



HAL
open science

Pose estimation of a single circle using default intrinsic calibration

Damien Mariyanayagam, Pierre Gurdjos, Sylvie Chambon, Vincent Charvillat, Florent Brunet

► **To cite this version:**

Damien Mariyanayagam, Pierre Gurdjos, Sylvie Chambon, Vincent Charvillat, Florent Brunet. Pose estimation of a single circle using default intrinsic calibration. Asian Conference on Computer Vision (ACCV 2018), Dec 2018, Perth, Australia. pp.1-6. hal-02092924

HAL Id: hal-02092924

<https://hal.science/hal-02092924>

Submitted on 8 Apr 2019

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.



Open Archive Toulouse Archive Ouverte

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible

This is an author's version published in:
<http://oatao.univ-toulouse.fr/22656>

To cite this version: Mariyanayagam, Damien and Gurdjos, Pierre and Chambon, Sylvie and Charvillat, Vincent and Brunet, Florent *Pose estimation of a single circle using default intrinsic calibration*. (2018) In: Asian Conference on Computer Vision (ACCV 2018), 2 December 2018 - 6 December 2018 (Perth, Australia).

Any correspondence concerning this service should be sent to the repository administrator: tech-oatao@listes-diff.inp-toulouse.fr

Pose estimation of a single circle using default intrinsic calibration

Damien MARIYANAYAGAM, Pierre GURDJOS,
Sylvie CHAMBON and Vincent CHARVILLAT
IRIT (INP-ENSEEIH), Toulouse, France
Email: dmariyan@enseeiht.fr

Florent BRUNET
Ubleam, Toulouse, France
Email: florent.brunet@ubleam.com

Abstract—Circular markers are planar markers which offer great performances for detection and pose estimation. For an uncalibrated camera with an unknown focal length, at least the images of at least two coplanar circles are generally required to recover their poses. Unfortunately, detecting more than one ellipse in the image must be tricky and time-consuming, especially regarding concentric circles. On the other hand, when the camera is calibrated, one circle suffices but the solution is twofold and can hardly be disambiguated. Our contribution is to put beyond this limit by dealing with the uncalibrated case of a camera seeing one circle and discussing how to remove the ambiguity. We propose a new problem formulation that enables to show how to detect geometric configurations in which the ambiguity can be removed. Furthermore, we introduce the notion of default camera intrinsics and show, using intensive empirical works, the surprising observation that very approximate calibration can lead to accurate circle pose estimation.

I. INTRODUCTION

The problem of estimating the pose of a camera (or dually of a 3D object) from a set of 2D projections in a single view has been widely studied in the computer vision literature for a long time [11]. The “minimal” problem i.e., which requires the minimal amount of information necessary, is known as the perspective-3-point-problem (P3P) and consists in recovering the pose of a calibrated camera from three 3D-2D point correspondences. Many solutions are available for the general case when more information is available. When the environment in the scene can be controlled, artificial features with known positions are very often deployed in the scene. They are used in a wide range of applications, especially when a reliable reference is needed to, e.g., in cluttered or textureless environments. The most popular artificial features are probably coplanar features [4] whose layout in 2D space defines a so-called planar marker. The mapping between a planar marker and its image is a 2D projective transformation known as (world-to-image) homography and can be estimated from at least four world-to-image correspondences (the most simple planar marker is a square). Once the camera is calibrated, the decomposition of the homography matrix allows to recover the pose of the camera (or dually that of the plane).

Other well-known artificial markers that have been recently investigated again are those consisting of coplanar circles [1], [5], [8], [13]. The knowledge of at least two circle images (without any information on their parameters on the support

plane) allows to compute a world-to-image homography without ambiguity for all the spatial configurations of two circles except one [5].

Given a single circle image, it is well-known that a twofold solution exists for the normal to the support plane (and so for the pose) *only if* the camera is calibrated [13]. In this work, our contribution is to put beyond this limit by dealing with the case of an uncalibrated camera seeing one circle. Actually, our starting point came from the surprising observation, learned from empirical works, that very approximate calibration can lead to accurate circle pose estimation. Our idea is to use default intrinsics by designing a generic camera model delivering a default focal length based on off-line calibration of several smartphone cameras.

Our first contribution is to run extensive experiments that assess how the inaccuracy of the calibration impacts the quality of the pose estimation. We found out that exact calibration may not be required as small variations on the focal length does not affect the reprojection error of other reference coplanar points especially when the marker is far from the camera. Our second contribution is to provide a new geometrical framework to state the pose problem in which the issue of how to remove the twofold ambiguity can be thoroughly investigated.

We review the related works in section II. Then in section III, we remind the problem of recovering the pose from the projection of a circle before introducing the solution proposed in section IV. The idea is to introduce a new way of computing the vanishing line (dual to the plane normal) from one circle image. Thanks to it, as the general method leads to two possible solutions, we show how under some assumptions about the geometric configuration we can recover the correct one. Then, as we suppose that we work with uncalibrated images, we explain how we select parameter values to obtain what we called default camera intrinsic parameters. Finally in section V, we evaluate our method in the context of augmented reality.

II. RELATED WORK

A lot of existing works suggest to use a set of features encoded in a planar pattern to simplify the pose estimation. Fiala et al. introduced a fiducial system [4] and proposed the case of a special planar square marker. Recent efficient algorithms allow to detect ellipses precisely and, in consequence,

circles become features worth of interest. The four projective and affine parameters of the world-to-image homography (the remaining four define a similarity on the world plane [6, p42]) can be recovered by detecting the images of two special point of the support plane, known as circular points (e.g., see [6, p52-53]) which are common to all circles. Gurdjos et al. [5] relied on the notion of pencil of circle images to formulate the problem of detecting the images of the circular points as a problem of intersection of lines, obtained from the degenerate members of the pencil. Kim et al. [8] proposed algebraic and geometric solutions in the case of concentric circles. Calvet et al. [2] described a whole fiducial system using concentric circles which allows to accurately detect the position of the image of the circles common center under highly challenging conditions. In a same vein, Huang et al. [7] proposed to use the common self-polar triangle of concentric circles.

When using circular markers it is also possible to simplify the model of the camera to only depend on a sole focal length parameter. Chen et al. [3] autocalibrate the focal length using two or more coplanar circles. The problem to solve contains two parameters: one ellipse and the focal length. Then, two correspondences between a circle and an ellipse are necessary to estimate the focal length. Based on the same method, Bergamasco et al. [1] designed a marker composed of small circles spread on the edge of two or more concentric rings. The image of each circle is used with a vote system to estimate the focal length and the image of the external rings.

Two circles on a planar marker (except if one encloses the other) is the minimum to fully estimate the homography without any other assumptions. However in some applications e.g., dealing with concentric circles, detecting the images of two or more circles can be tricky. First because the lack of points induces an inaccurate estimation and, secondly because it is time consuming. When the camera has already been calibrated, it is possible to compute the homography from one circle image with two ambiguities. Pagani et al. [9] introduced a method quite similar to the solution proposed by Chen et al. [3], where the ambiguity is solved by minimizing a distance between the image of the marker rectified and the expected pattern on all possible poses.

III. POSE ESTIMATION FROM THE IMAGE OF ONE CIRCLE

We remind here some geometrical background on the problem of pose estimation from the image of a single circle. We consider an Euclidean projective camera, represented by a 3×4 -matrix $P \sim KR [1 | \mathbf{T}]$, where the rotation matrix $R \in SO(3)$ ¹ and the translation vector $\mathbf{T} \in \mathbb{R}^3$ describe the pose of the camera, i.e., respectively its orientation and position in the object 3D frame. The upper triangular order-3 matrix K is the calibration matrix as defined in [6, page 157].

Assume that \mathcal{P} is a plane with equation $Z = 0$ in the world frame. The pose of \mathcal{P} in the camera frame is given by the vector $[\mathbf{N} = \mathbf{r}_3, -d]^\top$, where \mathbf{r}_3 , the third column of R , defines the unit norm \mathbf{N} of \mathcal{P} , and d is the orthogonal

distance to \mathcal{P} . The restriction to \mathcal{P} of the projection mapping is an homography whose matrix writes $H \sim KR [e_1 | e_2 | \mathbf{T}]$, where e_1 and e_2 are the first two columns of l . In the projective plane, any conic can be represented in 2D homogeneous coordinates by a real symmetric order-3 matrix. Under perspective projection, any circle of \mathcal{P} , assuming its quasi-affine invariance [6, p515] i.e., that all its points lie in front of the camera, is mapped under the homography H to an ellipse by the projection equation $C = H^{-\top}QH^\top$, where $Q \in \text{Sym}_3^2$ is the circle matrix and $C \in \text{Sym}_3$ is the ellipse matrix.

For reasons that will become clearer later, we want to parameterize the homography H , *from only the knowledge of the circle image C and the vanishing line \mathbf{v}_∞ of \mathcal{P}* . Let $S_{\mathcal{P}}$ be a similarity on the world plane that puts the circle Q into a unit circle centered at the origin and $S_{\mathcal{I}}$ be a similarity on the image plane \mathcal{P} that puts C into a canonical diagonal form $C' = \text{diag}(C'_{11}, C'_{22}, C'_{33})$. Using an approach similar to [2] with the notation $[u, v, 1]^\top \sim C^{-1}\mathbf{v}_\infty$, it can be shown that, under the assumption of a camera with square pixels, we have $H \sim S_{\mathcal{I}}^{-1}MS_{\mathcal{P}}$ where

$$M = \begin{bmatrix} -1 & C'_{22}uv & -u \\ 0 & -C'_{11}u^2 + 1 & -v \\ -C'_{11}u & C'_{22}v & 1 \end{bmatrix} \begin{bmatrix} r & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & s \end{bmatrix}$$

with $r = \left(-\frac{C'_{22}}{C'_{11}}(C'_{11}u^2 + C'_{22}v^2 + C'_{33})\right)^{1/2}$ (1)

and $s = \left(-C'_{22}(1 - C'_{11}u^2)\right)^{1/2}$

Note that the matrices $S_{\mathcal{P}}$ and $S_{\mathcal{I}}$ can be completely determined by the circle image C and M , except for an unknown 2D rotation around the circle centre on \mathcal{P} . Recovering this rotation is not the goal of this paper. Some simple solution like placing a visible mark on the edge of the marker works generally well in many cases.

Our main task will be to recover the vanishing line \mathbf{v}_∞ of the plane, as explained in the sequel. Note that the vector $\mathbf{x}_c = [u, v, 1]^\top$ defined above is that of the image of the circle centre which is the pole of \mathbf{v}_∞ w.r.t. the dual ellipse of C .

IV. SUPPORT PLANE'S VANISHING LINE ESTIMATION

We warn the reader that parts written in italics in this section requires a proof that is not provided due to lack of space. However all proofs will appear in an extended paper version.

A. A twofold solution in the calibrated case

In the case of calibrated image, an equivalent problem of computing the pose of its support plane \mathcal{P} is that of recovering the vanishing line of \mathcal{P} . Let Q be the matrix of a circle on a plane \mathcal{P} , and $\psi = H^\top\omega H$ be that of the back-projection onto \mathcal{P} of the image of the absolute conic [6, p. 81], where $\omega = K^{-\top}K^{-1}$. *It is easy to show that ψ represents also a virtual³ circle (as does ω).*

²Sym₃ refers to the space of order-3 real symmetric matrices.

³Virtual conics have positive definite matrices, so, no real points on them.

¹SO(3) refers to the 3D rotation group.

Let $\{\alpha_i\}_{i=1..3}$ denotes the set of generalized eigenvalues of the matrix-pair (Q, ψ) , i.e., the three roots of the characteristic equation $\det(Q - \alpha\psi) = 0$. The set of matrices $\{Q - \alpha\psi\}_{\alpha \in \mathbb{R} \cup \{\infty\}}$ defines a conic pencil [5] which includes three degenerate conics with matrices $D_i = Q - \alpha_i\psi$. These rank-2 matrices represent line-pairs and have the form $D_i = \mathbf{l}_a^i (\mathbf{l}_b^i)^\top + \mathbf{l}_b^i (\mathbf{l}_a^i)^\top$, where \mathbf{l}_a^i and \mathbf{l}_b^i are vectors of these lines. Such line-pair matrix D_i can be easily decomposed and vectors of its lines recovered albeit it is impossible to distinguish \mathbf{l}_a^i from \mathbf{l}_b^i . *It can be shown that the projective signatures⁴ of the three degenerate members always are (1, 1), (2, 0) and (0, 2).* Assume, without loss of generality, that the degenerate conic D_2 is the one with signature (1, 1). A **first key result** is that D_2 is a pair of two distinct real lines, one of which being the line at infinity $\mathbf{l}_\infty = [0, 0, 1]^\top$; the other one being denoted by \mathbf{l}_o . The other two degenerate conics D_1 and D_3 —with signatures (2, 0) and (0, 2)— are pairs of two conjugate complex lines. Consequently, the three (so-called) base points \mathbf{x}_i , where lines in a pair meet, are real. Moreover, their vectors are the generalized eigenvectors of (Q, ψ) and satisfy $D_i \mathbf{x}_i = 0$.

Similarly, in the image plane, if C denotes the image of the circle Q , the set of matrices $\{C - \beta\omega\}_{\beta \in \mathbb{R} \cup \{\infty\}}$ defines also a conic pencil whose members are the images of the pencil $\{Q - \alpha\psi\}_{\alpha \in \mathbb{R} \cup \{\infty\}}$. Hence, the line-pair in $\{C - \beta\omega\}$ that includes the image of \mathbf{l}_∞ i.e., the vanishing line \mathbf{v}_∞ , can always be identified since it is the only degenerate member with signature (1, 1). Nevertheless, at this step, it is impossible to distinguish \mathbf{v}_∞ from the other line \mathbf{v}_o , image of \mathbf{l}_o .

Assume that all matrices Q , C , ψ and ω are normalized to have a unit determinant. It is known that, in this case, parameters in pencils satisfy $\alpha = \beta$, so, the generalized eigenvalues of the matrix-pair (Q, ψ) are *exactly* the same as those of (C, ω) . *It can be shown that these eigenvalues can always be sorted such that $\lambda_1 \geq \lambda_2 \geq \lambda_3$, where $D_2 = C - \lambda_2\omega$ is the (sole) degenerate conic with signature (1, 1).* Remind that D_2 is the conic which contains \mathbf{v}_∞ plus \mathbf{v}_o , which are two *a priori* indistinguishable lines denoted by $\mathbf{v}_{1,2}$. Because the matrix D_2 is real, symmetric, rank-2 and order-3, *its generalized eigen-decomposition using the base point vectors $\mathbf{x}_1, \mathbf{x}_3 \in \mathbb{R}^3$ writes as following:*

$$D_2 = \begin{bmatrix} \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} & \frac{\mathbf{x}_3}{\|\mathbf{x}_3\|} \end{bmatrix} \begin{bmatrix} \lambda_1 - \lambda_2 & 0 \\ 0 & \lambda_3 - \lambda_2 \end{bmatrix} \begin{bmatrix} \frac{\mathbf{x}_1^\top}{\|\mathbf{x}_1\|} \\ \frac{\mathbf{x}_3^\top}{\|\mathbf{x}_3\|} \end{bmatrix} \quad (2)$$

from which *it can be shown that*

$$\mathbf{v}_{1,2} = \sqrt{\lambda_1 - \lambda_2} \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} \pm \sqrt{\lambda_2 - \lambda_3} \frac{\mathbf{x}_3}{\|\mathbf{x}_3\|} \quad (3)$$

The two solutions to the normal to \mathcal{P} are given by $\mathbf{N}_{1,2} = \mathbf{K}^\top \mathbf{v}_{1,2}$ in the camera frame, and (3) explains the known doublefold ambiguity in the plane pose [3].

⁴The signature of a conic is $\sigma(C) = (\max(p, n), \min(p, n))$, where p and n count the positive and negative eigenvalues of its (real) matrix C . It is left unchanged by projective transformations.

B. About removing the twofold ambiguity

We have seen that there are two solutions for the vanishing line (or the plane normal in the calibrated case) which are in general not distinguishable. In this section, we discuss whether known configurations allows the ambiguity to be removed. We extend the new theoretical framework proposed in §IV-A that involves the point \mathbf{q} (on the support plane \mathcal{P}) where the optical axis cuts \mathcal{P} plus the line \mathcal{L} obtained by intersecting \mathcal{P} and the principal plane⁵ of the camera (\mathcal{L} is orthogonal to the orthogonal projection of the optical axis onto \mathcal{P}). Now, let \mathcal{L}' denote the line parallel to \mathcal{L} through the circle centre. Within this geometrical framework, we can claim, for instance, that *a sufficient condition for the ambiguity to be solved is given by the two following conditions:*

- (i) \mathbf{q} and the orthogonal projection on \mathcal{P} of the camera centre lie on the same side of \mathcal{L}' ;
- (ii) the point, intersection of the orthogonal projection on \mathcal{P} of the optical axis and \mathcal{L}' , lies outside the circle centered at \mathbf{q} with same radius as Q .

Figure 1 illustrates this important result. We are convinced that future investigations using this framework can help to reveal more configurations in which the ambiguity can be removed. We are now giving more geometrical insights indicating how to determine such configurations, via three propositions. The first is the **second key result** which is the building brick of our approach:

Proposition 1 (second key result) *The line \mathbf{l}_o in D_2 separates the two base points \mathbf{x}_1 and \mathbf{x}_3 . Hence, denoting by $\bar{\mathbf{x}}$ the normalized vector $\bar{\mathbf{x}} = \mathbf{x}/x_3$, the following inequalities hold $(\mathbf{l}_\infty^\top \bar{\mathbf{x}}_1)(\mathbf{l}_\infty^\top \bar{\mathbf{x}}_3) > 0$ and $(\mathbf{l}_o^\top \bar{\mathbf{x}}_1)(\mathbf{l}_o^\top \bar{\mathbf{x}}_3) < 0$.*

These two inequalities hold under any affine transformation but not under a general projective transformation.

How the conditions in proposition 1 can be helpful in removing the plane pose ambiguity? Can we state a corollary saying that, in the image plane, under some known geometric configuration, we know which the line \mathbf{v}_o in $C - \lambda_2\omega$, image of \mathbf{l}_o , *always* separates points \mathbf{z}_1 and \mathbf{z}_3 , images of base points \mathbf{x}_1 and \mathbf{x}_3 , while the other does not? That is, if we *a priori* know $\text{sign}(\mathbf{v}_o^\top \bar{\mathbf{z}}_1)(\mathbf{v}_o^\top \bar{\mathbf{z}}_3)$ can we guarantee that $(\mathbf{v}_o^\top \bar{\mathbf{z}}_1)(\mathbf{v}_o^\top \bar{\mathbf{z}}_3) = -(\mathbf{v}_\infty^\top \bar{\mathbf{z}}_1)(\mathbf{v}_\infty^\top \bar{\mathbf{z}}_3)$? If yes, since the vectors of these base points are the generalized eigenvectors of (C, ω) associated to parameters λ_j , $j \in \{1, 3\}$ and can be straightforwardly computed, we could remove the ambiguity by choosing as vanishing line \mathbf{v}_∞ the “correct” line in $C - \lambda_2\omega$. We claim the following proposition for this corollary to hold, whose (omitted) proof directly follows from the properties of quasi-affineness w.r.t. the base points [6].

Proposition 2 *When \mathbf{x}_1 and \mathbf{x}_3 lie either both in front or both behind the camera i.e., on the same half-plane bounded by \mathcal{L} , we have $(\mathbf{v}_o^\top \bar{\mathbf{z}}_1)(\mathbf{v}_o^\top \bar{\mathbf{z}}_3) < 0$ and $(\mathbf{v}_\infty^\top \bar{\mathbf{z}}_1)(\mathbf{v}_\infty^\top \bar{\mathbf{z}}_3) > 0$. Otherwise $(\mathbf{v}_o^\top \bar{\mathbf{z}}_1)(\mathbf{v}_o^\top \bar{\mathbf{z}}_3) > 0$ and $(\mathbf{v}_\infty^\top \bar{\mathbf{z}}_1)(\mathbf{v}_\infty^\top \bar{\mathbf{z}}_3) < 0$.*

⁵The 3D plane through the camera centre and parallel to the image plane.

Now let us investigate a formal condition saying when \mathbf{x}_1 and \mathbf{x}_3 lie on the same half-plane bounded by \mathcal{L} . Consider an Euclidean representation of the projective world in which the origin is the point \mathbf{q} at which the optical axis cuts the plane \mathcal{P} . Let the X -axis be parallel to the line \mathcal{L} and the Y -axis is the orthogonal projection of the optical axis onto \mathcal{P} . Consequently, the Z -axis is directed by the normal to \mathcal{P} , as shown in figure 1. Let $\mathbf{C} = [0, -\cos\theta, \sin\theta]^\top$, $\theta \in [0, \frac{\pi}{2}[$, be the 3D cartesian coordinates of the camera centre, where $\pi - \theta$ is the angle between the Y -axis and the optical axis in the YZ -plane (note that we choose the scale such that the camera centre is at distance 1 from the origin). Therefore the direction of the optical axis is given by $-\mathbf{C}$.

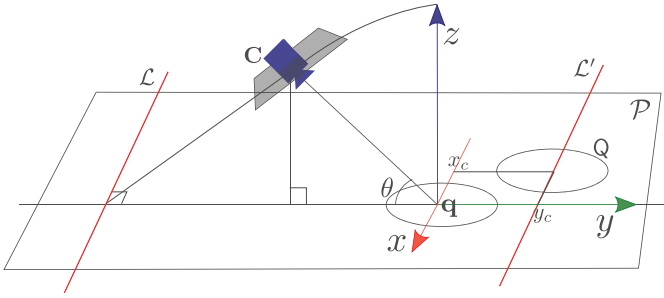


Fig. 1: Proposed parametrization for detecting the ambiguity.

In the 2D representation of the projective plane \mathcal{P} (i.e., of the XZ -plane), let the circle have centre (x_c, y_c) and radius R . Let $\mathbf{d} = [0, 1, \cos\theta]^\top$ is the vector of line \mathcal{L} . It can be shown, using a symbolic software like MAPLE⁶, that:

Proposition 3 Base points \mathbf{x}_1 and \mathbf{x}_3 lie, in the world plane, on the same side of \mathcal{L} if and only if

$$\cos\theta(y_c^2 - R^2)(y_c + \cos\theta) + y_c \cos\theta(1 + x_c^2) + x_c^2 + y_c^2 \leq 0 \quad (4)$$

Since $\cos\theta > 0$, if $y_c > 0$ and $y_c^2 - R^2 > 0$ then \mathbf{x}_1 and \mathbf{x}_3 lie on opposite sides of \mathcal{L} . The former inequality says that \mathbf{q} must lie on the same side of \mathcal{L}' , the line parallel to \mathcal{L} through the circle centre, as the orthogonal projection of the camera centre onto \mathcal{P} . The latter inequality says the point $(0, y_c)$ must lie outside the circle centered at $\mathbf{q}(0, 0)$ with same radius R as Q . As we are in the “otherwise” part of proposition 2, the vanishing line is given by the line that does not separate the image of the base points. Since $(0, y_c)$ represents the intersection of the orthogonal projection on \mathcal{P} of the optical axis and \mathcal{L}' , this is the result announced at the beginning of this section.

C. Defining default intrinsics for the camera

In the previous sections we have seen that, providing that the camera intrinsics are known, there is a twofold solution for the vanishing line. Recovering accurate intrinsics of a camera requires generally a calibration procedure. In many applications, the model of the camera can be simplified to reduce the number of parameters. A very common model is

that of a camera with square pixels and principal point at the centre of the image plane. Consequently, the focal length is the sole unknown, e.g., for self-calibration purposes [10]. The focal length value is sometimes available through EXIF data, stored in digital images or video files, through camera hardware on top level API (Android, iOS) or through data provided by manufacturer on websites. Focal length, denoted f , in pixels (what we need) can be obtained from this data if we find the field of view in angle or the focal length equivalent in 35mm. However the focal length is very often given in millimetre without the sensor size required to obtain the focal length in pixels.

We consider here the case where it is impossible to calibrate the camera by none of the methods mentioned above. So how to do? We propose to design a generic camera model delivering default intrinsics (i.e., focal length) and based on off-line calibration of several smartphone cameras. If a camera can generally take any focal length value, the optics and the sensor of smartphones are constrained by the device size and the desired field of view. Why doing that? We found out that surprisingly enough, that it is not necessary to have very accurate intrinsics to estimate the vanishing line given the image of a single circle. In fact, as shown in the experimental section V, this estimation is very robust to intrinsics fluctuation.

After calibrating a dozen of camera devices and obtaining data from manufacturers of twenty more smartphones, we estimate a gaussian model of the focal length equivalent in 35mm, as shown in figure 2. In our case we obtained

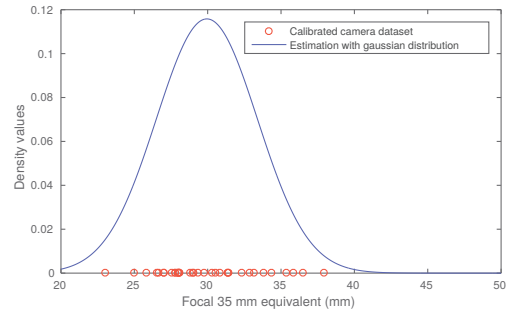


Fig. 2: Focal calibration of different camera parameters

experimentally an average focal length of $f_{35} = 29.95\text{mm}$ with a variance of $\sigma_{f_{35}}^2 = 11.86$. More precisely, we estimate a gaussian function (in blue) based of the focal values collected or estimated (in red) from different smartphone device brands.

V. EXPERIMENTAL RESULTS

A. Test Description

The goal of the test presented in this section is to evaluate the proposed method to estimate the pose of a camera. We performed those tests on synthetic and real images in the conditions illustrated in figure 3. In order to limit the poses used in experiments, we made some hypotheses. First, we suppose that the camera focus on the centre of the marker,

⁶<https://fr.maplesoft.com/>

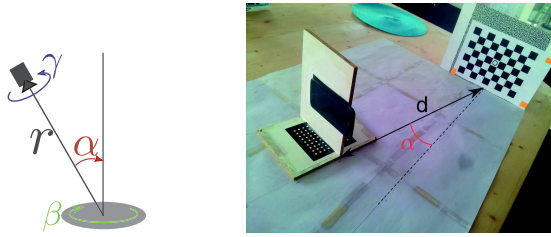


Fig. 3: Setup: reference chessboard and pose annotation

i.e. the principal axis of the camera passes through the centre of the marker, see figure 3. Then, the angle γ has been set to zero. In deed, we can simulate any angle by rotating the image using the assumption that the principal point is centred on the image and that the principal axis is orthogonal to the image plane. Finally, the angle β has been fixed to zero as estimating the 2D rotation around the plane normal is out of the scope of this article. The remaining variables whose variations are studied in our test are the angle α and the distance r .

We know that introducing generic camera parameter, as proposed in section IV-C, should have a negative impact on the accuracy of the pose estimation. Consequently, one of the objectives of this experiment is to evaluate the sensitivity of the proposed method to inaccurate camera focal parameter. The observation of the distribution of focal length of various smartphone camera, see figure 2, reveals that all 35mm focal equivalent are included in $[-30\%, +30\%]$ of the average value. So, five different values that span this range are used in the experiment: $\{0.7, 0.85, 1.0, 1.15, 1.3\}$. In order to generate synthetic images, we have simulated a synthetic camera of focal $\alpha_x = 1280$ and resolution of 1280×720 pixels. To obtain real images, we have used the camera of a smartphone which have been calibrated with openCV library⁷. In both cases, we suppose that ellipses have been firstly detected in the images, i.e. contour points are first detected and then ellipses are estimated [12]. We try to evaluate the impact of errors of this estimation to the quality of the results. In consequences, in our synthetic tests, we have also simulated noises on the detections of the ellipses, i.e. errors on the pixels that belong to the ellipse. More precisely, edge points of the ellipse have been translated with a zero mean gaussian variance of $\sigma_x = 1.0$.

Finally, we evaluate the quality of the results obtained by using three different measurements relative to the pose and the reprojection accuracy:

- a) Error on the normal of the plane relative to the camera;
- b) Error on the position of the marker;
- c) Error of reprojection of 3D points close to the marker.

Each curve illustrates the results obtained by applying a modifier on focal length used for pose estimation. The resulting errors are displayed as function of the distance r in the interval $[15 \times D, 50 \times D]$ where D is the diameter of the marker. This interval is related to the distances used for being able to detect and to recognize a marker for an augmented reality

⁷<https://opencv.org/>

application, i.e. the distance where the marker occupies, at least 80 pixels. We also show results for three different angle values, $\alpha \in \{15, 30, 45\}$, displayed in three sub-figures.

B. Analysis of the results

Results on synthetic images are presented in figure 4. In 4a, we show the error on the estimation of the orientation for the pose. We can notice that as the distance of the marker to the camera increases, the error on pose orientation also increases. This relation is even more remarkable when the angle is the lowest between the marker plane and the camera, i.e. the graph on the left. In 4b, we can see that in the calibrated case the accuracy in position stays low and does not depend on the distance to the camera and the angles between the marker plane and the camera. In the uncalibrated cases, as expected the detection of the ellipses becomes less accurate when the distance increases and, consequently, the quality of the estimation of the marker position is also affected. In fact, the error in position increases linearly when the distance increases. This observation is quite intuitive. In deed observing a marker with a zoom or taking its image closer leads to very similar shape of the marker. The error on the reprojection of 3D points, presented in 4c, illustrates that, with a focal length well estimated, the higher the distance, the higher the errors. Whereas, when the focal length is not well estimated, the higher the distance, the lower the error and, more important, this error is quite near the error obtained when the focal length is correctly estimated. It means that using generic parameter is not affecting the quality of the reprojection in a context where the marker is far from the camera.

The figure 5 allows us to present similar conclusions on real images. The 3D point reprojection error is presented. The error in calibrated case slightly increases with the distance as observed in figure 4a. When the marker is close to the camera, the error of reprojection when the camera is not correctly calibrated is high but it drastically decreases when the distance to the camera increases, and, finally, this error is of the same order as that obtained with the calibrated case. This observation is not really a surprise as the projection of a distant object loses its perspective with distance. Again, this result illustrates the interest of using generic camera parameter in augmented reality.

VI. CONCLUSION

In this paper, we introduced a method to estimate the pose of a camera from the image of a circular marker in a calibrated case. If, in general case, two solutions are found, some assumptions on geometric configuration can help to distinguish the correct pose. Moreover, we demonstrated the interest of using default camera parameters, in the context of augmented reality. In particular, the results presented showed that, in a case of a distant marker, the 3D reprojection errors is low enough. Future work would be to use more information in the marker environment to increase the stability of the detection of the marker and the pose estimation and to allow decoding from longer distance.

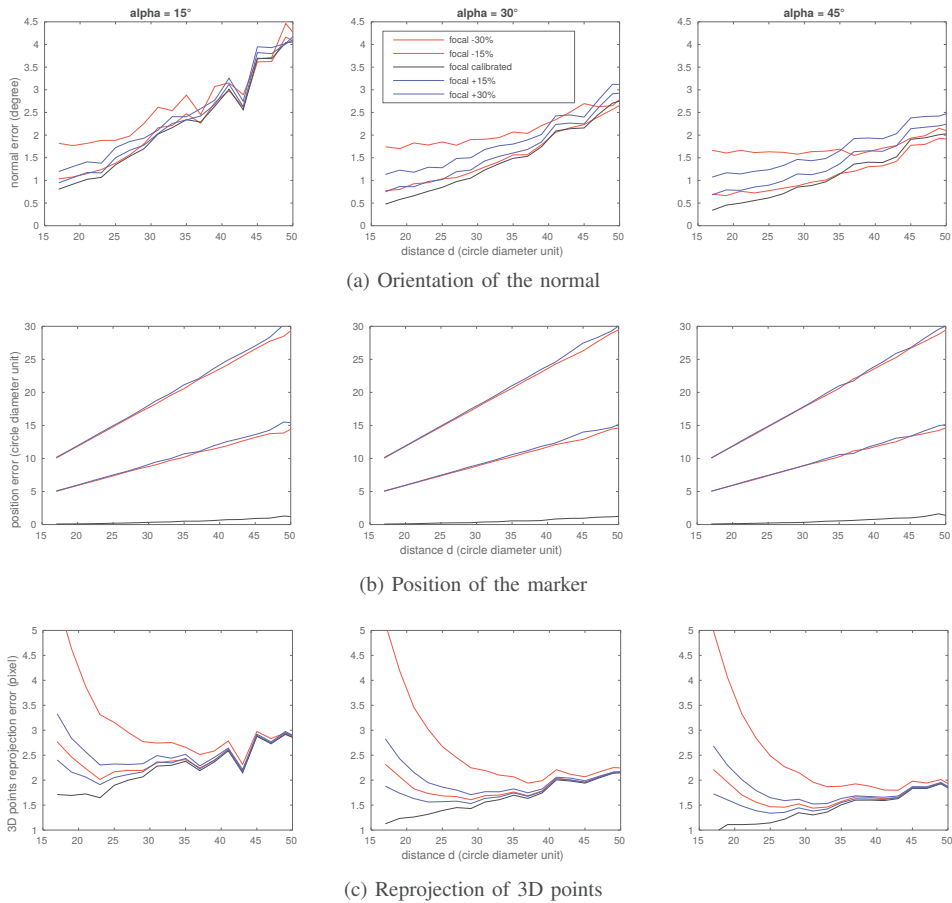


Fig. 4: Error with synthetic images.

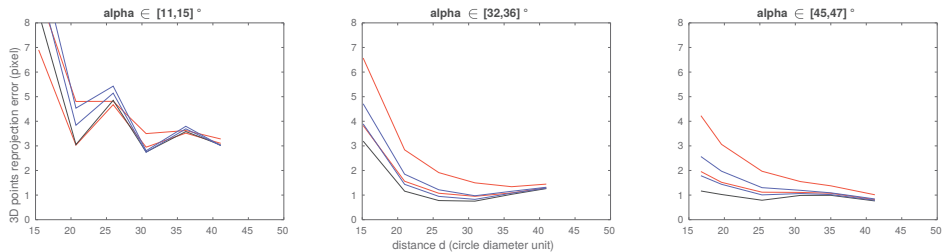


Fig. 5: Error with ellipses detected on real image: 3D points reprojection.

REFERENCES

- [1] F. Bergamasco, A. Albarelli, E. Rodolà, and A. Torsello. RUNE-Tag: A high accuracy fiducial marker with strong occlusion resilience. *CVPR*, 2011.
- [2] L. Calvet, P. Gurdjos, C. Griwodz, and S. Gasparini. Detection and Accurate Localization of Circular Fiducials under Highly Challenging Conditions. *CVPR*, 2016.
- [3] Q. Chen, H. Wu, and T. Wada. Camera calibration with two arbitrary coplanar circles. *ECCV*, 2004.
- [4] M. Fiala. Designing highly reliable fiducial markers. *PAMI*, 2010.
- [5] P. Gurdjos, P. Sturm, and Y. Wu. Euclidean structure from $N \geq 2$ parallel circles: Theory and algorithms. *ECCV*, 2006.
- [6] R. Hartley and A. Zisserman. *Multiple View Geometry*. 2003.
- [7] H. Huang, H. Zhang, and Y. M. Cheung. The common self-polar triangle of concentric circles and its application to camera calibration. *CVPR*, 2015.
- [8] J. S. Kim, P. Gurdjos, and I. S. Kweon. Geometric and algebraic constraints of projected concentric circles and their applications to camera calibration. *PAMI*, 2005.
- [9] A. Pagani, J. Koehle, and D. Stricker. Circular Markers for camera pose estimation. *Image Analysis for Multimedia Interactive Services*, 2011.
- [10] M. Pollefeys, R. Koch, and L. Van Gool. Self-Calibration and Metric Reconstruction in Spite of Varying and Unknown Internal Camera Parameters. *IJCV*, 1999.
- [11] R. Szeliski. *Computer Vision: Algorithms and Applications*. 2010.
- [12] Z. L. Szpak, W. Chojnacki, and A. van den Hengel. Guaranteed Ellipse Fitting with a Confidence Region and an Uncertainty Measure for Centre, Axes, and Orientation. *JMIV*, 2015.
- [13] H. Wu, Q. Chen, and T. Wada. Conic-based algorithm for visual line estimation from one image. *Automatic Face and Gesture Recognition*, 2004.