

## 1 HSfM: Hybrid Structure-from-Motion

### 1.1 问题：数据集的规模

数据集名称	数据集照片数量
Pantheon Int.	587
St. Peters Ext.	1155
Block	479
Block2	3482
Alamo	613
Piccadilly	2468
Vienna Cathedral	897
Trafalgar	5288

图 1: 主流数据集照片数量

## 2 3D Menagerie: Modeling the 3D Shape and Pose of Animals

### 2.1 问题：如何通过一张照片实现模型重建

作者通过使用 Artec 手持 3D 扫描仪扫描动物模型，创建动物的三维模型数据集。作者共进行了 41 次扫描，其中包括：1 只猫，5 只猎豹，8 只狮子，7 只老虎，2 只狗，1 只狐狸，1 只狼，1 只鬣狗，1 只鹿，1 匹马，6 只斑马，4 头母牛和 3 只河马。并且他们估计了一个缩放因子，使得不同的动物大小一致。作者还从 3D 人体数据集以及用图像创建的三维模型中汲取经验，使用一组 36 个手动点击的关键点用于辅助网格生成模型。

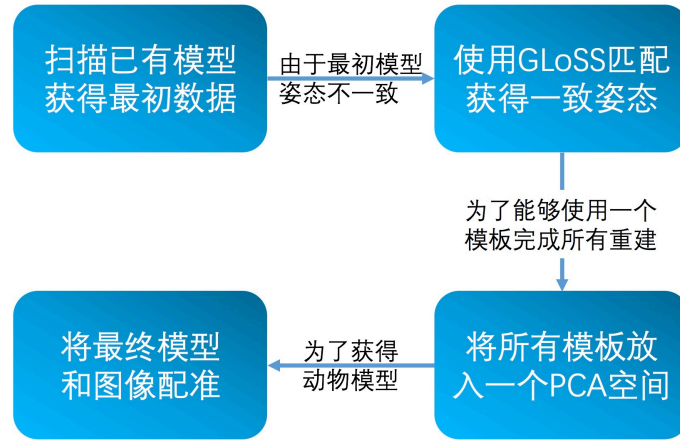


图 2: 三维模型生成流程图

然后使用 GLoSS 来实现模型的统一化。GLoSS 是一种 3D 关节模型生成方法，本体将每个部件进行了分割，并通过最小化部件接口处的缝合成本来组装每个部件。每个部分的形状变化是通过分析获得的，而不是通过学习得到的。这样的优点有两个：第一，这样的结果更接近于真实生活中的动物场景；第二，这允许我们将其应用于新的动物形状，而不需要先验的训练数据。

将模板初始化到扫描过程分两步进行。首先，作者使用基于梯度的方法优化 GLoSS 模型。使得 GLoSS 模型更靠近扫描模型。然后使用 ARAP 正则化对网格顶点进行无模型注册来以捕获细节。

得到一个标准模型以后，作者通过手动操作获取 19 个语义关键点加上尾尖的一点，将标准模型匹配到照片中。作者使用 TigDog 数据集中的图像和注释。对于没有注释的图像，作者手动点击所有动物相同的 20 个关键点，每个图像大约需要一分钟并且手动分割所有的图像。