RANSAC算法

1. 算法原理：

**RANSAC**(**RA**ndom **SA**mple **C**onsensus,随机采样一致)算法是从一组有”内点”和“外点”(outliers)组成的数据中正确估计数学模型参数的迭代算法。算法的核心思想是：在迭代K次的情况下，得到模型的准确率为***P***。当迭代次数K发生变化时，模性型的准确率***P***也会跟着发生变化。

算法的具体步骤如下：

在执行算法之前需要确定的参数：

1. 总的样本个数M，正确的样本点个数N。
2. 第一次进行模型拟合时所需的最少的内点个数x。
3. 模型迭代的次数K
4. 在迭代K次时得到的模型的准确率***P***。
5. 在每次判断数据的其他点（建立模型时没有用到的数据点）是否为内点时所需要的拟合误差δ。
6. 需要注意的是，“内点”的概率 p=N/M 通常是一个先验值。然后***P***是我们希望RANSAC得到正确模型的概率。如果事先不知道 p 的值，可以使用自适应迭代次数的方法。也就是一开始设定一个无穷大的迭代次数，然后每次更新模型参数估计的时候，用当前的“内点”个数/总的样本数M当成 p，进一步 来估算出迭代次数K。

二． 算法推导：

明确一点：在进行迭代时，每次选取的X个点都为内点时，在误差拟合δ范围内，用所有的内点拟合出来的模型一定是正确的。

1. 从样本大小为M且含有N个“内点”的数据样本中，随机选取一个点为“内点”的概率为SP=N/M，则选取x个点都为内点的概率为SPx，选取的x个点中至少有一个不为内点的概率为（1-SPx），当进行K次迭代时，每一次都至少有一个点不为内点的概率为（1-SPX）K。所以，K次迭代时，每次选取的x个点都为内点的概率为1-（1-SPX）K=***P***。所以，在选用X个内点迭代K次的情况下，得到的模型正确的概率是***P***。
2. 当不知道迭代次数时，可以通过***P***来求解K。：

1-P=（1-SPX）K。

=K ln（1-SPX）  
K= / ln（1-SPX）

1. 算法步骤：

①跟据样本的分布情况，先假设一个数据模型model，从数据样本中选取最少（对于直线拟合来说就是两个点，对于计算Homography矩阵就是4个点）的可以拟合出模型的x个内点，组成内点集合进行模型拟合。得到模型model的参数。

②用样本中的其他所有未参与模型拟合的样本点对①中拟合出来的模型进行验证，将验证结果在拟合误差δ的范围内的所有点添加到之前的内点集合中，组成新的内点集合。并记录下此时内点集合的内点个数。

③重新选取X个内点，重复步骤①、②。

④在迭代K次（已知K的情况下）以后，判断K次迭代中哪一次得到的内点集合的元素数目最多，则将该内点集合对应的模型model（内点个数最多）作为最终拟合的目标。

1. RANSAC算法评价

RANSAC 算法的优点：

是能鲁棒的估计模型参数。例如，他能从包含大量局外点的数据集中估计出高精度的参数。

RANSAC 算法的缺点：

它计算参数的迭代次数没有上限，如果设置迭代次数的上限，得到的结果可能不是最优的结果，甚至可能得到错误的结果。

RANSAC只有一定的概率得到的可信的模型，概率与迭代次数成正比。另一个缺点是它要求设置跟问题相关的阈值，

RANSAC只能从特定的数据集中估计出一个模型，如果存在两个（或多个）模型，RANSAC不能找到别的模型

1. 应用RANSAC算法拟合直线样例。

https://zhuanlan.zhihu.com/p/62238520