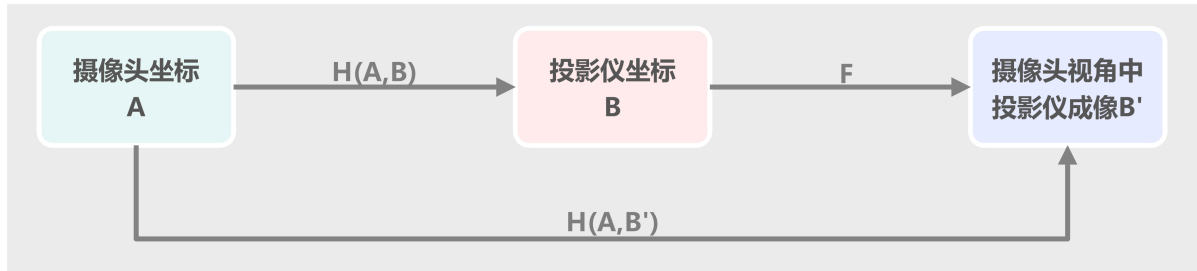


1. 自动校准第一阶段——对齐摄像头坐标与投影成像

摄像头坐标 A \times 当前单应性矩阵 $H(A, B)$ = 投影仪坐标 B

投影仪坐标 B 和摄像头中看到的成像 B' 有固定的设备偏差 F , $F(B) = B'$



我们的目标：

获取到摄像头坐标 A 和摄像头中投影成像 B' ，计算出一个合适的单应性矩阵 $H(A, B)$ 使得摄像头坐标 A 和投影成像 B' 重合。我们将目标情况下的单应性矩阵均一写为 H' ，即：

$$F(A \cdot H'(A, B)) = B' = A$$

我们把固定设备偏差 F 的具体作用形式，假设为一个 B 到 B' 的单应性矩阵 $H_F(B, B')$ 。那么我们通过一组 (B, B') 的数据对即可计算矩阵 H_F 。至此：

$$A \cdot H(A, B) \cdot H_F(B, B') = B' = A \cdot H(A, B')$$

在目标情况下 $A = B'$ 时， $H'(A, B')$ 为单位矩阵 I ：

$$A \cdot H'(A, B) \cdot H_F(B, B') = A = B' = A \cdot H'(A, B')$$

$$H'(A, B) \cdot H_F(B, B') = H'(A, B') = I$$

$$H'(A, B) = H_F^{-1}$$

其中 H_F 为假设的设备偏差矩阵，在任何情况下都固定不变。已知目标情况下的 $H'(A, B')$ 和 $H_F(B, B')$ 即可求得目标情况下的 $H'(A, B)$ 。

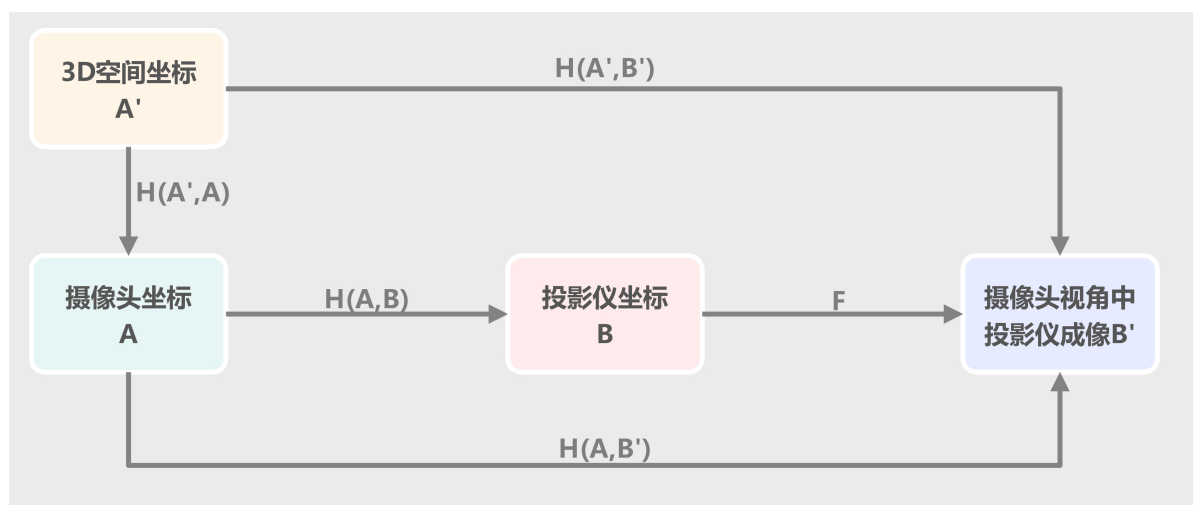
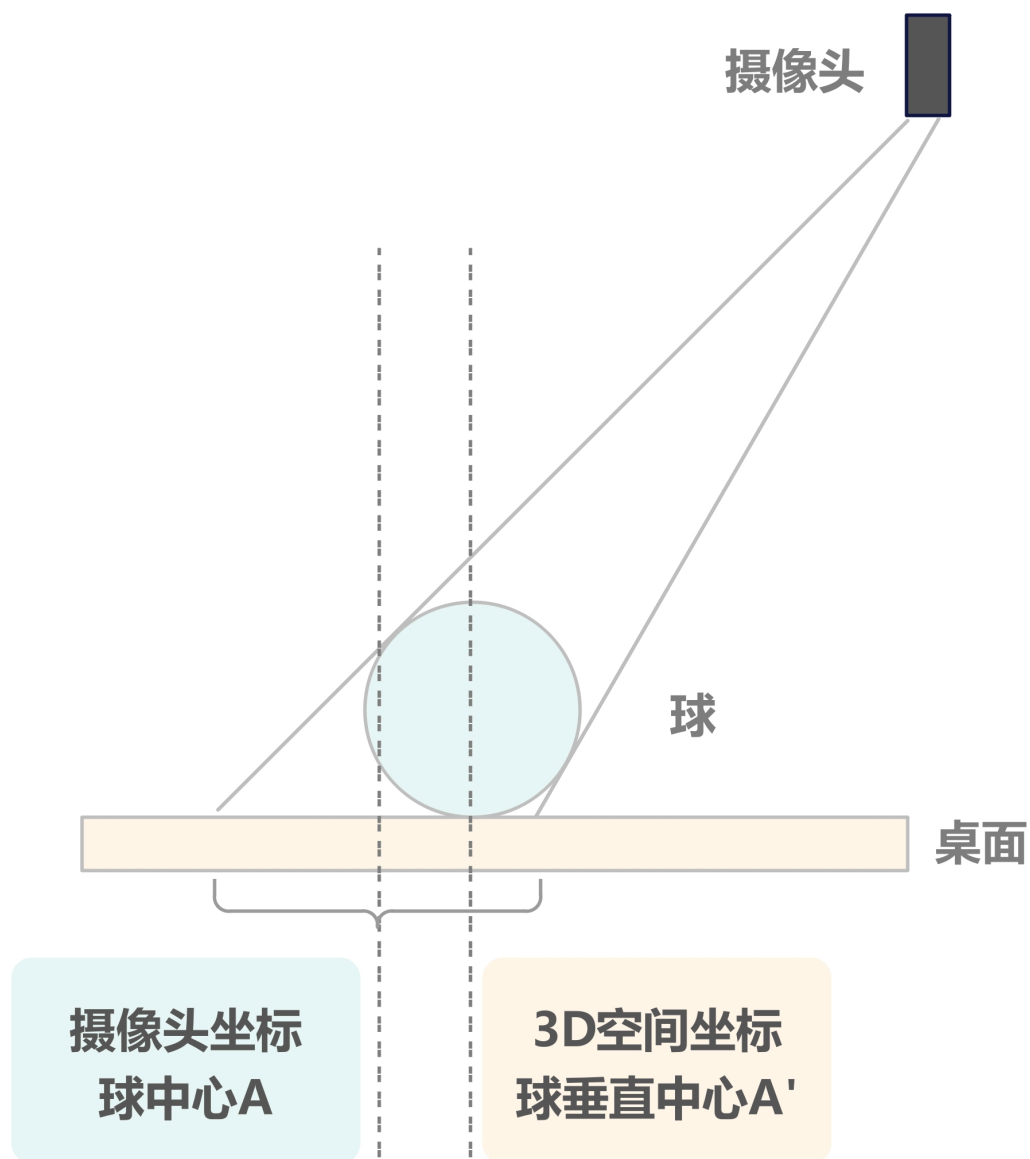
我们可以把流程划分为几个步骤：

1. 根据 B, B' 求设备偏差矩阵 H_F
2. 根据 H_F 求得 $H'(A, B)$

2. 自动校准第二阶段——解决摄像头视角与桌面坐标中球体中心的偏差

第一阶段的校准只能保证在摄像头视角中无偏差，然而投影成像在人眼中依然和球体坐标有偏差，是因为人眼观察到的真实的球体位置和摄像头观察到的球体位置不一致。

如图：左侧虚线为摄像头观察到的球体坐标 A ；右侧虚线为球体质心垂直于桌面的投影，也是球体真实坐标 A' 。



如果我们能够同时获取到 A, A', B, B' ，则可以获取到当前状态下四种坐标之间的映射关系 $H(A', A), H(A, B), F$ 和 $H(A', B')$ 。如果沿用第一阶段的假设， F 为线性且用 $H_F(B, B')$ 表示，则有：

$$A' \cdot H(A', A) \cdot H(A, B) \cdot H_F(B, B') = B' = A' \cdot H(A', B')$$

在目标情况下 $A' = B'$ 时，映射关系 $H(A', A)$ 与 $H_F(B, B')$ 为固定偏差不变， $H(A, B)$ 与 $H(A', B')$ 为目标情况下的新映射关系：

$$A' \cdot H(A', A) \cdot H'(A, B) \cdot H_F(B, B') = A' = A' \cdot H'(A', B')$$

$$H(A', A) \cdot H'(A, B) \cdot H_F(B, B') = I$$

$$H'(A, B) = H(A', A)^{-1} \cdot H_F(B, B')^{-1}$$

则可以通过计算 $H'(A', B')$ 求得最终应用的单应性矩阵 $H'(A, B)$ 。

我们在第一阶段中已经获取到 A, B, B' 和矩阵 $H(A, B), H_F$ ，如何精确获取到球体的3D空间坐标 A' 是第二阶段需要解决的问题。

2.1 直接获取3D空间坐标 A'

对于如何获取3D空间坐标 A' ，我们可以把流程划分为几个步骤：

1. 对环境（球桌与球）进行三维重建。

纯视觉方案：手持/固定摄像头环绕拍摄+3D高斯splatting算法

激光雷达方案：球台正上方固定的精度较高的激光雷达扫描

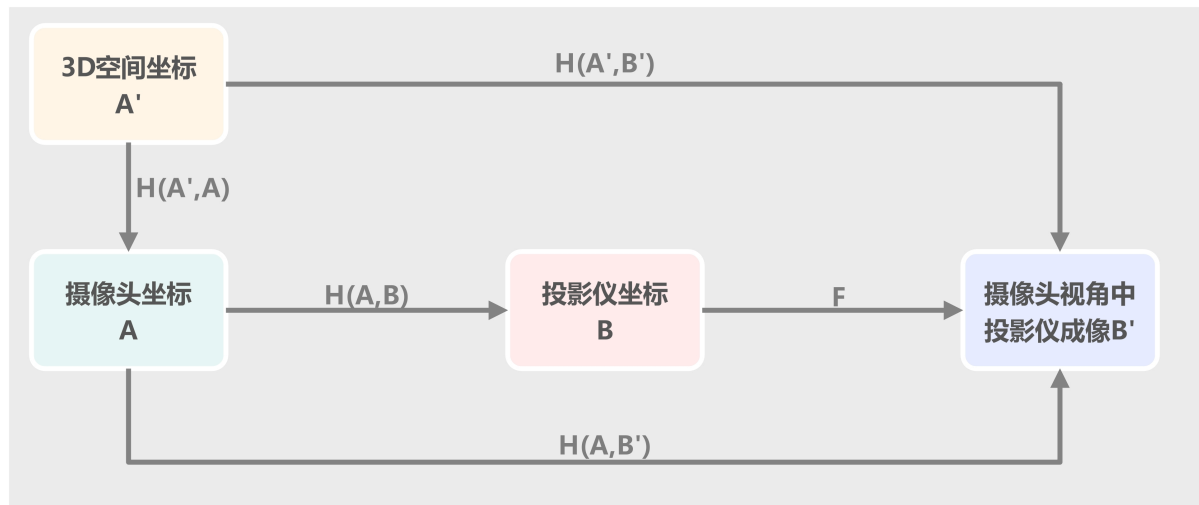
2. 从重建场景中分离出台球，计算台球质心；从重建场景中提取台球桌平面。

台球实例提取：3D检测 或 点云聚类+球面几何分析

球桌平面提取：

- 1)点云滤波，处理表面不光滑问题
- 2)RANSAC 平面拟合，从台球桌点云中提取平面
- 3)多次采样平面修正，对平面点集进行局部采样，求平均平面

3. 根据台球质心求其垂直于台球桌平面的投影，即3D空间坐标系中的 A' 。



2.2 近似拟合 3D空间坐标 A' 与 摄像头坐标 A 之间的关系 $H(A', A)$

获取3D空间坐标 A' 的成本过于高昂，我们希望能够直接使用深度学习模型，通过捕捉受到 $H(A', A)$ 影响的特征，间接估计出 $H(A', A)$ 。

事实上 3D空间坐标 A' 与 摄像头坐标 A 之间的关系取决于摄像头与桌台之间的相对位置关系（高度和水平位置关系）。我们可以通过一些恒定标定物在摄像头视角中的特征变化，来间接估计出摄像头与桌台之间的相对位置关系，从而间接估计出 摄像头坐标下 A 与 3D空间坐标下 A' 的映射关系。

例如固定尺寸的台球，在摄像头视角中的大小与位置反应出摄像头与台球之间的位置关系，间接反映了摄像头坐标下 A 与 3D 空间坐标下 A' 的映射关系。

假设我们能够成功训练一个模型，其输入摄像头视角下球的特征 $A_{feature}$ 输出映射关系 $H(A', A)$ 。

$$H(A', A) = Model(A_{feature})$$

在得到 $H(A', A)$ 后，第一阶段求解可知 $A, H(A, B), B, H_F(B, B'), H(A, B')$ ，可计算出 $H(A', B')$ 和 A' ：

$$H(A', A) \cdot H(A, B') = H(A', B')$$

$$A' = A \cdot H^{-1}(A', A) = B' \cdot H^{-1}(A', B')$$

则可以推断出 $A' = B'$ 时目标情况下的 $H'(A', B')$ ，存在：

$$H'(A', B') = H(A', A) \cdot H'(A, B) \cdot H_F(B, B')$$

通过计算可以求解目标情况下的 $H'(A, B)$ 。其中 $H(A', A)$ 和 $H_F(B, B')$ 不随着 $H(A, B)$ 的改变而改变，在每张桌台中认为是常量。

如果将这个模型拓展为端到端的模型，即输入目标情况下 ($A' = B'$ 时) 的 $A_{feature}$ ，同时输入第一阶段的变量 A, B, B' 模型可以输出最终我们想要的 $H'(A, B)$ 。

$$H'(A, B) = Model(A_{feature}, A, B, B')$$

然而在实践中，为了保证模型收敛，能够直接求解的变量尽量直接求解，我们需要尽可能简化深度学习模型。最有可能的做法是，获取手动校准好 (目标情况 $A' = B'$) 的 $(A_{feature}, H(A', A))$ 数据对，以球的特征 $A_{feature}$ 为输入，以映射矩阵 $H(A', A)$ 为输出来训练模型。

目标情况下 $H(A', A)$ 可以由公式得到，其中 $H'(A, B)$ 可以直接获取， $H_F(B, B')$ 可以由 B, B' 计算得到：

$$H(A', A) = H'(A, B)^{-1} \cdot H_F(B, B')^{-1}$$

以此可以获取 $(A_{feature}, H(A', A))$ 训练数据对。