UR5 / Multi-Camera Calibration

Gerard Kennedy & Zheyu Zhang

1 Overview

This document outlines the setup and method for calibrating a multi-camera rig with a UR5 robot arm.

2 Problem Formulation

The task is to develop a pipeline for intrinsic calibration of a multi-camera system, extrinsic calibration of the camera system relative to a global reference frame, and computation of the pose of a UR5 robot and its workspace (a table) relative to the same global reference frame. The global reference frame is defined as the base of the UR5 (see Figure 1).

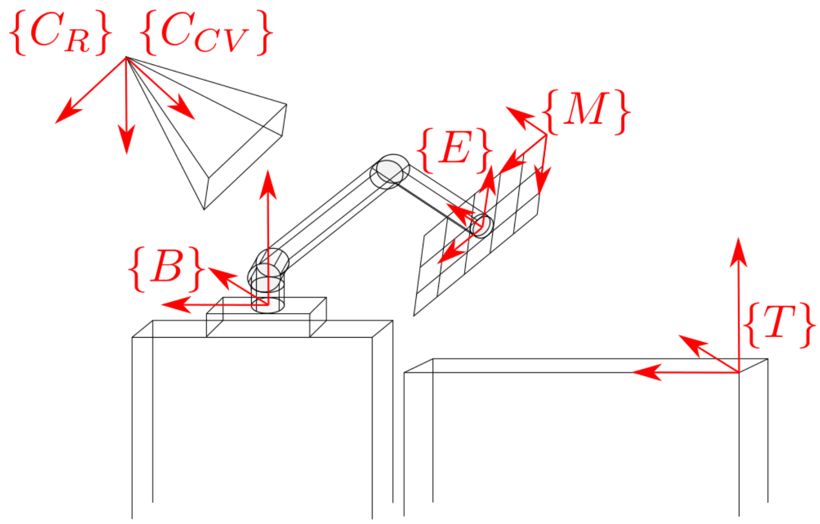

Figure 1: System setup. $\{C_{CV}\}$ refers to camera pose using the Computer Vision convention, C_R refers to camera pose using the Robotics convention, $\{M\}$ refers to the ChAruco Marker board, $\{E\}$ refers to end-effector pose, $\{B\}$ refers to the pose of the UR5 base, and $\{T\}$ refers to the pose of the table.

We want to find $\boldsymbol{^{B}X_{C_R}}, \boldsymbol{^{B}X_{T}}, \boldsymbol{^{B}X_{E}}$ (i.e. the pose of the camera, table, and end-effector with respect to the base of the UR5). These poses are emphasised in **bold** in this document.

From the setup we know:

- ${}^{M}X_T$ - the pose of the table with respect to the Marker board. This can be found by moving the end-effector so that the calibration board is on the table and measuring the offset.
- ${}^{M}X_E$ - the pose of the Marker board with respect to the end-effector when the board is on the table. This is an offset that can be measured, and is related to the provided CAD model.
- ${}^{C_{CV}}X_{C_R}$ - the transformation between the Computer Vision camera pose convention (z-axis in the camera focal direction) and the Robotics pose convention (z-axis is normal to the ground plane):

$$
{}^{C_{CV}}X_{C_R} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

- $\boldsymbol{{}^{B}X_E(\theta_i)}$ - the pose of the end-effector with respect to the UR5 base. This can be found via the UR5 forward kinematics given the UR5 joint angles $\boldsymbol{\theta} \in \mathbb{R}^{6\times 1}$ at frame i.

From the extrinsic camera calibration we know:

- ${}^{C_{CV}}X_M$ - the pose of the Marker board with respect to the camera. This is the output of the extrinsic calibration.

The known and required relative poses are summarised in Figure 2.

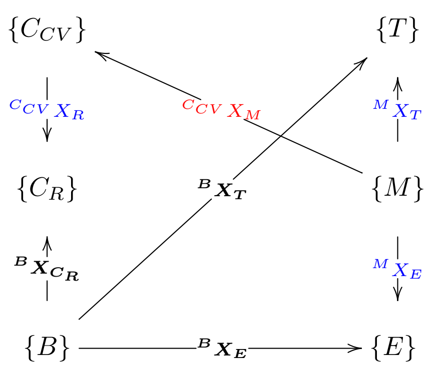

Figure 2: Reference Frames. The 3 required relative poses are shown in **bold**. The relative poses that can be obtained by measuring an offset or using a known transformation are shown in blue. The relative pose provided by the extrinsic calibration is shown in red.

The remaining poses can be found using the poses listed above. Using the calibration we can find the pose of the camera with respect to the UR5 base

$\boldsymbol{{}^{B}X_{C_R}}$:

$$
\begin{aligned}
{}^{C_{CV}}X_E &= {}^{C_{CV}}X_M \cdot {}^{M}X_E \\
{}^{E}X_{C_R} &= ({}^{C_{CV}}X_{C_R}^{-1} \cdot {}^{C_{CV}}X_E)^{-1} \\
{}^{B}X_{C_R}(\boldsymbol{\theta}_i) &= {}^{B}X_E(\boldsymbol{\theta}_i) \cdot {}^{E}X_{C_R} \\
\boldsymbol{{}^{B}X_{C_R}} &= \text{mean}({}^{B}X_{C_R}(\boldsymbol{\theta}_i))
\end{aligned} \tag{1}
$$

Two different approaches can be used to obtain $\boldsymbol{{}^{B}X_{T}}$ for each image frame.

Method 1 (using the forward kinematics):

$$
\begin{aligned}
{}^{E}X_{T_0} &= {}^{M}X_E^{-1} \cdot {}^{M}X_T \\
{}^{B}X_{T_0} &= {}^{B}X_E \cdot {}^{E}X_T
\end{aligned} \tag{2}
$$

Method 2 (using the extrinsic calibration):

$$
{}^{B}X_{T_1} = {}^{B}X_{C_R} \cdot {}^{C_R}X_{C_{CV}} \cdot {}^{C_{CV}}X_M \cdot {}^{M}X_T \tag{3}
$$

Using these two Methods we can test the calibration with the following error function.

$$
E = ||{}^{B}X_{T_0} - {}^{B}X_{T_1}||_F \tag{4}
$$

The final output $\boldsymbol{{}^{B}X_{T}}$ pose is computed as the mean of Equation (2):

$$
\boldsymbol{{}^{B}X_{T}} = \text{mean}({}^{B}X_{T_1}) \tag{5}
$$

The mean poses (Equations (1) and (5) are obtained by averaging translation values as usual, and averaging rotation over the associated Lie algebras $\mathfrak{so}(3)$. For example, for Equation (1):

$$
\begin{aligned}
\Omega_i &= \log({}^{B}R_C(\boldsymbol{\theta}_0)^{\top} \cdot {}^{B}R_C(\boldsymbol{\theta}_i)), R \in \mathbf{SO(3)} \\
\bar{\Omega} &= \frac{1}{N}\sum_{i=0}^{N} \Omega_i \\
{}^{B}\bar{R}_C &= {}^{B}R_C(\boldsymbol{\theta}_0)\exp(\bar{\Omega})
\end{aligned}
$$

Rotation averaging is performed iteratively until convergence using the Karcher (aka Geodesic L2) mean.

Figure 3 shows an example of the ChAruco board attached to the UR5 end-effector.

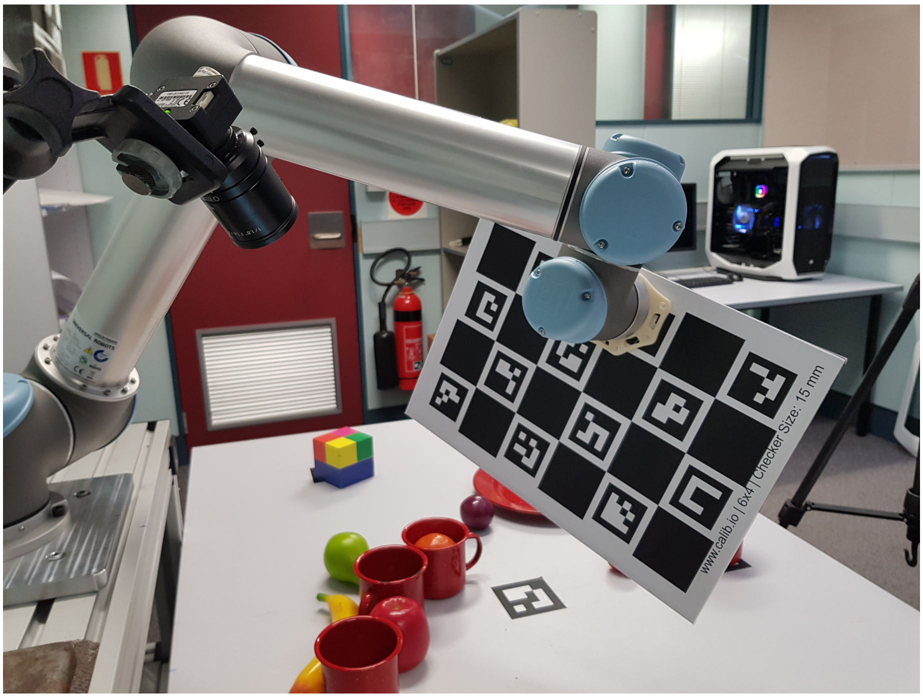

Figure 3: Example Setup.