

**中 国 航 天**

**中国空间技术研究院**

**硕士学位论文**

**论文题目：基于特征点检测的图像配准算法研究**

培养单位： 中国空间技术研究院西安分院

学 科： 信息与通信工程

研究方向： 航天器数据传输与处理

导 师： 周诠

研 究 生： 张晨光

完成时间： 2019年2月11日

**基于特征点检测的图像配准算法研究**

张晨光（数据传输与处理）

导师：周诠

**摘要**

图像配准是数字图像处理领域中的一项重要技术，是模式识别、图像融合、图像拼接中的关键步骤，在卫星监测地貌、水文，医学检测病灶、分析病情，军事中精确制导都有着应用广泛。图像配准这个问题自从被提出，受到了很多的关注，出现了很多种算法，涉及到了很多其他的基础算法如边缘检测，图像滤波，插值，图像变换等。关于图像配准的算法大致可以分为两类，一种是利用整体的灰度信息的算法，如互信息法等，还有一种是利用了图像的局部特征，如以特征点为基础的配准算法。特征点是指那些稳定的，不随尺度变换和旋转平移而变化的点。具有代表性的算法有SIFT算法，SURF算法，KAZE算法，ORB算法等。以特征点为基础的算法具有计算量小的特点，且更具有鲁棒性，对遮挡，噪声，图像形变的适应性更好。本文将重点放在了以特征点为基础的图像配准算法，尤其是SIFT算法，对其进行了改进，并探索了图像配准在图像集中的应用。本文主要研究内容有：

1. 介绍和分析了关于图像配准的研究现状和主要的技术路线，重点介绍了基于特征点的几种算法的流程，深刻认识了图像配准的原理，并在不同情况下从不同方面对不同算法进行了系统的比较，对主流算法有了清晰的认识。除了图像配准技术本身，本文还研究了与之相关的几种基础技术，如图像的坐标系问题，双目视觉问题，图像之间的变换关系。图像的表示与变换关系的表示关系到最终的配准效果。
2. 针对SIFT算法效果良好，但计算量过大，配准较慢和没有利用图像色彩信息的缺点，分别提出了两方面的改进。一种是的思路是利用图像下采样减小图像金字塔，从而成比例地减少运算量，达到快速匹配的效果。一种是将SIFT算法与显著性检测方法结合起来，在增加的计算量相当有限的情况下，检测出了更多更高质量的特征点，有利于拟合出更加准确的模型，另一方面，利用显著性区域检测可以实现图像主体部分的分割，本文对原有方法进行了改进，可以获得更加完整的主体部分，在此基础上的图像配准可以去除背景噪声和背景特征的干扰，这在只要求主体部分配准的情况下是有用的。
3. 提出了一种图像配准的新的应用方法：基于图像配准较小图像数据集的方法。这里的图像数据集特指图像中场景大概率相同或者相似的数据集，比如对于某个建筑物或者风景的多次连拍或者抓拍，它们之间的关系很可能只是视角的变化，同时保存批量的相似图像会加重存储系统的负担，所以可以使用图像配准描述出变换关系，将本地已有的图像作为待配准图像，将新增的相似图像作为参考图像，用保存变换矩阵代替保存新的图像。为了便于复原删去的图像，可以记录配准后的图像与参考图像的差值情况。同时，在数据集中同时存在着场景完全不同的图像，这时是不合适做图像配准的。为了方便判断两幅图像是否包含相同场景，是否需要配准，本文引入了哈希算法，分别计算两幅图像的哈希值，当它们的哈希值的汉明距离小于某个阈值时就可以认为可以使用图像配准。

目录

[第1章 绪论 1](#_Toc2173906)

[1.1. 课题研究背景及意义 1](#_Toc2173907)

[1.2. 图像配准技术概述 2](#_Toc2173908)

[1.3. 国内外研究现状 5](#_Toc2173909)

[1.2.1. 基于灰度信息法 6](#_Toc2173910)

[1.2.2. 基于变换域 8](#_Toc2173911)

[1.2.3. 基于特征法 9](#_Toc2173912)

[1.4. 本文的组织结构 16](#_Toc2173913)

[第2章 图像配准相关技术 17](#_Toc2173914)

[2.1. 数字成像基础 17](#_Toc2173915)

[2.1.1. 针孔相机模型 17](#_Toc2173916)

[2.1.2. 图像坐标系 18](#_Toc2173917)

[2.1.3. 齐次坐标 22](#_Toc2173918)

[2.2. 图像变换 23](#_Toc2173919)

[2.2.1. 刚体变换（Rigid Transform） 23](#_Toc2173920)

[2.3.2. 相似变换（Similarity Transformation） 24](#_Toc2173921)

[2.3.3. 仿射变换（Affine Transformation） 25](#_Toc2173922)

[2.3.4. 投影变换（Projective Transformation） 26](#_Toc2173923)

[2.3.5. 非线性变换 27](#_Toc2173924)

[2.3. 双目视觉 28](#_Toc2173925)

[2.3.1. 本质矩阵（Epipolar Geometry） 28](#_Toc2173926)

[2.3.2. 单应矩阵（Homography matrix） 29](#_Toc2173927)

[2.4. 图像配准评价标准 29](#_Toc2173928)

[2.5.1 准确率 29](#_Toc2173929)

[2.5.2 配准精度 30](#_Toc2173930)

[2.5.3 峰值信噪比 30](#_Toc2173931)

[2.5.4 相似度检测SSIM 31](#_Toc2173932)

[2.5.5 配准速度 32](#_Toc2173933)

[2.5.6 其他评价标准 32](#_Toc2173934)

[第3章 基于特征点的图像配准 33](#_Toc2173935)

[3.1. 配准流程 33](#_Toc2173936)

[3.1.1. 特征点提取 33](#_Toc2173937)

[3.1.2. 描述子生成 33](#_Toc2173938)

[3.1.3. 特征点匹配 33](#_Toc2173939)

[3.1.5. 坐标变换和插值 34](#_Toc2173940)

[3.2. 常用算法介绍 34](#_Toc2173941)

[3.1.1. SIFT 34](#_Toc2173942)

[3.1.2. SURF 36](#_Toc2173943)

[3.1.3. KAZE 38](#_Toc2173944)

[3.1.4. ORB、 40](#_Toc2173945)

[3.3. 算法对比 41](#_Toc2173946)

[第4章 基于SIFT算法的新算法 41](#_Toc2173947)

[4.1 SIFT的简化 41](#_Toc2173948)

[4.2 SIFT与显著性检测 42](#_Toc2173949)

[4.4 本章小结 45](#_Toc2173950)

[第5章 配准算法的应用 47](#_Toc2173951)

[5.1 哈希算法 47](#_Toc2173952)

[5.2 图像减法 48](#_Toc2173953)

[5.2 图像压缩 49](#_Toc2173954)

[5.3 实验 49](#_Toc2173955)

[第6章 总结与展望 52](#_Toc2173956)

[参考文献 54](#_Toc2173957)

[硕士学位论文原创性声明 66](#_Toc2173958)

# 第1章 绪论

## 课题研究背景及意义

在我们的话语体系中有很多俗语与视觉有关：百闻不如一见、眼见为实、眼睛是心灵的窗户，这些俗语都表现了视觉的重要性。我们看书看报，看电视，看电影，都离不开视觉的贡献。事实上，通过视觉，人和动物感知外界物体的大小、明暗、颜色、动静，获得对机体生存具有重要意义的各种信息，至少有80%以上的外界信息经视觉获得，视觉是人和动物最重要的感觉。生理学的研究也在不断地为我们揭示视觉的奥秘。D.H.Hubel和 T.N.Wiesel凭借在视觉系统的信息加工中的贡献在1981年获得了诺贝尔奖。[1][2]自然界或者人为的图经过晶状体等在视网膜上面所成的像，这就是所谓的图像。图像由光信号变为电信号，刺激中枢神经，大脑做出相应的反应。

与此同时，如何对进入人眼之前的图像加工也是重要的研究内容。数字图像处理起源于20世纪20年代，当时通过海底电缆从英国的伦敦到美国的纽约采用数字压缩技术传输了第一张数字照片。[3]

所谓数字图像，就是像素点的坐标和灰度值都是离散数值的一个二维函数。而数字图像处理的目的就是通过计算机进行计算得出图像的一些属性值或者使得图像更适于人眼观看或进一步处理。传统的数字图像处理包括图像去噪[4].、图像增强、图像分割[5].、图像压缩[6]等。而最近大热的计算机视觉可以理解为是高级的数字图像处理，它大多借助深度学习，实现计算机对于图像的理解和分析。实现图像理解是计算机视觉的终极目标[7]。张等借助残差网络实现了图像去噪，图像解压缩等[8]。而在数字图像处理之中，图像配准[9][10]是一项关键的技术。因为随着大数据时代的到来，人工智能、机器学习的火热发展，我们面对的是海量的数据，其中很大一部分是智能手机、监控摄像头等传感器获得的图像，我们可能需要将数目庞大的不同图像结合起来处理，找出其共同特征或者找出差异，比如图像匹配、图像融合。然而，不同的传感器获得的图像效果可能千差万别，从分辨率，到成像角度，到图像大小都有差异，所以，在进行具体的图像处理之前需要进行图像配准[11]。图像配准的输入是相同场景在不同时间，不同视角，通过相同或者不同的传感器得到的两幅图像，拟合出二者之间的变换关系，并将其中一幅图像（待配准图像或变形图像）变换到与另外一幅图像（参考图像）相同状态的技术。

这项技术在遥感，医用，制图，计算机视觉领域被广泛地应用。如增强现实技术中，借助图像配准实现视觉融合，增强对环境的感知[12].；

在医学分析中，有时候单靠单一医学图像很难准确识别病灶，这就需要不同的医学图像融合获取更多的信息，而图像融合的关键就是利用图像配准将不同属性的医学图像对齐。哈工大的付浩威针对配准中的非线性不适定问题进行了研究[13]。

在遥感分析中也是同样的道理，受到气流、云层、光照等影响，卫星或者飞机上的相机在每次拍摄时不能保证具有相同的成像条件，同时不同时期的地形地貌水文可能发生改变，需要利用多幅图像进行对比或者拼接，这都需要图像配准技术。

在模式识别领域 ，如车牌识别[14]、人脸识别[15]、运动目标跟踪和三维重建，虽然基于监督训练的方法已经被广泛应用，但基于传统图像配准的方法也仍然占据一席之地：在样本图像与模板库中的图像进行配准，如果配准成功就说明识别成功。

## 图像配准技术概述

图像配准的数学描述是两个二维函数直接的空间变换和灰度变换。我们定义两幅具有偏移关系（包括平移，缩放，旋转）的图像分别为参考图像和偏移图像，并利用二维数组 和 表示图像相应位置处的灰度值，则两幅图像在数学上有如下变换关系：



其中表示二维空间坐标变换，表示灰度或辐射变换，用于描述因为传感器类型的不同或辐射变形所引起的变换。配准的目标就是要找出最佳坐标和灰度变换参数。通常意义上的的配准只表示位置坐标的变换。灰度或辐射变换可以归为图像预处理部分。

图像配准在目标检测、模型重建、运动估计、特征匹配、肿瘤检测、病变定位、血管造影、地质勘探、航空侦察方面都有广泛的应用。每一种配准方法通常都针对某个具体问题而设计的，众多方法中，唯一的共性就是每个配准问题最终都要在变换空间中寻找一种最有效的变换，这种变换能够使两幅图像之间在某种意义上达到匹配。1992年，L．G.Brown在其论文[16]中对近三十年来的图像的配准技术进行了系统的总结，他将配准技术分为四个组成部分：特征空间、搜索空间、搜索策略和相似性度量。下面分别简单对这四个部分做一个介绍：

1. 特征空间

特征空间是与输入空间对应的。对输入的图像进行特征提取，提取出的多维向量构成特征空间。图像的处理大多建立在特征空间之上，因为相比于输入空间，特征空间数据量更少，达到了降维的效果，因此能够大大减小图像处理过程的运算量；且可以更好地表示图像的特定信息，特征对灰度变化、图像变形以及遮挡等有很好的适应能力。一般选取的特征可以是点、线、区域这种浅层次的特征，也可以是更深层次的高阶特征如统计信息（信息熵、矩等）。在深度学习中，一般使用神经网络进行特征的提取得到特征图。特征也可以可视化。

1. 搜索空间

搜索在目的是匹配。具体来讲是根据特征的描述向量在另外一幅图中搜索合适的匹配对。为了进行高效率的搜索，可以对搜索空间进行划分。搜索空间的最终目的是靠匹配对成功拟合出最优的空间变换模型和其参数值。这里的空间变换模型以变换范围可以分成三类：全局、局部、位移场，全局变换指图像中所有的像素可以用同一个矩阵表示，局部则表示图像的不同区域由不同的参数表示，其中对关键点位置进行参数变换，其他位置进行插值；以变换方式可以分成线性变换和非线性变换，线性变换包括刚体变换，仿射变换和投影变换。非线性变换一般用多项式函数表示。

1. 相似性度量

相似性指的是配准（变换）之后的结果与参考图像的相似性，根据对结果的评估从而对搜索进行调整。不同的特征空间使用的相似性度量一般不同，而相似性度量的值会影响变换参数的选取，所以它们之间的关系是相互影响的。常用的相似性度量有互信息、欧式距离、联合熵等。

1. 搜索策略

搜索的任务的找到最优的变换参数，在寻找的过程中使用不同的策略会对搜索的效果产生很大的影响。策略不同，其采用的代价函数（损失函数）也不同。最简单的搜索策略就是贪婪搜索和穷举法，但在实践中一般选取效率更高的牛顿法、梯度下降法、Powell法、Brent法、黄金分割法、抛物线法等。搜索策略直接决定了配准的速度，并且搜索空间和相似性度量也都对搜索策略的选取有影响。

图像配准与图像匹配联系紧密。图像匹配是在两幅图像之间搜索相似内容的过程，目仅仅是找到相似内容的位置，不包括矫正的过程。图像配准在图像匹配的基础上，还需要拟合出相似内容之间的变换关系（相似内容之间可能存在旋转、缩放等变换），从而将待配准图像变换到参考图像相同的形态。它的目的是达到两个图像的内容在拓扑上的对应和几何上的对齐；即建立两幅图像内容的对应关系，并使相对应的内容在位置上对齐。这是图像的一种广义匹配概念。在寻找相似内容的过程，二者使用的方法基本相同。

因此，能够进行配准的多幅图像具有一个共同的特点，就是相邻的多幅图像分量之间至少有一部分在逻辑上是相同的，即相邻的图像之间有一部分是重叠的，应反映地面上的同一景物的特征，这是实现图像配准的基本条件。如果确定了相邻的图像之间的空间几何函数关系即坐标关系，就可以实现图像在拓扑关系上的配准。但是确定所有的像素之间的坐标关系是很困难的，而且是不切实际的。所以人们通常的方法是确定足够数量的同名像素（即两幅或多幅图像中代表相同特征的像素）点，根据所采用的图像之间的数学模型，可以表示图像中其它所有像素之间的坐标关系，从而完成图像的配准。由此可见，图像之间的配准主要包括两方面的内容：其一，是确定足够数量的配准控制点（RCP，Registration Control Point）；其二，是根据这些配准控制点确定两幅或多幅图像的像素之间的坐标对应关系。

根据待配准图像之间的关系，Brown将图像配准分为多源图像配准、基于模板的图像配准、多角度图像配准、时间序列图像配准四大类[18]。

多源图像配准，是由异源成像传感器获得的同一场景图像之间的配准。典型应用是多传感器图像的融合，比如医学图像领域的CT、MRI、PET、SPECT图像融合；遥感图像领域中多波段的图像融合，比如雷达、微波或者多谱段等；场景分类如建筑物、道路车辆，植被等。方法特点是建立传感器之间的变换模型，利用灰度信息的预配准、物体形状和一些基本标志可使问题简化。

基于模板的图像配准，是在待配准图像中寻找标准的模板图像。典型应用是识别和定位模板，如地图、物体和目标等。遥感数据处理方面，定位和识别定义好的或者已知特征的场景，如飞机高速路、车站，停车场等。模式识别、字符识别、标识确认、波形分析等。方法特点是根据模板预先选定特征信息。

多角度图像配准，对从不同观察点获得的具有不同视场角度的图像间进行配准。典型应用是对图像的深度或形状进行重建。接收机视觉：从视角差异中构建三维深度和形状信息；目标物体的运动跟踪、序列图像分析。方法特点是变换多为透视变换，多用于视觉几何和表面属性等的假设条件，须考虑阻挡问题。

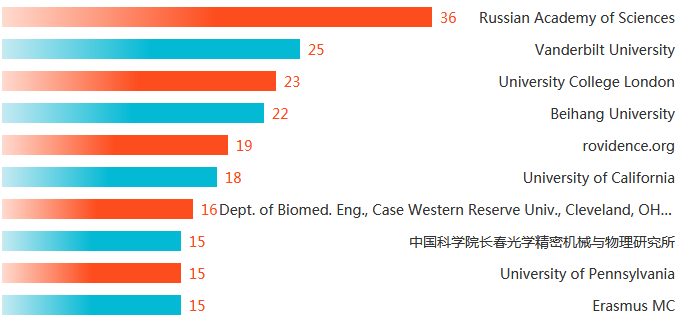
时间序列图像配准，是在不同时间或者不同条件下获得的同一场景图像间的配准。典型应用是监视和检测变化或增长。医学图像处理中数字剪影、血管造影、注射造影剂前后图像的配准、肿瘤检测、白内障检测。遥感数据处理中自然资源监控、核增长监控、市区增长监控。方法特点是需要容忍图像中部分内容的差异和形变对配准的影响，有时需要建立传感噪声和视点变换的模型。

## 国内外研究现状

图像配准最早在上个世纪七十年代出现，由美国提出并运用于导弹制导，极大提高了导弹的命中率，将弹着点平均误差半径缩小到十几米。在之后图像配准被国内外研究者继续研究，并在各个领域得到了广泛应用，如模式识别、医学诊断、计算机视觉等。图像配准是很多重要问题如图像拼接[17]、图像融合的关键步骤。



百度学术统计的图像配准(Image Registration)在各个年份的成果数。



百度学术统计的各个机构的研究成果数。

可以将图像配准分成两类：基于灰度的图像配准和基于特征的图像配准[18]。但是这是1992年的总结，在谢凤英[19]

等主编的资料中将图像配准分成三类：基于灰度信息、基于变换域和基于特征的图像配准方法。不同的图像配准方法总是对应于某种适用的图像变换模型[20]。

下面对这三种方法做一点介绍。

根据Institute of ScientificInformation(ISI)的调查数据表明，最近十年里至少有超过1000篇的学术论文在研究图像配准问题。2012年网上公布的美国申请专利中，图像配准相关的部分就超过50项，IBM和GE等大型跨国企业甚至有自己的工作组专门研究图像配准问题，而在最近的国际顶级学术会议上都有关于图像配准的专题讲座，这足以见得图像配准问题的关注程度。

1. 基于灰度信息法

基于灰度信息的方法顾名思义就是利用两幅图像之间灰度信息的相似性进行搜索寻找最值，最值确定参考图像和待配准图像之间的变换模型的参数。这种方法的关键是找到匹配位置，根据已知模板图像与另外一幅图像的子图进行滑动匹配。不同的匹配算法主要是相关性度量的准则不同。算法优点是不需要对图像进行复杂特征提取等预处理，缺点是运算量大，没有利用图像的颜色信息；不能直接用于具有非线性变换的图像配准；对噪声较敏感。光照强度或灰度强弱作为显著信息而不是局部的形状或者结构作为特征信息时效果更好，更受研究者们的欢迎。基于灰度信息的方法有互相关法；互信息法；序列相似度检测方法SSDA[21]。

平均绝对差算法（Mean Absolute Differences，MAD），由Leese在1971年提出，以模板与图像窗口的平均绝对差作为相似性度量。该算法是模式识别中常用的匹配算法，计算简单但计算量大，且容易受噪声影响。但是MAD仍然可以作为其他算法的补充，燕磊等[22]在2017年就将Oriented FAST和MAD结合起来用于图像配准。

绝对误差和算法（Sum of Absolute Differences，SAD）和MAD算法的思想几乎一样，只不过将相似性测度改成了模板与子图之间的L1距离，并且不再求均值。郭龙源博士等将SAD作为费用函数，在三个通道分别处理，对彩色图像的匹配进行了改进[23]。

而如果将SAD方法中的距离度量改为L2距离，就是误差平方和算法（Sum of Squared Differences，SSD）。如果对SSD中的度量公式求均值，那么就是平均误差平方和算法（Mean Square Differences，MSD）。

序贯相似性检测算法（Sequential Similiarity Detection Algorithm，SSDA），由Barnea和Silverman于1972年提出，是一种对于之前的传统模板匹配算法的改进，比MAD算法可以快几十到几百倍。在SSDA中，首先定义了绝对误差：



其中是搜索图，是搜索图的子图，表示模板。带有上划线的表示平均值。所以绝对误差就是模板与子图各自减去其自身的均值之后的差值。而加速的实现依赖于阈值的设定：子图与模板中随机选取的点之间的误差会累积，当当前子图产生的误差累加到超过阈值时，就会停止当前子图的计算。根据阈值选取的不同，SSDA可以大致分成三种：固定阈值SSDA算法、单调递增阈值序列的SSDA算法和自适应阈值序列的SSDA算法[24]。张维琪等利用OpenMP编程语言实现了多核多线程的自适应阈值SSDA的图像配准并行算法[25]。

针对SSDA受灰度值的线性变化影响大的缺陷，西北工业大学的贾凯提出了相邻像素差值矩阵的概念，在此基础上利用序贯相似度进行相关性比较，消除了灰度线性变化的影响[26]。Yijun Wang提出了基于小波金字塔作为搜索策略的改进自适应序列相似性检测算法，提高了匹配速度[27]。

hadamard变换算法（Sum of Absolute Transformed Difference，SATD），也是基于绝对值求和的算法，不同的是在SATD算法中对模板和子图的差值得到的矩阵先求其hadamard变换，再求绝对值之和，即SATD值，作为相似性测度。SATD值最小的子图则是最佳匹配。

归一化积相关算法（Normalized Cross Correlation， NCC），同样是利用子图与模板图之间的灰度信息，相关性度量使用归一化的互相关公式。NCC算法精度高、不易受灰度值的线性变化影响，但是计算速度较慢。

1948 年，Shannon 首次提出互信息的概念。后经多年发展，为解决异源医学图像配准问题，Viola[36]和 Collignon[37] 等人在1995年分别独立将互信息引入图像配准领域。利用图像配准位置处互信息最大原理，判断配准位置。优点是对异源图像配准效果较好，无需预处理以及鲁棒性强，缺点是计算量大，要求图像间有较大的重叠区，容易出现局部极值。该测度不需要对不同成像模式下图像灰度间的关系作任何假设，也不需要对图像进行分割或任何预处理，具有自动化程度高的特点，因此，最近几年将互信息作为图像配准过程的相似性测度，利用最大互信息法进行图像配准成为了图像处理领域的研究热点。国防科大的马政德提出并实现了一种基于互信息的刚性图像配准并行算法[38]。最大互信息算法是图像匹配的有力工具，但其仍然存在缺点，如插值引起的局部极值，容易导致误配准，空间信息利用不足等缺点，一些学者也提出了相应的改进方法。北京交通大学的阮求琦教授团队和大连理工的刘秀平团队也进行了相关研究，并且结合了局部特征研究。曹蹊渺还在图像配准结果和配准算法性能两方面评价的基础上提出了一种基于距离误差的配准结果客观评价算法[39]。曹源、武艳超以互信息作为相似性测度，提出改进的粒子群算法与烟花算法结合的混合算法来进行图像配准[40]。粒子群的惯性权重动态变化提高粒子的搜索范围，并利用烟花算法的高斯变异算子增加粒子的多样性，防止粒子群算法陷入局部极值，并提高了配准的精确度。

1. 基于变换域

该方法根据变换域内图像改变方式完成图像的登记。有人把空域和变换域的方法也认为是基于灰度信息的配准。

利用傅里叶变换，通过对比平移、旋转、缩放在时域和频域的对应关系，求得相应的变换参数。其好处在于其变换对于平移、旋转、缩放有天生的对应关系，利于 FFT 的快速实现，对与频域不相关的噪声有较强的鲁棒性。经傅立叶变换后相位性质表现成信噪比和最大峰值，与图像的空间域配准相比精度更高，配准性能更好。傅立叶变换方法充分利用相位性质对图像进行配准，非常适合于存在位移变化的图像进行配准。

此外，在傅里叶变换的基础上，在20世纪80年代前后出现了小波变换。小波变换优于傅里叶变换的地方是它在空域和时域是局部化的，在高频处窗口高而窄，可以精确地定位出突变信号的位置；在低频处窗口矮而宽，适应分析缓变信号的需要。这种特性被称为“变焦”,因而小波又被叫做数学显微镜[41]。这样有利于分析非平稳信号（如遥感图像）。在和图像分解和重建的塔形算法启发下,基于多分辨率分析框架,Mallat建立了离散正交小波的一种快速算法一一Mallat算法,通过与正交镜像滤波器（QMF）的卷积可以分解或重建给定的信号[42]。将小波变换技术运用到图像配准领域是目前科技工作者研究的热点。利用小波分解可以方便的产生各种分辨率的图像,我们可以综合利用这些图像进行图像配准方法的研究。

湖南大学伍君提出了基于小波变换提取边缘特征点的图像配准方法和基于小波变换与互信息测度相结合的图像配准方法[43]。辽宁科技大学的夏召红团队针对传统边缘特征检测的不足,引入了基于小波变换的边缘特征提取方法[44]。

1. 基于特征法

基于特征的图像配准方法能够将对整个图像的各种分析转化为对图像特征的分析，由于特征的数量相对于整幅图像来说只是一小部分，因此能够大大减小图像处理过程的运算量，而且特征对灰度变化、图像变形以及遮挡等有很好的适应能力，因此应用也更广泛。根据特征点进行图像配准涉及到特征提取、特征匹配以及对运动关系的求解。根据特征进行图像配准的方法是基于像素特点衍生出来的，在该方法中，涵盖图像信息的特征点都能够被加以使用，例如边缘化的特征点、角点。对于每一幅图像，都有其特有的特征。因此，通过提取图像中容易提取的特征，再通过相关矩阵等方法将两幅待配准图像中的特征进行匹配，去除未配对的特征。利用匹配特征坐标联立方程组，解出变换参数，因此基于特征法的图像配准大体上分为四步：特征提取、特征描述符生成、特征匹配、去除误匹配。

算法中常用的特征有直线、边缘、轮廓、重心、不变矩、特征点等。对于直线、边缘、轮廓等的提取，较为经典的算法包括 Hough 特征提取[48]、Canny算子边缘检测等[49]。基于 Hough 变换的特征提取算法，相关研究包括 H. Gonçalves;J. A. Gonçalves;L. Corte-Real;A. C. Teodoro 等人针对遥感图像提出名为 CHAIR 的方法[50]，该方法基于互相关和 Hough 变换，实验表明达到亚像素级精度，同时适用于异源图像，且对于 SAR 图像和可见光图像的配准也具有一定的潜力；还有Troglio G等人提出的将分水岭分割算法和广义Hough变换结合的自动特征提取算法，使用在了数据量庞大的行星图像配准分析中[51]。对于中心、不变矩特征提取较为经典的算法包括 Hu M K 提出的 Hu 不变矩特征[52]、基于Zernike 多项式的 Zernike 矩特征[53]等。相关研究如 Xiaolong Dai 和 Siamak Khorram[54]提出的将链码与不变矩相结合对特征进行描述，实现了对遥感图像的自动、鲁棒配准；Jan Flusser[55]利用图像中的不变矩特征，实现了目标识别和分类；Jan Flusser[56]和 Tomás Suk针对对称物体识别提出一组旋转、尺度、仿射不变的不变矩。

针对基于全局的特征配准算法计算量大、对图像质量和类别要求较高、在部分遮挡或相同区域较少时算法失效等缺陷，越来越多的学者将研究重点放到了对局部特征的研究上，一系列性能优异的算法相继被提出。局部特征包括局部区域特征和局部点特征。

基于区域的局部特征一般具有两个特点：第一，提取的特征区域要具有某种显著性；第二，提取的特征区域要具有仿射、尺度、旋转等不变性。目前较常用的局部区域检测算法包括：最大稳定极值区域（Maximally Stable Extremal Regions，MSERs）检测算法、基于边缘区域（Edge-Based Regions，EBRs）检测算法、基于密度极值区域（Intensity Extrema-Based Regions，IBRs）检测算法、显著性区域（Salient Region）检测算法等。

最大稳定极值区域[58]（MSERs）检测算法基于分水岭算法。此算法计算复杂度较低，且运用高效的合并-查找[59]（Union-Find）算法维护连接区域的列表和面积，因此经过加速可满足实时应用。Tuytelaars T, Van Gool L[60]等人提出了两种局部区域特征提取算法，即基于边缘区域（EBRs）检测算法和基于密度极值区域（IBRs）检测算法。EBRs 算法检测到的边缘可能发生断裂，从而造成特征区域的不匹配。为了弥补这方面的不足，他们又提出了 IBRs 算法。IBRs 算法利用图像中局部密度极值点和以极值点为中心向外辐射的“光线”上的特征点构造仿射不变区域。

Kadir 和 Brady在 2001 [61]年提出了显著性区域（Salient Region）检测算法，后在2004 年进行了进一步改进[62]。显著区域是图像中最能引起用户兴趣，最能表现图像内容的区域，通过计算图像区域与其周围区域在底层特征上的差别来决定显著度。在某一特定尺度下，不同的局部图像的信息熵有可能相等，但在不同尺度下信息熵会表现出不同的变化规律。根据这点，可以进行某些局部仿射不变特征的检测。

Mikolajczyk、Tuytelaars、Schmid 等人在文献[63]中对几种主流的仿射不变区域检测算子进行了性能测试，发现 MSERs 算法综合性能最优，Hessian-Affine算法[64][65]次之；Hessian-Affine算法[64][65][66]和 Harris-Affine 算法可以提取较多的不变区域，在复杂场景匹配方面具有优势；EBRs 算法适用于包含较多交叉边缘的图像；显著性区域算法的综合得分虽然较低，但在物体分类识别方面表现较好。

最早的对局部点特征的研究可追溯至 20 世纪 70 年代末。1977 年，Moravec[67]利用灰度自相关函数检测出“兴趣点”，即角点。其原理十分简单：平坦区域的像素灰度相差不多，而在多个方向都有亮度变化的位置则可判定为角点。缺点是对噪声敏感、不具有尺度和旋转不变性等。

Harris在1988 年提出了 Harris 角点提取算法[68]。其基本思想是观察一个假定的特征点周围小窗口内的方向性强度平均变化。由于灰度变化会出现在多个方向，因此要在所有可能的方向上进行计算。判定方法为首先获取平均强度变化最大值对应方向，接着检查位于它垂直方向的变化是否也很强烈，同时满足条件的便是一个角点。具体在计算中，用一个协方差矩阵表现所有方向上的强度变化率。此协方差矩阵的两个特征值表征了最大平均强度变化以及垂直方向上的平均强度变化：若这两个值比较小，那么位于平坦区域；若一个较大，另一个较小，则位于边界上；若两个值都较大，则位于角点处。因此，如图1-1，若两个特征值都大于某一阈值，则此点被判定为角点。Harris 角点对旋转和灰度变换具有不变性。

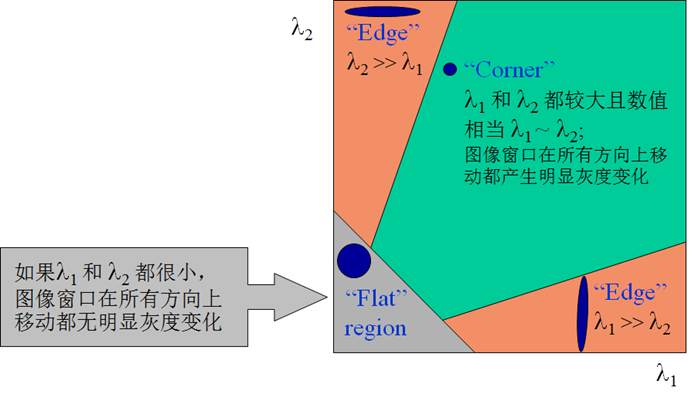


图1- 1：Harris角点检测，由协方差矩阵的特征值判断是否是角点

20 世纪 90 年代，Lindeberg[69][70][71][72][73]系统提出了图像的尺度空间理论，从理论上证明了构造尺度空间唯一的线性核是高斯核。通过分析图像中各个局部的尺度，极大地推动了尺度不变特征提取相关算法的发展，如 Mikolajczyk 和 Schmid[71]将 Harris 角点检测与高斯尺度空间相结合，提出了 Harris-Laplacian 算子，增加了 Harris 角点的尺度不变性。

在局部特征研究过程中，Lowe 做出了里程碑式的贡献。从 1999 年到 2004 年，Lowe 相继发表的三篇论文[75][76][77]，系统地提出并完善了 SIFT（Scale Invariant Feature Transform）算法。Lowe 利用构造高斯差分金字塔的方式，来快速求解高斯拉普拉斯空间中的极值点，加快了特征提取速度。在特征点周围 16\*16 的区域，通过梯度直方图统计的方式对特征点的方向进行归一化处理，并形成 128 维特征向量对特征点进行描述。这样的一系列算法，使得 SIFT 特征在图像旋转、尺度变换、一定的仿射变换和视角变换等条件下都具有良好的不变性。由于其优良的性能，直至现在依然在目标识别、图像拼接、目标跟踪等领域扮演着重要角色，甚至已经成为评判一种算法性能的标准。然而，算法仍存在一些问题，如阈值过多且难以确定和存在一定的错误匹配等。

2006 年，Bay[78]在 SIFT 算法的基础上，提出了另一个具有重要意义的局部特征，即 SURF（Speeded Up Robust Features）。Bay通过引入积分图像和模板近似，在特征描述阶段采用一阶Haar小波响应等改进，SURF算法在光照变化和视角变化不变性方面的性能接近SIFT算法，旋转和图像模糊不变性、鲁棒性优于SIFT算法，计算速度快3倍左右。在对特征点的描述方面，虽然在统计方式和描述子维数上与 SIFT 算法有所不同，但同样基于梯度直方图统计。

除 SURF 算法外，比较重要的类 SIFT 算法包括 PCA-SIFT[79] 、GLOH[80] 、ASIFT[81] 、ORB(oriented brief)[82]、CSIFT(colored SIFT)[83]等其他改进的SIFT算法[84][85]。

PCA-SIFT 利用了主分量分析技术（Principal Component Analysis，PCA）。在生成描述符时共分三步：第一步首先生成一个投影矩阵；第二步检测特征点；第三步利用投影矩阵对特征描述符降维，得到较低维度的描述符。因此，其核心思想为采用PCA 技术，通过寻找高维空间中起主要作用的分量，将高维空间的样本投影到低维的特征空间，从而降低了特征描述符的维度，提高了特征的匹配速度。但它的缺点是事先需要对大量典型图像进行学习，对于不同类型的图像，此算法不适用。

GLOH（Gradient Location-Orientation Histogram）描述子通过将特征点周围矩形

模板改为对数-极坐标系下的圆形模板，增强了 SIFT 描述子的鲁棒性和独特性。在生成梯度直方图时，由于维数较低，因此不用进行降维处理。其不足与 PCA-SIFT 一样，也需要事先对大量样本进行学习。

ASIFT（Affine-SIFT），是 2009 年提出的一种基于 SIFT 的算法，解决SIFT等特征提起算法在影像有较大倾斜角度时，同名点对提取较少的问题。在较大的仿射变换下依然可以实现同名点的匹配。近来，在手机等移动平台上已经有人成功实现了该算法，体现了算法的较好的应用前景。

Rosten 和 Drummond[86]于 2006 年提出了 FAST（Features from Accelerated Segment Test）角点检测算法，并在 2008 年进行了改进[87]。这种算法认为如果一个点周围的一定范围的圆形区域内有足够多连续点的灰度比中心点大或小，则此中心点被判定为角点。在一幅图像中，非角点往往是占多数，而且非角点检测要比角点检测容易得多，因此首先剔除掉非角点将大大提高角点检测速度。N=12时，编号为1，5，9，13的这4个圆周像素点中应该至少有三个像素点满足角点条件，圆心才有可能是角点。因此首先检查1和9像素点。如果和在 之间，则圆心肯定不是角点，否则再检查5和13像素点。如果这4个像素中至少有三个像素满足亮度高于或低于，则进一步检查圆周上其余像素点。同时，在判定过程中采用了机器学习中的 ID3 算法，加快了判定速度，在实际应用中可以达到实时性要求。

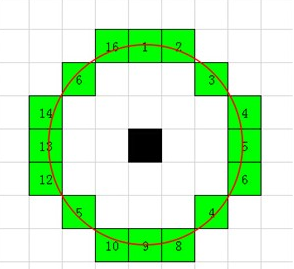


图1- 2：选取以像素点为中心的半径为3的离散化的Bresenham圆形区域，如果在圆周上有N个连续的像素的亮度比中心点大或小一个阈值，则判为角点。OpenCV中的N选取的是8。

Taylor、Rosten、Drummond[88][89]等人同样利用了机器学习并利用 FAST 特征检测提出了 HIPs 算法。经过对训练图集的学习和算法优化，使得这种算法的特征匹配速度达到了微秒级。

Alcantarilla、Bartoli、Davison[90]等人在 2012 年发表的论文中提出了 KAZE 特征。传统的 SIFT、SURF 等特征检测算法都是基于线性的高斯金字塔进行多尺度分解来消除噪声和提取显著特征点。但高斯分解是牺牲了局部精度为代价的，容易造成边界模糊和细节丢失。非线性的尺度分解有望解决这种问题，但传统方法基于正向欧拉法（forward Euler scheme）求解非线性扩散（Non-linear diffusion）方程时迭代收敛的步长太短，耗时长、计算复杂度高。由此，KAZE 算法的作者提出采用加性算子分裂算法(Additive Operator Splitting, AOS)来进行非线性扩散滤波，可以采用任意步长来构造稳定的非线性尺度空间。在论文的各项测试中，KAZE 算法比 SIFT、SURF 等算法具有更好的稳定性、旋转不变性、尺度不变性。唯一的不足是其计算复杂度过高，比SIFT 算法运算时间更长。如果可以在非线性尺度空间构造和特征提取、描述等方面进行优化，从而达到可以实时运算的要求，则 KAZE 算法一定会有更广阔的应用前景。

国内的研究也紧跟国际接轨，并且做出了一些贡献。西北工业大学李晖晖，郑平，杨宁，胡秀华针对参考图像和待配准图像对之间存在的仿射变换问题，提出了一种灵活通用的、基于 SIFT特征和角度相对距离的图像配准算法[91]。算法充分利用了图像正确匹配特征点对之间存在的角度关系，实现了特征点之间的精确匹配。潘建平等人基于SURF，引用相对距离理论，改进了图像配准算法[92]。上海海事大学的侯坚、张明提出一种蚁群优化的改进SIFT特征点的图像配准算法[93]。首先采用内核投影算法 Walsh-Hadamard 对 SIFT 特征描述子进行降维，然后采用优化的蚁群算法针对初匹配点进行提纯，提高了匹配率。北航的安勇，郑虹提出了基于K近邻点的SIFT特征点匹配[94]。根据配对点周围的点的相似度配对匹配的正确与否。电子科技大学的肖哲等针对大脑CT图像提出了结合SIFT特征和灰度特征的描述子并且通过欧式距离和余弦相似度得到匹配图[95]。

近几年国内学者针对图像配准的研究重心逐步转移到多模图像配准，提出了很多理论与方案。但是这些方法多基于像素级别的，而随着各个领域对图像质量要求的提高，如卫星遥感、医学领域需要更精确的配准，需要将配准精度控制在一个像素以内，即达到亚像素级。众多的学者对亚像素级配准的方法进行了研究与实验，提出了相关插值法、梯度法、小波变换法。相关插值法，这类方法的精度取决于插值算法的质量；梯度方法，其前提条件是图像灰度保持不变，微分算子对光照较为敏感；相位相关法利用FFT的方法提高配准的速度；小波变换法，使得求解配准参数的运算量较小，有效地提高效率。

基于特征点的配准技术研究已成为基于特征的配准研究领域中最热门和最突出的研究方向基于特征点的配准技术具有以下三个优点：

（1）在图像中提取出来的特征点在数量方面与图像的像素点相比有很大程度下降，运算量得到了大量消减。

（2）特征点在匹配过程中相似性度量函数能够很好地反馈坐标变化，因此能够使精度和准确率在很大程度上有所提升。

（3） 特征点可以适应图像的各种变化，即使图像遭受光照、旋转、尺度和噪声等影响，依然可以提取出质量较好的特征点适应性较强。

可以看出，针对图像配准的研究取得了一些重要的成果，但仍然存在一些亟待解决的问题：

(1)对于不同传感器图像与数据的配准技术。不同传感器获得的图像、图像灰度和特征信息往往有很大的不同，如SAR图像与可见光图像成像机理不同，同一地物或目标在灰度、统计特性等方面都有很大差异，很多研究者对于SAR与可见光图像配准也提出了一些方法。但大多利用河流、湖泊、道路等特征显著的图像进行实验，或在对于图像目标有先验知识情况下进行配准，但算法缺乏广泛的适应性。

(2)自动配准实现。自动配准是不需要人工干预、计算机可以根据既定的程序自动完成图像配准，但目前的研究中很多算法还需要一些人工的干预，无法实现全自动配准。

(3)亚像素级的图像配准技术实现。

(4)快速配准算法的实现。

(5)较大几何位置差别条件下的配准技术的实现。在待配准图像和参考图像之间有较大的尺度比例差别、较大的旋转角及较大的平移时，甚至还伴有各种图像的

畸变等情况下的配准也存在很大的困难。

## 本文的组织结构

图像配准作为图像处理中关键的预处理步骤，一直是国内外的研究者的课题之一。本文的主要研究内容是基于特征点的图像配准，研究目的是对该类方法有一个综合性的评测，对其实时性差的算法的特点进行改进，对一些没有充分利用图像信息，配准效果稍差的算法进行优化。

全文共分六章，各章节将按照如下安排：

第一章为绪论，先介绍了图像配准的研究背景和研究意义，然后简单解释了图像配准的定义及其与图像匹配的联系。介绍了图像配准的三类方法，并分别介绍了三类方法的国内外研究现状。

第二章介绍了图像配准相关的技术，如图像变换模型，描述图像关系的单应矩阵、本质矩阵，图像的坐标系等。这些都是与图像配准密切相关的知识点，掌握好这些技术才能更好的认识图像配准，更好的进行研究。

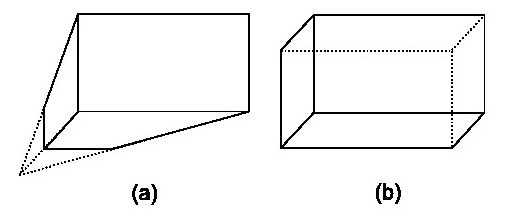
# 第2章 图像配准相关技术

## 数字成像基础

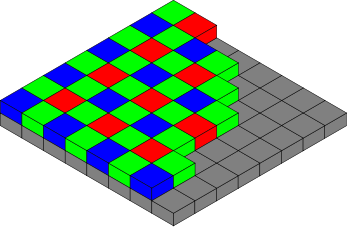
* + 1. 针孔相机模型

真实世界是三维立体的，但是数字图像是平面的，相机模型就是为了简化表示从三维物体到二维平面图像的成像过程。人眼看到物体实际上是物体反射光或者自己发光，所以无论是人眼成像还是相机成像的过程实际上都是投影的过程。最基础的投影变换就是正交变换和透视变换。使用正交变换的相机是正交投影照相机（Orthographic Camera），使用透视投影的相机称为透视投影照相机（Perspective Camera）。透视投影相机符合我们人眼观察物体近大远小的常识，如(a)，这种模型下我们认为光源是点光源，相机被抽象成一个点；而正交投影中的光源是平行光，这样得到的图像大小和物体实际大小是相同的，在三维空间内的平行线，投影到二维空间中也一定是平行的，如下图中的(b)。一般说来，对于制图、建模软件通常使用正交投影，这样不会因为投影而改变物体比例；而对于其他大多数应用，通常使用透视投影，因为这更接近人眼的观察效果[96]。

针孔相机模型(Pinhole camera model)就使用了透视投影，数学上大大降低了复杂性，不考虑像差、不考虑景深，并且假定等效透镜是薄透镜。所以说针孔相机模型仅仅是一种真实相机的成像过程的近似，甚至于我们可以说这是一种非常粗糙的近似，这使得这个模型对越符合针孔相机模型的真实相机近似程度越高，如网络摄像头、手机镜头、监控探头等等。



对于数字彩色图像的获取，还涉及到拜耳彩色滤波阵列。根据对于人眼的分析，存在红绿蓝三种感光细胞，所以我们可以把可见光分成三原色红绿蓝，根据三原色可以恢复出其他任意一种可见光。所以，我们只要分别获得每一个像素的三原色就可以得到真彩图像，每个三原色通道的获得可以使用滤波片。但是滤波片的工艺难度高，且造价高，需要三种滤波片严格对齐，于是，柯达公司的布莱斯·拜尔发明了一种新的滤波片排列方式，只需和图像像素大小相同的滤波片阵列进行一次滤波就可以得到三通道信息。下图是一种拜耳阵列：



这样的结果就是有的像素位置捕获到了红色通道的信息，有的是绿色通道信息，有的是蓝色通道信息。当这些像素位置正好和滤波片的颜色一样时当然可以无失真的捕获真实色彩，但是大多时候没有这么凑巧，我们仍然需要三原色来合成真实色彩。通过拜耳阵列就可以对每个三原色进行插值（去马赛克demosaicing），得到各个位置的其他二原色。而使用的插值方法其实是和常见的插值方法类似的，如最近邻插值，双线性插值。

* + 1. 图像坐标系

对于数字图像，我们的操作一般都是基于其像素的。为了表示每个像素，就需要知道其位置坐标，这就需要建立图像的坐标系。图像坐标系一般有四种：世界坐标系，相机坐标系，图像物理坐标系，像素坐标系。我们常用的是像素坐标系，但是了解像素坐标系与其他坐标系之间的关系有利于我们深入认识图像处理。下面对这几个坐标系做一个介绍。

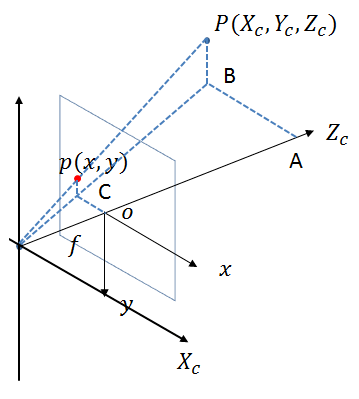
世界坐标系，也称客观坐标系，客观地描述了相机的位置和物体的位置，单位是米。坐标是三维的，用(X,Y,Z)表示。世界坐标系使用右手坐标系。

相机坐标系和世界坐标系相似，只不过更加规则化。相机坐标系规定了坐标原点是相机的光心，相机的光轴为Z轴，方向指向相机前方，X轴与Y轴与物体坐标系平行，单位为米。相机坐标系也使用右手坐标系。相机坐标系与世界坐标系之间是刚体变换，即只存在旋转和平移的关系：





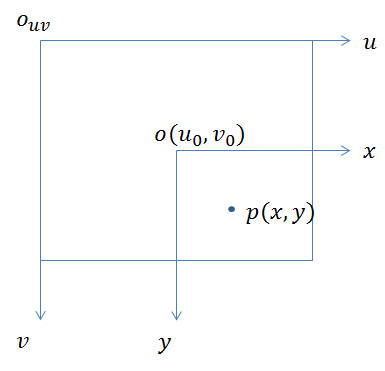
图像物理坐标系，也称为平面坐标系，原点在传感器中心，X轴与Y轴分别与图像的边界平行。用(x,y)表示其坐标。图像物理坐标系依然使用物理单位如毫米，依然使用右手坐标系。相机坐标系到图像物理坐标系的转换关系：







像素坐标系与其余坐标系最大的不同就是坐标原点的选取与单位的选取。坐标原点取在图像左上角，单位使用离散值，用像素的数目来衡量。相机坐标系到像素坐标系的转换有两个要素，一个是存在平移，一个是需要进行单位的转换：



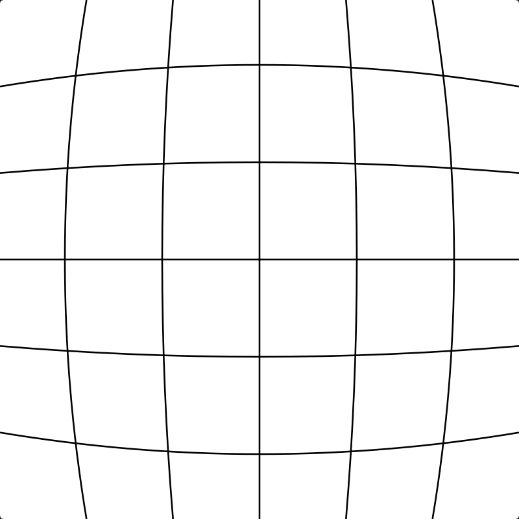
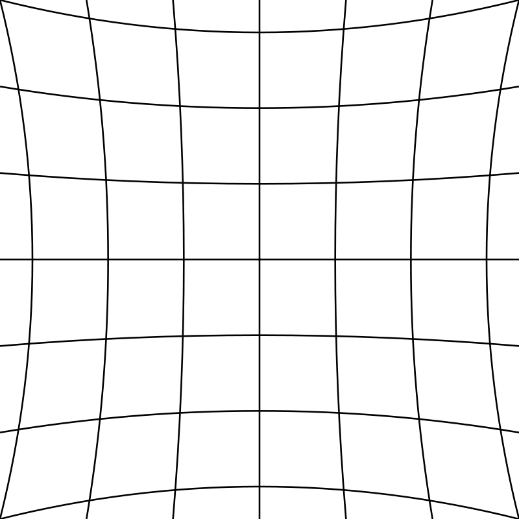




所以，我们可以一路从世界坐标转换到像素坐标，这也是数字图像的数字化生成过程。



第二行等式中的第一个矩阵是相机内参，包含x轴和y轴方向上的归一化焦距和图像中心与左上角的相对位置信息，第二个矩阵是相机外参，包括旋转向量和平移向量。相机的内参和外参可以通过张正友标定进行获取。另外，畸变也属于相机的内参，也可以在标定过程中解决。畸变主要是由镜头造成的，可以分为径向畸变和切向畸变。切向畸变产生的原因是感光元平面和透镜不平行。径向畸变产生的原因是光线在透镜中心和远离中心的地方的弯曲程度不同。径向畸变可以分成桶形畸变（Barrel distortion）和枕形畸变（Pincushion distortion）。桶形畸变中，物体得到不同程度的放大，偏离光轴越远，放大程度越小，就好像是图像被重新映射在了一个桶形的面或者球面上。最常见的桶形畸变发生在广角鱼眼镜头（Fisheye lenses）中，该镜头将宽度无限的物体映射到了有限的图像中。枕形畸变中，偏离光轴的部分放大效果越显著，光线向光心弯曲，就好像是枕头中被下压的部分。

桶形畸变 枕形畸变

* + 1. 齐次坐标

在平常数学运算中，我们一般使用笛卡尔坐标，但是在计算机图形学中使用的更多的是齐次坐标（Homogeneous Coordinates）。齐次坐标有三大优点。

一，还可以表示无穷远处的点。在欧式空间（Euclidean space）或者笛卡尔空间（Cartesian space）中，平行线是无法相交的，但是真实世界中因为透视关系，平行的铁轨在无穷远处也可以相交。而欧式空间中的坐标没有意义，所以齐次坐标就出现了。齐次坐标可以用维的坐标表示笛卡尔坐标系中维的坐标。齐次坐标对应笛卡尔坐标，是缩放系数。这样当取，就可以表示笛卡尔坐标系中的无穷大，对应无穷点（points at infinity）。一般的，取，得到归一化的齐次坐标。是齐次坐标，被称为增广矢量（augmented vector）。



二，齐次坐标可以区分点和向量。齐次坐标表示是计算机图形学的重要手段之一，它既能够用来明确区分向量和点，同时也更易用于进行仿射（线性）几何变换[97]。在笛卡尔坐标系中的一个坐标，我们无法判断它表示的是一个点还是一个从原点指向该点的向量。齐次坐标因为有一个冗余的维度（缩放系数），可以作为标志位，当它取1时表示坐标点，当它为0时表示向量。这是因为表达一个点比一个向量需要额外的信息：

对于单位基，向量可以由单位正交基线性表示：，那么点就是，如果将向量和点都写成矩阵相乘的方式， 就可以得到向量和点的齐次坐标形式。

三，齐次坐标允许平移、旋转、缩放和透视投影表示为矩阵与向量相乘的运算，而使用笛卡尔坐标，平移和透视投影不能表示成矩阵相乘。仿射变换其实就是线性变换（旋转缩放）与平移的叠加。笛卡尔坐标系不能用乘法表示仿射变换主要是因为平移变换。现在利用齐次就可以将平移变换中的矩阵相加转换为矩阵相乘：



至于为什么叫齐次，这是因为齐次坐标与笛卡尔坐标不是一一对应的。只要前个分量与缩放系数的比例相同就会映射到同一个笛卡尔坐标，即对齐次坐标每一个分量同等进行缩放，都表示的是笛卡尔坐标系中的同一个点，这就是尺度不变性（Scale Invariant），也就是齐次。

## 图像变换

图像配准的目的是将不同时间或者不同观察点或者不同传感器拍摄的同一场景的图像进行叠加，实际上其本质就是要找出参考图像与待配准图像坐标系之间的空间变换和灰度变换关系，而尤以图像之间的空间变换关系最为关键。

变换模型[98][99]是指根据待匹配图像与背景图像之间几何畸变的情况，所选择的能最佳拟合两幅图像之间变化的几何变换模型。可采用的变换模型有如下几种:刚性变换、仿射变换、透视变换和非线形变换等。

各种配准技术都要建立自己的变换模型，变换空间的选取与图像的变形特性有关，图像的几何变换可分成全局、局部两类。全局变换对整幅图像都有效，通常涉及矩阵代数，典型的变换运算有平移，旋转，缩放；局部变换有时又称为弹性映射，它允许变换参数存在对空间的依赖性。对于局部变换，由于局部变换随图像像素位置变化而变化，变化规则不完全一致，需要进行分段小区域处理。

经常用到的主要变换有刚体变换、仿射变换、投影变换、非线性变换。

### 2.2.1. 刚体变换（Rigid Transform）

变换前后两点间的距离依旧保持不变，即变换前后图像保持不变，则被称为刚体变换。 刚体变换可分解为平移变换、旋转变换和反转（镜像）变换。

刚体变换模型包括三个参数：分别对应了平移和旋转变换。二维空间中的点到的刚体变换可以表示为：



其中分别为图像在方向和方向的平移参数，是图像的选择角度。

刚体变换常常应用于较为理想的条件下或者用来检测算法的可靠性和鲁棒性。刚体变换中平移变换和旋转变换是不可逆的，也就是说对图像先进行平移后旋转和先旋转后平移，产生的效果是不同的。

### 2.3.2. 相似变换（Similarity Transformation）

在刚性变换模型中加入一个尺度变换参数就形成了相似变换模型，相似变换

也是一种典型的空间变换类型。经过相似变换后，图像中的直线仍然为直线，物

体的形状保持不变；但是由于尺度变换参数的加入，图像中任意两点间的距离会变化，物体大小也会随之发生改变。

二维空间中的点到的相似变换可以表示为：



可见，相似变换有四个参数： 。容易看出，刚性变换是相似变换

的一种特殊形式，刚性变换中 。

### 2.3.3. 仿射变换（Affine Transformation）

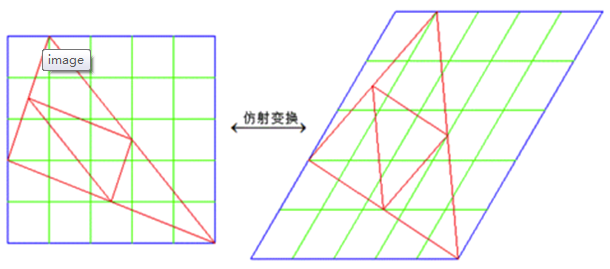
仿射变换是一种二维坐标到二维坐标之间的线性变换，保持二维图形的“平直性”（即变换后直线还是直线不会打弯，圆弧还是圆弧）和“平行性”（其实是指二维图形间的相对位置关系不变，平行线还是平行线，相交直线的交角不变）。

图2- 1：仿射变换，保持了二维图形的“平直性”和“平行性”，但会错切

图1- 3

相比于相似变换，仿射变换中的尺度变换参数在不同方向上不一定是均匀的，这就会形成图像的错切，就是允许图形任意倾斜，而且允许图形在两个方向任意伸缩的变换。仿射变换，可以保持原来的线共点、点共线的关系不变，保持原来相互平行的线仍然平行，保持原来的中点仍然是中点，保持原来在一直线上几段线段之间的比例关系不变。仿射变换可以通过一系列的原子变换的复合来实现，包括：平移（Translation）、缩放（Scale）、翻转（Flip）、旋转（Rotation）和剪切（Shear）。

仿射变换可以用下面公式表示：



可见，相似变换有六个参数：，容易看出，仿射变换是相似变换的一般形式，当时，仿射变换退化为相似变换。

### 2.3.4. 投影变换（Projective Transformation）

如果一幅图像中的直线经过后映射到另一幅图像上仍为直线，但平行关系基本不保持，则这种变换称为投影变换(Projective Transform )。在计算机三维图像中，投影可以看作是一种将三维坐标变换为二维坐标的方法。二维平面投影变换是关于齐次三维矢量的线性变换，在齐次坐标系下，二维平面上的投影变换具体可用非奇异矩阵形式来描述。

在欧式几何中，如果说在一个平面中任意两条直线必然相交于一点，是不能成立的，因为存在一个例外，即两条直线可以是平行的。这是欧式几何的麻烦之处。而换个说法，可以说平行的两条直线仍然相交于一点，只是它们相交于无限远处。这个在无限远处的交点，叫做理想点(ideal points)。由于增加了“理想点”，我们将传统的欧式空间拓展到了投影空间。在投影空间中，平面上任意两条直线的确是必然相交于一点的，即不再保持平行性。

正侧投影（Orthographic Projection）有些地方也翻译成正交投影。

正交是世界中的物体按照平行的光线投射到一张纸上(你的画布)，仿佛被压缩在上面，无论这个物体在世界中多远，投射结果还是原来的大小, 2d 游戏多用此种投影。如图2-2：

图2- 2：正侧投影，平行光的投影

透视投影（Perspective Projection）透视比较接近我们现实的方式了，眼睛形成一个夹角，离你的眼睛越近，东西越大，范围越小；反之离眼睛越远，东西越小，视野越开阔。它具有消失感、距离感、相同大小的形体呈现出有规律的变化等一系列的透视特性，能逼真地反映形体的空间形象。透视投影通常用于动画、视觉仿真以及其它许多具有真实性反映的方面。因此多用于 3D 中。如图2-3：

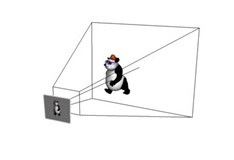


图2- 3：透视投影

事实上，无论是平移和旋转这样的基本变换形式，还是刚体变换、相似变换和仿射变换这样的复杂的空间变换，都是投影变换的某种特殊表示，都可以利用齐次坐标写成大小的投影变换矩阵的形式。

### 2.3.5. 非线性变换

非线性变换又称为弯曲变换(Curved Transform)，经过非线性变换，一幅图像上的直线映射到另一幅图像上不一定是直线，可能是曲线,在二维空间中，可以用以下公式表示：



 式中，表示把一幅图像映射到另一幅图像上的任意一种函数形式。多项式变换是典型的非线性变换，如二次、三次函数及样条函数，有时也使用指数函数。

表2- 1：各种变换模型比较

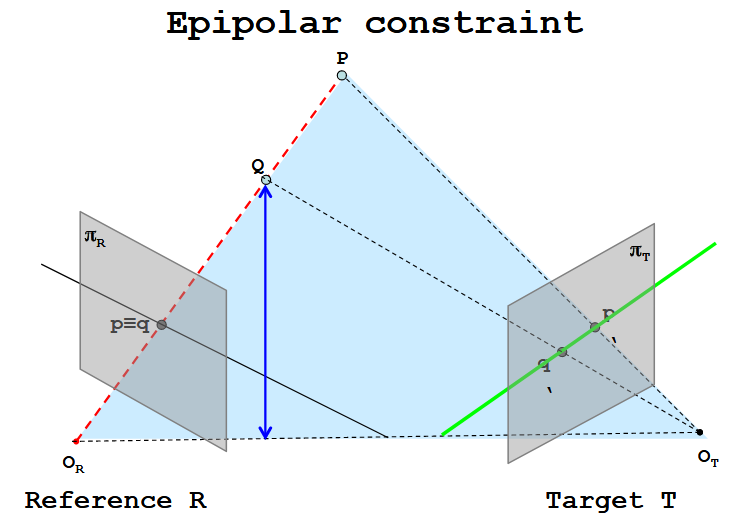
|  |  |  |  |
| --- | --- | --- | --- |
| 刚体变换 | 反转 | 旋转 | 平移 |
| 仿射变换 | 反转 | 旋转 | 平移 | 缩放 |
| 投影变换 | 反转 | 旋转 | 平移 | 缩放 | 投影 |
| 非线性变换 | 反转 | 旋转 | 平移 | 缩放 | 投影 | 扭曲 |

图像配准中，采用不同的图像变换说明了图像之间的映射关系将会不同。因此对于不同类型的图像，选取恰当的图像变换模型可以适当地提高图像的配准精度。表2-1总结了各种变换模型可以表示的变换，可以看出，非线性变换可以描述更多的变换，但也更加复杂。在对需配准的图像进行空间变换后，要对变换图像进行重新取样，以取得变换后的像素值。

## 双目视觉

图像配准中的输入图像很大一类属于视角变化，是相机相对位置变化造成的。同时在计算机图形学中也存在一种双目视觉，使用两个相机获得图像的景深信息。双目视觉涉及的知识主要的对极几何，而其实以本质矩阵和单应矩阵最为重要，表示了两幅来自不同相机之间的图像的像素之间的约束关系。

### 2.3.1. 本质矩阵（Epipolar Geometry）

****  
根据对极几何（Epipolar Geometry），同一个点在不同成像平面上的像素点存在极线约束，对极几何利用极平面的特殊性质将对应点约束在另外一幅图像的极线上。两个相机之间的关系用旋转矩阵R和平移向量T表示，相对于引入的先验信息。由极平面内三向量共面，可以推导出两个成像平面内点的约束关系，即本质矩阵E：



在本质矩阵中使用的坐标系是图像物理坐标系，如果将坐标系转化成利于图像处理的像素坐标系就可以得到基础矩阵F（Fundamental matrix）：



其中和分别是描述两个相机由图像物理坐标系到像素坐标系转化关系的仿射矩阵。

OpenCV工具库中函数findFundamentalMat可以找到基础矩阵Fundamental matrix。基础矩阵大小为，有8个自由度，需要至少4对匹配点对进行求解。

### 2.3.2. 单应矩阵（Homography matrix）

本质矩阵和基础矩阵中两个视图中点的对应关系是点对应直线（极线），而单应矩阵是点对应另外一幅视图中的点，约束力更强。之所以H矩阵约束力更强，是它不仅仅引入了两个镜头之间的旋转和平移变换，还包含了景深信息d。







为了保证两个视图之间的景深保持一致，拍摄画面最好是平面图像，或者是远景。更特别的，如果能保证平移向量为零，即相机只存在旋转运动，那么单应矩阵将更加简单。当以上条件都不满足时，就需要使用多个对应点拟合出一个最优解作为单应矩阵。

## 图像配准评价标准

### 2.5.1 准确率

准确率定义如下：



式中，为最终正确的匹配点对数目，是错误的匹配点数目。在理想情况下，图像的点通过变换后映射到的点为，和重合时被视为正确匹配点对。由于各种原因导致实际无法达到完全重合的效果，如果和在空间和尺度上非常近，其距离在规定范围内，则认为他们就是匹配正确的特征点对。

### 2.5.2 配准精度

配准精度是用来衡量配准结果的好坏，采用两幅图像所有匹配点(变换后坐标与参考图像对应点)坐标的均方根RMSE来衡量。



式中，为参考图像上的点的坐标，为待配准图像点的坐标，表示图像间的变换关系，为最终的匹配点数目。RMSE数值越小，则表示精度越高，效果越好。

### 2.5.3 峰值信噪比

峰值信噪比PSNR（Peak Signal to Noise Ratio）经常用作图像压缩等领域中信号重建质量的测量方法，它常简单地通过均方差（MSE）进行定义。两个m×n单色图像I和K，如果一个为另外一个的噪声近似，那么它们的的均方差定义为：



峰值信噪比定义为：



其中， 是表示图像点颜色的最大数值，如果每个采样点用 8 位表示，那么就是 255。更为通用的表示是，如果每个采样点用 B 位线性脉冲编调制表示，那么就是。

对于每点有RGB三个值的彩色图像来说峰值信噪比的定义类似，只是均方差是所有方差之和除以图像尺寸再除以 3。

PSNR是最普遍和使用最为广泛的一种图像客观评价指标，然而它是基于对应像素点间的误差，即基于误差敏感的图像质量评价。由于并未考虑到人眼的视觉特性（人眼对空间频率较低的对比差异敏感度较高，人眼对亮度对比差异的敏感度较色度高，人眼对一个区域的感知结果会受到其周围邻近区域的影响等），因而经常出现评价结果与人的主观感觉不一致的情况。

### 2.5.4 相似度检测SSIM

SSIM（Structural Similarity）结构相似性.、分别表示图像X和Y的均值，、分别表示图像X和Y的方差，表示图像X和Y的协方差。

SSIM分别从亮度、对比度、结构三方面度量图像相似性。







、、为常数，为了避免分母为0的情况，通常取，一般地取，则



SSIM取值范围[0,1]，值越大，表示图像失真越小。

SSIM特殊形式如下：



### 2.5.5 配准速度

配准速度是指从输入两幅图像到得到图像配准结果的整个过程所花费的时间，时间越长，则配准速度越慢，反之，速度越快。

### 2.5.6 其他评价标准

配准时间、配准率、算法复杂度、算法的可移植性、算法的适用性、图像数据对算法的影响等。

采用主客观结合来评价本算法的好坏，更加具有说服力。客观评价方法为归一化互信息( NMI) ，是信息论互信息的延伸，Studholme[100]等提出 NMI 评价配准结果，其值越大表示配准效果越好。主观评价方法为图像差值法，即参考图像减去配准后的图像得到差值图像，差值图像越黑表示配准结果越好。

# 第3章 基于特征点的图像配准

## 配准流程

### 3.1.1. 特征点提取

特征的选取至关重要，因为它是对图像信息的提取，基本上决定着算法的性能，好的特征可以提高算法匹配的速度和精度，减少不确定因素对结果造成的干扰。特征点的定义比较宽松，是图像中具有独特性和不变性的点。独特性是每个特征点具有唯一性；不变性指的是特征点应该具有尺度不变性，仿射不变性（旋转和平移不变），鲁棒性（对噪声、光照等）。不同图像配准算法之间主要的不同点就是如何提取特征点的方法。

### 3.1.2. 描述子生成

特征点所含的信息相对较少，但除了特征点坐标之外，为了保证尺度不变性和旋转不变性，还会包含尺度因子和主方向等。同时，为了具有鲁棒性，还会加入特征点邻域的一些统计信息，这些信息进行编码，就得到了描述子。描述子是特征点的数字化表示，有利于之后的匹配过程。不同图像配准算法之间主要的不同点除了提取特征点的方法之外就是生成描述子的方法。常用的生成描述子的方法有BEIEF可以生成二进制的字符串。

### 3.1.3. 特征点匹配

特征点的匹配是根据描述子的相似程度进行两幅图像之间特征点的匹配。匹配的过程与相似度准则的选取有关，而相似度准则的选取与描述子的形式有关。对于相似性测度目前常用的方法有两种：距离测度法、相似性函数法。距离测度法是对向量的距离求差的函数，主要有：明氏距离、欧氏距离、马氏距离等。相似性函数法主要是对向量的方向进行操作，主要有：夹角余弦法、相关系数法等。确定了相似性测度，根据如何在另外一幅图中搜索得到最相似的特征点又有不同的搜索策略。好的搜索策略可以提高算法效率，缩短运行时间。经常使用的有：穷尽搜索、分层搜索、Powell 算法、遗传算法、蚁群算法。

3.1.4. 变换模型估计

指根据待配准图像与参考图像之间的几何畸变的情况，选择能最佳拟合两幅图像之间变化的几何变换模型，可以分为全局映射模型和局部映射模型。其中，全局映射模型利用所有控制点信息进行全局参数估计；局部映射模型利用图像局部的特征分别进行局部参数估计。常见的变换模型包括仿射变换、透视变换、多项式变换等，其中最常用的是仿射变换和多项式变换。

### 3.1.5. 坐标变换和插值

将输入图像做对应的参数变换，使它与参考图像处于同一个坐标系下。由于图像变换后的坐标点不一定是整数，而计算机中存储的图像均为其灰度值的整数形式因此，因此就需要根据周围邻近点的灰度值进行重采样（插值处理操作）。常用的插值方法包括：最近邻插值、双线性插值、双三次插值、B样条插值、高斯插值。

采用最邻近像元法时工作量比较小，算法实现起来也比较简单，但是从采样后图像的质量来看，有明显的锯齿状。双三次卷积法对图像边缘有增强的作用，采样后图像的质量得到明显的改善，但是相应的计算量大大增加了。双线性插值法考虑到了采样点周围的像素点对它的影响，解决了最邻近像元法图像连续性不强的缺点，并且计算量与双三次卷积法相比也有所下降，采样后的整体效果比较令人满意。

## 常用算法介绍

自从SIFT算法提出以来，

### 3.1.1. SIFT

尺度不变特征转换(Scale-invariant feature transform或SIFT)是一种电脑视觉的算法用来侦测与描述影像中的局部性特征，它在空间尺度中寻找极值点，并提取出其位置、尺度、旋转不变量，此算法由 David Lowe在1999年所发表[75]，2004年完善总结[77]。在Mikolajczyk对包括SIFT算子在内的十种局部描述子所做的不变性对比实验中，SIFT及其扩展算法已被证实在同类描述子中具有最强的健壮性。

其应用范围包含物体辨识、机器人地图感知与导航、影像缝合、3D模型建立、手势辨识、影像追踪和动作比对。SIFT 特征是基于物体上的一些局部外观的兴趣点而与图像的大小和旋转无关。对于光线、噪声、细微视角改变的容忍度也相当高。基于这些特性，它们是高度显著而且相对容易提取，在庞大的特征数据库中，很容易辨识物体而且鲜有误认。此算法有其专利，专利拥有者为英属哥伦比亚大学。

极值点的检测在差分高斯金字塔中进行，搜索同尺度的9个邻域点和上下相邻尺度对应的9×2个点中的极值，再通过尺度空间函数进行三维二次函数曲线拟合寻找真正的极值点。尺度空间函数的泰勒展开式为，其中**是尺度空间函数，**是尺度空间函数在得到的初始特征点处的值，表示尺度空间上的点相对于初始特征点的偏移量。因为真实特征点应该是尺度空间函数上的极值，所以对求导并使之为0可解得真实特征点相对于初始特征点的偏移量。若大于阈值，表明真实特征点与初始特征点相距较大，需要通过重新插值对极值点进行拟合；同时可得偏移量处对应的尺度空间函数值，若其大于某一阈值，则认为其对应的特征点的对比度足够大而保留，否则进行剔除。

边缘点的特征是横跨边缘处的主曲率较大，而垂直边缘方向的主曲率较小。因为主曲率和2×2 Hessian矩阵的特征值大小成正比，可将比较主曲率的问题转化为比较特征值大小的问题。令*α*，*β*分别为***H***的最大和最小特征值，其比值，。，，有。是关于的递增函数，当即两个特征值大小相等时取最小值。可设定阈值，当，表明，即两个特征值大小的比值大于规定的阈值，出现了某个方向梯度很大，而另一个方向梯度很小的情况，表明特征点落在了边缘，属于边缘响应点，应该剔除。因此，计算得到特征点的Hessian矩阵的迹和行列式后再与阈值比较的方法可剔除边缘响应点。

在高斯差分尺度空间下可得稳定的极值点，这保证了特征点的尺度不变性。为了使特征点同时具有旋转不变性，还需要为特征点分配方向信息。特征点的方向信息是基于梯度的。数字图像中像素点的梯度幅值大小，梯度方向为。特征点自身的梯度方向并不直接作为特征点作为方向信息，而是对特征点的3*σ*邻域内所有像素的梯度方向信息进行直方图统计。具体做法是对邻域内的梯度方向进行量化处理，取直方图最高值对应的梯度方向作为特征点的主方向，为了增强鲁棒性，还可以选择直方图中峰值大于主峰值80%的次高峰作为辅方向。对特征点周围像素区域分块（4×4个），在每块内计算梯度直方图（8个方向），生成具有独特性的128维归一化向量作为特征点的描述子。

BFSIFT是基于SIFT的一种改进算法。BFSIFT将标准的高斯尺度空间通过双边滤波用各向异性尺度空间来代替。BFSIFT可以提高SIFT在SAR(Synthetic Aperture Radar)图像中的表现，SAR图像中有较多的斑点噪声(speckle noise).但是，双边滤波也是一个耗时的过程，所以无法使用在实时性图像匹配的应用中。

### 3.1.2. SURF

虽然SIFT算法已经被认为是最有效的，也是最常用的特征点提取的算法，但如果不借助于硬件的加速和专用图像处理器的配合，SIFT算法以现有的计算机仍然很难达到实时的程度。对于需要实时运算的场合，如基于特征点匹配的实时目标跟踪系统，每秒要处理8-24帧的图像，需要在毫秒级的时间内完成特征点的搜索、特征矢量生成、特征矢量匹配、目标锁定等工作。这样SIFT算法就很难适应这种需求了。SURF借鉴了SIFT中简化近似的思想，如果说SIFT算法中使用DOG对LOG进行了简化，提高了搜索特征点的速度，那么SURF算法则是对DoH的简化与近似。SURF把DoH中的高斯二阶微分模板进行了简化，使得模板对图像的滤波只需要进行几个简单的加减法运算，并且，这种运算与滤波器的尺度无关。实验证明，SURF算法较SIFT在运算速度上要快3倍左右。

下面总结一下SURF与SIFT的不同点，也正是因为这些不同点造成了SURF速度更快。

SURF对图像的近似与SIFT有很大不同。SIFT采用的是DOG图像，而SURF采用的是Hessian矩阵行列式近似值图像。我们可以利用Hessian矩阵行列式的极大值检测斑点。每个像素的Hessian矩阵：



特征点需要具备尺度无关性，所以在进行Hessian矩阵构造前，需要对其进行高斯滤波。这样，经过滤波后再进行Hessian的计算，其公式如下：



高斯平滑和求二阶导数这2种操作合在一起用一个高斯二阶微分模板代替。对高斯二阶微分模板进行简化，从而可以使用盒子滤波。同时，为了提高计算效率，还引入了积分图像的概念并可以应用于近似后的图像。一旦计算得到积分图，计算图像内任何矩形区域的像素值之和只需要三个加法。借助积分图像，图像与高斯二阶微分模板的滤波转化为对积分图像的加减运算，从而提高了速度。

SURF构建金字塔的方式也和SIFT不同。在SIFT算法中，不同组（octave）的图像大小是不一样的，下一组是上一组图像的降采样（1/4大小）；在每一组里面的几幅图像中，他们的大小是一样的，不同的是他们采用的尺度σ不同。而且在模糊的过程中，他们的高斯模板大小总是不变的，只是尺度σ改变。对于SURF算法，由于采用了盒子滤波和积分图像，所以，我们并不需要像SIFT算法那样去直接建立图像金字塔，而是采用不断增大盒子滤波模板的尺寸的间接方法。图像的大小总是不变的，改变的只是高斯模糊模板的尺寸，当然，尺度σ也是在改变的。对于SURF算法，由于采用了盒子滤波和积分图像，所以，我们并不需要像SIFT算法那样去直接建立图像金字塔，而是采用不断增大盒子滤波模板的尺寸的间接方法。图像的大小总是不变的，改变的只是高斯模糊模板的尺寸，当然，尺度σ也是在改变的。

SURF生成描述子的方法与SIFT不同。在SURF中，不统计其梯度直方图，而是统计特征点领域内的Harr小波特征。即以特征点为中心，计算半径为6(为特征点所在的尺度值)的邻域内，统计60度扇形内所有点在 (水平)和(垂直)方向的Haar小波响应总和(Haar小波边长取4)，并给这些响应值赋2的高斯权重系数，使得靠近特征点的响应贡献大，而远离特征点的响应贡献小。然后60度范围内的响应相加以形成新的矢量，Harr小波特征值反应了图像灰度变化的情况，那么这个主方向就是描述那些灰度变化特别剧烈的区域方向。在一个矩形区域来计算Haar小波响应。每个特征点就是16\*4=64维的向量，相比SIFT而言，少了一半，Haar小波响应也可以利用积分图像，这在特征匹配过程中会大大加快匹配速度。

### 3.1.3. KAZE

虽然SIFT，SURF等特征在图像特征提取方面已经取得比较好的成绩，但是这类特征都是基于一个通过高斯核进行的线性的尺度空间进行特征检测的，相同尺度下每个点的变换是一样的，由于高斯函数是低通滤波函数，会平滑图像边缘，以至图像损失掉许多细节信息。针对这一问题，ECCV2012，European Conference on Computer Vision(欧洲计算机视觉国际会议)中出现了一种比SIFT更稳定的特征检测算法KAZE[101]。作者提出了一种基于非线性尺度空间的特征点检测方法，该非线性尺度空间保证了图像边缘在尺度变化中信息损失量非常少，从而极大地保持了图像细节信息。

KAZE与SIFT、SURF最大的不同是建立了非线性的尺度空间。非线性尺度空间的构建主要基于非线性扩散滤波原理，非线性扩散滤波的基本公式是:



其中和分别表示散度和梯度。而函数表示扩散的传导函数，正是该函数的引入使得扩散能够适应图像局部特征。在这里该传导函数主要由梯度幅值控制，参数是尺度参数。

非线性滤波在图像方面的应用是由Perona 和 Malik 于1990年文章Scale-space and edge detection using annisotropic diffusion[102]提出来的，为了尽量使平滑在区域内而不是区域间，减少在边缘处的扩散，Perona提出了一个使得传导函数c依赖图像梯度幅值的函数：



其中是原始图像高斯平滑后的梯度。

由于非线性偏微分方程并没有解析解，只能通过一些数学方法近似求解该偏微分方程，本文主要使用AOS求解方案[103]。采用加性算子分裂算法(Additive Operator Splitting, AOS)来进行非线性扩散滤波，可以采用任意步长来构造稳定的非线性尺度空间。

文章采取的尺度空间的构建与SIFT尺度空间的构建比较相似，都是通过指数步长的系列组合(O个组，S个层)来离散化尺度空间的，但与SIFT不同的是文章中的各组图像都是对原图像的操作而不是对图像进行降采样，而且在同一尺度下各个像素点的变化是非线性的。各个层之间的尺度关系如下:



其中是基本尺度，表示组序号，表示所在的组的层序号，表示总的层数，为了进行非线性扩散滤波，需要将尺度空间中的尺度单元转换到时间单元概念，其转换如下：



然后根据AOS方法进行非线性扩散滤波得到每一层。

KAZE中特征点的检测与SIFT类似，而确定特征点主方向的方法与SURF类型，描述子的建立采用了M-SURF描述子。

KAZE的作者还提出了KAZE的加速版本AKAZE。AKAZE的得分和KAZE接近，但是AKAZE的速度加快了几个数量级。KAZE在构建非线性空间的过程中很耗时，在AKAZE中将Fast Explicit Diffusion(FED)加入到金字塔框架可以显著提升速度，且易于实施，更准确。在描述子方面，AKAZE使用了更高效的Modified Local Difference Binary(M-LDB)，从非线性尺度空间提取了梯度gradient和强度intensity信息。M-LDB是旋转和尺度不变的，并且对内存的要求更低。

### 3.1.4. ORB、

首先，ORB[104]算法来自于OpenCV 实验室，相比于SIFT和SURF，ORB在使用中不必担心专利的问题。但同时ORB在保证了一定性能的条件下做到了高效。在论文《ORB: An efficient alternative to SIFT or SURF》2011中，ORB在特征点检测和描述子生成方面都做了改进，结果是ORB比SIFT块两个数量级，比SURF块10倍，这样就不必借助GPU就可以实现算法的加速。

ORB的特征点检测和特征点的描述子分别使用了速度最快的FAST算法和BRIEF算法，但是又都针对他们的缺点进行了改善。对于FAST，使用图像金字塔方法实现了尺度不变性；使用Harris角点滤波器剔除检测到的边缘并对检测到的特征点评级（FAST算法对边缘敏感Large responses），可以挑出表现最好的特征点。FAST的一大缺陷是没有旋转不变性，在SIFT中通过特征点邻域的梯度直方图的统计得到方向信息，SURF则是利用Haar小波在邻域内进行近似计算。ORB的作者在Rosin对各种确定角点方向方法的分析基础上，选择了质心的方法。特征点的邻域是一个patch，那么patch一定会有一个质心(intensity weighted centroid)，特征点和质心一般不会重叠，即有一个offset，特征点到质心的连续构成的矢量就是该特征点的方向向量。至于patch到底是多大，文中没有明说，但我个人理解应该是和FAST算法时使用的半径大小是一致的。ORB作者在使用FAST算法时使用的是FAST-9，即半径大小是9个像素。

质心是由矩计算得到的。矩的计算公式：，那么质心。ORB基于BRIEF提出了两种具有主方向的描述子，一种是steer BRIEF，另外一种是rBRIEF。steer BRIEF使用旋转之前选择的tests坐标构成的矩阵与旋转矩阵相乘，得到旋转之后的tests。为了加快速度，steer ORB将角度离散化（间隔为12degrees），提前计算出了一个查找表。但是steer BRIEF改变了BRIEF原有的几个优良特性：1.BRIEF的二进制描述子方差大。2. BRIEF的二进制描述子的均值为0.5。

rBRIEF是对Steered BRIEF的改进，改进的目标是不仅要使得描述子对旋转不变，还要有大方差，0.5的均值，在tests之间还要保证低相关性。PCA可以从高维向量中提取出互不相关的特征向量，从而减少了冗余的信息。但是，二进制tests的空间过大，不便使用PCA，且为了得到256bits大小的无关的描述子，我们就必须在最初找到更高维的描述子，这就反而降低了效率。在ORB论文中，在所有可能的二进制tests中贪婪查找，找到满足三个条件的tests。在寻找最近邻点的时候，作者使用了multi-probe LSH局部哈希来提高效率。实验表明LSH比构建kd树还要快。

在outdoor dataset中，ORB优于SIFT和SURF，indoor set中也近似这样。SIFT在graffititype images中表现更好。

## 算法对比

# 第4章 基于SIFT算法的新算法

## 4.1 SIFT的简化

如果说SIFT算法中使用DOG对LOG进行了简化，提高了搜索特征点的速度，那么SURF算法则是对DoH的简化与近似。虽然SIFT算法已经被认为是最有效的，也是最常用的特征点提取的算法，但如果不借助于硬件的加速和专用图像处理器的配合，SIFT算法以现有的计算机仍然很难达到实时的程度。对于需要实时运算的场合，如基于特征点匹配的实时目标跟踪系统，每秒要处理8-24帧的图像，需要在毫秒级完成特征点的搜索、特征矢量生成、特征矢量匹配、目标锁定等工作。这样SIFT算法就很难适应这种需求了。SURF借鉴了SIFT中简化近似的思想，把DoH中的高斯二阶微分模板进行了简化，使得模板对图像的滤波只需要进行几个简单的加减法运算，并且，这种运算与滤波器的尺度无关。实验证明，SURF算法较SIFT在运算速度上要快3倍左右。SURF算法中要用到积分图像的概念。借助积分图像，图像与高斯二阶微分模板的滤波转化为对积分图像的加减运算。积分图像（Integral Image）的概念是由viola和Jones提出来的，而将类似积分图像用于盒子滤波是由Simard等人提出。

图像金字塔的构建是使用不同尺度因子的高斯卷积核对图像模糊来达到尺度不变性，卷积的计算量较大，计算量与图像的大小成正比。对于图像金字塔，计算量与金字塔的大小成正比。若能缩小金字塔的大小，就可以成比例地减少计算量。因为SIFT算法在构建金字塔的过程中使用了下采样，这里提前对实验图像进行下采样的预处理，考虑到下采样的插值方法和分别对待配准图像下采样还是同时对2幅图像下采样的情况进行了讨论。

若是单独对待配准图像进行下采样的预处理，在拟合得到投影变换模型之后，可以直接将变换作用在预处理之后的图像上，从而得到配准之后的图像；若是对待配准图像和参考图像都进行了下采样的预处理，则需要根据下采样的比例系数对拟合出的变换模型进行校准，得到适用于原图的变换关系。设任意1幅图像放大后的目标图像的宽为*D*w，高为*D*h，原图像宽为*S*w，高为*S*h，则目标像素点在原图像上对应的位置为: ，。当下采样比例为*s*时，待配准图像到参考图像的变换矩阵。因此，由中间变量进行修正可得到最终待配准图像与参考图像之间的变换矩阵，利用和插值（参考Rob Hess源码采用线性插值）可以实现配准[105]。

## 4.2 SIFT与显著性检测

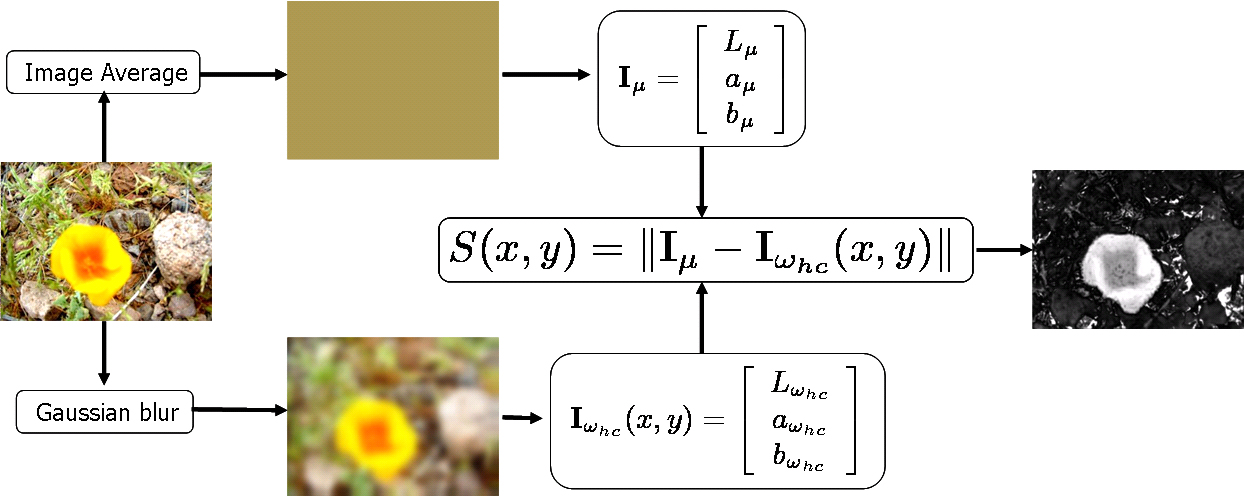
SIFT算法的关键是检测出具有不变性，且可以代表图像的关键信息的特征点，依靠两幅图像之间的特征点进行匹配，拟合出某个策略下的最优变换模型，所以特征点的质量和数量都很重要。

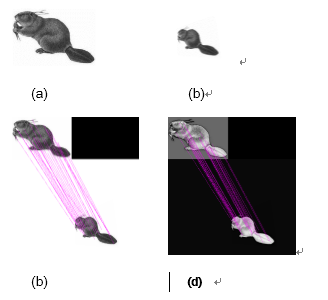
显著性区域检测方法模仿了人眼优先锁定感兴趣重要区域的特点，可以应用在目标检测和识别中，一定程度上减少搜索的工作量。将显著性检测方法和SIFT算法结合起来，前者作为后者的预处理，可以在一定程度上促使SIFT算法检测出更加稳定，高质量的特征点，从而有利于后续的变换模型估计和最终的配准效果。

显著性区域检测方法有多种，比如X Hou和L Zhang通过分析输入图像的对数谱来消除背景[106]，Yun Zhai提出了一种关于时空视频的注意点的技术，而Achanta的方法输出具有明确界定的显著性物体的全分辨率显著性图[107]。SIFT还有一个缺陷是算法的使用对象是灰度图像，即在图像处理的过程中丢失了图像的彩色信息。Frequency-tuned显著性检测方法使用了CIRLab颜色模型，利用了颜色和亮度的特征，并且易于实现。原图中每个像素的显著性的值表现为CIRLab颜色空间上当前像素和全局平均值的距离：



是图像特征均值向量,是高斯平滑之后的像素点(用的独立二值核) ，是二范数。FT算法可分为四个部分。 原始图像上的高斯模糊; 获取原始图像的LAB模型版本; 计算CIELab向量的平均值; 计算像素和平均值之间差值的L2范数。 完整过程如图所示。

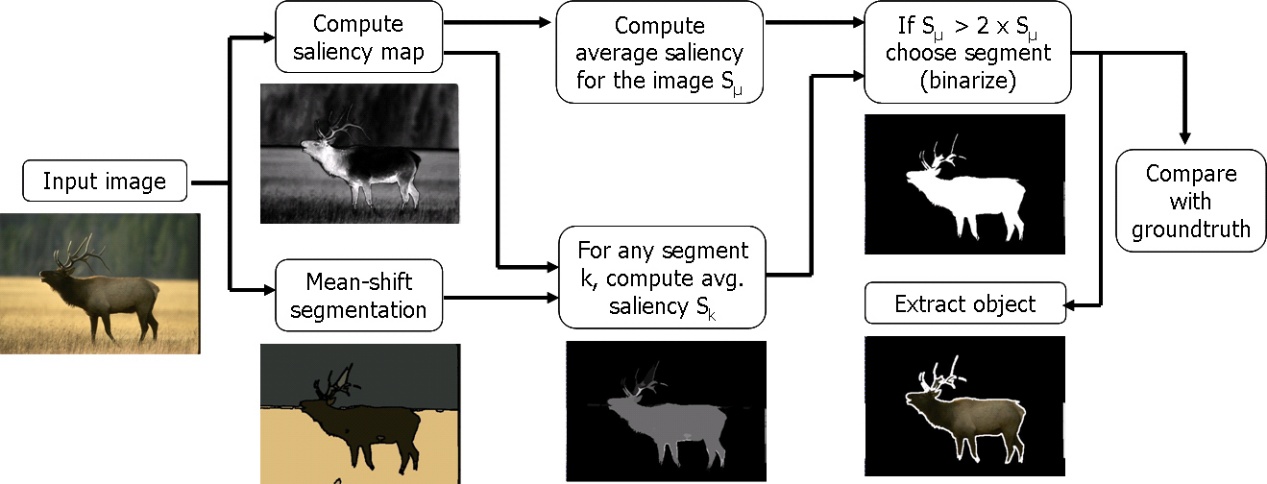




|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | N1 | N2 | Time/ms | PSNR/dB | RMSE |
| Original image | 116 | 95 | 1597 | 38.719 | 18.381 |
| saliency map | 353 | 167 | 2140 | 43.920 | 0.556 |

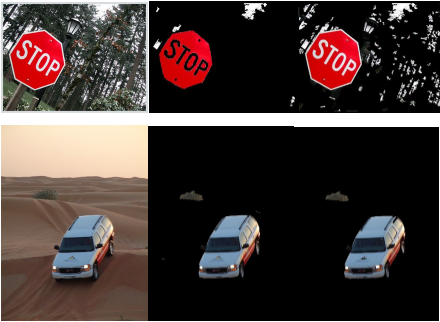
从实验结果可以看出，基于显著性图的SIFT检测获得了更多的特征点，实现的配准结果也提高了约5dB，而RMSE从18.381减少到0.556，代价是减速。

基于FT的显著性检测方法还可以实现目标物体的分割。分割原理如下图所示：

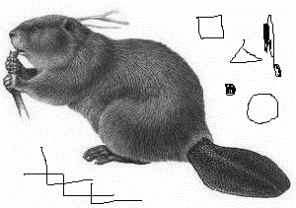


当我们的配准目标也是图像中的待检测分割目标时，可以将这一图像分割方法应用在图像配准中，从而减少图像中背景噪声对于特征点检测的影响，降低无意义的特征点的数量。本文对于分割算法做了一些改进，主要思路是使用原图的显著性均值对图像进行填充（这里选择将原图面积扩展为原图的四倍），使得图像中主体目标与填充后得到的新图整体的显著性的值差距更大，从而有利于后续的分割。实验表明，这种改进之后的分割方法可以得到更加完整的目标，即召回率更高。因为本文以图像配准为研究对象，在分割之后以其特征点数目等为指标。下面是实验结果：





可以看到填充背景后的分段保留了更多细节。 与原始方法相比，修改方法的结果更加完整。



三种分割方法的性能。 （a）是人工创建涂鸦背景的图像，（b）是直接分割的效果，（c）是用白色像素填充背景后的分割效果，（d）是分割后的分割效果， 背景填充了平均值。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **(a)** | **(b)** | **(c)** | **(d)** |
| **N** | 144 | 116 | 124 | 117 |
| **RMSE** | 16.483 | 1.476 | 12.651 | 6.460 |
| **Time/ms** | 1849 | 1615 | 2258 | 1622 |
| **PSNR/dB** | 36.140 | 25.616 | 15.851 | 27.098 |

## 4.4 本章小结

SURF (Speeded Up Robust Features)也是一种类似于SIFT的兴趣点检测及描述子算法。其通过Hessian矩阵的行列式来确定兴趣点位置，再根据兴趣点邻域点的Haar小波响应来确定描述子，SURF 的描述子基于 2D 离散小波变换响应Haar小波并且有效地利用了积分图。其描述子大小只有64维（也可以扩展到128维，效果更好），是一种非常优秀的兴趣点检测算法。SURF算法的概念及步骤均建立在SIFT之上，但详细的流程略有不同。

# 第5章 配准算法的应用

在向服务器上传或者向本地拷贝图像时，我们不希望无限制地允许图像集增大，对于重复性的图像我们希望保证其单一性，而不是无谓地占用内存空间，但直接比较两个图像的相似性有时过于复杂，一种可取的方法就是提前计算得到图像的哈希值，对于两个图像各自的哈希值的汉明距离为0，就认为两个图像完全一样；当汉明距离不为0但较小时，认为两幅图像相似性程度较高，这时可以利用配准算法估算出二者之间的变换关系，为了便于之后可能的复原，还可以记录待删除图像（即配准过程中的参考图像）与配准之后的图像的差值图像和其对于像素点的正负性。

## 5.1 哈希算法

哈希算法，又叫散列算法，本质上就是一种映射关系，通过约定的映射函数将原文本映射为较短的固定长度的二进制值。

加密哈希算法，MD5,SHA1,SHA256，在Python中导入hashlib模块，再调用函数就可以生成字符串或者图像对应的哈希值，但是加密哈希有一个特点就是文件一旦发生改变，即便变化很小，哈希值都会发生巨大的变化。这个特点在防止用户修改文件时很有帮助，这也是加密哈希名称的由来。算法还有一个特点，就是不可逆，不可逆根据哈希值反向推出原信息，这也有利于加密，网站保存我们的密码就不会直接保存原始密码，而是保存其哈希值，这样即使网站被黑，依然不会泄露密码。

但同时在很多时候我们希望相似图像或者相似文本具有相同的哈希值，这时候我们就需要采用其他的哈希算法。比较简单、易用的解决方案是采用感知哈希算法（Perceptual Hash Algorithm)。感知哈希算法是一类算法的总称，包括aHash（平均值哈希）、pHash（感知哈希）、dHash（差异值哈希）。顾名思义，感知哈希不是以严格的方式计算Hash值，而是以更加相对的方式计算哈希值，因为“相似”与否本来就是相对的而不是绝对意义上的。

平均值哈希方法计算简单，只需在归一化图像尺寸后，比较每个像素与整体灰度图像的灰度均值，从而就可以得到0、1的二进制字符串，作为原图像的哈希值。但是平均哈希算法的准确度不高，因为它只记录了每个像素与均值的大小情况。这里采用了一种称为差异值哈希的算法，在保证快速的同时，具有良好的精确性。

差异值哈希与平均值哈希大体相同，不同的是差异值哈希比较的是同一行相邻的两个像素的灰度值，得到0、1比特。不同图像的哈希值应该是相同的长度，这里采用的都是长度为16的十六进制字符串，转换成二进制就是长度为64的二进制字符串。因为灰度值比较的参考值不同，平均值哈希算法中将图像归一化为8x8的大小，差异值哈希算法将图像的宽归一化为9，将高归一化为8，这样最后按行拼接得到的二进制字符串长度都是64。除了归一化最终的哈希值长度，缩放图像的另外一个目的是对原图模糊处理，避免过多图像细节和噪声的影响，从更加宏观和整体的角度比较图像，同时较少计算量。

在信息论和密码学中，汉明距离是一种常用的距离和相似度的衡量方式，比较两个相同长度的二进制字符串对应字符，记录不相同的字符的总数作为汉明距离的值。在计算机中，汉明距离更加具有高效性，通过简单的异或运算就可以得到汉明距离的值。相比于加密哈希，差异值哈希方法对图像微小变化的鲁棒性更强，相似性越高的图像其哈希值之间的汉明距离越小。

## 5.2 图像减法

两幅图像之间的减法指的是两幅大小相同的图像按照对应像素做减法运算，可以简单地得到两幅图像之间像素级别的差异信息。两幅尺寸大小相同的图像做减法，本质上是矩阵的减法，只需保证矩阵的大小一致。在图像配准中，配准后的图像与参考图像就满足尺寸相同的特点，而且这两幅图像之间有相当一部分是重合的，配准的性能越好，重合度也就越高。利用图像减法可以去除冗余的重合部分，只记录差异的部分，从而减少信息的存储量，同时在需要原有图像时，也可以复原图像。而两幅图像的重合度或者说相似性可以利用差异值哈希算法来衡量，先分别计算两幅图像各自的哈希值，然后通过两个哈希值之间的汉明距离表示两幅图像的相似度。

但是需要注意的是，像素之间的差值可能为正也可能为负，为了将得到的差值转换为一幅新的差值图像，需要求取差值的绝对值，同时记录每个差值的正负性，形成一个新的掩膜图像，便于日后对于原图的复原。对于差值图像，如果配准后的图像与参考图像的重合度越高，差值图像中像素灰度值为0的部分会显著增多，即新的矩阵会变成一个稀疏矩阵，有利于做图像压缩，进一步减少图像占用的空间。

## 5.2 图像压缩

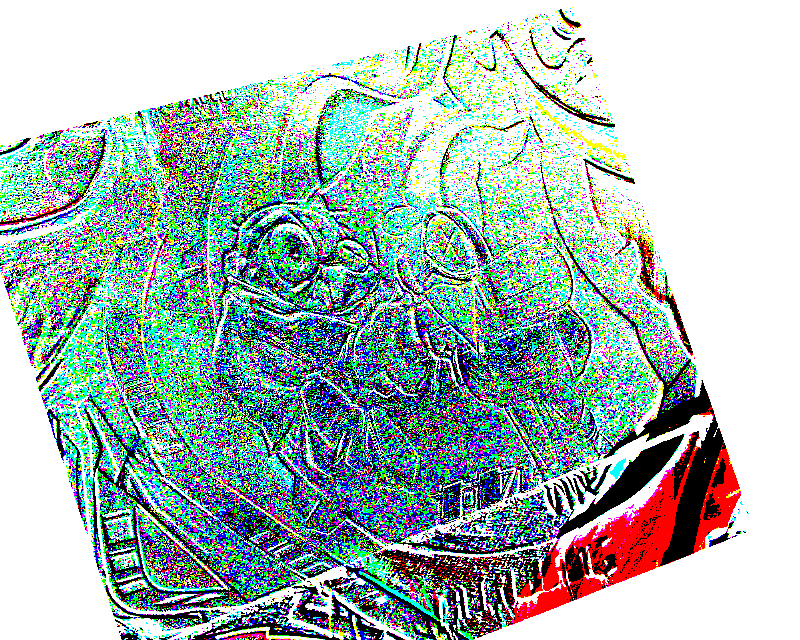
压缩算法是一种减少数据量的算法，在声音，视频等多媒体介质和传输中应用广泛。图像压缩是一种在尽量保证图像信息不受损失前提下，减少图像数据量的算法。图像压缩一般建立在编码冗余、像素间冗余和心理视觉冗余上。编码冗余是指编码技术没有使码元达到理论上可以表示的最大信息量，最佳编码应该使得每个像素所需的平均比特数最少；像素间冗余是由像素间相关性引起的，包括空间和时间（视频序列中）相关性，即不同像素表示的信息之间有重合；心理视觉冗余是指由于人眼的局限性，无法察觉一些失真和变化，那么这一部分就可以视作是冗余的。图像压缩系统一般包括两部分：编码器和解码器，分别完成压缩和解压缩这两个互补的操作。图像压缩是数字图像处理技术中用处最广泛和商业上最成功的技术之一，商业上的成功反过来推动了技术的演变。形成了一些著名的图像压缩上的工业和商业标准（也被称为图像格式、容器），规定了数据的排列方式和压缩类型及压缩过程。常见的标准有JPEG、JPEG2000、BMP、GIF、PNG等，其根据是否可以完全恢复原图像又可分为无损压缩、有损压缩。

## 5.3 实验





两个图像哈希值之间的汉明距离是15，说明有80%的内容是相近的，所以如果同时保存两张图像会在一定程度上造成信息冗余。对于明显存在某种变换的图像，可以使用图像配准拟合出二者之间的关系，用配准后的图像在一定程度上替换参考图像，甚至直接用拟合后的变换矩阵表示冗余的参考图像：变换矩阵与保留的待配准图像可以恢复参考图像中的主体部分，对于配准产生的误差与参考图像无法恢复的部分，借助图像减法记录这一部分信息。





# 第6章 总结与展望

图像配准在图像处理中是极具应用价值的一项技术，关系到后续的各种图像处理如图像融合，图像拼接，影响着医疗、遥感、军事等领域。图像配准一般指两幅图像之间的操作，一幅被称为待配准图像，另外一幅是参考图像。两幅图像之间的内容是同一个场景，但是存在着因为视角变化、光照变化等产生的形变。图像配准的目的就是找到一个恰当的模型来描述这种形变，从而将变形的图像校正到参考图像的形态。图像配准算法大致可以分为以灰度信息为基础的配准方法和基于局部特征的方法。本文的研究重点在基于特征点的图像配准，从两幅图像中具有旋转、平移、尺度不变性的特征点之间拟合出两幅图像之间的空间变换关系。本文研究了基于特征点配准算法的关键步骤和相关的技术，针对其不足之处做出了改进。

本文主要工作可以概括如下：

1. 从图像配准的背景和意义出发，从数学的角度严谨地认识了图像配准的原理，对各个环节有了充分的认识。按照文献，将图像配准分成三类，针对每一类介绍了其技术路线和国内外研究现状。
2. 研究了与图像相关的基本技术。从图像的成像模型到图像坐标系，到描述图像变换的模型和图配准的评价标准。估算出的模型的准确性直接关系到配准的效果。
3. 介绍了几种流行的基于特征点的配准算法如SIFT、SURF、KAZE，并研究了其关键的步骤。几种算法一脉相承又各具特色，编写代码对几种算法进行了比较。实验了几种算法在各种变换如旋转、平移、噪声下的配准性能。
4. 基于SIFT算法提出了两种改进的算法。一种是使用下采样减小图像金字塔的规模，降低了图像配准在特征点提取中的计算复杂度。一种使用了显著性检测的方法，弥补了SIFT算法没有利用图像灰度信息的特点，可以检测处的特征点在数量和质量上都有了提升，拟合出的变换模型质量也更高。将基于显著性检测的图像分割方法应用在图像配准中，并对其分割方法做出了改进，可以保留更多的主体信息。
5. 针对图像数据集中可能存在图像内容相似的图像，即存在适用于图像配准的图像族，提出了应用图像配准减小图像集的方法。方法首先计算图像的哈希值找出合适于图像配准的图像，记录图像与配准后图像的差异信息就可以舍弃该图像，从而减小占据的存储空间。

图像配准其实是一个很大的课题，涉及到的知识点众多。由于个人水平和时间有限，没有能做出特别满意的工作，未来还有很长的路要走。本文的待改进之处有：

1. 文章提出的一些思路还停留在实验阶段，各种优化工作做得不够完善，距离真正的实验还有一段距离。
2. 对图像配准算法本身的改进有限。现在图像算法的发展已经广泛使用了深度学习，如何将神经网络和深度学习应用在图像配准方向有待发掘。

# 参考文献

1. Hubel, David H., and Torsten N. Wiesel. "Receptive fields of single neurones in the cat's striate cortex." The Journal of physiology148.3 (1959):574-591.
2. Wiesel, Torsten N. "The postnatal development of the visual cortex and the influence of environment." Stockholm: Nobel Foundation (1981).
3. Rafael C. Gonzalez, Richard E. Woods (阮秋琦，阮宇智等译). 数字图像处理.第二版.北京:电子工业出版社,2003.pp:4-560.
4. 陈晓曦,王延杰,刘恋.小波阈值去噪法的深入研究[J].激光与红外,2012,42(01):105-110.
5. 易三莉,张桂芳,贺建峰,李思洁.基于最大类间方差的最大熵图像分割[J].计算机工程与科学,2018,40(10):1874-1881.
6. 王浩,梁煜,张为.一种均匀化稀疏表示的图像压缩感知算法[J/OL].西安交通大学学报,2019(02):1-6[2019-01-03].http://kns.cnki.net/kcms/detail/61.1069.T.20181206.1718.004.html.
7. T. F. Cootes. C. J. Taylor．Statistical models of appearance for computer vision：World Wide Web Publication，2004
8. Zhang K , Zuo W , Chen Y , et al. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising[J]. IEEE TRANSACTIONS ON IMAGE PROCESSING.
9. Schaffalitzky F, Zisserman A. Multi-view Matching for Unordered Image Sets, or “How Do I Organize My Holiday Snaps?”[J]. Proc.european Conf.comput.vision May Copenhagen Denmark, 2002, 2350:414-431.
10. Tuytelaars T , Gool L V . Matching Widely Separated Views Based on Affine Invariant Regions[J]. International Journal of Computer Vision, 2004, 59(1):61-85.
11. Zitova B, Flusser J. Image registration methods :a survey[J].Image and Vision Computing. 2003,21(11),pp:977-1000.
12. 卞贤掌,费海平,李世强.基于语义分割的增强现实图像配准技术[J].电子技术与软件工程,2018(23):79-82.
13. 付浩威. 基于全变分约束的医学图像配准研究[D].哈尔滨工业大学,2018.
14. 耿庆田,于繁华,王宇婷,赵宏伟,赵东.基于SIFT的车标识别算法[J].吉林大学学报(理学版),2018,56(03):639-644.
15. 陶筱娇.SIFT算法下的多表情人脸识别[J].信息与电脑(理论版),2018(23):127-128+131.
16. Lisa Gottesfeld Brown . A Survey of Image Registration Techniques . ACM Comput,ng Surveys[J] , 1992 , 24(4):325~376 .
17. Brown M , Lowe D G . Recognising panoramas[C]// IEEE International Conference on Computer Vision. IEEE Computer Society, 2003.
18. Gottesfeld Brown L . A survery of image registration techniques[J]. Acm Computing Surveys, 1992, 24(4):325-376.
19. 谢凤英 赵丹培. VISUAL C++数字图像处理(附光盘)[M]. 电子工业出版社, 2008.
20. 孙雅琳, 王菲. 基于边缘和互信息的红外与可见光图像配准[J]. 电子科技, 2010, 23(4):80-82.
21. 张强. 图像匹配算法研究[D]. 西安电子科技大学, 2006.
22. 燕磊, 孙启湲, 朱菲菲. 基于Oriented FAST和MAD的特征匹配图像自动配准[J]. 计算机与数字工程, 2017(09):179-182.
23. 郭龙源, 夏永泉, 杨静宇. 一种改进的彩色图像匹配算法[J]. 计算机工程与应用, 2007, 43(27):98-99.
24. 沈晓芳, 向建勇, 董维科. 基于图像边缘特征的SSDA算法[J]. 电子科技, 2009, 22(3):16-18.
25. 张维琪, 樊斐. 自适应SSDA图像匹配并行算法设计与实现[J]. 计算机工程与应用, 2014, 50(20):64-67.
26. 贾凯, 曲仕茹. 改进的SSDA图像匹配算法[J]. 测控技术, 2012, 31(10):47-50.
27. Wang Y , Wu D . An improved adaptive SSDA based on wavelet-pyramid[C]// American Institute of Physics Conference Series. AIP Publishing LLC, 2017.
28. （希）西奥多里德斯 等著，李晶皎 等译.模式识别. 电子工业出版社. 2006-12
29. T. F. Cootes. C. J. Taylor．Statistical models of appearance for computer vision：World Wide Web Publication，2004
30. Anuta P E. Spatial registration of multispectral and multitemporal digital imagery using fast Fourier transformation techniques. IEEE Trans. on Geosci. Electron. 1970.8(4),pp:355-368.
31. Bamea D I. Silverman H F. A class of algorithms for fast digital image registration.

IEEE Trans.on Computers.1972,21(2),pp:179-186.

1. Pratt W K(高荣坤等译).数字图像处理学.北京:科学出版社.1984.9.
2. Svedlow M, Mac Gillen C D, Anuta P E. Image registration:similarily measure and preprocessing method comparisons. IEEE Trans. on Aerospace and Electronic Systems.1978,AES-14(1),pp:141-150.
3. Rignot E, Kwok R, Curlander Jeta1. Automate multisensor registration: requirments and techniques. Photogrammetric Eng. Remote Sensing.1991,57(8),pp:1029-1038.
4. Flussr Jan. An adaptive method for image registration. Pattern Recognition. 1992. 25(1),pp:45-54.
5. P.A. Viola, Alignment by maximization of mutual information, Ph.D. thesis, Massachusetts Institute of Technology,Boston,MA,USA,1995
6. A.Collignon,F.Maes,D.Delaere,D.Vandermeulen,P.Suetens,andG.Marchal,Automed multi-modality image registration based on information theory, Information Processing in Medical Imaging, Y. Bizais, C. Barillot, and R. Di Paola,Eds.1995,Kluwer Academic Publishers,1997,187-198
7. 马政德，基于互信息的图像配准并行算法研究与实现，国防科技大学工学硕士论文，2007
8. 曹蹊渺，基于互信息的图像配准算法研究，北京交通大学硕士学位论文，2007
9. 曹源，武艳超，基于粒子群优化算法和烟花算法的图像配准[J].Internet Herth.2017:152-153
10. 靳济芳.VC++小波变换与工程实践.北京:人民邮电出版社.,2004, 第一章
11. S.Mallat. A Theory for Multiresolution Signal Decomposition: the Wavelet Representation.IEEE TRANSACtions on Pattern Analysis and Machine Intelligence,1989,11(7):674-693
12. 伍君，基于小波变换的图像配准方法研究，湖南大学硕士学位论文，2005
13. 夏召红，基于小波变换的图像配准，辽宁科技大学硕士学位论文，2008
14. Viola P.; Wells WM.. Alignment by Maximization of Mutual Information[J]. International Journal of Computer Vision , 1997 , 24 (2 ):137-154 .
15. Reddy B S, Chatterji B N. An FFT-based technique for translation, rotation, and scale-invariant image registration[J]. Image Processing, IEEE TRANSACtions on, 1996, 5(8): 1266-1271.
16. Chen Q, Defrise M, Deconinck F. Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition[J]. Pattern Analysis and Machine Intelligence, IEEE TRANSACtions on, 1994, 16(12): 1156-1168.
17. Hough P V C. Method and means for recognizing complex patterns: U.S. Patent 3,069,654[P]. 1962-12-18.
18. Lin M C, Canny J F. A fast algorithm for incremental distance calculation[C] // Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on. IEEE, 1991: 1008-1014.
19. H. Gonçalves; J. A. Gonçalves; L. Corte-Real; A. C. Teodoro. CHAIR: automatic image registration based on correlation and Hough transform [J]. International Journal of Remote Sensing , 2012 , 33 (24 ):7936-7968 .
20. Troglio, G.; Le Moigne, J.; Benediktsson. Automatic Extraction of Ellipsoidal Features for Planetary Image Registration [J]. Geoscience and Remote Sensing Letters, IEEE , 2012 , 9 (1):95-99 .
21. Hu M K. Visual pattern recognition by moment invariants[J]. Information Theory, IRE TRANSACtions on, 1962, 8(2): 179-187.
22. Khotanzad A, Hong Y H. Invariant image recognition by Zernike moments[J]. Pattern Analysis and Machine Intelligence, IEEE TRANSACtions on, 1990, 12(5): 489-497.
23. Dai X, Khorram S. A feature-based image registration algorithm using improved chain-code representation combined with invariant moments[J]. Geoscience and Remote Sensing, IEEE TRANSACtions on, 1999, 37(5): 2351-2362.
24. Flusser J. Moment invariants in image analysis[C]//proceedings of world academy of science, engineering and technology. 2006, 11(2): 196-201.
25. Flusser J, Suk T. Rotation moment invariants for recognition of symmetric objects[J]. Image Processing, IEEE TRANSACtions on, 2006, 15(12): 3784-3790.
26. Matas J, Chum O, Urban M, et al. Robust wide-baseline stereo from maximally stable extremal regions[J]. Image and vision computing, 2004, 22(10): 761-767.
27. Murphy-Chutorian E, Trivedi M M. N-tree Disjoint-Set Forests for Maximally Stable Extremal Regions[C]//BMVC. 2006: 739-748.
28. Kristensen F, Mac Lean W J. Real-time extraction of maximally stable extremal regions on an FPGA[C]//Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on. IEEE, 2007: 165-168.
29. Tuytelaars T, Van Gool L. Matching widely separated views based on affine invariant regions[J]. International journal of computer vision, 2004, 59(1): 61-85.
30. Kadir T, Brady M. Saliency, scale and image description[J]. International Journal of Computer Vision, 2001, 45(2): 83-105.
31. Kadir T, Zisserman A, Brady M. An affine invariant salient region detector [M]//Computer Vision-ECCV 2004. Springer Berlin Heidelberg, 2004: 228-241.
32. Mikolajczyk K, Tuytelaars T, Schmid C, et al. A comparison of affine region detectors[J]. International journal of computer vision, 2005, 65(1-2): 43-72.
33. Mikolajczyk K, Schmid C. An affine invariant interest point detector[M]//Computer Vision—ECCV 2002. Springer Berlin Heidelberg, 2002: 128-142.
34. Mikolajczyk K, Schmid C. Scale & affine invariant interest point detectors[J]. International journal of computer vision, 2004, 60(1): 63-86.
35. Schaffalitzky F, Zisserman A. Multi-view matching for unordered image sets, or “How do Iorganize my holiday snaps?”[M]//Computer Vision—ECCV 2002. Springer Berlin Heidelberg, 2002: 414-431.
36. HansP M. Towards automatic visual obstacle avoidance[C]//Proceedings of the 5th international joint conference on Artificial intelligence. 1977: 584-584.
37. Harris C, Stephens M. A combined corner and edge detector[C]//Alvey vision conference. 1988, 15: 50.
38. Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion[J]. Pattern Analysis and Machine Intelligence, IEEE TRANSACtions on, 1990, 12(7): 629-639.
39. Lindeberg T. Scale-space theory: A basic tool for analyzing structures at different scales[J]. Journal of applied statistics, 1994, 21(1-2): 225-270.
40. Lindeberg T, Gårding J. Shape-adapted smoothing in estimation of 3-D shape cues from affine deformations of local 2-D brightness structure[J]. Image and vision computing, 1997, 15(6): 415-434.
41. Lindeberg T. Edge detection and ridge detection with automatic scale selection[J]. International Journal of Computer Vision, 1998, 30(2): 117-156.
42. Lindeberg T. Feature detection with automatic scale selection[J]. International journal of computer vision, 1998, 30(2): 79-116.
43. Mikolajczyk K, Schmid C. An affine invariant interest point detector[M]//Computer Vision—ECCV 2002. Springer Berlin Heidelberg, 2002: 128-142.
44. Lowe D G. Object recognition from local scale-invariant features[C]//Computer vision, 1999. The proceedings of the seventh IEEE international conference on. Ieee, 1999, 2: 1150-1157.
45. Lowe D G. Local feature view clustering for 3D object recognition[C]//Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. IEEE, 2001, 1: I-682-I-688 vol. 1.
46. Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International journal of computer vision, 2004, 60(2): 91-110.
47. Bay H, Tuytelaars T, Van Gool L. SURF: Speeded up robust features[M]//Computer Vision–ECCV 2006. Springer Berlin Heidelberg, 2006: 404-417.
48. Ke Y, Sukthankar R. PCA-SIFT: A more distinctive representation for local image descriptors[C]//Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on. IEEE, 2004, 2: II-506-II-513 Vol. 2.
49. Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. Pattern Analysis and Machine Intelligence, IEEE TRANSACtions on, 2005, 27(10): 1615-1630.
50. Morel J M, Yu G. ASIFT: A new framework for fully affine invariant image comparison[J]. SIAM Journal on Imaging Sciences, 2009, 2(2): 438-469.
51. Rublee E,Rabaud V,Konolige K G,Bradski J R. ORB: an Efficient Alternative to SIFT or SURF[C]//IEEE International Conference on Computer Vision .Barcelona，Spain，2011: 2564-257
52. Abdel-Hakim A E,Farag A A.CSIFT:A SIFT Descriptor with Color Invariant Characteristics [C]//IEEE.Conference on Computer Vision and Pattern Recognition.New York,2006:1978-1983
53. Amintoosi M,Fathy M,Mozayani N,A Fast Image Registration Approach Based on SIFT Key-Points Applied to Super-Resolution[J].The Imaging Science Jounral,2012,60(4):185-201
54. Vinividyadharan, Subusurendran,et al. Automatic Image Registration Using SIFT-NCC [J].Ijca Special Issue on Advanced Computing and Communication Technologies for Hpc Applications,2012,ACCTHPCA(4):29-32
55. RostenE,Drummond T.Machine learning for high-speed cornerd etection [M] // Computer Vision–ECCV 2006. Springer Berlin Heidelberg, 2006: 430-443.
56. Rosten E, Porter R, Drummond T. Faster and better: A machine learning approach to corner detection[J]. Pattern Analysis and Machine Intelligence, IEEE TRANSACtions on, 2010, 32(1): 105-119.
57. Taylor S, Rosten E, Drummond T. Robust feature matching in 2.3 µs[C]//Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on. IEEE, 2009: 15-22.
58. Taylor S, Drummond T. Binary histogrammed intensity patches for efficient and robust matching[J]. International journal of computer vision, 2011, 94(2): 241-265.
59. Alcantarilla P F, Bartoli A, Davison A J. KAZE features[M]//Computer Vision–ECCV 2012. Springer Berlin Heidelberg, 2012: 214-227.
60. 李晖晖，郑平，杨宁，胡秀华. 基于SIFT 特征和角度相对距离的图像配准算法. [A]. 西北工业大学学报.2017.4.35(2):281-285
61. Pan J P, Hao J M, Zhao J P. Improved algorithm based on SURF for image registration[J]．Remote Sensing for Land and Resources，2017，29 ( 1 ) :110 -115．
62. 侯坚,张明,一种蚁群优化的改进SIFT特征点的图像配准算法[J].现代计算机 图形图像.2016.7
63. An Yong, Zheng Hong. SIFT Matching Method Based on K Nearest Neighbor Support Feature Points[J]. 2016 IEEE International Conference on Signal and Image Processing:64-68
64. Zhe Xiao, Linfang Yu, Zhen Qin, Huaigui Ren, Zhenwei Geng. A Point Matching Algorithm for Brain CT Images Based on SIFT and Gray Feature.[J] IEEE International Conference on Signal Processing 2016
65. 佚名. Three.js开发指南[M]. 2015.
66. Hill F , Kelley S . 计算机图形学:OpenGL版[M]. 清华大学出版社, 2009.
67. Bouchiha R, Besbes K. Automatic remote-sensing image registration using SURF[C]. 2010 The 3rd International Conference on Machine Vision (ICMV 2010), 2010: 406-410.
68. 阮芹彭刚李瑞. 基于特征点的图像配准与拼接技术研究[J]. 计算机与数字工程, 2011, 32(2): 141-144.
69. STUDHOLME C, HILL L G, HAWKES D. An overlap invariant entropy measure of 3D medical image align-ment［J］.Pattern Recognition，1988，32( 1) : 43-46．
70. Pablo F.Alcantarilla, Adrien Bartoli, Andrew J.Davision .KAZE Features .[J] ECCV.2012
71. P Perona , J Malik. Scale-Space and Edge Detection Using Anisotropic Diffusion. IEEE Computer Society , 1990 , 12 (7) :629-639
72. J.Weicket; B.M. T. HRomeny;M.A. Viergver. Efficient.Efficient and reliable schemes for nonlinear diffusion filtering. IEEE TRANSACtions on Image Processing.1998
73. Rublee E , Rabaud V , Konolige K , et al. ORB: An efficient alternative to SIFT or SURF[J]. 2011.
74. 张晨光,周诠,回征.基于SIFT特征点检测的低复杂度图像配准算法[J].扬州大学学报(自然科学版),2018,21(04):52-56.
75. Hou X, Zhang L. Saliency Detection: A Spectral Residual Approach[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2007:1-8.
76. Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection[C]// Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009:1597-1604.
77. L. G.. A survey of image registration techniques. ACM Computer, Surveys, Dec. 1992, Vol. 24, No. 4, p325-376
78. Bouchiha R, Besbes K. Automatic remote-sensing image registration using SURF[C]. 2010 The 3rd International Conference on Machine Vision (ICMV 2010), 2010: 406-410.
79. 阮芹彭刚李瑞. 基于特征点的图像配准与拼接技术研究[J]. 计算机与数字工程, 2011, 32(2): 141-144.
80. opencv2.3改用的Rob Hess的源码http://robwhess.github.io/openSIFT/, 牛津大学开发的VLFeat open source library http://www.vlfeat.org/
81. H.Bay, T. Tuytelaars, L. V. Gool, SURF:Speed Up Robust Features[J],ECCV,2006
82. Pablo F.Alcantarilla, Adrien Bartoli, Andrew J.Davision .KAZE Features .[J] ECCV.2012
83. P Perona , J Malik. Scale-Space and Edge Detection Using Anisotropic Diffusion. IEEE Computer Society , 1990 , 12 (7) :629-639
84. J.Weicket; B.M. T. HRomeny;M.A. Viergver. Efficient.Efficient and reliable schemes for nonlinear diffusion filtering. IEEE TRANSACtions on Image Processing.1998
85. M.A. Fischler and R.C. Bolles, “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography,” Comm. ACM, vol. 24, no. 6, pp. 381- 395,1981.
86. 李航.统计学习方法.清华大学出版社.2012.3.1
87. Jon Louis Bentley. Multidimensional binary search trees used for associative searching. Communications of the ACM. 1975.:509-517
88. Jeffrey S. Beis, David G. Lowe. Shape Indexing Using Approximate Nearest-Neighbour Search in High-Dimensional Spaces. IEEE Computer Society Conference on Computer Vision , 1997 :1000-1006
89. Deriche R, Giraudon G. A Computational Approach for Corner and Vertex Detection. International Journal of Computer Vision, 1993, 10(2): 101-124.
90. Schmid C, Mohr R and Baucklage C. Comparing and Evaluating Interest Points[J]. International Journal of Computer Vision, 2000, 37(2): 151-172.
91. STUDHOLME C, HILL L G, HAWKES D. An overlap invariant entropy measure of 3D medical image align-ment［J］.Pattern Recognition，1988，32( 1) : 43-46．
92. 余先川, 吕中华, 胡丹.遥感图像配准技术综述[J].光学精密工程, 2013, 21(11):2960-2972. Yu Xian-chuan,Lü Zhong-hua,Hu Dan. Review of remote sensing image registration technique[J].Optics and Precision Engineering, 2013, 21(11):2960-2972.
93. 董强,刘晶红,周前飞. 基于改进SURF的图像配准算法及其在遥感图像拼接中的应用[J/OL]. 吉林大学学报(工学版)：1-10.(2017-03-09).
94. Yin Wen Dong,Luan Wan,Zhao Ming Shi,Ming Lei Zhu. An Image Registration Algorithm Based on Improved SIFT Feature[J]. Applied Mechanics and Materials,2013,2560(347):.
95. 任克强,胡梦云. 基于改进SURF算子的彩色图像配准算法[J]. 电子测量与仪器学报,2016,30(05):748-756.
96. Prerana Mukherjee,Brejesh Lall. Saliency and KAZE features assisted object segmentation[J]. Image and Vision Computing,2017,:.
97. 耿则勋,徐志军,卢兰鑫,沈忱,曾德佳. 改进的视角鲁棒KAZE特征匹配算法[J]. 光学精密工程,2016,24(03):616-625.
98. 郭一洋,宋伟东,戴激光. 一种改进的多源遥感影像SIFT算法匹配策略[J]. 测绘工程,2017,26(01):26-31.
99. 许佳佳,张叶,张赫. 基于改进Harris-SIFT算子的快速图像配准算法[J]. 电子测量与仪器学报,2015,29(01):48-54.
100. MIKOLAJCZYK K, SCHMID C. A performance evaluation of local descriptors[J]. IEEE TRANSACtions on Pattern Analysis and Machine Intelligence, 2005, 27(10):1615-1630
101. J.M. Morel and G.Yu, ASIFT, A new framework for fully affine invariant image comparison. SIAM Journal on Imaging Sciences, 2009.

致 谢

白昼渐长，春江水暖，我的硕士研究生生涯和学生生涯也即将一同结束，人生马上要开启新的篇章，去往另外一个熟悉又陌生的城市。回首在西安读书的两年半时间，是我成长最快的两年半，在这里我得以浅尝科研的苦涩与甜蜜，推动了新的领域的大门，也收获了珍贵的友情。

在这里首先要衷心感谢的是我的研究员周诠导师。周老师是一个严厉与和善共存的人。周老师对我的教诲如沐春风，鞭策着我的学习，让我不断进步。并且在面临工作选择的问题上为我耐心分析，处处为学生着想。周老师的为人和科研精神值得我不停学习和追赶，也一定会让我受益终身。

同样需要感谢的还有我的师兄师姐和同学们。有幸和回征师兄还有呼延烺师兄合作了一个项目，培养了我的团队合作精神，并且在后续的科研中二位都对我知无不言，解决了我的一些棘手的问题。感谢我的同学赵瑞、孟正科、龚静文、何雯等同学，大家平常的生活和学习几乎都在一起，从西电到老区再到新区，无疑是一段值得怀念的时光。

最后需要特别感谢的是我的家人。他们不仅含辛茹苦把我培养大，还要为我处处担心。他们一直是我的引路人，是我的母亲劝说我攻读的研究生，而父亲还有外公外婆则一直在背后默默支持。如今他们年纪大了，遇事反而要找我们商量，我一定要让他们继续为我而骄傲。感谢我弟弟的成长与懂事，为我负担家里的大小事情。还要感谢是我的女朋友张晶，一个爱笑又爱哭的女孩。谢谢张晶为我的生活增添了许多色彩，也谢谢她对我的激励，让我不懈怠，更感谢她对我的宽容，原谅了我的种种疏忽和出气闯一闯的决定，希望我们能早日安居乐业。

人生路漫漫，诱惑有之，困境有之，高楼平地起和直堕深渊的事情每天都在上演，我也不知道等待我的会是怎样的一段故事。最终的我的梦想里少不了每一个我爱和爱我的人，只有安身立命，怀揣着梦想，因上努力。

硕士学位论文原创性声明

本人郑重声明：此处所提交的硕士学位论文《 》，是本人在导师指导下，在中国空间技术研究院西安分院攻读硕士学位期间独立进行研究工作所取得的成果。据本人所知，论文中除已注明部分外不包含他人已发表或撰写过的研究成果。对本文的研究工作做出重要贡献的个人和集体，均已在文中以明确方式注明。本声明的法律结果将完全由本人承担。

签字 日期： 年 月 日

硕士学位论文使用授权书

《 》系本人在中国空间技术研究院西安分院攻读硕士学位期间在导师指导下完成的硕士学位论文。本论文的研究成果归中国空间技术空间技术研究院西安分院所有，本论文的研究内容不得以其它单位的名义发表。本人同意院方保留并向有关部门送交论文的复印件和电子版本，允许论文被查阅和借阅。本人授权中国空间技术研究院西安分院，可以采用影印、缩印或其他复制手段保存论文，可以公布论文的全部或部分内容。

学生签名： 日期： 年 月 日

导师签名： 日期： 年 月 日

个人简历、在学期间发表的学术论文与研究成果

个人简历

1993年4月26日出生于山西省太谷县。

2012年9月进入电子科技大学通信工程专业，2016年7月本科毕业并获得工学学士学位。

2016年9月进入中国空间技术研究院攻读工学硕士至今。

发表的学术论文

1. 张晨光,周诠,回征.基于SIFT特征点检测的低复杂度图像配准算法[J].扬州大学学报(自然科学版),2018,21(04):52-56.
2. 《A novel image registration algorithm based on salient region and SIFT》以第一作者被The 3rd International Society for Artificial Intelligence and Robotics （ISAIR 2018）录用
3. 一种适于微纳卫星的分块数据压缩方法, 专利号：201810634416.7。发明人：周诠、 回征、张晨光、呼延烺。
4. 《基于SIFT特征点检测的一种改进的低复杂度图像配准算法》以第一作者被第十九届全国图象图形学学术会议（NCIG2018）录用。