Simplified Mirror-Based Camera Pose Computation via Rotation Averaging

Gucan Long[1] Laurent Kneip[2] Xin Li[1] Xiaohu Zhang[1] Qifeng Yu[1]
[1]College of Aerospace Science and Engineering, National University of Defense Technology, P.R. China.
[2]Research School of Engineering, Australian National University.

We propose a novel approach to compute the camera pose with respect to a reference object given only mirrored views. The latter originate from a planar mirror at different unknown poses. This problem is highly relevant in several extrinsic camera calibration scenarios, where the camera cannot see the reference object directly. Important examples are given by:

- A camera that is custom-mounted in the front or back of a car. The corresponding calibration task is commonly denoted *eye-body calibration*. Despite the fact that a 3D model of the vehicle might be readily available, it is not visible inside the camera's field of view, thus blocking a calibration based on regular object pose estimation.

- Multiple cameras pointing into different directions such that their fields of view are no longer overlapping. Deriving the relative pose from the absolute poses with respect to a commonly observed calibration target hence becomes impossible. The problem of calibrating non-overlapping cameras is often approached by solving the so-called *hand-eye calibration* problem. A daily-life example of a non-overlapping camera pair is given by smart-phone devices equipped with a frontback camera pair.

This paper is focussing on mirror-based calibration or pose estimation which is crucial whenever the camera motion is constrained, thus blocking the applicability of ego-motion based approaches. Examples are given by a car exerting planar motion, or a static setup such as surveillance cameras. However, in contrast to numerous existing methods, our approach does not employ the *fixed axis rotation constraint* discovered in [1], but represents a more elegant formulation as a rotation averaging problem. Our contributions are:

- A solution that solves for all mirrored poses individually, and then combines them within a single subsequent fusion stage.

- Fusion of mirrored poses as a simple rotation averaging problem [2]. This enables mirror-based pose estimation in closed-form and linear complexity under the chordal L2-metric, and outlier-robust iterative averaging under the L1-norm.

- A generalization of rotation averaging [2] to improper rotation matrices.

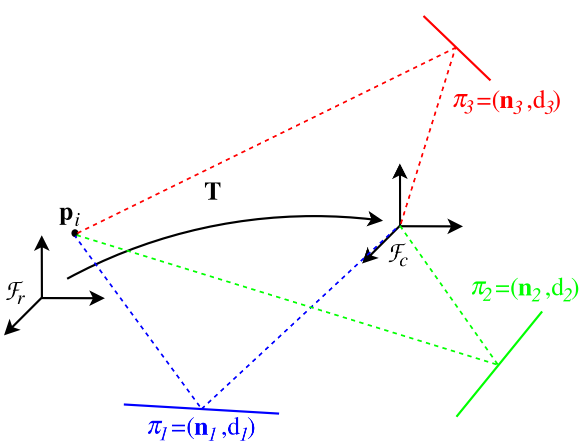

Figure 1: The geometry of the mirror-based camera pose estimation problem.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

Figure 2: Extrinsic calibration of a smart phone's camera-pair: Both front and back camera can see the chessboard only via mirror reflections. We apply our mirror based pose estimation to find both cameras' absolute pose with respect to the calibration target given only (at least 3) mirrored views of the object. By combining the absolute poses, we can then easily retrieve the relative pose between the non-overlapping pair of cameras.

The geometry of the problem is described in Figure 1. Let $\mathcal{F}_c$ be a camera frame without a direct view onto 3D points $[\mathbf{p}_1, \ldots, \mathbf{p}_m]$ defined inside a reference frame $\mathcal{F}_r$. Let $[\mathbf{n}_1, \ldots, \mathbf{n}_i]$ be different mirror orientations rendering the 3D points visible inside the camera's field of view. Let $\mathbf{T}$ describe the pose of the camera with respect to $\mathcal{F}_r$. Mirror-based camera pose estimation consists of computing $\mathbf{T}$ while the coordinates of the 3D points are known, and the parameters of all $\mathbf{n}_i$ are unknown. Let $\mathbf{R}$ be the unknown orientation of the camera. We may also define the pose (and orientation) of a virtual camera as the pose of a (left-handed) camera frame that returns the same observation than the real camera, however without employing a mirror. As shown in the paper, the real and the (improper) virtual orientation notably follow the rule

$$\tilde{\mathbf{R}}_i = \mathbf{R}\left(\mathbf{I} - 2\mathbf{n}_i\mathbf{n}_i^T\right),$$

with $\mathbf{I} - 2\mathbf{n}_i\mathbf{n}_i^T$ being a Householder matrix. We therefore find the unknown real camera orientation $\mathbf{R}$ from virtual ones by minimizing the sum of residual errors

$$C(\mathbf{R}, \mathbf{n}_i) = \sum_i^n \varepsilon(\tilde{\mathbf{R}}_i(\mathbf{I} - 2\mathbf{n}_i\mathbf{n}_i^T), \mathbf{R})^p,$$

which requires a simultaneous retrieval of the unknown mirror plane normal vectors $\mathbf{n}_i$. ε represents the geometric, chordal, or quaternion metric, and $p = 1$ or $p = 2$.

In our paper, we present two solutions:

- An efficient, closed-form solution based on the chordal L2-mean.

- A geometric L1-mean algorithm providing robustness with respect to outlier poses.

Our algorithms provide state-of-the-art accuracy, and we furthermore demonstrate how the method can be used successfully to calibrate the non-overlapping camera pair of a regular smart-phone.

[1] J. Gluckmann and S. K. Nayar. Catadioptric stereo using planar mirrors. *International Journal of Computer Vision (IJCV)*, 44(1):65–79, 2001.

[2] R. Hartley, J. Trumpf, Y. Dai, and H. Li. Rotation averaging. *International Journal of Computer Vision (IJCV)*, 103(3):267–305, 2013.