

**中 国 航 天**

**中国空间技术研究院**

**硕士学位论文文献综述**

**论文题目：图像配准及校准**

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： 2017年10月25日

目录

[第一章 绪论 6](#_Toc499804194)

[1.1课题研究背景及意义 6](#_Toc499804195)

[1.1.1 研究背景 6](#_Toc499804196)

[1.1.2 研究意义 7](#_Toc499804197)

[1.2国内外研究现状 8](#_Toc499804198)

[第二章 图像配准 12](#_Toc499804199)

[2.1 图像配准的分类 12](#_Toc499804200)

[2.1.1 根据待配准图像之间的关系 12](#_Toc499804201)

[2.1.2根据配准图像信息处理方法 13](#_Toc499804202)

[2.1.3根据选取的特征信息 14](#_Toc499804203)

[2.2 图像配准的原理 15](#_Toc499804204)

[2.2.1 图像配准问题表述 15](#_Toc499804205)

[2.2.2 图像配准数学模型 16](#_Toc499804206)

[2.3 图像变换 16](#_Toc499804207)

[2.3.1 刚体变换 17](#_Toc499804208)

[2.3.2 仿射变换 17](#_Toc499804209)

[2.3.3 投影变换 18](#_Toc499804210)

[2.3.4 非线性变换 20](#_Toc499804211)

[2.4图像配准流程 21](#_Toc499804212)

[2.5图像配准评价标准 22](#_Toc499804213)

[2.5.1 峰值信噪比 22](#_Toc499804214)

[2.5.2 相似度检测SSIM 23](#_Toc499804215)

[2.5.3 其他评价标准 24](#_Toc499804216)

[3.1 SIFT综述 25](#_Toc499804218)

[3.2 SIFT特点 25](#_Toc499804219)

[3.3 SIFT流程 26](#_Toc499804220)

[3.3.1 尺度空间的生成 27](#_Toc499804221)

[3.3.2 检测尺度空间极值点 29](#_Toc499804222)

[3.3.3 精确定位极值点 31](#_Toc499804223)

[3.3.4 求取特征点的主方向 33](#_Toc499804224)

[3.3.5 关键点描述子的生成 34](#_Toc499804225)

[3.3.6 根据SIFT进行Match 36](#_Toc499804226)

[3.4 本章总结 37](#_Toc499804227)

[第四章 SURF 38](#_Toc499804228)

[4.1 surf综述 38](#_Toc499804229)

[4.2 SURF特点 38](#_Toc499804230)

[4.3 Surf算法原理 39](#_Toc499804231)

[4.3.1 构建Hessian矩阵, 计算特征值α 39](#_Toc499804232)

[4.3.2 构造高斯金字塔 42](#_Toc499804233)

[4.3.3 定位特征点 44](#_Toc499804234)

[4.3.4 确定特征点主方向 45](#_Toc499804235)

[4.3.5 构造特征描述子 46](#_Toc499804236)

[4.4 RANSAC 47](#_Toc499804237)

[4.5 本章小结 48](#_Toc499804238)

[第五章 KAZE 特征 50](#_Toc499804239)

[5.1 KAZE综述 50](#_Toc499804240)

[5.2 KAZE原理 50](#_Toc499804241)

[5.2.1 非线性尺度空间构建 50](#_Toc499804242)

[5.2.2 特征点检测 51](#_Toc499804243)

[5.2.3 计算特征点主方向 52](#_Toc499804244)

[5.2.4 建立描述子 52](#_Toc499804245)

[5.3 本章总结 53](#_Toc499804246)

[第六章 总结 53](#_Toc499804247)

[参考文献 57](#_Toc499804248)

# 第一章 绪论

## 1.1课题研究背景及意义

### 1.1.1 研究背景

眼睛是心灵的窗户。我们很早就认识到视觉的重要性，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。

关于人眼成像的原理，实际上就是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。

随着技术的发展，我们已经可以通过各种传感器成像，比如遥感多光谱成像，毫米波成像，微波成像，同位素成像等，不同的领域利用不同的传感器，运用不同的图像处理技术，获得、突出对应的图像特征，从而实现特定的目标。然而，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。

图像配准就是将不同时间、不同传感器（成像设备）或不同条件下（天候、照度、摄像位置和角度等）获取的两幅或多幅图像进行匹配、叠加的过程。一般来说我们可以认为是将一张图匹配至另一张图的过程，其中一个叫参考图像(reference image)，一个是感知图像(sense image)。由于成像条件的不同，得到的图像也有所不同，所以我们需要图像配准技术来将图片进行合成。这项技术在遥感，医用，制图，计算机视觉领域被广泛地应用。

### 1.1.2 研究意义

   图像配准是许多应用问题必须的预处理步骤,比如:时序图像的变化检测或多模图像融合、图像拼接、图像镶嵌、图像检索、目标识别,这些问题遍及军事、遥感、医学、计算机视觉等多个领域。许多领域都需要图像配准,实际应用过程可能会有所不同,但其中关键的因素是类似的。概括地说,图像配准是对取自不同时间、不同传感器或者不同视角的同一场景的两幅或多幅图像进行最佳匹配的过程,包括像素灰度匹配和空间位置对齐。 图像配准是图像处理研究领域中的一个典型问题和技术难点，其目的在于比较或融合针对同一对象在不同条件下获取的图像，例如图像会来自不同的采集设备，取自不同的时间，不同的拍摄视角等等，有时也需要用到针对不同对象的图像配准问题。具体地说，对于一组图像数据集中的两幅图像，通过寻找一种空间变换把一幅图像映射到另一幅图像，使得两图中对应于空间同一位置的点一一对应起来，从而达到信息融合的目的。

  图像配准技术在计算机视觉、医学图像处理以及材料力学等领域都具有广泛的应用。根据具体应用的不同，有的侧重于通过变换结果融合两幅图像，有的侧重于研究变换本身以获得对象的一些力学属性。20世纪以来医学成像技术经历了从静态到动态，从形态到功能，从平面到立体的飞速发展。

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，但对于不同的应用领域，对图像类型的要求不同，就需要具体问题具体分析。

有研究者根据待配准图像之间的关系，将图像配准分为多源图像配准、基于模板的配准、多角度图像配准、时间序列图像配准四大类。

根据图像获取的不同方法，我们可以找到图像配准的几大问题：

1.      不同视角(Different viewpoints)。

同一事物在不同视角下会获得不同结果，通过配准技术，我们可以获得一个更大的二维视角甚至该事物的三维表达。

2. 不同时间(Different times)。

同样的事物在不同时间下也会有不同的效果。通过图像配准，我们就可以知道该事物根据时间的变化。

3.      不同传感器(Different sensors)。

通过配准不同传感器获得的图像，我们对所摄事物获得更多复杂和细节的信息。

4.      场景到模型配准(Scene to model registration)：

通过所摄场景定位到已知虚拟模型，我们可以对图像进行分析并和模型进行比较。

## 1.2国内外研究现状

图像配准技术最早出现在上世纪七十年代，由美国提出，期初运用于飞行器导航、导弹制导等领域，并得到了美国军方的大力支持和资金上的辅助。历经二十余年的研究，最终成功地将该项技术应用于潘兴II式导弹和战斧式巡航导弹上，保证其弹着点平均圆误差半径小于十几米，使得导弹的命中率得到极大地提高。八十年代之后，这项技术已经成为国内外广泛关注的焦点，中外学者对图像配准技术展开了深入的研究，并在各个领域取得广泛的应用，如模式识别自动导航、医学诊断、计算机视觉等领域。图像配准技术在各个领域都呈现出各自的特点和适应性，都是根据各自具体的应用背景，并需要结合实际情况而特殊定制，使得图像配准技术同时具有相似性和差异性。

从图像配准的提出到目前为止的几十年时间里，随着计算机科学技术的飞速发展和普及，国内外对图像配准技术的研究有了比较系统的发展，形成了一些成熟的算法。在基于图像特征的配准技术中，目前运用比较成熟的图像特征包括点特征、直线段、边缘、轮廓、闭合区域和统计矩等。特征提取算法可分为点特征提取算子（如Harris算子和Susan算子）、线特征提取算子（如LoG算子、Canny算子）和面特征算子。随着边缘检测和图像分割等技术的快速发展，基于轮廓、边缘和区域的配准方法正在逐渐成为研究的热点，如需要人工干预、特征点的获取比较困难等；基于灰度的图像配准直接利用整幅的灰度对两幅图像之间的相似性进行度量，基于灰度的配准方法的优势在于只对图像的灰度进行处理，可以避免主观因素的影响。常见的算法有最大互信息法、相关法、条件熵法、联合熵法等。

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。

在多模图像配准中，关于互信息图像配准技术的研究，已经成为该领域的热点课题，最大互信息算法是图像匹配的有力工具，但其仍然存在确定，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。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，如提高非刚性配准的速度等诸多问题都需要国内学者共同探讨。

下面依据科研分布情况对当前图像配准技术的研究现状进行阐述。通过对研究成果(比如:硕、博士论文)相对集中的科研机构或技术人员进行分析,可以很好地把握该领域的研究广度和深度。

1)UMCU(UniversityMedicalCenterUtrecht,Neth2erlands)图像科学所。以MaxA.Viergever教授为代表,研究兴趣包括计算机视觉和医学成像的所有方面。文献[6]综述了该所上世纪最后10年的研究情况,重点在图像配准、图像分割和可视化。

2)VU(VanderbiltUniversity,USA)电子工程与计算机科学系。以J.MichaelFitzpatrick教授为代表,研究兴趣在医学成像,医学图像处理,图像配准、图像制导手术。

3)SU(SyracuseUniversity,USA)电子工程与计算机科学系。以PramodK.Varshney教授为代表,研究兴趣在隐藏武器探测。文献[7]综述了其课题小组近10年的研究情况。

4)KCL(King’sCollegeLondon)医学成像科学组。以DerekL.G.Hill教授为代表,研究兴趣在图像配准、多模图像制导手术、MRI运动纠正、形状分析等,著有《医学图像配准》。

5)WSU(WrightStateUniversity,USA)计算机科学工程系。以ArdeshirGoshtasby教授为代表,近20年来,一直致力于图像配准技术研究。他是图像融合系统研究的发起者,并创立了以研究和发展图像配准及融合技术为目的的公司。

6)美国NASA空间飞行中心。以JacquelineLeMoigne教授为代表,为图像配准在遥感领域应用作了大量工作,涉及多种技术在该领域的应用,诸如小波多分辨率策略以及互相关、互信息、Hausdorff距离等相似度测量。

图像配准是许多应用问题必须的预处理步骤,比如:时序图像的变化检测或多模图像融合,这些问题遍及军事、遥感、医学、计算机视觉等多个领域。许多领域都需要图像配准,实际应用过程可能会有所不同,但其中关键的因素是类似的。概括地说,图像配准是对取自不同时间、不同传感器或者不同视角的同一场景的两幅或多幅图像进行最佳匹配的过程,包括像素灰度匹配和空间位置对齐。条分类准则,依次为:图像维数、配准特征的来源、变换模型、坐标变换域、交互性、优化策略、成像模式、配准对象、配准目标的特点等。可以参考Brown和Maintz的分类方法,将配准技术概括为8个方面,包括:配准对象、特征提取、特征匹配、变换模型、优化策略、坐标变换与插值、系统实现及算法评估。

图像配准一般可分为3种：1.不同分辨率，相同传感器的两幅图像之间的匹配，比如景象匹配末制导时，必须对实时图进行景象匹配。2.图像部分内容相同，相同传感器的两幅或多幅图像之间的配准，比如可见光图像与雷达图像（SAR）、红外图像，医学图像CT、PET、MRI等的匹配。

图像配准分为以下几个基本步骤：特征检测、特征匹配、变换模型参数计算、图像的变换及重采样口。其中，特征检测方法主要包括2种：基于区域灰度的特征检测和基于特征(点、线、斑状区域)的特征检测。其中基于特征的图像配准方法主要通过图像的特征匹配关系建立图像之间的映射关系，已经成为图像配准的热点研究课题。

在基于特征点检测的图像配准算法中，英国学者David Lowe于2004年总结完善，提出SIFT特征匹配算法(scale invariant feature transform)[1]。该算法通过提取稳定的特征，可以对发生平移、旋转、仿射变换、视角变换、光变化情况的图像进行匹配，因此得到了广泛运用。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提出SURF(speeded—up mbust features)[5]。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。

兴趣点检测相关研究：

1998 Lindberg介绍自动尺度选择的概念，允许检测图像中的兴趣点在它们的特征尺度上。他实验了Hessian矩阵的行列式和Laplacian(和矩阵的迹一致)检测团状结构。1998 Lowe提出用DOG近似LOG。2001 Mikolajczyk 和Schmid 重新定义了这个方法，名为Harris-Laplace和Hessian-Laplace。使用Harris或Hessian矩阵的行列式来选择特征点的闻之，使用Laplacian选择尺度。

使用HESSIAN矩阵的近似检测兴趣点。使用积分图像加快计算。2001 Viola and Jones 提出积分图像的概念。1998 Simard 提出的盒形计算框架使用积分图像。

此外Mikolajczyk（2005，2006）还做了一些算子的比较工作。从中可知：基于Hessian检测器比基于Harris检测器更稳定，重复检测性更好。此外，使用Hessian矩阵的行列式比使用它的迹更有优势。同时也发现使用类似于DOG的近似方法可以提高速度但只损失很小的精度。

描述符的相关研究:

图像特征点的描述符一个共同点是表达了兴趣点邻域内小尺度的特征分布。使得描述符的描述性更好，识别性更高。SIFT的特点正是掌握了空间域亮度模式的大量信息（基于直方图方法：8个方向的箱格，4\*4像素）。描述了特征点邻域内点的梯度方向信息，共128维。PCA-SIFT：36维，匹配速度更快，但区分度下降，并且延长了特征的计算时间。GLOH：区分度更高但是数据压缩花销时间太长。

2006 Grabner使用积分图像近似SIFT。可以达到和我们同样的速度。但是相比SIFT质量有所下降。（为SURF提供了重要信息积分图像）。

匹配算法：BBF（k-d tree）,balltrees, vocabulary trees, locality sensitine hashing.本文补充提出了，使用Hessian矩阵的迹来显著提高匹配速度。在低维描述符下，任何算法的匹配速度都很快。

# 第二章 图像配准

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，但对于不同的应用领域，对图像类型的要求不同，就需要具体问题具体分析。

## 2.1 图像配准的分类

### 2.1.1 根据待配准图像之间的关系

根据待配准图像之间的关系，可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。

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

### 2.1.2根据配准图像信息处理方法

（1）基于待配准图像灰度信息的配准方法

以整幅图像的灰度信息为依据，建立待配准图像和基准参考图像之间的相似性度量，利用某种搜索算法，寻找出使得相似度量达到最优值的变换模型参数。也称之为直接配准法；

（2）基于待配准图像变换域内信息的配准方法

通常是利用傅里叶变换为基础，进行频域内的配准。傅里叶变换能够用于具有平移、旋转、缩放的图像配准中是因为傅里叶变换具有平移、旋转、缩放的特点。

（a）对于两幅具有平移量的图像经过傅里叶变换后，他们的相位关系是不同的，因为在时域内平移量直接决定在傅里叶变换域内的相位关系；

（b)对于两幅具有旋转量的图像，在傅里叶变换域内的旋转量保持不变；

（c)对于两幅具有尺度缩放的图像，先进行坐标系转换，变换到对数坐标系当中，图像的缩放可以转换成图像平移进行处理。

（3）基于待配准图像特征信息的配准方法

### 2.1.3根据选取的特征信息

（a）基于特征点的匹配

一般所选取的特征点是相对于其领域表现出来具有某种奇异性的像素点。特征点往往容易被提取到，但是特征点所包含的信息相对较少，只能反映出其在图像中的位置坐标信息，所以在两幅图像中寻找匹配的特征点是关键所在。

（b）基于特征区域的匹配

在图像中寻找某些明显的区域信息作为特征区域，然而在实际应用中寻找到特征区域后，采用最多的还是区域的形心点，所以这类算法要求特征区域提取的精度非常高。

（c）基于特征边缘的匹配

图像中最明显的特征就是边缘，而边缘特征也是最好提取的特征之一。因此给予边缘的匹配方法鲁棒性较强，适用范围很广，但这类方法对特征边缘的提取要求很高，并且要求边缘信息全部用数学语言予以表述也较为困难。

基于特征的配准方法是目前最常用的配准方法之一，该算法只需要提取待配准图像中的点、线、边缘等特征信息，不需要其它辅助信息，在减少计算量、提高效率的同时，能够对图像灰度的变化有一定的鲁棒性。但是，由于该算法只采用了图像一小部分的特征信息，所以这种算法对特征提取和特征匹配的精度及准确性要求非常高，对错误非常敏感。

## 2.2 图像配准的原理

### 2.2.1 图像配准问题表述

多幅图像是由于采用了不同的传感器在同一时间，或者是由一个或数个传感器在不同时间对地面同一目标成像获得的。在这种多幅图像的处理应用中，完成多幅图像之间的配准是其中一个重要的环节。在许多图像配准的文献中，图像配准的数学上的定义如下：给定两幅待配准的图像 ，称其中之一比如 为基准图像，另一个 为待配准图像，则称图像配准为两幅图像在坐标位置和灰度级上的双重映射变换。

从推广的范畴考虑，图像配准是产生三维空间的图像集合或者匹配某一目标的过程。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。比如说，对于同一公路场景，它的高分辨率图像可能是一条宽带，带有明晰的道路边界，而其低分辨率图像则可能是一根模糊线条，利用公路和它环境的几何特征可以判断两幅图像是否在几何上对齐，即配准或广义匹配。而不是将一幅图像变换得与另一幅相等。由此可以将图像配准和图像匹配的相关概念定义如下。

定义1：图像配准是给定两幅待配准图像，对其中一幅图像作变换使得变换后的图像与另一个图像的内容在拓扑上相对应并几何上相对齐。图像配准可以称为图像的广义匹配。

定义2：狭义的图像匹配是指一幅图像在灰度属性上按像元逼近另一幅图像。

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。

图像配准的基本问题就是找出一种图像转换方法，用以纠正图像的形变。造成图像形变的原因多种多样，例如对于遥感图像而言，传感器噪声、由传感器视点变化或平台不稳定造成的透视变化、被拍摄物体的移动、变形或生长等变化、闪电和大气变化，以及阴影和云层遮盖都使图像产生不同形式的形变。

### 2.2.2 图像配准数学模型

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h表示二维空间坐标变换，g表示灰度或辐射变换，描述因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，灰度变换参数。通常意义的配准只关心位置坐标的变换。灰度或辐射变换可以归为图像预处理部分。

## 2.3 图像变换

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。

变换模型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。

经常用到的主要变换有刚体变换、放射变换、投影变换、非线性变换。

### 2.3.1 刚体变换

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（Rigid Transform）。 刚体变换可分解为平移变换、旋转变换和反转（镜像）变换。

### 2.3.2 仿射变换

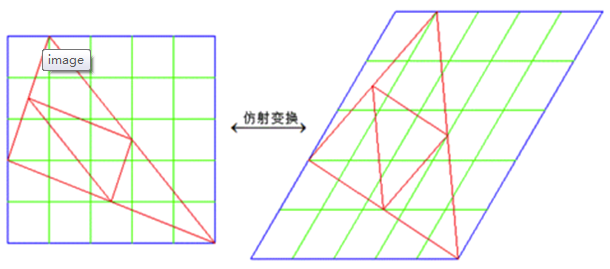
Affine Transformation是一种二维坐标到二维坐标之间的线性变换，保持二维图形的“平直性”（译注：straightness，即变换后直线还是直线不会打弯，圆弧还是圆弧）和“平行性”（译注：parallelness，其实是指保二维图形间的相对位置关系不变，平行线还是平行线，相交直线的交角不变）。

图2- 1

图1- 1

仿射变换，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：



其中 表示平移量，而参数 反映了图像旋转、缩放等变化。将参数计算出，即可得到两幅图像的坐标变换关系。

当变换矩阵为 时，就是平移变换。平移变换就是一种刚体变换。

### 2.3.3 投影变换

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 )。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3x3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。

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

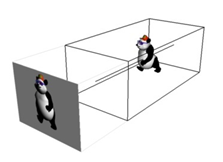
正交却是世界中的物体按照平行的光线投射到一张纸上(你的画布)，仿佛被压缩在上面，无论这个物体在世界中多远，投射结果还是原来的大小, 2d 游戏多用此

图2- 2

种投影。如下图：

透视投影（Perspective Projection）

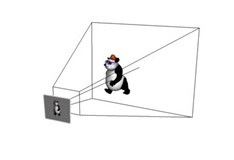
透视比较接近我们现实的方式了，你的眼睛发出的光形成一个夹角，离你的眼睛越近，东西越大，范围越小；反之离眼睛越远，东西越小，视野越开阔。因此多用于 3D 中。如下图：

图2- 3

### 2.3.4 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，F表示把一幅图像映射到另一幅图像上的任意一种函数函数形式。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

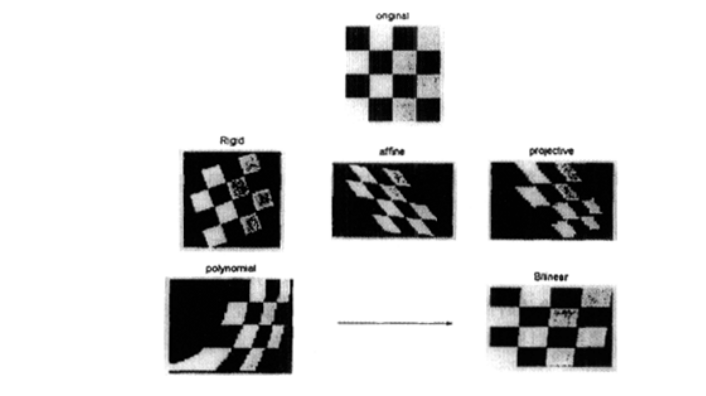
图像配准中，采用不同的图像变换说明了图像之间的映射关系将会不同。因此对于不同类型的图像，选取恰当的图像变换模型可以适当地提高图像的配准精度。在对需配准的图像进行空间变换后，要对变换图像进行重新取样，以取得变换后的像素值。

图2- 4

## 2.4图像配准流程

目前，很难找到一种普适的方法能够应对所有的配准情况，任何一种配准算法都必须考虑图像的成像原理、几何变形、噪声影响、配准精度等因素。不过，从原理上将，配准算法可以大致分为以下四个步骤：

（1）特征提取

采用人工或者自动的方法检测图像中的不变特征，如：闭合区域、边缘、轮廓、角点等。特征提取算法需要满足三个条件

（a）显著性，所提取的特征应该是比较明显的，分布广泛的、易于提取的特征；

（b）抗噪性，具有较强的噪声抑制能力且对成像条件的变化不敏感；

（c）一致性，能准确地检测出两幅图像的共有特征；

（2）特征匹配

通过特征描述算作及相似性度量来建立所提取的特征之间的对应关系。特征匹配常用到的区域灰度、特征向量空间分布和特征符号描述等信息。某些算法在进行特征匹配的同时也完成了变换模型参数的估计；

（3）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。

（4）坐标变换与插值

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，因此，需要考虑一定的插值处理操作。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；

## 2.5图像配准评价标准

### 2.5.1 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 位线性脉冲编码调制表示，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。

### 2.5.2 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：



### 2.5.3 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。

# 第三章 Scale Invariant Feature Transform(SIFT)

## 3.1 SIFT综述

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[1]，2004年完善总结[2]。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影像的大小和旋转无关。对于光线、噪声、些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撷取，在母数庞大的特征数据库中，很容易辨识物体而且鲜有误认。使用 SIFT特征描述对于部分物体遮蔽的侦测率也相当高，甚至只需要3个以上的SIFT物体特征就足以计算出位置与方位。在现今的电脑硬件速度下和小型的特征数据库条件下，辨识速度可接近即时运算。SIFT特征的信息量大，适合在海量数据库中快速准确匹配。

此算法有其专利，专利拥有者为英属哥伦比亚大学。

## 3.2 SIFT特点

SIFT算法的特点有：

1. SIFT特征是图像的局部特征，其对旋转、尺度缩放、亮度变化保持不变性，对视角变化、仿射变换、噪声也保持一定程度的稳定性；

2. 独特性（Distinctiveness）好，信息量丰富，适用于在海量特征数据库中进行快速、准确的匹配；

3. 多量性，即使少数的几个物体也可以产生大量的SIFT特征向量；

4. 高速性，经优化的SIFT匹配算法甚至可以达到实时的要求；

5. 可扩展性，可以很方便的与其他形式的特征向量进行联合。

SIFT算法可以解决的问题：

目标的自身状态、场景所处的环境和成像器材的成像特性等因素影响图像配准/目标识别跟踪的性能。而SIFT算法在一定程度上可解决：

1. 目标的旋转、缩放、平移（RST）

2. 图像仿射/投影变换（视点viewpoint）

3. 光照影响（illumination）

4. 目标遮挡（occlusion）

5. 杂物场景（clutter）

6. 噪声

SIFT在图像的不变特征提取方面拥有无与伦比的优势，但并不完美，仍然存在：

1. 实时性不高。2. 有时特征点较少。3. 对边缘光滑的目标无法准确提取特征点等缺点。对模糊的图像和边缘平滑的图像，检测出的特征点过少，对圆更是无能为力。

## 3.3 SIFT流程

Sift特征匹配算法主要包括两个阶段，一个是Sift特征的生成，即从多幅图像中提取对尺度缩放、旋转、亮度变化无关的特征向量；第二阶段是Sift特征向量的匹配。

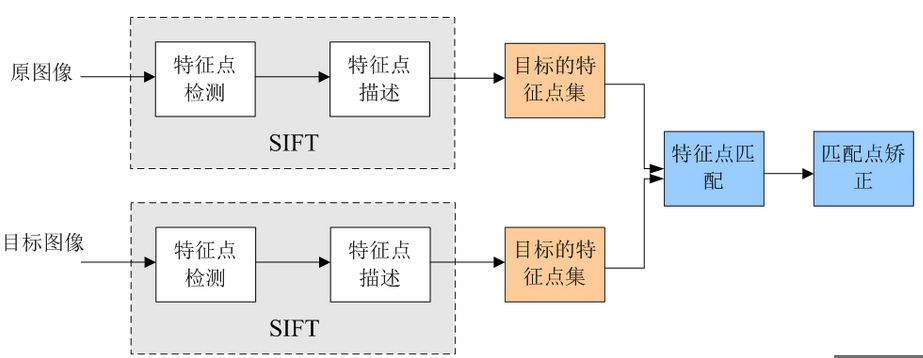


图3- 1

### 3.3.1 尺度空间的生成

在一定的范围内，无论物体是大还是小，人眼都可以分辨出来。然而计算机要有相同的能力却不是那么的容易，在未知的场景中，计算机视觉并不能提供物体的尺度大小，其中的一种方法是把物体不同尺度下的图像都提供给机器，让机器能够对物体在不同的尺度下有一个统一的认知。在建立统一认知的过程中，要考虑的就是在图像在不同的尺度下都存在的特征点。所以，尺度空间的生成是一个初始化操作，尺度空间理论目的是模拟图像数据的多尺度特征。

高斯卷积核是实现尺度变换的唯一线性核，于是一副二维图像的尺度空间定义为：



其中 是尺度可变高斯函数， ，在二维空间中，这个公式生成的曲面的等高线是从中心开始呈正态分布的同心圆。

 是空间坐标， 是尺度坐标。 大小决定图像的平滑程度，大尺度对应图像的概貌特征，尺度越大图像越模糊。小尺度对应图像的细节特征。大的 值对应粗糙尺度(低分辨率)，反之，对应精细尺度(高分辨率)。分布不为零的像 素组成的卷积矩阵与原始图像做变换。每个像素的值都是周围相邻像素值的加权平均。原始像素的值有最大的高斯分布值，所以有最大的权重，相邻像素随着距离原 始像素越来越远，其权重也越来越小。这样进行模糊处理比其它的均衡模糊滤波器更高地保留了边缘效果。理论上来讲，图像中每点的分布都不为零，这也就是说每个像素的计算都需要包含整幅图像。在实际应用中，在计算高斯函数的离散近似时，在大概3σ距离之外的像素都可以看作不起作用，这些像素的计算也就可以忽略。

构建尺度空间的目的是为了检测出在不同的尺度下都存在的特征点，而检测特征点较好的算子是 (高斯拉普拉斯,LoG)：



事实上，2002年Mikolajczyk在详细的实验比较中发现尺度归一化的高斯拉普拉斯函数 的极大值和极小值同其它的特征提取函数，例如：梯度，Hessian或Harris角特征比较，能够产生最稳定的图像特征。

使用LoG虽然能较好的检测到图像中的特征点，但是其运算量过大，通常可使用DoG（差分高斯，Difference of Gaussina）来近似计算LoG。这得益于Lindeberg早在1994年的发现[4]。

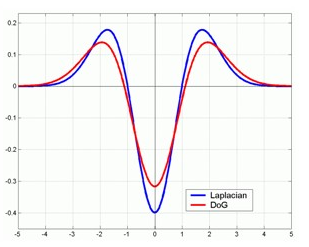


图3- 2

红色曲线表示的是高斯差分算子，而蓝色曲线表示的是高斯拉普拉斯算子

高斯差分尺度空间（DOG scale-space）：



利用不同尺度的高斯差分核与图像卷积生成。

关于尺度空间的理解说明：2kσ中的2是必须的，尺度空间是连续的。在 Lowe的论文中 ，将第0层的初始尺度定为1.6（最模糊），图片的初始尺度定为0.5（最清晰）. 在检测极值点前对原始图像的高斯平滑以致图像丢失高频信息，所以 Lowe 建议在建立尺度空间前首先对原始图像长宽扩展一倍，以保留原始图像信息，增加特征点数量。

图像金字塔的建立：对于一幅图像I,建立其在不同尺度(scale)的图像，也成为子八度（octave），这是为了scale-invariant，也就是在任何尺度都能够有对应的特征点，第一个子八度的scale为原图大小，后面每个octave为上一个octave降采样的结果，即原图的1/4（长宽分别减半），构成下一个子八度（高一层金字塔）。

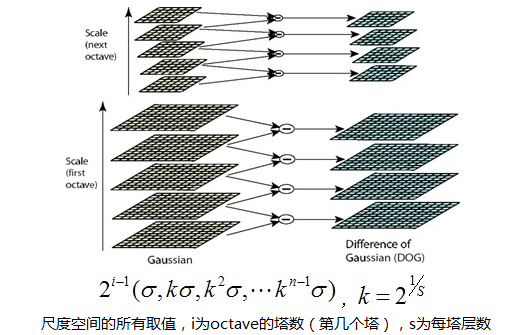


图3- 3

由图片size决定建几个塔，每塔几层图像(S一般为3-5层)。0塔的第0层是原始图像(或称double后的图像)，往上每一层是对其下一层进行Laplacian变换（高斯卷积，其中σ值渐大，例如可以是σ, k\*σ, k\*k\*σ…），直观上看来越往上图片越模糊。塔间的图片是降采样关系，例如1塔的第0层可以由0塔的第3层down sample得到，然后进行与0塔类似的高斯卷积操作。

### 3.3.2 检测尺度空间极值点

为了寻找尺度空间的极值点，每一个采样点要和它所有的相邻点比较，看其是否比它的图像域和尺度域的相邻点大或者小。如图所示，中间的检测点和它同尺度的8个相邻点和上下相邻尺度对应的9×2个点共26个点比较，以确保在尺度空间和二维图像空间都检测到极值点。 一个点如果在DOG尺度空间本层以及上下两层的26个领域中是最大或最小值时，就认为该点是图像在该尺度下的一个特征点, 如此可以保证检测到的关键点在尺度空间和二维图像空间上都是局部极值点。如图所示。

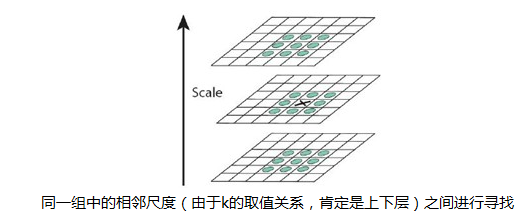


图3- 4

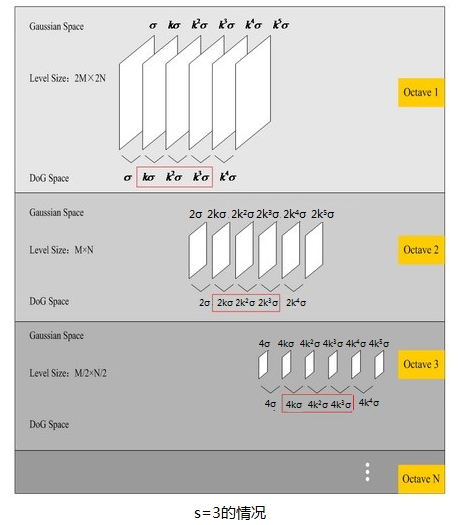


图3- 5

### 3.3.3 精确定位极值点

通过比较检测得到的DoG的局部极值点实在离散的空间搜索得到的，由于离散空间是对连续空间采样得到的结果，因此在离散空间找到的极值点不一定是真正意义上的极值点，因此要设法将不满足条件的点剔除掉[10]。

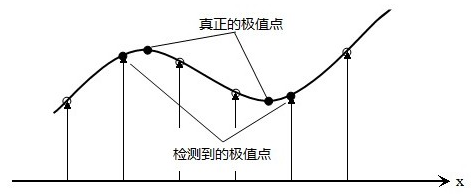


图3- 6

DOG值对噪声和边缘比较敏感，所以在第2步的尺度空间中检测到的局部极值点还要经过进一步的筛选，去除不稳定和错误检测出的极值点，另一点就是在构建高斯金字塔过程中采用了下采样的图像，在下采样图像中提取的极值点对应在原始图像中的确切位置，也是要在本步骤中解决的问题。

可以通过尺度空间DoG函数进行三维二次函数曲线拟合寻找极值点，这一步的本质是去掉DoG局部曲率非常不对称的点。

要剔除掉的不符合要求的点主要有两种：

1.低对比度的特征点；2.不稳定的边缘响应点。

对于低对比度的点：通过拟和三维二次函数以精确确定关键点的位置和尺度（达到亚像素精度），同时去除低对比度的关键点和不稳定的边缘响应点(因为DoG算子会产生较强的边缘响应)，以增强匹配稳定性、提高抗噪声能力，在这里使用近似Harris Corner检测器。

1.候选特征点空间尺度函数泰勒展开式如下：



对比度为 绝对值，是偏移量。因为候选点认为是极值点，所以对上式求导,其为0,得到精确的位置, 即



2.把公式代入，即在DoG Space的极值点处D(x)取值，只取前两项，得 ，设对比度的阈值为T

若 ，该特征点就保留下来，否则丢弃。

对于不稳定的边缘响应点：一个定义不好的高斯差分算子的极值在横跨边缘的地方有较大的主曲率，而在垂直边缘的方向有较小的主曲率。候选特征点的DoG函数D(x)的主曲率与一个2×2 的Hessian矩阵H的特征值成正比:



令α为H的较大特征值，β为较小的特征值，是 和 方向的梯度，则：



令 ， 是最大特征值与最小特征值的比值。则：



 的值在两个特征值相等的时候最小，随着的增大而增大，值越大，说明两个特征值的比值越大，即在某一个方向的梯度值越大，而在另一个方向的梯度值越小，而边缘恰恰就是这种情况。所以为了剔除边缘响应点，需要让该比值小于一定的阈值，因此，为了检测主曲率是否在某域值 下，只需检测



If  ，throw it out，在lowe的文章中[2]，取r=10.

### 3.3.4 求取特征点的主方向

上一步中确定了每幅图中的特征点，为每个特征点计算一个方向，依照这个方向做进一步的计算， 利用关键点邻域像素的梯度方向分布特性为每个关键点指定方向参数，使算子具备旋转不变性。

稳定的极值点是在不同尺度空间下提取的，这保证了关键点的尺度不变性。为关键点分配方向信息所要解决的问题是使得关键点对图像角度和旋转具有不变性。方向的分配是通过求每个极值点的梯度来实现的。

对于任一关键点，其梯度幅值表述为：



梯度方向为：



其中L所用的尺度为每个关键点各自所在的尺度。

分配给关键点的方向并不直接是关键点的梯度方向，而是按照一种梯度方向直方图的方式给出的。要采集其所在高斯金字塔图像3σ邻域窗口内像素的梯度和方向分布特征。

具体的方法是：计算以关键点为中心的邻域内所有点的梯度方向梯度直方图的范围是0～360度，其中每45度一个柱，总共8个柱, 或者每10度一个柱，总共36个柱。将梯度方向直方图中纵坐标最大的项代表的方向分配给当前关键点作为主方向，若在梯度直方图中存在一个相当于主峰值80%能量的峰值，则将这个方向认为是关键点的辅方向。辅方向的设计可以增强匹配的鲁棒性，Lowe指出，大概有15%的关键点具有辅方向，而恰恰是这15%的关键点对稳定匹配起到关键作用[9]。

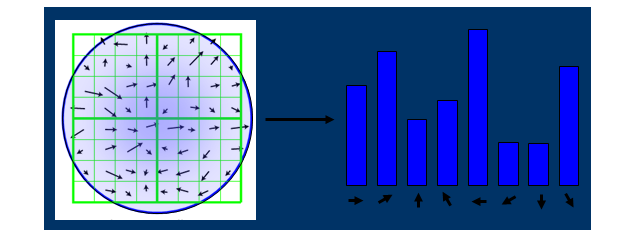


图3- 7

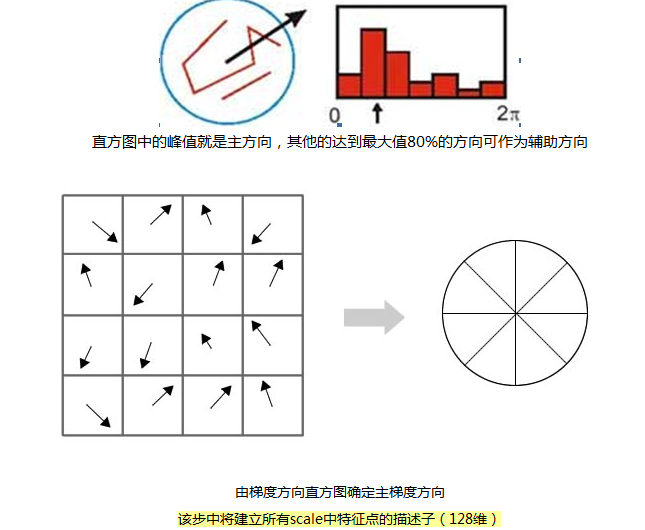


图3- 8

至此，图像的关键点已经检测完毕，每个关键点有三个信息 ：位置，所处尺度、方向，由此可以确定一个SIFT特征区域。一个SIFT特征区域由三个值表示，中心表示特征点位置，半径表示关键点的尺度，箭头表示主方向。

### 3.3.5 关键点描述子的生成

对关键点的描述是后续实现匹配的关键步骤，描述其实就是一种以数学方式定义关键的过程。描述子不但包含关键点，也包括关键点周围对其有贡献的邻域点。描述的思路是：对关键点周围像素区域分块，计算块内梯度直方图，生成具有独特性的向量，这个向量是该区域图像信息的一种抽象表述。

特征描述符的生成大致有三个步骤：

1. 校正旋转主方向，确保旋转不变性。

2. 生成描述子，最终形成一个128维的特征向量

3.归一化处理，将特征向量长度进行归一化处理，进一步去除光照的影响。

首先将坐标轴旋转为关键点的方向，以确保旋转不变性。以关键点为中心取8×8的窗口。图左部分的中央为当前关键点的位置，每个小格代表关键点邻域所在尺度空间的一个像素，利用公式求得每个像素的梯度幅值与梯度方向，箭头方向代表该像素的梯度方向，箭头长度代表梯度模值，然后用高斯窗口对其进行加权运算。

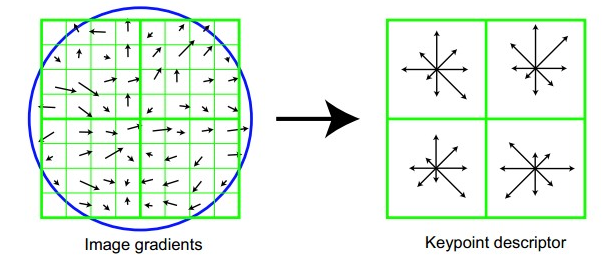


图3- 9

图中蓝色的圈代表高斯加权的范围（越靠近关键点的像素梯度方向信息贡献越大）。然后在每4×4的小块上计算8个方向的梯度方向直方图，绘制每个梯度方向的累加值，即可形成一个种子点，如图右部分示。此图中一个关键点由2×2共4个种子点组成，每个种子点有8个方向向量信息。这种邻域方向性信息联合的思想增强了算法抗噪声的能力，同时对于含有定位误差的特征匹配也提供了较好的容错性。

David G.Lowed[1]的实验结果表明：对每个关键点，采用4\*4\*8共128维向量的描述子进项关键点表征，综合效果最佳：计算keypoint周围的16\*16的window中每一个像素的梯度，而且使用高斯下降函数降低远离中心的权重。

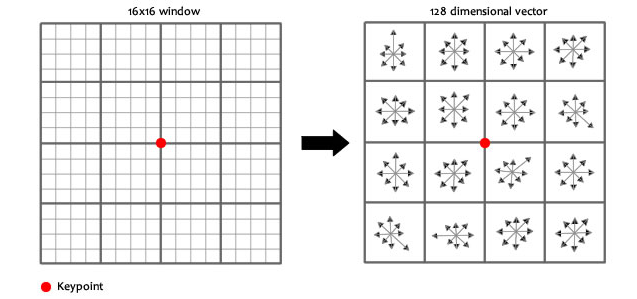


图3- 10

在每个4\*4的1/16象限中，通过加权梯度值加到直方图8个方向区间中的一个，计算出一个梯度方向直方图。这样就可以对每个feature形成一个4\*4\*8=128维的描述子，每一维都可以表示4\*4个格子中一个的scale/orientation. 将这个向量归一化之后，就进一步去除了光照的影响。

### 3.3.6 根据SIFT进行Match

生成了A、B两幅图的描述子，（分别是k1\*128维和k2\*128维），就将两图中各个scale（所有scale）的描述子进行匹配，匹配上128维即可表示两个特征点match上了[11]。

实际计算过程中，为了增强匹配的稳健性，Lowe建议对每个关键点使用4×4共16个种子点来描述，这样对于一个关键点就可以产生128个数据，即最终形成128维的SIFT特征向量。此时SIFT特征向量已经去除了尺度变化、旋转等几何变形因素的影响，再继续将特征向量的长度归一化，则可以进一步去除光照变化的影响。 当两幅图像的SIFT特征向量生成后，下一步我们采用关键点特征向量的欧式距离来作为两幅图像中关键点的相似性判定度量。取图像1中的某个关键点，并找出其与图像2中欧式距离最近的前两个关键点，在这两个关键点中，如果最近的距离除以次近的距离少于某个比例阈值，则接受这一对匹配点。降低这个比例阈值，SIFT匹配点数目会减少，但更加稳定。为了排除因为图像遮挡和背景混乱而产生的无匹配关系的关键点,Lowe提出了比较最近邻距离与次近邻距离的方法,距离比率ratio小于某个阈值的认为是正确匹配。因为对于错误匹配,由于特征空间的高维性,相似的距离可能有大量其他的错误匹配,从而它的ratio值比较高。Lowe推荐ratio的阈值为0.8。但作者对大量任意存在尺度、旋转和亮度变化的两幅图片进行匹配，结果表明ratio取值在0. 4~0. 6之间最佳，小于0. 4的很少有匹配点，大于0. 6的则存在大量错误匹配点。(如果这个地方你要改进，最好给出一个匹配率和ration之间的关系图，这样才有说服力)作者建议ratio的取值原则如下:

ratio=0. 4　对于准确度要求高的匹配；

ratio=0. 6　对于匹配点数目要求比较多的匹配；

ratio=0. 5　一般情况下。

也可按如下原则:当最近邻距离<200时ratio=0. 6，反之ratio=0. 4。ratio的取值策略能排除错误匹配点

## 3.4 本章总结

SIFT算法的实质是在不同的尺度空间上查找关键点(特征点)，并计算出关键点的方向。Sift算法就是用不同尺度（标准差）的高斯函数对图像进行平滑，然后比较平滑后图像的差别，差别大的像素就是特征明显的点。SIFT所查找到的关键点是一些十分突出，不会因光照，仿射变换和噪音等因素而变化的点，如角点、边缘点、暗区的亮点及亮区的暗点等。

Sift是David Lowe于1999年提出的局部特征描述子，并于2004年进行了更深入的发展和完善。Sift特征匹配算法可以处理两幅图像之间发生平移、旋转、仿射变换情况下的匹配问题，具有很强的匹配能力。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。

# 第四章 SURF

## 4.1 surf综述

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，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。

## 4.2 SURF特点

SURF特点：

1.使用积分图像完成图像卷积（相关）操作;

2，使用Hessian矩阵检测特征值；

3，使用基于分布的描述符（局部信息）。

## 4.3 Surf算法原理

### 4.3.1 构建Hessian矩阵, 计算特征值α

surf构造的金字塔图像与sift有很大不同，就是因为这些不同才加快了其检测的速度。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。首先是图像中某个像素点的Hessian矩阵：



海森矩阵的混合偏导数是海森矩阵非主对角线上的元素。假如他们是连续的，那么求导顺序没有区别，海森矩阵的行列式，可用于分辨 f 的临界点是属于鞍点还是极值点。

每一个像素点都可以求出一个Hessian矩阵。但是由于我们的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，经过滤波后在进行Hessian的计算，其公式如下：



其中， 为高斯滤波后图像 在各个方向的二阶导数，也可看成是高斯二阶微分与图像的卷积。是一幅图像在不同解析度下的表示，利用高斯核函数与图像函数在点 的卷积来实现。

二阶导数计算方法为：



其中。 为原始图像的灰度值， 是将高斯滤波处理后的图像。

为了找出图像中的特征点，需要对原图进行变换，在sift算法中，是在DOG图像中进行在surf算法中，该变换图就是原图每个像素的Hessian矩阵行列式的近似值构成的。公式为：



其中，0.9为原文作者给出的一个经验值。

由于求Hessian时要先高斯平滑，然后求二阶导数，这在离散的像素点是用模板卷积形成的，这2种操作合在一起用一个模板代替就可以了。又为了将模板与图产像的卷积转换为盒子滤波运算，我们对高斯二阶微分模板进行简化，使得简化后的模板只是由几个矩形区域组成，矩形区域内填充同一值，如下图所示，在简化模板中白色区域的值为正数，黑色区域的值为负数，灰度区域的值为0。，比如说y方向上的模板如下，以9\*9的模板为例：

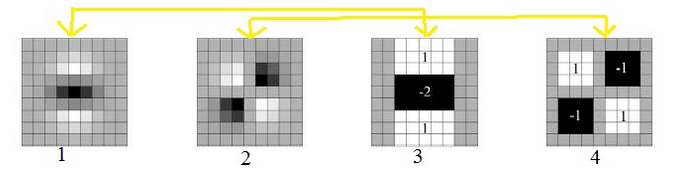


图4- 1

上图中1为y方向的先高斯滤波然后二阶求导的处理，近似处理为图3，图2位x和y方向上的先高斯滤波然后二阶混合偏导，近似为图4。

有了这个近似的模板以后，计算高斯滤波和二阶导数两个步骤就可以一个步骤完成，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，提高了速度。

积分图像的概念是由Viola和Jones提出的。就是当前的每个像素的灰度都是它与坐标原点（0,0）形成的对角线的矩形内的所有像素的灰度值之和。如下图：

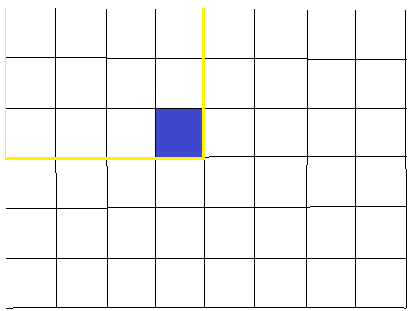


图4- 2

图中蓝色像素的积分图像的灰度值就是黄色框内的所有像素的灰度值之和。其他像素同样的计算方法。

有了积分图像的概念，在计算某个矩形框内的像素灰度值之和时，就可以很简单的得出了。如下图

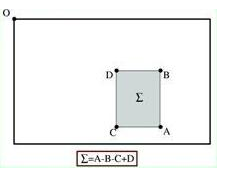


图4- 3

灰色矩形框内的灰度值之和就是A+D-C-B，就是很简单的一个矩形面积计算公式，A代表AO对角线矩形的面积，以此类推[6]。

这样，计算像素的H矩阵时，将会非常的方便，比如-2那一块，只需要知道4个顶点像素的积分图像的值，便可以求出那一块的灰度值之和。求积分图像，只需要对原图像所有像素进行一遍扫描。一旦积分图计算好了，计算图像内任何矩形区域的像素值的和只需要三个加法。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。

至此，我们求出了原图的Hessian矩阵，然后计算行列式的近似值就得出了我们需要寻找特征点的变换图像。

### 4.3.2 构造高斯金字塔

在计算视觉领域，尺度空间被象征性的表述为一个图像金字塔，其中，输入图像函数反复与高斯函数的核卷积并反复对其进行二次抽样，这种方法主要用于Sift算法的实现，但每层图像依赖于前一层图像，并且图像需要重设尺寸，因此，这种计算方法运算量较大。在sift算法中，每一组（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。

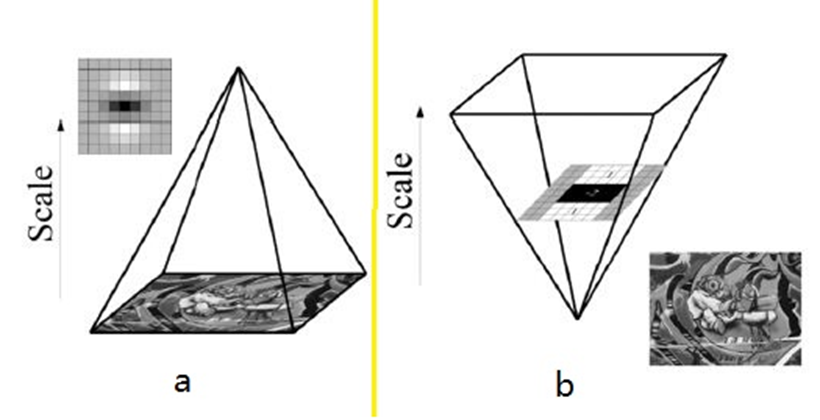


图4- 4

SURF算法申请增加图像核的尺寸，这也是SIFT算法与SURF算法在使用金字塔原理方面的不同。算法允许尺度空间多层图像同时被处理，不需对图像进行二次抽样，从而提高算法性能。图1(a)是传统方式建立一个如图所示的金字塔结构，图像的尺寸是变化的，并且会反复使用高斯函数对子层进行平滑处理，图1(b)说明

Surf算法使原始图像保持不变而只改变滤波器大小。因为surf算法没有了降采样的过程，因此处理速度得到提高。

与SIFT算法类似，我们需要将尺度空间划分为若干组（Octaves）。一个组代表了逐步放大的滤波模板对同一输入图像进行滤波的一系列响应图。每个组又由若干固定的层组成。由于积分图像离散化的原因，两个层之间的最小尺度变化量是由高斯二阶微分滤波器在微分方向上对正负斑点响应长度决定的，它是盒子滤波器模板尺寸的。对于的模板，它的。下一层的响应长度至少应该在 的基础上增加2个像素，以保证一边一个像素，即。这样模板的尺寸就为。以此类推，我们可以得到一个尺寸增大模板序列，它们的尺寸分别为：，滤波器尺寸依次增加6；黑色、白色区域的长度增加偶数个像素，以保证一个中心像素的存在。

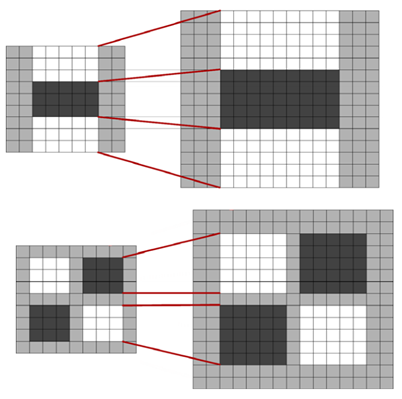


图4- 5

采用类似的方法来处理其他几组的模板序列。其方法是将滤波器尺寸增加量翻倍（6，12，24，38）。这样，可以得到第二组的滤波器尺寸，它们分别为15，27，39，51。第三组的滤波器尺寸为27，51，75，99。如果原始图像的尺寸仍然大于对应的滤波器尺寸，尺度空间的分析还可以进行第四组，其对应的模板尺寸分别为51，99，147和195。下图显示了第一组至第三组的滤波器尺寸变化。

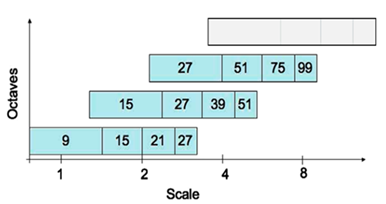


图4- 6

在通常尺度分析情况下，随着尺度的增大，被检测到的斑点数量迅速衰减。所以一般进行3-4组就可以了，与此同时，为了减少运算量，提高计算的速度，可以考虑在滤波时，将采样间隔设为2。

如前所述，我们使用 的模板对图像进行滤波，其结果作为最初始的尺度空间层（此时，尺度值为s=1.2，近似 的高斯微分），后续的层将通过逐步放大滤波模板尺寸，以及放大后的模板不断与图像进行滤波得到。由于采用盒子滤波和积分图像，滤波过程并不随着滤波模板尺寸的增加而使运算工作量增加。

### 4.3.3 定位特征点

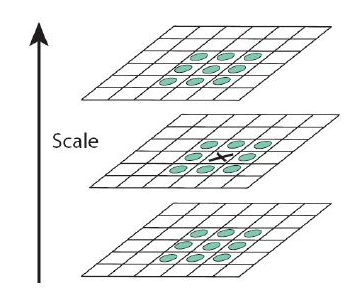
首先初步定为特征点，我们用了  邻域非最大值抑制。如下图，将经过hessian矩阵处理过的每个像素点与其3维邻域的26个点进行大小比较，如果它是这26个点中的最大值或者最小值，则保留下来，当做初步的特征点。

图4- 7

然后，跟sift算法类似，采用3维线性插值法，Hessian矩阵行列式的最大值在尺度和图像空间被插值，得到亚像素级的特征点，同时也去掉那些值小于一定阈值的点，增加极值使检测到的特征点数量减少，最终只有几个特征最强点会被检测出来。

### 4.3.4 确定特征点主方向

为了保证旋转不变性，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6s(S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x(水平)和y(垂直)方向的Haar小波响应总和(Haar小波边长取4s)，并给这些响应值赋高斯权重系数，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，遍历整个圆形区域，选择最长矢量的方向为该特征点的主方向。这样，通过特征点逐个进行计算，得到每一个特征点的主方向。该过程的示意图如下：

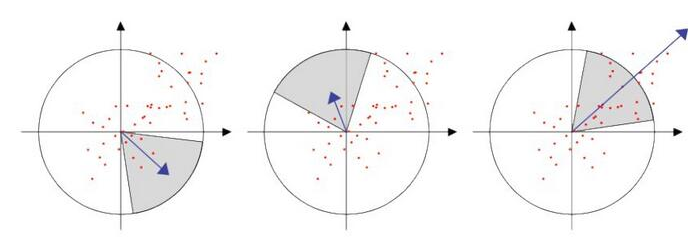


图4- 8

### 4.3.5 构造特征描述子

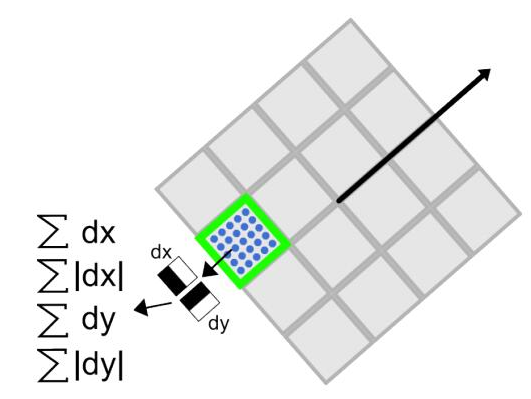
生成特征点描述子与确定特征点方向有些类似，它需要计算图像的Haar小波响应。不过，与主方向的确定不同的是，这次我们不是使用一个圆形区域，而是在一个矩形区域来计算Haar小波响应。在特征点周围取一个正方形框，框的边长为20s(s是所检测到该特征点所在的尺度)。该框带方向，方向当然就是第4步检测出来的主方向了。然后把该框分为16个子区域，每个子区域统计25个像素的水平方向和垂直方向的Haar小波特征，这里的水平和垂直方向都是相对主方向而言的。该Haar小波特征为水平方向值之和，水平方向绝对值之和，垂直方向之和，垂直方向绝对值之和。该过程的示意图如下所示：

图4- 9

这样每个小区域就有4个值，所以每个特征点就是16\*4=64维的向量，相比sift而言，少了一半，这在特征匹配过程中会大大加快匹配速度。

为了实现快速匹配，SURF在特征矢量中增加了一个新的变量，即特征点的拉普拉斯响应正负号。在特征点检测时，将Hessian矩阵的迹的正负号记录下来，作为特征矢量中的一个变量。这样做并不增加运算量，因为特征点检测进已经对Hessian矩阵的迹进行了计算。在特征匹配时，这个变量可以有效地节省搜索的时间，因为只有两个具有相同正负号的特征点才有可能匹配，对于正负号不同的特征点就不进行相似性计算。

简单地说，我们可以根据特征点的响应值符号，将特征点分成两组，一组是具有拉普拉斯正响应的特征点，一组是具有拉普拉斯负响应的特征点，匹配时，只有符号相同组中的特征点才能进行相互匹配。显然，这样可以节省特征点匹配的时间[6]。

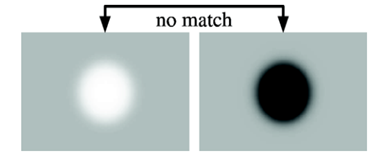


图4- 10 黑背景下的亮斑和白背景下的黑斑，因为它们的拉普拉斯响应正负号不同，不会对它们进行匹配

综上所述，Surf在速度上比sift要快许多，这主要得益于它的积分图技术，已经Hessian矩阵的利用减少了降采样过程，另外它得到的特征向量维数也比较少，有利于更快的进行特征点匹配。SURF采用Henssian矩阵获取图像局部最值还是十分稳定的，但是在求主方向阶段太过于依赖局部区域像素的梯度方向，有可能使得找到的主方向不准确，后面的特征向量提取以及匹配都严重依赖于主方向，即使不大偏差角度也可以造成后面特征匹配的放大误差，从而匹配不成功；另外图像金字塔的层取得不足够紧密也会使得尺度有误差，后面的特征向量提取同样依赖相应的尺度，在这个问题上我们只能采用折中解决方法：取适量的层然后进行插值。

## 4.4 RANSAC

对角点进行初始匹配后，所选定的角点并不能保证全部是正确的点，也可能有误点，因此，还需要进一步对所选定的角点进行精确匹配。RANSAC(RANdom Sample And Consensus) 随机抽样一致性算法是由Fischler和Bolles提出的一种鲁棒性的参数估计方法[14]。它的基本思想是在进行参数估计时，不是不加区分地对待所有可用的输入数据，而是首先针对具体问题设计出一个目标函数，然后迭代地估计该函数的参数值，利用这些初始参数值把所有的数据分为所谓的“内点”(Inliers，即满足估计参数的点)和“外点”(Outliers，即不满足估计参数的点)，最后反过来用所有的“内点”重新计算和估计函数的参数。

RANSAC的基本假设是：

（1）数据由“局内点”组成，例如：数据的分布可以用一些模型参数来解释；

（2）“局外点”是不能适应该模型的数据；

（3）除此之外的数据属于噪声。

局外点产生的原因有：噪声的极值；错误的测量方法；对数据的错误假设。使用RANSAC估计方法，可以最大限度地减少噪声及外点的影响。简单的最小二乘法不能找到适应于局内点的直线，原因是最小二乘法尽量去适应包括局外点在内的所有点。相反，RANSAC能得出一个仅仅用局内点计算出模型，并且概率还足够高。但是，RANSAC并不能保证结果一定正确，为了保证算法有足够高的合理概率，我们必须小心的选择算法的参数。

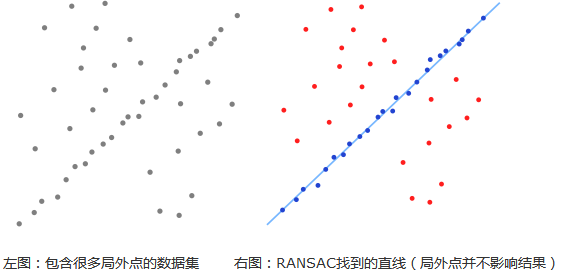


图4- 11

### 4.4.6 根据SIFT进行Match

RANSAC用于选定最佳角点的主要步骤为：

(1)、将初始提取的2行N列的N个角点变成3行N列的N个角点，第三行为全1；(2)、对初始提取的角点进行归一化；(3)、设置迭代最大次数；(4)、设置每次随机选取时最少的角点个数；(5)、每次随机选取由(4)指定个数的角点；(6)、判断由(5)中选取的角点，是否有部分共线，并设定最大循环次数；(7)、对选定的角点计算2-D单应性矩阵；(8)、选定一次内点并记录；(9)、对选定的内点进行判断；

(10)、如满足条件(9)则作为选出的最好一组内点，否则，依次循环(5)~(9)；

RANSAC算法由于在初始时是随机选取角点，因此存在不确定性，即使相同操作对两对同样的特征点进行RANSAC算法角点提取，每一次得出的结果也不一定就是相同的，但是最终的效果一般都是比较理想的。

### 4.4.6 RANSAC优缺点

RANSAC的优点是它能鲁棒的估计模型参数。例如，它能从包含大量局外点的数据集中估计出高精度的参数。RANSAC的缺点是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可信的模型，概率与迭代次数成正比。RANSAC的另一个缺点是它要求设置跟问题相关的阀值。

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。

经典RANSAC有以下三个主要的局限性：

（1） 效率：经典方法效率与子集大小、类内点比例以及数据集大小有关，因此在某些场景下效率较低。

（2） 精度：经典方法计算参数时选取最小子集是从效率的角度考虑，往往得到的是非最佳参数，在应用产参数 之前还需要再经过细化处理。

（3） 退化：经典方法的目标函数求取最大化的过程基于一个假设：“选取的最小子集中如果有类外点，那么在这种情况下估计的参数获得的目标函数（数据集中点的个数）往往较少”但这种情况在退化发生时有可能是不对的。

## 4.5 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Harr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。

SURF算法与ransac算法相结合可以剔除出错误的匹配对。RANSAC算法在SURF特征筛选中的主要流程是：

(1) 从样本集中随机抽选一个RANSAC样本，即4个匹配点对。

(2) 根据这4个匹配点对计算变换矩阵M。

(3) 根据样本集，变换矩阵M，和误差度量函数计算满足当前变换矩阵的一致集consensus，并返回一致集中元素个数。

(4) 根据当前一致集中元素个数判断是否最优(最大)一致集，若是则更新当前最优一致集。

(5) 更新当前错误概率p，若p大于允许的最小错误概率则重复(1)至(4)继续迭代，直到当前错误概率p小于最小错误概率。

# 第五章 KAZE 特征

## 5.1 KAZE综述

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[12]。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。

## 5.2 KAZE原理

### 5.2.1 非线性尺度空间构建

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[13]。

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。

### 5.2.2 特征点检测

特征点的检测与Sift特征点检测也非常相似，首先计算各点在本层，上下层3X3的立方体邻域内响应值是否为极值，再去掉重复点，再亚像素化精确位置，所不同的是各个像素点处响应值的计算。文章中响应值是采用多尺度下尺度归一化Hessian行列式计算的如下:



其中分别表示x方向2阶导数，y方向2阶导数，xy方向混合导数，其主要过程描述如下：

（1）计算各个像素点处响应值，响应值的计算如上

（2）查找在3x3空间邻域内的极大值

（3）删除重复特征点，两点在相邻层并且距离在尺度范围内认为是重复的。

（4）亚像素化，精确定位，亚像素化的过程与sift中的一样，通过泰勒表达式求导数，倒数为0的地方即为极值所在处，该位置即为偏移量：



可以通过解方程：



得到，而 即为对应矩阵。

### 5.2.3 计算特征点主方向

文章所用的特征点主方向计算方法与SURF算法如出一辙。主要过程如下：

（1） 以特征点为中心，半径为6s区域内的点求方向倒数,然后以标准差为2.5s的高斯函数进行加权

（2）建立一个 的扇形区域，该扇形以0.15弧度为步长进行旋转，然后计算落角度在该扇形区域内点的方向导数和

（3）获取所有扇形区域中使得方向导数和的平方和的最大的扇形，主方向即为该扇形区域方向导数和的反正切结果

### 5.2.4 建立描述子

文章中描述子的建立是采用M-SURF描述子，对于一个尺度为 的特征，取该特征点邻域的矩形，再将该矩形分成4x4个的小矩形区域(每个相邻子区域有个像素是重叠的);计算邻域内所有点的方向导数，并且以标准差为的高斯函数进行加权，然后统计每个子区域内的方向导数以及方向导数绝对值的和，对于每个子区域再通过标准差为的高斯函数进行加权，最终每个子区域就能形成一个4维特征,个子区域能够形成64维特征，如果想将64维特征扩展到128维特征，可以对方向导数的正负值分别统计。

## 5.3 本章总结

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。

# 第六章 总结

图像配准在图像处理，尤其是海量处理中是非常重要的一个步骤，它把来源、质量、大小、角度参差不齐的多幅图像进行变换，从而与一幅事先选定的标准图像进行配准，进而利于后续的各种图像处理如图像融合、图像拼接、图像镶嵌、目标识别。事实上，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、航空侦察等领域都有广泛的应用。图像配准的质量好坏直接关系到最终的图像处理效果。这使得图像配准成为极具应用价值的技术。

而在图像配准中，图像特征点的检测又是关键和难点。图像的特征提取起到了降维分析的目的。因此一个算法是否成功往往由它使用和定义的特征决定。因此特征提取最重要的一个特性是“可重复性”：同一场景的不同图像所提取的特征应该是相同的。对于尺度不变特征，尺度不变特征变换（Scale Invariant Feature Transform，SIFT）是一种图像特征提取与描述算法。SIFT算法由David.G.Lowe于1999年提出并在2004年进行了完善总结。SIFT算法可以处理两幅图像之间发生平移、旋转、尺度变化、光照变化情况下的特征匹配问题，并能在一定程度上对视角变化、仿射变化也具备较为稳定的特征匹配能力。针对SIFT不能实时运算的问题，Bay在2006年发表在ECCV大会上的SURF 算法，全称是 Speeded-Up Robust Features。该算子在保持 SIFT 算子优良性能特点的基础上，同时解决了 SIFT 计算复杂度高、耗时长的缺点，对兴趣点提取及其特征向量描述方面进行了改进，且计算速度得到提高。而利用RANSAC算法，可以进一步剔除一些错误的匹配对，加快处理速度，改进处理效果。ECCV2012中出现了一种比SIFT更稳定的特征检测算法KAZE，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。

国内在图像配准术领域的研究与国际上基本同步，国内的许多科研院所，如北京邮电大学、清华大学、中科院、西安电子科技大学、南京理工大学、天津大学、哈尔滨工业大学、湖南大学等，都在开展这方面的研究工作。国家的863计划、973计划和国家自然基金都也都对这一方向的研究给予了资助。

西安分院目前也有相关课题在进行图像配准方面的工作，周诠研究员于2005年开始着手关于图像配准方面的研究，目前获得已授权的与图像配准技术相关的专利申请有30余项。在周老师的指导下，其研究生冯新岗，李晓博，魏佳圆，朱厉洪，方海，刘娟妮、张茗茗等一直在丰富图像配准这方面的研究工作。

每种算法都有各自的优缺点，当处理样本不一样时，算法表现也不一样，并且特征点的提取其实不是越多越好，看到的好多论文只是片面地列举能表现自己改进的算法优点的数据，选择有利于自己的图片库。这是存在问题的。论文：A comparison of SIFT, PCA-SIFT and SURF 对三种方法给出了性能上的比较，源图片来源于Graffiti dataset，对原图像进行尺度、旋转、模糊、亮度变化、仿射变换等变化后，再与原图像进行匹配，统计匹配的效果。效果以可重复出现性为评价指标。由此可见，SIFT在尺度和旋转变换的情况下效果最好，SURF在亮度变化下匹配效果最好，在模糊方面优于SIFT，而尺度和旋转的变化不及SIFT，旋转不变上比SIFT差很多。速度上看，SURF是SIFT速度的3倍。但是该论文的评价标准没有量化，只是简单的用common、good评价。

所以，本课题的思路一方面是结合最新的研究成果，对已有算法进行优化和改进，如算法主动识别出完全不相关的图像，做出拒判，以期达到更加稳定，更加快速的效果；一方面是用多源、多样的图片库来验证算法，找到算法的不足，进一步调整，希望算法更加有普适性。而这些目标，不仅要求对计算机视觉有更深的认识，深刻理解最基本的比如SIFT算法，因为后期众多算法是SIFT的改进；此外还要对人类视觉有更深的理解，因为图像处理是利用计算机进行处理，但归根结底是服务于人眼的，这就依赖生物学的一些研究。此外，计算机视觉的研究表明，不同视角下物体往往可以通过一个透视矩（ 或）阵的变换而得到，但这种模型下的图像配准，以观察者的角度说，只配准了观察角度，以图像为观察对象的话，只对它进行了仿射变换，即只进行了旋转，缩放等变换，图像的颜色、分辨率、饱和度等依然没有和参考图像配准。还有，就目前看到的论文，更多的是针对两幅图像之间的配准，没有考虑到对多幅图像进行批量配准。批量配准不应该只是传统意义上的多次配准，应该引入机器学习的概念，在配准过程中不断迭代优化，找到更加快速、稳定的算法。

# 参考文献

1. David G.Lowe “Object Recognition from Local Scale-Invariant Features”. ICCV .1999: 1150 - 1157 vol.2
2. D.G.Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”[J], IJCV,January 5, 2004
3. 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/
4. T.Lindeberg, "Scale-space theory:A basic tool for analysing structures at different scales".1994
5. H.Bay, T. Tuytelaars, L. V. Gool, SURF:Speed Up Robust Features[J],ECCV,2006
6. 思维之际 —— As we may think.SURF算法与源码分析，http://www.cnblogs.com/ronny/p/4045979.html。http://www.cnblogs.com/ronny/p/4048213.html
7. Lindeberg的论文《Scale-space theory: A basic tool for analysing structures at different scales》
8. CSDN,jiutianhe博客,图像似度算法--SIFT算法详解http://blog.csdn.net/jiutianhe/article/details/39896931
9. RaySaint,UnderTheWood博客：SIFT算法研究 http://underthehood.blog.51cto.com/2531780/658350
10. 特征点检测学习\_1(sift算法) http://www.cnblogs.com/tornadomeet/archive/2012/08/16/2643168.html
11. 特征点检测学习\_2(surf算法) <http://www.cnblogs.com/tornadomeet/archive/2012/08/17/2644903.html>
12. Pablo F.Alcantarilla,Adrien Bartoli, Andrew J.Davision .KAZE Features.[J]ECCV.2012
13. J.Weicket;B.M.T.HRomeny;M.A.Viergver.Efficient.” Efficient and reliable schemes for nonlinear diffusion filtering”. IEEE Transactions on Image Processing.1998
14. 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.
15. 科学计算中的偏微分方程有限差分法 杨乐主编