Globally Optimal DLS Method for PnP Problem with Cayley parameterization

Gaku Nakano
g-nakano@cq.jp.nec.com

Information and Media Processing Labs, NEC Corporation
Kawasaki, Japan

The perspective-n-point (PnP) problem, which estimates 3D rotation and translation of a calibrated camera from n pairs of known 3D points and corresponding 2D image points, is a classical problem but still fundamental in the computer vision community. It is well studied that the PnP problem can be solved by at least three points [1]. If $n \geq 4$, the PnP problem becomes a nonlinear problem where the number of the solutions depend on n and the shape of the scene. This paper proposes an efficient, scalable, and globally optimal DLS method parameterized by Cayley representation, which has been regarded as a unsuitable parametrization due to its singularity.

First we derive a new optimality condition without Lagrange multipliers. Letting $\mathbf{p}_i = [x_i, y_i, z_i]^T$ be an i-th 3D point and $\mathbf{m}_i = [u_i, v_i, 1]^T$ be the corresponding calibrated image point in homogeneous coordinates, the PnP problem can be formulated as a nonlinear optimization

$$\begin{aligned} &\min_{\mathbf{R},\mathbf{t}} \quad \sum_{i=1}^{n} \left\| [\mathbf{m}_i]_\times (\mathbf{R}\mathbf{p}_i + \mathbf{t}) \right\|^2 \\ &\text{s.t.} \quad \mathbf{R}^T\mathbf{R} = \mathbf{I}, \quad \det(\mathbf{R}) = 1 \end{aligned} \tag{1}$$

where $[\ \ ]_\times$ denotes a matrix representation of the vector cross product. The optimal translation $\mathbf{t}$ can be expressed as a function of $\mathbf{R}$ since Eq. (1) is a linear least squares problem of $\mathbf{t}$. If we define $\mathbf{M} \in \mathbb{R}^{9\times 9}$ by a symmetric coefficient matrix computed from $\mathbf{p}_i$ and $\mathbf{m}_i$, Lagrange function of Eq. (1) can be written as

$$L(\mathbf{R}, \mathbf{S}, \lambda) = \frac{1}{2}\mathbf{r}^T\mathbf{M}\mathbf{r} - \frac{1}{2}\text{trace}\left(\mathbf{S}(\mathbf{R}^T\mathbf{R} - \mathbf{I})\right) - \lambda(\det(\mathbf{R}) - 1). \tag{2}$$

Here, $\mathbf{r}$ is a vector form of $\mathbf{R}$, λ is a Lagrange multiplier, and $\mathbf{S} \in \mathbb{R}^{3\times 3}$ is a symmetric matrix of Lagrange multipliers. Then, the first-order optimality condition is given by

$$\frac{\partial L}{\partial \mathbf{R}} = \text{mat}(\mathbf{M}\mathbf{r}) - \mathbf{R}\mathbf{S} - \lambda\mathbf{R} = \mathbf{0}, \tag{3}$$

$$\frac{\partial L}{\partial \mathbf{S}} = \mathbf{R}^T\mathbf{R} - \mathbf{I} = \mathbf{0}, \tag{4}$$

$$\frac{\partial L}{\partial \lambda} = \det(\mathbf{R}) - 1 = 0, \tag{5}$$

where mat$(\ \ )$ is a reshaping operator from a 9×1 vector to a 3×3 square matrix. Multiplying $\mathbf{R}^T$ from the left-hand and the right-hand side of Eq. (3), we have

$$\mathbf{R}^T \text{mat}(\mathbf{M}\mathbf{r}) = \mathbf{S} + \lambda\mathbf{I}, \tag{6}$$

$$\text{mat}(\mathbf{M}\mathbf{r})\mathbf{R}^T = \mathbf{R}(\mathbf{S} + \lambda\mathbf{I})\mathbf{R}^T. \tag{7}$$

Since $\mathbf{S} + \lambda\mathbf{I}$ is a symmetric matrix, the left-hand side of Eqs. (6) and (7) must be symmetric matrices. Hence, we obtain a new optimality condition where the Lagrange multipliers are eliminated:

$$\mathbf{P} = \mathbf{R}^T \text{mat}(\mathbf{M}\mathbf{r}) - \text{mat}(\mathbf{M}\mathbf{r})^T\mathbf{R} = \mathbf{0}, \tag{8}$$

$$\mathbf{Q} = \text{mat}(\mathbf{M}\mathbf{r})\mathbf{R}^T - \mathbf{R}\text{mat}(\mathbf{M}\mathbf{r})^T = \mathbf{0}. \tag{9}$$

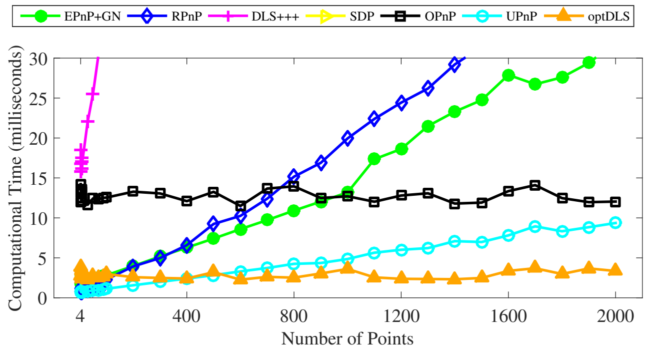

Figure 1: Computational time for varying $4 \leq n \leq 2000$ and fixed $\sigma = 2.0$

Let $P_{j,k}$ and $Q_{j,k}$ be the (j,k) element of $\mathbf{P}$ and $\mathbf{Q}$, respectively. Obviously the diagonal elements are zeros, $P_{j,j} = Q_{j,j} = 0$. On the other hand, the non-diagonal elements are second degree polynomials in $\mathbf{R}$. Due to the symmetry, $P_{j,k} = P_{k,j}$ and $Q_{j,k} = Q_{k,j}$, we have six polynomials in total:

$$\begin{aligned} P_{1,2} = 0, \quad P_{1,3} = 0, \quad P_{2,3} = 0, \\ Q_{1,2} = 0, \quad Q_{1,3} = 0, \quad Q_{2,3} = 0. \end{aligned} \tag{10}$$

Although Eq. (10) is derived from a general rotation parameterization, any parameterizations satisfying Eqs. (4), (5), and (10) are usable for solving the PnP problem. By using Kukelova *et al.*'s automatic Gröbner basis solver, we compared three types of parameterizations: general rotation matrix, quaternion, and Cayley parameterization. Table 1 is a comparison of the above parameterizations with existing methods. For efficiency and stability, this paper selects Cayley parameterization whose elimination template and action matrix are the smallest among the three representations.

The proposed method was evaluated on synthetic data with existing methods in terms of robustness against image noise and computational time. While the proposed method has the same robustness as the state-of-the-art, OPnP [5], the computational time is less than 3 msec for almost all cases. As shown in Fig. 1, it is the fastest especially for $n \geq 400$ points.

The conclusion is that the new optimality condition without the Lagrange multipliers can be solved by any types of rotation parameterizations. Furthermore, Cayley parameterization is suitable for realtime applications, such as augmented reality and visual SLAM, where hundreds or thousands of the points is not a rare situation.

[1] B. M. Haralick *et al*. Review and analysis of solutions of the three point perspective pose estimation problem, *INT J COMPUT VISION*, 13(3):331–356, 1994.

[2] J. A. Hesch and S. I. Roumeliotis. A direct least-squares (DLS) method for PnP. In *ICCV*, 2011.

[3] L. Kneip *et al*. UPnP: An Optimal O (n) Solution to the Absolute Pose Problem with Universal Applicability. In *ECCV*, 2014.

[4] Z. Kukelova *et al*. Automatic generator of minimal problem solvers. In *ECCV*, 2008.

[5] Y. Zheng *et al*. Revisiting the pnp problem: A fast, general and optimal solution. In *ICCV*, 2013.

Table 1: Comparison of rotation parameterizations

	Existing methods			Proposed params		
	DLS [2] (Cayley)	OPnP [5] (Non-unit Quat.)	UPnP [3] (Quaternion)	Rotation Matrix	Quaternion	Cayley
# of unknowns	3	4	4	9	4	3
# of equations	3	4	8	27	7	6
# of solutions	27	40	8	40	80	40
singularity	yes	no	no	no	no	yes
elim. templ.	120×120	348×376	141×149	1936×1976	630×710	124×164
action matrix	27×27	40×40	8×8	40×40	80×80	40×40