



Generic Primitive Detection in Point Clouds Using Novel Minimal Quadric Fits

Tolga Birdal, Benjamin Busam, Navab Nassir, Slobodan Ilic, Peter Sturm

► To cite this version:

Tolga Birdal, Benjamin Busam, Navab Nassir, Slobodan Ilic, Peter Sturm. Generic Primitive Detection in Point Clouds Using Novel Minimal Quadric Fits. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42 (6), pp.1333-1347. 10.1109/TPAMI.2019.2900309 . hal-02368932

HAL Id: hal-02368932

<https://inria.hal.science/hal-02368932>

Submitted on 18 Nov 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.

Generic Primitive Detection in Point Clouds Using Novel Minimal Quadric Fits

Tolga Birdal, Benjamin Busam, Nassir Navab, Slobodan Ilic, Peter Sturm, *Members, IEEE*

Abstract—We present a novel and effective method for detecting 3D primitives in cluttered, unorganized point clouds, without axillary segmentation or type specification. We consider the quadric surfaces for encapsulating the basic building blocks of our environments - planes, spheres, ellipsoids, cones or cylinders, in a unified fashion. Moreover, quadrics allow us to model higher degree of freedom shapes, such as hyperboloids or paraboloids that could be used in non-rigid settings.

We begin by contributing two novel quadric fits targeting 3D point sets that are endowed with tangent space information. Based upon the idea of aligning the quadric gradients with the surface normals, our first formulation is exact and requires as low as four oriented points. The second fit approximates the first, and reduces the computational effort. We theoretically analyze these fits with rigor, and give algebraic and geometric arguments. Next, by re-parameterizing the solution, we devise a new local Hough voting scheme on the null-space coefficients that is combined with RANSAC, reducing the complexity from $O(N^4)$ to $O(N^3)$ (three points). To the best of our knowledge, this is the first method capable of performing a generic cross-type multi-object primitive detection in difficult scenes without segmentation. Our extensive qualitative and quantitative results show that our method is efficient and flexible, as well as being accurate.

Index Terms—Quadrics, Surface Fitting, Implicit Surfaces, Point Clouds, 3D Surface Detection, Primitive Fitting, Minimal Problems

1 INTRODUCTION

Surface fitting and detection enjoys a rich history in computer vision and graphics communities. The problem is found particularly important because of the power of 3D surfaces to explain generic man-made structures omnipresent in every day life. Many of the constructed or manufactured objects and architecture that surrounds us are results of careful computer aided design (CAD). Some of the primary concerns of 3D computer vision, mapping and reconstruction, try to associate the visual cues acquired by various 2D / 3D sensors with those idealized CAD models, that are used in the assembly of our environments.

One family of approaches tries to find a direct rigid association between those CAD models and 3D scenes [1], trying to solve the six degree of freedom (DoF) pose estimation problem. While these approaches are quite successful as the only parameters to discover are rotations and translations, they require a huge number of CAD models to generically represent the real scenarios. To overcome this limitation, inspired by the fact that all CAD models are designed using a similar set of tools, a different line of research attempts to find common bases explaining a broad set of 3D objects, and tries to detect these bases instead of individual models. Such bases that are the common building blocks of our world, and typically termed *geometric primitives*. While the approaches using bases, significantly reduce the database size, usually, the bases undergo higher dimensional transformations compared to, for instance, rigid ones. Examples of the geometric primitives are

splines and nurbs surfaces defined by several control points, or *quadrics*, the three dimensional, nine-DoF, quadratic forms.

Thanks to their power to embody the most typical geometric primitives, such as planes, spheres, cylinders, cones or ellipsoids, quadrics themselves were of huge interest since 80s [2]. Some exemplary studies involve recovering 3D quadrics from image projections [3], fitting them to 3D point sets [4], or detecting special cases of the quadratic forms [5]. A majority of those works either put emphasis on fitting to a noisy, but isolated point set [4], [6], [7], or restrict the types of shapes under consideration (thereby reduce the DoF) to devise detectors robust to clutter and occlusions [8], [9], [10].

In this work, our aim is to unite the fitting and detection worlds and present an algorithm that can simultaneously estimate all parameters of a generic nine DoF quadric, which resides in a 3D cluttered environment and is viewed potentially from a single 3D sensor, introducing occlusions and partial visibility. We craft this algorithm in three stages: (1) First, we devise a new quadric fit. Unlike its ancestors, this one uses the extra information about the tangent space to increase the number of constraints instead of regularizing the solution. This fit requires only four oriented points. We show that such construction also has a regularization effect as a by-product. (2) We then thoroughly analyze its rank properties and devise a novel null-space Hough voting mechanism to reduce the four point case to three. Three points stands out to be the minimalist case developed so far. (3) We propose a variant of RANSAC that operates on our local bases, which are randomly posited. Per each local basis, we show how to make use of the fit and the voting to hypothesize a likely quadric. Finally, we use simple clustering heuristics to group and strengthen the candidate solutions. Our algorithm works purely on 3D point cloud data and does not depend upon any acquisition modality. Moreover, it makes assumptions neither about the type of the quadric that is present in the scene nor how many are visible.

This journal paper extends our recent CVPR publication [11].

- T. Birdal, B. Busam, N. Navab and S. Ilic are with the Department of Informatics, Technical University of Munich, Germany
E-mail: tolga.birdal@tum.de, navab@cs.tum.de
- T. Birdal and S. Ilic are with Siemens AG, Munich, Germany.
E-mail: tolga.birdal@siemens.com, slobodan.ilic@siemens.com
- B. Busam is with Framos GmbH, Munich, Germany.
E-mail: b.busam@framos.com
- P. Sturm is with INRIA, Grenoble, France
E-mail: peter.sturm@inria.fr

Manuscript received June XX, 2018.

by providing additionally the following:

- 1) qualitative and quantitative experiments to better grasp the behavior of the proposed fit,
- 2) algebraic and geometric theoretical analysis of the quadric fit
- 3) improved elaborate descriptions of the method as well as accompanying pseudocode.

2 RELATED WORK

2.1 Quadrics

Quadrics appear in various domains of vision, graphics and robotics. They are found to be one of the best local surface approximators in estimating differential properties [12]. Thus, point cloud normals and curvatures are oftentimes estimated with local quadrics [1], [13]. Yan *et al.* propose an iterative method for mesh segmentation by fitting local quadratic surfaces [14]. Yu presented a quadric-region based method for consistent point cloud segmentation [15]. Kukulova uses quadric intersections to solve minimal problems in computer vision [16]. Uto *et al.* [17] as well as Pas and Platt [18] use quadrics to localize grasp poses and in grasp planning. Quadrics have also been a significant center of attention in projective geometry and reconstruction [3], [19] to estimate algebraic properties of apparent contours. Finally, You and Zhang [20] used them in feature extraction from face data.

2.2 Primitive detection

Finding primitives in point clouds has kept the vision researchers busy for a lengthy period of time. Works belonging to this category treat the primitive shapes independently [10], giving rise to specific fitting algorithms for planes, spheres, cones, cylinders, etc. Planes, as the simplest forms, are the primary targets of the Hough-family [21]. Yet, detection of more general set of primitives made RANSAC the method of choice as shown by the prosperous Globfit [22]: a relational local to global RANSAC algorithm. Schnabel *et al.* [23] and Tran *et al.* [24] also focused on reliable estimation using RANSAC. Monszpart *et al.* [25] proposed a greedy heuristic improving upon its randomized counterparts in plane detection. Oesau *et al.* [26] proposes a tandem scheme for plane detection (by region-growing) and regularization to correct the imperfections of the hypotheses. The latter two improve upon Globfit by simultaneously extracting the planes and the primitive relations. The plane detection has recently been lifted to structural scales [27], [28]. An interesting application of primitives is given by Qui *et al.* who extract pipe runs using cylinder fitting [29]. The local Hough transform of Drost and Ilic [30] showed how the detection of primitives can be made more efficient by considering the local voting spaces. Authors give sphere, cylinder and plane specific formulations targeting point clouds. Lopez *et al.* [31] devise a robust ellipsoid fitting based on iterative non-linear optimization. Sveier *et al.* [32] suggest a conformal geometric algebra to spot planes, cylinders and spheres. Andrews' approach [5] deals with paraboloids and hyperboloids in CAD models. Even though this is slightly more generalized, paraboloids or hyperboloids are not the only geometric shapes described by quadrics. Methods in this category are quite successful in shape detection, yet they handle the primitives separately. This prevents automatic type detection, or generalized modeling of surfaces.

2.3 Quadric fitting

Since the 1990s generic quadric fitting is cast as a constrained optimization problem, where the solution is obtained from a Generalized Eigenvalue decomposition of a scatter matrix. Pioneering work has been done by Gabriel Taubin [4] in which a Taylor approximation to the geometric distance is made. This work has then been enhanced by 3L [7], fitting a local, explicit ribbon surface composed of three-level-sets of constant Euclidean distance to the object boundary. This fit implicitly used the local surface information. Later, Tasdizen [6] improved the local surface properties by incorporating the surface normals as regularizers. This allows for a good and stable fit. Recently, Beale *et al.* [33] introduced the use of a Bayesian prior to regularize the fit. All of these methods use at least nine or twelve [33] points. Moreover, they only use surface normals as regularizers - not as additional constraints and are also unable to deal with outliers in data. There are a few other studies [8], [34], improving these standard methods, but they involve either non-linear optimization [35] or share the common drawback of requiring nine independent constraints and no outlier treatment.

2.4 Quadric detection

Recovering general quadratic forms from cluttered and occluded scenes is a rather unexplored area. A promising direction was to represent quadrics with spline surfaces [36], but such approaches must tackle the increased number of control points, i.e. 8 for spheres, 12 for general quadrics [37], [38]. Segmentation is one way to overcome such difficulties [36], [39]. Besl and Jain suggested a variable order segmentation-based surface fitting. They, too, use an iterative procedure where the primitive order is raised incrementally [40]. This is not very different from performing individual primitive detection. Vaskevicius *et al.* [41] developed a noise-model aware quadric fitting and region-growing algorithm for segmented noisy scenes. However, segmentation, due to its nature, decouples the detection problem in two parts and introduces undesired errors especially under occlusions. Other works exploit genetic algorithms [42] but have the obvious drawback of inefficiency. QDEGSAC [43] proposed a six-point hierarchical RANSAC, but the paper misses out an evaluation or method description for a quadric fit. Petitjean [12] stressed the necessity of outlier aware quadric fitting however only ends up suggesting M-estimators for future research.

Finally, the remarkable performance of deep neural networks (DNN) for learning in 2D image domain [44], [45] have recently been extended to 3D point clouds [46], [47]. While 3D-PRNN [48] and PCPNet [49] are tailored for fitting 3D shape primitives and extracting differential surface properties respectively, their application to our problem of detecting nine-DoF primitives in real scenes containing noise, clutter and occlusions is not immediate. To the best of our knowledge, this remains to be open challenge. We would also like to stress that the community lacks comprehensive primitive detection datasets and this gives the learning algorithms a hard time to grasp all the shape variations of quadrics.

3 PRELIMINARIES

Definition 1. A quadric in 3D Euclidean space is a hypersurface defined by the zero set of a polynomial of degree two:

$$f(x, y, z) = Ax^2 + By^2 + Cz^2 + 2Dxy + 2Exz + 2Fyz + 2Gx + 2Hy + 2Iz + J = 0. \quad (1)$$

Alternatively, the vector notation $\mathbf{v}^T \mathbf{q} = 0$ is used, where:

$$\mathbf{q} = [A \ B \ C \ D \ E \ F \ G \ H \ I \ J]^T \quad (2)$$

$$\mathbf{v} = [x^2 \ y^2 \ z^2 \ 2xy \ 2xz \ 2yz \ 2x \ 2y \ 2z \ 1]^T$$

Using homogeneous coordinates, quadrics can be analyzed uniformly. The point $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$ lies on the quadric, if the projective algebraic equation over $\mathbb{RP}^3$ with $d_q(\mathbf{x}) := [\mathbf{x}^T \mathbf{1}] \mathbf{Q} [\mathbf{x}^T \mathbf{1}]^T = 0$ holds true, where the matrix $\mathbf{Q} \in \mathbb{R}^{4 \times 4}$ is defined by re-arranging the coefficients:

$$\mathbf{Q} = \begin{bmatrix} A & D & E & G \\ D & B & F & H \\ E & F & C & I \\ G & H & I & J \end{bmatrix}, \quad \nabla \mathbf{Q} = 2 \begin{bmatrix} A & D & E & G \\ D & B & F & H \\ E & F & C & I \end{bmatrix}. \quad (3)$$

$d_q(\mathbf{x})$ can be viewed as an algebraic distance function. Similar to the quadric equation, the gradient at a given point can be written as $\nabla \mathbf{Q}(\mathbf{x}) := \nabla \mathbf{Q}[\mathbf{x}^T \mathbf{1}]^T$. Quadrics are general implicit surfaces capable of representing cylinders, ellipsoids, cones, planes, hyperboloids, paraboloids and potentially the shapes interpolating any two of those. All together there are 17 sub-types [50]. Once $\mathbf{Q}$ is given, this type can be determined from an eigenvalue analysis of $\mathbf{Q}$ and its subspaces [51], [52]. Note that quadrics have constant second order derivatives and are practically smooth.

Definition 2. A quadric whose matrix is of rank 2 consists of the union of two planes: $\mathbf{Q} = \Pi_1 \Pi_2^T + \Pi_2 \Pi_1^T$, where Π_1 and Π_2 are the homogeneous 4-vectors representing the two planes. A quadric whose matrix is of rank one consists of a single plane: $\mathbf{Q} = \Pi \Pi^T$.

Definition 3. The polar plane Π of a point $\mathbf{p}$ with respect to a quadric $\mathbf{Q}$ is $\Pi = \mathbf{Q}\mathbf{p}$. Reciprocally, $\mathbf{p}$ is called the pole of plane Π .

Note that if $\mathbf{Q}\mathbf{p} = \mathbf{0}$, then the polar plane does not exist for $\mathbf{p}$; also note that for a point that lies on the quadric, the polar plane is the tangent plane in that point.

Definition 4. A quadric is called central if it possesses a finite center point $\mathbf{c}$ that is the pole of the plane at infinity: $\mathbf{Q}\mathbf{c} \sim [0 \ 0 \ 0 \ 1]^T$ e.g. ellipsoids, hyperboloids.

Definition 5. A dual quadric $\mathbf{Q}^* \sim \mathbf{Q}^{-1}$ is the locus of all planes $\{\Pi_i\}$ satisfying $\Pi_i^T \mathbf{Q}^{-1} \Pi_i = 0$.

Quadric dual space is formed by the Legendre transformation, mapping points to tangent planes as covectors. Given a hypersurface, the tangent space at each point gives a family of hyperplanes, and thus defines a dual hypersurface in the dual space. Every dual point represents a plane in the primal. Many operations such as fitting can be performed in either of the spaces [3]; or in the primal space using constraints of the dual. The latter forms a mixed approach, involving tangency constraints. Note that, knowing a point lies on the surface gives one constraint, and if, in addition, one knows the tangent plane at that point, then one gets two more constraints. In this paper, we will use the extra dual constraints to increase the rank of a linear system that solves for the quadric coefficients. This will in return allow us to perform fits with reduced number of points and thereby to lower the minimum number of required points. In Fig. 1(a) we provide combinations of primal and dual constraints each of which leads to a minimal case. Note that, if we have four points, and associated tangent planes, a fit can be formulated.

	# Pri.	# Dual		# Pri.	# Dual	VS
PD-0	9	0	Plane	1	1	0
PD-1	7	1	2-Planes	2	2	0
PD-2	5	2	Sphere	2	2	1
PD-3	4	3	Spheroid	2	2	3

(a)
(b)

Figure 1. (a) Number of constraints for a minimal fit in Primal(P) or Dual(D) spaces. PD-i refers to i^{th} combination. (b) Number of minimal constraints and voting space size for various quadrics.

4 QUADRIC FITTING TO 3D DATA

4.1 A new perspective to quadric fitting

State of the art direct solvers for quadric fitting rely either solely on point sets [4], or use surface normals as regularizers [6]. Both approaches require at least nine points, posing a strict requirement for practical considerations, i.e. using nine points bounds the possibility for RANSAC-like fitting algorithms as the space of potential samples is N_x^9 where N_x is the number of points. Here, we observe that typical real life point clouds make it easy to compute the surface normals (tangent space) and thus provide an additional cue. With this orientation information, we will now explain a closed form fitting requiring only four oriented points. Similar to gradient-one fitting [6], [53], our idea is to align the gradient vector of the quadric $\nabla \mathbf{Q}(\mathbf{x}_i)$ with the normal of the point cloud $\mathbf{n}_i \in \mathbb{R}^3$. However, unlike $\nabla \mathbf{1}$ [53], we opt to use a linear constraint to increase the rank rather than regularizing the solution. This is seemingly non-trivial as the vector-vector alignment brings a non-linear constraint either of the form:

$$\frac{\nabla \mathbf{Q}(\mathbf{x}_i)}{\|\nabla \mathbf{Q}(\mathbf{x}_i)\|} - \mathbf{n}_i = 0 \quad \text{or} \quad \frac{\nabla \mathbf{Q}(\mathbf{x}_i)}{\|\nabla \mathbf{Q}(\mathbf{x}_i)\|} \cdot \mathbf{n}_i = 1. \quad (4)$$

The non-linearity is caused by the normalization as it is hard to know the magnitude and thus the homogeneous scale in advance. We solve this issue by introducing a per normal homogeneous scale α_i among the unknowns and write:

$$\nabla \mathbf{Q}(\mathbf{x}_i) = \nabla \mathbf{v}_i^T \mathbf{q} = \alpha_i \mathbf{n}_i \quad (5)$$

Stacking this up for all N points $\mathbf{x}_i$ and normals $\mathbf{n}_i$ leads to:

$$\mathbf{A}' \begin{Bmatrix} \mathbf{M} \\ \mathbf{N} \end{Bmatrix} = \mathbf{0} \quad (6)$$

$$\mathbf{A}' = \begin{bmatrix} \mathbf{v}_1^T & 0 & 0 & \cdots & 0 \\ \mathbf{v}_2^T & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_n^T & 0 & 0 & \cdots & 0 \\ \nabla \mathbf{v}_1^T & -\mathbf{n}_1 & \mathbf{0}_3 & \cdots & \mathbf{0}_3 \\ \nabla \mathbf{v}_2^T & \mathbf{0}_3 & -\mathbf{n}_2 & \cdots & \mathbf{0}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \nabla \mathbf{v}_n^T & \mathbf{0}_3 & \mathbf{0}_3 & \cdots & -\mathbf{n}_n \end{bmatrix} \begin{bmatrix} A \\ B \\ \vdots \\ I \\ J \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

where $\nabla \mathbf{v}_i^T = \nabla \mathbf{v}(\mathbf{x}_i)^T$, $\mathbf{0}_3$ is a 3×1 column vector of zeros, $\mathbf{A}'$ is $4N \times (N + 10)$ and $\boldsymbol{\alpha} = \{\alpha_i\}$ are the unknown homogeneous scales. The solution containing quadric coefficients and individual scale factors lies in the null-space of $\mathbf{A}'$, and can be obtained accurately via Singular Value Decomposition.

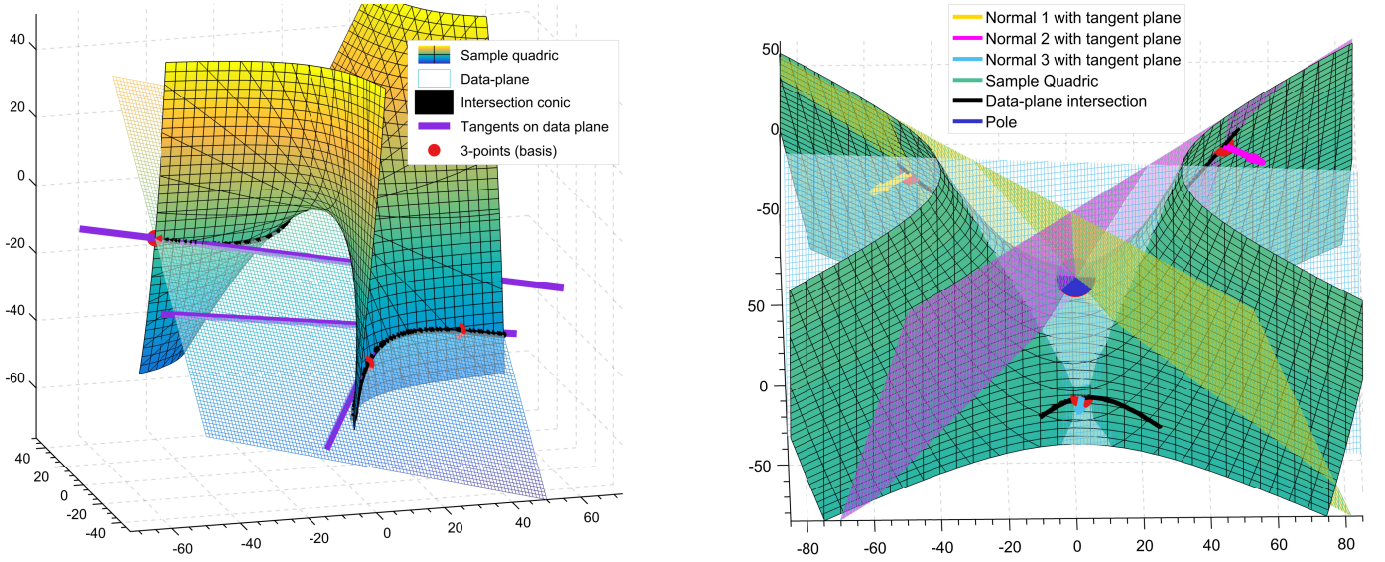


Figure 2. Illustration of the geometric intuition. **(left)** Visualizations on a sample quadric: the selected basis (three-oriented points); the data-plane; the conic of intersection between data-plane and the quadric; the lines on the data-plane that are tangent to the quadric. **(right)** Exemplary drawing that shows that the tangent planes to the basis points meet at the pole (see text).

Algorithm 1 Quadric fitting, full version.

Require: Unit normalized point set $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, Corresponding surface normals $\{\mathbf{n}_x, \mathbf{n}_y, \mathbf{n}_z\}$, A weight coefficient ω

Ensure: Quadric $\mathbf{q} = [A, B, C, 2D, 2E, 2F, 2G, 2H, 2I, J]^T$, Scale factors α

```

 $n = \text{numel}(\mathbf{x})$ 
 $\mathbf{1} = \text{ones}(n, 1)$ ;
 $\mathbf{0} = \text{zeros}(n, 1)$ ;
 $\mathbf{0}_{n \times n} = \text{zeros}(n, n)$ ;
 $\mathbf{X} = [\mathbf{x}^2, \mathbf{y}^2, \mathbf{z}^2, \mathbf{x} * \mathbf{y}, \mathbf{x} * \mathbf{z}, \mathbf{y} * \mathbf{z}, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{1}, \mathbf{0}_{n \times n}]$ ;
 $\mathbf{N} = [\text{diag}(\mathbf{n}_x); \text{diag}(\mathbf{n}_y); \text{diag}(\mathbf{n}_z)]$ ;
 $\mathbf{dX} = [2\mathbf{x}, \mathbf{0}, \mathbf{0}, \mathbf{y}, \mathbf{z}, \mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{0}; \dots$ 
 $\quad \mathbf{0}, 2\mathbf{y}, \mathbf{0}, \mathbf{x}, \mathbf{0}, \mathbf{z}, \mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{0}; \dots$ 
 $\quad \mathbf{0}, \mathbf{0}, 2\mathbf{z}, \mathbf{0}, \mathbf{x}, \mathbf{y}, \mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0}]$ ;
 $\mathbf{A} = [\mathbf{X}; \omega \cdot [\mathbf{dX}, -\mathbf{N}]]$ ;
 $[\sim, \sim, \mathbf{V}] = \text{svd}(\mathbf{A})$ ;
 $\mathbf{q} = \mathbf{V}(1 : 10, n + 10)$ ;
 $\alpha = \mathbf{V}(11 : (n + 10), n + 10)$ ;

```

Alg. 1 provides a MATLAB implementation of such fit. For a non-degenerate quadric, the following rank (rk) relations hold:

$$N = 1 \Rightarrow \text{rk}(\mathbf{A}) = 4, N = 2 \Rightarrow \text{rk}(\mathbf{A}) = 7 \quad (7)$$

$$N = 3 \Rightarrow \text{rk}(\mathbf{A}) = 9, \text{ and } N > 3 \Rightarrow \text{rk}(\mathbf{A}) = 10. \quad (8)$$

We will now further investigate on this interesting behavior.

4.2 Existence of a trivial solution for three-point case

The problem of estimating a quadric from three points and associated normals seems initially to be well-posed: when counting constraints and degrees of freedom, one obtains nine on each side (each point gives one constraint, each normal two, whereas a quadric has nine degrees of freedom). Yet, it turns out that our linear equation system always has a trivial solution besides the true one. This is summarized in Eq. 8 by providing the ranks for different cardinalities of bases. We now give further intuition and proof for this behavior:

Theorem 1. *Three-oriented point quadric fitting, as formulated, possesses a trivial solution besides the true solution, namely the plane spanned by the three data points. The fitting problem thus has at least a one-dimensional linear family of solutions, spanned by the true quadric and this trivial solution.*

Proof. In the following, let us call **data-plane**, the plane spanned by the three data points (coordinates only, i.e. not considering the associated normals). We illustrate this in Fig. 2 (left). As mentioned in § 4.1 above, any rank-1 quadric consists of a single plane Π and can be written as $\mathbf{Q} = \Pi \Pi^T$. Hence, for any point $\mathbf{U}$ on the plane and thus on the quadric, we have $\mathbf{Q}\mathbf{U} = \mathbf{0}$. In our formulation of the fitting problem, this amounts to $\mathbf{N}\mathbf{q} = \mathbf{0}$. $\mathbf{N}$ refers to all the gradient-normal correspondence equations, stacked together (lower part of Eq. 6). We also have $\mathbf{M}\mathbf{q} = \mathbf{0}$. due to the point lying on the quadric. This means that the following vector is a solution of the equation system: coefficients A to J are those of the rank-1 quadric and the three scalars α_i are zero. In other words, the trivial solution is identified as the rank-1 quadric consisting of the data-plane. Hence, the estimation problem admits at least a one-dimensional linear family of solutions, spanned by the true quadric and the rank-1 quadric of the data-plane. In some cases, the dimension of the family of solutions may be higher (such as when the true quadric is a plane). $\square$

4.2.1 A geometric explanation of the fact that the three-oriented-point problem is always under-constrained

Despite the analytical proof, it is puzzling that nine constraints on nine unknowns are never sufficient in our problem. Moreover, we may wonder if the existence of a trivial solution is due to our linear problem formulation or if it is generic. It turns out that this is generic and can be explained geometrically. To make it easier to imagine, our description will closely follow the Figures 2(a) and 2(b). Let us decompose the estimation of the quadric in two parts, the first part being the determination of the quadric's intersection with the data-plane. The intersection of any quadric with any plane is in general always a conic (shown as black

curve), be it real or imaginary (the only exception is when the quadric itself contains the data-plane entirely, in which case the “intersection” is the entire plane).

Let us examine which constraints we have at our disposal to estimate the intersection conic. First, the three data points lie on the conic. Second, we know the tangent planes at the data points, to the true quadric. Let us intersect the three tangent planes with the data plane – the resulting three lines (shown in purple) must be tangent to our conic (the only exception occurs when one or more of these tangent planes are identical to the data plane).

Hence, we know three points on the conic and three tangent lines – the problem of estimating the conic is thus in general overdetermined by one DoF. In other words, six of the nine constraints at our disposal for estimating the quadric are dedicated to estimating the five degrees of freedom of its intersection with the data plane. Hence, the remaining three constraints are not sufficient to complete the estimation of the quadric.

What are these three remaining constraints? They refer to the orientation of the tangent planes: each of the tangent planes is defined by an angle expressing the rotation about its intersection line with the data plane. This angle gives one piece of information on the quadric; for three oriented points we thus have our three remaining constraints.

Note that the three tangent planes to the quadric intersect in the quadric’s pole to the data plane (see Fig. 2(b)). Hence, we can determine this pole which, as shown in appendix, lies on the line joining the centers of the possible solutions for the quadric.

Let us also note that the fact that six pieces of information (three data points and three tangent lines to a conic in the data plane) only constrain five degrees of freedom means that these six pieces of information are not independent from one another: in the absence of noise or other errors, they must satisfy a consistency constraint (the fact that they define a conic). In the presence of noise, the input information will not satisfy this constraint, meaning that a perfect fit will not exist. This is different in most so-called minimal estimation problems in geometric computer vision (such as three-point pose estimation - P3P), where the computed solution is perfectly consistent with the input data. In our case, we can expect that the computed quadric will not satisfy all constraints exactly, i.e. will not necessarily be incident with all data points or be exactly tangent to the given tangent planes.

This gives room to different formulations for the problem, depending on how one quantifies the quality of fit. For instance, one possibility would be to impose that the quadric goes exactly through the data points, but that the tangency is only approximately fulfilled by computing the intersection conic in the data plane and minimizing some cost functions over the tangent lines.

4.3 Regularizing with gradient norm

Quadric fitting problem, like many others (e.g. calibration, projective reconstruction) is intrinsically of non-linear nature, meaning that a “true” Maximum Likelihood Estimation or Maximum A Posteriori solution, minimizing a geometric distance, cannot be achieved by a linear fit. However, our main objective in this stage is a sufficiently close and computationally efficient fit, using as few points as possible and upon which we can build our voting scheme. Despite its sparsity, for such purpose, formulation in § 4.1 still remains suboptimal since the unknowns in Eq. 6 scale linearly with N , leaving a large system to solve. In practice, analogous to gradient-one fitting [53], we could prefer unit-norm polynomial

Algorithm 2 MATLAB 10-liner for approx. quadric fitting.

Require: Unit normalized point set $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$, Corresponding surface normals $\{\mathbf{n}_x, \mathbf{n}_y, \mathbf{n}_z\}$, A weight coefficient ω

Ensure: Quadric $\mathbf{q} = [A, B, C, 2D, 2E, 2F, 2G, 2H, 2I, J]^T$

```

1 = ones(numel(x), 1);
0 = zeros(numel(x), 1);
X = [x^2, y^2, z^2, x * y, x * z, y * z, x, y, z, 1];
dX = [2x, 0, 0, y, z, 0, 1, 0, 0, 0; ...
      0, 2y, 0, x, 0, z, 0, 1, 0, 0; ...
      0, 0, 2z, 0, x, y, 0, 0, 1, 0];
A = [X; omega * dX];
N = [n_x; n_y; n_z];
b = [0; omega * N];
q = A\b;

```

gradients, and thus, can write $\alpha_i = 1$ or equivalently $\alpha_i \leftarrow \bar{\alpha}$, one common factor. This **soft constraint** will try to force zero set of the polynomial respect the local continuity of the data. Similar direction is also taken by [54], for the case of spheres. However, there, authors follow a two-step fitting process, solving first the gradient and then the positional constraint, whereas we formulate a single system solving for all the shape parameters simultaneously. Such regularization also saves us from solving the sensitive homogeneous system [20], and lets us re-write the system in a more compact form $\mathbf{A}\mathbf{q} = \mathbf{n}$:

$$\mathbf{n} = [0 \quad 0 \quad \dots \quad n_x^1 \quad n_y^1 \quad n_z^1 \quad n_x^2 \quad n_y^2 \quad n_z^2 \quad \dots]^T$$

$$\mathbf{q} = [A \quad B \quad C \quad D \quad E \quad F \quad G \quad H \quad I \quad J]^T$$

$$\mathbf{A} = \begin{bmatrix} x_1^2 & y_1^2 & z_1^2 & 2x_1y_1 & 2x_1z_1 & 2y_1z_1 & 2x_1 & 2y_1 & 2z_1 & 1 \\ x_2^2 & y_2^2 & z_2^2 & 2x_2y_2 & 2x_2z_2 & 2y_2z_2 & 2x_2 & 2y_2 & 2z_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 2x_1 & 0 & 0 & 2y_1 & 2z_1 & 0 & 2 & 0 & 0 & 0 \\ 0 & 2y_1 & 0 & 2x_1 & 0 & 2z_1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 2z_1 & 0 & 2x_1 & 2y_1 & 0 & 0 & 2 & 0 \\ 2x_2 & 0 & 0 & 2y_2 & 2z_2 & 0 & 2 & 0 & 0 & 0 \\ 0 & 2y_2 & 0 & 2x_2 & 0 & 2z_2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 2z_2 & 0 & 2x_2 & 2y_2 & 0 & 0 & 2 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \quad (9)$$

$\mathbf{A}$, only $4N \times 10$, is similar to the $\mathbf{A}'$ in § 4.1 and gets full rank for four or more oriented points. In fact, it is not hard to show that the equations in rows are linearly dependent, which is why we get diminishing returns when we add further constraints. Note that by removing the scale factors from the solution, we also solve the sign ambiguity problem, i.e. the solution to Eq. 6 can result in negated gradient vectors. To balance the contribution of normal induced constraints we introduce a scalar weight w , leading to the ten-liner MATLAB implementation as provided in Alg. 2.

In certain cases, to obtain a type-specific fit, a minor redesign of $\mathbf{A}$ tailored to the desired primitive suffices (see §. 6.3.4). If outliers corrupt the point set, a four-point RANSAC could be used. However, below, we present a more efficient way to calculate a solution to Eq. 9 rather than using a naive RANSAC on four-tuples by analyzing its solution space. The next section can also be used as a generic method to solve any fitting problem formulated as a linear system, more efficiently.

5 QUADRIC DETECTION IN POINT CLOUDS

We now factor in clutter and occlusions into our formulation and explain a new pipeline to detect quadrics in 3D data.

Definition 6. A basis $\mathbf{b}$ is a subset composed of a fixed number of scene points (b) and hypothesized to lie on the sought surface.

Our algorithm operates by iteratively selecting bases from an input scene. Once a basis is fixed, an under-determined quadric fit parameterizes the solution and attached to this basis, a local accumulator space is formed. All other points in the scene are then paired with this basis to vote for the potential primitive. To discover the optimal basis, we perform RANSAC, iteratively hypothesizing different basis candidates and voting locally for probable shapes. Subsequent to such joint RANSAC and voting, we verify resulting hypotheses with efficient two-stage clustering and score functions such that multiple quadrics can be detected without repeated executions of the algorithm. We will now describe, in detail, the voting and the bases selection, respectively.

5.1 Parameterizing the solution space

Linear system in Eq. 6 describes an outlier-free closed form fit. To treat the clutter in the scene, a direct RANSAC on nine-DoF quadric appears to be trivial. Yet, it has two drawbacks: 1) evaluating the error function many times is challenging, as it involves a scene-to-quadric overlap calculation in a geometric meaningful way. 2) even with the proposed fitting, selecting random four-tuples from the scene might be slow in practice.

An alternative to RANSAC is Hough voting. However, $\mathbf{q}$ has nine DoFs and is not discretization friendly. The complexity and size of this parameter space makes it hard to construct a voting space. Instead, we will now devise a local search. For this, let $\mathbf{q}$ be a solution to the linear system in (9) and $\mathbf{p}$ be a particular solution. $\mathbf{q}$ can be expressed by a linear combination of homogeneous solutions $\boldsymbol{\mu}_i$ as:

$$\begin{aligned} \mathbf{q} &= \mathbf{p} + \sum_i^{d_n} \lambda_i \boldsymbol{\mu}_i \\ &= \mathbf{p} + [\boldsymbol{\mu}_1 \quad \boldsymbol{\mu}_2 \quad \dots] [\lambda_1 \quad \lambda_2 \quad \dots]^T = \mathbf{p} + \mathbf{N}_A \boldsymbol{\lambda}. \end{aligned} \quad (10)$$

The dimensionality d_n of the null space $\mathbf{N}_A$ depends on the rank of $\mathbf{A}$, which is directly influenced by the number of points used: $d_n = 10 - rk(\mathbf{A})$. The exact solution could always be computed by including more points from the scene and validating them, i.e. by a local search. For that reason, the fitting can be split into distinct parts: first a parametric solution is computed, such as in Eq. 10, using a subset of points $\mathbf{b} = \{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ which lie on a quadric. We refer to subset $\mathbf{b}$ as the *basis*. Next, the coefficients $\boldsymbol{\lambda}$, and thus the solution, can be obtained by searching for other point(s) $(\mathbf{x}_{m+1}, \dots, \mathbf{x}_{m+k})$ which lie on the same surface as $\mathbf{b}$.

Proposition 1. If two point sets $\mathbf{b} = (\mathbf{x}_1, \dots, \mathbf{x}_m)$ and $\mathbf{X} = (\mathbf{x}_{m+1}, \dots, \mathbf{x}_{m+k})$ lie on the same quadric with parameters $\mathbf{q}$, then the coefficients $\boldsymbol{\lambda} = [\lambda_1 \quad \lambda_2 \quad \dots]^T$ of the solution space (10) are given by the solution of the system:

$$(\mathbf{A}_k \mathbf{N}_A) \boldsymbol{\lambda} = \mathbf{n}_k - \mathbf{A}_k \mathbf{p} \quad (11)$$

where $\mathbf{A}_k, \mathbf{n}_k$ are the linear constraints of the latter set $\mathbf{b}'$ in form of (9), $\mathbf{p}$ is a particular solution and $\mathbf{N}_A$ is a stacked null-space basis as in (10), obtained from $\mathbf{b}$.

Proof. Let $\mathbf{q}$ be a quadric solution for the point set $(\mathbf{x}_1, \dots, \mathbf{x}_m)$ and let $(\mathbf{A}_k, \mathbf{n}_k)$ represent the $4k$ quadric constraints for the k points $\mathbf{X} = (\mathbf{x}_{m+1}, \dots, \mathbf{x}_{m+k})$ in form of (6) with the same parameters $\mathbf{q}$. As $\mathbf{x}_i \in \mathbf{X}$ by definition lies on the same quadric $\mathbf{q}$, it also satisfies $\mathbf{A}_k \mathbf{q} = \mathbf{n}_k$. Inserting Eq. 10 into this, we get:

$$\mathbf{A}_k (\mathbf{p} + \mathbf{N}_A \boldsymbol{\lambda}) = \mathbf{n}_k \quad (12)$$

$$(\mathbf{A}_k \mathbf{N}_A) \boldsymbol{\lambda} = \mathbf{n}_k - \mathbf{A}_k \mathbf{p} \quad (13)$$

Solving Eq. 13 for $\boldsymbol{\lambda}$ requires a multiplication of a $4k \times 10$ matrix with a $10 \times m$ one and ultimately solving a system of $4k$ equations in m unknowns. Once $\mathbf{N}_A$ and $\mathbf{p}$ are precomputed, it is much more efficient to evaluate Eq. 11 for $k < m$ rather than re-solving the system (9). This resembles updating the solution online for a stream of points. For our case, the amount of streamed points will depend on the size of the basis, as explained below.

5.2 Local voting for quadric detection

Given a fixed basis composed of b points ($b > 0$) as in Fig. 3, a parametric solution can be described. The actual solution can then be found quickly by using Prop. 1 by incorporating new points lying on the same quadric as the basis. Thus, the problem of quadric detection is de-coupled into 1) finding a proper basis and 2) searching for compatible scene points. In this section, we assume the basis is correctly found and explain the search by voting. For a fixed basis $\mathbf{b}_i$ on a quadric, we form the null-space decomposition of the under-determined system $\mathbf{A}_i \mathbf{q} = \mathbf{n}_i$. We then sample further points from the scene and compute the required coefficients $\boldsymbol{\lambda}$. Thanks to Prop. 1, this can be done efficiently. Sample points lying on the same quadric as the basis (inliers) generate the same $\boldsymbol{\lambda}$ whereas outliers will produce different values. Therefore we propose to construct a voting space on $\boldsymbol{\lambda}$ attached to basis $\mathbf{b}_i$ and cast votes to maximize the consensus, only up to the locality of the basis. Fig. 3 illustrates this configuration. The size of the voting space is a design choice and depends on the size of the basis $\mathbf{b}_i$ vs. the DoFs desired to be recovered (see Fig. 1(b)).

While many choices for the basis cardinality are possible (and the formulation in § 5.1 allows for all), we find from Fig. 1(a) that using a three-point basis is advantageous for a generic quadric fit - having three dual points, reduces the minimum number of required primal (incidence) constraints to only four. And by the rank analysis given in Eq. 8 we see that it is possible to trade one point off to 1D local search as opposed to two-point vs 3D search for the five-point case.

5.3 Efficient computation of voting parameters for a 1D voting space

Adding a fourth sample point $\mathbf{x}_4$ completes $rk(\mathbf{A}) = 10$ and a unique solution can be computed, as described above. Yet, as we will select multiple $\mathbf{x}_4$ candidates per basis, hypothesized in a RANSAC loop, an efficient scheme is required, i.e. once again, it is undesirable to re-solve the system in Eq. 9 for each incoming $\mathbf{x}_4$ tied to the basis. It turns out that the solution can be obtained directly from Eq. 10:

Proposition 2. If the null-space is one dimensional (with only 1 unknown) it holds $\boldsymbol{\lambda} \mathbf{N}_A = \lambda_1 \boldsymbol{\mu}_1$ and the computation in Prop. 1 reduces to the explicit form:

$$\lambda_1 = \frac{\mathbf{A}_1 \mathbf{N}_A}{\|\mathbf{A}_1 \mathbf{N}_A\|^2} \cdot (\mathbf{n}_1 - \mathbf{A}_1 \mathbf{p}) \quad (14)$$

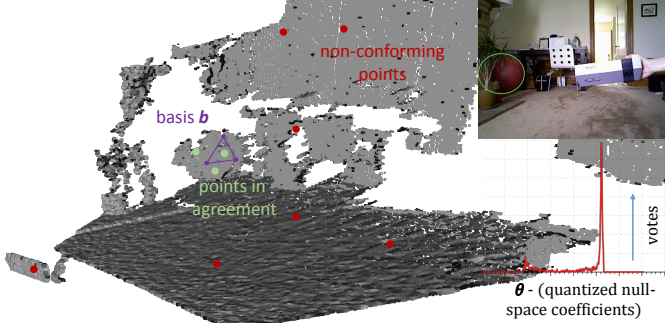


Figure 3. Once a basis is randomly hypothesized, we look for the points on the same surface by casting votes on the null-space. The sought pilates ball (likely quadric) is marked on the image and below that lies the corresponding filled accumulator by KDE [55].

Proof. Let us re-write Eq. [13] in terms of the null space vectors: $\lambda_1(\mathbf{A}_1\boldsymbol{\mu}_1) = \mathbf{n}_1 - \mathbf{A}_1\mathbf{p}$. A solution λ_1 can be obtained via Moore-Penrose pseudoinverse as $\lambda_1 = (\mathbf{A}_1\boldsymbol{\mu}_1)^+(\mathbf{n}_1 - \mathbf{A}_1\mathbf{p})$. Because for one-dimensional null spaces, $\mathbf{A}_1\boldsymbol{\mu}_1$ is a vector ($\mathbf{v}$), for which the $^+$ operator is defined as: $\mathbf{v}^+ = \mathbf{v}/(\mathbf{v}^T\mathbf{v})$. Substituting this in Eq. [11] gives Eq. [14] $\square$

Prop. [2] enables a very quick computation of the parameter hypothesis in the case of an additional single oriented point. A MATLAB implementation takes ca. $30\mu s$ per λ . Note that for the minimal system we propose, four incidence (primal) and three tangent plane alignment (dual) constraints are sufficient. This means that the normal of the fourth sample point does not contribute to the set of constraints for a *minimal* fit. Hence, we use this piece of information for the verification of the fit. We only accept to vote a candidate quadric if the gradient of the fitted surface agrees with the surface normal of the fourth point:

$$\frac{\nabla \mathbf{Q}(\mathbf{x}_4)}{\|\nabla \mathbf{Q}(\mathbf{x}_4)\|} \cdot \mathbf{n}(\mathbf{x}_4) > \tau_n. \quad (15)$$

We typically set $\tau_n \approx 0.85$ in order to tolerate certain noise.

5.4 Quantizing λ for voting

Unfortunately, λ is not quantization-friendly, as it is unbounded and has a non-linear effect on the quadric shape (Fig. [5.2]). Thus, we seek to find a geometrically meaningful transformation to a bounded and well behaving space so that quantization would lead to little bias and artifacts. From a geometric perspective, each column of $\mathbf{N}_A$ in Eq. [10] is multiplied by the same coefficient λ , corresponding to the slope of a high dimensional line in the solution space. Thus, it could as well be viewed as a rotation. For 1D null-space, we set:

$$\theta = \text{atan2}\left(\frac{y_2 - y_1}{x_2 - x_1}\right) \quad (16)$$

where $[x_1, y_1, \dots]^T = \mathbf{p}$ and $[x_2, y_2, \dots]^T$ is obtained by moving in the direction $\mathbf{N}_A$ from the particular solution $\mathbf{p}$ by an offset λ . This new angle θ is bounded and thus easy to vote for. As the null-space dimension grows, λ starts to represent hyperplanes, still preserving the geometric meaning, i.e. for $d > 1$, different $\theta = \{\theta_i\}$ can be found.

Even though θ behaves better than λ for voting, we still can not guarantee a unimodal distribution such that a single peak can be

1. Simple $\tan^{-1}(\lambda)$ could work but would be more limited in the range.

Algorithm 3 Combined RANSAC & Local Voting.

Require: Unit normalized point set $\mathbf{P}$, Corresponding surface normals $\mathbf{N}$, A weight coefficient ω , Minimum vote threshold s_{min}
Ensure: Quadrics $\mathbf{Q} = \{\mathbf{q}_i\}$
 $(\mathbf{S}, \mathbf{N}) \leftarrow \text{Sample scene } (\mathbf{P}, \mathbf{N})$.
while !satisfied **do** $\triangleright$ seek the best global candidates
 $\mathbf{b}_i \leftarrow \text{Pick a random 3 point-basis from } (\mathbf{S}, \mathbf{N})$.
 $(\mathbf{A}, \mathbf{n}) \leftarrow \text{Form system in eq. (4) using } \mathbf{b}_i$.
 $(\mathbf{p}, \boldsymbol{\mu}) \leftarrow \text{Perform null space decomposition - eq.7.}$
 $\mathbf{V} \leftarrow \text{Initial voting space of length \# bins}$
 $\Lambda \leftarrow \{\}$
for all $\mathbf{p}_i$ in $\mathbf{P}$ **do** $\triangleright$ local voting
 Compute λ_i by including $\mathbf{p}_i$ using eq. 11.
 if $(1 - \frac{\nabla \mathbf{Q}(\mathbf{p}_i)}{\|\nabla \mathbf{Q}(\mathbf{p}_i)\|} \cdot \mathbf{n}_i < \tau)$ **then** $\triangleright$ goodness of fit
 Quantize: $\theta \leftarrow \tan^{-1}(\lambda_i)$ (using atan2).
 $\mathbf{V}[\theta] \leftarrow \mathbf{V}[\theta] + 1$ $\triangleright$ accumulate
 $\Lambda[\theta] \leftarrow \Lambda[\theta] \cup \lambda_i$
 end if
end for
 $\theta^* \leftarrow \text{argmax}_j \mathbf{V}_j$ $\triangleright$ best candidate in quantized space
if $|\Lambda[\theta^*]| > s_{min}$ **then**
 $\lambda_{best} \leftarrow \sum_k \Lambda[\theta^*][k] / |\Lambda[\theta^*]|$ $\triangleright$ best local coefficient
 $\mathbf{q} \leftarrow \mathbf{p} + \lambda_{best}\boldsymbol{\mu}$ $\triangleright$ best local solution
 $\mathbf{Q} \leftarrow \{\mathbf{Q}, \mathbf{q}\}$
end if
end while
 $\mathbf{Q} \leftarrow \text{mean of the clusters in } \mathbf{Q} \text{ using distance } d_{close}$
 $\mathbf{Q} \leftarrow \text{mean of the clusters in } \mathbf{Q} \text{ using distance } d_{far}$
 $\mathbf{Q} \leftarrow \text{sort}(\text{score}(\mathbf{Q}))$

identified unambiguously. Nevertheless, thanks to the local voting, the case that one distribution is noisy or *misty* will be handled when other random bases are selected. It is more likely that the peaks coming from different bases are concentrated around the same mode, rather than a single peak of one accumulator. Besides, we have empirically observed that in many real cases, even when the distribution is amodal, a single peak is prominent when the sampled fourth is in a reasonable vicinity of the basis.

5.5 Hypotheses aggregation

Up until now, we have described how to find plausible quadrics given local triplet bases. As mentioned, to discover the basis lying on the surface, we employ RANSAC [56], where each triplet might generate a hypothesis to be verified. Many of those will be similar as well as dissimilar. Thus, the final stage of the algorithm aggregates the potential detections to reduce the number of candidate surfaces and to increase the per quadric confidence. Not to sacrifice further speed, we run an agglomerative clustering similar to [1] in a coarse to fine manner: First a fine (*close*) but fast distance measure helps to cluster the obvious hypotheses. Second, a coarse (*far*) one is executed on these cluster centers.

Definition 7. Our distance computation is two-fold: Whenever two quadrics are close, we approximate their distances as in Eq. [17] (d_{close}), where $\mathbf{I} \in \mathbb{R}^{4 \times 4}$ is the identity matrix and $\mathbb{1} : \mathbb{R} \rightarrow \{0, 1\}$ the indicator function. We use the pseudoinverse just to handle singular configurations. If the shapes are far, such manifold-distance becomes erroneous and we use a globally consistent metric. To do so, we define a more geometric-meaningful distance using the points on the scene itself (d_{far}):

$$\begin{aligned} d_{close}(\mathbf{Q}_1, \mathbf{Q}_2) &:= \mathbb{1}(\|\mathbf{q}_1 - \mathbf{q}_2\|_1 < \tau) \cdot \|\mathbf{Q}_1 \mathbf{Q}_2^+ - \mathbf{I}\|_F \\ d_{far}(\mathbf{Q}_1, \mathbf{Q}_2) &:= \end{aligned} \quad (17)$$

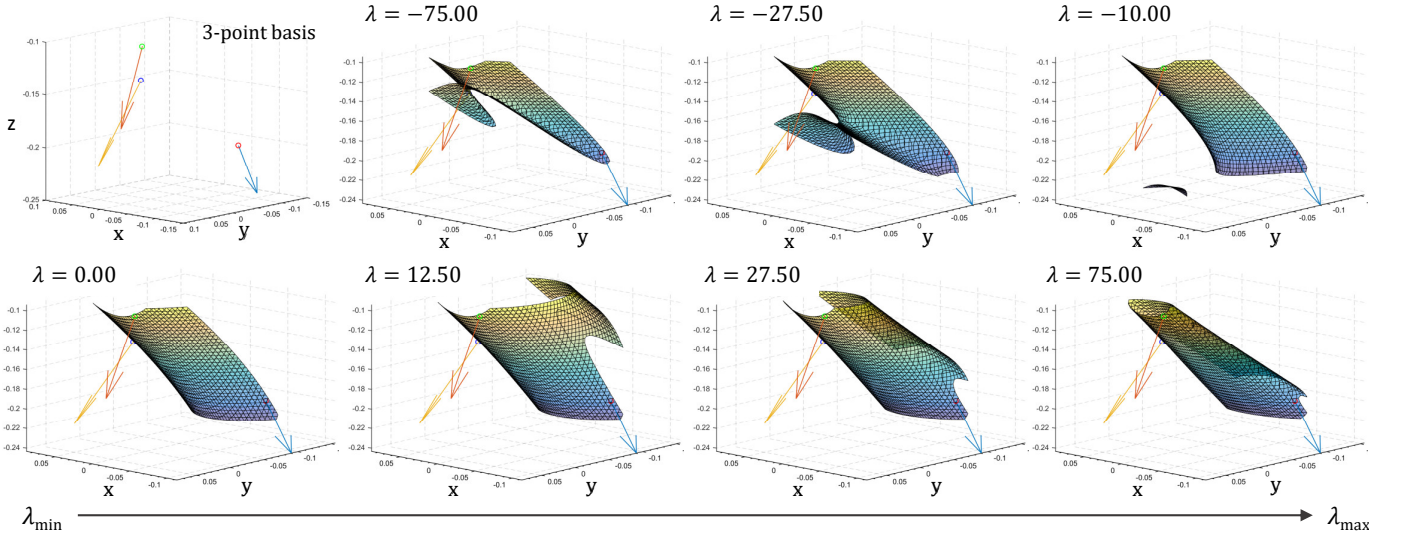


Figure 4. Effect of λ on the surface geometry. We compute null-space decomposition for a fixed basis and vary λ from -75 to 75 to generate different solutions $\mathbf{q}$ along the line in the solution space. The plot presents the transition of the surface controlled by λ .

$$1 - \frac{1}{K} \sum_{i=1}^K \mathbb{1}(|\mathbf{x}_i^T \mathbf{Q}_1 \mathbf{x}_i| < \tau) \cdot \mathbb{1}(|\mathbf{x}_i^T \mathbf{Q}_2 \mathbf{x}_i| < \tau) \cdot \mathbb{1}(1 - \mathbf{n}_i \cdot \nabla \mathbf{Q}_1(\mathbf{x}_i) < \tau_n) \cdot \mathbb{1}(1 - \mathbf{n}_i \cdot \nabla \mathbf{Q}_2(\mathbf{x}_i) < \tau_n).$$

$\{\mathbf{x}_i\}$ denote the K scene samples.

Note that, algebraic but efficient d_{close} lacks geometric meaning, while slower d_{far} can, to a certain extent, explain the geometry. Finally, the quadrics are sorted w.r.t. their scores, evaluated pseudo-geometrically by point and normal-gradient compatibility according to Def. 8.

Definition 8. The score of a quadric is defined to be:

$$S_{\mathbf{Q}, \mathbf{x}} = \frac{1}{K} \sum_{i=1}^K \mathbb{1}(|\mathbf{x}_i^T \mathbf{Q} \mathbf{x}_i| < \tau) \mathbb{1}(1 - \mathbf{n}_i \cdot \nabla \mathbf{Q}(\mathbf{x}_i) < \tau_n)$$

While other distance metrics, such as spectral decompositions are possible, we found these to be sufficient in our experiments. The final algorithm is summarized in Alg. 3.

6 EXPERIMENTAL EVALUATION AND DISCUSSIONS

6.1 Implementation details

Prior to operation, we normalize the point coordinates to lie in a unit ball to increase the numerical stability [57]. Next, we down-sample the scene using a spatial voxel-grid enforcing a minimum distance of $\tau_s \cdot \text{diam}(\mathbf{X})$ between the samples ($\tau_s = 0.03$) [58]. The required surface normals are computed by the local plane fitting [59]. As planes are singular quadrics and occupy large spaces of 3D scenes, we remove them. To do so, we convert our algorithm to a type specific plane detector, which happens to be a similar algorithm to [30]. Next, influenced by the smoothness of quadrics, we use Difference of Normals (DoN) [60] to prune the points not located on smooth regions. What follows is an iterative selection of triplets to conduct the three-point RANSAC: We first randomly draw the initial point of the basis $\mathbf{x}_1$. Once $\mathbf{x}_1$ is fixed, we query the points in a large enough vicinity, whose normals differ enough to form the three-point basis $\mathbf{b}$. The rest of the points are then randomly selected respecting these criteria. To avoid

degenerate configurations, we skip the basis if it does not result in a rank-9 matrix $\mathbf{A}$. In addition, to reducing the bias towards repeating bases, we hash the seen triplets and avoid duplicates.

6.2 Synthetic tests of fitting and ablation studies

To assess the accuracy of the proposed fitting, we generate a synthetic test set of multitudes of random quadrics and compare our method with the fitting procedures of Taubin [4], Tasdizen [6], Andrews [5], and Beale [33]. We propose two variants: **Ours full** will refer to Alg. 1, whereas **Ours** is the regularized one (Alg. 2).

6.2.1 Quantitative assessments

Prior to run, we add Gaussian noise to the ground-truth vertices with $\sigma = [0\% - 5\%]$ relative to the size s of the quadric. At each noise level, ten random quadrics are tested. We perform not single but twenty fits per set. For the constrained fitting method [5] we pre-specified the type, which might not be possible in a real application. We then record and report the average point-to-mesh distance and the angle deviation as well as the runtime performances in Fig. 5. Although, our fit is designed to use a minimal number of points, it also proves robust when more points are added and is among the top fitters for the distance and angle errors. In addition, Fig. 5c shows that the errors on the gradient magnitudes obtained by our quadrics. We achieve the least errors, showing that gradient norms align well with the ground truth, favoring the validity of our approximation/regularization. Next, looking at the noise assessments, we see that our full method performs the best on low noise levels but quickly destabilizes. This is because the system might be biased to compute correct norms rather than the solution and it has increased parameters. We believe the reason for our compact fit to work well is the soft constraint where the common scale factor acts as a weighted regularizer towards special quadrics. When this constraint cannot be satisfied, the solution settles for a very acceptable shape.

In a further test, we include the six neighboring points of each of seven query points to perform a standard Taubin-fit. We call this *Taubin-42*. Fig. 6a shows that while the error of our method is

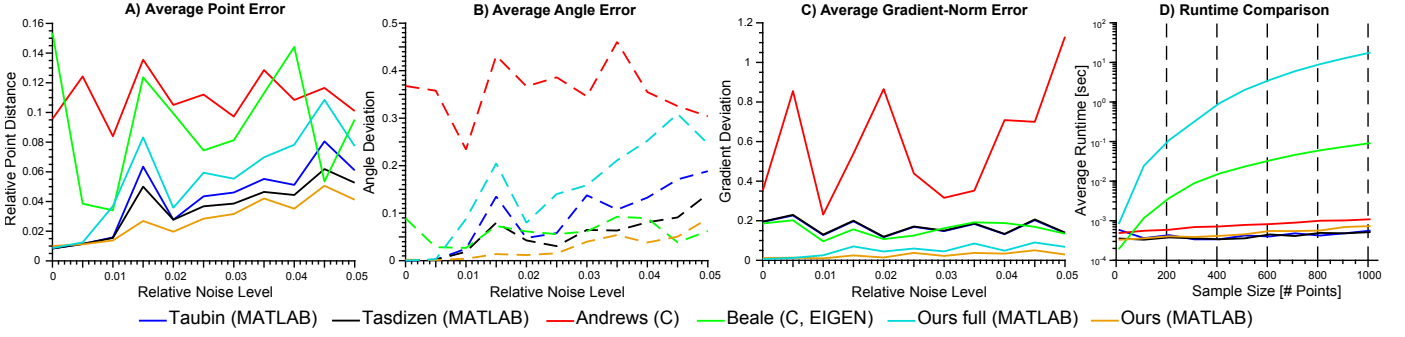


Figure 5. Synthetic evaluations. The plot depicts mean geometric errors on points (a) and mean angular errors (b) for different quadric fitting methods. The per point error is measured as the average point-to-mesh distance from every ground truth vertex to the fitted quadric. The angular error (dashed) is computed as the negated dot product between quadric gradient and the ground truth normal. Moreover, (c) shows the average error of the gradient norm compared to the ground truth and (d) gives speed and detection rate on synthetic data.

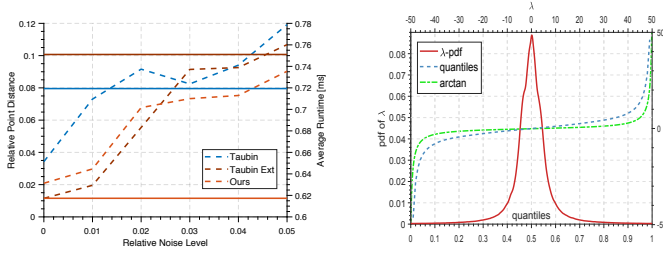


Figure 6. a (left). Effect of extended point neighborhood to the fitting. b (right). Statistical distribution of the solution-space coefficient and our quantization function: PDF (red curve) and inverse CDF (dashed blue-curve) of λ over collected data, and $\tan^{-1}$ function (green-line). Note that our quantization function is capable of explaining the empirical data.

on par with *Taubin-42*, we are more robust at higher noise values and more efficient with a runtime advantage of ca. 22%.

Since for a visually appealing fit, the normal alignment is crucial, we next present a qualitative evaluation.

6.2.2 Qualitative assessments

We synthesized a random saddle quadric and performed a random point sampling over its surface. Next, we added Gaussian noise on the sample points and computed the normals. To resolve the sign ambiguity, each normal is flipped in the direction of ground truth gradient. We plot the results of the fitting in Fig. 7. Even in presence of little noise only some methods fail to estimate the correct geometry, mostly due to the bias towards certain shape [5], [33]. Our approach is able to recover the correct surface even in presence of a severe noise. Also the effect of our regularization is visible on the last column, which possesses the best visual quality.

It is of interest to see whether our regularized fit can estimate correct surface normals as well as direction. Thus, a second test was performed to qualitatively observe the gradient properties in more detail. For this, a series of randomly generated quadrics is fitted by Taubin’s and our method and the gradients are analyzed both in terms of magnitude and phase, as shown in Fig. 9. Due to our explicit treatment of the gradients, it can be clearly seen that the gradient direction is recovered better. Moreover, the right side of Fig. 9 also shows that our approximate approach yields the expected results, while the full method could sometimes generate inconsistent gradient signs, as the scale factors are estimated individually. Finally, it is qualitatively visible in Fig. 9 that the

magnitudes recovered by our method are compatible to the ground truth. Such improvement without sacrificing gradient quality validates the regularizing nature of our approach.

6.2.3 Is atan2 a valid transformation for λ

To assess the practical validity of the quantization, we collect a set of 2.5 million oriented point triplets from several scenes and use them as bases to form the underdetermined system $\mathbf{A}$. We then sample the fourth point from those scenes, compute λ and establish the probability distribution $p(\lambda)$ for the whole collection to calculate the quantiles, mapping λ to bins via the inverse CDF. A similar procedure has been applied to cross ratios in [61]. We plot the findings together with the atan2 function in Fig. 6b and show that the empirical distribution and atan2 follow similar trends, justifying that our quantizer is well suited to the data.

6.2.4 Effect of weighting on the fit

We now investigate the effect of weighting parameter w on the fit. For a selection of eight noisy points, located on three different synthetic quadrics, we vary w and plot, in Fig. 8a, the geometric errors attained by Alg. 2, against the ground truth and Taubin fit. While too low of w hurts our fit, there is a large range of values $w \in [0.08, 1.0]$, where we can outperform [4].

6.2.5 How do voting spaces look like?

To provide insights on the local voting spaces of the angles θ , we sample different random bases on four synthetic quadrics as embedded in Fig. 8b, and collect the votes along with the quantized bins. These accumulators are shown in the same figure, each with a different color. It is observed that, the voting spaces are myst-free and a only single mode emerges, thanks to the maximum distance threshold selected between the basis and the paired point. It is still possible to obtain multiple modes if the threshold is unrealistically picked. The consensus votes correspond to the true shape, and erroneous votes spread randomly.

6.3 Real experiments on quadric detection

Besides synthetic tests where self evaluation was possible, we assess the quality of generic primitive detection, on 3 real datasets:

- 1) **Our Dataset** First, because there are no broadly accepted datasets on quadric detection, we opt to collect our own. To do so, we use an accurate phase-shift stereo structured light scanner and capture 35 3D scenes of 5 different objects within clutter and occlusions. Our objects are three bending

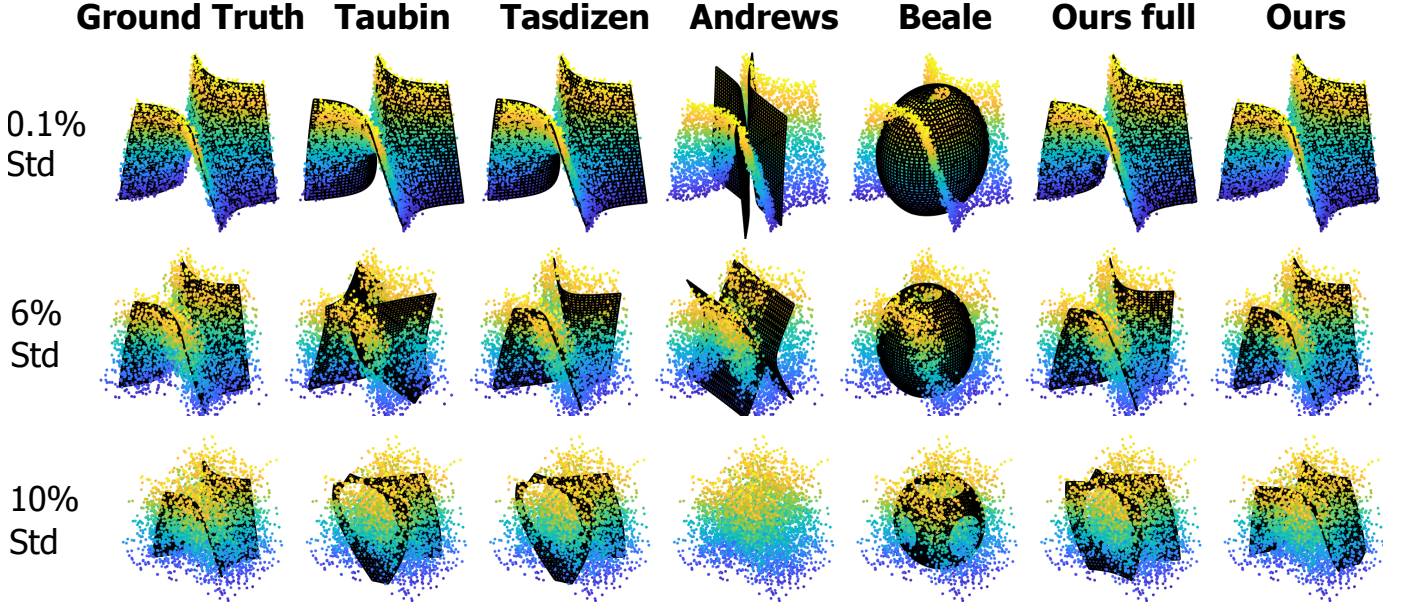


Figure 7. Synthetic tests at various noise levels for different fitting methods. Gaussian noise is added to the point coordinates as well as the estimated normal. The standard deviation varies from 0.11% of the visible quadric size to 10%.

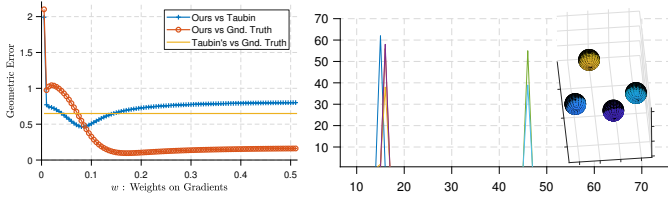


Figure 8. **a** (left). Effect of the weight w on the quality of the fit (Alg. 2). **b** (right). Voting spaces due to different bases selection (θ vs # votes).

papers, helmet, paper towel and cylindrical spray bottle. Other objects are included to create clutter. To obtain the ground truth, for each scene, we generated a visually acceptable set of quadrics using 1) [23] when shapes represent known primitives 2) by segmenting the cloud manually and performing a fit, when the quadric type is not available. Each scene then contains 1-3 ground truth quadrics. This dataset has low noise, but a high amount of clutter and partial visibility due to the FOV limitations of the sensor.

- 2) **Large Objects** Kinect sensor is widely accepted in computer vision community. Thus, it is desirable to see the performance of our generic and type-specific fit approaches on the Kinect depth images. To this end, we adapt the large objects RGB-D scan dataset of [62]. From this dataset, we sample only the scenes containing objects, that could roughly be explained by geometric primitives. These scenes include apples, globes, footballs, or other small balls. Tab. 1 summarizes the objects used. Example detections are also shown in Fig. 12. We also augment this dataset with a Pilates Ball sequence that we collect. This sequence involves a lot of partial visibility, clutter and fast motions (see appendix).
- 3) **Cylinders** Finally, we use a subset of the ITODD dataset [63], designed to evaluate object pose estimators. Our subset, *Cylinders*, includes 14 scenes of varying number of cylinders, from one to ten, as shown in Fig. 13a. Again, we use RGB images only to ease the visual perception.

6.3.1 Evaluations on detection accuracy

To assess the detection accuracy, we manually count the number of detected quadrics aligning with the ground truth in *Our Dataset*. We compared the four-point and three-point algorithms, both of which we propose. We also tried the naive nine-point RANSAC algorithm (with [4]), but found it to be infeasible when the initial hypotheses of the inlier set is not available. Fig. 10 visualizes the detected quadrics both on our dataset and on the 3D data captured by Kinect. Fig. 11(a) presents our accuracy over different sampling rates and the runtime performance. Our three-point method is on par with the four-point variant in terms of detection accuracy, while being significantly faster. Next, we also evaluate our detector on the large objects dataset of [62] without further tuning. Tab. 1 shows 100% accuracy in locating a frontally appearing ellipsoidal rugby ball over a 1337 frame sequence without type prior. While such scenes are not particularly difficult, it is noteworthy that we manage to generate the similar quadric repeatedly at each frame within 5% of the quadric diameter.

6.3.2 How fast is it?

As our speed is influenced by the factors of closed form fitting, RANSAC and local voting, we evaluate the fit and detection separately. Fig. 5d shows the runtime of fitting part. Our method scales linearly due to the solution of an $4N \times 10$ system, but it is the fastest approach when < 300 points are used. Thus, it is more preferred for a minimal fit. Fig. 11(a) then presents the order of magnitude speed gain, when our four-point C++ version is replaced by three-points without accuracy loss. Although the final runtime is in the range of 1-2 seconds, our three-point algorithm is still the fastest known method in segmentation free detection.

6.3.3 How accurate is the fit?

To evaluate the pose accuracy on real objects, we use closed geometric objects of known size from the aforementioned datasets and report the distribution of the errors, and their statistics. We choose *football* and *pilates ball 1* as it is easy to know their geometric

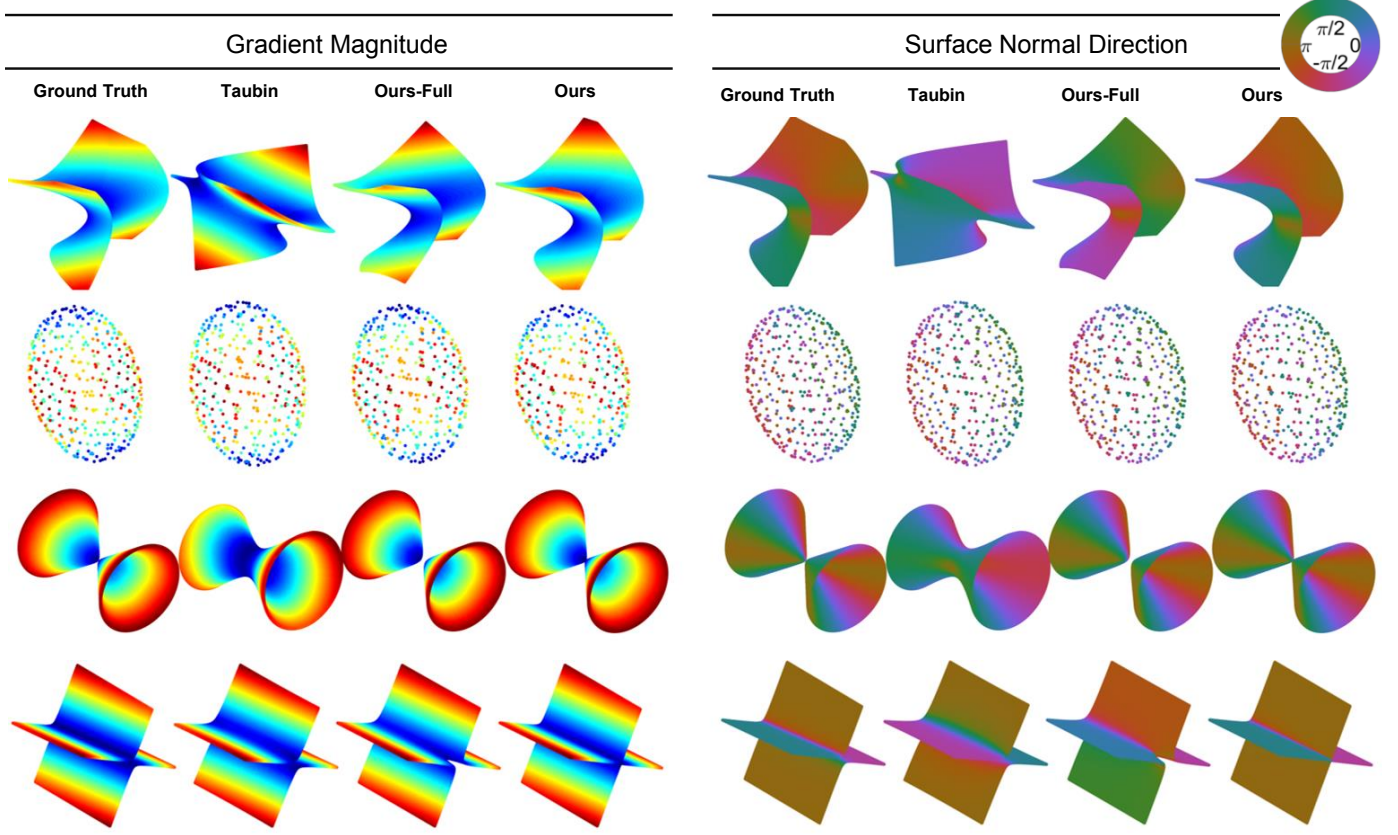


Figure 9. Qualitative evaluation of surface normals. Randomly generated quadrics are used as ground truth and fitting is performed. The estimation results of gradient magnitude and angle (phase) is color coded on the surface. For color selection, we use a jet-like temperature map for the gradient magnitude, where blue denotes the lowest and red denotes the highest magnitudes. For the phase, an angular map as show in the color-bar is used. The ideal case is given in ground-truth against which the methods compete.

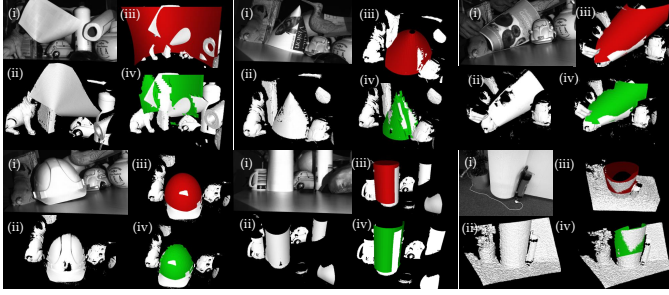


Figure 10. (i) Images captured by an industrial structured light sensor and Kinect (last image). (ii) 3D scene. (iii) Detected quadric, shown without clipping. (iv) Quadric in (iii) clipped to the points it lies on.

properties (center and radius). We compare the radius to the true value while the center is compared to the one estimated from a non-linear refinement of the sphere. Our results are depicted in Fig. 11(b). Note that the errors successfully remain about the used sampling rates ($\tau \approx 0.05$), which is as best as we could get.

6.3.4 Type-specific detection

It is remarkably easy to convert our algorithm to a type specific one by re-designing matrix $\mathbf{A}$. Here, we propose a sphere-specific detector. Let us write any sphere in the following matrix form:

$$\mathbf{Q}_s = \begin{bmatrix} \mathbf{I}_3 & -\mathbf{c} \\ -\mathbf{c}^T & \|\mathbf{c}\|^2 - r^2 \end{bmatrix} \quad (18)$$

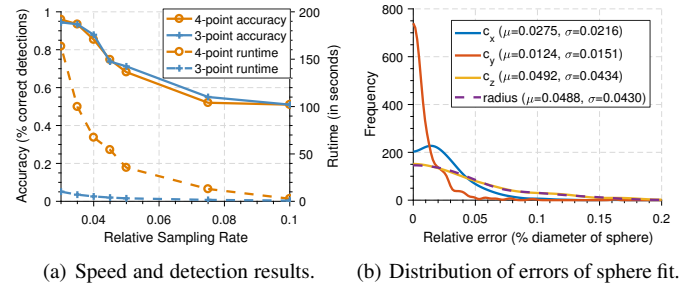


Figure 11. Experiments on real datasets. Our three-point variant clearly outperforms the four-point algorithm in terms of speed, while not sacrificing quality. For three-point method, the mean errors indicate that the localization accuracy is on the levels of the sampling rate $\tau \approx 0.05$, agreeing to the success of our algorithm.

where $\mathbf{c} = (c_x, c_y, c_z)$ and r are the geometric parameters (center and radius) of the sphere. Rotation does not affect spheres and our $\mathbf{A}\mathbf{q} = \mathbf{b}$ formulation in § 4.3 then simplifies to:

$$\mathbf{A} = \begin{bmatrix} \|\mathbf{x}_1\|^2 & 2x_1 & 2y_1 & 2z_1 & 1 \\ \|\mathbf{x}_2\|^2 & 2x_2 & 2y_2 & 2z_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 2\mathbf{x}_1 & \mathbf{I}_3 & \mathbf{0}_3 \\ 2\mathbf{x}_2 & \mathbf{I}_3 & \mathbf{0}_3 \\ \vdots & \vdots & \vdots \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \mathbf{n}_1 \\ \mathbf{n}_2 \\ \vdots \end{bmatrix}, \mathbf{q} = \begin{bmatrix} A \\ B \\ C \\ D \\ E \end{bmatrix} \quad (19)$$

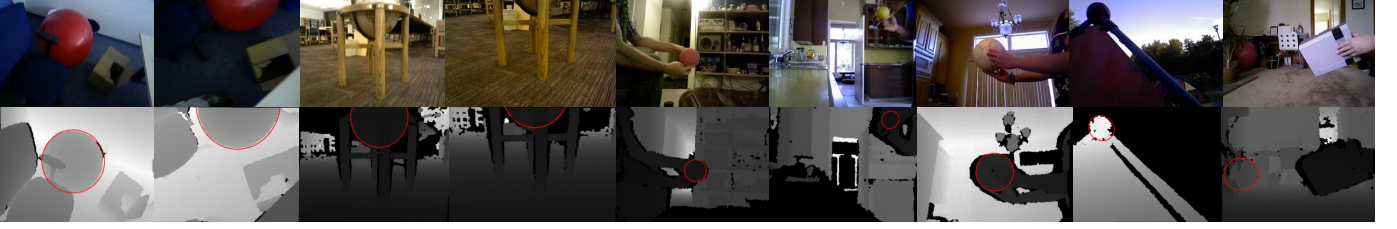


Figure 12. Qualitative visualizations of sphere detection in the wild: Our algorithm successfully detects primitives in difficult scenarios including clutter, occlusions and large distances. Note that the sphere is detected in 3D only using the point clouds of depth images and we draw the apparent contour of the quadric. The RGB pictures are also included in the top row to ease the visual perception.

	Dataset	# Objects	Type	Occlusion	Accuracy
Pilates Ball 1	Ours	580	Generic	Yes	94.40%
Rugby Ball	[62]	1337	Generic	No	100.00%
Pilates Ball 2	[62]	1412	Sphere	Yes	100.00%
Big Globe	[62]	2612	Sphere	Yes	90.70%
Small Globe	[62]	379	Sphere	Yes	56.90%
Apple	[62]	577	Sphere	Yes	99.60%
Football	[62]	1145	Sphere	Yes	100.00%
Orange Ball	[62]	270	Sphere	Yes	93.30%

Table 1
Detection accuracy on real datasets.

Due to the geometric interpretability, at the scoring phases, we can use the point-to-sphere distance as:

$$d_{\text{point} \rightarrow \text{sphere}}(\mathbf{p}, \mathbf{q}) = |(\|\mathbf{p} - \mathbf{c}\|_2 - r)| \quad (20)$$

where $\mathbf{p}$ is the point to compute the distance. A sphere to sphere distance (used in clustering) can be obtained by:

$$d_{\text{sphere} \rightarrow \text{sphere}}(\mathbf{q}_1, \mathbf{q}_2) = \frac{1}{2}(|r_1 - r_2| + \|\mathbf{c}_1 - \mathbf{c}_2\|_2) \quad (21)$$

Center and radius of the sphere can always be obtained from the quadric form $\mathbf{Q}_s$ as described in eq. [18]

Note that $rk(\mathbf{A}) = 4$ if one point is available, leaving only one free parameter which forms a single dimensional null-space. Geometrically, this means that the radius cannot be resolved from a single point. Yet, by fixing another point, one can vote locally as explained in § 5.2. While at this stage Drost and Ilic [30] prefer to vote for radius explicitly, we vote for the null-space coefficient. The difference is that [30] involves trigonometric computations before the voting stage, but vote linearly for the geometric parameter, whereas we keep linearity until the voting stage but vote for the non-linear angle θ corresponding to λ . Our approach evaluates far less trig functions (only one atan2). We plug this specific fit into our detector without changing other parts and evaluate it on scenes from [62] which contains spherical everyday objects. Tab. [1] summarizes the dataset and reports our accuracy while Fig. [12] qualitatively shows that our sphere-specific detector can indeed operate in challenging real scenarios. Our algorithm is able to detect a sphere on many difficult cases, as long as the sphere is partially visible. We also do not have to specify the radius as unlike many Hough transform based methods. Note that, due to reduced basis size ($b = |\mathbf{b}| = 1$) this type specific fit can meet real-time criteria.

6.3.5 Comparison to model based detectors

The literature is overwhelmed by the number of 3d model based pose estimation methods. Hence, we decide to compare our model-

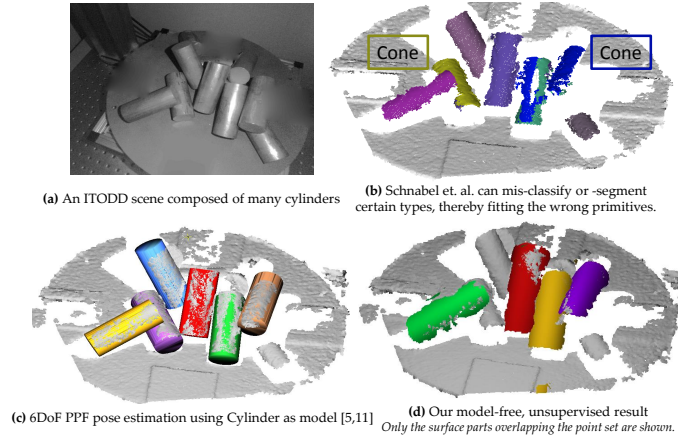


Figure 13. Multiple cylinder detection in clutter and occlusions: Our approach is type agnostic and uninformed about cylinders.

PPF3D	PPF3D-E	PPF3D-E-2D	S2D [65]	RANSAC [66]	Ours
72%	73%	74%	24%	86%	41.9%

Table 2
Results on ITODD [63] cylinders: Even without looking for a cylinder, we outperform the model based [65].

free approach to the model based ones. For that, we take the cylinders subset of the recent ITODD dataset [63] and run our generic quadric detector without training or specifying the type. Visuals of different methods are presented in Fig. [13] whereas detection performance are reported in Tab. [2]. Our task is not to explicitly estimate the pose. Thus, we manually accept a hypothesis if ICP [64] converges to a visually pleasing outcome. Note, multiple models are an important source of confusion for us, as we vote on generic quadrics. However, our algorithm outperforms certain detectors, even when we are solving a more generic problem as our shapes are allowed to deform into geometries other than cylinders.

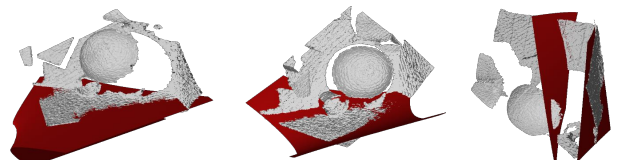


Figure 14. When planes gather a majority of the votes, our method approximates them by a close non-degenerate quadric surface. This is similar to representing a large planar football field using Earth's surface.

7 DISCUSSION AND CONCLUSIONS

We presented a fast and robust pipeline for generic primitive detection in noisy and cluttered 3D scenes. Our first contribution is a novel, linear fitting formulation for oriented point quadruplets. We thoroughly analyzed this fit and devised an efficient null-space voting which uses three pieces of point primitives plus a simple local search instead of a full four oriented point fit. Together, the fitting and voting, establish the minimalist cases known up to now - three oriented points, potentially paving the way towards real-time operation. While our detector targets a generic surface, we can, optionally, convert to a type-specific fit to boost speed and accuracy.

Unless made specific, our method is surpassed by type-specific fits in detection rate since solving the generic problem is more difficult. It remains an open issue to bring the performance to the levels of type-specific fits. Nevertheless, if the design matrix A targets a specific type, we perform even better. Degenerate cases are also difficult for us as shown in Fig. 14, but we always find a non-degenerate configuration good-enough to approximate the primitive.

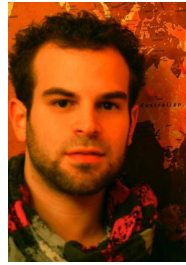
ACKNOWLEDGMENTS

The authors would like to thank Bertram Drost and Maximilian Baust for fruitful discussions.

REFERENCES

- [1] T. Birdal and S. Ilic, "Point pair features based object detection and pose estimation revisited," in *3D Vision (3DV)*. IEEE, 2015.
- [2] J. R. Miller, "Analysis of quadric-surface-based solid models," *IEEE Computer Graphics and Applications*, vol. 8, no. 1, pp. 28–42, Jan 1988.
- [3] G. Cross and A. Zisserman, "Quadric reconstruction from dual-space geometry," in *International Conference on Computer Vision*, 1998.
- [4] G. Taubin, "Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 11, pp. 1115–1138, 1991.
- [5] J. Andrews and C. H. Séquin, "Type-constrained direct fitting of quadric surfaces," *Computer-Aided Design and Applications*, 2014.
- [6] T. Tasdizen, "Robust and repeatable fitting of implicit polynomial curves to point data sets and to intensity images," Ph.D. dissertation, Brown University, 2001.
- [7] M. M. Blane, Z. Lei, H. Çivi, and D. B. Cooper, "The 3l algorithm for fitting implicit polynomial curves and surfaces to data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000.
- [8] S. Allaire, J.-J. Jacq, V. Burdin, C. Roux, and C. Couture, "Type-constrained robust fitting of quadrics with application to the 3d morphological characterization of saddle-shaped articular surfaces," in *International Conference on Computer Vision*. IEEE, 2007.
- [9] J. Andrews, "User-guided inverse 3d modeling," EECSS Department, University of California, Berkeley, Tech. Rep., May 2013.
- [10] K. Georgiev, M. Al-Hami, and R. Lakaemper, "Real-time 3d scene description using spheres, cones and cylinders," *arXiv preprint arXiv:1603.03856*, 2016.
- [11] T. Birdal, B. Busam, N. Navab, S. Ilic, and P. Sturm, "A minimalist approach to type-agnostic detection of quadrics in point clouds," in *Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2018.
- [12] S. Petitjean, "A survey of methods for recovering quadrics in triangle meshes," *ACM Computing Surveys (CSUR)*, 2002.
- [13] H. Zhao, D. Yuan, H. Zhu, and J. Yin, "3-d point cloud normal estimation based on fitting algebraic spheres," in *Image Processing (ICIP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 2589–2592.
- [14] D.-M. Yan, Y. Liu, and W. Wang, "Quadric surface extraction by variational shape approximation," in *International Conference on Geometric Modeling and Processing*. Springer, 2006, pp. 73–86.
- [15] Y. Liu, "Robust segmentation of raw point clouds into consistent surfaces," *Science China Technological Sciences*, 2016.
- [16] Z. Kukulova, J. Heller, and A. Fitzgibbon, "Efficient intersection of three quadrics and applications in computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [17] S. Uto, T. Tsuji, K. Harada, R. Kurazume, and T. Hasegawa, "Grasp planning using quadric surface approximation for parallel grippers," in *International Conference on Robotics and Biomimetics (ROBIO)*, 2013.
- [18] A. t. Pas and R. Platt, "Localizing grasp affordances in 3-d points clouds using taubin quadric fitting," *arXiv preprint arXiv:1311.3192*, 2013.
- [19] P. Gay, C. Rubino, V. Bansal, and A. Del Bue, "Probabilistic structure from motion with objects (psfmo)," in *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.
- [20] S. You and D. Zhang, "Think locally, fit globally: Robust and fast 3d shape matching via adaptive algebraic fitting," *Neurocomputing*, 2017.
- [21] D. Borrmann, J. Elseberg, K. Lengemann, and A. Nüchter, "The 3d hough transform for plane detection in point clouds: A review and a new accumulator design," *3D Research*, vol. 2, 2011.
- [22] Y. Li, X. Wu, Y. Chrysathou, A. Sharf, D. Cohen-Or, and N. J. Mitra, "Globfit: Consistently fitting primitives by discovering global relations," in *ACM Transactions on Graphics (TOG)*, 2011.
- [23] R. Schnabel, R. Wahl, and R. Klein, "Efficient ransac for point-cloud shape detection," in *Computer graphics forum*, 2007.
- [24] T.-T. Tran, V.-T. Cao, and D. Laurendeau, "Extraction of reliable primitives from unorganized point clouds," *3D Research*, 2015.
- [25] A. Monszpart, N. Mellado, G. J. Brostow, and N. J. Mitra, "Rapter: rebuilding man-made scenes with regular arrangements of planes," *ACM Trans. Graph.*, vol. 34, no. 4, pp. 103–1, 2015.
- [26] S. Oesau, F. Lafarge, and P. Alliez, "Planar shape detection and regularization in tandem," in *Computer Graphics Forum*, vol. 35, no. 1. Wiley Online Library, 2016, pp. 203–215.
- [27] H. Fang, F. Lafarge, and M. Desbrun, "Planar shape detection at structural scales," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [28] T. Czerniawski, B. Sankaran, M. Nahangi, C. Haas, and F. Leite, "6d dbscan-based segmentation of building point clouds for planar object classification," *Automation in Construction*, vol. 88, pp. 44–58, 2018.
- [29] R. Qiu, Q.-Y. Zhou, and U. Neumann, "Pipe-run extraction and reconstruction from point clouds," in *ECCV*. Springer, 2014.
- [30] B. Drost and S. Ilic, "Local hough transform for 3d primitive detection," in *3D Vision (3DV), International Conference on*. IEEE, 2015.
- [31] E. López-Rubio, K. Thurnhofer-Hemsi, Ó. D. de Cózar-Macías, E. B. Blázquez-Parra, J. Muñoz-Pérez, and I. L. de Guevara-López, "Robust fitting of ellipsoids by separating interior and exterior points during optimization," *Journal of Mathematical Imaging and Vision*, 2016.
- [32] A. Sveier, A. L. Kleppe, L. Tingelstad, and O. Egeland, "Object detection in point clouds using conformal geometric algebra," *Advances in Applied Clifford Algebras*, pp. 1–16, 2017.
- [33] D. Beale, Y.-L. Yang, N. Campbell, D. Cosker, and P. Hall, "Fitting quadrics with a bayesian prior," *Computational Visual Media*, 2016.
- [34] K.-i. Kanatani, "Further improving geometric fitting," in *International Conference on 3-D Digital Imaging and Modeling*. IEEE, 2005.
- [35] D.-M. Yan, W. Wang, Y. Liu, and Z. Yang, "Variational mesh segmentation via quadric surface fitting," *Computer-Aided Design*, 2012.
- [36] T. Mörwald, A. Richtsfeld, J. Prankl, M. Zillich, and M. Vincze, "Geometric data abstraction using b-splines for range image segmentation," in *Robotics and Automation, International Conference on*. IEEE, 2013.
- [37] K. Qin, "Representing quadric surfaces using nurbs surfaces," *Journal of Computer Science and Technology*, vol. 12, 1997.
- [38] K. Qin, W. Wang, and Z. Tang, "Representing spheres and ellipsoids using periodic nurbs surfaces with fewer control vertices," in *Pacific Graphics. Pacific Conference*. IEEE, 1998.
- [39] A. Makhmal, F. Thomas, and A. P. Gracia, "Grasping unknown objects in clutter by superquadric representation," *arXiv preprint arXiv:1710.02121*, 2017.
- [40] P. J. Besl and R. C. Jain, "Segmentation through variable-order surface fitting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 2, pp. 167–192, 1988.
- [41] N. Vaskevicius, K. Pathak, R. Pascanu, and A. Birk, "Extraction of quadrics from noisy point-clouds using a sensor noise model," in *Robotics and Automation, International Conference on*. IEEE, 2010.
- [42] P. F. Gotardo, K. L. Boyer, O. R. P. Bellon, and L. Silva, "Robust extraction of planar and quadric surfaces from range images," in *International Conference on Pattern Recognition (ICPR)*. IEEE, 2004.
- [43] J.-M. Frahm and M. Pollefeys, "Ransac for (quasi-) degenerate data (qdgsac)," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 1. IEEE, 2006, pp. 453–460.
- [44] L. Minto, G. Pagnutti, and P. Zanuttigh, "Scene segmentation driven by deep learning and surface fitting," in *ECCV Workshops*, 2016.
- [45] A. Garcia-Garcia, S. Orts-Escobedo, S. Oprea, V. Villena-Martinez, and J. Garcia-Rodriguez, "A review on deep learning techniques applied to semantic segmentation," *arXiv preprint arXiv:1704.06857*, 2017.

- [46] H. Deng, T. Birdal, and S. Ilic, "Ppfnet: Global context aware local features for robust 3d point matching," in *Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2018.
- [47] —, "Ppf-foldnet: Unsupervised learning of rotation invariant 3d local descriptors," in *The European Conference on Computer Vision (ECCV)*, September 2018.
- [48] C. Zou, E. Yumer, J. Yang, D. Ceylan, and D. Hoiem, "3d-prnn: Generating shape primitives with recurrent neural networks," in *Proc. ICCV*, vol. 2, 2017.
- [49] P. Guerrero, Y. Kleiman, M. Ovsjanikov, and N. J. Mitra, "PCPNet: Learning local shape properties from raw point clouds," *Computer Graphics Forum*, vol. 37, no. 2, pp. 75–85, 2018.
- [50] E. W. Weisstein, "Quadratic surface," MathWorld - A Wolfram Web Resource, 2017.
- [51] D. Eberly, "Classifying quadrics using exact arithmetic," Geometric Tools, LLC, Tech. Rep., 2008.
- [52] B. Le, "Classification of quadrics," 2009.
- [53] T. Tasdizen, J.-P. Tarel, and D. B. Cooper, "Algebraic curves that work better," in *Computer Vision and Pattern Recognition (CVPR)*, 1999.
- [54] G. Guennebaud, M. Germann, and M. Gross, "Dynamic sampling and rendering of algebraic point set surfaces," in *Computer Graphics Forum*, vol. 27, no. 2. Wiley Online Library, 2008, pp. 653–662.
- [55] M. Rosenblatt *et al.*, "Remarks on some nonparametric estimates of a density function," *The Annals of Mathematical Statistics*, 1956.
- [56] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, 1981.
- [57] R. I. Hartley, "In defense of the eight-point algorithm," *IEEE Transactions on pattern analysis and machine intelligence*, 1997.
- [58] T. Birdal and S. Ilic, "A point sampling algorithm for 3d matching of irregular geometries," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sept 2017, pp. 6871–6878.
- [59] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle, *Surface reconstruction from unorganized points*. ACM, 1992, vol. 26.2.
- [60] Y. Ioannou, B. Taati, R. Harrap, and M. Greenspan, "Difference of normals as a multi-scale operator in unorganized point clouds," in *3DIMPVT, 2012, International Conference on*. IEEE, 2012.
- [61] T. Birdal, I. Dobryden, and S. Ilic, "X-tag: A fiducial tag for flexible and accurate bundle adjustment," in *IEEE Conference on 3D Vision*, 2016.
- [62] S. Choi, Q.-Y. Zhou, S. Miller, and V. Koltun, "A large dataset of object scans," *arXiv:1602.02481*, 2016.
- [63] B. Drost, M. Ulrich, P. Bergmann, P. Hartinger, and C. Steger, "Introducing mvtec itodd - a dataset for 3d object recognition in industry," in *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.
- [64] A. W. Fitzgibbon, "Robust registration of 2d and 3d point sets," *Image and Vision Computing*, vol. 21, no. 13, pp. 1145–1153, 2003.
- [65] M. Ulrich, C. Wiedemann, and C. Steger, "Combining scale-space and similarity-based aspect graphs for fast 3d object recognition," *IEEE transactions on pattern analysis and machine intelligence*, 2012.
- [66] C. Papazov and D. Burschka, "An efficient ransac for 3d object recognition in noisy and occluded scenes," in *ACCV*, 2010.



Benjamin Busam is a PhD candidate with the Chair for Computer Aided Medical Procedures, Technical University of Munich and Head of Research at FRAMOS Imaging Systems. He finished his Bachelors in Mathematics at TUM in 2011. In his subsequent postgraduate programme, he studied Mathematics and Physics at ParisTech, France and at University of Melbourne, Australia, before he graduated with distinction at TU Munich in 2014. In continuation to his Master's thesis on "Projective Geometry and 3D point cloud matching", for which he received the EMVA Young Professional Award 2015, he now focuses his research and development on 2D/3D object detection and pose estimation for both outside-in and inside-out tracking.



Nassir Navab is a professor of computer science and founding director of Computer Aided Medical Procedures and Augmented Reality (CAMP) laboratories at Technische Universität München (TUM) and Johns Hopkins University (JHU). He received the Ph.D. degree from INRIA and University of Paris XI, France, and enjoyed two years of postdoctoral fellowship at MIT Media Laboratory before joining Siemens Corporate Research (SCR) in 1994. At SCR, he was a distinguished member and received the Siemens Inventor of the Year Award in 2001. In 2012, he was elected as a fellow member of MICCAI society. He received the 10 year lasting Impact Award of IEEE ISMAR in 2015. He holds 45 granted US patents and over 40 European ones and served on the program and organizational committee of over 80 international conferences including CVPR, ICCV, ECCV, MICCAI, and ISMAR. His current research interests include robotic imaging, computer aided surgery, computer vision and augmented reality.



Slobodan Ilic is currently senior key expert research scientist at Siemens Corporate Technology in Munich, Perlach. He is also a visiting researcher and lecturer at Computer Science Department of TUM and closely works with the CAMP Chair. From 2009 until end of 2013 he was leading the Computer Vision Group of CAMP at TUM, and before that he was a senior researcher at Deutsche Telekom Laboratories in Berlin. In 2005 he obtained his PhD at EPFL in Switzerland under supervision of Pascal Fua. His research interests include: 3D reconstruction, deformable surface modelling and tracking, real-time object detection and tracking, human pose estimation and semantic segmentation.



Tolga Birdal is a PhD candidate at the Chair for Computer Aided Medical Procedures / TU Munich and a Doktorand at Siemens AG. He completed his Bachelors as an Electronics Engineer at Sabanci University in 2008. Subsequently, he studied Computational Science and Engineering at TU Munich. In continuation to his Master's thesis on 3D Deformable Surface Recovery Using RGBD Cameras, he now focuses his research and development on large object detection, pose estimation and reconstruction. Recently, he is awarded Ernst von Siemens Scholarship 2015 and received the EMVA Young Professional Award 2016 for his PhD work.



Peter Sturm obtained MSc degrees from INPG (National Polytechnical Institute of Grenoble, France) and the University of Karlsruhe, both in 1994, and a Ph.D. degree from INPG in 1997, with Long Quan as advisor. His Ph.D. thesis received the SPECIF award. After a two-year postdoc at Reading University, working with Steve Maybank, he joined Inria on a permanent research position as Chargé de Recherche in 1999. Since 2006, he is Directeur de Recherche (the INRIA equivalent of Professor). Since 2015, he is Deputy Scientific Director of Inria. His main research topics have been in Computer Vision, and specifically related to camera (self-)calibration, 3D reconstruction and motion estimation. In 2011, Peter joined the STEEP research team, which is working toward contributing to sustainable development and on the use of integrated land use and transportation models for urban areas, in particular.