

Towards Calibrating a Pan-Tilt-Zoom Camera Network

Sudipta N. Sinha and Marc Pollefeys

Department of Computer Science,
University of North Carolina at Chapel Hill.

Abstract. In this paper we discuss the problem of recovering the calibration of a network of pan-tilt-zoom cameras. The intrinsic parameters of each camera over its full range of zoom settings are estimated through a two step procedure. We first determine the intrinsic parameters at the camera's lowest zoom setting very accurately by capturing an extended panorama. Our model includes two parameters of radial distortion. The camera intrinsics and radial distortion parameters are then determined at discrete steps in a monotonically increasing zoom sequence that spans the full zoom range of the cameras. Both steps are fully automatic and do not assume any knowledge of the scene structure. We validate our approach by calibrating two different types of pan tilt zoom cameras placed in an outdoor environment. We also show the high-resolution panoramic mosaics built from the images captured during this process. We present an approach for accurate computation of the epipolar geometry based on the full panorama instead of individual image pairs. Finally, we briefly discuss how this can be used to compute the extrinsics for all the cameras and how our approach can be used in the context of active camera networks.

1 Introduction

Active pan-tilt-zoom cameras in wide-area surveillance systems allows one to monitor a large area using only a few cameras. We refer to an unknown arbitrary configuration of cameras in such an environment as a camera network. During operation such cameras can act like high-resolution omnidirectional sensors, which can potentially track activities over a large area and capture high-resolution imagery around the tracked objects. Active camera networks could also be used for 3D modeling of large scenes and reconstruction of events and activities within a large area. This paper describes a fully automatic method for calibrating a network of such pan tilt zoom cameras, that does not require physical access to the cameras or the space in their field of view. Both intrinsics and extrinsic parameters are estimated from images captured within the camera's full range of pan tilt and zoom configurations. Our method is inherently feature-based, but does not require a calibration object or specific structures in the scene.

Past work on active camera calibration has mostly been done in a laboratory setup using calibration targets and LEDs or at least in a controlled environment. Some of these include active zoom lens calibration by Willson et. al. [9, 7, 10], self-calibration from purely rotating cameras by deAgapito [2], and more recently pan-tilt camera calibration by Davis et. al. [4]. Our approach towards zoom calibration is simpler than that of Wilson [10] who computed both focal length and radial distortion at many different zoom settings [10] and is similar to that of Collins et. al. [3], who calibrated a pan-tilt-zoom active camera system in an outdoor environment. However we extend the lens distortion model proposed by Collins [3] who assumed constant radial distortion and estimated it only at a particular zoom level and modelled its variation using a magnification factor. We actually estimate the radial distortion caused by optical zoom of the camera.

Thus in our method the camera intrinsics are estimated by robustly computing homographies induced by rotating and zooming the camera. Intrinsic parameters and radial distortion is first estimated for the lowest zoom setting of the camera by computing homographies between multiple images acquired by a rotating camera. Using bundle adjustment [8], the homography model is extended to take radial distortion into account. This allow us to build a panorama of the complete scene with sub-pixel alignment error. We then use an image sequence from the full zoom range of the camera to estimate the variation of its intrinsics with zoom. The epipolar geometry between two cameras is then robustly computed from the respective panoramas, as this provides a good distribution of features over a large field-of-view. A method to recover full metric calibration of all the cameras in the network from a sufficient number of fundamental matrices between camera-pairs using only pair-wise matches, is described in [1]. The paper is organised as follows. Section 2 introduces the camera model while Section 3 explains the calibration procedure. Section 4 addresses the construction of panoramas while calibration results are presented in Section 5. We conclude with discussions and scope for future work in Section 6.

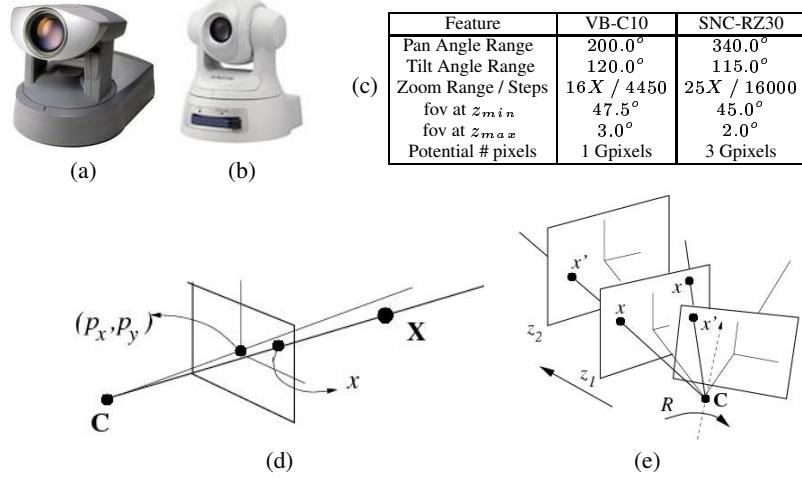


Fig. 1. The cameras we use: (a) Canon VB-C10. (b) Sony SNC-RZ30. (c) Relevant camera specifications. (d) Our pin-hole camera model. (e) Camera rotation and zoom.

2 Theory and Background

2.1 Camera Model

We chose to use a simple pan-tilt-zoom (PTZ) camera model and make a tradeoff for simplicity over exactness in our choice, similar to [2, 3]. Our model assumes that the center of rotation of the camera is fixed and coincides with the camera's center of projection during operation. Some inexpensive PTZ cameras may violate such an assumption depending on its pan-tilt mechanism and more general models [4, 9] have been proposed to deal with them. However when cameras are used outdoors or in large environments, the deviation of the center is negligible compared to the average distance of the observed features, which are typically distant. Our experiments with the Canon VB-C10 and Sony SNC-RZ30 surveillance cameras (see Fig. 1 for relevant specifications) have shown this model to be reasonably accurate.

$$\mathbf{P} = \mathbf{K}[\mathbf{R} - \mathbf{R}\mathbf{t}] \quad \mathbf{K} = \begin{pmatrix} \alpha f_y & s & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{pmatrix} \quad (1)$$

In the pin-hole camera model (see Fig. 1(d)) for the perspective camera, a point \mathbf{X} , in 3D projective space P^3 projects to a point \mathbf{x} , on the 2D projective plane P^2 (the image plane). This can be represented by a mapping $f : P^3 \rightarrow P^2$ such that $\mathbf{x} = \mathbf{P}\mathbf{X}$, \mathbf{P} being the 3×4 rank-3 camera projection matrix. This matrix \mathbf{P} can be decomposed as shown in Eq. 1 where \mathbf{K} represents the camera intrinsics while \mathbf{R} and \mathbf{t} represents the camera position and orientation with respect to the world coordinate system. The matrix \mathbf{K} can be expressed in terms of α , s , f_y , p_x and p_y (Eq. 1), where α and s are the camera's $x:y$ pixel aspect ratio and skew respectively; f_y its focal length f , measured

in pixel in the y directions; (p_x, p_y) its principal point in the image. Since we model the camera's pan and tilt movements by pure rotations about its projection center \mathbf{C} , we choose it as the world origin and set $t = 0$. Our goal is to estimate the unknown parameters of a model for $\mathbf{K}_{\mathbf{p}, \mathbf{t}, \mathbf{z}}$ that provides the intrinsics for any (pan= p ; tilt= t ; zoom= z) configuration within the admissible PTZ ranges. The principal point (p_x, p_y) and focal length f depend only on the camera's zoom, and not its state of pan and tilt. Hence the unknown intrinsics we wish to estimate are of the form in Eq. 2.

$$\mathbf{K}(\mathbf{z}) = \begin{pmatrix} \alpha f_y(z) & s & p_x(z) \\ 0 & f_y(z) & p_y(z) \\ 0 & 0 & 1 \end{pmatrix} \quad (2)$$

where α and s , (we assume $s=0$) are constants for a particular camera and z is its zoom level. Most cameras deviate from a real pin-hole model due to effects of radial distortion which becomes more prominent as the focal length of the camera decreases. The 3D point \mathbf{X} which projects to $\mathbf{x} = (\tilde{x}, \tilde{y}, 1)$ under the pin-hole model actually gets imaged at (x_d, y_d) due to radial distortion as shown in Eq. 3. $\tilde{r} = \sqrt{\tilde{x}^2 + \tilde{y}^2}$ is the radial distance of \mathbf{x} from the center of distortion (x_c, y_c) and $\mathbf{L}(\tilde{\mathbf{r}})$ is a distortion factor determined by \tilde{r} . The function $\mathbf{L}(\mathbf{r})$ is represented as $\mathbf{L}(\mathbf{r}) = 1 + \kappa_1 \mathbf{r}^2 + \kappa_2 \mathbf{r}^4 + \kappa_3 \mathbf{r}^6 + \dots$ and $(\kappa_1, \kappa_2, \kappa_3, x_c, y_c)$ is the parametric model for radial distortion. Since the effect of radial distortion diminishes with increasing focal length, our model for radial distortion, \mathcal{R} is also a function of zoom.

$$\begin{pmatrix} x_d \\ y_d \end{pmatrix} = \mathbf{L}(\tilde{\mathbf{r}}) \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} \quad \mathcal{R} = (\kappa_1(\mathbf{z}), \kappa_2(\mathbf{z}), \kappa_3(\mathbf{z}), \mathbf{x}_c(\mathbf{z}), \mathbf{y}_c(\mathbf{z})) \quad (3)$$

We determine calibration over the full zoom range by estimating $\mathbf{K}(\mathbf{z})$ and $\mathcal{R}(z)$ (as shown in Eqs. 2 and 3) at equal steps of zoom on a logarithmic scale, between z_{min} and z_{max} , the minimum and maximum optical zoom levels respectively. For the VB-C10 and the SNC-RZ30, (z_{min}, z_{max}) was $(0, 120)$ and $(0, 101)$ respectively. We linearly interpolate between these discrete levels to obtain intrinsics at any state of zoom.

2.2 Rotating and Zooming Cameras

Here we consider the case of a rotating and zooming camera. Let x and x' be the images of \mathbf{X} taken at two different instants by a camera that is either zooming or rotating or both. These points, x and x' are related to X as $x = \mathbf{K}[\mathbf{R} \ t]\mathbf{X}$ and $x' = \mathbf{K}'[\mathbf{R}' \ t']\mathbf{X}$. Hence, $x' = \mathbf{K}'\mathbf{R}'\mathbf{R}^{-1}\mathbf{K}^{-1}\mathbf{x}$. In our model, the intrinsics remain the same for pure rotation at constant zoom, and hence this equation reduces to $x' = \mathbf{K}'\mathbf{R}_{rel}\mathbf{K}^{-1}\mathbf{x}$ where $\mathbf{R}_{rel} = \mathbf{R}'\mathbf{R}^{-1}$ represents the relative camera rotation about its projection center between the two views and \mathbf{K} is the camera intrinsic matrix for that particular zoom level. Similarly for a zooming camera with fixed center of projection, $x' = \mathbf{K}'\mathbf{K}^{-1}\mathbf{x}$. These homographies are represented by \mathbf{H}_{rot} and \mathbf{H}_{zoom} (see Eq. 4).

$$\mathbf{H}_{rot} = \mathbf{K}'\mathbf{R}\mathbf{K}^{-1} \quad \mathbf{H}_{zoom} = \mathbf{K}'\mathbf{K}^{-1} \quad (4)$$

3 Our Calibration Procedure

An overview of the two phase calibration procedure is shown in Fig. 2. The notation used here is as follows. I_i are images acquired by the rotating camera. H_i and V_j represent the homographies between horizontal and vertical adjacent pairs in the image grid. T_i represents the homographies with respect to a reference image I_r . The optimal homographies \tilde{T}_i , computed through *Bundle I*, are used to obtain an approximate intrinsics $K(z_{min})$ at the lowest zoom. \tilde{R} represents camera rotation matrices and $R_d(z)$ stands for the radial distortion model parameters at zoom level z .

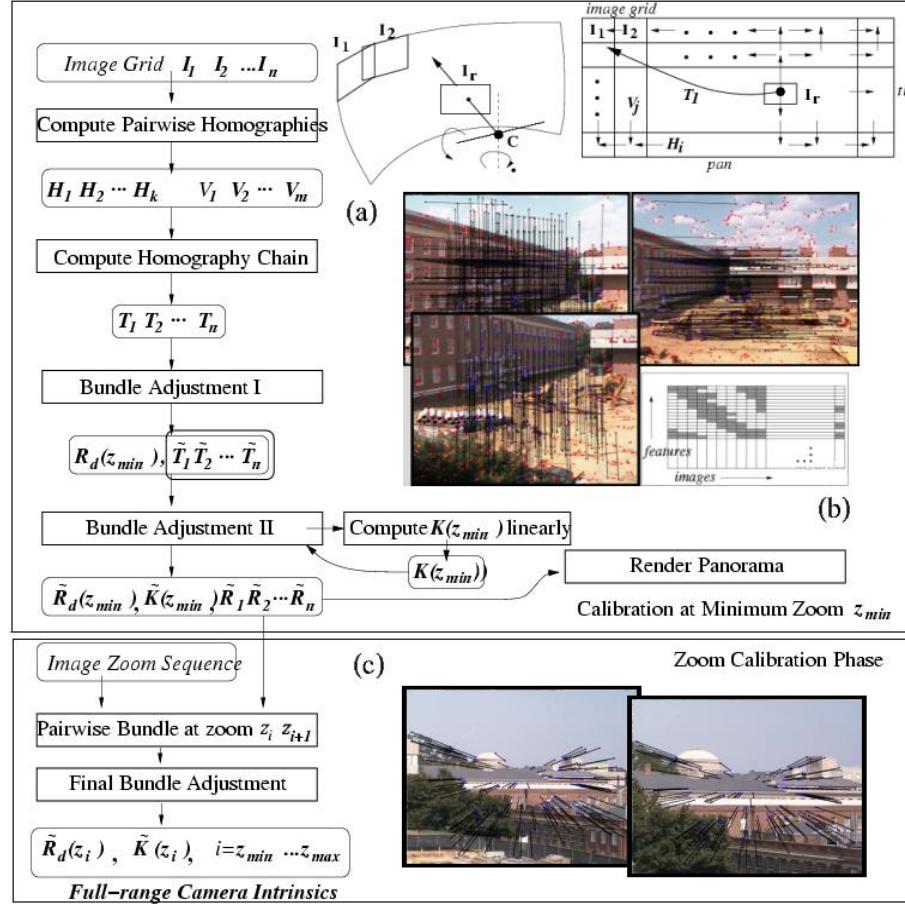


Fig. 2. Overview of the intrinsics calibration procedure. (a) The image grid I_i captured during rotation at fixed zoom and the mosaic computed with respect to I_r . (b) Horizontally and vertically adjacent images in the grid shown with the corresponding matches. Feature lists built from features visible in two or more images are illustrated. (c) Successive images from a zoom sequence shown with the corresponding matches.

3.1 Computing Intrinsic at Minimum Zoom

The first step towards computing intrinsic is determining the intrinsic at minimum zoom. This is done from images captured from a rotating camera by computing homographies \mathbf{H}_{rot} (see Eq. 4) between each of them and a chosen reference image. During a capture phase, images are acquired in a spherical grid (see Fig. 2(a)) for certain discrete pan and tilt steps. Fig. 5 shows 18 images captured at 6 pan and 3 tilt steps. The homographies between every adjacent horizontal pair of images, H_i and between every adjacent vertical pair, V_j in the grid are computed as described in [5] (Chap.3, page 108). Fig. 2(b) shows a horizontal image pair and a vertically pair with the respective matched features. One of the images, I_r is chosen as the reference image and homographies, T_i are computed for every image I_i , by composing a sequence of transformations, $(\dots H_a, H_b \dots V_c, V_d \dots)$ along a connected path between I_i to I_r in the image grid as illustrated in Fig. 2(a). An accurate estimate of T_i 's for all the images would allow multi-image alignment in the image plane of I_r . Since residual errors accumulate over the composed homographies, the final mosaic obtained by aligning all the images, contains significant registration errors.

Global image alignment and sub-pixel registration is achieved through a bundle adjustment (we use a sparse implementation to efficiently deal with a large number of images) that also estimates the radial distortion parameters. The bundle adjustment [8] is initialized using the set of computed homographies T_i and by building a global feature list (see Fig. 2(b)) from the horizontal and vertical pairwise matches. Bundle Adjustment performs global minimization which produces the maximum likelihood estimation of the model parameters when the image error is assumed to be zero-mean Gaussian noise. This is first used to estimate all the homographies T_i , the radial distortion parameters $\mathcal{R}(z_{\min})$ and the panorama features X^j that minimizes the mean square error between the observed image features and the reprojected points. (see Eq. 5). We call this *Bundle I* in Fig. 2.

$$\min_{T_i, \mathcal{R}(z_{\min}), X^j} \sum_{k=1}^m \sum_{i=1}^n D(x_i^j, \mathcal{R}(T_i X^j))^2 \quad (5)$$

The accurate homographies from this bundle, \tilde{T}_i are used to estimate $K(z_{\min})$ using the Hartley's linear algorithm for computing intrinsic for a purely rotating camera [11]. This value of $K(z_{\min})$ is used to initialize another bundle, *Bundle II*, which refines the estimates of $K(z_{\min})$, $\mathcal{R}(z_{\min})$ by minimizing the reprojection error given by Eq. 6.

$$\min_{K(z_{\min}), R_i, \mathcal{R}(z_{\min}), X^j} \sum_{k=1}^m \sum_{i=1}^n D(x_i^j, K(\mathcal{R}(R_i X^j)))^2 \quad (6)$$

Every 2D homogeneous feature point, X^j is projected to a unit cube and parameterized as (a_i, b_i) where the third coordinate is set to +/-1 depending on the particular cube face that X^j projects to.

3.2 Zoom Sequence Calibration

Full range zoom calibration can be achieved by building a mosaic and repeating the process described in Sec. 3.1 at multiple steps within the camera's zoom range. How-

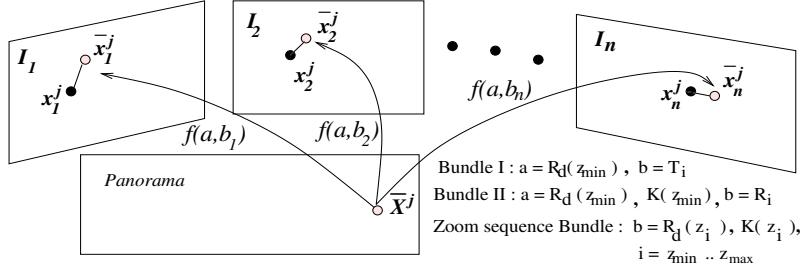


Fig. 3. Minimizing the reprojection error in a Bundle Adjustment. $f(a, b_i)$ represents the model parameters. Common parameters are denoted by a and view-dependent parameters by b . The values of a, b in different bundle adjustment phases are shown.

ever Hartley [11], showed that more images are required in the mosaic as the camera progressively zooms in and its effective field-of-view decreases making the approach time-consuming. Instead we compute intrinsics over the zoom range using an image sequence, captured in a fixed direction with the camera progressively zooming in. We then estimate homographies H_{zoom} , (see Eq. 4) between every image pair within this sequence. This requires fewer images and is considerably faster. The first phase of zoom calibration estimates $K(z_i)$ and $R(z_i)$, the intrinsics at zoom level, z_i for every step in the zoom sequence. This is done using a pairwise bundle adjustment that works with images at zoom steps z_i, z_{i-1} , using the value of $K(z_{i-1})$ and $R(z_{i-1})$ estimated in the previous iteration to minimize the reprojection error in Eq. 7. A full bundle adjustment then refines the estimated parameters by minimizing the same reprojection error over all the images in the sequence. Estimating radial distortion only from a zoom sequence has inherent ambiguities since a distortion at a particular zoom can be compensated by a radial function at another zoom. We avoid this ambiguity by keeping the intrinsics computed at the lowest zoom fixed in the bundle adjustment step. The uncertainties of the estimated parameters are used to determine the zoom levels at which the effect of radial distortion, namely coefficients κ_2 and then κ_1 , becomes negligible. Wilson [9] shows that the image center is hard to estimate accurately. We chose to constrain the principal point to be the same as the center of radial distortion in our zooming camera model.

$$\min_{K(z_i), R(z_i), X^j} \sum_{k=1}^m \sum_{i=1}^n D(x_i^j, K(R(X^j)))^2 \quad (7)$$

3.3 Constructing Panoramas

Sec. 3.1 describes the approach, similar to that of [6] for achieving sub-pixel accuracy in multi-image alignment and creating mosaics from images acquired by a rotating camera at its lowest value of zoom. We could build mosaics at large focal lengths too (see Fig. 4). Since the unknown focal length f is computed through the same bundle, the cube-map face is chosen to be of size $2f \times 2f$, since this preserves the pixel resolution of the original images. We have rendered panoramas with a single cube-map face at resolutions of $6k \times 6k$ pixels, from 119 images ($f=3120$ pixels approx. 5X zoom) in about

20-25 mins. Panoramas created at different resolutions would allow background image synthesis for any pan, tilt and zoom at an appropriate resolution and allow background differencing, change detection and high precision calculation of PTZ settings even if the camera control is not repeatable. A tile-based representation is adopted for these large images to allow efficient out-of-core processing. We use a simple blending function, (radially weighted averages) to blend overlapping pixels in the generated cubemaps. Better blending functions can remove the artifacts caused by the high dynamic range in the scene, but these require estimates of the camera's response function.



Fig. 4. The front face of the computed cube-maps: (a) Radial distortion was ignored in the camera model (Note that straight lines in the world are not imaged as straight lines). (b) Accurate panorama created after incorporating radial distortion. (c) High-resolution panorama (6000×6000 pixels) built from 119 images at 5X zoom. Note the zoomed-in regions of the panorama, displayed in the original scale.

3.4 Computing Extrinsic

We propose to estimate the extrinsic parameters for all the cameras in the network by computing the pairwise epipolar geometry for sufficient camera pairs. The fundamental matrix for a pair of PTZ cameras can be computed more robustly from the panoramas we build. Since the cameras are located in the same environment, the panoramic images are guaranteed to contain overlapping regions which are visible in both views. In spite of the presence of moving objects in the scene which gives rise to outliers, a lot of spatially well distributed static features are guaranteed to exist in most environments. This allows the use of completely unsynchronized or asynchronous cameras which acquire images completely independent of each other. The fundamental matrix corresponding to a view-pair is computed by the robust RANSAC-based algorithm described in [5] (Chap 10, page 275). Although extracted features for each camera are represented by panorama coordinates, feature correspondences are determined by comparing pixels in the original images. Each panorama feature is mapped back into one of the acquired images and its local neighbourhood in those images are compared using a normalized cross-correlation measure. A complete metric calibration of all the cameras can be obtained from a sufficient number of fundamental matrices and pairwise matches using the method described in [1]. This involves computing a projective reconstruction and upgrading it to a metric reconstruction using self-calibration. However metric cameras can also be directly obtained from the fundamental matrices via the essential matrix [5] (Chap 8, page 226), since the corresponding intrinsic parameters have already been computed. In this paper we compute extrinsics for a short baseline camera-pair using this simple method. However for larger camera networks, the method described in [1] is more appropriate.

4 Experimental Results

Here we present results from fully calibrating two Canon VB-C10 and two Sony SNC-RZ30 pan tilt zoom cameras in an outdoor environment. The cameras are placed near two adjacent windows about 3-4 meters apart looking out at a construction site roughly 100×120 meters in area. This setup reduced each camera's available field of view for pan to only 150° . Hence only the front face of the cubemaps we build are interesting and hence shown. Fig. 5 shows the images used in calibration and Fig. 4 the resulting panoramas. The recovered intrinsics for the four cameras as a function of zoom are shown in Fig. 6(a). The principal point was found to move in a straight line for difference zoom sequences. The motion was most noticeable at high zooms. The VB-C10 had a linear mapping of focal length to zoom whereas the SNC-RZ30's focal length was non-linear. The pixel aspect ratio of the VB-C10's and SNC-RZ30's were found to be 1.09 and 0.973 respectively while the skew was assumed to be zero. Repeated zoom sequence calibration for the same camera from different datasets (Fig. 6(b)&(c)) showed the focal length estimation to be quite repeatable. The coefficients of radial distortion in our model, κ_1 and κ_2 were estimated along with their respective uncertainties. These uncertainties were used to clamp κ_1 and κ_2 to zero at particular zoom steps during the pairwise bundle adjustment described in Sec 3.2. The mean reprojection error from the final zoom sequence bundle for 35-40 images, with roughly 200-300 feature matches

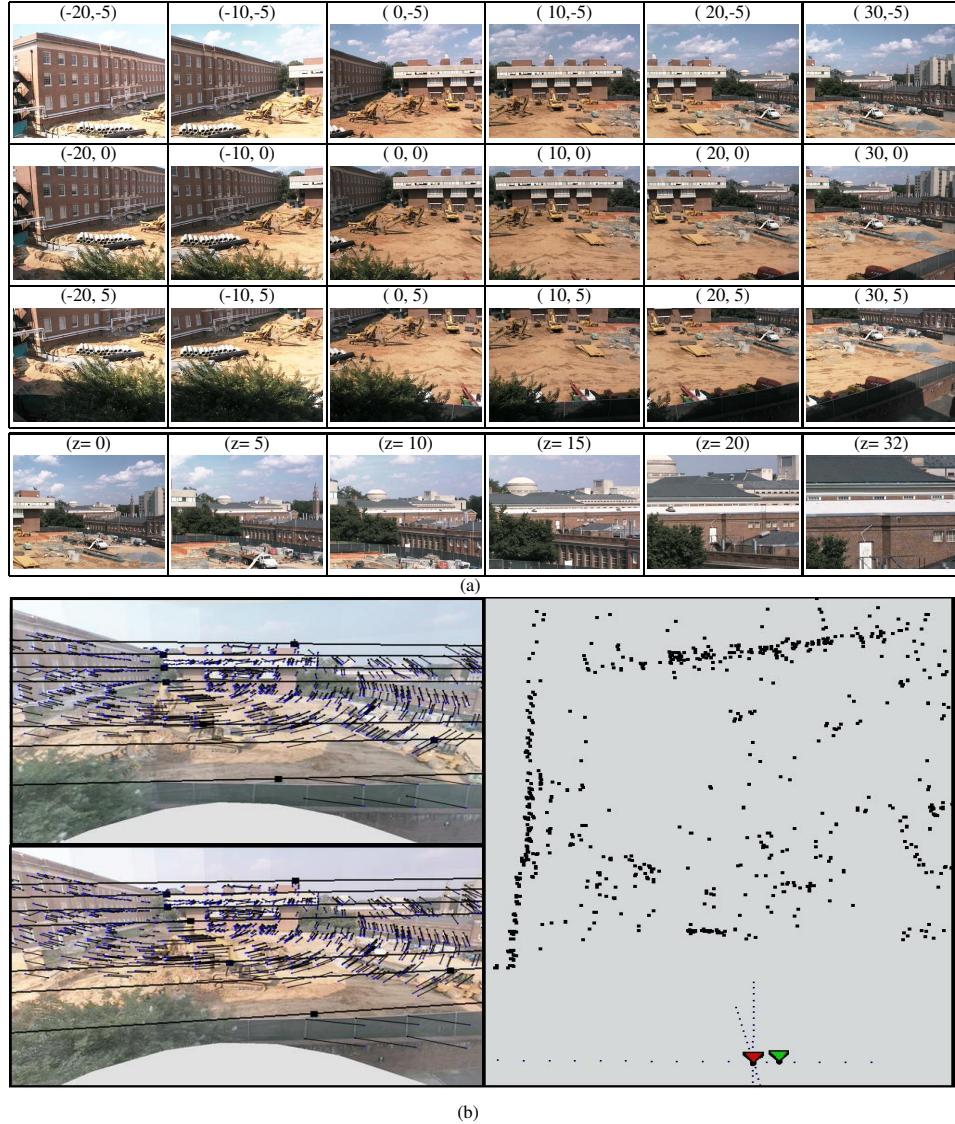


Fig. 5. (a) Top 3 rows: 18 images captured at a fixed zoom ($z=0$), pan and tilt angles are in degrees. The bottom row shows 6 frames from a zoom sequence of 36 images for fixed pan & tilt angles. (b) The front faces of panoramic cubemaps built from each of two cameras placed on adjoining windows 3-4 meters apart are shown. The corresponding features and a few corresponding epipolar lines are shown. An orthographic top-view of the recovered cameras and some of the reconstructed feature points are shown upto an ambiguity of scale.

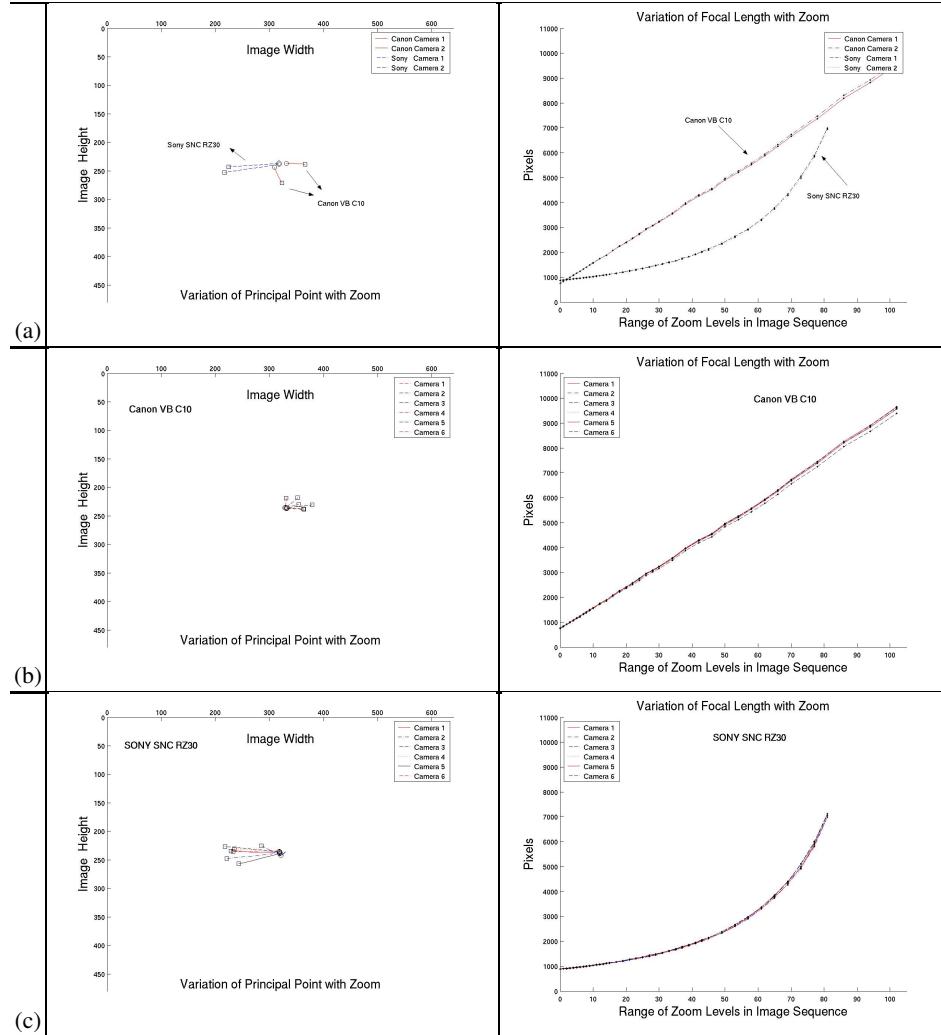


Fig. 6. (a) The variation of the principal point and focal length with zoom is shown for each of the four cameras in our experiments. (b) Calibration results of a Canon VB-C10 from six different image sequences. (c) Calibration results of a Sony SNC-RZ30 from six different images sequences.

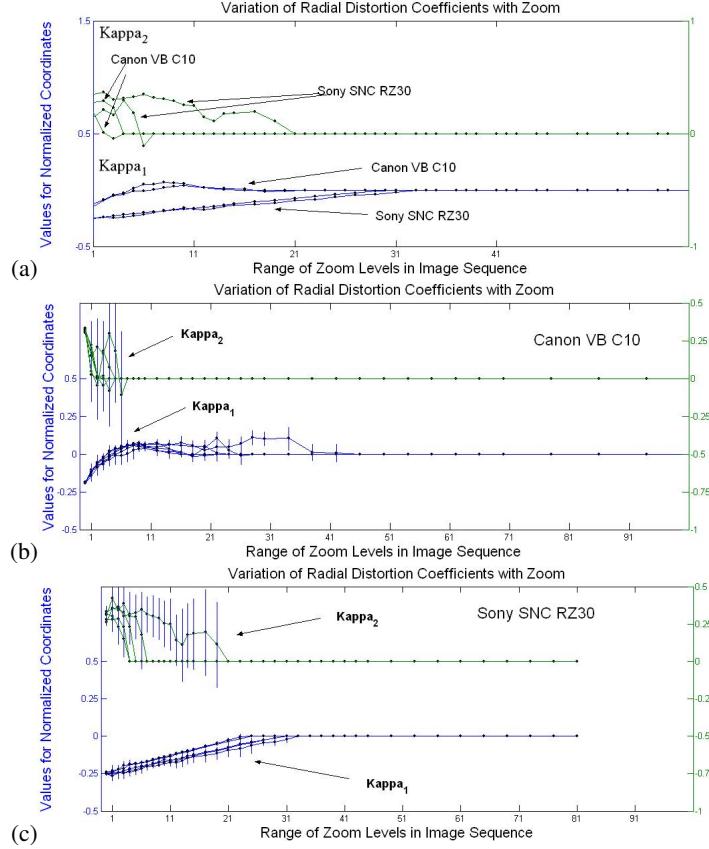


Fig. 7. (a) The variation of the radial distortion coefficients, κ_1 and κ_2 , with zoom is shown for our 4 cameras. The uncertainties associated with κ_1 and κ_2 are shown using error bars. κ_2 has larger uncertainty compared to κ_1 which is estimated for a longer zoom range. (b) and (c) shows the variation of κ_1 and κ_2 for repeated calibration of a Canon and Sony camera, each using 6 different image sequences.

for every successive pair was within 0.43 pixels.

For computing extrinsics, at this stage we concentrate on robustly estimating the fundamental matrix from the cubemap panoramas. Our results for a single camera pair are shown in Fig. 5. The corresponding epipolar lines and the feature correspondences extracted by our method are displayed. The metric cameras recovered via the essential matrix and the reconstructed points are shown on the right.

5 Conclusions

We have presented an automatic method for calibrating a network of active PTZ camera typically used in surveillance systems. The camera intrinsics are estimated over its full

range of pan, tilt, zoom by robustly computing homographies between images acquired by a rotating and zooming camera. Our calibration algorithm also computes accurate panoramas by building a mosaic from image acquired by the camera rotating at fixed zoom. The camera extrinsics are then robustly determined by estimating the epipolar geometry from a pair of panoramas. In future we will work on maintaining calibration of active cameras in operation and determine the degree of repeatability of their PTZ mechanisms. A lack of repeatability will be addressed by building an efficient closed-loop system, that re-estimates the calibration everytime the camera moves, by registering its images with a pre-computed calibrated panorama of the background. The pan-tilt-zoom camera described here, can potentially produce extremely detailed omnidirectional images by virtue of their long zoom range. Since stitching the whole mosaic is expensive at high resolutions, we are currently exploring an efficient multi resolution framework for rendering very large panoramas using a hierarchy of zooms.

Acknowledgements

This work was partially supported by the NSF Career award IIS 0237533 and the DARPA/DOI award NBCH 1030015.

References

1. S.N. Sinha and M. Pollefeys. *Camera Network Calibration from Dynamic Silhouettes*, to appear at CVPR04.
2. L. de Agapito, R. Hartley, and E. Hayman. *Linear selfcalibration of a rotating and zooming camera*. In Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition, pages 15–21, 1999.
3. R.T. Collins and Y. Tsin. *Calibration of an outdoor active camera system*. In Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition, pages 528–534. IEEE Computer Society, June 1999.
4. J. Davis and X. Chen, *Calibrating pan-tilt cameras in wide-area surveillance networks*. In Proc. of ICCV 2003, Vol 1, page 144-150, October 2003.
5. R. Hartley and A. Zisserman, *Multiple View Geometry In Computer Vision*, Cambridge University Press, 2000.
6. H. S. Sawhney, S. Hsu, and R. Kumar. *Robust video mosaicing through topology inference and local to global alignment*. In Proc. European Conference on Computer Vision, 1998.
7. B. J. Tordoff and D. W. Murray. *Violating rotating camera geometry: The effect of radial distortion on self-calibration*. In Proc. International Conference on Pattern Recognition, 2000.
8. B. Triggs, P. McLauchlan, R. Hartley, A. Fitzgibbon, *Bundle Adjustment: A Modern Synthesis*, In B. Triggs, A. Zisserman, R. Szeliski (Eds.), *Vision Algorithms: Theory and Practice*, LNCS Vol.1883, pp.298-372, Springer-Verlag, 2000.
9. R. Willson and S. Shafer, *What is the Center of the Image?* in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 670-671, 1993.
10. R.G. Willson. *Modeling and Calibration of Automated Zoom Lenses*. PhD thesis, Carnegie Mellon University, 1994.
11. R. I. Hartley. *Self-calibration of stationary cameras*. International Journal of Computer Vision, 22(1):5–23, 1997.