

Omnidirectional Camera Model and Epipolar Geometry Estimation by RANSAC with Bucketing*

Branislav Mičušík and Tomáš Pajdla

Center for Machine Perception, Dept. of Cybernetics, Czech Technical
University, Karlovo nám. 13, 121 35 Prague, Czech Republic
{micusb1,pajdla}@cmp.felk.cvut.cz

Abstract. We present a robust method of image points sampling used in RANSAC for a class of omnidirectional cameras (view angle above 180°) possessing central projection to obtain simultaneous estimation of a camera model and epipolar geometry. We focus on problem arising in RANSAC based estimation technique for omnidirectional images when the most of correspondences are established near the center of view field. Such correspondences satisfy the camera model for almost any degree of an image formation non-linearity. They are often selected in RANSAC as inliers, estimation stops prematurely, the most informative points near the border of the view field are not used, and incorrect camera model is estimated. We show that a remedy to this problem is achieved by not using points near the center of the view field circle for camera model estimation and controlling the points sampling in RANSAC. The camera model estimation is done from image correspondences only, without any calibration objects or any assumption about the scene except for rigidity. We demonstrate our method in real experiments with high quality but cheap and widely available Nikon FC-E8 fish-eye lens.

1 Introduction

Recently, high quality, but cheap and widely available, lenses, e.g. Nikon FC-E8 or Sigma 8mm-f4-EX fish-eye converters, and curved mirrors, e.g. [15], providing the view angle above 180° have appeared. Cameras with so large view angle, called omnidirectional cameras, are especially appropriate in applications, (e.g. surveillance, tracking, structure from motion, navigation, etc.) where more stable ego-motion estimation is required. Using such cameras in a stereo pair calls for searching of correspondences, camera model calibration, epipolar geometry estimation, and 3D reconstruction all that to be done analogically as for standard directional cameras [7].

In this work we show, see Fig.1, how the points should be sampled by a RANSAC estimation technique to obtain an unbiased simultaneous estimate of the

* This research was supported by the following projects: CTU 8306813, GAČR 102/01/0971, MSM 212300013, MŠMT KONTAKT 22-2003-04, BeNoGo IST-2001-39184.

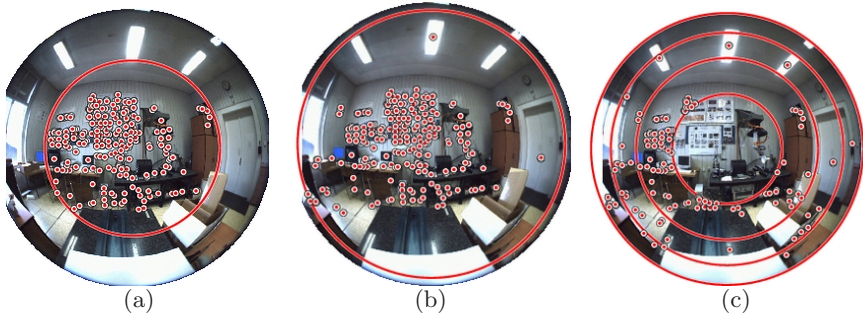


Fig. 1. Inliers detection. 1280×1024 images were acquired by Nikon FC-E8 fish-eye converter. Correspondences were obtained by [9]. (a) *Wrong* model. All points were used in model estimation by RANSAC. The model, however, suits only to the points near the center of the view field circle and other points were marked as outliers. (b) *Correct* model. Only points near the boundary of the view field circle were used for computing of the model. The model suits to points near the center as well as to points near the boundary. (c) Image zones used for points sampling in RANSAC for computing the correct model.

camera model and epipolar geometry for omnidirectional cameras with lenses as well for mirrors providing view angle above 180° and possessing central projection. We assume that point correspondences, information about the view field of the lens, and its corresponding view angle are available.

Previous work on the estimation of camera model with lens distortion includes methods that use some knowledge about the observed scene, e.g. calibration patterns [3, 13] and plumb line methods [4, 5, 16], or calibrate cameras from point correspondences only, e.g. [6, 14, 18].

Fitzgibbon [6] dealt with the problem of lens nonlinearity estimation in the context of camera self-calibration and structure from motion. His method, however, cannot be directly used for omnidirectional cameras with view angle above 180° because it represents images by points in which rays of a camera intersect an image plane. We extended in [10] the method [6] for omnidirectional cameras, derived appropriate omnidirectional camera model incorporating lens nonlinearity, and suggested an algorithm for estimating of the model from epipolar geometry.

The structure of the paper is the following. The omnidirectional camera model and its simultaneous estimation with epipolar geometry is reviewed in Section 2. The robust bucketing technique based on RANSAC for outliers detection is introduced in Section 3. Experiments and summary are given in Sections 4 and 5.

2 Omnidirectional Camera Model

For cameras with view angle above 180° , see Fig.2, images of all scene points \mathbf{X} cannot be represented by intersections of camera rays with a single image plane. For that reason, we will represent rays of the image as a set of unit vectors in \mathbb{R}^3 such that *one* vector corresponds just to *one* image of a scene point.

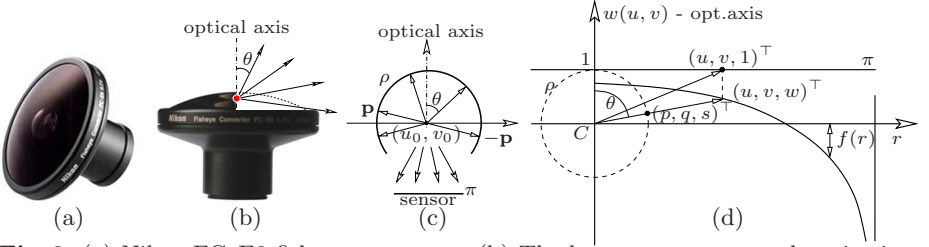


Fig. 2. (a) Nikon FC-E8 fish-eye converter. (b) The lens possesses central projection, thus all rays emanate from its optical center, which is shown as a dot. (c) Notice, that the image taken by the lens to the planar sensor π can be represented by intersecting camera rays with a spherical retina ρ . (d) The diagram of the construction of mapping f from the sensor plane π to the spherical retina ρ . The point $(u, v, 1)^\top$ in the image plane π is transformed by $f(\cdot)$ to $(u, v, w)^\top$ and then normalized to unit length, and thus projected on the sphere ρ .

Let us assume that $\mathbf{u} = (u, v)^\top$ are coordinates of a point in an image with the origin of the coordinate system in the center of the view field circle $(u_0, v_0)^\top$. Remember, that it is not always the center of the image. Let us further assume that a nonlinear function g , which assigns 2D image coordinates to 3D vectors, can be expressed as

$$g(\mathbf{u}) = g(u, v) = (u \ v \ f(u, v))^\top, \quad (1)$$

where $f(\mathbf{u})$ is rotationally symmetric function w.r.t. the point $(u_0, v_0)^\top$. Function f can have various forms determined by lens or mirror construction [3, 8]. For Nikon FC-E8 fish-eye lens we use the division model [10]

$$\theta = \frac{ar}{1 + br^2}, \quad r = \frac{a - \sqrt{a^2 - 4b\theta^2}}{2b\theta}, \quad (2)$$

where θ is the angle between a ray and the optical axis, and $r = \sqrt{u^2 + v^2}$ is the radius of a point in the image plane w.r.t. $(u_0, v_0)^\top$, and a, b are parameters of the model. Using $f(\mathbf{u}) = \frac{r}{\tan \theta}$, see Fig.2, 3D vector \mathbf{p} with unit length can be expressed up to scale as

$$\mathbf{p} \simeq g(\mathbf{u}) = \begin{pmatrix} \mathbf{u} \\ w \end{pmatrix} = \begin{pmatrix} \mathbf{u} \\ f(\mathbf{u}, a, b) \end{pmatrix} = \begin{pmatrix} \mathbf{u} \\ \frac{r}{\tan \theta} \end{pmatrix} = \begin{pmatrix} \mathbf{u} \\ \frac{r}{\tan \frac{ar}{1 + br^2}} \end{pmatrix}. \quad (3)$$

The equation (3) captures the relationship between the image point \mathbf{u} and the 3D vector \mathbf{p} emanating from the optical center towards a scene point.

2.1 Model estimation from epipolar geometry

Function $f(\mathbf{u}, a, b)$ in (3) is a two parametric non-linear function, which can be expanded to Taylor series with respect to a and b in a_0 and b_0 , see [10] for more details. The vector \mathbf{p} can be then written, using (3), as

$$\mathbf{p} \simeq \left[\begin{pmatrix} \mathbf{u} \\ f(\cdot) - a_0 f_a(\cdot) - b_0 f_b(\cdot) \end{pmatrix} + a \begin{pmatrix} \mathbf{0} \\ f_a(\cdot) \end{pmatrix} + b \begin{pmatrix} \mathbf{0} \\ f_b(\cdot) \end{pmatrix} \right] = \mathbf{x} + a\mathbf{s} + b\mathbf{t},$$

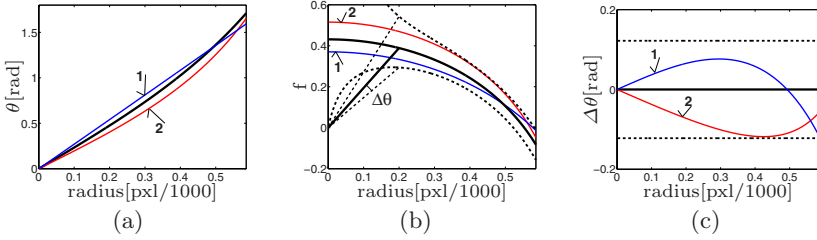


Fig. 3. Model fitting with a tolerance $\Delta\theta$. (a) The graph $\theta = f(r)$ for ground truth data (black thick curve) and two models satisfying the tolerance (red and blue curves). Parameters a and b can vary for models satisfying the tolerance. (b) The area between dashed curves is determined by the error. In this area, all models satisfying the tolerance must lie. (c) The angular error for both models with respect to the ground truth.

where \mathbf{x} , \mathbf{s} , and \mathbf{t} are known vectors computed from image coordinates, a and b are unknown parameters, and f_a , f_b are partial derivatives of $f(\cdot)$ w.r.t. a and b . The epipolar constraint for vectors \mathbf{p}' in the left and \mathbf{p} in the right image that correspond to the same scene point reads as

$$\begin{aligned} \mathbf{p}'^\top \mathbf{F} \mathbf{p} &= 0 \\ (\mathbf{x}' + a\mathbf{s}' + b\mathbf{t}')^\top \mathbf{F} (\mathbf{x} + a\mathbf{s} + b\mathbf{t}) &= 0 \end{aligned}$$

After arranging of unknown parameters into the vector \mathbf{h} we obtain

$$(\mathbf{D}_1 + a\mathbf{D}_2 + a^2\mathbf{D}_3)\mathbf{h} = 0, \quad (4)$$

where matrices \mathbf{D}_i are known [10] and vector \mathbf{h} is

$$\mathbf{h} = [f_1 \ f_2 \ f_3 \ f_4 \ f_5 \ f_6 \ f_7 \ f_8 \ f_9 \ bf_3 \ bf_6 \ bf_7 \ bf_8 \ bf_9 \ b^2f_9]^\top$$

with f_i being elements of the fundamental matrix. Equation (4) represents the Quadratic Eigenvalue Problem (QEP) [2, 17], which can be solved by MATLAB using the function `polyeig`. Parameters a , b , and matrix \mathbf{F} can be thus computed simultaneously.

We recover parameters of model (3), and thus angles between rays and the optical axis, which is equivalent to recovering an essential matrix, and therefore *calibrated* camera. We used angular error, i.e. angle between a ray and its corresponding epipolar plane [12], to measure the quality of the estimate of epipolar geometry.

Knowing that the field of view is circular, the view angle equals θ_m , the radius of the view field circle equals R , and from (2), parameter a can be then expressed as $a = \frac{(1+bR^2)\theta_m}{R}$. Thus (3) can be linearized to a *one parametric* model, and a 9-points RANSAC as a pre-test to detect most of outliers can be used like in [6]. To obtain better estimate, *two parametric* model with a priori knowledge $a_0 = \frac{\theta_m}{R}$, $b_0 = 0$, can be used in the 15-points RANSAC estimation.

3 Using Bucketing in RANSAC

There are outliers and noise in correspondences, and therefore we need to use a robust estimation technique, e.g. RANSAC [7], for model estimation and outliers

detection. We propose a strategy for point sampling, similar to bucketing [19], in order to obtain a good estimate in a reasonable time.

As it was described before, angle between a ray and its corresponding epipolar plane is used as the criterion of the estimation quality, call it the angular error. Ideally it should be zero but we admit some tolerance in real situations. The tolerance in the angular error propagates into the tolerance in camera model parameters, see Fig.3. The region, in which lie models that satisfy a certain tolerance $\Delta\theta$ is narrowing with increasing the radius of points in the image, see Fig.3b. Since $f(0) = \frac{1}{a}$ [10], the points near the center $(u_0, v_0)^\top$ will affect only parameter a . There is a large tolerance in parameter a because the tolerance region near the center $(u_0, v_0)^\top$ is large. If the majority of points is near the center, RANSAC finds the model with high number of inliers there and stops prematurely, see Fig.1a, because RANSAC is stopped if the probability of finding more inliers [7] drops below a defined threshold, usually 5%. On the other hand, there may exist model with more inliers that suit to the points near the center as well as to the points near to the boundary of the view field circle, see Fig.1b.

In [11], the division model (2) is fitted to the ground truth data, compared with other commonly used models, and it is analyzed how many points are needed and where they should be located in the image to fit the model from a minimum subset of points with a sufficient accuracy. As it is shown there, points near the center $(u_0, v_0)^\top$ have no special contribution to the final model fitting and the most informative points lie near the boundary of the view field circle.

Therefore, to obtain the correct model, it is necessary to exclude points near the center $(u_0, v_0)^\top$ a priori from the RANSAC sampling. The rest of the image, as Fig.1c shows, is split into three zones with equal areas from which the same number of points are randomly chosen by RANSAC. This helps to avoid the degenerate configurations, strongly biased estimates, and it decreases the number of RANSAC iterations.

The zone sampling can be replaced by a more general method if *i*) it is assured that probability of choosing a point near the boundary is larger than the probability of choosing a point near the center $(u_0, v_0)^\top$ and *ii*) the selected points are not close to each other. The method built on this assumptions is more robust but due to its complexity is out of range of this paper. The zone method is simple, leads to good results and is sufficient for most of practical applications.

As it was mentioned before, our model can be reduced to a *one-parametric* model using $a = \frac{(1+bR^2)\theta_m}{R}$, where R is radius¹ corresponding to the maximum view angle θ_m ². A priori known values R and θ_m fix all models with various b to the point $[R \ \theta_m]$, see [11] for more details. The resulting model has only one degree of freedom and thus smaller possibility to fit outliers. Using the approximate knowledge of a reduces the minimal set to be sampled by RANSAC from 15 to 9 correspondences and makes sampling faster. It is natural to use the 9-points RANSAC as a pre-test that excludes most disturbing outliers before the full, and more accurate, 15-points RANSAC is applied.

¹ R can be obtained by circle fitting to the view field boundary in the image.

² It is provided by lens manufacturer.

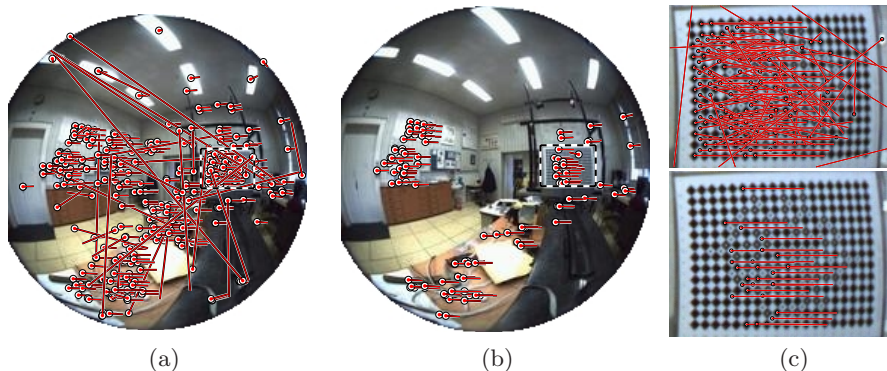


Fig. 4. Using of the proposed method for outliers detection in omnidirectional images. (a) Tentative correspondences between a pair of the omnidirectional images found by technique [9]. Circles mark points in the first image, lines join them to their matches in the next one. (b) Detected inliers. (c) A detail view of an object with repetitive texture. Top: input tentative correspondences. Bottom: inliers detected by our method.

4 Experiments

In this section, the proposed method is applied to real data. First, it is shown how the method succeeds in outliers detection, see Fig.4. Tentative correspondences were subjected to the 9-points followed by the 15-points RANSAC in order to detect inliers. It is difficult for matching algorithms that do not use the correct omnidirectional camera model and epipolar geometry to find correct matches between images, especially in images with repetitive texture, as Fig.4c shows. As a remedy to this problem the proposed method can be used, camera model parameters, epipolar geometry and outliers can be obtained.

Next experiments show camera models parameters and cameras trajectories (up to the magnitudes of translation vectors) estimation, i.e. structure from motion. Relative camera rotations and directions of translations used for trajectories estimations were computed from essential matrices [7]. For obtaining the magnitudes of the translation vectors we would need to reconstruct the observed scene. It was not the task of this paper. Instead, we assumed the unit length of translation vectors. Correspondences were obtained by the commercial program *boujou* [1] in this experiment.

The first “structure from motion” experiment shows a rotating omnidirectional camera, see Fig.5. The camera was mounted on a turntable such that the trajectory of its optical center was circular. Images were acquired every 10° , 36 images in total. First, one \bar{a} and one \bar{b} were estimated as the median of all a 's, b 's computed for every consecutive pair of images in the whole sequence. The 9-points RANSAC as a pre-test to detect most of outliers and then the 15-points RANSAC were performed to compute parameters a , b for every pair. Matrices F were then computed for each pair using the same \bar{a} , \bar{b} . Fig.5 shows the trajectory computed from the estimated essential matrices.

The next experiment calibrates the omnidirectional camera from its translation in the direction perpendicular to its optical axis, see Fig.6a. The average

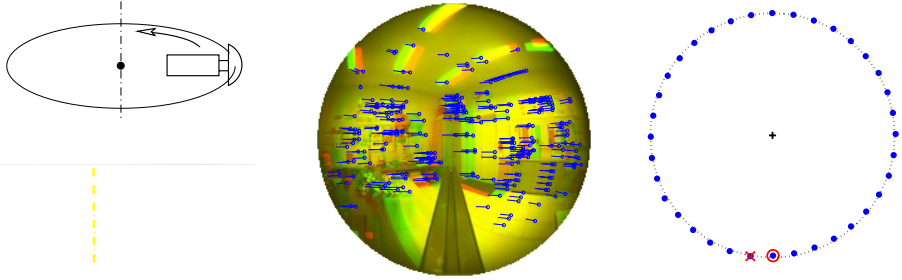


Fig. 5. Left: Nikon FC-E8 fish-eye converter mounted on the PULNIX TM1001 digital camera with resolution 1017×1008 pixels is rotated along a circle. The experiment by courtesy of J.Šivic. Middle: Correspondences between two consecutive images. Circles mark points in first image, lines join them to the matches in the next one. The images are superimposed in red and green channel. Right: Estimated trajectory.

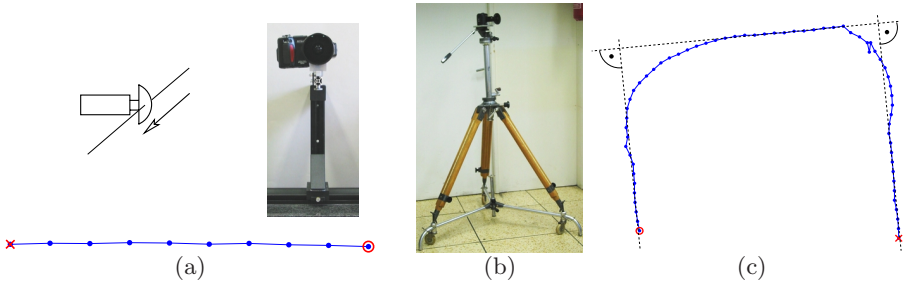


Fig. 6. (a) Side motion of the camera. The real setup is shown on the top. The estimated trajectory is depicted on the bottom. (b) The setup of the general motion experiment. (c) The estimated trajectory with starting and ending directions that are mutually parallel.

angular differences between estimated and true motion directions for every pair is 0.4° .

The last experiment shows the calibration from a general motion in a plane. Fig.6b shows a mobile tripod with a camera. Fig.6c shows an estimated U-shaped trajectory with right angles. Discontinuities of the trajectory are caused by hand driven motion of the mobile tripod. Naturally it has no effect on the final estimate and the final trajectory has really right angles.

5 Conclusion

The paper presented robust simultaneous estimation of the omnidirectional camera model and epipolar geometry. As the main contribution, the paper shows how the points should be sampled in RANSAC to avoid degenerate configurations and biased estimates. It was shown that the points near the center of the view field circle can be discarded and the final model computed only from points near the boundary of the view field circle. The suggested technique allows to incorporate an omnidirectional camera model into a 9-points RANSAC followed by the

15-points RANSAC for camera model, essential matrix estimation, and outliers detection.

Real experiments suggest that our method is useful for structure from motion with sufficient accuracy as a starting point for a bundle adjustment.

References

1. 2d3 Ltd. *Boujou*. 2000. <http://www.2d3.com>.
2. Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems : A Practical Guide*. SIAM, Philadelphia, 2000.
3. H. Bakstein and T. Pajdla. Panoramic mosaicing with a 180° field of view lens. In *Proc. of the IEEE Workshop on Omnidirectional Vision*, pages 60–67, 2002.
4. C. Bräuer-Burchardt and K. Voss. A new algorithm to correct fish-eye- and strong wide-angle-lens-distortion from single images. In *Proc. ICIP*, pages 225–228, 2001.
5. F. Devernay and O. Faugeras. Automatic calibration and removal of distortion from scenes of structured environments, July 1995.
6. A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proc. CVPR*, volume 1, pages 125–132, 2001.
7. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2000.
8. J. Kumler and M. Bauer. *Fisheye lens designs and their relative performance*. <http://www.coastalopt.com/fisheyep.pdf>.
9. J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *Proc. of the British Machine Vision Conference*, volume 1, pages 384–393, UK, Sep. 2002. BMVA.
10. B. Mičušík and T. Pajdla. Estimation of omnidirectional camera model from epipolar geometry. In *Proc. CVPR*. IEEE Computer Society, June 2003.
11. B. Mičušík and T. Pajdla. Using RANSAC for omnidirectional camera model fitting. In *Proc. of the 8th Computer Vision Winter Workshop*, pages 153–158, Valtice, Czech Republic, Feb. 2003. Czech Pattern Recognition Society.
12. J. Oliensis. Exact two-image structure from motion. *PAMI*, 24(12):1618–1633, 2002.
13. S. Shah and J. K. Aggarwal. Intrinsic parameter calibration procedure for a (high distortion) fish-eye lens camera with distortion model and accuracy estimation. *Pattern Recognition*, 29(11):1775–1788, November 1996.
14. G. P. Stein. Lens distortion calibrating using point correspondences. In *Proc. CVPR*, pages 602–609, 1997.
15. T. Svoboda and T. Pajdla. Epipolar geometry for central catadioptric cameras. *International Journal of Computer Vision*, 49(1):23–37, August 2002.
16. R. Swaminathan and S. K. Nayar. Nonmetric calibration of wide-angle lenses and polycameras. *PAMI*, 22(10):1172–1178, 2000.
17. F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Review*, 43(2):235–286, 2001.
18. Y. Xiong and K. Turkowski. Creating image-based VR using a self-calibrating fisheye lens. In *Proc. CVPR*, pages 237–243, 1997.
19. Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78(1-2):87–119, 1995.