International Journal of Computer Vision
https://doi.org/10.1007/s11263-018-1139-6

An Exact Robust Method to Localize a Known Sphere by Means of One Image

Rudi Penne[1] · **Bart Ribbens**[2] · **Pedro Roios**[2]

Received: 22 November 2017 / Accepted: 27 November 2018

Abstract
In this article we provide a very robust algorithm to compute the position of the center of a sphere with known radius from one image by a calibrated camera. To our knowledge it is the first time that an exact sphere localization formula is published that only uses the (pixel) area and the ellipse center of the sphere image. Other authors either derived an approximation formula or followed the less robust and more time consuming procedure of fitting an ellipse through the detected edge pixels. Our method is analytic and deterministic, making use of the unique positive real tool of a cubic equation. We observe that the proposed area method is significantly more accurate and precise than an ellipse fitting method. Furthermore, we investigate in what conditions for sphere images the proposed exact method is preferable to the robust approximation method. These observations are validated by virtual, synthetic and real experiments.

Keywords Extrinsic calibration · Robust localization · Projective geometry

1 Introduction

A spherical object is easy to recognize in an image and its silhouette is visible independent from any orientation the image has been taken. This has motivated various authors to use images of one or more spheres in camera calibration, both intrinsic and extrinsic (Teramoto and Xu 2002; Lu and Payandeh 2010; Zhang et al. 2007; Penna 1991; Agrawal and Davis 2003; Daucher et al. 1994; Sun et al. 2015; Penne et al. 2015; Guan et al. 2015). In Sun et al. (2015) it is claimed that their sphere image based calibration improves the 2% accuracy tollerance of the standard method of Zhang (2000). Often, the accurate localization of the centers of the viewed spheres is a crucial step in these calibration procedures. For instance, in Sun et al. (2015) the position of the sphere is recovered after the initial computation of the linear intrinsic calibration parameters in order to remove radial distortion by an iterative procedure. Another illustration is given by Guan et al. (2015), where the localization of the sphere centers is used as a first step to establish the extrinsic calibration of a network of calibrated cameras by means of the orthogonal procrustes analysis of Arun et al. (1987). See Lu and Payandeh (2010) for another algorithm to compute the extrinsic camera calibration from sphere locations.

In this article we present a new accurate method for localizing the center of a sphere with known radius relative to the camera reference system by means of one calibrated image of this sphere. Consequently, this center can be considered as a robust 3D feature point that has been determined avoiding image corner detection, stereovision or range sensors, that is useful in issues such as ego-motion or relative camera pose determination. Furthermore, the position of the spheres might be interesting in itself in applications such as sound source localization (cf. `LMS Soundbrush`), or for mobile robots collecting golf balls.

Commonly, image based methods for sphere localization or other sphere analysis purposes rely on the segmentation of the elliptic boundary of the sphere image and need a well fitted ellipse for these edge points. Examples of ellipse fitting can be found in Xie and Ji (2002) and Yin and Chen (1994), but state of the art algorithms are essentially based on Fitzgibbon et al. (1996). For these methods an edge detector such as given by Canny (1986) is required in a preprocessing phase, implying that the resulting quality of the fitted ellipse is sen-

Communicated by Daniel Cremers.

✉ Rudi Penne
rudi.penne@uantwerpen.be

[1] Department of Mathematics, Faculty of Applied Engineering, University of Antwerp, 2020 Antwerp, Belgium

[2] Faculty of Applied Engineering, University of Antwerp, 2020 Antwerp, Belgium

Published online: 05 December 2018

sitive to image segmentation errors. A meritorious attempt to avoid a treshold for contour point extraction is given by Ouellet and Hébert (2008). Our method does not require any ellipse fitting of the sphere image edge as it only uses the ellipse center (center of gravity of the image blob) and its area (pixel area of this blob), which are more robust image measurements and avoid strict edge detection. Such a robust area method for sphere localization has already been done by Guan et al. (2015). But the main contribution of our paper is to provide an exact formula for the sphere center, while Guan et al. (2015) approximates its position (assuming a sphere radius very small relative to the depth of its position). To our knowledge, it is the first time that an analytic formula for a sphere center has been derived from the area and the center of one image (provided intrinsic camera calibration and a known sphere radius).

A second addition to the presentation in Guan et al. (2015) is an alternative way to validate the robustness of the area method for sphere localization. In Guan et al. (2015) the authors compare their approximating area formula to the method of Agrawal and Davis (2003) that is in essence also designed to intrinsically calibrate the camera, while we chose to go in competition with a method that directly locates the sphere center for a calibrated camera by means of the ellipse fitted to the sphere image. This direct ellipse fitting method locates a sphere more accurately than the reference method of Agrawal and Davis (2003) that is used by Guan et al. (2015), but it is still outperformed by a robust area method anyway. A second difference in approach is that Guan et al. (2015) evaluates its method by means of an experiment where the ground truth is given by manual measurements, while we provide a more reliable ground truth by mounting the camera on a translating robot arm. In addition, we use realistic image render software to obtain synthetic digital sphere images such that the performance can be evaluated under different controlled circumstances with the availability of an exact ground truth.

The paper is composed as follows. In Sect. 2 we list the geometric properties of the perspective projection of sphere that will be needed in the sequel. In particular, we highlight an important formula relating the elliptic parameters of a ball image to the intrinsics of a pinhole camera (Eq. 1 in Proposition 2). This section contains no new material, referring to Daucher et al. (1994), Penne et al. (2015) and Beardsley et al. (1992).

In Sect. 3 we derive an analytic formula (Eq. 6) for the center of the sphere in camera coordinates, using the knowlegde of the ball radius A, the center of the elliptic sphere image M and the (half) major axis a of this ellipse, relying on Proposition 1 of Sect. 2. Any method that segments the sphere image and succeeds in determining the center M and the major size a of the ellipse, can benefit from this formula. Equation (6) will be our reference for computing the sphere position by ellipse fitting and will be compared with the robust methods. Although the computations for deriving this formula are straightforward, we did not find any reference for it.

Section 4 presents the main contribution of this paper. Here we explain how the size a of the major axis of the elliptic image of a sphere can be computed from the ellipse area A, ellipse center M and the intrinsic parameters of the camera. This is established due to Eq. (1) and by solving a cubic equation in b^2, with b the (half) minor ellipse axis (Eq. 9). In combination with Eq. (6), the result of this section yields an exact analytic procedure to localize a sphere with known radius. We were pleased to have the opportunity to use old school algebra, such as the Cardano–Tartaglia formula and the rule of Descartes.

In Sect. 5 we present Junzhi's approximation formula of Guan et al. (2015), that also localizes the sphere center by means of the robust measurements M and A only.
Finally, in Sects. 6 and 7 we validate the robustness of both area methods (the proposed exact method and Junzhi's method) and compare the performance with sphere localization relying on ellipse fitting. In order to have a reliable ground truth and to be able to investigate the sensitivity of the methods regarding calibration errors and radial distortion, we chose to generate synthetic data by software that renders realistic digital images. We can conclude that the robust area methods outperform the ellipse fitting method in most cases, both in accuracy and in precision. For balls that are small relative to the position depth the approximation method of Junzhi performs equally well as our exact method, because in this case the approximation error has the same order of magnitude as the image noise. A difference in quality between both area methods (in favor of the proposed exact method) only becomes apparent for larger spheres, as we show by means of a geometric simulation. The observations of Sect. 6 are confirmed by real experiments presented in Sect. 7. In these experiments we acquired images by means of an RGB camera rigidly attached to a robot manipulator, controlled by pure translations of the toolcenter. Consequently, we have access to the translational distance of the attached camera, without any need for computing the hand-eye calibration. The considered sphere localization methods can be evaluated by computing the relative sphere distance for image pairs and by comparing it to the ground truth provided as the distance between the two corresponding toolcenter positions.

2 Some Fundamental Facts About the Perspective Projection of a Sphere

Let $\mathcal{B}$ be a sphere centered at P with radius R. A pinhole camera acquires the image of a sphere by a perspective projection from the camera center O upon the sensor plane γ. Alternatively, if $\mathcal{K}$ is the cone with top O tangent to the given

sphere, this perspective image $\mathcal{E}$ is obtained by intersecting $\mathcal{K}$ by the sensor plane γ. We also assume that we have a complete image of the sphere, implying that the sensor plane γ separates the camera O from the viewed ball. It was already observed by Apollonius in his *Konika* (225 A.D) that the resulting perspective image (conic section) is an ellipse.

By symmetry it is easily seen that the points of tangency where the tangents from O meet the sphere $\mathcal{B}$, make a circle on this sphere in a plane perpendicular to the cone axis OP. This implies that $\mathcal{K}$ is in fact a *circular* cone. Hence, the *Theorem of Dandelin–Quételet* (Dandelin 1822) applies. This theorem forms the basis of the geometric properties of the elliptic image of a sphere.

Theorem 1 (Dandelin–Quételet) *Let γ be a plane that intersects a circular cone $\mathcal{K}$ in a closed curve $\mathcal{E}$, and B_1 and B_2 be the two spheres* (Dandelin balls) *that are simulataneously tangent to the cone $\mathcal{K}$ and the plane $\mathcal{E}$ (see Fig. 1). Then $\mathcal{E}$ is an ellipse and the Dandelin balls touch the plane γ in the two foci of $\mathcal{E}$.*

Proof We refer to Fig. 1. □

In Penne et al. (2015) the authors demonstrate the crucial role of Theorem 1 in deriving the following property of sphere images (under perspective projection), that will be useful for our purposes. The observation also appeared in Daucher et al. (1994) without a proof.

Proposition 1 *Let γ be a plane that intersects the circular cone $\mathcal{K}$ in an elliptic section $\mathcal{E}$. If α is the plane through the cone axis s of $\mathcal{K}$, and perpendicular to γ, then α contains the major axis of the ellipse $\mathcal{E}$. Consequently, if C is the orthogonal projection of O on γ, and if Q is the perspective projection of the sphere center in γ, then both C and Q are collinear with this major axis.*

Notice that this is a Euclidean property. The concept of 'major axis of an ellipse' is even not an affine invariant. Therefore, in order to apply Proposition 1 for a pinhole camera, the elliptic projection of the sphere must be described in calibrated image coordinates (world units), or with respect to (undistorted) square pixels.

Finally we arrive at an important formula that relates the elliptic image of a sphere to the intrinsics of a pinhole camera. To our knowledge, this formula first appeared in Beardsley et al. (1992). Formula 1 has also been proven in full detail in Penne et al. (2015), as a consequence of Theorem 1 and using the concept of reciprocal ellipse-hyperbola pairs.

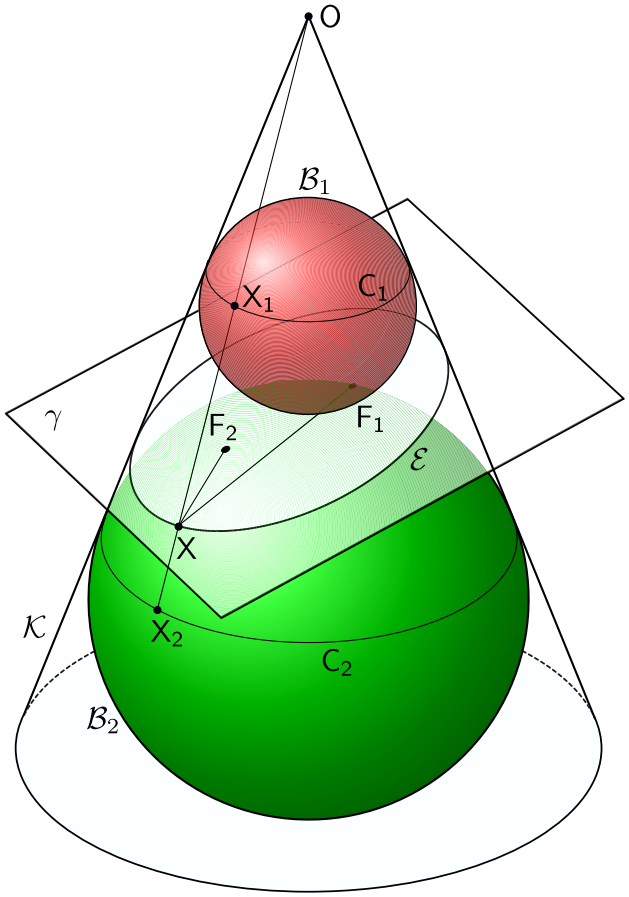

Fig. 1 The proof of Theorem 1. In this diagram one can see that $d(X, F_1) = d(X, X_1)$ as two "tangency distances" for a sphere. For the same reason, $d(X, F_2) = d(X, X_2)$, whence $d(X, F_1) + d(X, F_2) = d(X_1, X_2) =$ constant. This implies that the points $X \in \gamma \cap \mathcal{K}$ lie on the ellipse with foci F_1 and F_2

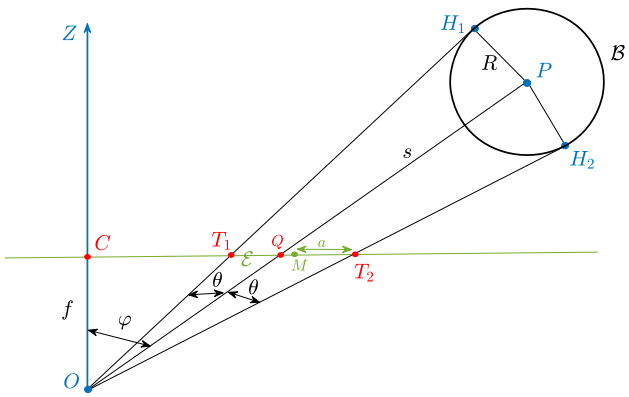

Fig. 2 Due to Proposition 1 we can represent the cone axis $s = OP$, the major ellipse axis T_1T_2 and the Z-axis in one plane

Proposition 2 *Assume an elliptic image $\mathcal{E}$ of a ball by a perfect pinhole camera with square pixels. Let δ_M denote the distance between the principal point C and the center M of $\mathcal{E}$, let a denote the half major axis and b denote the half minor axis of $\mathcal{E}$ (all in pixel units). Then we can relate the focal length f of the camera to a, b and δ_M by means of the following formula:*

$$f^2(a^2 - b^2) = b^2(\delta_M^2 - a^2 + b^2). \tag{1}$$

3 The Coordinates of the Sphere Center Relative to the Pinhole Camera Reference System

Suppose we take an image of a sphere $\mathcal{B}$ with known radius R. If we neglect or remove nonlinear lens distortions, this ball image is obtained by a perspective projection from the pinhole center O upon the sensor plane γ. So we may assume that the image of $\mathcal{B}$ is an ellipse $\mathcal{E}$ in the image plane γ (see Sect. 2). Furthermore, we assume that we know the parameters of the elliptic image $\mathcal{E}$, e.g. by some ellipse fitting algorithm in combination with an edge detector. As a matter of fact, we only need the value $2a$ of the major axis and the position of the center M of $\mathcal{E}$. In our experiments, we used the ellipse fitting algorithm of Fitzgibbon et al. (1996), but we also want to mention (Ho and Chen 1995; Ballard 1981; Xie and Ji 2002; Yin and Chen 1994).

Under these conditions highschool maths suffices to locate the observed sphere in the standard camera reference frame. The origin of this frame is given by the pinhole center O, axis Z coincides with the focal axis (pointing towards the scene), axes X and Y are parallel to the pixel axes, and the unit equals the world unit (chosen to measure the ball radius). A possible solution is given below. For the convenience of the reader we first present a list of the used notations (Fig. 2).

Notations

- $\mathcal{K}$: the cone with top O and tangent to the ball $\mathcal{B}$. Notice that the ellipse $\mathcal{E}$ is the intersection of $\mathcal{K}$ with the image plane γ.
- s: the central axis (of symmetry) of the cone $\mathcal{K}$.
- Q: the intersection of s with the image plane, so the virtual image of the ball center P.
- T_1 and T_2: left and right end points of the major axis of the ellipse $\mathcal{E}$.
- $2a$: the length of the major axis of $\mathcal{E}$ (px).
- $2b$: the length of the minor axis of $\mathcal{E}$ (px).
- $M = (u_M, v_M)$: center of ellipse $\mathcal{E}$ (px).
- $C = (u_0, v_0)$: principal point of the camera image (px).
- δ_M: distance between C and M (px).
- f: focal length (px).

Notice that we can restrict our arguments and computations to the plane α through the camera center O and the line CM (in the image plane). The plane α contains both the focal axis Z and the cone axis s (as a consequence of Proposition 1). Figure 2 shows the situation as restricted to plane α.

Important angles (defined in the plane α, see Fig. 2)

- φ: the angle between the focal axis Z and the cone axis s
- θ: the (half) cone aperture: the angle between the cone axis s and OT_1 or OT_2 (s bisects $T_1\hat{O}T_2$)

Strategy for localizing sphere center P

1. Considering the right-angled triangles OT_1C and OT_2C, we obtain two relations for the angles φ and θ, from which φ and θ can be obtained:

$$f\tan(\varphi + \theta) = \delta_M + a \tag{2}$$
$$f\tan(\varphi - \theta) = \delta_M - a \tag{3}$$

2. From angle θ we can derive the distance $|OP|$ (in world units). Indeed, if H_2 denotes the point of tangency of OT_2 on the sphere, the following relation is observed in the right-angle triangle OPH_2 (Fig. 2):

$$|OP| = \frac{R}{\sin\theta} = R\frac{\sqrt{\tan^2\theta + 1}}{\tan\theta} \tag{4}$$

with R the known ball radius.

3. From angle φ, using Proposition 1, we derive the image of the sphere center $Q = (u_1, v_1)$ (in pixel units):

$$Q = C + \frac{f \cdot \tan\varphi}{\delta_M}(M - C) \tag{5}$$

with the coordinates of the principal point C and distance $\delta_M = |CM|$ given in pixel units.

Conclusion We recover the position P (in world units) by

$$P = \frac{|OP|}{|OQ|} \cdot (u_1 - u_0, v_1 - v_0, f) \tag{6}$$

where we can compute $|OQ|$ by Pythagoras' theorem:

$$|OQ| = \sqrt{(u_1 - u_0)^2 + (v_1 - v_0)^2 + f^2} \tag{7}$$

with $C = (u_0, v_0)$ and $Q = (u_1, u_2)$.

This algorithm for locating the sphere center will be referred to as the *ellipse fitting method*, because it requires the knowledge of the elliptic ball image. It will be compared

to the proposed robust method as well as to the approximating robust method by Guan et al. (2015). We believe that this ellipse fitting method is a better reference for validation than the method of Agrawal and Davis (2003) (as done by Guan et al. (2015)), because it is a straightforward localization, while the procedure of Agrawal and Davis (2003) is mainly designed to intrinsically calibrate the camera.

4 Robust Localization of a Sphere Center Relative to a Calibrated Camera

In this section we avoid the computation of the elliptic boundary of the ball image $\mathcal{E}$. Now we only assume the knowledge of the center M and the area $A = \pi ab$ of $\mathcal{E}$. Because M is the mean and A is the number of the pixels belonging to $\mathcal{E}$, both can be considered as robust measurements. Furthermore, A (and consequently M) can be calculated by means of normalized grey values (between 0 and 1), that can be interpreted as membership probabilities with respect to $\mathcal{E}$, such that on the edge no hard decisions (thresholds) are necessary. This strategy appeared to be appropriate in our experiments, since we acquired sphere images under controlled background light. When other light sources are present, the grey values of the pixels in the neighborhood of the edge of the sphere images might be caused by light reflections, such that they do not represent the uncertainty due to image noise.

In any case, we assume that the area A is known to us, and so, using Eq. (1), two equations relating a and b are available. Once we determined a from these equations, the computations of the previous section apply for locating the center of the ball, since the ellipse center M is supposed to be known as well.

We give the details. From Eq. (1) we deduce:

$$a^2 = b^2 \frac{\delta_M^2 + f^2 + b^2}{f^2 + b^2} \tag{8}$$

Substituting in $A^2 = \pi^2 a^2 b^2$, we find the following cubic equation for b^2:

$$\pi^2 b^6 + \pi^2(\delta_M^2 + f^2)b^4 - A^2 b^2 - f^2 A^2 = 0 \tag{9}$$

This cubic equation in the unknown b^2 can be solved by the method of Cardano–Tartaglia (Barbeau 1989). Because we know that the ellipse $\mathcal{E}$ does exist (as the observed ball image), we know that Eq. (9) has at least one positive real root b^2 (this is also implied by Proposition 3). Once we have computed b, we find a from the known area A, and the procedure of Sect. 3 applies to localize the ball center P.

Fortunately, due to the next proposition, this algebraic method does not generate multiple solutions, avoiding the task to get rid of false solutions for b.

Proposition 3 *The cubic equation* $\pi^2 X^3 + \pi^2(\delta_M^2 + f^2)X^2 - A^2 X - f^2 A^2 = 0$ *has exactly one positive real root.*

Proof We can use the rule of Descartes to determine the positive real roots of a polynomial, because the signs of the coefficients are well determined in our case (Barbeau 1989). This sign sequence is given by $+\ +\ -\ -$. So the number v of sign switches equals 1. According to Descartes' rule, the number n of positive real roots of a given polynomial obeys the following rules:

1. $n \leq v$
2. $v - n$ is even

The fact that $v = 1$ leaves us with only one possibility: $n = 1$. □

5 An Approximation Formula for Robust Sphere Localization

In Guan et al. (2015) the authors suggest an alternative algorithm for the localizing of a sphere with known radius by means of one image. They also avoid ellipse fitting and make only use of the robust parameters A (area) and M (ellipse center) of the ball image. Their approach differs from ours by the fact that their formula is not exact but rather an approximation. More precisely, they postulate two assumptions:

1. The projection Q of the ball center P approximately coincides with the center M of the ball image: $M \approx Q$.
2. The size of the ball is very small relative to its distance to the sensor: $R << |OP|$.

Notice that second assumption yields a small ball image, implying a small distance $|M - Q|$, which tends to the first assumption. Also, the first assumption is justified for sphere images close to the image center. In spite of the restriction dictated by these assumptions, the experiments in Guan et al. (2015) validate an accurate estimation of the sphere position, even outperforming the exact conic-based method of Agrawal and Davis (2003). Furthermore, Guan et al. (2015) obtains pleasing results in multi-camera calibration while using their sphere localization. So we believe it is necessary to compare our exact robust sphere localization to their approximation for sphere position.

In our notations (Sect. 3) the approximation formula of Junzhi is given by (we refer to Guan et al. 2015 for the details):

$$\begin{aligned} P &\approx \frac{|OP|}{|OM|} \cdot (u_M - u_0, v_M - v_0, f) \\ &\approx R\sqrt{\frac{\pi}{A\cos\varphi}}(u_M - u_0, v_M - v_0, f). \end{aligned} \tag{10}$$

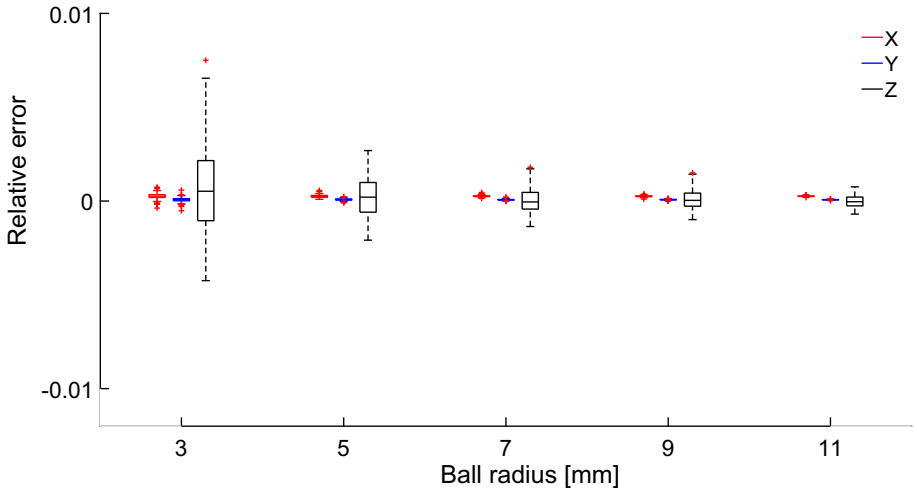

Fig. 3 For each chosen radius we sample a uniform data set of 100 spheres in the defined work space. Under controlled circumstances (no distortion, correct calibration) the proposed exact robust method appears to be very accurate, with improving stability for larger radii

The algorithm for obtaining this approximated sphere location is definitely simpler than the proposed method as presented in Sect. 4, although the time profit is not significant. Indeed, both algorithms need to obtain the same parameters of the sphere image (the area A and the ellipse center M). The only additional computational effort by the proposed algorithm is solving a cubic equation for obtaining the projected sphere center Q, which is not significant. However, the approximation of Junzhi is even more robust than the exact robust algorithm, because the approximated sphere center is a backprojection of the ellipse center M (hardly sensitive to noise), while the proposed method suffers from noise propogation in the computation of Q. This will be illustrated in the experiments of Sect. 6, but we will also show that from the moment the sphere images are sufficiently large, the exact method outperforms the approximation, even at large image noise levels.

6 Validation by Simulation

In this section we evaluate the accuracy and the precision of the proposed exact robust method for localizing the center of a sphere (Sect. 4) from by means of simulations. We compare our performance to the approximation formula of Junzhi (Sect. 5) and to the ellipse-fitting method as presented in Sect. 3. In order to prepare the sphere image for ellipse-fitting, we experimented with several standard available edge detectors and decided for the popular Canny edge detector (Canny 1986) because of its excellent performance for our type of images that have been acquired with background light for the spheres (in the simulations as well as in the real experiments). Due to the sharp edges of the sphere images, the Canny edge detector is preferable to more recent and advanced algorithms such as Xie and Tu (2015). For the actual ellipse fitting we preferred the algorithm of Fitzgibbon et al. (1996), that appeared to give the best results when a sharp thin edge is available. In essence, the proposed method and the ellipse fitting method both reconstruct the sphere center in the camera reference system by means of Eq. (6). The difference is that the proposed method computes the center M of the ellipse by averaging the sphere image pixels, and the half major axis a by means of the ellipse area A and solving Eq. (9). We repeat that the measurements A and M are robust and obtained in a fast and simple way, with no need for edge detection or ellipse fitting. We also evaluate the robust method of Guan et al. (2015), restricting to the robust measurements M and A as well, but recovering the sphere center by means of the approximation given by Eq. (10).

In this experiment we rendered realistic images of spheres by the `Maxon` software `Cinema 4D` (https://www.maxon.net/en-us/products/cinema-4d/overview/), simulating image noise by quantization. Consequently, we have perfect knowledge about the ground truth (spatial coordinates of the sphere centers). Furthermore, we are able to compare the three methods in perfect circumstances (perfectly calibrated camera) and we can easily investigate the sensitivity with respect to independently varying parameters. We simulated a camera with a resolution of 1024 by 768 pixels on a 4×3 mm sensor. We chose a FOV equal to $18.925° \times 14.250°$, corresponding to a focal length of 12 mm or 3072 px. The principal point is taken to be the geometric midpoint of the sensor. The observed spheres are situated in a workspace bounded by $-100\,\text{mm} \leq X \leq 100\,\text{mm}$, $-75\,\text{mm} \leq Y \leq 75\,\text{mm}$ and $700\,\text{mm} \leq Z \leq 850\,\text{mm}$. The statistics in this section have been computed by a sample of 100 sphere images. The measured errors on the computed coordinates or on the distance of the sphere center to the camera are always presented relative to the distance of the ground truth centers to the camera.

The boxplots in Fig. 3 only validate the proposed method, showing the computed spatial coordinates for the centers of spheres randomly distributed in the defined workspace, con-

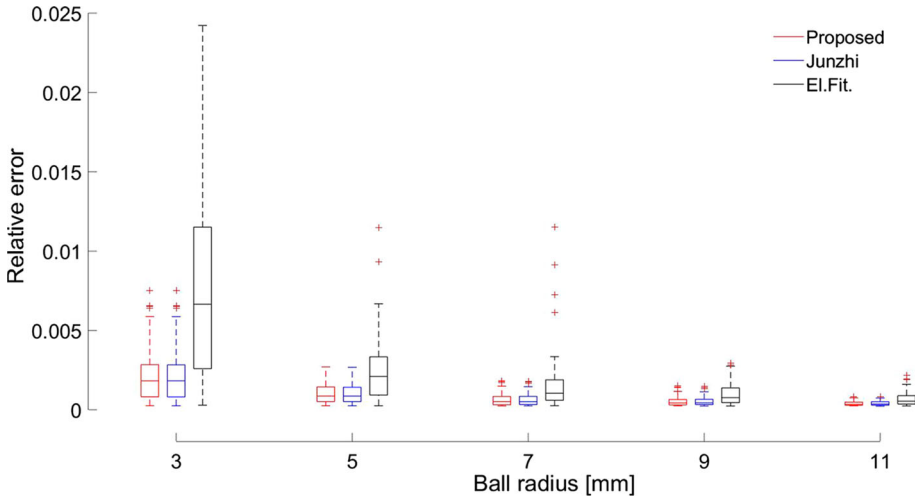

Fig. 4 Under the controlled circumstances (no distortion, correct calibration) the robust methods outperform the ellipse fitting method, especially for smaller radii

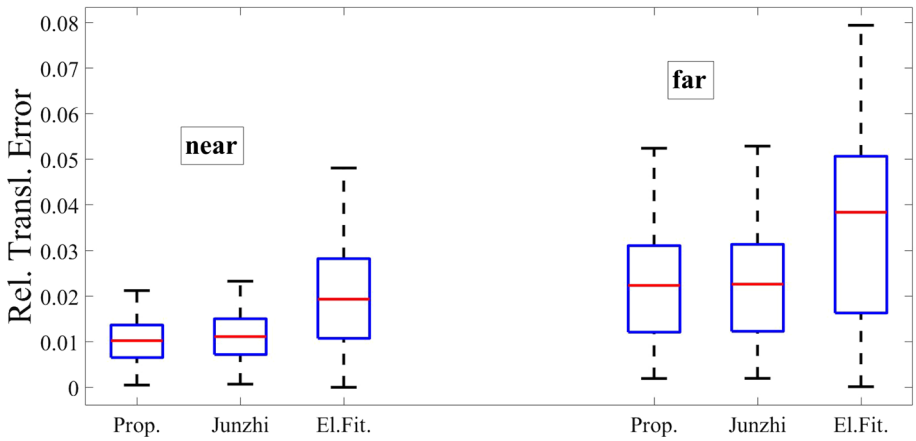

Fig. 5 The presence of radial distortion has a significant negative effect on the accuracy of the three considered methods. The stability of the area methods seems to be lost in case of large distortion

sidering samples for each choice out of 5 possible radii. We observe that our method is very accurate and precise, with the less best results for the computed Z-coordinates, with no relative errors larger than 0.5%. Our method is more precise when using larger spheres (implying larger sphere images). The robust methods clearly perform better than the ellipse fitting method, due to the sensitivity of the latter for edge uncertainties. The sensitivity is more manifest for smaller balls. We also observe that the approximation method performs equally well as our exact method. This is explained by the fact that the used sphere radii are small relative to the sphere distances, (as a consequence of the FOV of the simulated camera), a circumstance in favour of the approximation method (Guan et al. 2015).

In Fig. 4 we compare the three methods for different ball sizes, randomly positioned on a depth $Z = 700$ mm. For layout reasons we only present the distance error between the ground truth and the computed centers.

Next, we compare the performance of the location algorithms under calibration uncertainties. In Fig. 5 we investigate the performance of the three methods in competition on images that are not (or not sufficiently) calibrated for nonlinear lens distortion. More precisely, the ball images rendered by `Cinema 4D` are radially distorted in the sense that each pixel has been moved away from the image center such that its distance r to this image center increases to $r \cdot (1 + kr^2)$ (one distortion parameter k). We observe that the ignorance or the incomplete removal of radial distortion is very nefast for the computation of the sphere position, leading quickly to several whole percentages of relative error. For small values of k the robust area methods seem to have lost their leading position with respect to the ellipse fitting method, but the increasing of k affects the ellipse fitting more for the worse than the other methods.

Finally, we investigated the effect of a calibration error for the focal length (Fig. 6) and for the principal point (Fig. 7). Although the robust methods seem to lose their leading position with respect to the ellipse fitting method when the errors on the calibration intrinsics increase, both in accuracy and in precision, the damage appears to be restricted, not exceeding 0.5% of relative error, hence at a much smaller extent than the distortion effects.

In order to assess the benefit of the proposed exact method with respect to Junzhi's approximation, we produced syn-

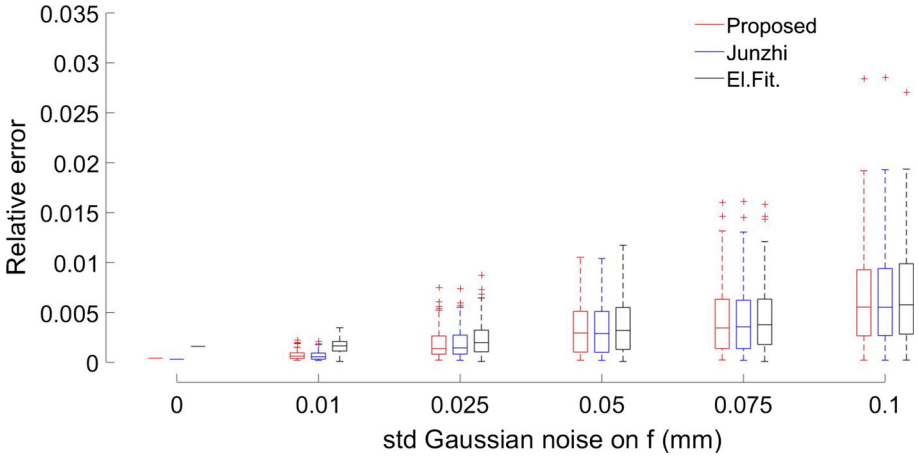

Fig. 6 The three algorithms operate on digitized rendered but undistorted images of randomly postioned spheres with radius 11 mm. The principal point is exactly known, but the correct focal length of 12 mm is disturbed by additive Gaussian noise before it is used in the computation

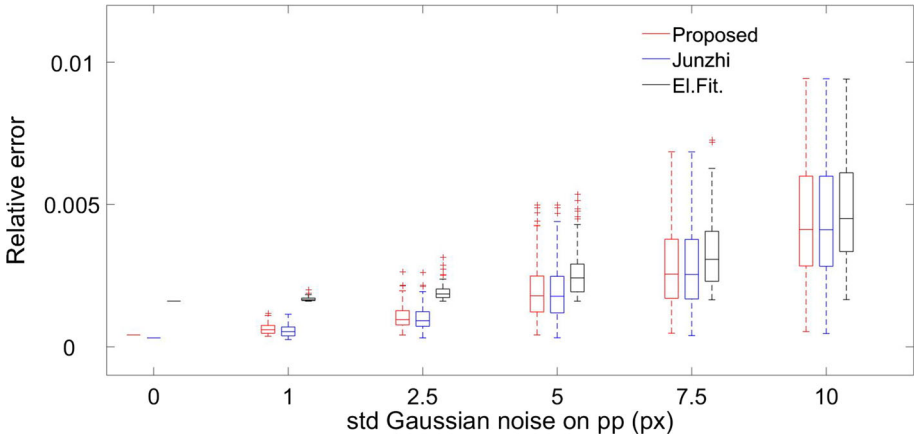

Fig. 7 The three algorithms operate on digitized rendered but undistorted images of randomly postioned spheres with radius 11 mm. In the computation the correct focal length has been used ($f = 12$ mm), but the principal point is disturbed by additive Gaussian noise up to $\sigma = 10$ px

thetic data by perspectively projecting spheres with radius 50, randomly placed at depth 200 and at a distance of 240 to the focal axis, using a focal length $f = 3$. Points on these sphere image contour are disturbed by Gaussian noise, with σ increasing up to 0.1, that is more than 3% of the focal length. More precisely, we produced a run of 40 sphere images at each noise level. An example of a synthetic sphere image at the largest considered noise level is given in Fig. 8. Because the sphere images are large and away from the image center, the circumstances are less in favour for the approximation method. Indeed, the difference between the ellipse center M and the projection of the sphere center Q is not negligible anymore. In Fig. 9 we displayed the success of our method to recover the sphere center projection Q for several noise levels. The approximation error $Q \approx M$ in Junzhi's method is hardly affected by pixel noise due to the robustness of M, but is about a factor 10 larger than the error in our exacty method for computing Q.

By considering spheres images similar to Fig. 8, the computation of Q in the exact method is significantly more accurate than the approximation $Q \approx M$. On the other hand, the final sphere location of the exact method depends on the accuracy of the computed projection Q, and hence, it is

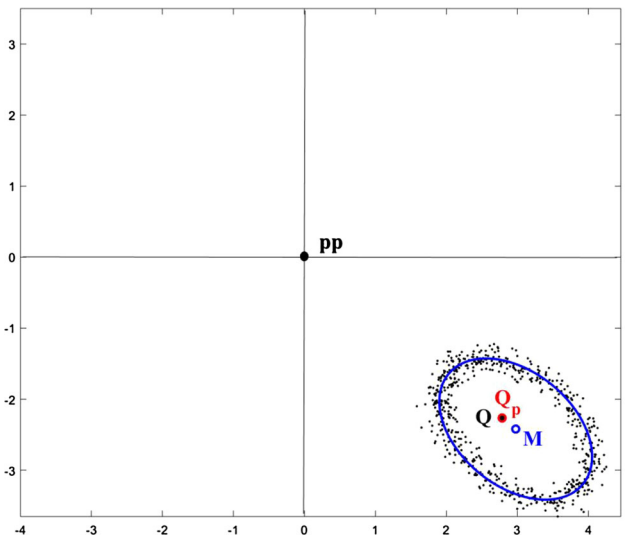

Fig. 8 For large sphere images, away from the image center, the difference between the ellipse center M and the projection Q of the sphere center becomes clear. The proposed method yields an accurate computation Q_p voor Q, while Junzhi's method assumes $Q \approx M$

definitely sensitive for noise, while Junzhi's method is less affected by noise. This phenomenon has been visualized in

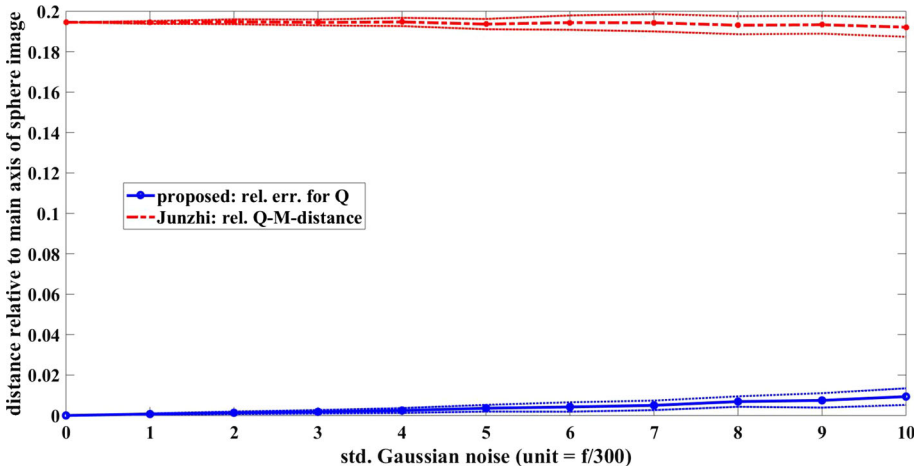

Fig. 9 The error by the proposed method for computing Q is surely affected by pixel noise, but remains significantly smaller that the approximation error $Q \approx M$ (which is less affected by noise). The error is expressed by relating the pixel distance to the length of the main ellipse axis of the sphere image, in order to have a resolution-independent error measure

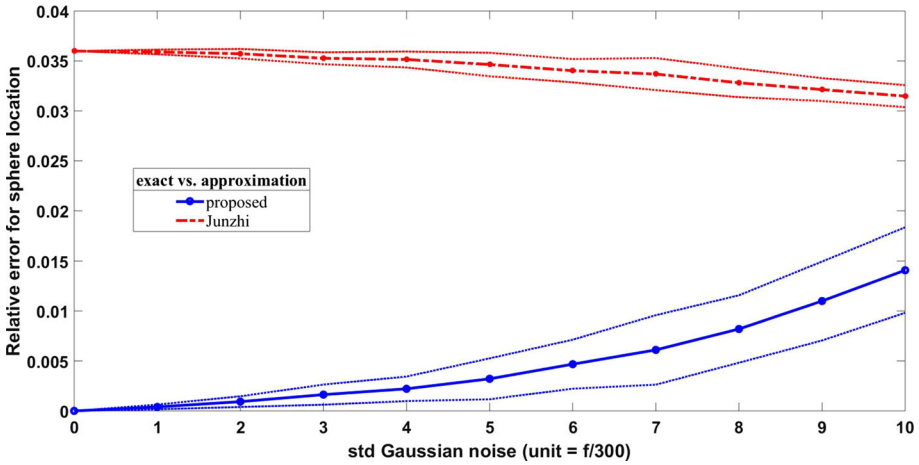

Fig. 10 The location error of both methods is expressed relative to the camera-sphere distance

Fig. 10. The proposed method clearly loses some of its benefit while image noise has been increased, both with respect to accuracy as to precision. However, for the considered sphere positions, even when assuming image noise with $\sigma = f/30$, the approximation error is at least twice the error of the proposed method.

Finally, the trade off between the simple approximation formula and the proposed analytic formula involving the solution of a cubic equation, is perhaps most transparant in Fig. 11. This figure reports a synthetic experiment with varying sphere radii (from 15 to 60). For the several sphere radii we present the mean and the standard deviation of the relative error for the sphere center as computed by Junzhi's approximation formula and by the proposed method in a run of 40 random spheres at a fixed depth of 200 and a focal axis offset of 240, where the boundary points of the projected spheres are disturbed by Gaussian image noise with standard deviation equal to $f/30$.

Initially, both methods bcome more accurate for larger spheres, because the proportion of the constant image noise relative to the size of the sphere images decreases. The interesting observation is that Junzhi's method has a turning point for spheres with a radius of about 20% of the sphere depth. The approximation error of Junzhi's method becomes more significant for larger spheres, yielding circumstances where the proposed method is preferred.

7 Validation by Real Experiments

In this section we report some real experiments testing the proposed method, in comparison with Junzhi's method and the ellipse fitting method. To this end, we acquired images of a professional biljard ball with a diameter of 61.5 mm by an RGB camera with a 1280×1024 resolution and an 8 mm lens. The camera was rigidly mounted to the end effector of an articulated robot arm (Fig. 12). The ball was kept fixed and viewed from several camera positions that were controlled by the robot while allowing only pure translations for the end effector. Consequently, the ground truth for the distance between two relative postions of the fixed ball with respect to the translating camera is provided by the translational distance between two distinct tool center positions of the robot. So, we validate a sphere location method indirectly by com-

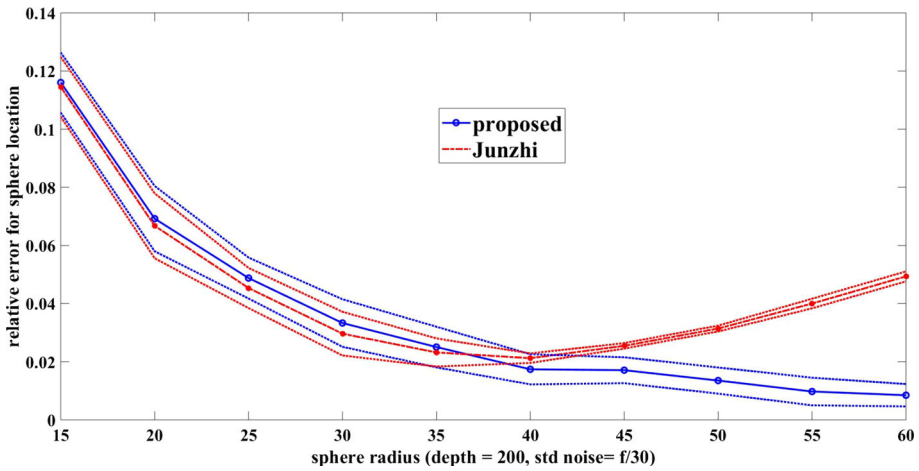

Fig. 11 For a fixed image noise level (Gaussian noise with $\sigma = f/3$) and a fixed sphere distance (depth 200, offset 240 from focal axis), we plot the sphere location error for both robust methods in function of an increasing spere radius. The location error is considered relative to the sphere distance

puting the distance between two sphere positions belonging to a pair of images and comparing it to the known translational distance.

The camera was calibrated by a standard toolbox, based on the classical chess pattern method of Zhang (2000), removing radial distortion by minimizing the reprojection error.

In the experiment we took 30 images nearby the viewed ball (at depths of less than 40 cm), and as many images at a larger distance (at depths between 80 and 90 cm), thus generating datasets for two depth levels of the measured spheres (Fig. 13). At both depth levels we can check 435 image pairs and compare the proposed method to the ellipse fitting method and Junzhi's approximation.

Figure 14 shows the results of the three methods performing on the undistorted images, and this for both depth levels. For each image pair at each depth level we measure the error of the distance between the computed sphere positions with respect to the ground truth as provided by the robot manipulator. Errors are always presented relative to the mean distance of the considered pair of spheres (as obtained by the proposed method). The means and standard deviations of these relative errors can also been found in Table 1.

The real experiments confirm the observations in the simulations of Sect. 6. Indeed, the robust area methods perform very stable in computing the postion of a sphere within a 1% accuracy (relative to the distance of the viewed sphere). The area methods outperform the ellipse fitting method both in accuracy and precision, while in both aspects the proposed method is a little bit better than Junzhi's method. As expected, the gain of the exact area method with respect to the approximating area method increases for larger sphere images. For all methods we observe a loss of quality of the resulting sphere localization as the viewed sphere tends to be at a larger distance to the camera, even if the quality is considered relative to this distance.

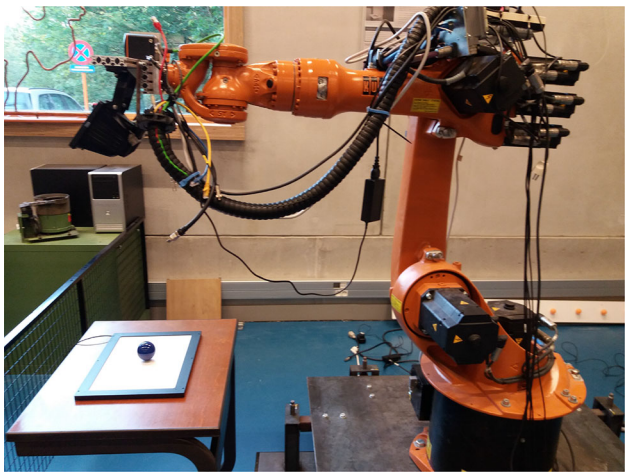

Fig. 12 We provide ground truth for a sphere location method by way of the known translational distance of the RGB-camera that is attached to a robot arm controlled such that the end effector only performs pure translations

In Fig. 15 we applied the sphere location methods without having removed the radial distortion. As predicted by the synthetic experiments, radial distortion appears to have a significant effect on the accuracy of the computed sphere postion. This effect is the largest for the ellipse fitting method, and the smallest for the proposed area method.

Finally, Fig. 16 reports the "lazy user modus", assuming square pixels, a principal point in the geometric image center, taking the focal length as claimed by the manufacturer and ignoring nonlinear lens distortions. As predicted by the synthetic experiments, variations on the pinhole parameters have far less impact compared to the presence of lens distortion. For applications with spheres on reasonable distance, allowing relative errors up to 2%, the proposed method can be used without the need for any calibration procedure.

Table 1 An overview of the performance of the three considered methods under different measuring conditions

Depth level	Near		
Method	Proposed	Junzhi	Ellipse fitting
Undistorted	0.0089 ± 0.0041	0.0099 ± 0.0045	0.0171 ± 0.0100
Distorted	0.0101 ± 0.0046	0.0111 ± 0.0050	0.0199 ± 0.0114
Lazy cal.	0.0109 ± 0.0049	0.0120 ± 0.0054	0.0216 ± 0.0125
Depth level	Far		
Method	Proposition	Junzhi	Ellipse fitting
Undistorted	0.0203 ± 0.0109	0.0206 ± 0.0110	0.0318 ± 0.0190
Distorted	0.0225 ± 0.0117	0.0227 ± 0.0117	0.0360 ± 0.0208
Lazy cal.	0.0241 ± 0.0126	0.0244 ± 0.0127	0.0390 ± 0.0226

The performance is presented by the mean and the standard deviation of the relative errors of 435 image pairs with respect to their distance to the camera

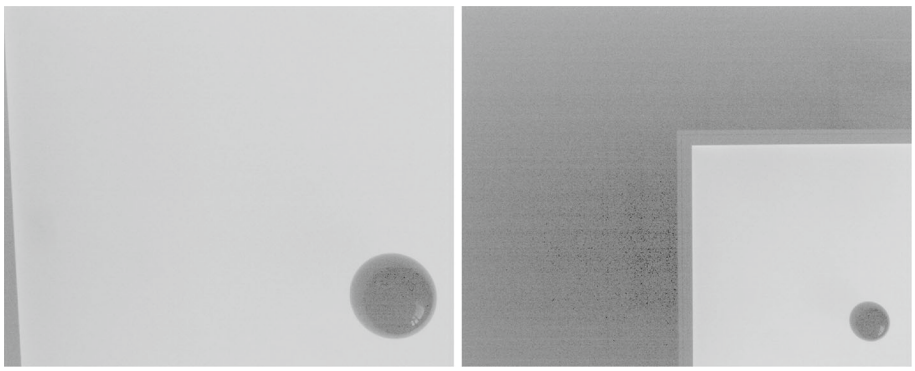

Fig. 13 One sample image for both considered depth levels of the viewed sphere

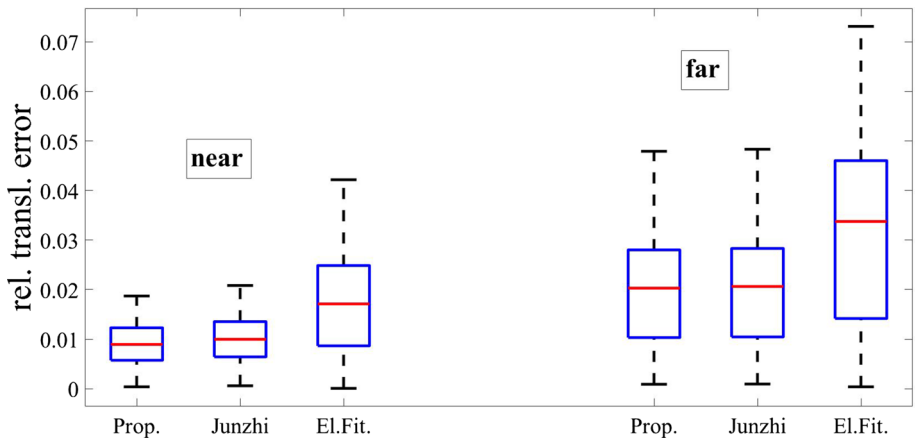

Fig. 14 These boxplots show the relative translation errors of the proposed exact method compared to Junzhi's approximation and the ellipse fitting method, performed on 435 image pairs at two depth levels for the measured sphere. The pinhole and lens distortion parameters have been obtained with a classical chess pattern procedure

8 Conclusions and Future Work

We deduced a formula for the position of a sphere center with known radius, using the center M and the major size a of the elliptic image, and the intrinsics of the pinhole camera (Eq. 6). Further, using basic polynomial algebra and Eq. (1), we showed how to obtain a from the ellipse area A, yielding a robust localization method. This formula does not depend on assumptions for approximation as it is the case in the existing area formula of Junzhi. From our synthetic and real experiments we conclude:

- For a reliably calibrated camera, the robust area methods are more accurate and precise than an ellipse fitting method.
- For spheres with a small size relatively to their depth, the approximation error of Junzhi is compensated by the image noise. Only in case of large sphere images (radius at leat 20% the depth of the sphere location relative to the

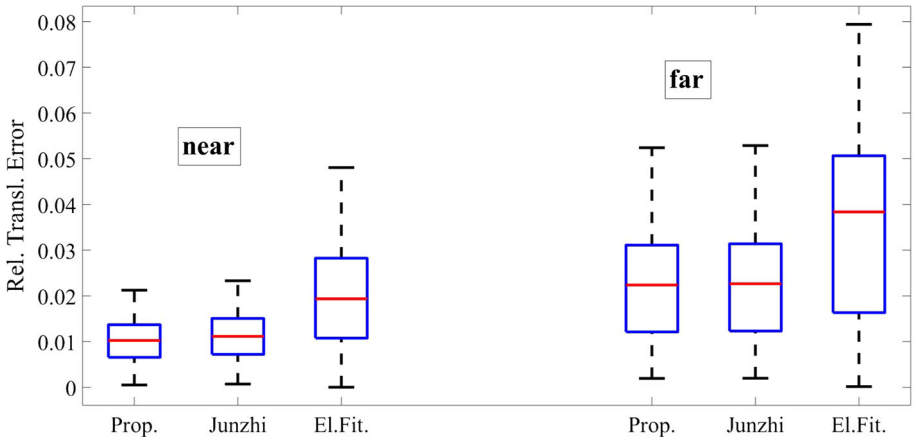

Fig. 15 These boxplots show the relative translation errors of the proposed exact method compared to Junzhi's approximation and the ellipse fitting method, performed on 435 image pairs at two depth levels for the measured sphere. The pinhole parameters have been obtained with a classical chess pattern procedure, but no radial distortion has been removed

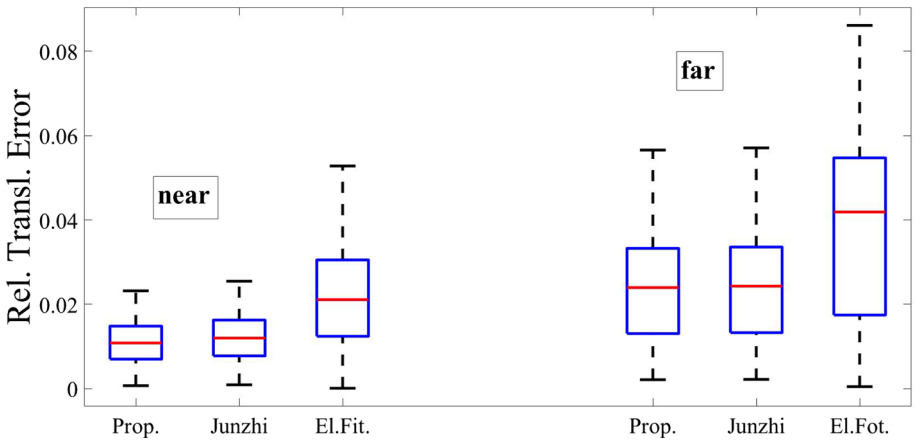

Fig. 16 These boxplots show the relative translation errors of the proposed exact method compared to Junzhi's approximation and the ellipse fitting method, performed on 435 image pairs at two depth levels for the measured sphere. The lens distortion was not corrected, the principal point was identified with the geometric image center, and the focal length was taken equal to 1509.4 px corresponding to 8 mm

camera), the proposed exact method appears to be more accurate.

- All methods are seriously affected by radial distortion. Our experiments confirm that the removal of radial distortion is necessary in applications that require very accurate sphere localization. In this context we refer to Agrawal and Davis (2003) for a sphere based method to remove radial distortion.
- On the other hand, inaccuracies on the pinhole parameters (principal point, focal length) have less impact than the presence of radial distortion. For some applications, the robust area methods deliver reliable sphere positions when using cameras without any calibration preprocessing.

In the future we hope to extend our ellipse area computations to a robust sphere based calibration method, upgrading the algorithm of Agrawal and Davis (2003), in order to obtain a global ellipse area based strategy for extrinsic and intrinsic calibration (including the removal of nonlinear lens distortion).

References

Agrawal, M., & Davis, L. S. (2003). Camera calibration using spheres: A semi-definite programming approach. In: *IEEE international conference on computer vision* (pp. 782–789).

Arun, K. S., Huang, T. S., & Blostein, S. D. (1987). Least-squares fitting of two 3-D point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *9*(5), 698–700.

Ballard, D. H. (1981). Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, *13*(2), 111–122.

Barbeau, E. (1989). *Polynomials*. New York: Springer.

Beardsley, P., Murray, D., & Zisserman, A. (1992). Camera calibration using multiple images. In *Proceedings of the second European conference on computer vision, ECCV* (pp. 312–320).

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *8*(6), 679–698.

Dandelin, G. P. (1822). Mémoire sur quelques propriétés remarquables de la focale parabolique. *Nouveaux mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles T*, *II*, 171–202.

Daucher, D., Dhome, M., & Lapreste, J. (1994). Camera calibration from spheres images. In *Proceedings European conference computer vision* (pp. 449–454).

Fitzgibbon, A., Pilu, M., & Fisher, R. (1996). Direct least-square fitting of ellipses. In *Proceedings international conference on pattern recognition* (pp. 253–257).

Guan, J., Deboeverie, F., Slembrouck, M., van Haerenborgh, D., van Cauwelaert, D., Veelaert, P., et al. (2015). Extrinsic calibration of

camera networks using a sphere. *Sensors*, *15*(8), 18985–19005. https://doi.org/10.3390/s150818985.

Ho, C., & Chen, L. (1995). A fast ellipse/circle detector using geometric symmetry. *Pattern Recognition*, *28*(1), 117–124.

Lu, Y., & Payandeh, S. (2010). On the sensitivity analysis of camera calibration from images of spheres. *Computer Vision and Image Understanding*, *114*(1), 8–20.

Maxon. (2009). Cinema 4D. https://www.maxon.net/en-us/products/cinema-4d/overview/.

Ouellet, J. N., & Hébert, P. (2008). Precise ellipse estimation without contour point extraction. *Machine Vision and Applications*, *21*(1), 59. https://doi.org/10.1007/s00138-008-0141-3.

Penna, M. (1991). Camera calibration: A quick and easy way to determine the scale factor. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *13*(12), 1240–1245.

Penne, R., Ribbens, B., Mertens, L., & Levrie, P. (2015). What does one image of one ball tell us about the focal length? *Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Science*, *9386*, 501–509.

Sun, J., Chen, X., Gong, Z., Liu, Z., & Zhao, Y. (2015). Accurate camera calibration with distortion models using sphere images. *Optics Laser Technology*, *65*, 83–87.

Teramoto, H., & Xu, G. (2002). Camera calibration by a single image of balls: From conics to the absolute conic. In *Asian conference on computer vision* (pp. 499–506).

Xie, S., & Tu, Z. (2015). Holistically-nested edge detection. CoRR. **abs/1504.06375**. http://arxiv.org/abs/1504.06375.

Xie, Y., & Ji, Q. (2002). A new efficient ellipse detection method. In *International conference on pattern recognition* (pp. 957–960).

Yin, P., & Chen, L. H. (1994). A new method for ellipse detection using symmetry. *Journal of Electronic Imaging*, *3*, 20–29.

Zhang, H., Wong, K., & Zhang, G. (2007). Camera calibration from images of spheres. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(3), 499–503.

Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(11), 1330–1334.