

前 言

复杂环境下的飞行器自主导航是实现飞行器高效侦察、精确打击、应急救援等重要任务的关键，对提升飞行器智能化和自主化水平具有重要的作用。随着人工智能、传感器技术与图像处理技术的发展，视觉导航技术已成为飞行器自主导航的关键技术，其核心在于通过飞行器载图像传感器实时感知环境，并基于景象匹配与目标识别为飞行器提供定位导航支持，为飞行器在复杂环境下的自主定位与路径规划提供高效的解决方案。

红外视觉导航和合成孔径雷达(synthetic aperture radar, SAR)视觉导航是飞行器视觉导航的两大主流方向。本书基于作者课题组多年研究工作的系统总结，以红外视觉导航和 SAR 视觉导航两大技术体系为主线，系统阐述视觉导航的理论基础、技术方法与工程实践，涵盖传统算法与智能算法的协同创新。

全书包括 8 章，内容安排如下：第 1 章在阐述飞行器视觉导航原理的基础上，介绍视觉导航的关键技术；第 2 章主要介绍红外视觉导航所涉及的基本概念、原理、相关技术；第 3 章和第 4 章针对红外基准图制备中的红外图像仿真问题，分别介绍基于红外特性建模的红外图像仿真方法和基于深度学习的红外图像仿真方法；第 5 章主要介绍 SAR 视觉导航所涉及的基本概念、原理、相关技术；第 6 章主要介绍 SAR 景象匹配适配区选取方法；第 7 章主要介绍异源图像匹配方法；第 8 章主要介绍 SAR 典型目标检测方法。

在本书的撰写过程中，参阅了大量国内外文献资料，在此对这些文献的作者表示衷心的感谢。

谨以本书献给所有致力于飞行器视觉导航研究的同仁。由于视觉导航技术快速迭代，加上作者水平有限，书中难免存在不妥之处，恳请读者不吝指正。

作 者

2025 年春于西安

目 录

前言

第 1 章 概论	1
1.1 视觉导航原理	1
1.1.1 基本原理	1
1.1.2 主要类别	2
1.2 视觉导航关键技术	3
1.2.1 基准图制备	4
1.2.2 图像匹配算法	6
1.2.3 目标检测算法	12
1.3 本书主要内容	16
参考文献	17
第 2 章 红外视觉导航	20
2.1 引言	20
2.2 红外成像原理	21
2.2.1 黑体辐射定律	21
2.2.2 实际物体的辐射	23
2.2.3 红外成像窗口	24
2.2.4 红外热成像仪	24
2.3 红外图像特性分析	26
2.3.1 红外图像整体特性	26
2.3.2 红外图像背景特性	27
2.3.3 红外图像目标特性	29
2.4 基于模板匹配的前视红外目标识别	30
2.4.1 问题描述	30
2.4.2 算法描述	31
2.4.3 实验结果与分析	37
2.5 基于知识模型的前视红外目标识别	38
2.5.1 机场跑道	39
2.5.2 桥梁	39

2.5.3	油库	40
2.5.4	雷达天线罩	41
2.5.5	冷却塔	42
2.6	基于深度学习的红外目标检测	43
2.6.1	红外目标检测数据集的制作	43
2.6.2	红外目标检测模型	46
2.6.3	实验结果与分析	48
	参考文献	51
第 3 章	基于红外特性建模的红外图像仿真	53
3.1	引言	53
3.2	地物红外辐射特性分析	54
3.2.1	地物表面温度影响因素	55
3.2.2	建筑物红外特性分析	58
3.3	红外仿真软件	60
3.3.1	SE-Workbench-IR 软件	60
3.3.2	Vega 软件	64
3.4	红外仿真图像质量评估	65
3.5	基于 SE-Workbench-IR 的红外图像仿真	68
3.5.1	仿真流程	68
3.5.2	实验结果与分析	73
	参考文献	75
第 4 章	基于深度学习的红外图像仿真	76
4.1	引言	76
4.2	生成对抗模型	77
4.2.1	基本原理	77
4.2.2	条件生成对抗网络	79
4.2.3	经典模型	82
4.3	面向匹配任务的可见光/红外图像转换算法	83
4.3.1	算法描述	84
4.3.2	实验结果与分析	88
4.3.3	算法讨论	94
4.3.4	小结	97
4.4	单边无配对的可见光/红外图像转换算法	97
4.4.1	网络设计	98

4.4.2	损失函数设计	101
4.4.3	实验结果与分析	104
4.4.4	小结	111
4.5	判别器引导的可见光/红外图像转换算法	111
4.5.1	算法描述	112
4.5.2	实验结果与分析	117
4.5.3	小结	120
	参考文献	120
第 5 章	SAR 视觉导航	123
5.1	引言	123
5.2	SAR 成像原理	124
5.2.1	微波后向散射	124
5.2.2	雷达成像仪	127
5.3	SAR 图像特性分析	132
5.3.1	相干斑	132
5.3.2	几何特性	133
5.3.3	典型目标类型	135
5.4	SAR 景象匹配	136
5.5	SAR 寻的制导	138
	参考文献	144
第 6 章	SAR 景象匹配适配区选取	145
6.1	引言	145
6.2	景象匹配适配区特征参数	146
6.2.1	基于信息量的指标	146
6.2.2	基于稳定性的指标	147
6.2.3	基于重复度的指标	148
6.2.4	基于地形特征的指标	149
6.3	面向成像差异的 SAR 景象匹配适配区选取法	151
6.3.1	典型地物稳定性分析	151
6.3.2	稳定区域提取	153
6.3.3	小结	156
6.4	基于卷积神经网络的 SAR 景象匹配适配区选取法	156
6.4.1	训练样本集构建	157
6.4.2	卷积神经网络子区分类模型	160

6.4.3 实验结果与分析	163
6.4.4 小结	166
参考文献	166
第 7 章 异源图像匹配	168
7.1 引言	168
7.2 基于端到端孪生 U-Net 的异源图像模板匹配算法	169
7.2.1 算法描述	169
7.2.2 实验结果与分析	172
7.2.3 小结	178
7.3 基于改进 SuperPoint 点特征与视觉线性 Transformer 的异源图像 匹配算法	178
7.3.1 算法描述	178
7.3.2 实验结果与分析	184
7.3.3 小结	187
7.4 基于结构特征边缘的异源图像匹配算法	187
7.4.1 最佳尺度确定	188
7.4.2 基于改进的自适应结构张量粗匹配	190
7.4.3 基于线特征的精匹配	192
7.4.4 实验结果与分析	196
7.4.5 小结	198
7.5 基于封闭均匀区域的 SAR 异源图像匹配算法	199
7.5.1 基于多尺度非线性扩散滤波的均匀区域提取	200
7.5.2 基于多边形拟合的区域粗匹配	201
7.5.3 基于重合度的区域精匹配	203
7.5.4 基于区域质心构建变换模型	205
7.5.5 实验结果与分析	205
7.5.6 小结	209
参考文献	209
第 8 章 SAR 典型目标检测	211
8.1 引言	211
8.2 基于二维经验模态分解的 SAR 图像建筑物检测算法	211
8.2.1 经验模态分解算法	212
8.2.2 特征区域提取	213
8.2.3 基于特征区域组合的建筑物检测	217

8.2.4 实验结果与分析	219
8.2.5 小结	220
8.3 基于 Ghost 卷积与注意力机制的 SAR 图像建筑物检测算法	221
8.3.1 算法描述	221
8.3.2 实验结果与分析	227
8.3.3 小结	231
8.4 基于深度可分离逆残差网络与双注意力机制的 SAR 图像飞机 检测算法	231
8.4.1 算法描述	231
8.4.2 实验结果与分析	237
8.4.3 小结	240
8.5 基于深层次多尺度特征融合卷积神经网络的 SAR 图像舰船 目标检测算法	240
8.5.1 算法描述	241
8.5.2 实验结果与分析	245
8.5.3 小结	248
参考文献	249

第1章 概 论

1.1 视觉导航原理

飞行器自主导航是指飞行器在没有人工干预或在较少人工干预的情况下，利用飞行器载传感器和导航算法，完成飞行器的定位、路径规划和航迹调整，从而实现飞行器的自主飞行。如何在复杂环境下实现飞行器的精确导航一直是精确制导领域研究的热点问题，是实现飞行器高效侦察、精确打击、应急救援等重要任务的关键，对提升飞行器智能化和自主化水平具有重要的作用。

1.1.1 基本原理

飞行器导航模式主要包括惯性导航、卫星导航、视觉导航等。惯性导航基于惯性测量单元测量，解算确定飞行器在惯性坐标系中的地理位置，自主性好，短时精度高，但定位误差会随时间累积。卫星导航基于全球定位系统或北斗卫星导航系统，通过接收和解算卫星信号得到高精度的位置信息，误差不随时间积累，但自主性差，容易受干扰。视觉导航基于飞行器载图像传感器获取飞行路径上或目标区域的图像，通过图像处理算法获取飞行器位置或目标位置，为飞行控制提供导航信息，是一种模拟人类视觉定位功能的导航模式，使飞行器在飞行过程中具备视觉能力，具有高自主、低成本、稳定性强等优势，但在飞行器运动变化剧烈、光照强度变化大、目标场景复杂等情况下，导航精度会受到影响。

实际应用中，单一的导航模式无法满足高精度的导航要求，通常采用以惯性导航为主、其他导航为辅的组合导航模式。视觉导航是一种常用的辅助导航模式，可在卫星导航受干扰、地形环境受限、地磁特征不明显等条件下辅助惯性导航，从而大大提高飞行器导航的精度和自主性，其示意图如图 1.1 所示。

视觉导航的工作过程如下。飞行前，基于卫星遥感正射影像、目标先验知识和约束条件，制备基准图并装载在飞行器载计算机上；飞行中，利用飞行器载图像传感器采集实时图，与预先存储的基准图进行实时匹配运算，获得飞行器位置信息或目标位置信息，实现对飞行器的导航、制导与控制。基于上述过程，视觉导航通常又称为景象匹配导航^[1,2]。

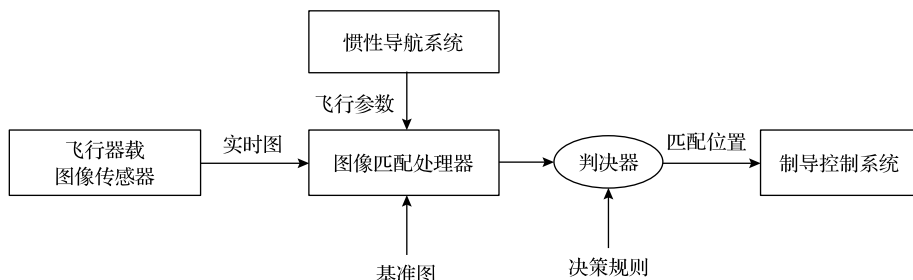


图 1.1 视觉导航示意图

1.1.2 主要类别

根据不同的分类方法,飞行器视觉导航具有多种不同的类别。根据导航任务的不同,视觉导航可分为下视景象匹配与前视景象匹配^[3]。根据实时图成像所采用的频段,视觉导航可分为可见光成像导航、红外成像导航、合成孔径雷达(synthetic aperture radar, SAR)成像导航(在精确制导领域,常称其为可见光成像制导、红外成像制导、SAR 成像制导)。上述任务和实时图各种类别的不同组合,形成了多种不同的景象匹配导航方式。

1. 下视景象匹配

下视景象匹配的主要特点是:基准图为飞行路径上的卫星遥感正射图像,每个像素具有确定的地理位置信息,实时图像为飞行器下方图像,匹配的目的是确定实时图在基准图中的相对位置,从而获得实时图成像时刻飞行器的地理位置信息。如图 1.2 所示,理想情况下,实时图中心所对应的匹配位置应位于基准图预定航线上某一位置;实际情况下,匹配位置与预定位置可能存在一定的偏移,根据此偏移量生成制导指令,可以对飞行器进行航迹调整,使之回到正确的航线上。下视景象匹配模式中,实时图和基准图主要是可见光图像或 SAR 图像。

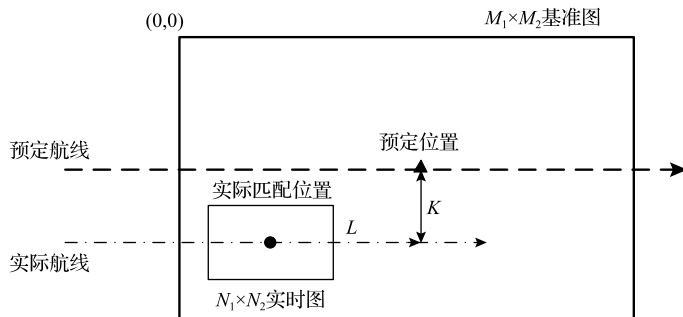


图 1.2 下视景象匹配位置与偏移量

2. 前视景象匹配

前视景象匹配和下视景象匹配的不同之处在于,飞行器载传感器的成像角度由向下看变成向前看,其主要特点是:基准图为目标在某一视点和角度下的模板图像,实时图为飞行器前下方包含目标在内的目标区域图像,匹配的目的在于确定目标模板图像在实时图中的位置,获得飞行器与目标的相对关系。如图 1.3 所示,理想情况下,探测器光轴 OO_C 应与 OO_t 重合,目标应位于探测器视场中心,匹配位置应位于实时图中心。实际情况,匹配位置可能并不位于实时图中心,而是存在一定的偏移量,因此需要根据此偏移量计算使光轴指向目标的航向偏移角 $\Delta\varphi$ 和俯仰偏移角 $\Delta\psi$ 。

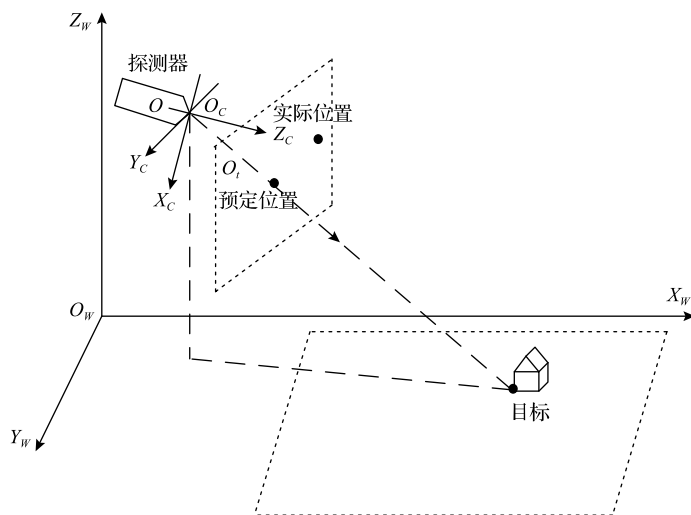


图 1.3 前视景象匹配位置与偏移量

前视景象匹配模式中,实时图和基准图可以是可见光图像、红外图像或 SAR 图像。前视景象匹配的实质是对飞行器前下方目标区域图像的典型目标进行自动目标识别 (automatic target recognition, ATR), 因此又称其为前视寻的制导。

1.2 视觉导航关键技术

随着红外成像技术、SAR 成像技术与图像处理技术的发展,视觉导航技术已成为发展精确制导武器等无人自主飞行器的关键技术,通过对目标区域进行成像,对获取的红外图像或 SAR 图像进行景象匹配与目标识别,为飞行器提供定位导航支持,可以有效提高飞行器的导航性能和命中精度。

视觉导航的关键技术包括基准图制备、图像匹配和目标检测等,对导航系统

工作的可靠性、鲁棒性和精确性具有很大的影响。基准图制备是视觉导航的基础,没有高适配性的基准图,就无法实现高精度的景象匹配和飞行器定位。图像匹配是视觉导航的核心,解决实时图与基准图正确匹配的问题,是飞行器精确定位的前提与依据。目标检测是前视景象匹配的关键,代表飞行器视觉导航的发展趋势。在人工智能高速发展的今天,上述技术经历了从传统算法到深度学习算法的发展和进化。

1.2.1 基准图制备

在下视景象匹配和前视景象匹配任务中,基准图制备的处理对象、处理内容和成果形式分别具有各自的特点。

1. 下视景象匹配的基准图制备

在下视景象匹配任务中,基准图为飞行路径上的卫星遥感正射图像,以可见光图像或 SAR 图像为主。

飞行器飞行过程中可能会经过包含各种地物景象的区域,如建筑物、道路、河道等特征明显且容易匹配的区域(称为适配区),或植被、荒漠、海洋、山区等特征不明显且难以匹配的区域(称为非适配区)。为了使景象匹配结果能为修正惯性导航偏差提供参考信息,从而实现飞行器的精确定位,在执行飞行任务之前,需要对卫星遥感正射图像进行适配性分析,选取从起飞点到降落点之间的适配区,并规划出一条经过尽可能多适配区的飞行航迹。

适配区选取就是按照一定的要求或者准则,选取景象信息量大、稳定性好且具有唯一性的图像作为景象匹配基准图,以保证飞行过程中获得的实时图与预先存储于飞行器载计算机中的基准图能够成功匹配。

适配区选取主要通过人工选取或自动选取进行。人工选取方法容易受到操作人员的经验、专业素质等因素的影响,难以客观快速地找出满足要求的适配区。在自动选取方法方面,传统方法是通过对图像特征指标进行分析,按照一定准则确定满足要求的适配区。由于图像特征指标之间相互依存和互相制约,采用单一特征指标作为依据选取适配区的方法,只能反映图像某一方面的适配性,不能得到令人满意的结果。因此,在选取适配区时,应该综合多种影响因素恰当选取某些特征指标,构建能够全面反映图像适配性的综合特征量。

在机器学习领域,适配区选取可归结为模式识别中的分类问题,即输入子区为多维统计特征向量,输出为该子区是否适合匹配的二值问题,因此可以运用模式识别理论,采用有监督学习方法,通过计算各子区特征,建立特征与子区适配性之间的映射关系。在深度学习领域,在构建地物类别代表性好、样本数量足够大、正负样本数量均衡的训练样本数据集的基础上,通过设计和训练深度学习模

型,可以自动学习适配区的适配性特征,对输入子区进行适配区与非适配区的自动预测。

2. 前视景象匹配的基准图制备

在前视景象匹配任务中,基准图为目标在某一视点和角度下的模板图像。由于飞行器载图像传感器相对于目标的成像位置和成像角度在不断变化,实时图中目标的位置和尺寸也在不断变化,每次模板匹配时所需基准图均不相同,因此需要制备航迹上不同距离不同角度的基准图序列,匹配时利用惯性导航系统提供的粗略的飞行器位置和方位信息,从基准图序列中选取合适的基准图进行景象匹配^[4]。

前视景象匹配中最常应用的是前视红外目标匹配识别。在这种应用中,实时图为红外图像,基准图通常是可见光图像,或是由可见光图像反演生成的红外图像。

当飞行器飞行高度较低时,实时图中的目标表现特征以侧面为主,此时基准图的制作需要根据目标区域的卫星遥感可见光正射影像、测绘地形数据、目标与场景的典型地物特征、飞行器视角参数等信息,利用坐标变换对目标进行视角变换,生成目标区域可见光三维场景,由卫星视角的目标顶部影像生成得到飞行器视角的目标侧面影像,然后利用红外辐射特性对目标进行灰度变换,得到所需视角的目标区域红外仿真图像,截取得到目标红外基准图,作为景象匹配所需模板,流程如图 1.4 所示。

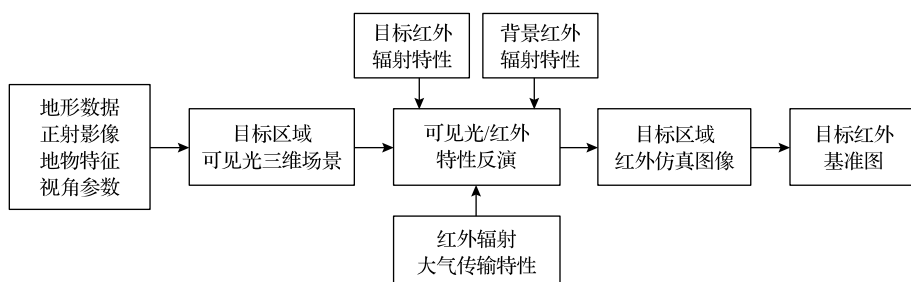


图 1.4 前视红外基准图制备

当飞行器飞行高度较高时,实时图中的目标以顶面为主,基准图制备不需要生成目标区域三维场景,成果形式主要有以下三种。

(1) 可见光基准图。基于目标区域的可见光卫星遥感正射影像,按照不同距离和不同视角要求,制备目标基准图序列。

(2) 特征基准图。对目标区域的可见光卫星遥感正射影像进行特征提取,如边缘、轮廓、结构等在可见光图像和红外图像中都比较稳定的共性地物特征,生成

特征影像，从中截取生成目标基准图序列。

(3) 红外基准图。从目标区域的可见光卫星遥感正射影像出发，根据对典型目标的红外辐射特性分析，采用红外图像反演方法仿真生成目标区域的红外图像，然后截取生成目标红外基准图序列。采用红外基准图，可以在景象匹配中实现红外实时图和红外基准图的直接匹配，充分使用目标的灰度信息。

由此可见，红外图像反演是前视红外基准图制备中非常重要的技术。目前，主流的红外图像反演技术可以分为两大类：一是基于红外特性建模的图像仿真技术，二是基于深度学习的红外图像仿真技术。前者是在建立与呈现场景红外特性相关数学模型的基础上，通过计算机仿真显示场景的红外仿真效果；后者是利用深度学习算法对大量可见光场景及其红外特性效果进行训练，得到场景的可见光与红外图像间的映射模型，而后依据此模型生成可见光场景的红外特性，并以图像的形式显示出来。随着人工智能的发展，基于深度学习的红外图像仿真技术逐渐成为发展趋势，得到了广泛的应用。

1.2.2 图像匹配算法

图像匹配的本质是寻找同一场景的不同时间、不同条件、不同传感器获得的图像在空间上的对应关系，参与匹配的图像可能存在平移、仿射等线性变换和各种非线性变换。在精确制导领域，惯性导航系统能提供高度、航向等辅助信息，可以对实时图进行校正处理，使得校正后的实时图与基准图之间的尺度差异和视角差异相差不大，图像间变换可以用平移变换和仿射变换等简单模型拟合。在无人机视觉定位领域，无人机所拍摄的实时图一般为大倾角影像，与卫星遥感正射影像所制备的基准图具有较大的视角差异，这种差异需要用投影变换和非线性变换等复杂模型进行拟合。

按照分类标准不同，图像匹配可以按照以下几种方法进行分类。

(1) 按照待匹配图像的来源是否相同可将图像匹配分为同源图像匹配和异源图像匹配。由于异源图像具有不同的成像原理，同一场景下异源图像具有较大差异，如可见光图像和 SAR 图像，给图像匹配带来较大困难^[5]。

(2) 按照匹配方法可将图像匹配分为基于模板的匹配和基于特征的匹配，其中基于模板的匹配主要用来解决图像间存在平移情况下的匹配；基于特征的匹配可分为点、线、区域等特征，可解决仿射变换、投影变换和非线性变换等情况下的匹配。

(3) 按照采用人为手工设计特征(hand-crafted feature)或采用深度神经网络进行自动特征提取，可将图像匹配分为传统的图像匹配和基于深度学习的匹配。后者需要大规模已标注的图像匹配训练样本集支撑。

下面主要从模板匹配和特征匹配的分类角度，按照从传统算法到深度学习算

法的发展历程，对图像匹配算法进行介绍。

1. 模板匹配算法

模板匹配是景象匹配领域的代表性算法，其基本原理是通过图像对间的相似度计算获取匹配位置，即模板图像遍历搜索图像中的子图，与子图计算相似度，将与模板图像相似度最高的子图所在位置作为匹配结果。模板匹配性能主要取决于相似性度量、搜索策略和匹配后处理等。

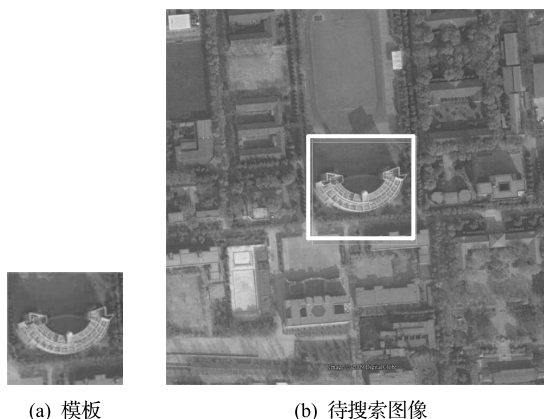
1) 传统的模板匹配算法

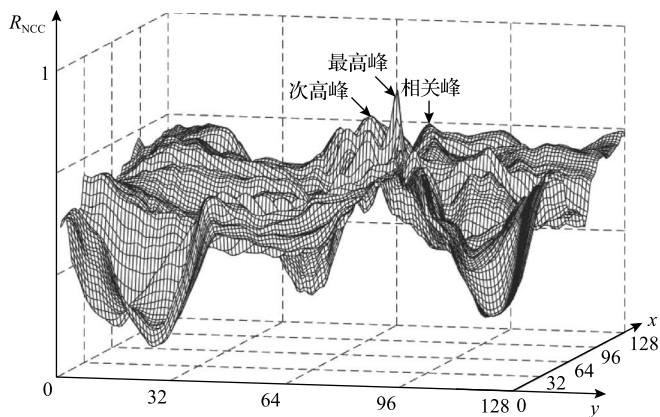
模板匹配采用的相似性度量主要有归一化积相关系数、豪斯多夫 (Hausdorff) 距离、差方和 (sum of squared difference, SSD)、互信息 (mutual information) 和相位相关 (phase correlation, PC) 等，以及为适应图像灰度差异、形变等因素提出的最近邻相似性 (best-buddies similarity, BBS)^[6]、形变多样性相似度 (deformable diversity similarity, DDIS)^[7] 等相似性度量。

归一化积相关算法是比较经典常用的模板匹配算法，采用去均值和归一化来尽可能消除实时图和基准图之间灰度差异对匹配性能的影响，如式 (1-1) 所示：

$$R_{NCC}(u, v) = \frac{\sum_x \sum_y (f_1(x, y) - \mu_1)(f_2(x - u, y - v) - \mu_2)}{\sqrt{\sum_x \sum_y (f_1(x, y) - \mu_1)^2} \sqrt{\sum_x \sum_y (f_2(x - u, y - v) - \mu_2)^2}} \quad (1-1)$$

其中， R_{NCC} 称为归一化积相关系数，显然有 $0 \leq R_{NCC} \leq 1$ ，当 $R_{NCC} = 1$ 时度量值取得极大值，说明两幅图像的信息完全一致，因此根据 R_{NCC} 的大小可以判断图像的相似程度。归一化积相关系数可用如图 1.5 (c) 所示的相关曲面表示，相关函数最大值 (图中相关曲面最高峰) 所对应的坐标值 (u, v) 就是所求的匹配位置。进行





(c) 相关曲面

图 1.5 模板匹配示意图

模板匹配后，相关矩阵主峰所在位置即为理论上的匹配位置。实际情况下，基准图和实时图存在较大差异，真实匹配点往往落在相关矩阵的次高峰上，从而导致匹配失败。

模板匹配常常建立在图像子区灰度信息的直接比较上，使用从图像中提取得到的特征进行匹配更具有图像差异不变性。因此，基于特征的模板匹配在实际中得到了广泛的应用，该类方法的共同点在于，首先对模板和测试图像进行特征提取，得到相应的特征图像，然后对两幅特征图像进行模板匹配，特征可以是边缘、梯度、形状和轮廓等。

模板匹配算法计算过程相对简单，容易理解，可解决大部分平移变换问题，在同源图像匹配时有较好的匹配效果，同时可处理存在一定范围内的形变、遮挡和颜色差异的情况。但是，当待匹配图像对为异源图像或存在复杂几何形变、尺度变化等差异时，模板匹配很难获得准确的匹配结果。因此，选取鲁棒的模板匹配算法和根据模板匹配结果确定最佳匹配位置是该类算法的关键所在。

由于模板匹配需要遍历待搜索图像，运算量随图像尺寸的增加迅速增大，给视觉导航系统的实时性带来影响。因此，在实际应用中，可以采用多种策略提高匹配速度以满足实时性要求，如采用基于快速傅里叶变换 (fast Fourier transform, FFT) 的相位相关法、金字塔快速匹配法等。

2) 基于深度学习的模板匹配算法

随着人工智能技术的发展，基于深度学习的模板匹配算法应运而生，由于深度学习可以从大量训练数据中利用学习—反馈—调整的策略抽取出高层次的语义信息，因此体现出了比传统模板匹配算法更强的适应性。该类算法一般包括特征网络和度量网络两部分，特征网络学习将图像块转换成特征表示，通常使用孪生

(siamese)结构,包括共享权重的孪生网络(适用于同源图像匹配)和不共享权重的伪孪生网络(适用于异源图像匹配),度量网络用于学习相似性度量准则,神经网络的使用使得相似性度量方式更加多样化。

基于斑块(patch)特征的 MatchNet 模型^[8]是深度学习图像匹配的经典模型之一,将匹配问题转换成了分类问题,级联实现生成描述子的孪生网络和全卷积的决策网络,在同源图像匹配上取得了较好效果。MatchNet 模型的网络结构如图 1.6 所示,由于网络结构的优越性,许多基于度量学习的图像匹配算法都采用了 MatchNet 框架,在特征学习和相似性度量方面进行改进。

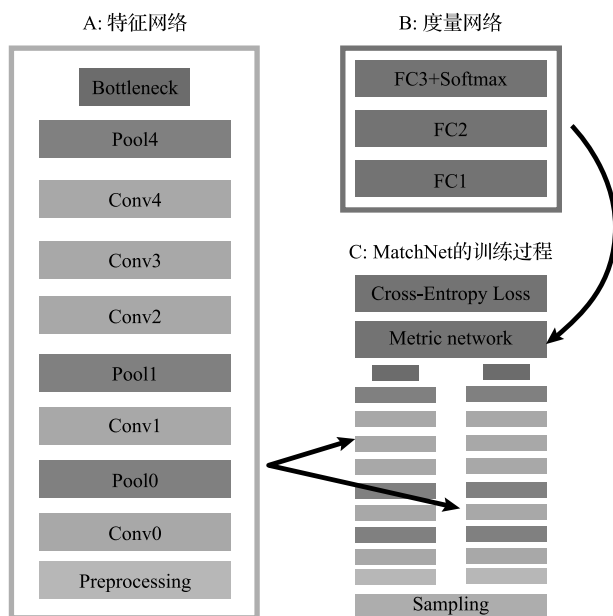


图 1.6 MatchNet 模型的网络结构

Merkle 等^[9]采用孪生全卷积网络提取深度共同特征,使用扩展卷积来增加感受野大小,并通过点积计算产生相似图,度量得到可见光图像和 SAR 图像之间的偏移。DeepDesc 模型^[10]使用卷积网络学习图像块的特征表示,并用 L_2 距离对相似度进行度量,该方法使用了成对的正负样本和铰链损失函数对孪生网络进行训练,并使用难例挖掘来降低正负样本间不平衡的影响。CorrASL 模型^[11]是端到端的深度学习框架,使用基于特征空间的互相关计算和学习的特征得到对应的热度图,提高模板图像在搜索图像中的地理定位精度。

对于异源图像匹配、大视差图像匹配问题,还可将深度学习的生成对抗网络引入图像匹配流程,减小待匹配图像之间的灰度差异和视角差异。例如,在可见光图像和 SAR 图像的匹配中,可将可见光图像反演为同场景的 SAR 图像,然后

再与 SAR 图像进行匹配；在卫星影像和无人机影像的匹配中，可将无人机影像反演为卫星视角影像，然后再与卫星影像进行匹配。

2. 特征匹配算法

特征匹配算法首先从匹配图像中提取特征，然后通过对特征的描述和匹配实现图像匹配。与模板匹配算法相比，特征匹配算法具有更强的环境适应能力，尤其适合于不同传感器图像和成像条件变化较大情况下的图像匹配，因此在飞行器视觉导航中获得了广泛的应用。

1) 传统的特征匹配算法

图像匹配中常用的特征包括点特征(如角点、高曲率点等)、线特征(如直线段、边缘等)、面特征(如轮廓、闭合区域等)。点特征可以直接用于图像匹配，线特征和面特征需要经过某种变换或处理得到可以用于匹配的关键点。

基于特征描述子的图像匹配算法首先检测图像的关键点，并对其建立具有尺度不变性和旋转不变性的特征描述子，然后进行关键点的匹配，最后利用匹配的关键点计算相应的空间变换模型。其步骤如图 1.7 所示，无论采用何种关键点，都需要经过关键点提取、关键点描述和关键点匹配三个主要步骤，最终匹配的关键点即可作为图像匹配所需的控制点。

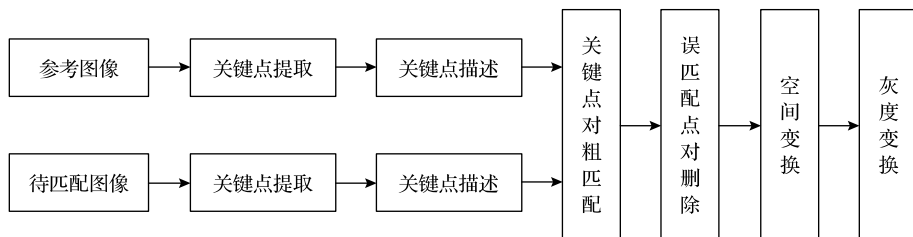


图 1.7 基于特征描述子的图像匹配的一般过程

尺度不变特征变换(scale invariant feature transform, SIFT)特征是一种代表性的关键点特征，这种特征对图像的尺度变化和旋转变化具有不变性，而且对光照变化和图像变化具有较强的适应性^[12]。这种关键点可以是灰度变化的局部极值点，含有显著的结构信息，也可以没有实际的直观视觉意义，但却在某种角度、某个尺度上含有丰富的易于匹配的信息。SIFT 的提出使得点特征突破了角点检测的局限，同时 SIFT 算法也在后续发展中根据实际情况发展出不同的变体，例如，针对 SAR 图像匹配提出的 SAR-SIFT 特征^[13]，针对可见光图像与 SAR 图像匹配提出的尺度不变特征变换(optical-to-SAR scale invariant feature transform, OS-SIFT)特征^[14]，针对 SIFT 算法时间复杂度高提出的加速稳健特征(speeded-up robust features, SURF)等^[15]。点特征匹配可以解决大部分图像匹配问题，但在待