

# 相机标定文档

MF1923010 丁明宇

## Abstract

在图像测量过程以及机器视觉应用中，为确定空间物体表面某点的三维几何位置与其在图像中对应点之间的相互关系，必须建立相机成像的几何模型，这些几何模型参数就是相机参数。在大多数条件下这些参数必须通过实验与计算才能得到，这个求解参数的过程就称之为相机标定（或摄像机标定）。无论是在图像测量或者机器视觉应用中，相机参数的标定都是非常关键的环节，其标定结果的精度及算法的稳定性直接影响相机工作产生结果的准确性。因此，做好相机标定是做好后续工作的前提，提高标定精度是科研工作的重点所在。

相机的参数标定可以看作是机器学习中的回归问题，我们要回归的参数便是相机的内外参。

相机标定方法有：传统相机标定法、主动视觉相机标定方法、相机自标定法。

本文档针对相机姿态（ $R$ 、 $T$ ）的标定，着重对 1999 年由张正友教授提出的经典的张氏相机标定法和 1999 提出的光束平差法（**Bundle Adjustment**，以下简称 **BA 算法**）进行介绍以及实验结果的展示分析。

## I . Introduction

传统相机标定法需要使用尺寸已知的标定物，通过建立标定物上坐标已知的点与其图像点之间的对应，利用一定的算法获得相机模型的内外参数。根据标定物的不同可分为三维标定物和平面型标定物。三维标定物可由单幅图像进行标定，标定精度较高，但高精密三维标定物的加工和维护较困难。平面型标定物比三维标定物制作简单，精度易保证，但标定时必须采用两幅或两幅以上的图像。传统相机标定法在标定过程中始终需要标定物，且标定物的制作精度会影响标定

结果。同时有些场合不适合放置标定物也限制了传统相机标定法的应用。

目前出现的自标定算法中主要是利用相机运动的约束。相机的运动约束条件太强，因此使得其在实际中并不实用。利用场景约束主要是利用场景中的一些平行或者正交的信息。其中空间平行线在相机图像平面上的交点被称为消失点，它是射影几何中一个非常重要的特征，所以很多学者研究了基于消失点的相机自标定方法。自标定方法灵活性强，可对相机进行在线定标。但由于它是基于绝对二次曲线或曲面的方法，其算法鲁棒性差。

基于主动视觉的相机标定法是指已知相机的某些运动信息对相机进行标定。该方法不需要标定物，但需要控制相机做某些特殊运动，利用这种运动的特殊性可以计算出相机内部参数。基于主动视觉的相机标定法的优点是算法简单，往往能够获得线性解，故鲁棒性较高，缺点是系统的成本高、实验设备昂贵、实验条件要求高，而且不适合于运动参数未知或无法控制的场合。

## 1. 张氏标定法<sup>[1]</sup>

### 1.1 原理简介

张氏标定法是张正友博士在 1999 年发表在国际顶级会议 ICCV 上的论文《Flexible Camera Calibration By Viewing a Plane From Unknown Orientations》中，提出的一种利用平面棋盘格进行相机标定的实用方法。该方法介于摄影标定法和自标定法之间，既克服了摄影标定法需要的高精度三维标定物的缺点，又解决了自标定法鲁棒性差的难题。标定过程仅需使用一个打印出来的棋盘格，并从不同方向拍摄几组图片即可，任何人都可以自己制作标定图案，不仅实用灵活方便，而且精度很高，鲁棒性好。因此很快被全世界广泛采用，极大的促进了三维计算机视觉从实验室走向真实世界的进程。

相机标定的目的是：建立相机成像几何模型并矫正透镜畸变。建立相机成像几何模型：计算机视觉的首要任务就是要通过拍摄到的图像信息获取到物体在真实三维世界里相对应的信息，于是，建立物体从三维世界映射到相机成像平面这一过程中的几何模型就显得尤为重要，而这一过程最关键的部分就是要得到相机的内参和外参。在进行这一步的介绍之前，我们要明确几个重要的概念：

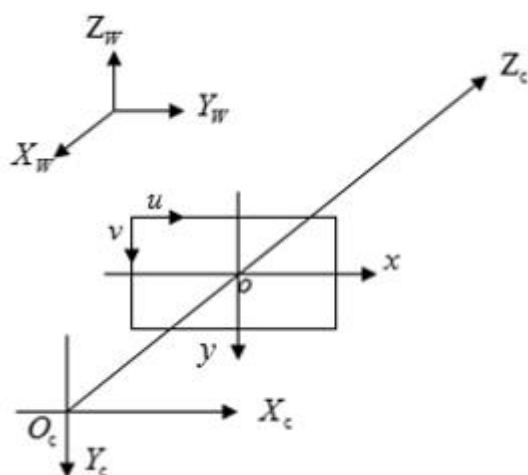
**世界坐标系(world coordinate system)：**用户定义的三维世界的坐标系，为了描述目标物在真实世界里的位置而被引入。单位为 m。

**相机坐标系(camera coordinate system)：**在相机上建立的坐标系，为了从相机的角度描述物体位置而定义，作为沟通世界坐标系和图像/像素坐标系的中间一环。单位为 m。

**图像坐标系(image coordinate system)：**为了描述成像过程中物体从相机坐标系到图像坐标系的投影透射关系而引入，方便进一步得到像素坐标系下的坐标。单位为 m。

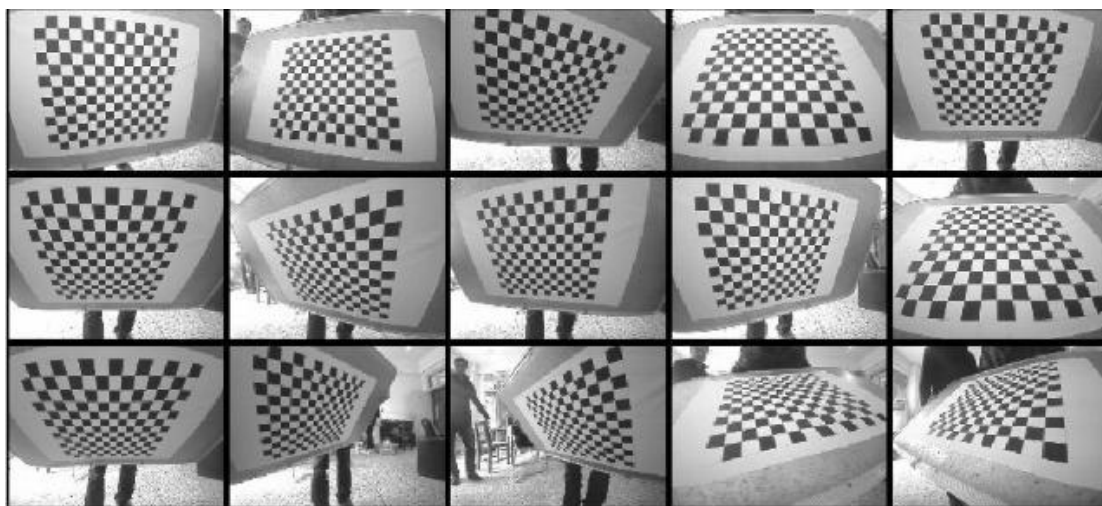
**像素坐标系(pixel coordinate system)：**为了描述物体成像后的像点在数字图像上（相片）的坐标而引入，是我们真正从相机内读取到的信息所在的坐标系。单位为个（像素数目）。

这四个坐标系之间的关系如下图所示：



其中世界坐标系： $X_w$ 、 $Y_w$ 、 $Z_w$ 。相机坐标系： $X_c$ 、 $Y_c$ 、 $Z_c$ 。图像坐标系： $x$ 、 $y$ 。像素坐标系： $u$ 、 $v$ 。相机坐标系的  $z$  轴与光轴重合，且垂直于图像坐标系平面并通过图像坐标系的原点，相机坐标系与图像坐标系之间的距离为焦距  $f$ （也即图像坐标系原点与焦点重合）。像素坐标系平面  $u-v$  和图像坐标系平面  $x-y$  重合，但像素坐标系原点位于图中左上角（之所以这么定义，目的是从存储信息的首地址开始读写）。

在这里我们先引入“棋盘”的概念：棋盘是一块由黑白方块间隔组成的标定板，我们用它来作为相机标定的标定物（从真实世界映射到数字图像内的对象）。之所以我们用棋盘作为标定物是因为平面棋盘模式更容易处理（相对于复杂的三维物体），但与此同时，二维物体相对于三维物体会缺少一部分信息，于是我们会多次改变棋盘的方位来捕捉图像，以求获得更丰富的坐标信息。如下图所示，是相机在不同方向下拍摄的同一个棋盘图像。



下面将依次对刚体进行一系列变换，使之从世界坐标系进行仿射变换、投影透射，最终得到像素坐标系下的离散图像点，过程中会逐步引入各参数矩阵。

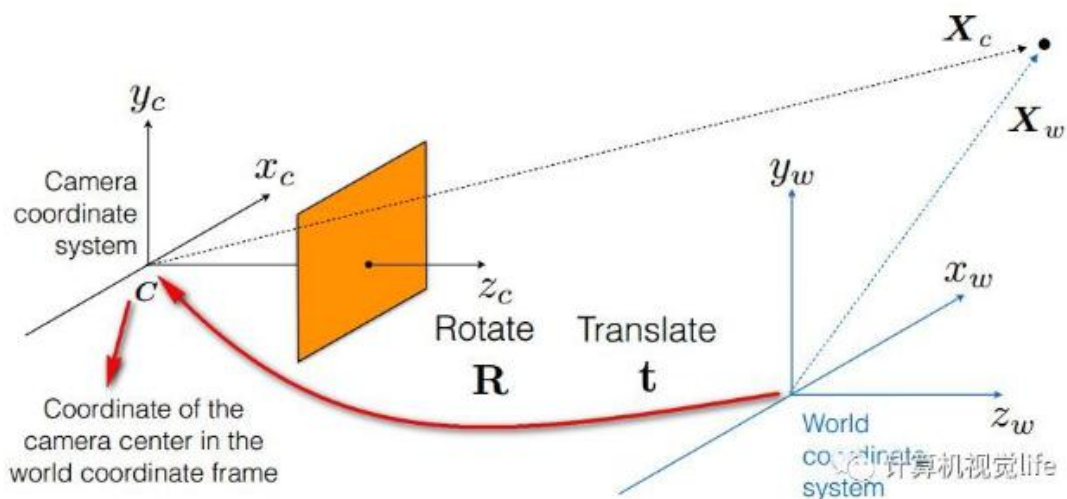
### 1.1.1 从世界坐标系到相机坐标系

刚体从世界坐标系转换到相机坐标系的过程，可以通过旋转和平移来得到，我们将其变换矩阵由一个旋转矩阵和平移向量组合成的齐次坐标矩阵来表示：

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0_3^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & r_3 & t \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

其中， $R$  为旋转矩阵， $t$  为平移向量，因为假定在世界坐标系中物点所在平面过世界坐标系原点且与  $Z_w$  轴垂直，所以  $z_w=0$ ，可直接转换成式 1 的形式。其中变换矩阵即为前文提到的外参矩阵，之所称之为外参矩阵可以理解为只与相机外部参数有关，且外参矩阵随刚体位置的变化而变化。

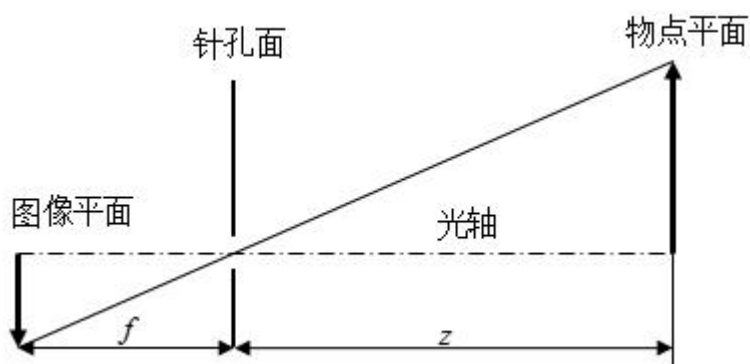
下图表示了用  $R$ ， $t$  将上述世界坐标系转换到相机坐标系的过程。



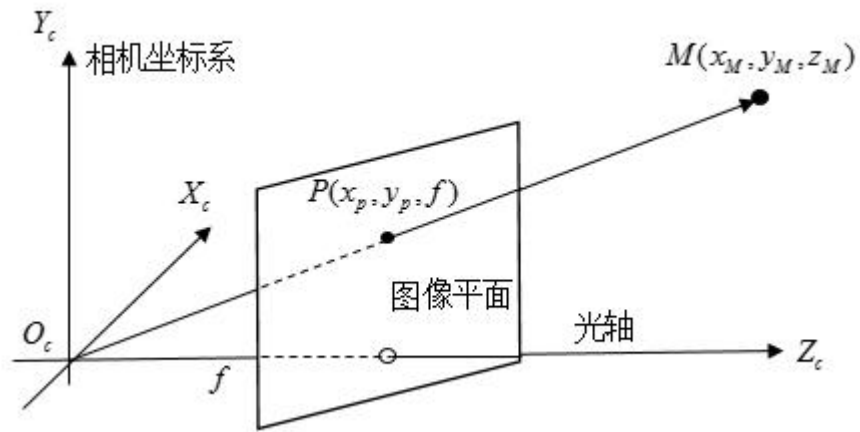
### 1.1.2 从相机坐标系到理想图像坐标系（不考虑畸变）

这一过程进行了从三维坐标到二维坐标的转换，也即投影透视过程。我们还是拿针孔成像来说明。

成像过程如图二所示：针孔面（相机坐标系）在图像平面（图像坐标系）和物点平面（棋盘平面）之间，所成图像为倒立实像。



但是为了在数学上更方便描述，我们将相机坐标系和图像坐标系位置对调，变成图三所示的布置方式：



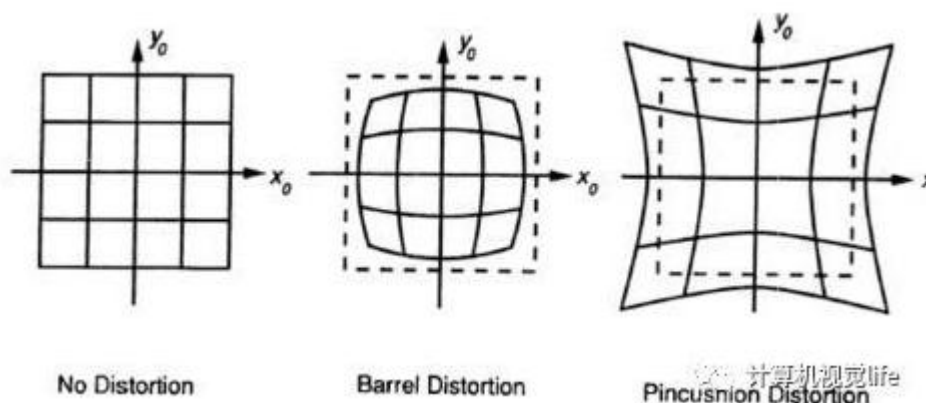
此时，假设相机坐标系中有一点  $M$ ，则在理想图像坐标系下（无畸变）的成像点  $P$  的齐次坐标表示形式为：

$$z_M \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_M \\ y_M \\ z_M \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_M \\ y_M \\ z_M \\ 1 \end{bmatrix}$$

### 1.1.3 从理想图像坐标系到实际图像坐标系（考虑畸变）

透镜的畸变主要分为径向畸变和切向畸变，还有薄透镜畸变等等，但都没有径向和切向畸变影响显著，所以我们在这里只考虑径向和切向畸变。

径向畸变是由于透镜形状的制造工艺导致。且越向透镜边缘移动径向畸变越严重。下图所示是径向畸变的两种类型：桶形畸变和枕形畸变。



实际情况中我们常用  $r=0$  处的泰勒级数展开的前几项来近似描述径向畸变。

矫正径向畸变前后的坐标关系为：

$$\begin{cases} x_{corr} = x_p(1 + k_1r^2 + k_2r^4 + k_3r^6) \\ y_{corr} = y_p(1 + k_1r^2 + k_2r^4 + k_3r^6) \end{cases}$$

由此可知对于径向畸变，我们有 3 个畸变参数需要求解。

切向畸变是由于透镜和 CMOS 或者 CCD 的安装位置误差导致。因此，如果存在切向畸变，一个矩形被投影到成像平面上时，很可能会变成一个梯形。切向畸变需要两个额外的畸变参数来描述，矫正前后的坐标关系为：

$$\begin{cases} x_{corr} = x_p + [2p_1x_py_p + p_2(r^2 + 2x_p^2)] \\ y_{corr} = y_p + [p_1(r^2 + 2y_p^2) + 2p_2x_px_p] \end{cases}$$

由此可知对于切向畸变，我们有 2 个畸变参数需要求解。

综上，我们一共需要 5 个畸变参数（ $k_1$ 、 $k_2$ 、 $k_3$ 、 $p_1$  和  $p_2$ ）来描述透镜畸变。

#### 1.1.4 从实际图像坐标系到像素坐标系

由于定义的像素坐标系原点与图像坐标系原点不重合，假设像素坐标系原点在图像坐标系下的坐标为  $(u_0, v_0)$ ，每个像素点在图像坐标系  $x$  轴、 $y$  轴方向的尺寸为： $dx$ 、 $dy$ ，且像点在实际图像坐标系下的坐标为  $(xc, yc)$ ，于是可得



到像点在像素坐标系下的坐标为：

$$u = \frac{x_c}{d_x} + u_0, v = \frac{y_c}{d_y} + v_0$$

化为齐次坐标表示形式可得：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix}$$

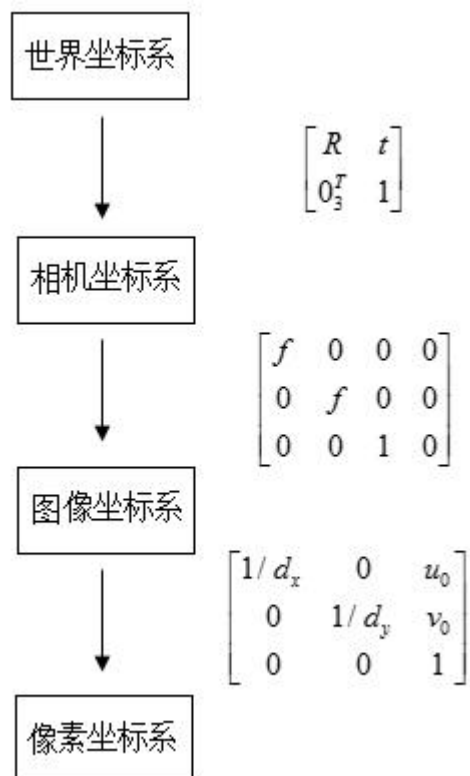
公式 2 中 $(x_p, y_p)$ 与公式 5 中 $(x_c, y_c)$ 相同，都是图像坐标系下的坐标。

若暂不考虑透镜畸变，则将式 2 与式 5 的转换矩阵相乘即为内参矩阵  $M$ ：

$$M = \begin{bmatrix} f/d_x & 0 & u_0 \\ 0 & f/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

之所以称之为内参矩阵可以理解为矩阵内各值只与相机内部参数有关，且不随物体位置变化而变化。

最后用一幅图来总结从世界坐标系到像素坐标系（不考虑畸变）的转换关系：



有了坐标系的转换关系，我们便可以学习出我们所需要的相机参数。

#### 问题 1: 学习外在参数

我们的目标是对一个与给定场景相关的摄像机的位置 and 方向进行复原。有时这被称为多点透视(Perspective n Point, PnP)问题或者外方位问题。一个通用应用是增强现实，在增强现实中我们需要知道这种关系来渲染出现在现实场景中稳定部分的虚拟对象。

该问题可以以一种更为形式化的方式加以描述: 假设一个已知对象有  $n$  个不同的 3D 点，它们在图像中的对应投影点，以及已知的内在参数。我们的目标是对旋转矩阵和平移矢量进行估计，这两者可以将目标坐标系中的点映射为摄像机坐标系中的点，这是最大似然学习问题。

#### 问题 2: 学习内在参数

我们的目标是对内在参数  $A$  进行估计， $A$  将通过光心的光线方向与图像平面

上的坐标相关联。这一估计过程称为校准。如果我们希望使用摄像机构建现实世界的 3D 模型，内在参数知识非常重要。

校准问题可以以一种更形式化的方式加以描述:给定一个已知 3D 对象，该对象有 1 个不同的 3D 点及其在图像中的对应投影点，对内在参数估计，这也是一个最大似然学习问题。

## 1.2 实验部分

本次实验部分并不是简单的复现张氏标定法的过程，而是将张氏标定法应用于估计相机的位姿上。实验估计两个不同相机（视角）的位姿，方法是用两个视角的相机拍摄同一个棋盘格若干次，每次拍摄都分别应用张氏标定法计算每一次拍摄的内外参，并根据两个视角标定的外参转换为两个相机之间的旋转矩阵  $R$  和平移矢量  $T$ 。

### 1.2.1 标定物选取

本次实验的标定物可以选择普通的棋盘格，棋盘格的尺寸要尽量大一些，使得两个相机拍摄的棋盘格清晰易分辨，能够准确检测到棋盘格角点。本次实验我们选用的标定物如下图所示：

针对多视角相机的外参标定问题，我们采用一种先粗标定再优化的方式，有效解决二维棋盘格标定物在多相机系统中存在的自遮挡问题。在不需要三维标定物的情况下，利用多相机之间的拓扑关系同步优化所有相机外参，可以得到精度可观的标定结果。

为了实现上述目的，整个实验由三个主要阶段组成。

**第一阶段为相机内参预标定阶段。**这一阶段的目的是对相机的内部参数进行标定，内参主要包括相机畸变、焦距、主点位置等。标定结果为相机的径向畸变参数 $(k_1, k_2)$ 及内参矩阵 $K = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}$ ，这一阶段得到的参数将参与之后的

外参标定阶段的投影计算过程，作为常量使用。

**第二阶段为相机外参粗标定阶段。**这一阶段的目的是对所有相机在世界坐标系下的外部参数进行粗略的标定，以作为下一个阶段中非线性优化过程的初始参数。在这一阶段，对于所有相机围成一圈，焦点在同一块中心区域的多相机系统，选定其中一个相机的相机坐标系作为世界坐标系。根据相机的相邻关系，每两个相邻相机之间标定外参 $(R_{ij}, t_{ij})$ ，标定方法为：

相机 1 和相机 2 同时拍摄同一张棋盘格标定物，则可得到两个相机相对于棋盘格坐标系的外参 $(R_{p1}, t_{p1})$ 、 $(R_{p2}, t_{p2})$ ，相机 2 相对于相机 1 的外参矩阵为：

$$RT_{12} = \begin{pmatrix} R_{p1} & t_{p1} \\ 0 & 1 \end{pmatrix}^{-1} * \begin{pmatrix} R_{p2} & t_{p2} \\ 0 & 1 \end{pmatrix},$$

这里默认相邻两个相机在至少一种棋盘格姿态下一定能同时检测到棋盘格角点。对于两个不相邻的相机，要求它们之间的外参则存在两种情况：一是这两个

相机的视野中能同时看到同一姿态的棋盘格，这种情况下则不需要利用中间相机作为媒介传递外参而可以直接计算这两个相机的外参；二是这两个相机的视野中不能同时看到同一姿态的棋盘格，这种情况下则需要利用到相机的传递关系。若相机 1 与相机 2 相邻，相机 2 与相机 3 相邻，则已知相机 2 相对于相机 1 的外参  $(R_{12}, t_{12})$ ，相机 3 相对于相机 2 的外参  $(R_{23}, t_{23})$ ，则相机 3 相对于相机 1 的外参矩阵为：

$$RT_{13} = \begin{pmatrix} R_{12} & t_{12} \\ 0 & 1 \end{pmatrix} * \begin{pmatrix} R_{23} & t_{23} \\ 0 & 1 \end{pmatrix}$$

利用这种传递关系便可得到所有相机相对于世界坐标系的外参  $(R_{iw}, t_{iw})$ 。

**第三阶段为外参的非线性优化阶段。**这一阶段的目的是对第二阶段获得的初始外参进一步精细化。对于新拍摄的一组标定图像，检测图像上的角点。若一个棋盘格姿态在某些相机上能检测到角点，则根据相机外参矩阵可以将其投影到世界坐标系下。

若一个棋盘格姿态在参照相机拍摄的图像上即可检测到角点，则可以直接使用该角点在世界坐标系下的坐标。

利用上述方法获得棋盘格所有姿态下的角点世界坐标，下面只需对每一个棋盘格姿态利用第一阶段标定获得的相机内参及第二阶段标定获得的相机外参重投影回每个相机拍摄的对应的图像上，这样即可获得重投影点  $\hat{m}(K_i, R_i, t_i, M_j)$ 。假设我们有棋盘格模型的  $n$  个图像，模型平面上有  $m$  个点，且图像上的噪声服从独立同一分布，则极大似然估计可以通过最小化以下函数得到：

$$\sum_{i=1}^n \sum_{j=1}^m \left\| m_{ij} - \hat{m}(K_i, R_i, t_i, k_{i1}, k_{i2}, M_j) \right\|^2$$

可以使用 Levenberg-Marquardt 算法进行迭代来求最优解，这个算法需要一个初始值，初始值从第二阶段获取。

### **具体实验步骤：**

实验设备为八个同一型号的工业相机按编号围成一圈，焦点在同一块中心区域。

#### **步骤 1) 内参预标定**

对八个相机分别预标定内部参数。具体做法为每一个相机分别拍摄 10 组左右不同姿态棋盘格焦点图片，放入 matlab 内置工具箱 Camera Calibrator 进行标定，标定结果存储为 matlab 表格，作为常量使用。

#### **步骤 2) 外参标定数据采集**

这一步需要采集两次数据，第一次用于相机外参粗标定，以两个相邻相机为单位，在两者视线之内拍摄 10 组以上不同姿态棋盘格图像；第二次用于外参优化过程，采集时需要所有相机同步拍摄十组以上不同姿态棋盘格角点，确保每个相机至少在一组姿态中能检测到角点，最好将棋盘格绕中心轴匀速旋转一圈，所有相机同步连续采集，采集的图像数据以 PositionXX\_CameraXX 格式命名。

#### **步骤 3) 相机外参粗标定**

选定相机 0 的相机坐标系作为整个系统的世界坐标系，对第一次采集的数据，

利用每一对相机之间 Position 相同的图像标定这一对相机之间的外参。求得每两个相邻相机的外参之后，利用相机相邻关系求出所有相机在世界坐标系中的外参，这一步获得的外参作为优化初值。

#### 步骤 4) 角点重投影

对于第二次采集的数据，检测相机 0 拍摄的图片中是否能检测到角点并记录下来，若有，则保存对应的角点图像坐标坐标；若没有，则顺序检测相机 1 到相机 7 对应姿态拍摄的图像中是否角点，若检测到角点，则通过该相机与相机 0 之间的外参及两个相机的内参将角点重投影到相机 0 所在的图像上，并停止检测剩余的相机。这一步可以确保相机 0 的所有图像上均有焦点存在。利用所有相机的内外参将相机 0 图像上的角点重投影到每个相机相对应的图像上，获得所有相机的重投影点。

#### 步骤 5) 外参优化

通过第四步获得所有标定图像上的重投影点  $m(K_i, R_i, t_i, k_{i1}, k_{i2}, M_j)$ ，检测相机 1 到相机 7 拍摄的所有图像中的角点，若检测不到，则跳过这一图像，若检测到角点  $m_{ij}$ ，则计算以下代价函数：

$$\delta = \sum_{i=1}^n \sum_{j=1}^m \left\| m_{ij} - \hat{m}(K_i, R_i, t_i, k_{i1}, k_{i2}, M_j) \right\|^2$$

通过使用 Levenberg-Marquardt 算法进行迭代最小化代价函数  $\delta$  来求  $(R_i, t_i)$  的最优解。

## 步骤 6) 标定结果验证

获得优化的外参之后，需要对结果的准确性进行验证，这里提供两种验证的方法：第一种方法是所有参与标定的相机再重新拍摄一组验证数据集，根据标定的外参按照步骤 4 相同的逻辑重投影角点，定量计算重投影角点与检测角点的误差；第二种方法是利用虚拟相机仿真系统（已完成），通过在这个仿真系统中手动设置相机内外参及拍摄“虚拟”标定数据，放入标定算法中标定外参，通过比较标定外参与设置的外参之间的误差定量衡量标定结果的准确性。