

Accurate Fiducial Mapping For Pose Estimation Using Manifold Optimization

Xiao Hu¹, Jakobsen Jakob¹, Knudsen Per¹, Wei Jiang²

¹National Space Institute, Technical University of Denmark

²School of Electronic and Information Engineering Beijing Jiaotong University



DTU Space
National Space Institute

Problem

The accurate pose estimation for moving objects within a given workspace is one of the most fundamental tasks for many applications.

A typical solution is to use a motion capture system such as Vicon^a, Optitrack^b which can provide pose with high precision. However, these systems are usually costly and can only work indoor with limited coverage.

Another common solution is the integration system with high-end Inertial Navigation System (INS) and a Real Time Kinematic (RTK) Global Positioning System (GPS). Nonetheless, its performance will deteriorate dramatically in clustering environments and even stop to work in the GPS denied places.

^a<http://www.vicon.com>

^b<http://www.optitrack.com>

Contributions

We propose an approach to build a map of fiducial markers based on manifold optimization and then extend the fiducial map for pose estimation. Process pose fusion and optimization directly on manifold to avoid numeric problems, singularities and maintain smoothness. The fiducial map based pose estimation system is cost-effective, lightweight and can work both indoor and outdoor.

Preliminaries

G: the global coordinate system.

M: the local coordinate system of fiducial marker. The local coordinate system of i^{th} marker is represented by M_i .

C: the camera coordinate system. Similarly, C_k denotes the camera coordinate system of the k^{th} image.

$SO(3)$ is defined with (1):

$$SO(3) = R \in \mathbb{R}^{3 \times 3} | RR^T = I_{3 \times 3}, \det(R) = 1 \quad (1)$$

The Special Euclidean Group $SE(3)$ is further defined (2):

$$SE(3) = \left\{ T = \begin{bmatrix} R & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4} | R \in SO(3), \mathbf{t} \in \mathbb{R}^3 \right\} \quad (2)$$

An additive increment in the vector space will associate to an multiplication increment on manifold, which follows the below approximation

$$\exp((\xi + \delta\xi)^\wedge) \approx \exp((\mathcal{J}\delta\xi)^\wedge) \exp(\xi^\wedge) \quad (3)$$

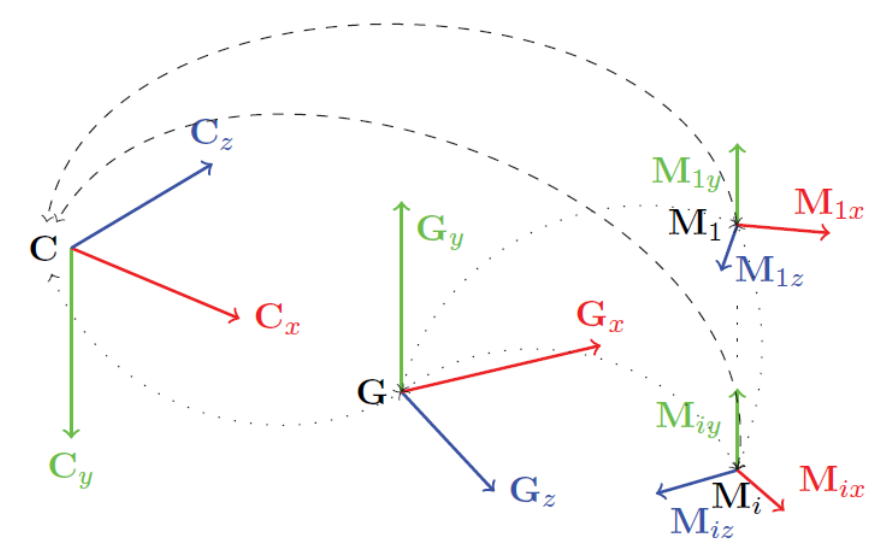


Figure 1: Coordinate Systems. Dashed connections are transformation matrices directly measured and dotted connections are transformation matrices to be optimized.

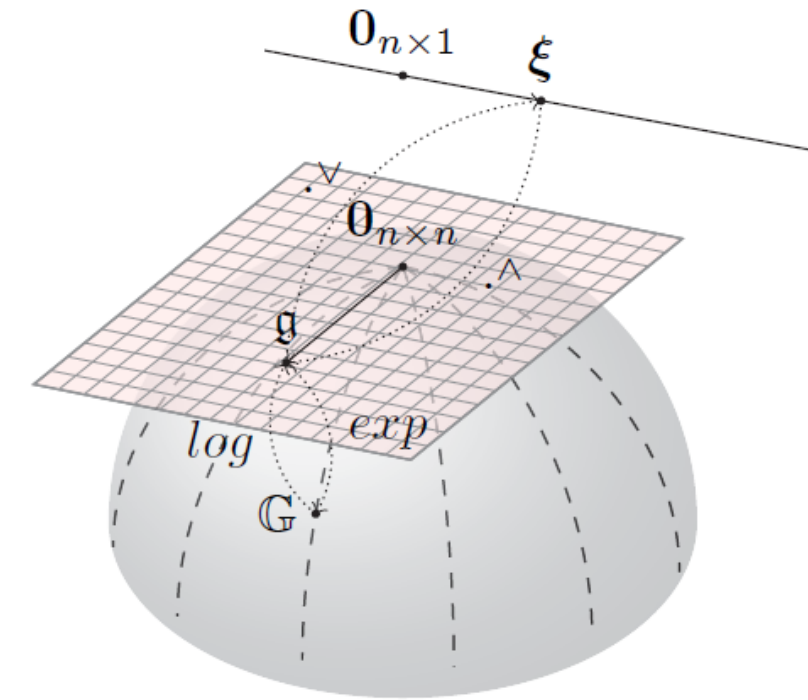


Figure 2: Mapping relationship among matrix Lie Group, Lie Algebra and the vector space.

The pinhole camera model which is appropriate to most cameras is used in this work. We assume that the camera is supposed to be well-calibrated and the intrinsic parameters are precise since no optimization will be made on them.

References

Methods

This section details each step of the proposed method with a systematic chart shown in Fig. 3.

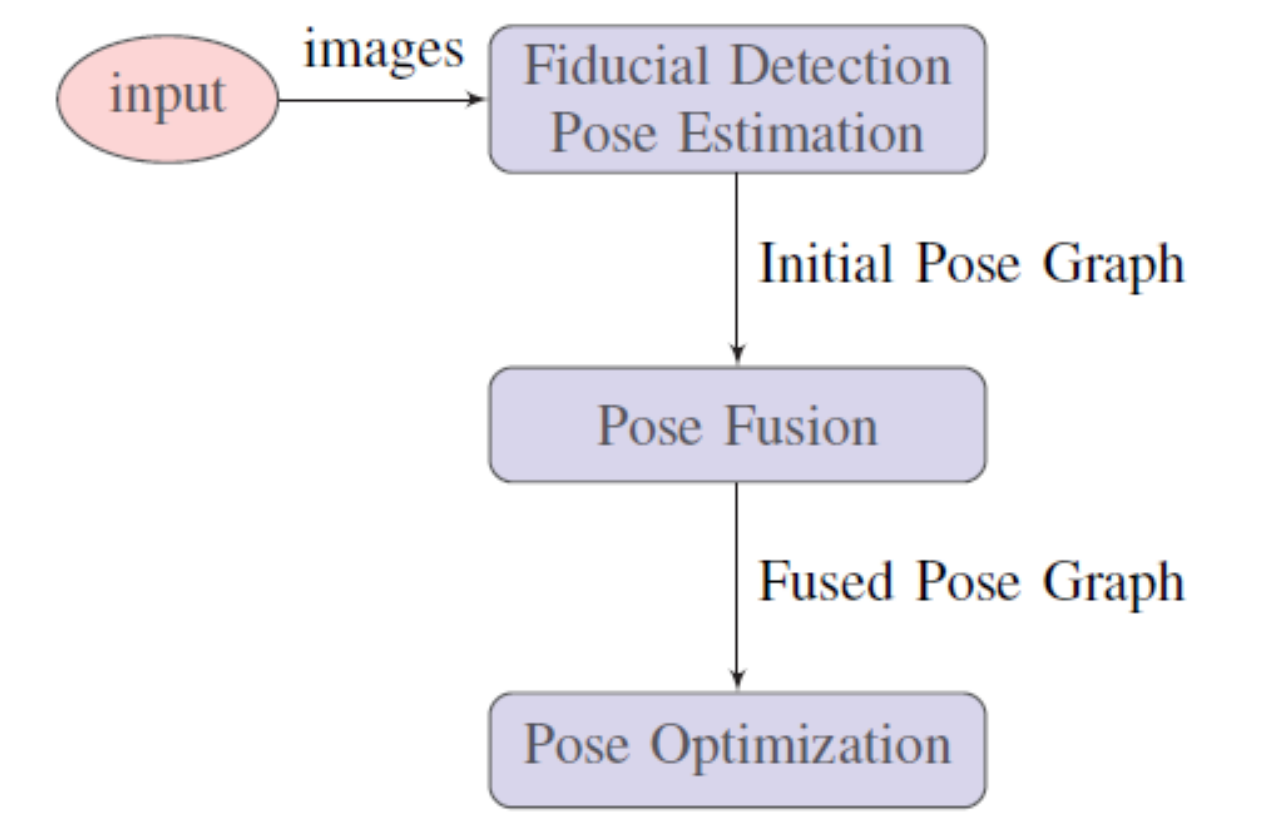


Figure 3: Flowchart of proposed approach.

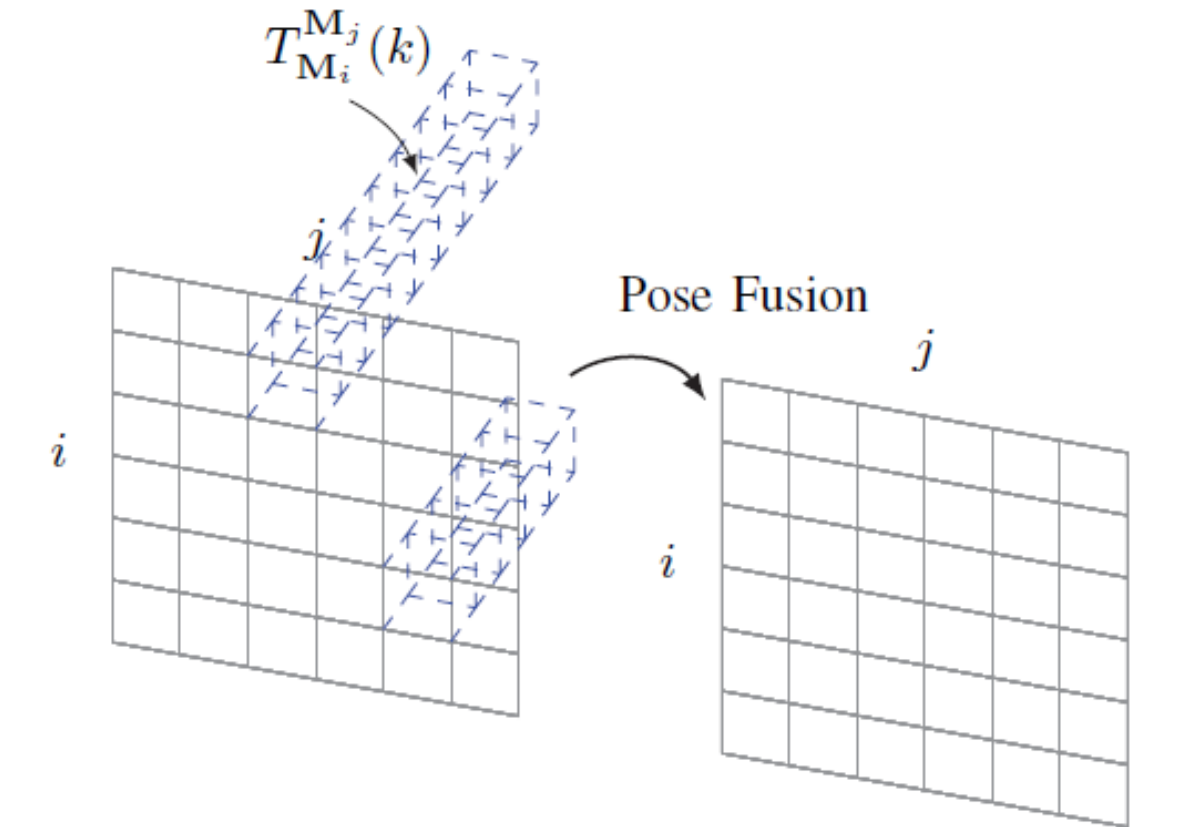


Figure 4: Schematic diagram of the established initial pose graph pg and pose graph pg_f after pose fusion process.

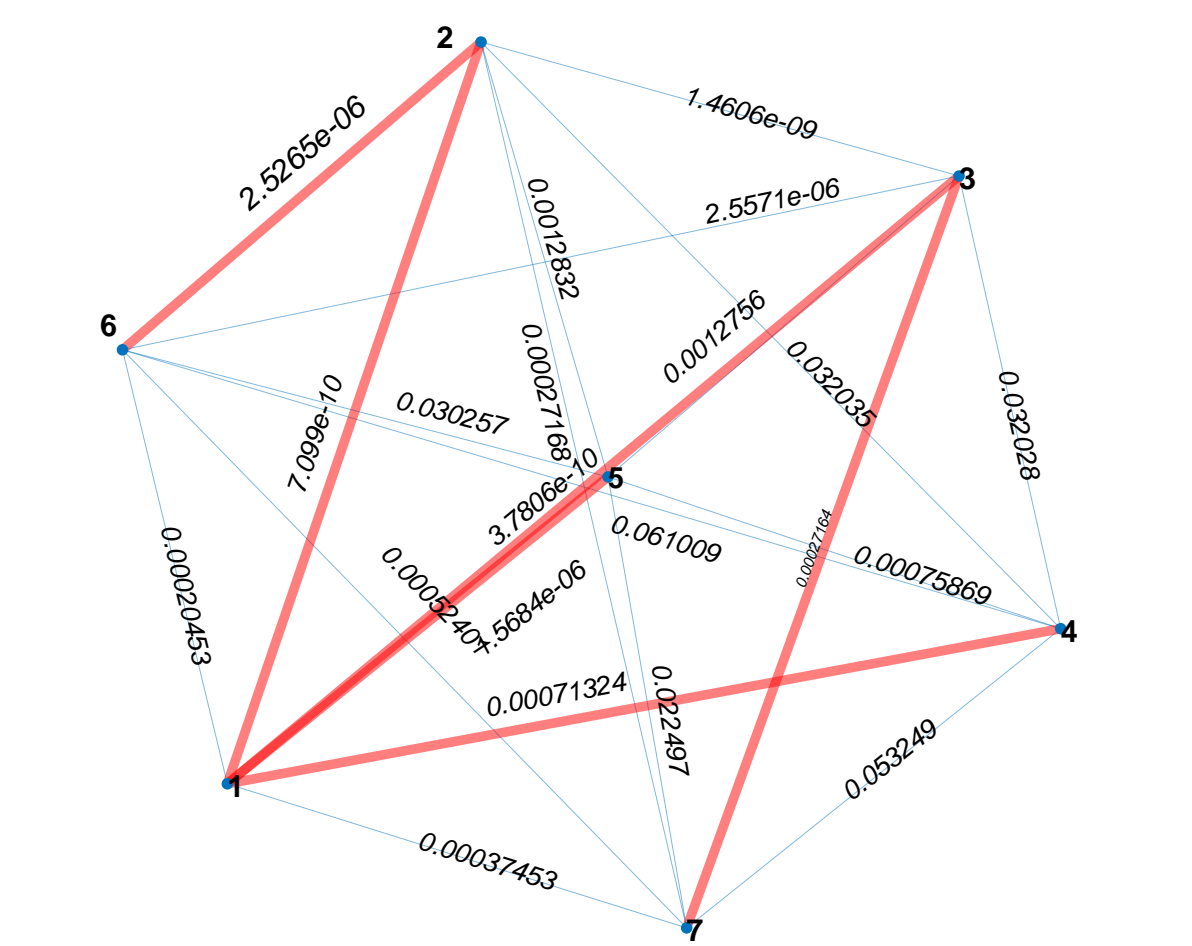


Figure 5: Undirected Graph with blue dots denoting markers, solid light blue edges denoting costs. Red thick edges describe the minimum spanning tree obtained.

ArUco detector implemented

in OpenCV for fiducial detection and estimate pose with the RPNP algorithm [?].

The relative pose from i^{th} marker to j^{th} marker, if both viewed in k^{th} camera frame, can be further computed as

$$T_{M_i}^{M_j^{-1}} = T_{M_i}^{C_k^{-1}} T_{M_j}^{C_k}$$