$, $P_2 \in (-10^{-5}, 10^{-5})$. These bounds are highly conservative as they include distortions far stronger that those found in wide-angle imaging sensors. The search also requires an initial starting point. We assumed the system to be distortion-free at the start of the search, i.e., $C_3 = C_5 = P_1 = P_2 = 0$.

During the search iterations, computation of ξ_1 and ξ_2 is straightforward, however, ξ_3 requires computing $\hat{\mathbf{q}}$ (see 10), for which there is no closed-form solution. Hence, $\hat{\mathbf{q}}$ needs to be estimated via a $2D$ search in the image space. For every incorrectly selected (noisy) image point, there exists a distorted image point along the radial direction. We can thus reduce the $2D$ search to a $1D$ search problem. Our calibration implementation takes under 30 seconds on a 300 MHz Pentium to estimate the radial and tangential distortion parameters.

Experimental results show that including the optical center parameters in the nonlinear search make the system unstable in the presence of noise. This was also observed in [4], even though the severity of distortions, as well as the noise levels, was much lower. Hence, we recommend nesting the estimation of $\{C_3, C_5, P_1, P_2\}$ within a coarse-to-fine search for the optical center (x_p, y_p).

5 Experiments with Synthetic and Real Images

To evaluate the robustness of our calibration technique, it is imperative to test it in the presence of noise. Noise enters the system from three main sources: human error in selecting points in the image, finite image resolution, and the fact that lines in the scene may not be perfectly straight. It is difficult to quantify the robustness of any nonmetric calibration method using only real images due to lack of ground truth. Hence, our experiments included simulated synthetic images as well as real images.

5.1 Simulation Results

Points were randomly sampled from synthetically generated lines with random orientations and positions (Fig. 4a). These points were then distorted with known distortion parameters (Fig. 4b). To simulate erroneous point selection, uniform noise in the interval $(-w, +w)$ was added to the points (Fig. 4c). These points were used to estimate the distortion parameters. The estimated parameters were used to undistort the noiseless, distorted image points (Fig. 4b) to recover the undistorted image points (Fig. 4d).

Measuring the deviation of the estimated parameters from the known (simulated) parameters does not give an intuitive feel of what distortions remain in the image. A good measure of accuracy is distance between the true perspective image points (Fig. 4a) and the undistorted image points (Fig. 4d). Our measure of error is the average distance, over all points used in the simulation. We tested each objective function (ξ_1, ξ_2, ξ_3) using various random line sets $\mathcal{L}$, distortion parameters $\mathcal{S}$, and several levels of noise ($w \in [0, 5]$ pixels).

TABLE 1
Errors Measured on Running Simulations Using ξ_1 (a), ξ_2 (b), and ξ_3 (c)

Distortion Coefficients					**Average Error (pixels)**			
$\mathcal{L}$	C_3	C_4	P_1	P_2	$w = 0$	$w = 1$	$w = 2$	$w = 5$
#1	10^5	10^9	10^5	10^5	0.000	3.360	13.973	42.521
	10^5	10^9	0.000	0.000	0.000	3.264	13.917	42.574
#2	10^5	10^9	10^5	10^5	0.000	12.095	39.567	66.817
	10^5	10^9	0.000	0.000	0.000	12.184	39.616	66.849

(a)

Distortion Coefficients					**Average Error (pixels)**			
$\mathcal{L}$	C_3	C_5	P_1	P_2	$w = 0$	$w = 1$	$w = 2$	$w = 5$
#1	10^5	10^9	10^5	10^5	0.000	0.356	2.473	12.383
	10^5	10^9	0.000	0.000	0.000	0.396	2.272	12.373
#2	10^5	10^9	10^5	10^5	0.000	1.618	5.448	28.639
	10^5	10^9	0.000	0.000	0.000	1.592	5.550	28.711

(b)

Distortion Coefficients					**Average Error (pixels)**			
$\mathcal{L}$	C_3	C_5	P_1	P_2	$w = 0$	$w = 1$	$w = 2$	$w = 5$
#1	10^5	10^9	10^5	10^5	0.002	0.363	0.390	0.398
	10^5	10^9	0.000	0.00	0.003	0.328	0.273	0.318
#2	10^5	10^9	10^5	10^5	0.008	0.663	0.773	0.502
	10^5	10^9	0.000	0.000	0.006	0.529	0.734	0.330

(c)

Although ξ_2 shows improvements over ξ_1, it is still very sensitive to noise. However, note the striking improvement over both ξ_1 and ξ_2.

Tables 1a, 1b, and 1c show the errors present in the recovered undistorted points using the sum of squares (ξ_1), normalized sum of squares (ξ_2) and the noise estimation method (ξ_3), respectively. Notice the sharp degradation in accuracy with increasing noise in cases of ξ_1 and ξ_2 (Tables 1a and 1b). However, ξ_2 does perform better than ξ_1 for certain noise levels. In contrast ξ_3 is much more robust and can yield subpixel accuracy even for high noise levels, as can be seen from Table 1c and Table 2. Even for large levels of noise ($w = 5$ pixels) the error is below 5 pixels, in spite of the nonlinear magnification of noise by the distortion model.

Recovery of the optical center is implemented as a coarse-to-fine exhaustive search around the image center. The search was done using a 5 x 5 grid at intervals of 10, 5, and 2 pixels. As Table 3 indicates, fine searches in the presence of noise can result in inaccurate solutions. This was also observed by Brown in [4], in spite of using data with much lesser noise. In contrast, coarse-to-fine searches appear to give better results. The time taken to recover all six distortion parameters $\{C_3, C_5, P_1, P_2, x_p, y_p\}$ is linear in the number of grid points used. The run time for the complete calibration algorithm for a 5 x 5 grid is about 20 minutes on a 300 MHz Pentium II machine.

5.2 Results with Real Images

We tested our algorithm based on ξ_3 using images taken with two different camera systems. To test robustness over a wide range of conditions, we used a low distortion camera (1/2″ CCD Sony XC-75 camera with a Computar 3.6mm lens) as well as an inexpensive board camera (1/3″ Computar EMH200-L25 CCD board camera with a 2.5mm lens) with severe lens distortions. The calibration of the sensors was done using a set of about 10 lines and a total of about 250 points. For unbiased results, the chosen scene lines should be uniformly distributed within the image. The estimated distortion parameters were used to undistort the acquired images. Fig. 5a illustrates an image acquired with the Computar system. This was undistorted using function ξ_3 (see Fig. 5b). As can be seen, straight lines in the scene now map to straight lines in image.

TABLE 2
Detailed Experimental Results for ξ_3

Distortion Coefficients					Average Error (pixels)			
$\mathcal{L}$	C_3	C_5	P_1	P_2	$w=0$	$w=1$	$w=2$	$w=5$
#1	10^5	10^9	10^5	10^5	0.002	0.428	0.522	0.391
	10^5	10^9	0.000	0.000	0.004	0.344	0.382	0.246
	10^5	10^{10}	0.000	0.000	0.281	0.348	0.579	2.818
	10^5	10^{10}	10^6	10^6	0.007	0.278	0.623	2.782
#2	10^5	10^9	10^5	10^5	0.000	0.151	0.015	0.068
	10^5	10^9	0.000	0.000	0.003	0.305	0.339	0.221
	10^5	10^{10}	0.000	0.000	0.029	0.152	0.345	1.591
	10^5	10^{10}	10^6	10^6	0.068	0.192	0.339	1.701
#3	10^5	10^9	10^5	10^5	0.000	0.501	0.574	0.590
	10^5	10^9	0.000	0.000	0.007	0.329	0.330	0.337
	10^5	10^{10}	0.000	0.000	0.043	0.444	0.488	2.356
	10^5	10^{10}	10^6	10^6	0.009	0.415	0.645	2.368

6 Polycameras

We define a polycamera as a tight cluster of cameras that together capture a large field of view. Unlike multiple cameras used in stereo, the cameras that comprise a polycamera are configured to have minimally overlapping fields of view. However, the finite size of sensors makes it difficult to maintain a single viewpoint. We therefore relax the single viewpoint constraint, but ensure that the individual viewpoints of the cameras are close enough to produce a seamless mosaic for objects beyond some minimum distance from the cluster center. We call this distance the *minimum working distance* of the polycamera.

In spirit, the idea of using multiple sensors is similar to that of Nalwa's system [9], which requires careful arrangement of mirrors and the cameras. In contrast, our system avoids the need for accurate positioning of the imaging components. Another sensor along these lines is the *Dodeca* by Immersive Media [8]. It uses 11 imaging sensors, arranged on a sphere tessellated as a dodecahedron. The large number of sensors increases the cluster size, thus causing greater parallax between views and increasing its minimum working distance. We would like to use the least number of cameras to capture a desired field of view. Wide-angle systems aid us in this respect. However, these systems suffer from radial and tangential distortions. To calibrate for these effects, one can use the methods proposed earlier in the paper.

TABLE 3
Results on Estimation of Optical Center (x_p, y_p)

Distortion Coefficients						Average Error (pixels)		
$\mathcal{L}$	C_3	C_5	P_1	P_2	Grid	$w=0$	$w=1$	$w=2$
#1	10^5	10^9	10^5	10^5	2	0.002	4.232	9.014
	10^5	10^9	10^5	10^5	5	0.002	0.363	10.220
	10^5	10^9	10^5	10^5	10	0.002	0.363	0.390
#2	10^5	10^9	10^5	10^5	2	0.008	4.271	3.792
	10^5	10^9	10^5	10^5	5	0.008	0.663	12.017
	10^5	10^9	10^5	10^5	10	0.008	0.663	0.773

6.1 Polycamera Configurations

We present configurations in which wide-angle sensors can be arranged to provide panoramic, hemispherical and spherical fields of view. Six sensors can be arranged on the sides of a cube so as to capture a spherical field of view. Each sensor is rotated by 90 degrees around its optical axis with respect to the adjacent side's sensor. This ensures a view overlap between adjacent sensors (Fig. 6a). The advantage of this configuration over the Dodeca is the use of fewer sensors while acquiring the same spherical field of view.

Another configuration includes four sensors oriented 90 degrees apart to capture a 360 degree panoramic field of view. Sensors with over 90 degrees horizontal field of view can be used in such configurations (Fig. 6b). A third configuration uses three imaging sensors arranged on the sides of a pyramid (Fig. 6c). Depending on the field of view of the individual sensors, a complete hemisphere to a partial hemisphere may be captured. The advantage of this configuration is that, with monochrome sensors, video signals from the three sensors can be combined into a single RGB signal to be used with a color frame buffer. This reduces the hardware and computational power needed for acquisition and processing.

7 Nonsingular Viewpoints and the Minimum Working Distance

In general, it is difficult to maintain a single viewpoint in polycameras. This causes parallax between overlapping views (Fig. 7b). Fig. 7a shows the disparity of a point Q when projected from two views, centered at C_1 and C_2, onto a single surface. However, depending on the resolution at which multiple views are combined, beyond some finite distance (*minimum working distance*), parallax effects become negligible. Now, we derive the constraint to estimate the minimum working distance for a polycamera.

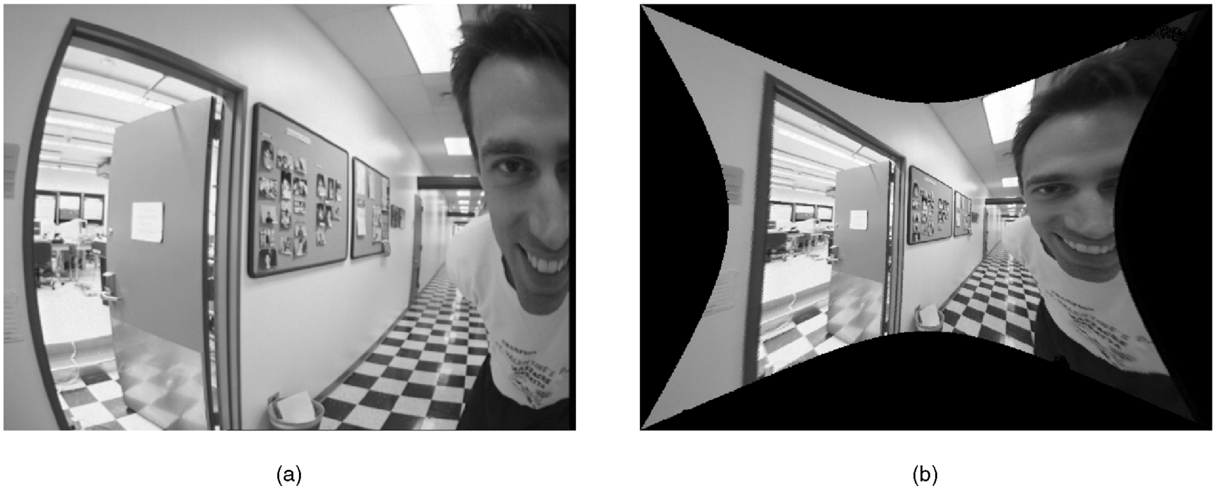
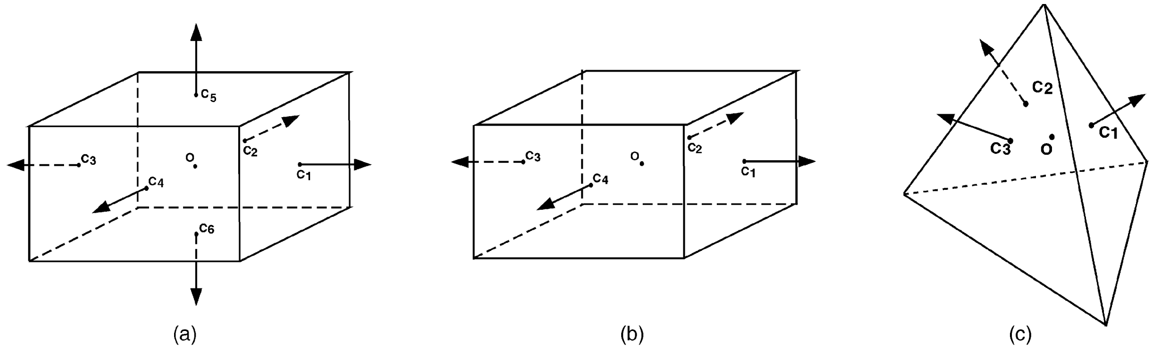
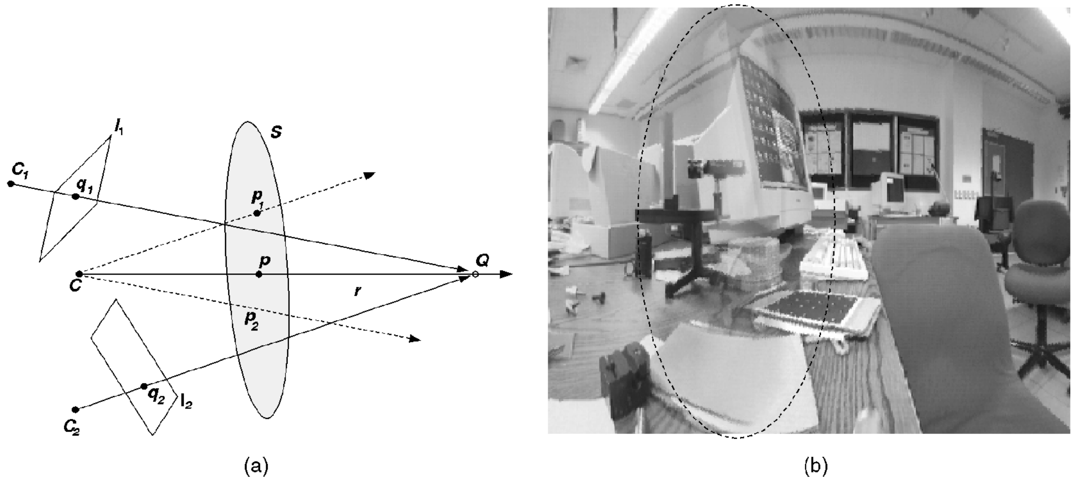


Fig. 5. (c) Image produced by a Computar 2.5mm lens and a Computar $1/3''$ CCD board camera. (b) Distortion parameters recovered via the minimization of ξ_3 are used to map (a) to perspective image. Notice that straight lines in the scene, such as door edges, map to straight lines in the undistorted images.



(a) (b) (c)

Fig. 6. (a) A spherical FOV polycamera configured using six wide-angle sensors. Each sensor is located on the sides of a cube such that their axes are mutually orthogonal. (b) A panoramic polycamera configured using four wide-angle sensors. Each camera has a horizontal field of view of over 90 degrees. (c) Three wide-angle sensors arranged on the sides of a pyramid.



(a) (b)

Fig. 7. (a) Two cameras, C_1 and C_2, image the scene point Q at points q_1 and q_2, respectively. Their projections onto the surface S are p_1 and p_2. p is the perspective projection of Q on S. The distance d at which the disparity between p_1 and p_2 falls below the threshold ϵ is the *minimum working distance*. (b) Segment of panorama clearly showing parallax effects for scene points within the minimum working distance.

Consider a projection surface S centered at O. Points p on this surface are parameterized in terms of the viewing direction (ϕ, θ) from O. A point Q in the scene along this ray direction at a distance d from O is given by:

$$\mathbf{Q} = \begin{pmatrix} d\cos(\theta)\cos(\phi) & d\sin(\theta) & d\cos(\theta)\sin(\phi) & 1 \end{pmatrix}^T. \qquad (12)$$

Let the relative orientation of a camera C_i with respect to O be D_i and its perspective projection matrix be P_i. $\mathcal{P}$ is a map of points imaged by C_i onto S. The map of the scene point Q imaged by C_i onto S then is:

$$p_i = \mathcal{P}(P_i \cdot D_i \cdot Q). \qquad (13)$$

The constraint that the parallax between two views falls below the threshold ϵ is then given by:

$$\|p_i - p\| \leq \frac{\epsilon}{2}. \qquad (14)$$

To estimate the minimum working distance for the entire cluster, we need to estimate d for all ray direction which project onto S. The maximum distance along all ray directions is the minimum working distance.

8 A REAL-Time PANORAMIC POLYCAMERA

Fig. 8 shows the polycamera consisting of four Computar EMH200-L25 board cameras with 2.5mm lenses placed approximately 90 degrees apart. Each camera has about 115 degree horizontal field of view. This ensures overlap between adjacent views as well as a complete 360 degree field of view. The complete sensor is enclosed in a cylinder that is 7 cm tall and 7.5 cm in diameter.

8.1 Minimum Working Distance

Fig. 9 illustrates the minimum working distance computed for the above polycamera. The four sensors were assumed to be coplanar and are 0.02m apart. The minimum working distance estimated was approximately 4.0m. Brighter points in the map correspond to

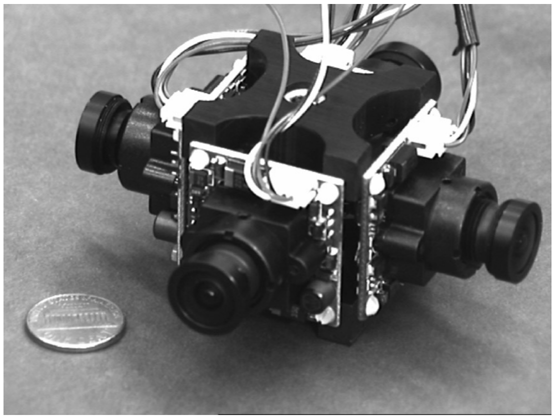


Fig. 8. A panoramic polycamera configured using four $1/3''$ CCD Computar EMH200-L25 board cameras with 2.5mm lenses. Each camera has a horizontal field of view of about 115 degrees, ensuring a 360 degree panoramic field of view.

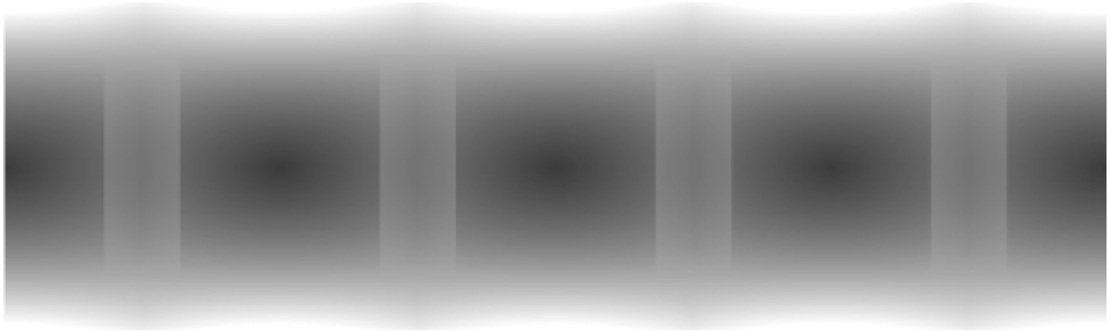


Fig. 9. The minimum distance map for a cylindrical polycamera for a parallax error of less than half a pixel. Brighter regions correspond to larger minimum working distances. The upper limit on the minimum working distance was about 4.0 meters.

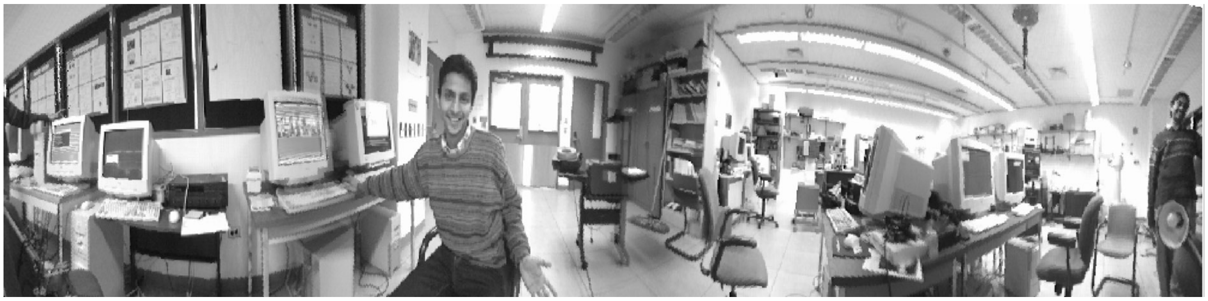


Fig. 10. Cylindrical panoramic video stream generated using the polycamera shown in Fig. 8. The panorama is computed using a look-up table, which is constructed taking into account the relative orientations of the four wide-angle cameras as well as their distortion parameters.

higher minimum working distances. As expected, the distance is minimum along the camera's optical axis and increases as we progress along any direction orthogonal to this axis. The sharp edge in the distance map corresponds to the overlapped field of view. In this region, we consider the maximum of the two minimum working distances computed for each view.

8.2 Panorama Generation

Prior to blending adjacent views, we need to undistort the image streams. We therefore calibrate each camera using objective function ξ_3 (11). Point correspondences in overlapping undistorted views were then used to estimate relative orientation (assumed to be purely rotational) between pairs of sensors. To account for the differences in the gains of the four cameras, we used the blending algorithm described by Szeliski [16], where each contributing pixel value is weighted by its proximity to the view boundary. The map between pixels in the panorama and the acquired images is stored as a static lookup table. Four video streams captured simultaneously using four Matrox boards and a 400 MHz Pentium-II PC were used to create a real-time 1,000 x 480 panorama at 15Hz (see Fig. 10).

9 Summary

We presented a new method to calibrate imaging systems for radial and decentering distortions. The distortion parameters were estimated by minimizing an objective function designed to handle large noise levels. Also, we proposed the idea of a polycamera (closely packed camera clusters) and presented polycamera configurations for various fields of view. We defined the minimum working distance for polycameras and described ways to compute it. Finally, results for a four sensor panoramic polycamera were presented.

Acknowledgments

This work was supported in part by the VSMA effort of DARPA's Image Understanding Program and an ONR/DARPA MURI grant under ONR contract no. N00014-97-1-0553.

References

[1] S. Becker and V.M. Bove, "Semi-Automatic 3-D Model Extraction from Uncalibrated 2-D Camera Views," *Proc. SPIE Visual Data Exploration and Analysis II,* vol. 2,410, pp. 447-461, Feb. 1995.

[2] M. Born and E. Wolf, *Principles of Optics.* Permagon Press, 1965.

[3] D.C. Brown, "Decentering Distortion of Lenses," *Photogrammetric Eng.,* vol. 32, no. 3, pp. 444-462, May 1966.

[4] D.C. Brown, "Close Range Camera Calibration," *Photogrammetric Eng.,* vol. 37, no. 8, pp. 855-866, Aug. 1971.

[5] A. Conrady, "Decentering Lens Systems," *Monthly Notices of the Royal Astronomical Soc.,* vol. 79, pp. 384-390, 1919.

[6] A. Goshtasby, "Correction of Image Deformation from Lens Distortion Using Bezier Patches," *Computer Vision, Graphics, and Image Processing,* vol. 47, pp. 385-394, 1989.

[7] S.B. Kang, "Semi-Automatic Methods for Recovering Radial Distortion Parameters from a Single Image," Technical Reports Series CRL 97/3, DEC, Cambridge Research Labs, May 1997.

[8] D. McCutchen, "Method and Apparatus for Dodecahedral Imaging," at United States Patent no. 5,023,725 (http://www.immersivemedia.com/), 1991.

[9] V. Nalwa, "A True Omnidirectional Viewer," technical report, Bell Laboratories, Holmdel, N.J., Feb. 1996.

[10] J.A. Nelder and R.A. Mead, "A Simplex Method for Function Minimization," *Computer J.,* vol. 7, pp. 308-313, 1965.

[11] H.S. Sawhney and R. Kumar, "True Multi-Image Alignment and Its Application to Mosaicing and Lens Distortion Correction," *Proc. 1997 Conf. Computer Vision and Pattern Recognition,* pp. 450-456, 1997.

[12] H.S. Sawhney and R. Kumar, "True Multi-Image Alignment and Its Application to Mosaicing and Lens Distortion Correction," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, no. 3, pp. 235-243, Mar. 1999.

[13] G.P. Stein, "Internal Camera Calibration Using Rotation and Geometric Shapes," AITR-1,426, master's thesis, MIT, AI Laboratory, 1993.

[14] G.P. Stein, "Accurate Internal Camera Calibration Using Rotation, with Analysis of Sources of Error," *Proc. Fifth Int'l Conf. Computer Vision,* pp. 230-236, 1995.

[15] G.P. Stein, "Lens Distortion Calibration Using Point Correspondences," *Proc. 1997 Conf. Computer Vision and Pattern Recognition,* pp. 143-148, June 1997.

[16] R. Szeliski, "Video Mosaics for Virtual Environments," *IEEE Computer Graphics and Applications,* vol. 16, no. 2, pp. 22-30, Mar. 1996.

[17] R.Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision," *Int'l J. Robotics and Automation,* vol. 3, no. 4, pp. 323-344, Aug. 1987.

[18] "VNI," *IMSL Fortran Numerical Libraries,* 1998.

[19] J. Weng, P. Cohen, and M. Herniou, "Camera Calibration with Distortion Models and Accuracy Evaluation," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 14, no. 10, pp. 965-980, Oct. 1992.