无人机桥梁裂缝检测研究

第一章 绪论

1. 1 研究背景与研究意义

无人机是一种空中飞行的机器人，和普通机器人相比更加灵活，适合工作在普通机器人无法到达的地方。桥梁通常建立在大河大江、高山峻岭的地方，为了安全，需要定期检查桥梁缺陷以动态监测桥梁状态，而桥梁检测任务往往较为危险，不适合人工作业。无人机不仅可以灵活飞行，也能够携带传感器对桥梁进行多维度融合的无损检测，因此使用无人机进行桥梁检测具有操作灵活、使用安全、检测全面等特点。

* + 1. SLAM简介

SLAM(Simultaneous Localization and Mapping，实时定位与地图构建)是机器人自主导航和定位的关键技术，机器人在未知环境从任意位置开始运动，在运动过程中实时地依靠自身位置估计和地图进行定位,同时基于估计的位置建造增量式地图，实现机器人的自主定位和导航。

* + 1. 图像镶嵌累计误差简介

全景图像镶嵌技术是指将一组内容上存在交叠的图像序列经过配准并无缝融合成一幅宽视场、高分辨率的图像。在如医学影像、军事、地质勘探等领域中，由于相机获取的单幅图像存在视场的限制，目标往往不能被容纳到单幅图像中，这就需要对目标进行多次不同视角的拍摄，经过镶嵌后得到目标的整体影像。在全景图像镶嵌的过程中，由于单幅图像之间镶嵌误差客观存在，多幅图像镶嵌时就会出现累积误差，随着镶嵌数量的增加，误差会越来越大，导致最后的全景图像出现较大的偏差，对于类似无人机这样大量数据采集的设备来说，图像镶嵌中累积误差问题就显得十分突出。

* + 1. 桥梁裂缝识别与评价简介

桥梁的缺陷主要来自与长期使用过程中产生的裂缝，因此对与裂缝的处理是桥梁检测的关键。无人机获取的是桥梁的图片，经过处理之后，需要识别图片中的裂缝以定位和的评价，所以裂缝识别是裂缝检测的关键技术。

无人机可以安全简单地定期对桥梁进行自主检测，裂缝经过识别和定位后被可视化的展示出来，这时需要对裂缝进行自动化评价，主要是对裂缝进行定量评价，对不同时间内获取的裂缝进行定量评价，建立裂缝动态评价数据库，也为之后的裂缝评价打下基础。

* 1. 国内外研究现状
     1. SLAM研究概况

SLAM是一种概念，涉及方方面面的技术，概括起来主要是四个部分：

a). 基本理论

SLAM的基本理论可以分为滤波器和优化方法两大类。滤波器如卡尔曼滤波(EKF)粒子滤波(PF)等，属于早期的SLAM理论，但在13年出现了Random Finite Set的方法。优化方法主要是姿态图的优化，今年来优化方法的使用日渐增多。

b). 传感器

机器人使用传感器赶至人类世界，传感器的选择和安装决定了数学模型的具体形式。早期的SLAM使用激光传感器，现在多使用视觉相机，声纳等水下传感器，特别是几年来价格低廉的的深度相机使得SLAM的研究门槛下降了许多。在传感器方面，研究方向主要有新兴传感器的发明，安装方式的改进，传感器数据处理的改进等方面。

c.) 地图构建

建立地图是一个重要的研究方向，现存的地图主要有以路标标点组成的地图，卡尔曼滤波中构建和使用的就是路标地图。网格地图，也就是常见的点云地图，这种地图精度高，位置信息十分丰富，对导航很友善，但是网格地图的存储需要耗费大量的存储空间，如3D点云地图将所有空间点都保存起来，对于导航来说，只需告诉机器人障碍物的位置即可，但是对于需要构建三维空间的情况是十分友善的。拓扑地图，是一种紧凑型地图。拓扑地图借鉴图论中的概念，将地图抽象为点和边，更加符合人类的思维习惯。混合地图，这种地图就是把根据需要把上述地图糅合到一起，提供不同类别信息的地图。

d.) 回环检测

回环检测，又称闭环检测(Loop closure detection)，指机器人能够识别出历史场景的能力，倘若检测成功，就能显著减小累计误差，属于误差控制范围。

早期的回环检测使用词袋模型(Bag-of-Word)，其实是一个模型识别问题，现在有很多机器学习方法适用于回环检测，但由于SLAM实时性要求，所以现在的SLAM更侧重与在线学习方法。

SLAM还有很多高级研究方向，如动态场景，语义地图，多机器人协作等等。

* + 1. 图像镶嵌累计误差概况

在全景图像镶嵌的过程中，由于单幅图像之间镶嵌误差客观存在，多幅图像镶嵌时就会出现累积误差。图像镶嵌主要有两个步骤：图像配准和图像融合。近年来，研究人员在图像配准和图像融合方面做了很多研究。Lowe提出的基于尺度不变性的特征检测已经普遍运用于图像镶嵌。David G. Lowe 使用简化的SIFT(尺度不变特征转换，Scale-invariant feature transform)算法进行图像配准，降低了SIFT特征点的维度，提高了算法的运行速度，避免了大量的计算。林曾等[等人用SIFT不变性和动态规划的方法进行图像配准。李妍等人通过生物基因算法对图像配准进行优化，但生物基因算法的实时性不高。李艳 基于直线段检测算法的SIFT特征对图像进行配准，能够消除部分误匹配，对提高模型变换的精度起到一定作用。李小娟等人运用粒子群算法对图像配准和融合进行优化，但粒子群算法耗时很高。朱霞 使用双线性插值的图像融合方法对配准后的图像进行融合，但没有考虑到多幅图像拼接后图像扭曲无法正常插值的情况。现有的图像镶嵌中各种图像配准、图像融合的基本方法、类型的研究现状都得到了很好的总结。但是对图像镶嵌中的误差问题却很少有人涉及，对于类似无人机这样大量数据采集的设备来说，图像镶嵌中累积误差问题就显得十分突出。

* + 1. 裂缝识别与评价简介

现存的裂缝识别和检测方法主要是针对路面的检测，如基于不同滤波器的图像滤波方法，基于形态学的裂缝去噪方法，如李刚提出的基于灰色系统理论的路面图像裂缝检测算法，胡世昆提出的基于数字图像处理技术的路面裂缝检测算法。

裂缝评价的研究少之又少，多半是提取出裂缝认为去评测裂缝的情况，只有少部分根据裂缝的形状和面积进行定量的分析，如李刚的基于灰色系统理论的路面图像裂缝检测算法中将裂缝分为网状裂缝和线性裂缝，对于网状裂缝计算器外接面积，对于线性裂缝，计算器长宽比。这些数据能够在一定程度上反应裂缝的情况，但对于建立一个动态的裂缝评价系统来说还是有些不足。

* 1. 本文研究的内容和完成的工作

本文使用无人机传感器（彩色和深度摄像头）获取彩色和深度图像，基于RGBDSLAM对桥梁进行二维和三维的检测，RGBDSLAM不仅能对桥梁进行致密的三维重建，还能给出无人机在飞行过程中拍摄的轨迹和图像，给后期的桥梁分析工作提供更加精确的数据。对于二维图像，使用图像镶嵌技术将桥梁的关键部分拼接成一幅全景图像，尽可能的展示出桥梁关键部分的完整情况。之后对裂缝进行检测和定位以及定量评价，建立动态的桥梁裂缝评价机制。

* 1. 本文结构

本文对无人机桥梁裂缝检测进行研究，文章结构如下：

1. 对研究背景和研究意义进行了阐述，并对本文中所用的相关技术，如SLAM，图像镶嵌误差处理等进行了介绍，并叙述了国内外的研究现状。
2. 详细分析和实现了RGBDSLAM的各个步骤，并给出了实验室环境下的SLAM实现结果。
3. 对基于反馈机制的图像镶嵌误差处理进行了分析和实现，提出了解决的新思路。
4. 介绍和分析了本文选择的裂缝识别技术，提出了基于将为和基尼系数的裂缝评价技术。
5. 对本文的工作进行总结，并对未来工作进行展望，提出了可以继续工作的部分。

第二章 图像镶嵌累计误差处理

2.1. 特征点粗匹配

图像粗配准中，首先选用SIFT(Scale-invariant feature transform，尺度不变特征转换算法)提取两幅图像中的特征点，根据两幅特征点之间的关系进行特征点之间粗匹配。特征点粗匹配需要两个步骤，互匹配唯一性原则和匹配点对筛选。

2.1.1互匹配唯一性原则

SIFT算法具有良好的性能，能够在两幅图中寻找出大量的特征点，由于特征点的描述子带有比如位置、尺度、方向等信息，因此特征点的之间具有很大差异，直接利用特征点进行特征点对的选择和匹配具有很强的可行性，本文采用任意两个点的欧式距离来描述两个点之间的关系，并使用互匹配唯一性约束原则结合最近邻和次近邻距离之比以减少误匹配。

对分别属于两幅待匹配图像的两个特征点A和B，它们的描述子分别为和，其中和表示为描述子的某一维度值，则这两个点之间的欧氏距离定义如下：



SIFT算法生成的描述子为128维，这里*n*的值为128，dis为这两个点之间的欧氏距离。

互匹配唯一性约束原则：分别属于两幅图的两个点特征点A和B为待选匹配点对当且仅当A，B互为对方欧氏距离下的最近点。





如图1所示，对第一幅图中的每个特征点*Xi*，根据它距离第二幅图中所有特征点的距离建立一个列向量*Axi*，如式4。并根据欧式距离从小到大排序，得到第一幅图和第二幅图的欧式距离矩阵*M*，如公式5所示。



 

图1 互匹配唯一性约束原则 图2不同待配准图像阈值与匹配点对数量曲线



分别属于两幅图的特征点*Xn*和*Ym*，在执行互匹配唯一性原则时需验证以下两点：

1. *Ym*是*Xn*的最近邻点。这里由于矩阵*Axn*是有序的，所以矩阵*Axn*中第一个点就是*Xn*的最近邻点。
2. *Xn*是*Ym*的最近邻点。因为矩阵*M*的第一行表示的是以列表为序号的第一幅图中点的最近邻点。所以当*Ym*是*Xn*的最近邻点时，需要遍历矩阵*M*的第一行，找出以*Ym*为最近邻点的第一幅图中的所有点*Xk，Xn，Xh*，如果点*Xn*与*Ym*的距离是这些点中最近的（这里使用邻接表统计），那么*Xn*是*Ym*的最近邻点。

假设两幅图分别有*N*和*M*个特征点，互匹配唯一性原则的时间复杂度和空间复杂度分别如式6和式7所示：





如果互匹配唯一性原则成立，那么*Xn*和*Ym*为待选匹配点对。下面对所有的待选匹配点对进行筛选以确定粗匹配点对。

2.1.2 动态阈值分析

互匹配唯一性原则将两幅图中的特征点进行配对后，通过验证第一幅图中已经匹配的特征点的最近邻点和次近邻点的比值对待匹配点对进行筛选，这个比值叫做阈值。

当阈值设定在一定范围内，粗匹配下的匹配点对数达到一定数量后，就能给之后的精匹配提供良好的数据基础，称这个数量为期望匹配点数量*M*。这个值的设定和单张图片的重叠比有关。但是在阈值大于一定值之后，匹配精度得不到保证，阈值的调节不能超过最大阈值*tm*。图2中，在固定阈值*T*下，曲线*K1*，*K2*，*K3*对应的匹配点对分别为*m1*，*m2*，*m3*。当*t* = *T*，*m1* > *m2* > *M* > *m3，*对于三条曲线有：

1. 曲线*K3*的匹配点对数量*m3*不足期望值*M*。
2. 曲线*K2*的匹配点对数量*m2*略大于*M*。

3. 曲线*K1*的匹配点对数量*m1*远远大于*M*。

阈值*T*对于曲线*K3*是不符合条件的，需要调节阈值到*t3*，匹配点对数量才能略大于*M*，且匹配精度有保证；而曲线*K1*的匹配点对数量*m1*远远大于*M*，如果将阈值调节为*t1*，那么匹配点对数量就能刚好大于*M*，且匹配精度得到了提高。对曲线*K2*，当前阈值*T*基本符合条件，无需调节。

由上分析可知，对不同的两幅待匹配图像，需要选择不同的阈值进行匹配点对的筛选，以达到匹配精度和匹配点对数量的保证。这个保证精度和数量的阈值称为最佳阈值*TB*。在此分析的基础上，提出一种反馈机制，对不同的待匹配图像，算法可根据图像质量自行调节到最佳阈值*TB*，这样的反馈机制有着良好的自适应性。

2.1.3 反馈机制

文献[4]对图像最佳重叠比的概念做了分析，当两张图片之间的重叠比例达到一定数值时，图像镶嵌的误差可以降到最低。由于无人机是一种可控设备，无人机在执行任务时，可以根据需求控制其拍摄频率和其他参数，从而将图片重叠比控制在一定范围内。这里假设单张图片的重叠比为50%，后面会讨论当单张图片重叠比达不到或超过这个比例时对结果的影响。

建立一种反馈机制来动态调节粗匹配的阈值，运用二分法对反馈机制进行加速，能够快速的确定合适的阈值。

反馈机制描述如下：

1. 设定一个较小的阈值*Ts*；
2. 设定一个最大阈值*Tm*；
3. 根据图片拼接的数量动态计算图片最佳重叠比，将其换算成最佳匹配点对数量*Num*。
4. 二分法确定本次粗匹配的最佳阈值。
5. 当前阈值下的匹配点对即为粗匹配下的最佳匹配点对。

对于步骤4，这里给出二分法的步骤：

1. 确定当前二分的上下界，初始条件上界为最大阈值*Tm*，下界为最小阈值*Ts*。
2. 计算目标阈值为最大阈值和最小阈值的中间值，*tmp = (Tm+Ts)/2*。计算目标阈值下的匹配点对的数量。
3. 如果当前匹配点对为最佳匹配点对的95%~105%之间，二分结束，当前阈值为最佳阈值。如果不足最佳匹配点对的95%，将二分的下界改为*tmp*，如果大于最佳匹配点对的105%，将二分的上界改为*tmp*。
4. 重复步骤c直到步骤c停止或二分的上下界交叉。

2.1.4 动态阈值设定

现有文献中，一般将固定阈值设在0.6，本文取其大约70%，即0.4作为最小阈值*Ts*。最大阈值可以适当调整，本文设定为0.8。



图3 图像镶嵌重叠区域变化

对于最佳匹配点对数量，随着图片镶嵌数量的增加，重叠区域占总体的比例逐渐减小，如图3所示。图3 中（未经镶嵌的单幅待镶嵌图像之间的重叠比为50%，本文提及的单幅图像指的是未经任何镶嵌的单幅源图像）。图中一个小矩形代表单幅图像的50%的区域。。阴影部分表示本次镶嵌时两幅图的重叠区域。在单幅待镶嵌图像之间的重叠比为50%的条件下，计算任意数量图像拼接时阴影部分的面积，设单个小矩形的权重为1，那么单幅图像的权重为2。

设有A1，A2，A3,……,An共N幅图像进行镶嵌，相邻图像之间存在50%的重叠区域，计算最后一次镶嵌时阴影部分的面积。



根据容斥原理可以得到：





在初始化时计算每张图片的特征点数量，并记录每张图片的拼接历史，即当前待拼接图像是经过哪些图片拼接而成的，对A1，A2，A3,…,An共N幅图像，特征点数量分别为n1，n2，n3,…,nn。这样在每次计算最佳匹配点数量时就可以根据待拼接图像的拼接历史快速的换算出本次拼接的最佳匹配点数量。

在对大量图片数据进行实验统计后，本文将初始最佳匹配点对设置为230。

镶嵌第*N*幅图时，最佳匹配点对的计算如下：



公式11中代表历史拼接图像的匹配点对数量。

2.1.5 最佳重叠比

本文假设待镶嵌图片序列中相邻的单张图片的重叠比为50%，由于无人机飞行时环境的变化，这个比例可能得不到保证。下面讨论重叠不足50%和超过50%情况下，应用本文方法对粗匹配结果的影响。



图4 最佳重叠比分析

如图4所示，三个椭圆从大到小分别代表最佳重叠比大于50%，等于50%，小于50%下正确的匹配点对集合，匹配点对随重叠度增大而增加。匹配点对*A，B，C*分别属于的集合也如图所示。

当重叠比不足50%时，原本属于匹配点对的*B*被算入匹配点对中，成为误匹配点，只有*A*属于正确的匹配点对，此时粗匹配下的误匹配点增加。

当重叠比大于50%时，*B*点原本为正确的匹配点对，重叠比增大时它不会成为误匹配点，对结果无影响。

可见只有当重叠比小于50%时，本文的假设会增加误匹配数量，此时可以适当控制无人机的飞行参数保证重叠比不会过小。

在误匹配增加的情况下，良好的精匹配算法可以将绝大部分误匹配剔除。

2.2 特征点精匹配

在射影变换模型中，只需要4对匹配点就可以求出摄影变换矩阵，但是这是建立在理想情况下的，在实际图像镶嵌中，因为误差不可避免，那么这4个匹配点对只要有一个有误差，摄影变换矩阵就会产生很大误差。如果增加匹配点对的数量，单个匹配点对的误差就能被其他匹配点对分摊，所以在消除明显的误匹配之后，匹配点对数量的增加能够减小误差。

2.2.1 反馈机制

随机抽样一致算法(RANSAC，Random Sample Consensus, 随机抽样一致)对于明显的误匹配非常敏感，当图像质量不佳，匹配点对少的情况下，适当增加阈值可以增加有效匹配点对的数量；当匹配点对很多，适当减小阈值可以增加匹配点对的精度，也能保证匹配点对数量。如此反馈下的随机抽样一致算法有着良好的自适应性。算法最后选择的匹配点对集为最佳匹配点对集。反馈机制下描述如下：

1. 建立一个初始阈值*Ts*。
2. 设定一个步长*step*。
3. 设定一个最大阈值*Tm*。
4. 动态计算一个最佳匹配点对数量*Num*。
5. 从最小阈值开始，根据RANSAC算法得到的匹配点对数量*N*和最佳匹配点对*Num*进行比较，如果当前阈值下的匹配点对数量不小于最佳匹配点对*Num*，当前匹配点对就是最佳匹配点对集，反之在当前阈值上增加步长。直到满足最佳匹配点对条件或当前阈值大于最大阈值。
6. 当前匹配点对集为最佳匹配点对集。

对于上述反馈机制，可以参照粗匹配中的反馈机制将二分法应用在精匹配的反馈机制上。

2.2.2 动态阈值设定

经过了粗匹配时建立了反馈机制下的互匹配唯一性约束原则，保证了匹配点对的数量和质量。匹配点对中大部分都是有效的匹配点对，因为这里匹配点对的数量是两幅图的，而两幅图是否经过匹配不影响当本次匹配点对的数量，只与当前镶嵌的两幅图有关，无论这两幅图是当前镶嵌图本身还是的它的子图，这里取匹配点对的70%作为最佳匹配点对的数量。



由公式12计算出最佳匹配点对数量。

现有文献中一般取RANSAC阈值为1，这里分别取该固定阈值的10%和200%分别作为初始阈值*Ts*和最大阈值*Tm*，取步长为初始阈值的十分之一。即：



根据公式13和实时性要求，可以适当改变步长，动态的调整时间和精度上的平衡。

2.3 实验结果

由于图像拼接涉及到图像配准和图像融合两个步骤，每个步骤都有多种解决方案，现如今主流的方法是利用最短欧式距离初步确定匹配点对，用RANSAC方法去除误匹配点对，最后利用利普拉斯融合对图像进行融合，这套方法对三张(含三张)以下图片镶嵌有良好的效果，但是在多幅图像镶嵌中效果很差。本文主要的改进在于最短欧式距离的改进(互匹配唯一性原则)的和粗匹配以及精匹配下的反馈机制。

试验分别对九幅图像用本文改进的方法和主流方法进行图像镶嵌，统计每次镶嵌时粗匹配和精匹配的匹配点对数量和阈值变化，之后对整体结果用SSIM(structural similarity index measurement system,结构相似性评价系统)评价[14]。

图5是无反馈机制和反馈机制方法下九张图片镶嵌的结果，无反馈机制时，由于图像变换矩阵的不准确，图像出现了很多明显的黑缝，许多细节部分出现了小错位，图像边缘很毛躁，采用反馈机制后，图像细节对接准确，基本无镶嵌痕迹。



（a）无反馈机制 （b）反馈机制

图5 无反馈机制与反馈机制的全景图像镶嵌结果比较



（a）匹配点对数量对比 （b）阈值变化对比

图6 粗匹配下两种方法匹配点对数量和阈值的比较

分别统计每次图像镶嵌时粗匹配和精匹配的匹配点对数量和阈值的变化。下面是对图4镶嵌过程中数据统计结果。

图6中(a)中，在反馈机制下阈值根据图像镶嵌的次数与图片质量进行自我调节，随着图像镶嵌次数的增加，无反馈机制方法找到的匹配点对数量快速减少，导致匹配后期精匹配的数据太少。而在反馈机制下，由于阈值随图片镶嵌次数动态调节，算法找到的匹配点对数量可以基本保持平稳，且阈值处于最大阈值控制范围内，保证了匹配的精度。由(b)可以看出，反馈机制下阈值根据图片质量不断自我调节，以保证匹配点对的数量和质量。

精匹配与粗匹配类似，控制阈值动态变化来调节匹配点对的数量，限于篇幅不再赘述。下面利用SSIM对镶嵌结果进行整体评价。

SSIM算法是一种基于结构相似度图像评价方法，算法从对比度，结构相似度和亮度三个方面考察图像。SSIM方法评价函数为：



公式14中，*x，y*分别代表源图像和待评价图像，*l(x，y)*表示图像的亮度，*c(x，y)*表示图像的对比度，*s(x，y)*表示图像的结构相似度，得分越高表示结构相似度越大，最大值为1。

由于镶嵌后的图像没有原图像进行比较，本文采用镶嵌前源图像分别与结果图像做SSIM评价，将所有源图像的SSIM值求和后的结果作为镶嵌结果图像的SSIM评价值。假设有*n*-1幅待图像镶嵌，编号分别为1到*n*-1，镶嵌结果图像的编号为*n*。



公式15中SSIM（Final）表示镶嵌结果图像的总SSIM值，而SSIM(*i，n*)表示第*i*幅图像与镶嵌结果图像的SSIM值。

表1 两种方法下SSIM评价结果

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SSIM** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **Final** |
| **反馈** | 0.0885 | 0.0868 | 0.0814 | 0.0908 | 0.0991 | 0.1204 | 0.1307 | 0.1320 | 0.1174 | 0.9471 |
| **无反馈** | 0.0767 | 0.0793 | 0.0869 | 0.0763 | 0.0912 | 0.0970 | 0.1159 | 0.1277 | 0.1167 | 0.8677 |

可以看出，对比无反馈机制的镶嵌方法，反馈机制下每一幅待镶嵌的源图像与最后的镶嵌结果图像的SSIM值都有提升，而总体的SSIM值也有明显提升，可以看出反馈机制在全景图像镶嵌中起到了很好的优化作用。

第三章 RGBDSLAM技术

本章对RGBDSLAM中各个步骤做出详细分析，并基于华硕Xtion摄像头给出一个实验室环境下的实验结果。

3.1 图像变换为点云

Xtion获取的图像分为两个部分，即彩色图像部分和深度图像部分，彩色图像给出了被拍摄物体的色彩和细节信息，深度图像则提供被拍摄物体相对于相机的深度信息。RGBDSLAM重建的地图是致密地图，即点云表示下的三维地图，要生成三维地图，首先需要将Xtion获取的两部分图像结合起来，第一步是将二位的图像转换成三维信息，所以这一步是2D到3D的转换。

根据三维点云地图的假设，地图由一个点云来描述： ，每个点由6个分量组成，，分别表示图像颜色和空间信息。颜色信息方面，由彩色图像记录，空间位置信息可由图像、相机模型和机器人姿态计算出来。

RGBDSLAM使用常规相机的针孔相机模型，



图1

空间点 和它对应图像中的像素坐标 的关系为：







 为相机在 两个轴上的焦距， 指相机的光圈中心， 为深度缩放因子，整理可得：







相机在生产之后，就会产生四个不变的参数 ，称之为相机内参矩阵 ，相机内参可以由相机标定来去获取。取得相机内参之后，就可以用矩阵模型来描述每个点的空间位置和图片像素坐标，也就得到了图片上的2D点映射到空间3D点的公式。



其中， 和 是相机的姿态，为旋转矩阵，代表位移矢量，在单幅点云的条件下，认为相机没有发生旋转和平移，设置为单位矩阵 ，设置成零矩阵，是深度图中数据和实际距离的比例，Xtion中深度图的单位是毫米，将设置成1000。

3.2 图像特征提取与匹配

SLAM问题是由建图和定位两个问题构成，数据是机器人获取的类图像数据，机器人根据数据来建图和定位，给定两幅存在重叠的图像，SLAM系统需要根据这两幅图像重叠的部分计算出这两幅图像的关系，将这两幅图像组合在一起绘制成地图，这就是绘图的部分。根据图像之间的关系推导出相机姿态的转移关系，如平移和旋转矩阵，就可以得到相机在地图中的位置，这就是定位。

RGBDSLAM使用基于特征的方法来求解，首先需要求解出图像中的特征点。

假设视频存在重叠部分的两帧 和 ，获取两组对应的特征点：



存在旋转矩阵和位移矢量，使得



由于误差的存在，等号是不成立的，此时通过最小化误差求解，。

这里使用ICP或PnP求解。

特征点的选择多种多样，本文采用ORB特征描述子，ORB采用FAST（features from accelerated segment test）算法来检测特征点，使用BRIEF算法来计算一个特征点的描述子。假设我们已经得到相邻两帧对应的的ORB特征描述子，采用第二章中相同的方法进行匹配和误匹配的消除。使用PnP我们可以得到位移矩阵和旋转矩阵。

3.3 点云拼接

对点云进行拼接实际上是对点云进行变换，根据我们计算出来的旋转平移矩阵，可以精确的描述两个点云之间的关系。变换是通过变换矩阵来实现的，变换矩阵的描述如下：



左上部分是旋转矩阵，是一个正交矩阵，右上部分是位移矢量。左下部分 是缩放矢量，由于摄像头不启动变焦缩放功能，所以取成0，右下是用来齐次变换的1构成的矩阵。



这样就能将一块点云阵变换到另一块上，即实现了点云的拼接。在代码实现上，本文使用罗德里格斯变换将旋转向量变换成矩阵，然后和平移矩阵一起构成变换矩阵。

3.4 视觉里程计

视觉里程计就是使用机器人获取新的数据和已有数据进行匹配，估计机器人的与活动，然后将整个过程累加起来，形成完成的过程示意图。



由于匹配过程中可能发生匹配点过少，旋转与位移矩阵不准确的情况，所以限制匹配最小数量min\_match和RANSAC中最少局内点min\_points。

在移动过程中是连续采集图像的，所以两帧之间不会相差太多，规定一个最大相隔距离max\_distance。两帧之间的度量公式为：



3.5 姿态图优化

累计误差是所有计算机模拟或重建技术中都必须解决的问题，SLAM问题也同样存在着累计误差的问题，在视觉里程计中，后续的传感器姿态以来与前续传感器姿态，由于这个姿态是计算出来的，想要保证地图和定位的准确，就必须保证每次匹配都准确无误，但误差是不肯避免的。在建立SLAM系统中，我们希望在其中加入累积误差处理的部分，加强整个系统的鲁棒性和可信度。

在视觉里程计中，如果不仅仅是考虑两帧信息和关系，而是把所有帧的信息都一同考虑在内，那么一个简单的里程计就变成了一个全SLAM（full slam）问题。

姿态，是传感器在位置和朝向信息。姿态图，就是传感器姿态构成的图，这里的图是图论意义上的图，一个图由节点与边构成。



在SLAM的场景下，节点 表示相机姿态，由四元数或矩阵表示：



边 指两个节点 之间的变换：



姿态图示意图如下：



由于边 中存在误差，使得所有的边计算出来的数据不一致，因此可以优化不一致误差，也就是目标函数为最小化不一致误差的和：



 表示 的误差。是优化变量，给定初始值，根据目标函数对的梯度，调整使得取得极小值或是可以接受的值。



通常使用LM方法优化这种非线性平方误差函数，由于图的稀疏性，近些年稀疏的BA（Bundle Adjustment）求解方法被广泛应用求解这个问题。

3.6 回环检测

试想一下，如果机器人能够识别历史场景，那么对于场景中的某一部分就会有个重新的认识，对于误差的处理有着很重要的意义。

回环检测就是这么一种让机器人识别出历史场景的手段，机器人在不断接受环境给予的新数据的同时，也不断的将新数据和现有的数据进行匹配。但是机器人接受的数据量是十分巨大的，数据量多到我们并不需要将每一帧的都提取出来，而是根据精度要求设定一个步长，根据步长每隔一定数量的帧数，对帧进行提取。尽管是这样，在一定时间内还是有大量的数据累计，况且图片的处理需要耗费大量的时间和计算资源，所以需要找到一种高效的回环检测算法来解决这个问题。

基于词袋模型的回环检测算法是目前应用最为广泛的一种，词袋模型是机器学习中将图片进行分类的一种手段。