

Towards high-precision lens distortion correction

Rafael Grompone von Gioi, Pascal Monasse, Jean-Michel Morel, Zhongwei Tang

▶ **To cite this version:**

Rafael Grompone von Gioi, Pascal Monasse, Jean-Michel Morel, Zhongwei Tang. Towards high-precision lens distortion correction. ICIP 2010, Sep 2010, Hong Kong, China. pp.4237-4240, 10.1109/ICIP.2010.5651928 . hal-00654417

HAL Id: hal-00654417
https://enpc.hal.science/hal-00654417

Submitted on 22 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TOWARDS HIGH-PRECISION LENS DISTORTION CORRECTION

R. Grompone von Gioi[†]*, P. Monasse*[‡]*, J.-M. Morel*[†] *and Z. Tang*[†]

[†]CMLA, ENS-Cachan, France [‡]IMAGINE, LIGM-Université Paris Est, France

ABSTRACT

This paper points out and attempts to remedy a serious discrepancy in results obtained by global calibration methods: The re-projection error can be rendered very small by these methods, but we show that the optical distortion correction is far less accurate. This discrepancy can only be explained by internal error compensations in the global methods that leave undetected the inadequacy of the distortion model. This fact led us to design a model-free distortion correction method where the distortion can be any image domain diffeomorphism. The obtained precision compares favorably to the distortion given by state of the art global calibration and reaches a RMSE of 0.08 pixels. Nonetheless, we also show that this accuracy can still be improved.

Index Terms— Planar textured pattern, virtual pinhole camera, nonparametric, high-precision, lens distortion

1. INTRODUCTION

This paper presents a small step forward in a research programme whose aim is to devise a highly accurate camera calibration method. By highly accurate, we mean that the residual error between the camera and its numerical model obtained by calibration should be far smaller than the pixel size. At first sight, this problem seemed to have been solved adequately by recent global calibration methods. The celebrated Lavest et al. method [13] measures the non-flatness of a pattern and yields a remarkably small re-projection error of about 0.02 pixels, which outperforms the precision of other methods. The experiments described below will actually confirm this figure. For the goals of computer vision, this precision would be more than sufficient. Yet, this paper describes a seriously discrepant accuracy measurement contradicting this hasty conclusion. As in [1] we shall use the following fact:

Theorem 1[6] *a camera follows the pinhole model if and only if the projection of every straight line in space onto the camera is a line.*

Following this theorem, the accuracy criterion used herewith directly measures the straightness of straight lines. We shall see that this straightness criterion gives a RMSE in the order of 0.2 pixel, which contradicts the 0.02 pixel re-projection accuracy. This significant discrepancy means that, in the global optimization process, errors in the external and internal camera parameter are being compensated by opposite errors in the distortion model. Thus, an inaccurate distortion model can pass undetected. Such facts raise a solid objection to global calibration methods, which estimate simultaneously the lens distortion and the camera parameters. This paper reconsiders the whole calibration chain and examines an alternative way to guarantee a high accuracy. A useful tool toward this goal will be proposed and carefully tested. It is a direct nonparametric distortion correction method. By nonparametric, we mean that the distortion model allows for any diffeomorphism.

We shall follow the usual assumption that a real camera deviates from the ideal pinhole model [9] by a lens distortion [3]. Thus, distortion correction is a fundamental step in multiview geometry applications such as 3D reconstruction. The above mentioned error measurement discrepancy may explain why three categories of distortion correction methods still coexist: classic pattern-based methods [19, 20, 13], plumb lines methods [6, 11, 17, 1] whose goal is to correct the distortion only, and methods based on enlarged epipolar geometry [14, 8, 12, 18]. Yet, these methods are all parametric and depend on the *a priori* choice of a distortion model with a fixed number of parameters. This *per se* is a drawback: such calibration methods require several trials and a manual model selection. Most methods assume a radial distortion modeled as a low-order even polynomial [19, 16] (models with more parameters are needed for wide-angle or fish-eye lens [2, 6]).

The proposed distortion correction method does not belong to any of the three categories. Indeed, it is nonparametric, non-iterative and model-free. Like most methods in the second and third category, the method decouples the distortion estimation from the calibration of camera internal parameters, thus avoiding any error compensation between them [10, 14].

Our plan is as follows. Section 2 gives the necessary definitions of the real camera and the pinhole model. It explains why a distortion correction up to a homography is sufficient for 3D applications, and defines the concept of *virtual pinhole camera*. The proposed nonparametric distortion correction is detailed in section 3, and is followed by experimental results in section 4. The last section 5 discusses how the high accuracy quest could be pursued.

2. THE VIRTUAL PINHOLE CAMERA

The classic camera model is

$$\hat{C} := \mathcal{D}KR[I \mid -T] \tag{1}$$

where T is a 3D point representing the optic center of camera, R is a 3D rotation representing camera orientation, K is a 3×3 upper triangular matrix containing the camera internal

parameters, and $\mathcal{D}$ is a diffeomorphism of the image domain representing the non-linear lens distortion. By pinhole camera model we mean a distortion-free camera model

$$C := KR[I \mid -T]. \tag{2}$$

The nonparametric method will estimate the distortion up to an arbitrary invertible homography H: $\tilde{\mathcal{D}} = \mathcal{D}H$. The correction precision evaluation will be based on the straightness of corrected lines, which is preserved by any homography. Applying $\tilde{\mathcal{D}}^{-1}$ on $\hat{C}$, yields

$$\begin{aligned}\tilde{C} &= \tilde{\mathcal{D}}^{-1}\mathcal{D}KR[I|-T] \\ &= H^{-1}\mathcal{D}^{-1}\mathcal{D}KR[I|-T] = H^{-1}KR[I \mid -T]. \end{aligned}\tag{3}$$

Thus, inverting the distortion on all images produced by the camera yields a new camera model which becomes pinhole. H, K being invertible, the decomposition $H^{-1}K = \tilde{K}R'$ is unique by QR decomposition. So after distortion correction (up to a homography) we have $\tilde{C} = \tilde{K}R'R[I \mid -T] = \tilde{K}\tilde{R}[I \mid -T]$, which we call the *virtual pinhole camera obtained after distortion correction*. The orientation and internal parameters of this virtual model do not match the physics of the actual camera, but yield a virtual pinhole camera that can be used to the very same purposes. Indeed, consider several positions of the physical camera inducing as many camera models $C_i = \mathcal{D}KR_i[I \mid -T_i]$. Applying the correction $\tilde{\mathcal{D}}^{-1}$ to all images obtained from these camera positions yields virtual pinhole cameras $\tilde{C}_i = \tilde{K}\tilde{R}_i[I \mid -T_i]$, which maintains the same relative orientations. From these cameras the whole 3D scene can be reconstructed by standard methods, up to a 3D similarity.

3. NONPARAMETRIC DISTORTION CORRECTION

3.1. The experimental set up

A nonparametric method requires the use of a highly textured planar pattern obtained by printing a textured image and pasting it on a very flat object (a mirror was used in the experiments). Two photos of the pattern are taken by the camera in a fixed lens configuration (e.g., with fixed zoom and focus). Ideally, the whole captor must be covered by the whole pattern.

3.2. Feature points

The distortion is estimated (up to a homography) as the diffeomorphism mapping the original digital pattern to a photograph of it. This requires a dense registration, which is obtained by an improved version of the SIFT [15] method providing high-precision localization about 0.03 pixels [4]. Denote by I the original digital pattern, by P the printed pattern, and by v the photo of P. The set of SIFT points matching from I to v is denoted by I_v and the corresponding points in v by v_I. Since the pattern is planar, there is a planar homography H such that $\mathcal{D}HI_v = v_I$. Knowing $\mathcal{D}H$ permits to synthesize a virtual pinhole camera by applying $(\mathcal{D}H)^{-1}$ on $\hat{C}$, as shown in Eq. (3).

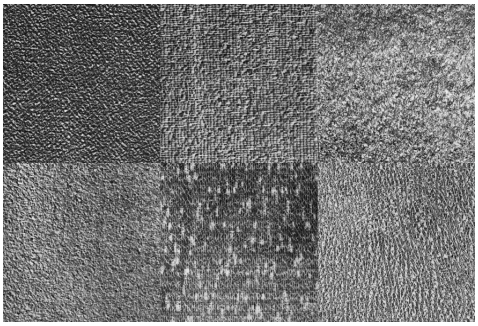


Fig. 1: *Digital pattern:* 1761×1174 *pixels.*

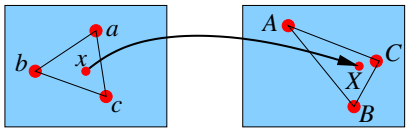


Fig. 2: *Local approximation of distortion by the affine transformation between corresponding Delaunay triangles. Point x is mapped to point X, by the affine transformation that maps triangle abc to triangle ABC.*

3.3. Triangulation and affine interpolation

The correspondences (I_v, v_I) actually only define the distortion field $\mathcal{D}H$ on the SIFT points I_v. The distortion being very smooth and the SIFT points dense enough, an affine interpolation is sufficient. This interpolation is performed after the image domain has been partitioned by a Delaunay triangulation of the SIFT points in v and I respectively (Fig. 2). Fig. 1 shows the texture pattern, selected to yield a maximal density of reproducible SIFT points at fine scales.

3.4. Outliers elimination: a loop validation

The few wrong SIFT matches (*outliers*) are nonetheless a serious problem. In our case, precisely because of the lens distortion, matching points are not related by a homography, and directly applying RANSAC [7] would not work. The problem is solved by a *loop validation* (Fig. 3). Consider two similar photos of the pattern, u and v (obtained by moving slightly the camera between two successive snapshots). With straightforward notation we have $u_I = \mathcal{D}H_u I_u$ and $v_I = \mathcal{D}H_v I_v$ (since the same camera and configuration are used, $\mathcal{D}$ does not change). The points v_I can be projected back on I by the distortion field from u to I, obtaining $I_{uv} = (\mathcal{D}H_u)^{-1}v_I$. It follows that I_v and I_{uv} are related by a homography (without distortion) because

$$I_{uv} = (\mathcal{D}H_u)^{-1}\mathcal{D}H_v I_v = H_u^{-1}H_v I_v. \tag{4}$$

This homography can be estimated by RANSAC algorithm and all the outliers not compatible with the homography are eliminated.

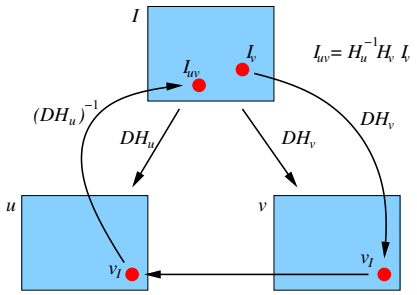
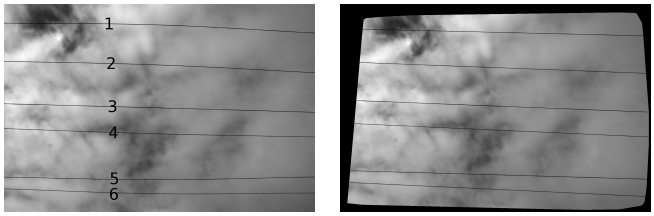


Fig. 3: *The loop validation used to remove outliers.*



(a) distorted image of tightly stretched strings

(b) corrected image by the nonparametric method

Fig. 5: *Distorted lines marked by numbers in (a) will be used to evaluate the precision after distortion correction.*

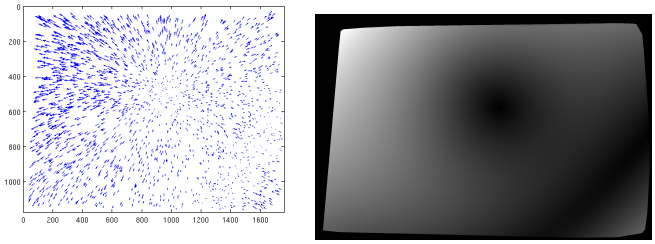
3.5. Algorithm summary

1. Take two slightly different photos of a textured planar pattern with constant camera settings;
2. apply high precision SIFT between the original digital pattern and both photographs;
3. eliminate outliers by the loop validation step;
4. interpolate the remaining matches to get a dense reverse distortion field;
5. by applying the reverse distortion field to all images produced by the real camera, the camera is converted into a virtual pinhole camera.

4. EXPERIMENTS

The experiments were made with a Canon EOS 30D reflex camera and an EFS $18-55$mm lens. The minimal focal length (18mm) was chosen to produce a fairly large distortion. The RAW images were demosaicked by summing up the four pixels of each 2×2 Bayer cell, obtaining a half-size image. Fig. 1 shows the digital pattern I that was used.

Fig. 4a shows a subsampling of the resulting distortion field, after the validation loop, and Fig. 4b shows the modulus of the interpolated distortion field on the discrete image domain. The distortion field is not circular symmetric, which is natural, the distortion being estimated up to an unknown homography.



(a) (b)

Fig. 4: *(a) The (subsampled) distortion field directly defined on "inlier" correspondences after the loop validation. (b) Modulus of the dense distortion field obtained by affine interpolation.*

To check the quality of the correction, we built a pattern with tightly stretched strings, that guarantees straightness, see Fig. 5a. The distortion is visible near the border of the image. Fig. 5b shows the image corrected by our nonparametric method. The lines numbered in Fig. 5a were used to evaluate the distortion error and compare it with the error left by the Lavest et al. algorithm [13]. On each corrected line, subpixel precision edge points were obtained by Devernay's algorithm [5]. Then, their regression line was computed and the RMS (root-mean-square) distance from each edge point to the line was used as the error measure. Table 1 shows the results. The proposed nonparametric method shows an improvement over the Lavest et al. method. But this difference is not just quantitative. The Lavest et al. result is somewhat final. Indeed, it already includes a correction of the non-flatness of the pattern while the nonparametric method does not.

RMSE (in pixels)		
	our method	Lavest method
line 1	0.095	0.170
line 2	0.066	0.083
line 3	0.067	0.100
line 4	0.059	0.078
line 5	0.079	0.089
line 6	0.126	0.178

Table 1: *Each distorted line marked by a number in Fig. 5a, is corrected either by the Lavest et al. method or by the proposed nonparametric method. The edge points are detected by Devernay's algorithm. The distortion error is computed as the root-mean-square distance (in pixels) from the edge points to their regression line.*

Figure 6 plots the straightness error of the proposed method along line number one (that is, the distance between edge points to the regression line). One can see small oscillations easily explained by noise and aliasing, but also a global tendency which can only be due to the non-flatness of the pattern. A parallel deterministic tendency observed on the other lines confirms this explanation. Fig. 7 shows that a flatness error of 100 μm (the thickness of a normal paper sheet) can produce the observed tendency. This non-flatness effect is stronger near the border of the image because the angle-of-view is larger. Simple physical measurements by the classic *ruler method* confirmed that the pattern showed a non-flatness of this amount.

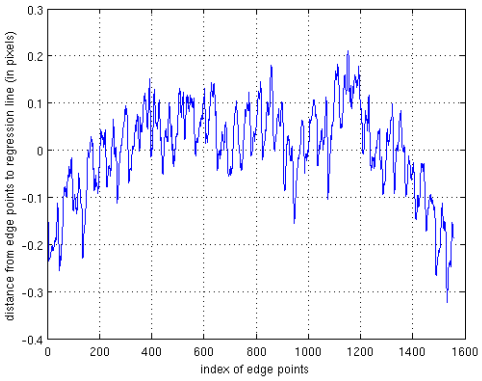


Fig. 6: *The distance in pixels from the edge points to their regression line on line 1 in Fig. 5a, after correction by the proposed method.*

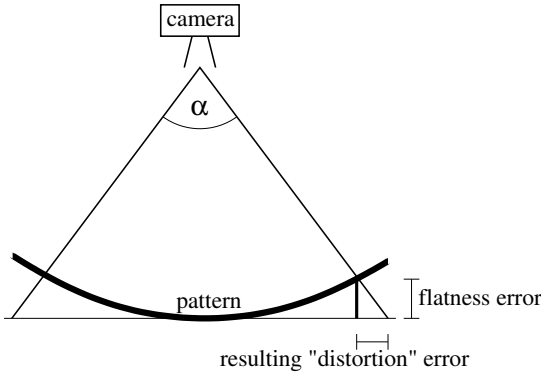


Fig. 7: *A flatness error in the pattern can be mistaken for a "distortion". In the experiments, with* $\alpha \approx 65^\circ$ *and a flatness error about 100 μm, the produced "distortion" error is about 64 μm (*flatness error $\times \tan \frac{\alpha}{2}$*). Our pattern has the size* 406×271 *mm and produces a* 1761×1174 *image, then one pixel corresponds to 230 μm. Thus the observed error would be approximately 0.3 pixel.*

5. DISCUSSION

The above experimental setting suggests two ways to eventually reach a still higher precision. The first way would simply be to use very flat patterns. But this raises the problem of making smart patterns. An alternative strategy was suggested in [1], using accurate straight objects, like the tightly stretched strings we already used here. Yet, it seems advisable to try to keep the strength and beauty of the Lavest et al. method, which is to estimate and correct the pattern's shape by the calibration process itself. An iterative method could be envisaged where, first, the distortion is corrected by the proposed nonparametric method and, second, the physical shape of the pattern is computed by the Lavest et al. method *with no distortion model*. Using this correction the distortion would be recomputed, and so on. This is, however, a complex process, which will require a heavier procedure and a mathematical analysis.

6. ACKNOWLEDGMENTS

The authors wish to thank Jean-Marc Lavest and Eric Royer (Université de Clermont-Ferrand) for lending us their calibration software and their gracious help. Part of this work was funded by the Agence Nationale de la Recherche, Callisto project (ANR-09-CORD-003).

7. REFERENCES

[1] L. Alvarez and J. R. Sendra, "An algebraic approach to lens distortion by line rectification," *Journal of Mathematical Imaging and Vision*, vol. 35, pp. 36–50, 2009.

[2] S. S. Brandt and J. Kannala, "A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses," *IEEE TPAMI*, vol. 28, pp. 1335–1340, 2006.

[3] D. C. Brown, "Close-range camera calibration," *Photogrammetric Engineering*, vol. 37, pp. 855–866, 1971.

[4] T. Buades, Y. Lou, J.-M. Morel and Z. Tang, "A note on multi-image denoising," *LNLA*, 2009.

[5] F. Devernay, "A Non-Maxima Suppression Method for Edge Detection with Sub-Pixel Accuracy," *INRIA Rapport de recherche*, No. 2724, November 1995.

[6] O. Faugeras, F. Devernay, "Straight lines have to be straight," *Mach. Vision Appli.*, vol. 13, pp. 14–24, 2001.

[7] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. Of the ACM*, vol. 24, pp. 381–395, 1981.

[8] A. Fitzgibbon, "Simultaneous linear estimation of multiple view geometry and lens distortion," *ICPR*, vol. 1, pp. 125–132, 2001.

[9] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.

[10] M. Herniou, J. Weng, P. Cohen, "Camera calibration with distortion models and accuracy evaluation," *IEEE TPAMI*, vol. 14, pp. 965–980, 1992.

[11] V. Hlaváč, T. Pajdla, T. Werner, "Correcting radial lens distortion without knowledge of 3-d structure," *Research Report, Czech Technical University*, 1997.

[12] Z. Kukelova, T. Pajdla, "Two minimal problems for cameras with radial distortion," *OMNIVIS*, 2007.

[13] J.-M. Lavest, M. Viala and M. Dhome, "Do We Really Need an Accurate Calibration Pattern to Achieve a Reliable Camera Calibration?," *ECCV*, vol. 1, pp. 158–174, 1998.

[14] H. Li, R. Hartley, "A non-iterative method for correcting lens distortion from nine point correspondences," *OmniVis*, 2005.

[15] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," *IJCV*, vol. 60(2), pp. 91–110, 2004.

[16] S. Ma, G. Wei, "Implicit and explicit camera calibration: theory and experiments," *IEEE TPAMI*, vol. 16, 1994.

[17] B. Prescott and G. F. Mclean, "Line-based correction of radial lens distortion," *Graphical Models and Image Processing*, vol. 59, pp. 39–47, 1997.

[18] G.P. Stein, "Lens distortion calibration using point correspondences," *CVPR*, vol. 602–608, 1997.

[19] R.Y. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *IEEE Journal of Robotics and Automation*, Vol. RA-3, 1987.

[20] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," *ICCV*, pp. 666–673, 1999.