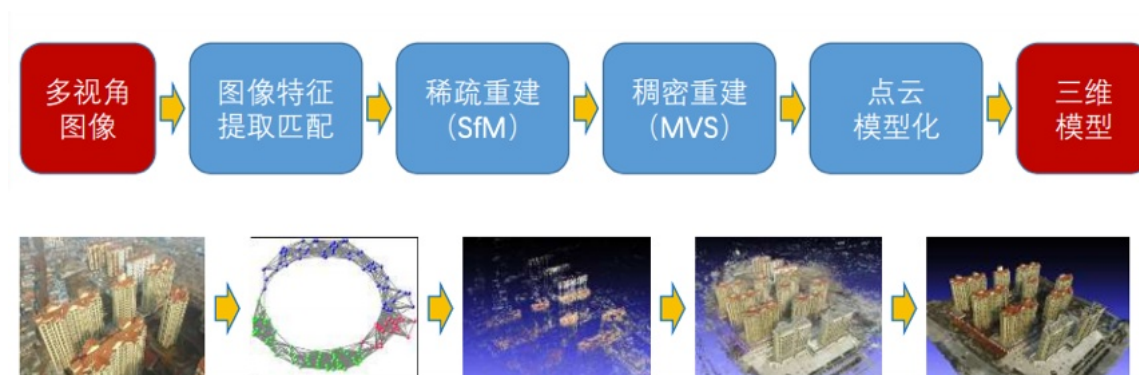


SfM核心的思想，以及最终目标，就是通过算法，对摄像机内参和外参的恢复

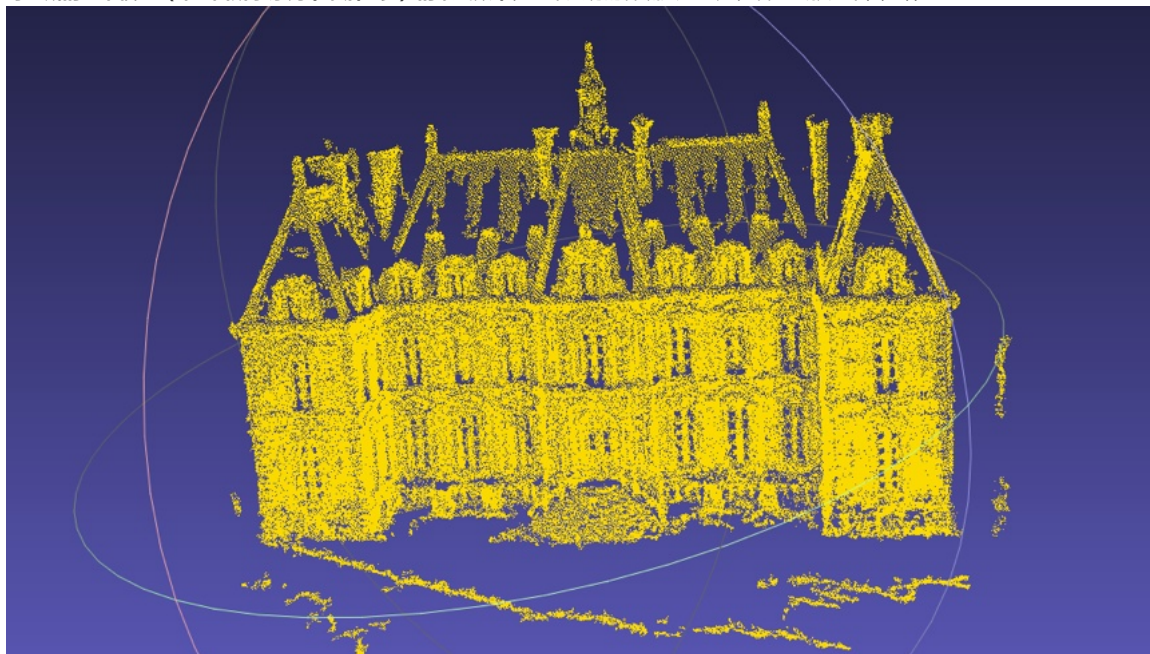
SfM综述



从上图我们可以直观感受一个完整的三维重建的过程。

众所周知，图像是由像素点组成的矩阵构成。因此，图像特征提取匹配，相当于SfM流程中的预处理，旨在从motion中，通过分析不同角度图片中所呈现的像素矩阵的差异，从而推导这一组二维图像反应的三维结构的空间位置信息。

所以SfM的核心思想，就是从原始的照片中，计算出涵盖三维位置信息的稀疏点云。这也是SfM的重要所在，因为重建出的位置关系的正确性，直接主导生成的三维模型（即上面所示序列中最后一步）的最终效果，一个成功的稀疏点云重建，看上去就如下图一样：



在这之后的稠密重建（MVS）和模型化过程，通俗地讲，就是在之前既有的三维点云的位置信息之上，添加我们人眼所能直观感受到的视觉信息，诸如颜色和轮廓。举个形象的例子，如果我们要开发一款游戏引擎，SfM就是模型框架的绘制过程，MVS就是在这之上的贴图部分。由此可以得出，一个好的三维重建有赖于SfM的准确性。

basic pipeline

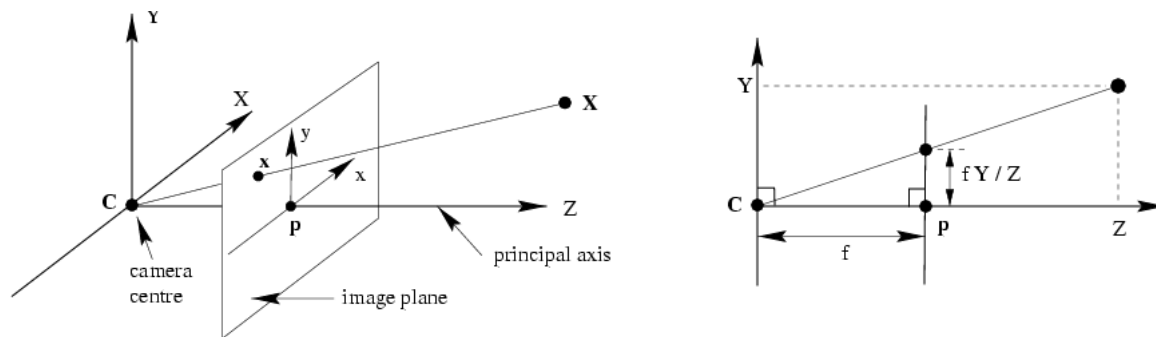
一个基本的SfM pipeline（基本流程）可以描述为：对每张图片检测特征点（feature point），对每对图片中的特征点进行匹配，只保留满足几何约束的匹配，最后执行一个迭代式的、鲁棒的SfM方法来恢复摄像机的内参（intrinsic parameter）和外参(extrinsic parameter)。

Intrinsic Parameter and Extrinsic Parameter

The intrinsic parameters, also known as internal parameters, are the parameters intrinsic to the camera itself, such as the focal length and lens distortion

按照定义，**Intrinsic Parameter(内参)**，就是相机本身固有的参数，例如焦距和镜头失真。**Extrinsic Parameter (外参)**，也称为外部参数或摄影机姿态，是用于描述摄影机与其外部世界之间的转换的参数、背景。

这涉及到摄像机的标定过程(Camera calibration),以经典的小孔成像原理图为例：



标定过程，简单来说是从世界坐标系换到图像坐标系的过程，也就是求最终的投影矩阵 (PP) 的过程

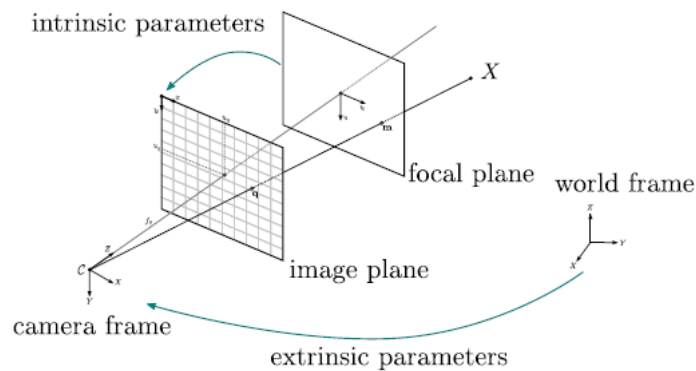
其中：

- CC 点表示camera centre，即相机的中心点，也是相机坐标系的中心点；
- ZZ 轴表示principal axis，即相机的主轴；
- pp 点所在的平面表示image plane，即相机的像平面，也就是图片坐标系所在的二维平面；
- pp 点表示principal point，即主点，主轴与像平面相交的点；
- CC 点到 pp 点的距离，也就是右边图中的 ff 表示focal length，即相机的焦距；

我们需要知道：

- 像平面上的 xx 和 yy 坐标轴是与相机坐标系上的 XX 和 YY 坐标轴互相平行的；
- 相机坐标系是以 XX, YY, ZZ (大写) 三个轴组成的且原点在 CC 点，度量值为米 (m) ；
- 像平面坐标系是以 xx, yy (小写) 两个轴组成的且原点在 pp 点，度量值为米 (m) ；
- 图像坐标系一般指图片相对坐标系，在这里可以认为和像平面坐标系在一个平面上，不过原点是在图片的角上，而且度量值为像素的个数 (pixel) ；

内参和外参的关系如下图：

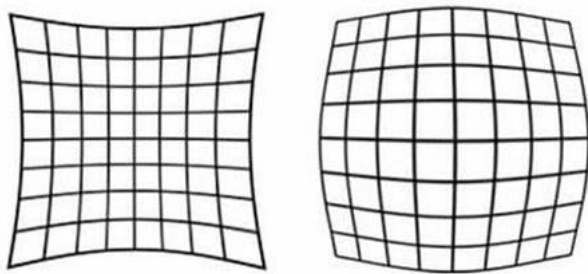


为了表达镜头焦距和失真状况，**相机内参**包括如下影响因素：

1. 焦距 (focal length)
2. 投影中心 (principle point)
3. 倾斜系数 (inclination factor)
4. 畸变系数 (distortion)

其中畸变系数，在x和y方向上需要分别讨论，反映了畸变图像x和y坐标轴的畸变情况

例如下图就是两种典型的畸变方式（径向畸变和切向畸变），如果你是一个摄影爱好者，那么在你的专业相机上也可以调整拍摄图像的畸变情况。



但数字图像是用矩阵的形式输入，因此我们可以通过矩阵的表达，综合考虑上述影响因素，用一个精简的数学公式表达内参

$$f_x = s_x * ff_y = s_y * f$$

所以，将上述关系定义一个内参矩阵即为

$$\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

用python代码表达内参可以写作

```
f_x=s_x*f
f_y=s_y*f
```

其中：

- f 是相机的焦距
上面矩阵中的分别用像素宽度和高度表示此值。
- 变量 s_x 和 s_y 是比例因子（用于引入表达失真状况，简化需要考虑在内的因素）（scale factors, relating pixels to distance, often equal to 1），是像素大小（宽度或高度）与相机空间中使用的单位之间的比率。没有失真（位置变化）的时候，该参数通常近似于1。该参数在x和y轴上要分开讨论，例如，若像素宽度是相机空间x轴上使用的单位大小的一半，则 $s_x = 2$ 。

为了表达相机相对于外部世界的姿态，**相机外参**包括如下参数：

1. 旋转矩阵(rotation matrix)
2. 平移矩阵(Translation Matrix)

旋转矩阵和平移矩阵共同描述了如何把点从世界坐标系转换到摄像机坐标系，依然可以用公式表达如下：

$$[R|t] = [R| -RC]$$

其中R和t分表表示旋转矩阵和平移矩阵

C是一个位置因子，可以由R和t算出，计算如下：

$$C = -R^{-1}t = -R^T t$$

零空间点在世界坐标系下的坐标为

$$X = (x, y, z, 1)^T$$

相机坐标系为

$$X_c = (x_c, y_c, z_c, 1)^T$$

因此我们可以这样描述外参矩阵：

$$X_c = \begin{bmatrix} R & -R\tilde{C} \\ 0^T & 1 \end{bmatrix} X$$

籍此，可以定义一个标准的相机模型（camera models (intrinsic + extrinsic(pose))

```
// Setup a simple camera at origin
// camera P = K[R/t], t = -RC
Mat3 K;
K << 1000, 0, 500,
```

```
0, 1000, 500,  
0, 0, 1;  
PinholeCamera cam(K, Mat3::Identity(), Vec3::Zero());
```

reference

- [1] Seitz S M, Szeliski R, Snavely N. Photo Tourism: Exploring Photo Collections in 3D[J]. Acm Transactions on Graphics, 2006, 25(3):835-846.
- [2] Agarwal S, Snavely N, Simon I, et al. Building Rome in a day[J]. Communications of the Acm, 2011, 54(10):105-112.
- [3] Snavely K N. Scene reconstruction and visualization from internet photo collections[M]. University of Washington, 2008.
- [4] Frahm J M, Fite-Georgel P, Gallup D, et al. Building Rome on a cloudless day[C]