CENTER FOR
MACHINE PERCEPTION

CZECH TECHNICAL
UNIVERSITY

REPRINT

# 3D Reconstruction from 360 x 360 Mosaics[1]

{bakstein, pajdla}@cmp.felk.cvut.cz

Hynek Bakstein and Tomáš Pajdla, 3D Reconstruction from 360 x 360 Mosaics, in Proceedings of the CVPR 2001, to be published





Center for Machine Perception, Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University
Technická 2, 166 27 Prague 6, Czech Republic
fax +420 2 2435 7385, phone +420 2 2435 7637, www: http://cmp.felk.cvut.cz

# 3D Reconstruction from 360 x 360 Mosaics

Hynek Bakstein    Tomáš Pajdla

Center for Machine Perception, Dept. of Cybernetics
Faculty of Electrical Eng., Czech Technical University
121 35 Prague, Czech Republic
{bakstein,pajdla}@cmp.felk.cvut.cz

## Abstract

*We are studying the geometry of a 360 x 360 mosaic image formation. A 360 x 360 mosaic camera model and a calibration procedure are proposed. It is shown that only one point correspondence is needed in order to acquire epipolarily rectified images. The 360 x 360 mosaic camera model is therefore determined by only one intrinsic parameter. It is shown that the relation between coordinates estimated with different values of intrinsic 360 x 360 mosaic camera parameters is a scaling of all scene point coordinates with additional nonlinear changes in the $z$ coordinates of the scene points. Experimental results verifying the reconstruction of real scene points are presented.*

## 1. Introduction

Recently, there has been an increased interest in the development of panoramic sensors. Several approaches have used a mirror to enlarge the field of view [1, 2, 4, 5, 9, 8]. Other methods employed moving parts to capture the whole panoramic scene [11, 9, 3]. Finally, there are systems combining both mirrors and camera movement [6].

Some of these systems can be modeled by the central projection [8]. However, many of the proposed methods lead to a noncentral camera. A noncentral camera is a camera where the light rays forming the image do not intersect in one common point. Instead, for example, they are tangent to a circle, or intersect a common line. Recently, attention was paid to the geometrical properties of noncentral cameras. It has been shown that stereo geometries of noncentral cameras exist and a generalized epipolar geometry with double ruled quadric epipolar surfaces has been described [7, 10]. The 360 x 360 mosaic camera model, its calibration procedure, and the relation between scene reconstructions from uncalibrated 360 x 360 mosaic images are presented in this paper.

The 360 x 360 [6] mosaic is an example of a noncentral camera where the light rays are tangent to a circle, see Figure 1. Let us explain the geometry of the light rays that form the mosaic. A plane $\pi$ is rotated on a circular path $\mathcal{C}$ with a radius $r$. The plane $\pi$ is perpendicular to the plane $\delta$, in which the circle $\mathcal{C}$ lies, and it is tangent to the circle $\mathcal{C}$. Let us take at each rotation position all the light rays that lie in the plane $\pi$ and intersect the point where the plane $\pi$ touches the circle $\mathcal{C}$. Because each point outside the circle $\mathcal{C}$ can be observed by two light rays, the 360 x 360 mosaic camera provides a complete spherical mosaic for both left and right eye after rotation of the plane $\pi$ about 360°. See [6] for more details about the 360 x 360 mosaic.

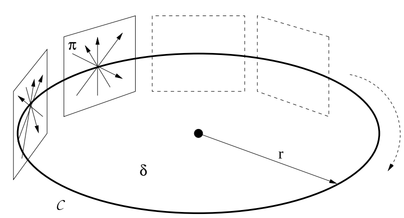


**Figure 1. 360 x 360 mosaic camera geometry.**

We can imagine that the plane $\pi$ represents a 1D omnidirectional camera [8] which is rotated on a circular path. This 1D camera and the motion together compose the 360 x 360 mosaic camera. Let us mention two possible realizations of the 1D omnidirectional camera. The first consists of a 2D CCD camera with a telecentric lens observing a conical mirror with a 90° angle at the apex, depicted in Figure 2(a). The second is a 2D CCD camera equipped with a fish eye lens with a field of view of at least 180°, e.g. Nikon fish

*This work was supported by the following grants: GAČR 102/01/0971, OMNIVIEWS IST-1999-29017, MŠMT KONTAKT 2001/09, and MŠMT 212300013. The authors would also like to thank Vladimír Smutný and Radim Šára for suggestions.

eye converter shown in Figure 2(b). Light rays from both camera realizations will lie in the plane $\pi$. The proposed 360 x 360 mosaic camera model assumes that for the realization with the mirror, a circle on the mirror is projected onto a circle in the image acquired by the CCD camera. This can be achieved when the mirror base is parallel to the image plane and the camera has squared pixels or when the camera is calibrated. When a fish eye lens is used, the central camera has to be calibrated first. For the development of the following theory we assume that the 2D CCD camera and the corresponding optics were calibrated and thus the 1D omnidirectional camera was calibrated too.

(a) (b)

**Figure 2. Experimental setup. (a) telecentric camera and a conical mirror, (b) central camera with Nikon fish eye converter.**

## 2. 3D Reconstruction from 360 x 360 Mosaic

### 2.1. 360 x 360 Mosaic Camera Model

A central camera's image coordinates $(u, v)$ directly produce a vector $\mathbf{p} = (u, v, 1)$ representing the corresponding light ray in some camera centered coordinate system with a center $C = (0, 0, 0)$. This construction cannot be utilized in case of the 360 x 360 mosaic camera, where the light rays do not intersect at one common point. Some representation of general lines in 3D space has to be chosen. We describe the light ray by a point

$$C = (r\cos\alpha, r\sin\alpha, 0)$$

on the viewing circle $\mathcal{C}$ and a vector

$$\mathbf{p} = (-sin(\alpha)\cos(\beta), \cos(\alpha)\cos(\beta), \sin(\beta)) \ .$$

Angles $\alpha$ and $\beta$ are shown in Figures 3 and 4.

For each point in the scene, outside the viewing circle, there are two light rays which are tangent to $\mathcal{C}$ and intersect in a scene point $P$, see Figure 3. There are three angles which define each pair of light rays intersecting in one scene point $P$: $\alpha$, $\alpha'$ and $\beta$, see Figures 3 and 4. Provided with the angles identifying the light rays, 3D coordinates of the scene point $P$ can be computed. It can be observed in Figure 3 that the $z$ coordinate of the point depends only on the angle $\beta$ and $x$ and $y$ coordinates depend only on the angles $\alpha$, $\alpha'$. Therefore, a planar view of the 360 x 360 mosaic camera geometry in the $xy$ plane can be used to illustrate the computation of the $x$ and $y$ coordinates of the scene points, as it is depicted in Figure 4.

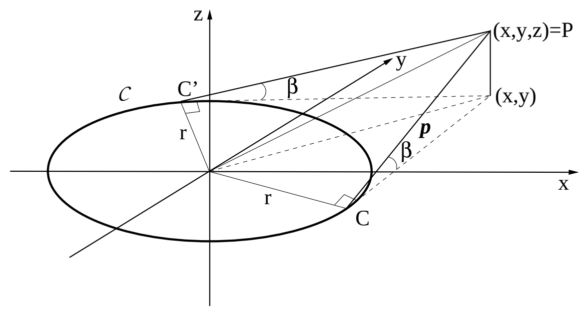


**Figure 3. 360 x 360 mosaic camera geometry.**

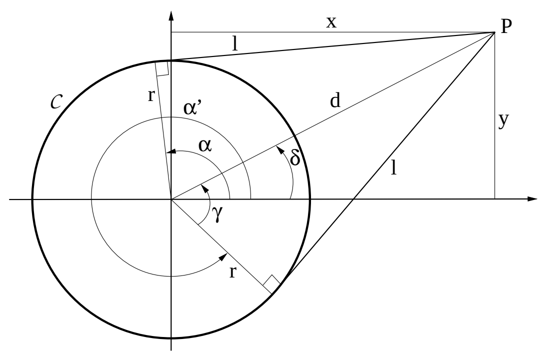


**Figure 4. 360 x 360 mosaic camera geometry - a view in the $\mathrm{xy}$ plane.**

Using basic trigonometric rules and identities, the following expressions for the $x$, $y$ and $z$ can be derived:

$$x = r\frac{\cos\delta}{\cos\gamma} \quad (1)$$

$$y = r\frac{\sin\delta}{\cos\gamma} \quad (2)$$

$$z = r\tan\gamma\tan\beta \ , \quad (3)$$

where $r$ is the radius of the viewing circle, $\gamma = \frac{\alpha-\alpha'}{2}$, and $\delta = \frac{\alpha+\alpha'}{2}$.

An important issue is the relation between the angles $\alpha$, $\alpha'$, and $\beta$ and the pixel coordinates $(u, v)$ in the mosaic images. To find it, two questions have to be answered: first, what is the relation between pixels and degrees and second, which pixels correspond to zero angles, i.e. from which image coordinates are the angles measured, see Figure 5.

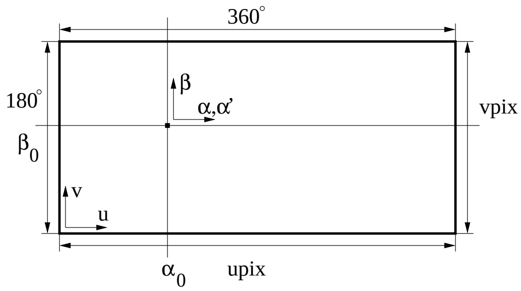


**Figure 5. Mosaic image coordinate system $(\mathbf{u}, \mathbf{v})$ and its relation to the 360 x 360 mosaic camera parameters - angles $\alpha$, $\alpha'$, and $\beta$ and their zero values $\alpha_0$ and $\beta_0$.**

To answer the first question is easy because our 1D omnidirectional camera as well as its motion are calibrated. The angles $\alpha$ and $\alpha'$ correspond to horizontal image axis, while the angle $\beta$ corresponds to the vertical axis. The images contain a full panoramic view of the scene in the horizontal direction and each contains half of the full panoramic view in the vertical direction. Therefore, under the assumption of a constant speed of rotation, the relation between degrees and pixels for the angles $\alpha$ and $\alpha'$ can be computed as $\frac{upix}{360}$ and for the angle $\beta$ as $\frac{vpix}{180}$, where $upix$ and $vpix$ denote the number of pixels in horizontal and vertical directions respectively.

The second question is more complex. The relation between pixel coordinates in the images $(u, v)$ and $(u', v)$ and the angles can be described as:

$$\alpha = \frac{360u}{upix} - \alpha_0 \tag{4}$$

$$\alpha' = \frac{360u'}{upix} - \alpha'_0 \tag{5}$$

$$\beta = \frac{180v}{vpix} - \beta_0 \ . \tag{6}$$

Note that $\alpha_0 = \alpha'_0$ which follows from the fact that unprimed as well as primed rays lie in the same plane $\pi$. The change of the zero $\alpha$ angle corresponds to a rotation of the 360 x 360 mosaic camera around the $z$ axis, see Figures 3 and 4. Therefore it can be incorporated into the extrinsic camera parameters.

As previously stated, each point outside the viewing circle $\mathcal{C}$ can be seen by two light rays. Therefore, two mosaic images can be acquired. In each rotation step, two sets of the light rays, belonging to the two mosaic images, are created by splitting the plane $\pi$ into two half planes by some line $l$ in $\pi$. If the line $l$ is perpendicular to the plane $\delta$, the images of the corresponding points are on the same rows in the two mosaic images. Such a pair of images is called a rectified pair. In a practical realization with a telecentric camera and a conical mirror, the light rays are imaged onto a circle which is split by the line $l'$, parallel to the vertical axis of the image and passing through the image of the tip of the cone, into two half circles, see Figure 7.

Due to the 360 x 360 mosaic camera setup, (the image plane of the CCD camera may be rotated inside the plane $\pi$) the line $l'$ does not have to be perpendicular to the plane $\delta$. This creates an angle $\beta'$ between the line $l'$ and a line $l$, see Figure 6. The images are not rectified, moreover the corresponding points may lie in the same image, as depicted in Figure 6(b). The same effect is obtained if the elevation angle $\beta$ is measured from a wrong zero value. Label I in Figure 6(a) denotes the correct division of rays by the line $l$, label II corresponds to a wrong division by $l'$. When dividing the rays into four parts, marked 1–4 in the figure, then, in the correct case, parts 1,2 and 3,4 compose the right eye and the left eye mosaic respectively. In the incorrect case, part 1 is moved to the top of the left eye mosaic and part 3 to the bottom of the right eye mosaic. Therefore, the point $A$ with the corresponding point $A'$ are both contained in the same mosaic image.

It is interesting that only one known point correspondence is required to estimate the value of the elevation angle $\beta_0$ that corresponds to the division by the line $l$. Denoting the elevation angles of the corresponding points $\beta_1$, $\beta_2$ and some incorrect zero elevation angle $\beta'_0$, the correct zero elevation angle $\beta_0$ can be computed as:

$$\beta_0 = \beta'_0 + \frac{(\beta_1 - \beta_2)}{2} \ . \tag{7}$$

Because the corresponding points should lie on the same image rows, the images have to be shifted in opposite directions until they are rectified. Rows which were moved from the top of one image will appear on top of the other image and symmetrically the same occurs with the bottom rows of the images. This is due to the fact that the 360 x 360 images have the topology of a torus [11]. After the rectification, it can be assumed that $\beta_0$ corresponds to the middle row of the image mosaics. A known $\beta_0$ angle is assumed in the following section.

### 2.2. 360 x 360 Mosaic Camera Calibration

In the following text, the term *360 x 360 mosaic camera calibration* refers to the estimation of unknown parameters of the proposed 360 x 360 mosaic camera model, the radius $r$ and a coordinate system change.

Provided with known coordinates of the scene points and their images, the 360 x 360 mosaic camera parameters can be estimated. It is convenient to express the relation between the scene point coordinates, the corresponding image coordinates, and the (unknown) 360 x 360 mosaic camera

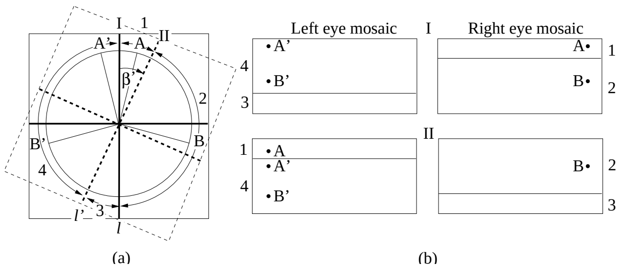


**Figure 6. Effect of a change of the zero elevation angle, see text for a description.**

parameters in the following form:

$$\mathbf{A}\mathbf{h} = 0 \ , \tag{8}$$

where matrix $\mathbf{A}$ contains known point coordinates and vector $\mathbf{h}$ is composed from the unknown 360 x 360 mosaic camera parameters. There is only one intrinsic parameter, the radius $r$, needed to be estimated in order to express light rays as vectors in a 360 x 360 mosaic camera centered coordinate system. A rigid motion transformation is then used to identify these vectors with vectors in a scene coordinate system.

Denoting coordinates of points in the 360 x 360 mosaic camera centered coordinate system as $\mathbf{x}_i = (x_i, y_i, z_i, 1)^T$, the relation between coordinates in the scene coordinate system $\tilde{\mathbf{x}}_i = (\tilde{x}_i, \tilde{y}_i, \tilde{z}_i, 1)^T$ can be generally expressed as:

$$\omega_i \mathbf{x}_i = \mathbf{H}\tilde{\mathbf{x}}_i \ . \tag{9}$$

The $4 \times 4$ matrix $\mathbf{H}$ represents both intrinsic and extrinsic 360 x 360 mosaic camera parameters and can be decomposed into matrices $\mathbf{K}$ and $\mathbf{M}$, $\mathbf{H} = \mathbf{K}\mathbf{M}$, where

$$\mathbf{K} = \begin{pmatrix} r & 0 & 0 & 0 \\ 0 & r & 0 & 0 \\ 0 & 0 & r & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{10}$$

contains 360 x 360 mosaic intrinsic camera parameters, the radius $r$, and the matrix

$$\mathbf{M} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \tag{11}$$

contains the extrinsic parameters, where the $3 \times 3$ matrix $\mathbf{R}$ stands for the rotation and the $3 \times 1$ vector $\mathbf{t}$ for the translation. Notice that in our case the fourth row of the matrix $\mathbf{H}$ equals $(0, 0, 0, 1)$, therefore we can write (8) after some rearranging in the form:

$$\begin{pmatrix} \tilde{\mathbf{x}}_i^T & \mathbf{0}^T & \mathbf{0}^T & -x_i \\ \mathbf{0}^T & \tilde{\mathbf{x}}_i^T & \mathbf{0}^T & -y_i \\ \mathbf{0}^T & \mathbf{0}^T & \tilde{\mathbf{x}}_i^T & -z_i \\ & & \vdots & \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ \vdots \\ h_{34} \\ 1 \end{pmatrix} = 0 \ . \tag{12}$$

This is just an expanded form of (8). A correct matrix $\mathbf{A}$ has rank 12, because the last column is a linear combination of the first, the fifth, and the ninth column. However, due to the presence of noise in real data, the matrix $\mathbf{A}$ may have rank 13. Unknown parameters forming the vector $\mathbf{h}$ can be found using standard methods, such as SVD.

### 2.3. Classes of Reconstructions

Another question, which should be answered, is what kinds or classes of reconstructions are possible when no calibration points are provided and whether something can be reconstructed at all. It can be observed from (1)-(3) that if the radius of the 360 x 360 mosaic camera is changed from $r$ to $r'$, then the coordinates of the reconstructed scene point $\mathbf{x}' = (x', y', z')$ can be computed from the original coordinates $\mathbf{x} = (x, y, z)$ as:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \frac{r'}{r} & 0 & 0 \\ 0 & \frac{r'}{r} & 0 \\ 0 & 0 & \frac{r'}{r} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \ . \tag{13}$$

When no point correspondence is provided and the images are not rectified, the error in estimation of the zero elevation angle $\beta_0$ leads to a nonlinear change in the $z$ coordinate of the reconstructed point. If the difference between $\beta_0$ and estimated zero elevation angle $\beta_0'$ is $\beta' = \beta_0' - \beta_0$, the following relation holds:

$$z' = d\frac{z - d\tan\beta'}{\tan\beta' z + d} \ , \tag{14}$$

where $d = \sqrt{x^2 + y^2 - r^2}$ as depicted in Figure 4. However, we can always rectify images when we can reconstruct a single point because a one point correspondence is sufficient for the rectification.

Also, the zero rotation angle $\alpha_0$ can be one of the parameters of the 360 x 360 mosaic camera. It determines the orientation of the 360 x 360 mosaic camera with respect to the 360 x 360 mosaic camera coordinate system. In this case, it's incorrect estimation results in a rotation around the $z$ axis by an angle $\alpha_0$. Therefore, it can be covered by a rigid motion transformation of the coordinates. Putting it all together, the transformation describing the relation between reconstructed points and points measured in the 360 x 360 mosaic camera centered coordinate system is just a scaling of all scene point coordinates (13). If the images are not rectified then a nonlinear transformation (14) of the $z$ coordinates of the points is added, however, only one point correspondence is needed in order to rectify the images.

In general, the 1D omnidirectional camera does not have to be calibrated and the motion does not have to be circular. Instead, they can be changed accordingly by an unknown projective transformation. Then the 360 x 360 mosaic camera will provide a projective reconstruction of the scene instead of a similarity reconstruction.

## 3. Experimental Results

A slice camera [6] was used in the experiments. The 360 x 360 mosaic camera was composed of a conical mirror with the apex angle 90° and a telecentric lens mounted on a standard 2D CCD camera. Both the 2D CCD camera and the mirror were mounted on a motorized turntable, see Figure 2(a), and 10 known calibration points were deployed in the scene.

The 2D CCD camera–mirror rig was rotated on a circular path with the mirror symmetry axis perpendicular to the axis of rotation. An image was captured at each rotation step and one circle of pixels was extracted from the image. The circle was split into two half circles which were straightened into two half slices composing the two mosaics, the left half for the left eye projection and the right half for the right eye projection, see Figure 7. The turntable was rotated with

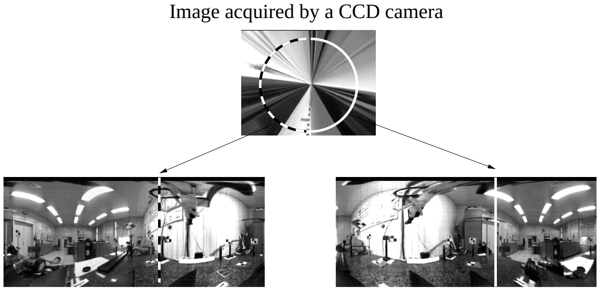


**Figure 7. Mosaic composition from slices.**

0.2° angular step which resulted in 1800 snapshots from which one circle, which was one pixel wide, was extracted. The resolution of the two panoramic images was 1800 pixels in the horizontal and 878 pixels in the vertical direction. The resulting mosaic images for the right eye and the left eye projection are shown in Figure 9 and Figure 10 respectively. The images are aligned so that the most distant points (the door on the back wall) lie in the same columns. Such an alignment allows the disparity in the image pairs to be clearly seen and was done only for visualization purposes.

A total number of 3 image pairs were selected from all experiments and images of the calibration points were detected manually. Each of these image pairs was acquired with a different radius of viewing circle. Then the scene coordinates of the calibration points were computed. Figure 8 shows the reconstructed points together with their coordinates measured manually in the scene coordinate system. Reconstruction error (the Euclidean distance between

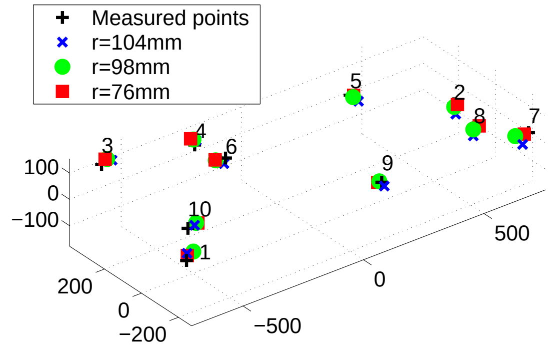


**Figure 8. Reconstructed points from three experiments with different radii compared with measured points.**

the points measured manually in the scene and coordinates of the scene points computed from manually detected image coordinates) for each calibration point is shown in Figure 11. Their respective point number is written below each point. The distance affects the resulting error. Generally, the longer the distance from the 360 x 360 mosaic camera center, the larger is the influence of the correspondence detection errors on the precision of the reconstruction. For example, one pixel error in the detection of points in the images corresponds to approximately 15 mm error for the nearest scene points and 35 mm error for the most distant points. In this experiment we did not explicitly calibrated the 1D omnidirectional camera because it involves the calibration of the whole setup. Instead we used a square pixel camera and carefully aligned all components of the setup. An explicitly calibrated 1D camera would also improve the precision of the reconstruction.

Figure 12 illustrates the reconstruction error before and after the 360 x 360 mosaic calibration. Radius $r = 104mm$ was measured by hand and used as the initial 360 x 360 mosaic camera parameter. Calibration matrix $\mathbf{H}$ was computed from the odd calibration points and results were verified on the even points. It can be observed that the overall size of the error decreased although some points have a bigger error than before the calibration. This may be due to the fact that only five points were used for the calibration and their error (caused by noise, 360 x 360 mosaic camera assembly etc.) influenced the result. More calibration points would improve the calibration precision.

## 4. Conclusion

A 360 x 360 mosaic camera model was presented in this paper. A calibration procedure was derived and a relation between reconstructions from uncalibrated views was de-

**Figure 9. Right eye mosaic.**

**Figure 10. Left eye mosaic.**

termined. It turns out that this 360 x 360 mosaic camera has a simpler model than a perspective camera because there is only one intrinsic parameter for a rectified pair of images and two otherwise. It is also shown that only one point correspondence is required for the rectification. The Relation between scene points reconstructed from an uncalibrated 360 x 360 mosaic camera and measured points in the scene is a similarity instead of a general projective transformation as it is the case of a perspective camera. The experimental results corrolate the presented theory.

## References


[1] S. Baker and S. K. Nayar. A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, 35(2):175–196, 1999.

[2] A. M. Bruckstein and T. J. Richardson. Omniview cameras with curved surface mirrors. In *IEEE Workshop on Omnidirectional Vision (OMNIVIS'00)*, pages 79–84, June 2000.

[3] J. Chai, S. B. Kang, and H.-Y. Shum. Rendering with non-uniform approximate concentric mosaics. In *Second Workshop on Structure from Multiple Images of Large Scale Environments, SMILE*, July 2000.

[4] C. Geyer and K. Daniilidis. A unifying theory for central panoramic systems and practical implications. In D. Vernon, editor, *European Conference on Computer Vision ECCV 2000*, volume 2, pages 445–462, June-July 2000.

[5] R. A. Hicks and R. Bajcsy. Catadioptric sensors that approximate wide-angle perspective projections. In *IEEE Workshop on Omnidirectional Vision (OMNIVIS'00)*, pages 97–103, June 2000.

[6] S. K. Nayar and A. Karmarkar. 360 x 360 mosaics. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'00)*, volume 2, pages 388–395, June 2000.


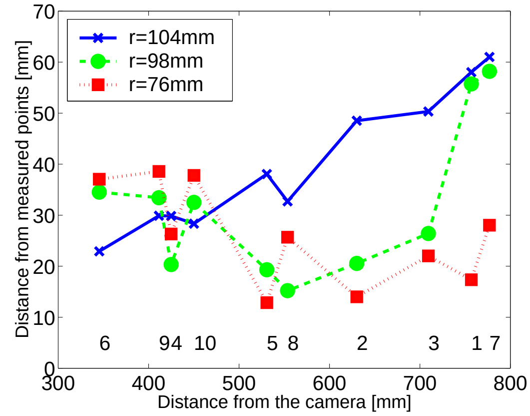


**Figure 11. Reconstruction error, see text for a description.**

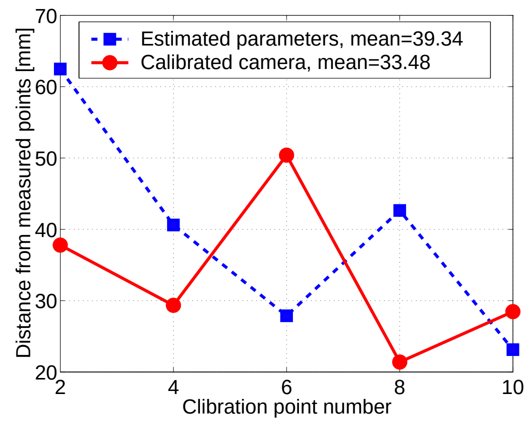


**Figure 12. Reconstruction before (radius $r = 104mm$ was measured by hand) and after the 360 x 360 mosaic calibration.**


[7] T. Pajdla. Epipolar geometry of some non-classical cameras. In B. Likar, editor, *Proceedings of Computer Vision Winter Workshop*, pages 223–233, February 2001.

[8] T. Pajdla, T. Svoboda, and V. Hlaváč. Epipolar geometry of central panoramic cameras. In R. Benosman and S. B. Kang, editors, *Panoramic Vision : Sensors, Theory, and Applications*. Springer Verlag, Berlin, Germany, 1st edition, 2000.

[9] S. Peleg, M. Ben-Ezra, and Y. Pritch. Omnistereo: Panoramic stereo imaging. *PAMI*, 23(3):279–290, March 2001.

[10] S. Seitz. The space of all stereo images. In J. Little and D. Lowe, editors, *ICCV'01: Proc. Eighth IEEE International Conference on Computer Vision*, volume 1, pages 26–35, July 2001.

[11] H.-Y. Shum, A. Kalai, and S. M. Seitz. Omnivergent stereo. In *Proc. of the International Conference on Computer Vision (ICCV'99)*, volume 1, pages 22–29, September 1999.