

# SfM中文流程

结构从运动 (Structure from Motion, SfM) 是一种从多视图图像中恢复三维场景结构和相机姿态的方法。以下是 SfM 的具体流程：

## 1. 图像预处理

在开始重建之前，通常会进行图像的预处理，包括调整图像的亮度、对比度、去噪等，以提高特征提取的效果。

## 2. 特征检测与描述

从每张图像中提取稳定的特征点，以便在不同视角下找到匹配。常用的特征检测和描述算法包括：

- SIFT (Scale-Invariant Feature Transform)
- ORB (Oriented FAST and Rotated BRIEF)
- SURF (Speeded-Up Robust Features)

这些算法可以提取具有尺度和旋转不变性的特征点，使得即使在不同视角或光照条件下拍摄的图像中，也能找到相同的点。

## 3. 特征匹配

在所有图像对之间进行特征匹配。常用的方法有：

- 直接计算特征向量之间的欧氏距离，匹配距离最小的点对。
- 使用最近邻匹配（例如最近邻与次近邻的比值）来提高匹配的准确性。
- 使用 RANSAC（随机抽样一致性）算法剔除错误匹配的点对。

**结果：**得到每对图像之间的特征匹配关系（包括正确匹配的特征点对）。

## 4. 初始相机姿态估计和三维点计算

选择一对视差较大的图像作为初始图像对，估计它们之间的相机姿态并进行初步三维点云重建。具体方法如下：

1. **基础矩阵/本质矩阵计算：**通过匹配点计算基础矩阵（ $F$ ）或本质矩阵（ $E$ ）。
2. **相机姿态分解：**将本质矩阵分解为旋转矩阵和平移向量。
3. **三角化初步三维点：**利用两个相机的姿态，通过三角化方法计算出匹配点的三维坐标。

## 5. 增量式重建

在初始的两视图重建基础上，逐步引入新的图像并进行增量重建：

1. **相机姿态估计**：对于每一个新的图像，根据已重建的三维点和相机姿态，使用 PnP (Perspective-n-Point) 方法估计新的相机姿态。
2. **三角化新的三维点**：利用新相机与已知相机的视角差异，继续通过三角化增加新的三维点。
3. **重复上述步骤**，直到所有图像的相机姿态都被估计，并且所有特征点的三维位置都被恢复。

## 6. 捆绑调整 (Bundle Adjustment)

捆绑调整是一种全局优化方法，用于在增量式重建后对所有相机姿态和三维点位置进行优化，以最小化重投影误差。优化目标函数为每个三维点的重投影误差：

$$\min \sum_{i,j} \|\text{proj}(P_i, X_j) - x_{i,j}\|^2 \quad (1)$$

其中  $P_i$  是相机的外参矩阵， $X_j$  是三维点， $x_{i,j}$  是图像中三维点  $X_j$  在相机  $i$  上的重投影坐标， $\text{proj}$  表示投影函数。

捆绑调整会进一步提高整个重建的精度，使得三维点云更加准确。

## 7. 稀疏点云生成

在捆绑调整完成后，所有的三维点和相机姿态已经得到优化，形成稀疏的三维点云。该点云包含场景中的关键点位置，并可以大致反映场景的形状结构。

## 8. （可选）尺度和坐标系调整

如果有真实世界的尺度信息（例如 GPS 数据），可以对点云进行尺度和坐标系的调整，以获得更真实的三维重建。否则，SfM 的点云通常是无尺度的，即不知道绝对的大小单位。

## 9. 输出结果

SfM 的输出通常包括以下几部分：

- 稀疏点云：由三维关键点组成，反映场景的基本结构。
- 相机姿态：每张图像对应的相机位置和方向。
- 内外参信息：包括相机的内参矩阵（如果有）和每张图像的外参（姿态）。

## 总结

1. 特征提取和匹配：找到图像间的对应点。
2. 初始相机姿态估计：选择初始图像对，估计相机姿态和三维点。

3. **增量式重建**：逐步添加图像，估计新图像的相机姿态和三维点。
4. **捆绑调整**：全局优化，最小化重投影误差。
5. **生成稀疏点云和相机姿态**：完成场景的稀疏三维重建。

SfM 通常生成稀疏点云，后续可以结合 **MVS**（多视图立体匹配）生成更密集的三维模型。