

A Note on Principal Point Estimability *

Alberto Ruiz[†]

Pedro E. López-de-Teruel[‡]

Ginés García-Mateos[†]

[†]Dpto. de Informática y Sistemas, [‡]Dpto. de Ingeniería y Tecnología de Computadores
Facultad de Informática, Campus de Espinardo, Universidad de Murcia, Spain

aruiz@um.es, pedroe@dittec.um.es, ginesgm@um.es

Abstract

We provide elementary geometric arguments to show that the principal point of cameras with small to moderate field of view cannot be reliably estimated from natural, noisy images (the problem is ill-posed). We also show that in robot navigation and other noisy geometric vision applications the exact location of the principal point is irrelevant in practice. In these cases satisfactory metric structure can be recovered from minimal information by efficient algorithms using simplified camera models.

1. Introduction

Visual reconstruction is an ill-posed inverse problem. Projective geometry is a powerful theoretical framework which provides elegant and practical algorithms. But even in high precision applications there is uncertainty, and robust estimation must be extensively used. In noisy situations (e.g. robot navigation, where sub-pixel feature detection is meaningless) some kind of regularization is required. This can be achieved by constraining camera parameters. It makes no sense to keep the aspect ratio, the skew, or the principal point (pp) free through all the processing stages.

There is some controversy in the recent literature on the required precision of the intrinsic parameters in order to obtain acceptable reconstructions. For instance, in [5] we find: “Zero skew and unit aspect ratio are two assumptions well satisfied by modern cameras. And it is also well known that reconstruction is not sensitive to the position of the principal point [1]”. However, in [3] (p.461, in the context of auto-calibration), we find a more cautious statement: “An assumption of zero skew is quite natural and is a safe assumption for most imaging conditions. However, an assumption of known principal point is much less tenable”. In [1] a simple formula is proposed to extract focal lengths from the

fundamental matrix under the assumption of square pixel and known principal point. The method is considered seriously in [2] (p.560, proposition 11.11) and it is recommended as initialization for the case of non-constant intrinsic parameters (p.572). However, in [3] (p.481) they point out some degenerate cases and conclude that “Generally speaking, our opinion is that this method is of doubtful value as a means of computing focal lengths”.

Motivated by robot navigation applications, where metric structure must be reconstructed in real time from minimal information, in this paper we analyze the estimability and relevance of the principal point.

2. Principal Point Estimability

A camera can be considered as a device for the measurement of angles between optical rays. This is required for 3D reconstruction, essentially based on triangulation. In contrast with the ordinary protractor used at school, which measures angles directly over a *circle arc*, a camera measures angles using intersections with the image *plane*. The relation between angles and measurements in the “curved image” of the protractor is linear. However, in the camera the angle α between the ray defined by an image point x and the perpendicular reference position o (the pp) is given by the nonlinear expression $\alpha = \arctan((x - o)/f)$. As x is farther from the pp a given angular distance corresponds to a larger pixel distance. While the protractor is “isotropic”, the optical axis is a “special” direction for the camera: near the pp there is greater linearity between angles and image positions. Thus, the position of the pp is obviously required to correctly compute angles between pixels.

Directly or indirectly, the position of the pp can only be computed by discovering the region in the image where a fixed angular distance corresponds to a minimum pixel distance. Unfortunately, the derivative of $\arctan(x/f)$ is nearly constant along the whole pixel domain of cameras with moderate field of view. Fig. 1 shows that the tangent function is *numerically* indistinguishable from a linear function in a 60 degree range. As an example, a typical

*This work has been partially supported by the Spanish MCYT grants DPI-2001-0469-C03-01 and TIC-2000-1151-C07-03.

camera such as the Mitsubishi 300 E, used in many robotic applications, has only a 44 degree field of view.

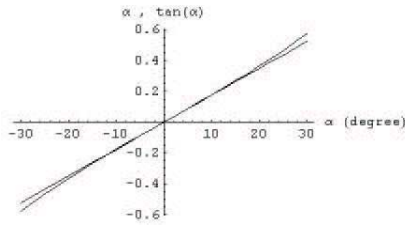


Figure 1. Comparison between $\tan(\alpha)$ and α .

It is clearly not feasible to accurately discover the minimum value of such derivative from noisy measurements (even if all the other camera parameters are perfectly known, which is not usually the case). When standard calibration algorithms [4] are applied to a fixed camera with different calibration sequences, they show acceptable stability of the focal lengths; but pp estimations have large variability¹. The problem is ill-conditioned. In fact, the pp is tightly coupled to camera orientation: the rotation origin corresponds to the optical axis. When the variations in the estimated pp are translated to angles we find that they are of the same order and consistent with (compensating) variations in the estimated camera orientation.

In cameras with a moderate field of view, any estimated pp is probably just an optimization artifact in the RQ factorization of the projection matrix.

3. Irrelevance of exact principal point location

Fortunately, the exact position of the pp is not required to compute an excellent approximation to the angle between any pair of directions. The angle defined by two given pixels x_1 and x_2 can be expressed up to second order by:

$$\begin{aligned} \alpha_1 - \alpha_2 &= \arctan\left(\frac{x_1 - o}{f}\right) - \arctan\left(\frac{x_2 - o}{f}\right) = \\ &= \frac{x_1 - x_2}{f} + O(x_i - o)^3 \end{aligned}$$

This means that the computed angle does not depend on the *exact* position of the pp if the optical rays are not very far from the optical axis (the error is of order 3). There is no quadratic component in the approximation, so nonlinearity between angles and pixels is not strong. For cameras with moderate field of view, this suggests that we can fix a nominal zero angle reference (e.g., in the image center) without the need of exact coincidence with the true pp .

¹ This phenomenon is recognized for instance in [2] (p.576, evaluating self-calibration): “We also note that the principal point is not estimated very accurately. However, as shown by the 3D reconstruction errors discussed next, this lack of accuracy is not of primary importance”.

Let us study, for different positions of the true pp , the difference between the true angle α , defined by two pixels o and x , and the angle α_e , estimated under the assumption that the pp is exactly at o (see Fig. 2).

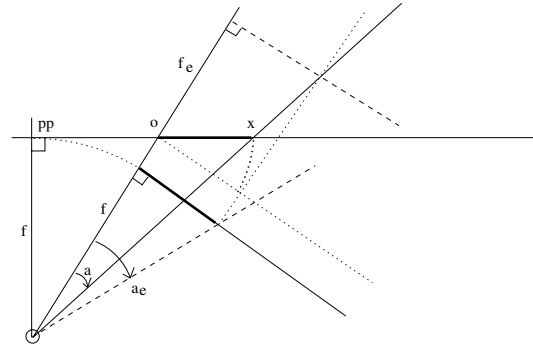


Figure 2. Angle measurement error produced by assuming that the true pp is at o (see text for details).

If the camera center and the focal length are kept fixed, the error in the pp is compensated by some rotation of the image plane and some pixel translation, to get approximate coincidence of rays. For a given pixel distance (thick line) the estimated angle α_e is greater when we believe that the pp is at o . Note that the angle can be correctly measured if an *effective focal* f_e is used (see below), but it depends on pixel position. In cameras with moderate field of view, small rotations are in practice indistinguishable from pixel translations, so it is expected that angular differences can be estimated without exact knowledge of the position of the pp . The true angle can be written as:

$$\alpha = \arctan\left(\frac{x - \varepsilon}{f}\right) + \arctan\left(\frac{\varepsilon}{f}\right)$$

where $\varepsilon = pp - o$ is the discrepancy between the true position of the pp and its assumed position o . The estimated angle is:

$$\alpha_e = \arctan\left(\frac{x}{f}\right)$$

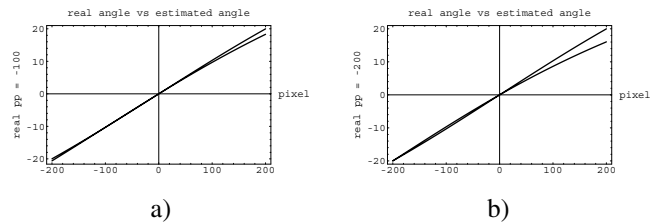


Figure 3. Comparison between α and α_e for (a) $\varepsilon = 10$ degree (100 pixel) and (b) $\varepsilon = 20$ degree (200 pixel). The image width is 400 pixel and the field of view is 40 degree.

The difference is negligible for $|\varepsilon| < 5$ degree. It is small for $|\varepsilon| < 10$ degree (1/4 of the whole image in Fig. 3.a). (Radial distortion probably produces greater errors.) The difference is noticeable for $|\varepsilon| > 20$ degree (pp outside the image in Fig. 3.b).

The analysis above assumes that the focal length is known and fixed. But in most cases it must be estimated, directly or indirectly, by optimization of correspondences between angles and pixels. A wrong pp can be compensated up to a certain extent by re-estimation of the the focal length. Actually, calibration algorithms compute an *effective* focal length f_e .² For a given error $\varepsilon = pp - o$, the effective focal f_e at a given pixel x can be expressed by:

$$f_e(x) = \frac{x}{\tan \left[\arctan \left(\frac{x-\varepsilon}{f} \right) + \arctan \left(\frac{\varepsilon}{f} \right) \right]} = \frac{\varepsilon^2 + f^2 - \varepsilon x}{f}$$

It is linear in x , so the average value for the whole image is taken at o , the center of the pixel array:

$$f_e = \frac{\varepsilon^2}{f} + f$$

This effective focal length uniformly distributes over the whole image the error between real angles and estimated angles, without regard of the pp true location. Fig. 4 compares the true angle and the angle estimated with f_e for different $\varepsilon = pp - o$. Now the maximum difference is not large even if the true pp is outside the image. This kind of “minimum error” result is expected from standard calibration algorithms.

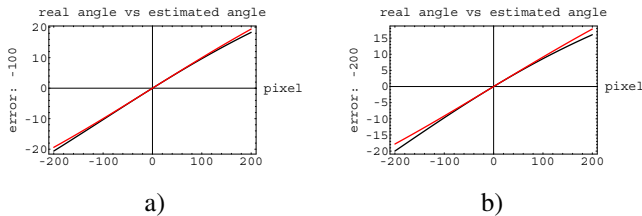


Figure 4. Comparison between α and α_e using f_e , for (a) $\varepsilon = 10$ degree (1/4 image). (b) $\varepsilon = 20$ degree (1/2 image)

In consequence, in cameras with small field of view (e.g., 20 degree) the pp is irrelevant in practice.

In cameras with moderate field of view (e.g., 40 degree), a nominal pp can be established at the image center, giving rise to negligible errors in applications with other uncertainty sources (e.g. robot navigation).

²If the pp is very far from the origin in one axis and close to it in the other, different effective focal lengths $f_{e,x}$ and $f_{e,y}$ could be required.

In wider fields of view, the principal point can be safely fixed at the estimation given by any standard calibration method. Errors in the pp of order of 20 degree are not relevant in robot navigation and other real world applications (see footnote 1).

4. Simplified Camera Model

For applications where high accuracy is not needed and cameras with moderate field of view (40-50 degree) we propose a simplified camera model K with just a single relevant intrinsic parameter: the *effective focal length* (common for both axes): $K = \text{diag}(f, f, 1)$ [5, 1].

The assumptions of zero skew and unit aspect ratio can be easily verified by checking if the image of a sphere is a circle. If it is (with high approximation), you can safely assume square pixel and nominal pp at the image center³. Reconstruction errors will be negligible in applications with other uncertainty sources.

The following subsections present experimental support for the proposed simplified camera model. In all cases the image size is 384×288 pixels and the horizontal field of view is approximately 40 degree.

4.1. Ordinary camera calibration

We can find 3D to 2D projective transformations, computed with different assumptions about the pp , which are indistinguishable in practice. The reprojection of points in the calibration object shown in Fig. 5, using a camera matrix estimated with a linear algorithm, produces a maximum error of 1.68 pixels. The principal point appears to be at pixel (202,147).

Fixing the pp at (150,100) and constraining to square pixel, the camera matrix estimated by Newton’s method produces a maximum reprojection error of 3.34 pixel. In the same conditions, with the pp at (250,200), the maximum error is 3.6 pixels (see Fig. 5.a). The arbitrary assumption of the pp in a 100 pixel range (1/4 of the image) around the estimated position produces some degradation but the results are still acceptable for many applications. Finally, if we fix the pp at a (100,50) we obtain the less satisfactory maximum reprojection error of 4.98 pixel.

4.2. Homography “Extrapolation”

Under the assumption of known pp and common focal length the full camera matrix can be obtained from calibration points in a single plane. This is particularly convenient for robot navigation, as explained in subsection 4.4. It is important to ascertain if this approach is robust against wrong

³Otherwise a linear transformation can be applied in the acquisition stage, in the same fashion as radial distortion correction.

assumptions about the pp . We have estimated a full projection matrix just from the “right” plane in the calibration device shown in Fig. 5, assuming the simplified camera K . If we fix the pp at the moderately wrong location (150,100), the reprojection on the “left” plane is still satisfactory (Fig. 5.b). Only when we fix the pp at an extremely wrong location (300,250) the degradation of extrapolated homography becomes unacceptable (Fig. 5.c).

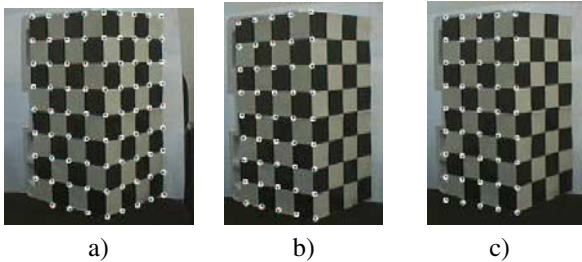


Figure 5. Reprojection results for illustrative arbitrary locations of the pp (see text for details).

4.3. Image Mosaicing

If the field of view is not large, small pan or tilt camera rotations can be satisfactorily modeled as affine transformations (essentially translations), without any projective component. This is illustrated in Fig. 6, where panoramic compositions using the projective warping:

$$H = \begin{pmatrix} 1.02528 & 0.0624325 & 74.2953 \\ -0.0517635 & 1.02853 & 8.24218 \\ 0.0000586617 & 0.0000186711 & 1. \end{pmatrix}$$

and affine warping:

$$A = \begin{pmatrix} 0.99749 & 0.0595305 & 75.6672 \\ -0.0598746 & 1.01175 & 10.2106 \\ 0. & 0. & 1. \end{pmatrix}$$

are virtually indistinguishable. The camera behaves almost as a protractor; there is no influence of the principal point.



Figure 6. Similar panoramic compositions from pure rotations using projective (left) and affine (right) warping.

4.4. Robot self-calibration

In robotic applications we find camera vibrations, irregularities in the ground and other noise sources. Reconstruction with millimetric precision is not usually necessary.

Simplified camera models are appropriate for robust navigation and self-location. Under reasonable assumptions, the full projection matrix allows analytical estimation of the effective focal and the extrinsic parameters from minimal information. Fig. 7.a shows relevant visual features detected in an indoor robot navigation scene. Just from one segment and point correspondence in three images (and odometric information) we can compute acceptable metric reconstructions of 3D space (Fig. 7.b) and self-localize the robot (Fig. 7.c).

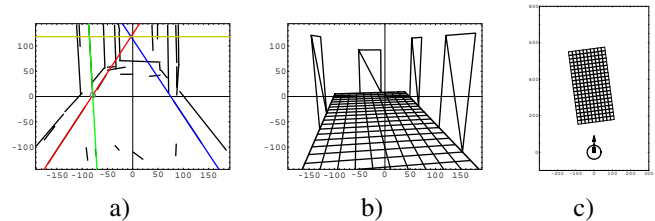


Figure 7. a) Relevant visual features in a structured robot navigation scene. b) conceptual reconstruction of a new scene and different view point. c) robot self-location. Common sense heuristics are used to infer the supporting plane of the visual features of interest.

5. Conclusion

The principal point of cameras with moderate field of view cannot be accurately estimated from noisy images. In robot navigation and related applications the principal point should not be considered as a free parameter. It can be safely fixed nominally at the image center. In such applications simplified camera models are extremely convenient, and do not increase the reconstruction errors originated by other, more serious sources of uncertainty.

References

- [1] S. Bougnoux. From projective to euclidean space under any practical situation. In *Proc. 6th Int. Conf. Comput. Vision*, pages 790–796, 1998.
- [2] O. Faugeras, Q.-T. Luong, and T. Papadopoulos. *The Geometry of Multiple Images*. MIT Press, 2001.
- [3] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [4] R. 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*, 3(4):323–344, 1987.
- [5] G. Xu, J.-I. Terai, and H.-Y. Shum. A linear algorithm for camera self-calibration, motion and structure recovery for multi-planar scenes from two perspective images. In *IEEE Conf. Computer Vision Pattern Recognition (CVPR)*, volume 2, pages 474–479, 2000.