A Globally Optimal Method for the PnP Problem with MRP Rotation Parameterization

Manolis Lourakis
Institute of Computer Science
Foundation for Research and Technology – Hellas
P.O. Box 1385, Heraklion, Crete, 711 10, GR
Email: surname@ics.forth.gr

George Terzakis
Rovco
The Quorum, Bond Street, Bristol, BS1 3AE, UK
Email: name.surname@rovco.com

***Abstract*—The perspective-n-point (PnP) problem is of fundamental importance in computer vision. A global optimality condition for PnP that is independent of a particular rotation parameterization was recently developed by Nakano. This paper puts forward a direct least squares, algebraic PnP solution that extends Nakano's work by combining his optimality condition with the modified Rodrigues parameters (MRPs) for parameterizing rotation. The result is a system of polynomials that is solved using the Gröbner basis approach. An MRP vector has twice the rotational range of the classical Rodrigues (i.e., Cayley) vector used by Nakano to represent rotation. The proposed solution provides strong guarantees that the full rotation singularity associated with MRPs is avoided. Furthermore, detailed experiments provide evidence that our solution attains accuracy that is indistinguishable from Nakano's Cayley-based method with a moderate increase in computational cost.**

I. Introduction

The perspective-n-point (PnP) problem concerns the recovery of 6D pose given the central projections of $n \geq 3$ known 3D points on a calibrated camera. It arises often in vision and robotics applications involving localization, pose tracking and multi-view 3D reconstruction, hence a considerable amount of research has been devoted to its study.

An optimality condition that is independent of any rotation parameterization was recently developed by Nakano [1]. He coupled it with the Cayley (a.k.a. Rodrigues) rotation representation and proposed a globally optimal direct least squares solution for PnP. Building upon Nakano's condition, this paper proposes to parameterize rotation with the modified Rodrigues parameters (MRPs) and develops a new globally optimal solution to PnP. The MRPs are a minimal rotation parameterization defined by the stereographic projection of the unit quaternion sphere onto a three-dimensional hyperplane. They can also be considered as a generalization of the classical Rodrigues parameters.

The contribution of this work is threefold. First, it is the first to corroborate the applicability of MRPs to the development of a minimal solver to a geometric vision problem. Second, it demonstrates that by using MRPs, the domain of validity of the Cayley representation employed in [1] can be doubled, with no impact on accuracy and only a moderate impact on computational performance. Third, it provides a principled way of dealing with singularities in the representation of rotation. The proposed solution inherits from [1] several attractive properties, such as applicability to arbitrary number of points n, global optimality, linear complexity and completeness in the sense that it returns all solutions irrespective of the value of n and the spatial arrangement of points in 3D.

The remainder of the paper is organized as follows. Section II provides an overview of previous work on the PnP problem. Section III gives a brief introduction to Gröbner basis solvers while Section IV discusses rotation parameterization and MRPs in particular. Nakano's optimality condition is presented in Section V, whereas Section VI shows how it can be used in conjunction with MRPs. Since the MRP representation is inherently singular at $\pm 2\pi$ (i.e. can not represent full rotations), Section VII explains how this singularity can be avoided. Experimental results are presented in Section VIII and the paper is concluded in Section IX.

II. Previous Work

The P3P problem, which is the minimal variant of PnP for $n = 3$, was first studied by Grunert [2] in the 19th century. Other solutions were later proposed by Fischler and Bolles [3], Gao et al. [4] and Kneip et al. [5]. P3P is known to admit up to four different solutions, whereas in practice it usually has just two. As a result, a fourth point is often used for disambiguation, giving rise to the P4P problem [6], [7].

For the generic (i.e., $n \geq 3$) PnP problem, Lu et al. [8] proposed an iterative method which is initialized with a weak perspective approximation and refines the rotation matrix via successive solutions to the absolute orientation problem [9]. More recently, Lepetit et al. [10] developed EPnP, in which an initial pose estimate is obtained by rotating and translating the 3D points in their eigenspace and thereafter, solving the least squares (LS) formulation without the orthonormality constraints. The estimate can be optionally improved iteratively. Although EPnP can yield very good results, it nonetheless relies on an unconstrained LS estimate, which can be skewed by noise in the data. It can also get trapped in local minima, particularly for small size inputs. Improvements to the basic EPnP algorithm have been proposed in [11]. Very recently, we developed SQPnP, a globally optimal algorithm that casts PnP as a quadratically constrained quadratic program [12].

The turn of attention during the last 15 years to algebraic geometry methods in conjunction with minimal parameterization schemes for rotation has led to rapid developments on the PnP problem. For example, the Direct Least Squares (DLS) algorithm by Hesch and Roumeliotis [13] employs the Cayley transform [14] to parametrize the rotation matrix and solve with resultants a quartic polynomial system obtained by the first order optimality conditions. DLS inherits a singularity for rotations equal to π from its use of the Cayley transform [15]. A very similar solution was later proposed by Zheng et al. [16], with improved rotation parameterization for singularity avoidance and a Gröbner polynomial solver. In principle, both methods require a number of elimination steps to recover up to 40 solutions which are later substituted in the cost function to determine the global minimum. Kneip et al. [17] proposed UPnP that is applicable to both central and non-central cameras.

Public availability of the automatic solver generator [18] simplified the use of Gröbner basis solvers and has facilitated the development of several approaches applying them to the PnP problem. For example, Bujnak et al. [7], [19] designed solutions for the P3P and P4P with unknown focal length problems and reported relatively short execution times. A more recent work that makes use of the automatic generator is that by Nakano [1]. He derives a rotation optimality condition without Lagrange multipliers and proposes a globally optimal DLS method, parameterized by the Cayley representation. Nakano's condition is presented in more detail in Section V.

III. Gröbner Basis Solvers

Several problems in geometric computer vision that relate to relative or absolute camera pose estimation can be solved using a minimal number of point correspondences. Such minimal solutions are very important in practical applications since they can be combined with random sampling schemes such as RANSAC [3] in order to identify outliers and robustly estimate pose [20]. The size of a minimal set is directly related to the total number of samples that need to be drawn in order to find a solution with acceptable confidence. As a result, techniques for solving systems of algebraic equations that arise in the solution of minimal or near minimal problems have attracted considerable interest in recent years. These systems are characterized by many unknowns and high degrees that render them challenging for general polynomial system Gröbner solvers. Fortunately, many vision problems are characterized by the property that the monomials appearing in the set of input polynomials for non-degenerate image measurements remain the same [18]. This, in turn, implies that general-purpose solvers are not necessary for dealing with them. Instead, special solvers tailored to such particularities have demonstrated to be more efficient and robust [18], [21].

The Gröbner basis method [22], [23] is a technique for solving systems of polynomial equations that has found increased use in computer vision due to its effectiveness in solving systems of algebraic equations that arise in the solution of minimal problems. It is applied using the action matrix approach that aims to transform the problem of finding the solutions to a system of polynomials to that of finding the eigenvectors of a special action matrix. In simpler terms, this means that once the action matrix is found, the system of polynomials can be solved through its eigendecomposition for which established numerical techniques are available.

The dimension of the action matrix directly relates to the number of solutions to the original system. The elements of the action matrix are polynomial combinations of the coefficients of the input polynomials. These polynomial combinations are found in an offline, preprocessing stage that is performed only once for a certain problem and aims to generate an elimination template by multiplying the original equations with different monomials. Gauss-Jordan elimination applied to the elimination template generates the action matrix. Online operation of the solver involves populating the elimination template using the input coefficients, computing its Gauss-Jordan elimination to extract the action matrix and extracting the solution from the latter using eigendecomposition.

The Gröbner basis method was first introduced in the field of computer vision by Stewénius [24], who applied it manually as did all early adopters of this method that followed. However, applying the Gröbner basis method manually is a laborious and error-prone task, which led Kukelova et al. [18] to develop a generator that automatically constructs elimination templates based on symbolic computer algebra. An automatic generator accelerates the development of Gröbner basis solvers, extends their applicability and makes their design simpler and repeatable. Availability of the generator [18] sparked more interest in the topic and instigated the study of additional minimal solvers, e.g. [7], [19], [1]. Improvements to the automatic generator were later proposed in [21].

IV. Parameterization of Rotation

We begin by providing some background on the Cayley representation of a rotation. The Cayley transform [14] parameterizes a rotation matrix with the three-parameter Rodrigues or Gibbs vector $\mathbf{g}$ by mapping the skew-symmetric matrix $[\mathbf{g}]_\times$ of vector $\mathbf{g}$ to a rotation matrix:

$$\begin{aligned}\mathbf{R} =& (\mathbf{I}_3 - [\mathbf{g}]_\times)^{-1}(\mathbf{I}_3 + [\mathbf{g}]_\times) \\ =& (\mathbf{I}_3 + [\mathbf{g}]_\times)(\mathbf{I}_3 - [\mathbf{g}]_\times)^{-1}.\end{aligned} \tag{1}$$

The Cayley transform is often invertible (i.e., when $\mathbf{I}_3+[\mathbf{g}]_\times$ is invertible) and the inverse is given by $(\mathbf{I}_3-[\mathbf{g}]_\times)(\mathbf{I}_3+[\mathbf{g}]_\times)^{-1}$.

A widely used representation of orientation is provided by unit quaternions. Quaternions are a redundant but free of singularities parameterization. They are defined as $\mathbf{q} = (q_s,\ \mathbf{q}_v)^\mathsf{T}$ with $\|\mathbf{q}\| = 1$ and can be seen as vectors that lie on the unit sphere in $\mathbb{R}^4$ [25], [15]. Components q_s and $\mathbf{q}_v$ are respectively the quaternion's scalar and vector parts. Antipodal unit quaternions, i.e. $\mathbf{q}$ and $-\mathbf{q}$, represent the same rotation. This implies that the mapping between quaternions and 3D orientations is 2:1, a property known as the "double cover" or the "2-fold symmetry".

The Rodrigues vector $\mathbf{g}$ is related to the representations of the principal axis – angle as well as the quaternion with

$$\mathbf{g} = \mathbf{e}\tan\left(\frac{\Phi}{2}\right) = \frac{\mathbf{q}_v}{q_s}, \tag{2}$$

where $\mathbf{e}$, Φ are the principal axis and angle, whereas $(q_s,\ \mathbf{q}_v)$ is the quaternion [25]. From the last formula is clear that the Rodrigues vector representation is singular and discontinuous when the principal rotation angle Φ equals π. This is the price paid for the Rodrigues vector providing a one-to-one mapping for rotations, i.e. both $\mathbf{q}$ and $-\mathbf{q}$ map to the same vector $\mathbf{g}$. The inverse relation, i.e. the unit quaternion in terms of the Rodrigues vector is given by

$$\mathbf{q} = \frac{1}{\sqrt{1+\|\mathbf{g}\|^2}} \begin{pmatrix} 1 \\ \mathbf{g} \end{pmatrix}. \tag{3}$$

We next focus on the modified Rodrigues parameters (MRPs) which are a minimal rotation parameterization. They are the stereographic coordinates of quaternions and, as such, they are mapped rationally and bijectively to the quaternion sphere. Rotation parameterization with the MRP vector ψ originated in the field of aerospace engineering [25], [26] and was recently introduced to computer vision by [15]. Vector ψ is related to the representations of the principal axis – angle as well as that of the quaternion with the following expressions

$$\psi = \mathbf{e} \tan\left(\frac{\Phi}{4}\right) = \frac{\mathbf{q}_v}{1+q_s}, \tag{4}$$

where $\mathbf{e}$, Φ are again the principal axis and angle, and $\mathbf{q}_v$, q_s are respectively the vector and scalar parts of quaternion $\mathbf{q}$. MRPs double the range of the classical Rodrigues parameters and encounter a singularity at a principal rotation of $\Phi = \pm 2\pi$.

Conversely, the unit quaternion in terms of ψ is [15]:

$$q_s = \frac{1-\|\psi\|^2}{1+\|\psi\|^2}, \quad \mathbf{q}_v = \frac{2\psi}{1+\|\psi\|^2} \tag{5}$$

From eqs. (2) and (5) it is clear that the Rodrigues vector is related to the MRPs via $\mathbf{g} = 2\psi/(1-\|\psi\|^2)$.

Just like quaternions $\mathbf{q}$ and $-\mathbf{q}$ refer to the same orientation, there are two possible MRP vectors that represent the same physical orientation; one is called the "shadow" MRP of the other and is denoted next as ψ_s. However, their coordinates are not opposite but are related by the following expression:

$$\psi_s = -\frac{\psi}{\|\psi\|^2} \tag{6}$$

If an MRP vector corresponds to a principal rotation Φ, its shadow vector corresponds to rotation $\Phi - 2\pi$. Both represent the same orientation, but become singular at different angles; notice that ψ_s is singular for $\Phi = 0$ and zero for $\Phi = \pm 2\pi$.

V. Nakano's Optimality Condition

Nakano developed in [1] an optimality condition for PnP that is independent of the rotation parameterization. In this section, we provide a brief overview of his approach and refer the interested reader to the original publication for details.

Let $\mathbf{M}_i \in \mathbb{R}^3$, $i \in \{1,\ldots,n\}$ be a set of known Euclidean world points and $\mathbf{m}_i$ their corresponding homogeneous normalized projections on the $Z=1$ plane of an unknown camera coordinate frame. PnP seeks to recover the rotation matrix $\mathbf{R}$ and translation vector $\mathbf{t}$ which are such that the homogeneous image points $\mathbf{R}\,\mathbf{M}_i + \mathbf{t}$ are in agreement with the known projections $\mathbf{m}_i$. Various cost functions can be used to quantify this agreement, based on either algebraic or geometric distances (e.g. reprojection error). Nakano employed an algebraic cost function which demands that $\mathbf{m}_i$ and $\mathbf{R}\,\mathbf{M}_i + \mathbf{t}$ are parallel:

$$\begin{aligned} &\underset{\mathbf{R},\,\mathbf{t}}{\text{minimize}} \sum_{i=1}^{n} \|[\mathbf{m}_i]_\times (\mathbf{R}\,\mathbf{M}_i + \mathbf{t})\|^2 \\ &\text{s.t.} \quad \mathbf{R}^\mathsf{T}\mathbf{R} = \mathbf{I}_3 \ \text{ and } \ \det(\mathbf{R}) = 1, \end{aligned} \tag{7}$$

where $\|\ \|$ indicates the Euclidean norm, $\mathbf{I}_3$ is the 3×3 identity matrix and $[\mathbf{m}]_\times$ denotes the skew-symmetric cross product matrix associated with the vector $\mathbf{m}$, specifically:

$$[\mathbf{m}]_\times = \begin{bmatrix} 0 & -m_z & m_y \\ m_z & 0 & -m_x \\ -m_y & m_x & 0 \end{bmatrix}. \tag{8}$$

Algebraic manipulation after directly expressing the translation vector in terms of rotation, converts (7) to the following non-linear problem subject to the orthonormality and unit determinant constraints for the rotation [1]:

$$\begin{aligned} &\underset{\mathbf{R}}{\text{minimize}} \ \mathbf{r}^\mathsf{T}\,\mathbf{M}\,\mathbf{r} \\ &\text{s.t.} \quad \mathbf{R}^\mathsf{T}\mathbf{R} = \mathbf{I}_3 \ \text{ and } \ \det(\mathbf{R}) = 1, \end{aligned} \tag{9}$$

where $\mathbf{M}$ is a symmetric $\mathbb{R}^{9\times9}$ matrix and $\mathbf{r} \in \mathbb{R}^9$ the vector formed by stacking the rows of $\mathbf{R}$. For future use, we will denote the inverse operation as $\text{mat}(\mathbf{r}) = \mathbf{R}$.

Using the first-order optimality condition, Nakano showed in [1] that an optimality condition for the constrained minimization (9) is given by the following pair of equations:

$$\begin{aligned} \mathbf{P} &\equiv \mathbf{R}^\mathsf{T}\,\text{mat}(\mathbf{M}\mathbf{r}) - \text{mat}(\mathbf{M}\mathbf{r})^\mathsf{T}\,\mathbf{R} = \mathbf{0} \\ \mathbf{Q} &\equiv \text{mat}(\mathbf{M}\mathbf{r})\,\mathbf{R}^\mathsf{T} - \mathbf{R}\,\text{mat}(\mathbf{M}\mathbf{r})^\mathsf{T} = \mathbf{0}. \end{aligned} \tag{10}$$

The diagonal elements of both 3×3 matrices $\mathbf{P}$ and $\mathbf{Q}$ are zero, whereas their off-diagonal elements are second degree polynomials in $\mathbb{R}$. Owing to symmetry, eq. (10) amounts to a total of six polynomials corresponding to the elements of $\mathbf{P}$ and $\mathbf{Q}$ that lie above the main diagonal, i.e.

$$\begin{aligned} &P_{1,2} = 0 \quad P_{1,3} = 0 \quad P_{2,3} = 0 \\ &Q_{1,2} = 0 \quad Q_{1,3} = 0 \quad Q_{2,3} = 0, \end{aligned} \tag{11}$$

where $P_{i,j}$, $Q_{i,j}$ are the (i,j) elements of matrices $\mathbf{P}$ and $\mathbf{Q}$.

From the above discussion, solving (9) is equivalent to solving the system of polynomials (11). Nakano studied three different rotation parameterizations for solving this system, namely a rotation matrix, a quaternion, and Cayley's. Considering that the computational cost and numerical stability of a solver employing Gröbner basis techniques depends on the size of the elimination and action matrices, it is preferable to employ parameterizations that give rise to the smallest such matrices. Hence, observing that the Cayley parameterization results in the most compact representation, Nakano proposed its use for dealing with PnP and showed experimentally that it achieves accuracy similar to [16].

VI. Optimal DLS PnP with MRPs

By inspecting eqs. (2) and (4), it is clear that employing the Cayley transform formula of eq. (1) with MRP vector ψ yields a rotation with half the angle of rotation represented by ψ. Hence, the rotation matrix corresponding to ψ is [25], [15]

$$\begin{aligned}\mathbf{R} =&(\mathbf{I}_3 - [\psi]_\times)^{-2}(\mathbf{I}_3 + [\psi]_\times)^2 \\ =&\mathbf{I}_3 + \frac{4(1-\|\psi\|^2)[\psi]_\times + 8[\psi]_\times^2}{(1+\|\psi\|^2)^2}.\end{aligned} \quad (12)$$

The denominator $(1+\|\psi\|^2)^2$ in eq. (12) can be eliminated by multiplying both sides with it. The resulting scaled matrix $\mathbf{I}_3(1+\|\psi\|^2)^2 + 4(1-\|\psi\|^2)[\psi]_\times + 8[\psi]_\times^2$ can be substituted in place of $\mathbf{R}$ in eqs. (11). This is because equations (11) are homogeneous, therefore they are not affected by scaling with a non-zero factor. The ensuing polynomial system in the components of vector ψ involves six multivariate equations of maximum degree eight and is solved with the Gröbner basis method using the automatic generator of [18].

The size of the elimination template computed by the generator for the MRP parameterization is 321×401 and that of the corresponding action matrix 80×80. For comparison, Nakano found that the elimination template and the action matrix were 124×164 and 40×40 for the Cayley parameterization and respectively 630×710 and 80×80 when rotation was parameterized with the quaternion.

In theory, the MRP parameterization can yield up to 80 solutions for rotation, however we have empirically observed that in practice, the non-imaginary ones are usually around 40. The best solution can be chosen as that minimizing either eq. (9) or the reprojection error. Expectedly, due to the 2-fold symmetry of the MRP parameterization, the solutions found form shadow pairs. From eq. (4), it is evident that

$$\text{If } \ \Phi \leq \pi \ \text{ then } \ ||\psi|| \leq 1 \ \text{ else } \ ||\psi|| > 1.$$

Thus, to disambiguate the shadow pair solutions, it suffices to consider only those with a ψ magnitude not exceeding one, i.e. those inside or on the unit sphere in $\mathbb{R}^3$ where the representation is respectively 1:1 and 2:1. Given an estimate of rotation, the translation can be computed in closed form, as in [1]. Cheirality is enforced by requiring that the 3D points transformed with the estimated pose lie in front of the camera.

It is noted at this point that the elimination template for the MRP parameterization would be smaller if its generation incorporated symmetry elimination techniques such as those presented in [27], [28]. However, as implementations of such methods are not publicly available, we have not applied them in the context of the present work.

VII. Avoiding Singularities

As already mentioned, the Cayley and MRP parameterizations of rotation are singular for rotations $\pm\pi$ and $\pm 2\pi$, respectively. To cope with singularities in the Cayley parameterization, Hesch and Roumeliotis [13] have performed rotations of the input 3D points around the three principal axes by $\pi/2$, solved the three resulting distinct PnP problems and selected the solution achieving the best score. Interestingly, they appear to have been unaware of the method of sequential rotations which was developed by Shuster to deal with singularities in the QUEST attitude estimation algorithm [29] that makes implicit use of the Cayley vector [30], [31].

The key idea behind sequential rotations is to solve for a rotation representing orientation with respect to a rotated reference frame away from the singularity and then convert the estimated rotation back to the original frame. More concretely, assuming that $\mathbf{R}_a^\pi$ is a prior rotation by π about one of the principal coordinate axes (sub. $a \in \{x, y, z\}$ specifies the axis) applied to the input 3D points, the sought rotation can be seen as the product of $\mathbf{R}_a^\pi$ followed by a "surplus rotation" away from π, i.e. $\mathbf{R} = \mathbf{S}_a \mathbf{R}_a^\pi$. For $\mathbf{M}$ of (9), a pre-rotation by $\mathbf{R}_a^\pi$ transforms it as $\mathbf{A}\mathbf{M}\mathbf{A}^\mathsf{T}$, where $\mathbf{A}$ is the block diagonal matrix

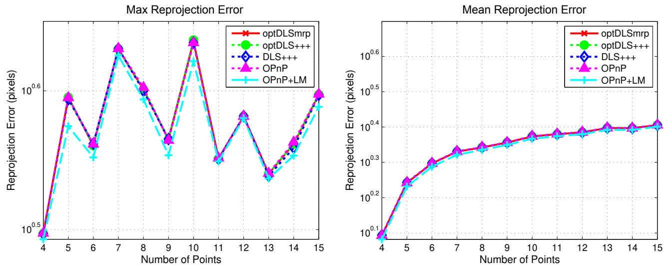
$$\mathbf{A} = \begin{bmatrix} \mathbf{R}_a^\pi & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{R}_a^\pi & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{R}_a^\pi \end{bmatrix}$$

and $\mathbf{0}_3$ is the 3×3 zero matrix. It is shown in [30] that the angle of rotation of at least one of matrices $\mathbf{S}_x$, $\mathbf{S}_y$, $\mathbf{S}_z$ and $\mathbf{R}$ is at most $2\pi/3$. In other words, at least one of these four matrices is away from the singularity at 2π, hence at least one of the systems (11) for the corresponding PnP problems can be safely solved to provide $\mathbf{S}_a$. Note that the fourth matrix corresponds to no pre-rotation, i.e. the latter equals the identity $\mathbf{I}_3$. The usefulness of sequential rotations lies in that an acceptable rotated frame is found without an initial orientation estimate.

VIII. Experiments

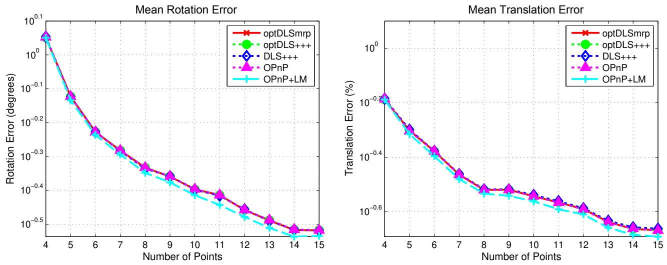
This section reports results from the comparison of the proposed MRP solver with Matlab implementations of state-of-the art algebraic solutions to PnP, namely optDLS [1], DLS [13] and OPnP [16]. We employed the original authors' own implementations of DLS and OPnP, whereas we generated the optDLS solver ourselves and wrote a wrapper for selecting the best of the solutions it computes using as criterion the reprojection error. The same criterion disambiguates the solutions of our solver. For better performance, in the code generated for optDLS and our proposed solver by the automatic Gröbner generator [18], we substituted the calls to `rref` for the Gauss-Jordan elimination with equivalent invocations of `mldivide`. Furthermore, rather than naively performing matrix multiplications, matrix $\mathbf{M}$ in eq. (9) is efficiently constructed by exploiting its structure (cf. supplement of [1]). As we verified in [12], OPnP and optDLS are among the best performing PnP solvers currently available.

Both optDLS and DLS include the three rotations by $\pi/2$ preprocessing to avoid the Cayley singularity. To stress this, the two methods are indicated in the plots with the '+++' suffix, i.e. optDLS+++ and DLS+++. Similarly, the proposed MRP solver is applied using sequential rotations equal to π as explained in Sec. VII, and is henceforth denoted by optDLSmrp. Non-linear minimization of the reprojection error with Levenberg-Marquardt is also initiated at the solution computed by OPnP. This corresponds to the maximum likelihood



(a) Max RMS reprojection error (b) Mean RMS reprojection error

Fig. 1. Plots of maximum (left) and mean (right) root mean square (RMS) reprojection error for 500 executions of each PnP solver on n random points, $4 \leq n \leq 15$. For each n, the points are repeatedly sampled from a previously generated point population contaminated with additive zero-mean Gaussian noise. Each plot represents the results obtained by points drawn from a different population and whose projections were contaminated with additive zero-mean Gaussian noise of standard deviation $\sigma{=}2$ pixels. Notice the logarithmic vertical axis scales.



(a) Rotation error (b) Translation error

Fig. 2. Plots of the mean rotation (left) and translation (right) errors for the experiments of Fig. 1. The plots summarize 500 trials with n random points, $4 \leq n \leq 15$ and additive zero-mean Gaussian noise of standard deviation $\sigma{=}2$ pixels.

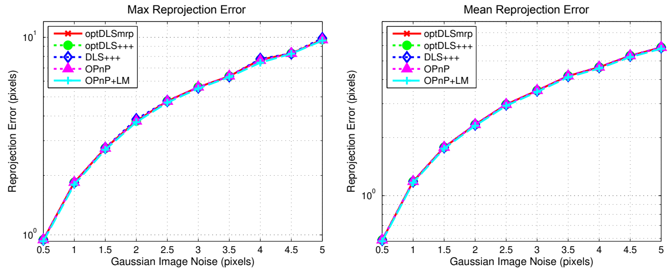
estimate and is indicated with OPnP+LM in the plots. No such root polishing is applied to the output of any other method.

We adopted the evaluation protocol of [16] and employed their testing framework[1] with limited modifications. For the convenience of the reader, we repeat next the basic assumptions. The dimensions of the simulated camera are 640×480 pixels and its focal length 800 pixels. The simulated 3D points are randomly distributed in a rectangular parallelepiped with size $[-2, 2] \times [-2, 2] \times [4, 8]$. The ground truth translation $\mathbf{t}_{true}$ is chosen equal to the centroid of the simulated 3D points and the ground truth translation $\mathbf{R}_{true}$ is generated randomly.

The rotational error is measured as the maximum angle (in degrees) between corresponding columns of the true and estimated rotation matrices, i.e. $\max_{i=1}^{3}\{\frac{\arccos\left((\mathbf{r}^i_{true})^\top \mathbf{r}^i\right)180}{\pi}\}$, while the translational error is measured by the percent of the relative difference between the true and estimated translation vectors, i.e. $\frac{\|\mathbf{t}_{true}-\mathbf{t}\|100}{\|\mathbf{t}\|}$.

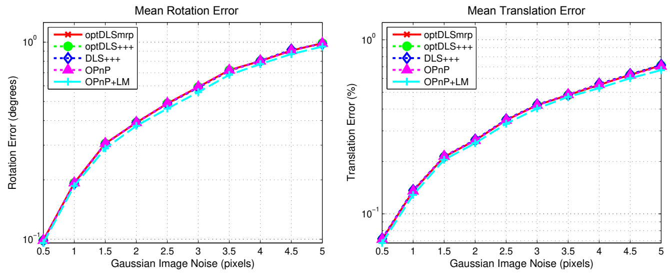
In the first set of experiments, the number of points n is increased from 4 to 15 in steps of 1, and zero-mean Gaussian noise with $\sigma = 2$ pixels is added to the image projections. For each value of n, 500 independent test sets are generated. Figure 1(a) shows the maximum root mean

[1] Downloaded from https://sites.google.com/site/yinqiangzheng/.



(a) Max RMS reprojection error (b) Mean RMS reprojection error

Fig. 3. Plots of maximum (left) and mean (right) root mean square (RMS) reprojection error for 500 executions of each PnP solver on $n = 10$ random points, contaminated with additive zero-mean Gaussian noise with standard deviation $0.5 \leq \sigma \leq 5$ pixels. For each σ, 500 trials are performed.



(a) Rotation error (b) Translation error

Fig. 4. Plots of the mean rotation (left) and translation (right) errors for the experiments of Fig. 3. The plots summarize 500 trials with $n = 10$ random points and additive zero-mean Gaussian noise of standard deviation $0.5 \leq \sigma \leq 5$ pixels.

square (RMS) reprojection errors across all executions of the compared methods. The motivation for plotting the maximum of the RMS reprojection errors is that it will indicate whether these methods consistently approach the minimum for a given n. The same is not true if the mean is used in place of the maximum, as any spikes will be averaged out. Nevertheless, for completeness, plots showing the corresponding mean RMS reprojection errors are also included in Fig. 1(b). Mean rotation and translation errors for the same experiments are in Figure 2.

A second set of experiments focuses on the performance of the methods when the number of points is kept fixed to 10 and the noise standard deviation is varied from 0.5 to 5 pixels in increments of 0.5. Again, 500 trials are performed for each noise level. Figure 3 shows the maximum and mean RMS reprojection errors across all executions of the compared methods, whereas the corresponding mean rotation and translation errors are in Figure 4.

It is clear from the plots that the results obtained by our optDLSmrp are practically indistinguishable from those obtained with the rest of the algebraic methods. Point sets of smaller size are generally more susceptible to noise, an observation that has previously been made by others [32]. Finally, the reprojection error achieved by OPnP+LM is the smallest, which is expectable as it corresponds to the maximum likelihood estimate that minimizes a geometrically meaningful error function.

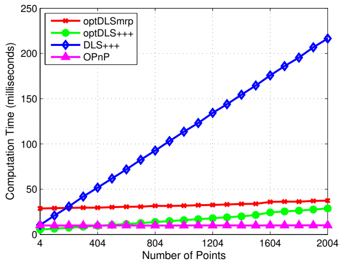


Fig. 5. Average execution times (in milliseconds) of the various solvers. The plot summarizes 500 trials for $n \in \{4, 104, 204, \dots, 2004\}$ random points and additive zero-mean Gaussian noise of standard deviation $\sigma{=}2$ pixels.

The computational cost of Matlab implementations of the various methods is illustrated in Figure 5, which shows their average execution time for fixed noise with $\sigma = 2$ and varying numbers of points n that increase from 4 to 2004 in increments of 100. For each value of n, 500 trials are performed. Compared to optDLS, optDLSmrp requires between about 1.3 to 5 times more time, for large and small sets respectively. Both achieve sublinear computation time.

IX. Conclusion

This paper has presented a globally optimal direct least squares method for the PnP problem. It couples Nakano's optimality condition with the modified Rodrigues parameters (MRPs) representation of rotation to derive a system of polynomial equations. The Gröbner basis solution of this system is combined with a scheme that avoids the full rotation singularity of MRPs to obtain a novel solution to PnP. Comprehensive experiments with simulated data have demonstrated the effectiveness and robustness of the proposed approach. An implementation is at https://github.com/mlourakis/optDLSmrp/.

Acknowledgment

M. Lourakis has been funded by the EU's H2020 programme under GAs 826506 (sustAGE) and 101017151 (FELICE).

References

[1] G. Nakano, "Globally optimal DLS method for PnP problem with Cayley parameterization," in *British Machine Vision Conference*, 2015, pp. 78.1–78.11.
[2] J. Grunert, "Das pothenotische Problem in erweiterter Gestalt nebst über seine Anwendungen in Geodäsie," *Grunerts Archiv für Mathematik und Physik*, 1841.
[3] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[4] X.-S. Gao, X.-R. Hou, J. Tang, and H.-F. Cheng, "Complete solution classification for the perspective-three-point problem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 930–943, 2003.
[5] L. Kneip, D. Scaramuzza, and R. Siegwart, "A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2011, pp. 2969–2976.
[6] R. Horaud, B. Conio, and O. Leboulleux, "An analytic solution for the perspective 4-point problem," *Computer Vision, Graphics, and Image Processing*, vol. 47, no. 1, pp. 33–44, 1989.
[7] M. Bujnak, Z. Kukelova, and T. Pajdla, "A general solution to the P4P problem for camera with unknown focal length," in *Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–8.
[8] C.-P. Lu, G. D. Hager, and E. Mjolsness, "Fast and globally convergent pose estimation from video images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 6, pp. 610–622, 2000.
[9] B. K. Horn, H. M. Hilden, and S. Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices," *JOSA A*, vol. 5, no. 7, pp. 1127–1135, 1988.
[10] V. Lepetit, F. Moreno-Noguer, and P. Fua, "EPnP: An accurate O(n) solution to the PnP problem," *International Journal of Computer Vision*, vol. 81, no. 2, p. 155, 2009.
[11] S. Urban, J. Leitloff, and S. Hinz, "MLPnP - a real-time maximum likelihood solution to the perspective-n-point problem," in *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 3, 2016, pp. 131–138.
[12] G. Terzakis and M. Lourakis, "A consistently fast and globally optimal solution to the perspective-n-point problem," in *European Conference on Computer Vision*. Springer Nature Switzerland, 2020, pp. 1–17.
[13] J. A. Hesch and S. I. Roumeliotis, "A direct least-squares (DLS) method for PnP," in *International Conference on Computer Vision*. IEEE, 2011, pp. 383–390.
[14] A. Cayley, "Sur quelques propriétés des déterminants gauches," *Journal für die reine und angewandte Mathematik*, vol. 32, pp. 119–123, 1846.
[15] G. Terzakis, M. Lourakis, and D. Ait-Boudaoud, "Modified Rodrigues parameters: an efficient representation of orientation in 3D vision and graphics," *Journal of Mathematical Imaging and Vision*, vol. 60, no. 3, pp. 422–442, 2018.
[16] Y. Zheng, Y. Kuang, S. Sugimoto, K. Åström, and M. Okutomi, "Revisiting the PnP problem: A fast, general and optimal solution," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 2344–2351.
[17] L. Kneip, H. Li, and Y. Seo, "UPnP: An optimal O(n) solution to the absolute pose problem with universal applicability," in *European Conference on Computer Vision*, 2014, pp. 127–142.
[18] Z. Kukelova, M. Bujnak, and T. Pajdla, "Automatic generator of minimal problem solvers," in *European Conference on Computer Vision*. Springer, 2008, pp. 302–315.
[19] M. Bujnak, Z. Kukelova, and T. Pajdla, "New efficient solution to the absolute pose problem for camera with unknown focal length and radial distortion," in *Asian Conference on Computer Vision*. Springer, 2010, pp. 11–24.
[20] M. Lourakis and X. Zabulis, "Model-based pose estimation for rigid objects," in *International Conference on Computer Vision Systems*. Springer, 2013, pp. 83–92.
[21] V. Larsson, K. Åström, and M. Oskarsson, "Efficient solvers for minimal problems by syzygy-based reduction," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2383–2392.
[22] D. Cox, J. Little, and D. O'Shea, *Using Algebraic Geometry*, ser. Graduate Texts in Mathematics. Springer New York, 2005, vol. 185.
[23] B. Buchberger and M. Kauers, "Groebner basis," *Scholarpedia*, vol. 5, no. 10, p. 7763, 2010.
[24] H. Stewénius, "Gröbner basis methods for minimal problems in computer vision," Ph.D. dissertation, Lund University, Sweden, 2005.
[25] M. Shuster, "A survey of attitude representations," *Journal of the Astronautical Sciences*, vol. 41, no. 4, pp. 439–517, 1993.
[26] H. Schaub and J. L. Junkins, "Stereographic orientation parameters for attitude dynamics: A generalization of the Rodrigues parameters," *Journal of the Astronautical Sciences*, vol. 44, no. 1, pp. 1–19, 1996.
[27] E. Ask, Y. Kuang, and K. Åström, "Exploiting p-fold symmetries for faster polynomial equation solving," in *International Conference on Pattern Recognition*, 2012, pp. 3232–3235.
[28] V. Larsson and K. Åström, "Uncovering symmetries in polynomial systems," in *European Conference on Computer Vision*, 2016, pp. 252–267.
[29] M. Shuster and S. Oh, "Three-axis attitude determination from vector observations," *Journal of Guidance, Control, and Dynamics*, vol. 4, no. 1, pp. 70–77, 1981.
[30] M. Shuster and G. Natanson, "Quaternion computation from a geometric point of view," *Journal of the Astronautical Sciences*, vol. 41, pp. 545–556, 1993.
[31] D. Mortari, M. Angelucci, and F. Markley, "Singularity and attitude estimation," in *Advances in the Astronautical Sciences*, vol. 105, 2000, pp. 479–493.
[32] E. Rosten, G. Reitmayr, and T. Drummond, "Improved RANSAC performance using simple, iterative minimal-set solvers," *CoRR*, vol. abs/1007.1432, 2010.