

浙江大学

硕士学位论文



论文题目 基于特征的图像配准算法研究

作者姓名 吴 铮

指导教师 朱 淼 良 教授

钱 徽 副教授

学科(专业) 计算机应用

所在学院 计算机学院

提交日期 2006 年 3 月

摘要

图像配准问题是图像处理里的一个经典问题，在计算机视觉、模式识别、医学图像处理 and 遥感信息处理方面有着广泛的应用。配准问题的目的就是将同一场景的不同图像对齐或匹配，消除存在的几何畸变。

本文在回顾现有的主要配准算法上，提出了一种新的综合了高级特征（曲线特征）和低级特征（点特征）的混合配准算法。此算法充分利用图像中特征曲线的稳定性和点特征定位的精确性，提高了配准算法鲁棒性。曲线的提取以图分割理论为基础，结合了图像中灰度、纹理和空间的多种信息，得到在各种畸变下仍然能保持相对稳定的区域边界为待配准曲线。通过构造一系列曲线局部几何描述算子，利用统计其分布特性来避开直接建立曲线上点的对应关系的难题。在经过曲线配准的粗配准步骤后，我们提出了利用 Gabor 滤波器提取特征点，用快速 NCC 计算建立匹配点对，并在最小平方误差的意义下进一步求解几何变换参数的方法。为了减少计算量，我们建立了高斯金字塔模型，并使用了只在顶层信号中采用曲线配准，在其他层次中采用特征点配准的逐步求精的策略。实验证明，本文提出的算法与传统的基于相关的算法相比，针对复杂场景的图像以及存在复杂畸变的情况下仍能有效的估计出几何变换参数。

关键词：图像配准，图分割，Gabor 滤波器，高斯金字塔结构

ABSTRACT

Image registration is a classic problem in image processing, and it is extensively applied to the field of computer vision, pattern recognition, medical image processing and remote sensing information processing. Basically, the goal of image registration is to align images of same scene by removal of the potential geometrical distortion existed.

In this paper, we propose a new hybrid registration algorithm combined with high level features (salient curve) and low level features (critical points), based on the analysis of related works. The hybrid algorithm improves the robustness effectively by using both the stableness of salient curves in image and accuracy of critical points. The curves that are stable under various distortions are extracted in Graph-cut Theory, which is fully using gray level, texture and spatial information of image. Through construction of a series of local geometry descriptors and statistical analysis of their distributions, we avoid the difficulties of direct point matching problem. When the coarse registration is finished, we further adopt Gabor filters to locate several critical points and establish the correspondence by fast normalized correlation coefficients. Then we solve the transform parameters in the means of least square error method. In order to reduce the computation, we build Gaussian pyramid structure. Curve matching is only applied in the top level while the point matching is used in the rest level so that we can solve the transform parameters in a coarse-to-fine procedure. The experiment shows that the algorithm we proposed is much effective than the usual correlation-based algorithm, even the images have complex scene and the distortion is great.

Key Words: Image Registration, Graph cut, Gabor filter, Gaussian Pyramid

第一章 绪论

1.1 图像配准问题的背景

图像配准问题是计算机视觉、图像处理中的基本问题，有两种对应的问题原型：一是两副（或者多幅）来自不同采集设备或者不同视角的图像需要对比，经过配准步骤可以得出两副图像的差别所在；二是根据一幅已知的模板图像在另外一幅图像中搜索类似模板的结构。这两种基本需求在多个领域的很多实际问题上都有所反映：如，不同传感器的信息融合；不同时间、场地条件下获得图像的差异检测；成像系统和物体场景变化情况下获得的图像的三维信息获取；图像中的模式或目标识别等等。

简单来说，图像配准就是将同一场景的不同图像“对齐”或进行广义的匹配，以消除存在的几何畸变。对同一场景使用相同或不同的传感器(成像设备)，在不同条件下(天候、照度、摄像位置和角度等)获取的两个或多个图像一般都会有差异。同一场景的多次成像的差别可以表现在：不同的分辨率、不同的灰度属性、不同的位置(平移和旋转)、不同的尺度、不同的非线性变形等等。多种成像模式产生的图像(称之为多模态图像)会表现出不同的分辨率、不同的灰度属性等等差异。图像配准的主要目的是消除几何上的差异，对于灰度信息的差异并不关心。当然，正是因为图像畸变中的灰度差异的存在，给几何校正增大了难度。

图像配准的应用领域概括起来主要有以下几个方面：

- 1、计算机视觉和模式识别。包括图像分割、物体识别、形状重建、运动跟踪和特征识别。
- 2、医学图像分析。包括医学成像信息诊断，生物医学信号处理等。
- 3、遥感信息处理。包括特定目标的定位和识别等。

根据 Brown^[1]的总结，图像配准在应用上可以粗略地归为四类：

- 多模态配准(multimodal registration)

问题描述：由不同传感器获得同一场景上的图像的配准。

典型应用：多模态图像的信息融合，指导进一步的图像分割。

方法特点：通常需要建立传感器模型和变换模型；

由于灰度属性或对比度可能有很大的差异，有时需要灰度的预配准；

利用手工设定的基准点可以简化问题。

应用例子 1：医学图像领域，CT，MRI，PET，SPECT 图像信息融合。

应用例子 2：遥感图像领域，多电磁波段图像信息融合，如微波、雷达、可视的或多谱段的等；场景分类，如分类建筑物、道路、车辆植被等。

- 模板配准

问题描述：在图像中寻找具有指定模板结构的匹配问题。

典型应用：在图像中识别和定位模板样式，例如地图、物体、目标物等。

方法特征：基于模式，预先选定特征，已知物体属性或高级特征进行匹配。

应用例子 1：遥感数据处理，定位和识别已知特征的目标物体如飞机场、高速公路、车站、停车场等。

应用例子 2：字符识别，签名认证，波形分析等。

- 视角配准

分类：对从不同观察视角获得的图像进行配准。

典型应用：深度或形状重建。

方法特征：变形多为透视变换；

常应用视觉几何和表面属性等方面的假设条件；

典型的方案是特征相关；

必须考虑阻挡问题。

应用例子 1：从视角差异中构建三维深度和形状信息。

应用例子 2：运动跟踪，序列图像分析。

- 时间序列配准

分类：对同一场景上不同时间或不同条件下获得的图像的配准。

典型应用：检测和监视图像内容变化。

方法特征：需要容忍图像中部分内容的差异和形变对配准造成的影响；

有时需要建立传感噪声和视点变换的模型。

应用例子 1：SAR 图像、医学图像处理，数字剪影血管造影(DSA)，注射造影剂前后的图像配准。

应用例子 2：遥感数据处理，自然资源监控。

对于图像配准技术的分类和比较评判通常很困难，因为每一个配准技术通常是针对某一个具体的应用而设计的，而对于那些特定的应用问题来说，并没有哪一个技术是必须的和唯一的。但是，所有以上阐述的配准问题，归根结底都是寻找一种变换，在这种变换下，两副图像之间可以建立像素点的对应关系。通常来说，寻找这种最优变换模式的方法取决于两副图像之间畸变的类型。常见的有以下图像畸变的类型：

- 1、信号噪音
- 2、视角变化带来的透视变换
- 3、物体变化包括物体的运动和变形
- 4、光照变化包括阴影
- 5、不同传感器输出同一物体的不同成像

1.2 图像配准方法的基本框架

在 Brown 综述^[1]中, 配准方法被分解成四个部分的组合:

- 1、特征空间
- 2、搜索空间
- 3、搜索策略
- 4、相似性度量

1.2.1 特征空间

特征空间指从图像中提取用于匹配的特征。

图像配准中使用的图像特征有着重要的意义, 因为它们通常决定了这个方法适合于什么样的图像。基本的像素灰度值可以是一种特征, 还可以包括常见的特征如边缘、曲线、曲面; 显著的视觉特征如拐角, 交叉线, 高曲率的点等; 统计特征, 如不变矩以及高层语义的描述算了等等。图像特征几乎是所有计算机视觉和图像处理任务都会涉及到的基本因素, 在图像配准上它会影响:

1. 成像设备和景物中哪种特征会对算法敏感(通常选择的特征都会降低图像噪音和畸变的干扰)
2. 图像的何种特征被匹配(比如需要的是匹配景物的结构而不是纹理信息)
3. 算法执行的效率和搜索代价

因此, 好的特征选择将会消除畸变噪音的干扰, 降低参与计算的数据量又不失充分的表达图像内容的内部结构。

1.2.2 搜索空间

搜索空间是一系列可以配准图像的变换操作的集合, 其中以几何变换为主要因素。

图像的几何形变可以分为三类: 全局的、局部的和位移场形式的。全局的变换指整副图像的畸变都可以用统一的参数矩阵来描述。典型的全局几何变换包括以下的一种或几种简单变换的组合: 平移、旋转、各向同性或各向异性的缩放、二次或三次多项式变换等等。局部变换允许变换参数有位置依赖性, 在图像中不同的位置上具有不同的变换参数矩阵, 一般变换参数只是定义在特定的关键点上, 而在其他区域进行插值。位移场方法, 有时候又称为光流场法, 使用一个连续函数优化机制, 为图像中的每一点计算出一个独立的偏移量, 并使用某种规整化机制进行约束。一般来说, 变换类型在某个特定应用场合可以作为一种先验知识, 将搜索空间限制在几种可能出现的特定变换。当没有先验知识参考时, 就必须考虑所有可能出现的变换形式。先验知识参考时, 就必须考虑所有可能出现的变换形式。

1.2.3 搜索策略

搜索策略决定了下一步可选的变换的依据，用于最优变换的评估。

由于很多配准算法伴随着庞大的计算量，常规的贪婪搜索法实际上无法实现，因此有效的搜索策略成为一个不容忽视的问题。给定一组对应特征和特征间参数化的变形，由搜索空间和相似性度量方法决定了搜索策略。常用的搜索策略有：松弛匹配、动态规划、牛顿法、最速下降法、共扼梯度法等等。

1.2.4 相似性度量

相似性度量决定了每一次变换优劣评估的结果。

相似性度量往往和提取的图像特征相关，它给出的评价价值将直接决定了配准变换的选择。在从每幅图像中提取图像特征后，由相似性度量函数来计算决定在当前所取的变换模型下图像是否被正确匹配了。通常，配准算法抗干扰的能力是由特征提取和相似性度量所共同决定的。最常见的相似性度量方法为特征空间的欧式距离，一般这样的相似性度量必须考虑归一化。

以上四个方面的组合，就构成了配准算法的基本框架。几乎目前所有的配准算法都可以分解成这样四个步骤。

1.3 图像配准方法的区别

尽管图像配准方法有着共同的框架，如前所述的四方面组合。但是不同算法在四个方面的选择主要是根据特定的应用背景，具体的说是图像变换类型、图像内容变化类型和参数计算方法所决定的。这里所说的图像变换类型指如何定义将一幅图像的像素点位置映射到另外一幅图像中的新位置上；图像内容变化是指图像对应像素的灰度值以及坐标的畸变；参数计算方法指计算出最优变换参数所采用的算法。

具体区别归纳起来如表 1.1 所示。

在此要特别指出，本文涉及到的图像配准技术主要针对 2 维平面数据。随着成像仪器设备的发展，目前出现了更多的 3D 体数据的配准需求，因此 3D 数据的配准也成为目前学术界的研究热点之一。然而，2D 数据的配准还没有得到完全解决，且 2D 数据的配准技术可以很容易扩展到 3D 数据上，所以本文的探讨基础和试验都建立在 2D 数据上。2D 数据配准算法和 3D 数据配准算法的区别在本文中不做讨论。

表 1.1 配准算法的区别

映射关系	
全局	映射关系由单一函数构成，函数的自变量为图像像素的坐标位置，且在图像上所有像素点都套用此公式。换句话说函数的参数不依赖于像素位置
局部	映射关系由多个函数构成，在不同的像素位置上有不同的函数形式
畸变类型	
全局	在整副图像域上灰度发生类似的改变，比如整副图像上的椒盐噪声
局部	变化取决于图像像素的坐标，比如投影变换下畸变程度决定于物体不同部位距离成像平面的深度
计算方法	
全局	使用整副图像的信息来计算变换参数
局部	只使用局部信息来分别计算局部变换参数

1.4 图像配准技术研究现状

在过去的几十年里，图像采集设备发展惊人，大量图像信息的获取增加了对自动配准技术的需求。尽管早在 1983 年 Ghaffary^[2]就发表了关于配准的综述，但是内容基本上是针对相关算法(Correlation)的。第一篇影响深远的综述性文章由 1992 年 Brown^[1]发表。特别针对医学图像配准的在文献^[3,4,5]中有阐述，而在文献^[6,7,8]中概括了遥感图像配准的方法。另一方面，根据 Institute of Scientific Information (ISI)的调查数据表明，最近 10 年里至少有超过 1000 篇的学术论文在研究图像配准问题。那些在 92 年之前就发表的算法，很多至今仍然在使用。2002 年网上公布的美国专利申请中，图像配准相关的部分就有超过 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医学图像配准问题。而在最近的国际顶级学术会议上（2004 年 CVPR 会议和 2005 年 ICIP 会议）都有关于配准的专题讲座，这足可见图像配准问题的关注程度。

一方面配准问题持热，另外一方面也说明仍然有许多技术难题没有解决且急需解决。配准问题的定义本身很简单，然而图像内部结构的复杂性使得配准的精确性、鲁棒性很难得到准确的检验。这也解释了为什么上千种的算法中，几乎找不到通用的、一劳永逸的算法。

1.5 本文的工作和组织

第一章，绪论。概述图像配准问题的背景和基本术语介绍。

第二章，配准理论和基本方法。查阅了国内外图像配准相关领域的大量文献资料，对图像配准及其相关技术进行了深入研究和总结。按技术类别对图像配准方法进行了详细的综述，分析了各种方法的适用性、优点和不足之处。

第三章，基于局部几何特征分布的曲线配准。对原有的基于边缘特征配准技术做出改进，引进增强鲁棒性的基于图分割理论的曲线提取算法。设计出基于曲线局部几何描述算子的配准方法，并给出试验证明。

第四章，基于特征点匹配算法。在曲线配准之后附加点匹配算法，利用 Gabor 滤波提取图像中的关键点，利用 SVD 方法求解点集之间的变换参数，并给出试验证明。

第五章，混合配准算法设计以及优化。提出综合曲线配准和点配准的由粗到精的分层配准策略，包括多分辨率模型、快速 NCC 计算和并行计算的探讨，并给出实现步骤。

第六章，总结和展望。

第二章 配准理论和基本方法

2.1 配准问题定义

图像配准问题可以定义为两副图像像素座标和灰度值上的双重映射。

设图像分别为 I_1 、 I_2 ， $I_1(x, y)$ 、 $I_2(x, y)$ 分别代表像素点的灰度值。如上所述的映射关系为：

$$I_2(x, y) = g(I_1(f(x, y))) \quad (2.1)$$

其中 f 为 2D 空间几何变换， g 为灰度变换；

配准问题就是要寻找最优的变换 f 、 g ，在此变换下将两副图像最大程度得对齐，从而可以进一步处理图像。通常情况下灰度变换关系的校正并不是必要的，所以寻找空间几何变换关系 f 便成为配准的关键所在。于是上式可改写成更为一般的表示方式：

$$I_2(x, y) = I_1(f(x, y)) \quad (2.2)$$

常见的简单变换形式有：

平移： $I_2(x, y) = I_1(x + a, y + b)$

旋转： $I_2(x, y) = I_1(x \cdot \cos \theta - y \cdot \sin \theta, x \cdot \sin \theta + y \cdot \cos \theta)$

亮度： $I_2(x, y) = a \cdot I_1(x, y) + b$ ，如果是不同光照条件影响，亮度变换会更复杂

噪音： $I_2(x, y) = I_1(x, y) + n(x, y); n(x, y) \sim (0, \sigma)$

平滑和锐化：这是原图像与点扩散函数的卷积 $I_2(x, y) = I_1(x, y) * n(x, y)$

2.2 空间几何变换

最常见的几何变换包括简单刚体变换，仿射变换，投影变换，透视变换和全局多项式变换。简单刚体变换指物体或者镜头的运动不影响成像物体保持相对尺寸和形状，一般由平移、旋转和尺度 3 种变换合成。仿射变换比刚体变换更具有一般性，在更复杂的畸变下仍然能保持很好的数学特性。投影、透视变换针对的是物体在不同距离条件下成

像所产生的畸变，通常需要知道物体成像时离镜头的具体距离才能有效的配准。多项式变换则是全局变换下最一般的情况。如果需要考虑图像内非刚性物体的因素，则弹性变换模型更加适用。

以下将具体介绍几种重要的变换的数学模型^[1]。

a) 刚体、仿射变换模型

仿射变换是配准中最常用的一类转换模型，由平移、旋转、尺度的笛卡尔变换构成，属于全局刚体变换，具有将平行线转换成平行线和有限点映射到有限点的一般特性。

仿射变换通常具有四个参数： t_x ， t_y ， s ， σ 。若点 (x_1, y_1) ， (x_2, y_2) 分别为基准图像与模板图像中对应的两点，则有：

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} + s \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \quad (2.3)$$

仿射变换式：

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} + \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \quad (2.4)$$

则是更一般的形式，除了尺度、平移、旋转外，它还考虑了畸变中的倾斜、纵横比变化。

b) 透视变换

透视变换是三维空间到二维空间的映射。

若 (x_0, y_0, z_0) 为三维场景内的一点，那么它在二维平面的对应点 (x_i, y_i) 如下表示：

$$x_i = \frac{-fx_0}{z_0 - f} \quad (2.5)$$

$$y_i = \frac{-fy_0}{z_0 - f} \quad (2.6)$$

其中 f 表示相机镜头焦距。

c) 投影变换

若场景由三维空间中一相对于图像平面倾斜的平面构成，这时便需要投影变换将场景平面映射到水平的图像平面上。

以 (x_p, y_p) ， (x_i, y_i) 分别表示场景平面与图像平面内的对应两点，两者的映射关系如下表示：

$$x_i = \frac{a_{11}x_p + a_{12}y_p + a_{13}}{a_{31}x_p + a_{32}y_p + a_{33}} \quad (2.7)$$

$$y_i = \frac{a_{21}x_p + a_{22}y_p + a_{23}}{a_{31}x_p + a_{32}y_p + a_{33}} \quad (2.8)$$

2.3 配准算法综述

在第一章中曾提到关于配准算法的基本框架，包括特征空间、搜索空间、搜索策略和相似性度量方法。按照这四个组成部分侧重点比较，可以给出图 2.1 的基本算法分类图示。

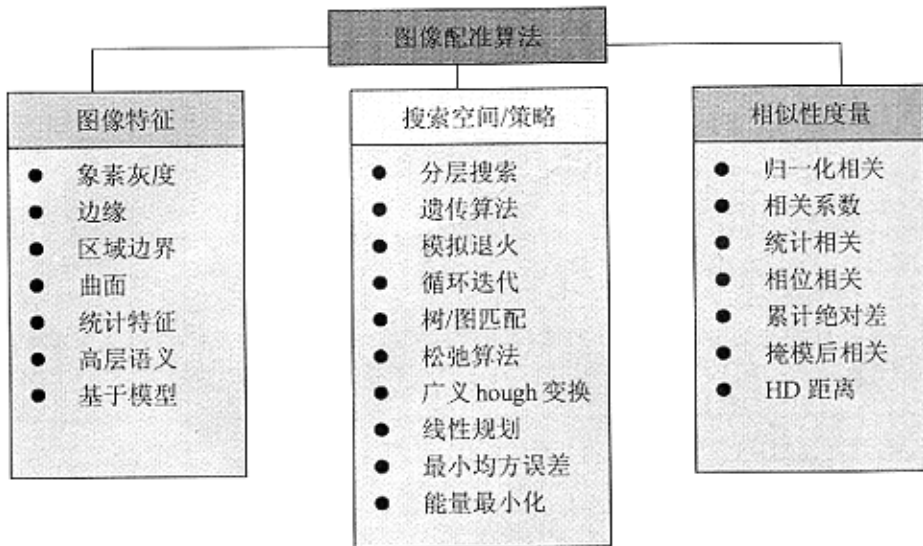


图 2.1 基本配准算法分类

2.3.1 基于空间域互相关的方法

空间域互相关是图像配准中最早出现的一类算法。相关运算在图像配准上的应用通常定义如下：给定图像 I 中的窗口 $W(dim_r \times dim_c)$ ，在图像 I' 中的对应窗口位置通过计算如下式子的最大值得到相关极值点：

$$S(t) = F(M, M'_t) \quad (2.9)$$

其中 t 是 M' 与 M 之间可能存在的位移, F 是某种相似性度量函数。 $F: \mathbb{R}^2 \rightarrow \mathbb{R}$

给出 M' 与 M 的两个实例窗口 W' 和 W , 两种典型的相似性度量函数 F 是:

$$F(W_1, W_2) = \sum_p W_1[p] W_2[p] \quad (2.10)$$

$$F(W_1, W_2) = - \sum_p (W_1[p] - W_2[p])^2 \quad (2.11)$$

第一种叫做 Cross-Correlations, 第二种叫做 SSD (Sum of Squared Differences), 通常都采用归一化后的计算。在经典的归一化互相关 (Cross Correlation) 算法和它的变形^[10]中, 相似性度量定义为(2.10)式, 其最大峰值的位置为检测到的平移量。如果需要亚像素的检测精度, 那么 CC 矩阵就需要做插值。

经典的归一化相关系数的计算定义如下:

$$\rho = \frac{\left| \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (A_{ij} - \bar{A}) \cdot (B_{ij} - \bar{B}) \right|}{\sqrt{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (A_{ij} - \bar{A})^2 \cdot \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (B_{ij} - \bar{B})^2}}, \quad 0 \leq \rho \leq 1 \quad (2.12)$$

其中 \bar{A} 、 \bar{B} 分别是 A 、 B 两幅图像内的灰度均值。

与相关紧密联系的是贯序相似性检测算法(SSDA)^[11], 它使用贯序搜索策略和较互相关更为简单的距离度量。通过(2.11)式, 在当前点的匹配窗口内, 按像素逐个累加模板图像和目标图像的灰度差值, 同时记录累加点数。若在累加过程中灰度差值的累加值达到了预先设定的门限值, 则停止累加, 转而计算下一点。从而省去大量的非匹配位置处的无用计算。当所有点都计算完后, 取最大累加点数的位置为匹配点。这种方法比 CC 计算准确性差, 但是计算速度更快。

基于相关计算的算法缺点主要有三个方面:

1. 相关计算通常只用矩形滑动窗口, 这种计算思路往往只适合局部平移的变换类型。如果一个图像的畸变稍微复杂一些, 比如加轻微的旋转, 相关计算的矩形窗口就很难同时覆盖原始图像和畸变图像的相同区域。Goshtabsy^[12]提出了圆形窗口用于检测旋转问题, 然而它仍然不能解决更复杂的变换情况。对一些更为复杂的几何畸变, Berthilsson^[13]在窗口区域内对每种可能的变形都尝试做相关计算; Simpler^[14]提出分而治之的系统, 对投影变换和镜头产生的畸变使用不同的相关算法。然而, 随着可供选择的变换增多, 算法负担会急剧上升。
2. 基于相关的算法大多强调特征的匹配步骤而淡化特征的提取, 通常这些算法都不去尝试检测图像中的显著结构特征, 特征对应关系的建立是通过直接利用整副图

像域或者一个窗口大小内的图像信息^[9,10]。当图像中含有自相关结构时，比如重复的纹理结构、或者大片的平滑区域，相关计算会很容易输出错误的匹配结果。

3. 各种经典的互相关计算直接利用图像像素灰度，没有分析图像的内容，所以对变灰度的变换非常敏感，在改变光照条件和成像条件下往往此类算法都是失败的。

普通互相关计算的结果曲面往往较为平坦，峰值不明显。锐化峰值的手段有多种，Pratt^[9]在配准前先为图像滤波以增强配准算法抗噪音能力；Van Wie^[15]和 Anuta^[16]采用了基于边缘的相关计算，使用分别提取边缘后的二值图像而不是原始的灰度图。这种思想的扩展，叫做向量相关法，计算相关时采用的是窗口信息的各种表达方式（比如不变矩描述算子^[12]）。

2.3.2 基于 HD 距离的方法

Huttenlocher^[17]提出了基于 Hausdorff 距离的相似性度量，发现在有突变像素位置存在时，HD 距离度量甚过互相关计算。之后 10 年内又衍生出各种扩展 HD 距离计算，并广泛用于图像配准中。

● Hausdorff 距离

给出两个有限点集： $A: \{a_1, a_2, \dots, a_m\}$ ， $B: \{b_1, b_2, \dots, b_n\}$ ，则 Hausdorff Distance 定义为：

$$D_H = \max(D_{AB}, D_{BA}) \quad (2.13)$$

其中 D_{AB} 定义为点集 A 到 B 的距离：

$$D_{AB} = \max_{a \in A} (d_{a,B}) \quad (2.14)$$

$$d_{a,B} = \min_{b \in B} (d_{a,b}) \quad (2.15)$$

● 部分 HD 距离

$$D_{AB} = K^{th}_{a \in A} (d_{a,B}) \quad (2.16)$$

求取部分 Hausdorff 距离，首先求出点集 A 中所有点到点集 B 的距离，然后将这些距离由小到大排序，其中序号 k 的距离即为 K_{th} HD 距离，它排除了一些外来点的干扰。其中 $k = f \times N_A$ ， N_A 表示 A 集合的点数，f 为给定分数，也有自适应计算 f 方法。

● MHD 距离

$$D_{AB} = \frac{1}{N_a} \sum_{a \in A} d_{a,B} \quad (2.17)$$

MHD 用取平均的方法代替 HD 中的取极值。Dubuisson^[18]将 MHD 与传统的 HD 距离以及它的简单变化做对比，证明 MHD 在目标识别上能更好的抵抗遮挡物的影响。

- **Censord HD 距离**

$$D_{\alpha,\beta} = \max(D_{\alpha,\beta}(A, B), D_{\alpha,\beta}(B, A)) \quad (2.18)$$

$$D_{\alpha,\beta}(A, B) = Q_{\beta, a \in A}(D_{\alpha}(a, B)) \quad (2.19)$$

$$D_{\alpha}(a, B) = Q_{\alpha}(d(a, b), b \in B) \quad (2.20)$$

其中 Q 函数即类似与部分 HD 距离的定义，从小到大排序所有距离值；

$\alpha, \beta \in (0, 1)$ 即为部分 HD 定义中的 f 。

Azencott^[19]认为 CHD 经过两次取“部分”操作后比传统 HD 能更好的抵抗噪声。

- **M-HD 有向距离**

$$h_M(A, B) = \frac{1}{N_A} \sum_{a \in A} \rho(d_B(a)) \quad (2.21)$$

$$\rho(x) = \begin{cases} |x|, & |x| \leq \tau \\ \tau, & |x| > \tau \end{cases}, \quad \tau \text{ 是阈值，用于减少遮挡点的影响}$$

- **LTS-HD 有向距离定义为：**

$$h_{LTS}(A, B) = \frac{1}{H} \sum_{i=1}^H d_B(a)_{(i)} \quad (2.22)$$

其中 $H=h \times N$ ，类似与部分 HD 的概念；

$d_B(a)_{(i)}$ 为从小到大排序的第 i 个值。

Sim^[20]提出以上 2 种修正的 HD 定义，均提高了匹配问题中抗遮挡点的能力。

2.3.3 基于 FFT 的方法

一个引入频域配准技术的直接原因是将一般的空域相关求极值转化到频域求极值可以进一步提高检测速度：

$$\iint f(x, y)g(x+a, y+b)dxdy \Leftrightarrow F(u, v)G^*(u, v) \quad (2.23)$$

其中 $F(u, v), G(u, v)$ 分别是原函数 f, g 的傅立叶变换， $*$ 表示共扼

频域配准技术的特性主要有以下几点：

1、较大的位移检测范围

理论上，当两幅图像仅存在位移变化时，频域相关算法能检测的 x , y 方向的范围分别为图像长宽一半。

2、较尖锐的检测峰值

只有当两幅图像相关时，检测结果才为一 δ 函数。从而有尖锐的检测峰值，所以能实现图像的精确匹配，我们可以利用这一点来区别两幅图像是否相关。传统相关方法容易出现多个峰值，而且最大峰值的定位不是最准确。相位相关的峰值比较少，而最大峰值相对来说比较高，比较容易判断。

为了获得亚像素精度，还可以用曲面拟合的方法来确定最高峰值的位置^[21]。

假设 $PC(x, y) = ax^2 + by^2 + cxy + dx + ey + f$ ，作为 2 阶曲面拟合方程；

构造 $AX = B$ 方程组求解上式中的 6 个参数

$$A = \begin{bmatrix} x_0^2 & y_0^2 & x_0 y_0 & x_0 & y_0 & 1 \\ x_1^2 & y_1^2 & x_1 y_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_8^2 & y_8^2 & x_8 y_8 & x_8 & y_8 & 1 \end{bmatrix} \quad X = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} \quad B = \begin{bmatrix} PC_0 \\ PC_1 \\ \vdots \\ PC_8 \end{bmatrix} \quad (2.24)$$

B 中的值为相位卷积结果，取最高峰值点和周围 8 个领域点。用回归方法可以求得 X 的解。然后 $PC(x, y)$ 求偏导可以得到精确的最高峰位置：

$$x = \frac{2db - ce}{c^2 - 4ab} \quad (2.25)$$

$$y = \frac{2ae - dc}{c^2 - 4ab} \quad (2.26)$$

3、对图像灰度依赖小

可证明，当两幅图像间还存在某一灰度差或仅有灰度翻转时，这种差别在检测结果中只表现为在 δ 函数上加一恒量。显然，这并不影响检测效果，由于频域相关法对图像灰度依赖小，它抗图像遮挡的能力亦很强。另外，由于归一化过程白化了噪声，使得相位相关同样对固定的校准误差、光照的均匀变化有一定抗干扰能力。

4、图像旋转对频域匹配效果影响较大

实验证明，图像旋转大于 5 度时，其相关信噪比衰减为零。用低通滤波的频域相关算法可明显减小几何失真及噪声对匹配的影响，但精度会降低。要获得更高的匹配概率和匹配精度，应进行图像旋转的校正。

假设 $f_2(x, y)$ 是 $f_1(x, y)$ 经过平移 (x_0, y_0) 和旋转 θ_0 得到的图像, 即:

$$f_2(x, y) = f_1(x \cos \theta_0 + y \sin \theta_0 - x_0, -x \sin \theta_0 + y \cos \theta_0 - y_0) \quad (2.27)$$

傅立叶变换后可以得到:

$$F_2(u, v) = e^{-j2\pi(x_0 u + y_0 v)} F_1(u \cos \theta_0 + v \sin \theta_0, -u \sin \theta_0 + v \cos \theta_0) \quad (2.28)$$

$$\|F_2(u, v)\| = \|F_1(u \cos \theta_0 + v \sin \theta_0, -u \sin \theta_0 + v \cos \theta_0)\| \quad (2.29)$$

可见两幅图像频域上的能量是相同的, 只差一个旋转角度, 因而在变换到极坐标系后角度差就可以用相位相关的方法得到:

$$M_1(\rho, \theta) = M_2(\rho, \theta + \theta_0) \quad (2.30)$$

其中 M 为图像的频域能量。

同样, 利用 Log-Polar 坐标变换可以将尺度和旋转角度转化为 Log-Polar 坐标下的相位相关计算^[22,23,24]。

2.3.4 基于互信息量的方法

互信息量方法是目前在多模态图像配准中使用最多的一种, 被广泛用于医学图像和遥感图像。互信息 (Mutual Information, 简称 MI) 的概念源自信息论, 用于衡量两组数据在统计意义上的依赖性, 所以特别适合多模态图像的配准。

以下给出互信息基本定义:

1. 定义自信息的数学期望为平均信息量, 即

$$H(X) = E\left[\log \frac{1}{P(a_i)}\right] = -\sum_{i=1}^q P(a_i) \log P(a_i) \quad (2.31)$$

此即信源的信息熵, 代表了信源输出后所提供的平均信息量, 或信源输出前的平均不确定度。

2. 定义平均互信息

已知随机变量 A 、 B 的联合概率密度为 $P_{AB}(a, b)$, 则互信息量为一个变量包含另一个变量的信息量。

$$I(A, B) = H(A) + H(B) - H(A, B) \quad (2.32)$$

$$H(A, B) = -\sum P_{AB}(a, b) \log P_{AB}(a, b) \quad (2.33)$$

3. 定义条件熵

$$H(A/B) = -\sum_{a,b} P_{AB}(a,b) \log P_{AB}(a/b) \quad (2.34)$$

$$H(B/A) = -\sum_{a,b} P_{AB}(a,b) \log P_{AB}(b/a) \quad (2.35)$$

以上两式分别是已知 B 的 A 的条件熵，和已知 A 的 B 的条件熵，所以互信息也是随机变量 A 的不确定度减去已知 B 的 A 的不确定度。

由互信息的概念可知，若两副图像的互信息越大，说明两副图的相似性越高。最早在文献^[25]中此技术被应用到 MRI 图像配准和 3D 模型与真实场景的匹配上，MI 通过梯度下降法达到极大值。

寻找 MI 函数下的极值，是一个多维函数的优化问题，这里的维数指得是畸变类型参数的自由度。最简单的解决办法是贪婪搜索，尽管它很费时，但是如果只有简单畸变类型存在时也会经常采用。在自由度比较高的时候，需要一些更复杂的优化思想。常见的有 Gauss-Newton 数值下降法^[26]、Levenberg-Marquardt 优化法^[27]，而 Maes 使用了 Powell 的多维搜索策略^[28]。Unser^[29,30,31]使用了各种策略来解决 MI 配准中单步更新的准确性，他们使用了 Parzen 窗口来计算联合概率密度，用 Marquardt-Levenberg 方法来最大化 MI。为了提高计算速度他们也使用了图像多分辨率结构。Ritter^[32]使用了层次搜索策略和模拟退火来寻找 MI 的极值。Studholme^[33]比较了三种信息量度量：联合熵，MI 和归一化 MI，他们使用离散直方图来取代以前使用的 Parzen 窗口，最大值通过多分辨率的爬山算法获得。文献^[36]给出了基本 MI 配准和使用多分辨率的版本的比较。文献^[37]比较了 MI 和其他 6 种包括 NCC 和梯度相关的配准算法。

2.3.5 基于特征匹配的方法

2.3.5.1 特征提取

基于特征的方法往往提取图像中明显的区域块（森林、湖泊、田地）、线结构（区域边界、海岸线、公路、河流）和关键点（区域拐点、线交叉点、曲线上高曲率点）作为特征。这些特征要求足够显著，容易在各种畸变条件下稳定地被检测到。由于同直接利用像素灰度信息的互相关算法相比，特征提取包含了高层信号信息，所以该类算法对光照、噪音的抗干扰能力较强。

● 区域特征

区域特征可以是任意封闭区域投影^[38,39]，水坝湖泊^[40,41]，建筑^[42]，森林^[43]，城市^[44]或者阴影^[45]。区域特征一般是通过图像分割得到的^[46]，因而分割的精度对接下来的配准

步骤影响很大。Goshtasby^[38]提出了一种修正分割过程以提高配准精度的算法。图像分割的步骤和配准步骤循环执行；在每一次循环中，粗估计的物体对应关系被重新用来更新分割参数，最终能获得亚像素精度的配准。

● 线特征

线特征可以用线段^[42,47,48]，物体边界^[49,50,51]，海岸线^[52,53]，公路^[54]或者医学成像上延长的解剖结构^[55]。线对应关系通常用线端点或者中间点对应关系来表示。标准边缘检测算法比如 Canny 算子^[56]，或者基于 LOG(Laplacian of Gaussian)算子^[57]经常被用来提取线特征。不同边缘检测算子在医学多模态图像中应用的效果由 Maintz^[58,59]阐述。

● 点特征

点特征包括线交叉点^[60,61]、公路交叉点^[44,62]、区域中心点^[63]、局部曲率断点^[65,66]、小波变换的局部极值点^[69,70]等。其中很多算法都涉及到所谓“拐点”的检测，因为拐点往往有几何不变性且很容易被人们视觉上注意到。

Kitchen 和 Rosenfeld^[71]提出了关于图像二阶偏微分函数的关键点检测方法，先计算图像与垂直图像梯度方向平面的交线，再计算该交线上各点的曲率，由曲率的局部极值点作为关键点。类似的，Dreschler 和 Nagel^[72]搜索的是高斯曲率，Beaudet^[73]搜索图像的 Hessian 行列式值的局部极值。然而，基于图像函数二阶导的关键点检测对噪音是敏感的。Forstner^[74]提出了更加鲁棒且只与一阶导相关的检测算法，著名的 Harris^[75]算子被试验证明似乎更加有效并沿用至今。

检测出的关键点可能会很多，这样在配准时候就会增加计算量。不少人又提出如何选择一个子集又不影响最终配准效果的算法。Goshtasby^[40]只使用了原点集的凸集，然而这种凸集在原始图像和畸变图像中可能并不一致。Lavine^[78]使用了最小生成树技术来选择子集，Ehlers^[64]则用归并方法选择成簇的点集。

2.3.5.2 特征匹配

● 使用特征点空间关系的方法

假设原始图像和待配准图像的特征都通过各自提取的特征点集(Control Points)来表示。目标是寻找点集之间的对应关系，可以通过特征点对的空间关系或者领域的不变描述算子来获得。

Goshtasby^[40]描述了基于图匹配的算法。通过计算畸变图像中的特征在某参数下反变换后落入原始图像相应位置一定邻域范围内的“有效”特征数目来为该设定参数打分，得分最高的参数便是最优变换参数。Stockman^[60]提出聚类技术在配准上的应用。对每一对(CP)的变换参数视作搜索参数空间的一个点，在对图像中所有 CP 对做变换参数估计

后, 最优参数将在参数空间中形成聚类, 代表能够最大限度得正确映射最多的 CP 对, 而错误的参数估计往往在参数空间形成稀疏分布。聚类在检测出后, 由其重心代表最优变换参数。由于聚类的鲁棒性, 使得局部突变不会影响整体配准效果。Barrow^[79]引进了切面匹配的配准算法, 用广义距离函数定义图像中两两对应线特征的距离, 并最小化距离函数。

● 使用不变算子的方法

另外一种使用特征空间关系的方法是利用在变换下保持不变的描述算子。这些描述算子要求具有不变性(原始图像和变换后的图像中相应的特征描述要保持一致)、独特性(不同特征的描述要不同)、稳定性(在微小未知突变干扰下特征描述不能改变太大)、独立性(如果特征描述算子是矢量形式, 分量应当保持相互独立)。然而, 这四个特性一般很难同时兼顾, 实际运用中往往才用折中的办法。

通过不变描述算子, 原始图和变换图中最相似的特征被视作一组对应的特征。在描述算子空间中搜索最佳特征匹配时往往采用最小距离加阈值限定的方法。更加鲁棒的方法^[80]可以采用匹配概率系数。

最简单的特征描述算子是一个邻域内的图像灰度本身^[76,81]。Sester^[43]选择森林区域作为区域特征, 用最小包围盒、孔洞数、紧支性等方法来描述。曲线的链码表示^{[49][50]}也被用于不变描述算子, 并用链码相关算法来获得对应匹配关系。Skea^[82]用圆来表示非共线的三个 CP 点。Suk^[83]提出用多边形表示的区域特征, 并扩展^[84]了这一思想。Hu^[85]的不变矩组在描述区域特征上使用得非常频繁, 适合图像中的封闭区域。Flusser^[86]得到一组仿射不变量用于 SPOT 和 Landsat 图像配准, 最近又介绍了对图像模糊也保持不变的矩不变量^[89], 并用于 SPOT 和 AVHRR 卫星图像。他们还得出模糊、旋转均不变的算子^[90], 将其推导出的一般形式用于 3D 图像配准^[91]。

● 松弛算法

有大量的配准算法是基于解决 CLP (Consistent Labeling Problem) 的松弛算法: 用原始图像的特征标注畸变图像中特征, 循环执行直到找到一种稳定的标注方法^[92]。Wang^[47]扩展了经典的松弛算法, 用尖锐性、对比度和斜率特征来描述特征点, 允许平移和旋转, 但是计算量非常大。Medioni^[94]采用线特征和多种线段描述算子(坐标, 朝向, 平均对比度等)。Ton^[63]利用归并算法加速了经典算法的效率。在文献^[99]给出了不同松弛算法的比较。对 CLP 问题的另外一个解决方法是回溯法, 用递归的思想来产生标注, 可以用于配准^[92]。

● 弹性变形方法

当局部畸变影响较大或者发生复杂形变, 难以用简单几何变换参数估计时, 特征的不变描述算子很难直接获得, 特征间的对应关系也就难以建立。于是可以采用弹性可变模型的方法来配准。此时, 图像看作可在定义的内力和外力下收缩或者延展, 获得配准

是通过循环中最小化能量范函^[100,101]。为了提高速度通常都要采用多分辨率结构。力的定义经常与图像中信息相关，或者是灰度信息、或者是梯度信息。

2.4 配准算法研究趋势

尽管配准问题已经持续做了数十年，自动配准仍然是个没有彻底解决的问题。本节介绍了多种经典算法，当前的热点是瞄准更加复杂的非线性变换和局部畸变的配准问题、多模态配准和 N 维数据的配准。

当经过非线性变换和局部畸变的图像需要配准时，我们主要面对着两个基本的问题：如何匹配控制点和使用什么样的映射函数。即便我们能从理论上选择合适的映射函数，但是控制点匹配从本质上是沒有通用解决方案的，因为图像间的畸变几乎是任意的，且我们很难区分像素点上的灰度变化是几何变形还是内容的改变。

多模态配准中，互信息量已经成为一个标准算法。但是，作为非依赖图像特征的算法，互信息模型有很多局限。很多学者开始尝试结合互信息和图像特征的方法，且用多分辨率模型来加速算法。但是，当图像间的平移、尺度差别比较大的时候，这些算法都要么失败要么效率非常低。显然，基于特征的算法将会是重点研究方向。

N 维数据配准的主要困难在于它的计算量。尽管当今计算机硬件发展迅猛，但是算法效率的指标从来没有降低过。而且一般配准算法都会用循环递归的方法逐步求精，更加提高了对算法效率的要求。

2.5 本章小结

本章首先阐述了配准问题的数学定义和问题难点介绍。在统一的算法框架下，根据算法侧重点的不同，分别介绍了最常见的基于空间域相关的算法、基于 HD 距离的算法、基于 FFT 的算法、基于互信息量的算法与基于特征匹配的算法。正如前文所述，各种算法都有各自适应的应用背景，且算法鲁棒性很难直接相互比较。下文将在前人工作基础上，针对一般的复杂背景图像提出更加鲁棒的基于图像特征的算法，并给出试验设计和结果。

第三章 基于局部几何特征分布的曲线配准

3.1 算法描述

直接建立图像间的灰度映射关系在很多应用中很困难,但是图像的几何结构信息往往在畸变前后保持着稳定性,比如区域的边界在光照变化下基本保持不变。所以选择基于曲线的特征匹配方式比直接利用灰度信息更加鲁棒。另一方面,图像中曲线的提取也是一个非常困难的问题。通常借用的边缘检测算子在遇到纹理密集的图像往往会产生许多伪边。所以在复杂背景的图片配准问题上,不能简单得用边缘检测算子来提取区域边界,并且我们希望提取的边界越连续越好。本文将描述基于曲线特征的图像配准算法,并重点阐述曲线提取,曲线局部几何特征提取,匹配策略三个关键步骤。算法框架如图 3.1 所示。

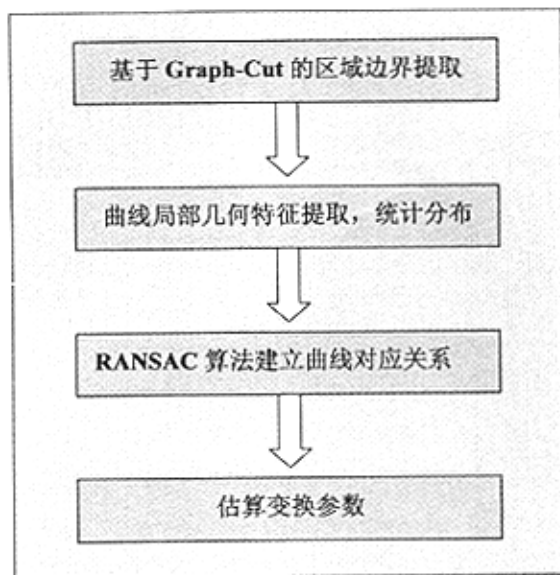


图 3.1 基于局部几何特征分布的曲线配准

3.2 图像曲线提取

3.2.1 Graph Cut 问题框架

Graph Cut 是近几年来在图像分割中兴起的一种新的思想,主要是利用图论中的概念

在整体和局部综合考虑下分割图像的一种框架^[102]。基本思想如下：

定义图 $G = \{V, E, W\}$ ，其中 V 代表顶点集合， E 代表边集。在图像上 V 就对应每个像素点， E 对应了像素点之间的连接， W 代表连接强度，它决定了像素是否在归类中属于同一类。给定图像的 N 个像素，分割问题就可以转化为将 V （点集）分解成 K 个无交集的子集，也就是 $V = \bigcup_{l=1}^K V_l$ ，且 $V_l \cap V_k = \emptyset$ ， $k \neq l$ ；符合条件的一种分割就记做：

$$\Gamma_V^K = \{V_1, \dots, V_K\}。$$

现在记 P, Q 为 V 的子集， P 和 Q 的连接关系 $\text{link}(P, Q)$ 为所有 P 中的点与 Q 中的点的连接强度之和：

$$\text{links}(P, Q) = \sum_{p \in P, q \in Q} W(p, q) \quad (3.1)$$

再定义一个子集 P 的深度为 P 到所有顶点的连接关系：

$$\text{deg}(P) = \text{links}(P, V) \quad (3.2)$$

因此子集 P, Q 的相对连接关系为：

$$\text{linkratio}(P, Q) = \frac{\text{links}(P, Q)}{\text{deg}(P)} \quad (3.3)$$

特别的， $\text{linkratio}(P, P)$ 表示 P 内部的连接强度， $\text{linkratio}(P, V/P)$ 表示 P 与外部的连接强度。一个好的分割，应当是分割出来的子集内部连接越紧密越好，子集之间的连接越松散越好，这是一个非常直观的想法。由此可以定义出分割好坏的一个定量标准：

$$\text{knassoc}(\Gamma_V^K) = \frac{1}{K} \sum_{l=1}^K \text{linkratio}(V_l, V_l) \quad (3.4)$$

上式表示 K 个分类的内部连接平均值，应该值越大表示分割越合理。

$$\text{kncut}(\Gamma_V^K) = \frac{1}{K} \sum_{l=1}^K \text{linkratio}(V_l, V \setminus V_l) \quad (3.5)$$

上式表示 K 个分类的各自与外部连接的平均值，应该值越小表示分割越合理。

显然有 $\text{kncut}(\Gamma_V^K) + \text{knassoc}(\Gamma_V^K) = 1$ ，所以我们只需要优化其中一个目标函数即可。

具体联系到图像上，首先我们定义深度矩阵：

$$D = \text{Diag}(W1_N) \quad (3.6)$$

其中 Diag 表示根据参数矢量构造对角矩阵， 1_N 表示 N 行都是 1 的列向量。

我们用 $N \times K$ 的矩阵 X 代表分割 Γ_V^K ， $X = [X_1, X_2, \dots, X_K]$ ， X_l 是 V_l 的一个二值标注向量，所以：

$$X(i, l) = \text{istrue}(i \in V_l), i \in V, l \in [K] \quad (3.7)$$

由于每个像素被分配且只分配到 K 类中的一类，所以又可以加上约束条件：

$$X : X1_K = 1_N \quad (3.8)$$

重新改写 link 和 deg 函数，我们得到：

$$links(V_i, V_l) = X_l^T W X_l \quad (3.9)$$

$$deg(V_l) = X_l^T D X_l \quad (3.10)$$

综合以上定义，得出 Graph Cut 的一个图像分割框架即，关于 X 变量的函数优化问题：

$$\text{Maximize } \varepsilon(X) = \frac{1}{K} \sum_{l=1}^K \frac{X_l^T W X_l}{X_l^T D X_l} \quad (3.11)$$

$$\text{约束条件为：} X \in \{0,1\}^{N \times K}, X1_K = 1_N \quad (3.12)$$

求解这个分配问题是 NP 难问题，可以转化为求一个广义特征值求解问题，Stella 给出了近似解的求法^[103]。

3.2.2 连接强度矩阵 W 的建立

在目标函数定义下来后，Graph Cut 中最重要的就是决定如何选择构造矩阵 W 。合理的构造 W 应当能反应图像中各点的正确分类，如图 3.2 所示。

Shi 给出了如下 W 计算方法^[102]，

$$w_{ij} = e^{\frac{-\|F(i)-F(j)\|_2^2}{\sigma_f}} * \begin{cases} e^{\frac{-\|X(i)-X(j)\|_2^2}{\sigma_x}}, & \text{if } \|X(i)-X(j)\|_2 < r \\ 0, & \text{otherwise} \end{cases} \quad (3.13)$$

上式 $F(\cdot)$ 代表图像点灰度， $X(\cdot)$ 代表图像点位置；表示如果两图在邻域内足够接近，那么他们间的连接强度取决于两者的灰度接近程度。注意到由于邻域条件的约束， W 矩阵是个稀疏矩阵，大部分项都是 0；Stella^[104]还考虑了多尺度下的连接关系，将连接矩阵修改为多个尺度下各个连接矩阵的加权和。

当然， W 矩阵的最优设计应当在正确分割之后才计算，但是 W 本身又是为了求解分割问题的。所以这里似乎存在着鸡生蛋还是蛋生鸡的问题。一般来说做图像分割的学者都会以循环的方式：生成 W 矩阵求解分割问题—>根据分割结果更新 W 矩阵。如果 W 矩阵的计算鲁棒性较好，一般不需要迭代很多次（考虑到求解时的计算量，多次迭代也是难以忍受的）。

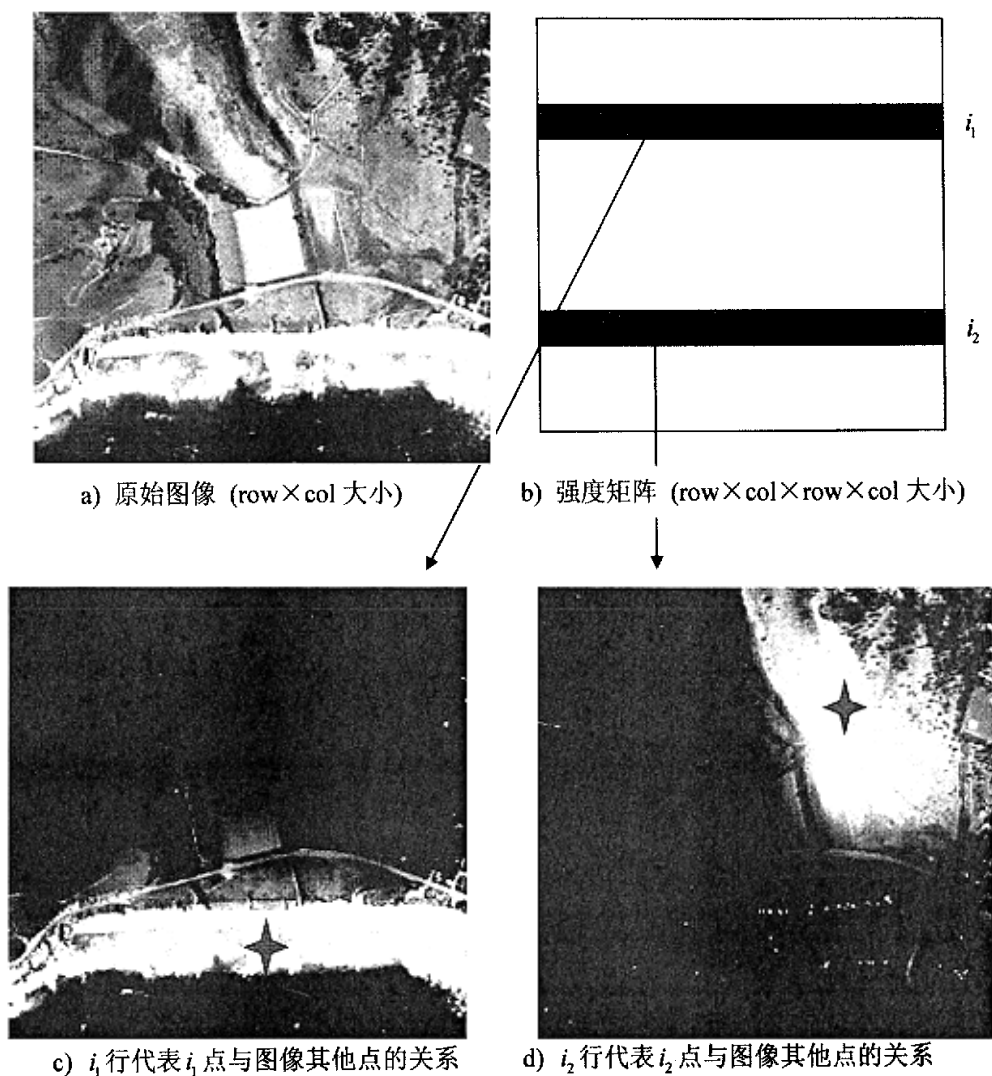


图 3.2 强度矩阵含义示例

由于图像的复杂性,仅仅利用灰度信息并不能得到满意的分割,可以综合颜色(灰度)、纹理、边缘的各种信息。本文采用的 W 计算如下:

$$w_{ij} = \begin{cases} e^{-\frac{\|I(i)-I(j)\|_2^2}{\sigma_I}} * e^{-\frac{\|T(i)-T(j)\|_2^2}{\sigma_T}} * e^{-\frac{\|X(i)-X(j)\|_2^2}{\sigma_X}}, & \text{if } \|X(i)-X(j)\|_2 < r \text{ and } \max OE_\theta < t \\ e^{-\frac{\|hist_I(i)-hist_I(j)\|_2^2}{\sigma_I}} * e^{-\frac{\|hist_T(i)-hist_T(j)\|_2^2}{\sigma_T}} * e^{-\frac{\|X(i)-X(j)\|_2^2}{\sigma_X}}, & \text{if } \|X(i)-X(j)\|_2 < r \text{ and } \max OE_\theta > t \\ 0, & \text{if } \|X(i)-X(j)\|_2 > r \end{cases} \quad (3.14)$$

$$\text{其中 } OE_{\theta,s} = (I * f_{\theta,s}^e)^2 + (I * f_{\theta,s}^o)^2 \quad (3.15)$$

$OE_{\theta,s}$ 表示一种针对复合型边缘（包括阶越值、峰值型和屋顶型）的能量， $f_{\theta,s}^e$ 和 $f_{\theta,s}^o$ 是一组偶对称和奇对称的滤波器^[105]。

当响应的能量小于一定阈值 t 时，表示在参考的区域内没有复杂边缘存在，此时 W_y 按照(3.14)式中第一个情况计算。此时在任意图像一点上，在一个小半径的圆域内计算该点统计信息， I, T, X 分别代表平均灰度、平均纹理能量，和该点位置；显然，灰度越接近，纹理越相似，位置又足够近的点对连接关系应当是强的。

当响应的能量超过阈值时，表示在参考的区域内有复杂边缘存在，此时 W_y 按照(3.14)式中第二个情况计算。此时在任意图像一点上，同样以一个小半径为圆域内计算。此时用 $hist_I, hist_T$ 分别代替 I, T ，即用直方图代替平均值。因为在出现复杂边缘的不平滑区域时，直方图统计特性比平均值更为准确。

现在还需要定义纹理计算方式和直方图求差公式如下：

1. 纹理特征的提取有多种选择。本文仍借用式(3.15)的形式，以每一点在一组滤波器下的响应能量为特征矢量的各个分量。在实际运用中为了减少计算量，可以利用 K 平均聚类的方法事先训练出一批典型纹理分类及其特征矢量，构成标准纹理库，则新计算得到的特征矢量通过 L2 距离映射到与之最接近的纹理类型上。(3.14)式中的纹理特征最终用纹理类型标号来表示。基于 K 平均聚类的纹理特征训练方法详见附录 A。

2. 直方图求差公式采用如下常见的归一化距离函数：

$$\chi^2(g, h) = \frac{1}{2} \sum \frac{(g_i - h_i)^2}{g_i + h_i} \quad (3.16)$$

其中 g, h 为直方图， g_i, h_i 为他们对应的 Bin。

3.2.3 算法设计和实验结果

本文中完整的 Graph Cut 算法步骤如下：

1. 按照式子(3.15)计算每一点的滤波响应。以矢量形式保存计算结果，作为该点纹理特征。记录分量最大值和对应的角度参数。
2. 根据纹理类型图库，用 L2 距离标注每一点从属的纹理分类。

3. 计算限定邻域内任意两点的灰度距离、纹理距离、欧式距离(或者是直方图距离)。
4. 按照(3.14)计算 W 强度矩阵。
5. 按照(3.6)计算深度矩阵 D 。
6. 构造广义特征根问题：

$$WV = \lambda DV \quad (3.17)$$

求解其对应最大 K 个特征值的特征向量，即求得一个 K 类的分割。求解此广义特征根算法一般输出的是连续解，求得近似的离散解可按照 Stella 的转轴逐步逼近的方法^[103]。

下图给出基于 Graph Cut 的区域边界提取效果和 Matlab 版本 Canny 算子提取边缘的对比，由于 Graph Cut 基本上是出于图像分割的角度来提取区域边界的，所以更能从整体效果上保证对伪边的压制。实际输出时，以分割区域边界（除去图像边界）记做一条完整的曲线，所以最终输出的所有曲线都不会有分叉，便于进一步处理。（注意：此时边界利用是冗余的，因为不同区域会共用边界）

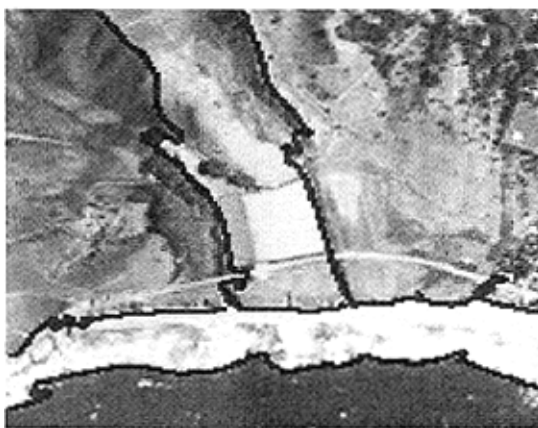
由 Graph Cut 输出的边界将作为配准算法的输入。当然区域边界可能有数条，如何寻找两副图像中曲线的对应关系还需要一定的匹配算法来保证。另外，虽然我们的目的是寻找区域的边界，但是如何划分区域、划分到什么程度并没有严格的定义。幸运的是，我们的目的是寻找待配准图像中的对应曲线，并不是图像分割问题，所以即使遗漏边界或者输出过多的边界曲线，只要两副图像的输出有可比性，对配准算法仍然是有效的。



a) 原始航拍图像



b) Graph Cut 输出的区域边界



c) Graph Cut 的分割结果



d) 5 个最大的特征向量显示

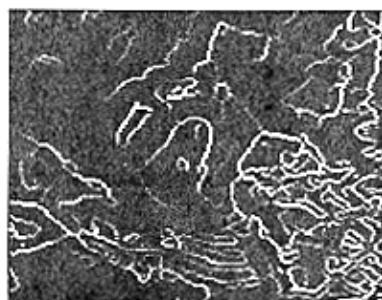


e) Canny 算子输出的边缘

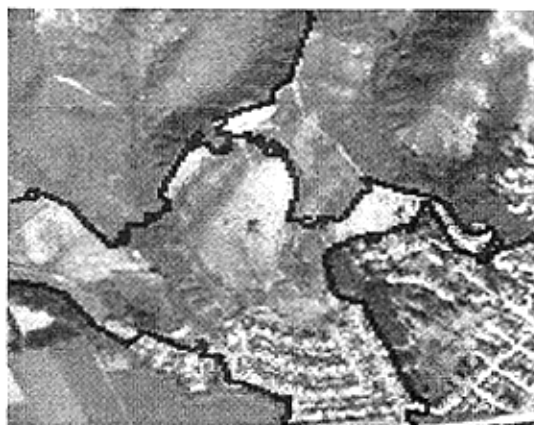
图 3.3 Graph-Cut vs Canny, Case 1



a) 原始航拍图像



b) Graph Cut 输出的区域边界



c) Graph Cut 的分割结果



d) 5 个最大的特征向量显示



e) Canny 算子输出的边缘

图 3.4 Graph-Cut vs Canny, Case 2

3.3 基于曲线局部几何特征的变换参数估计

本节提出的配准算法,并不直接建立精确的曲线上的点与另外一条曲线上的点的对应关系,而用曲线的某种几何描述的分布来相互比较。当然,这样的配准算法就产生了第一个前提:有稳定的曲线存在,且被正确检测到。在此前提下,我们给出几种全局几何变换模型下应用本节配准算法的方法。

设原始图像曲线上的点为 P , 对应畸变图像曲线上的点为 \hat{P} , 变换模型统一为:

$$\hat{P} = TP$$

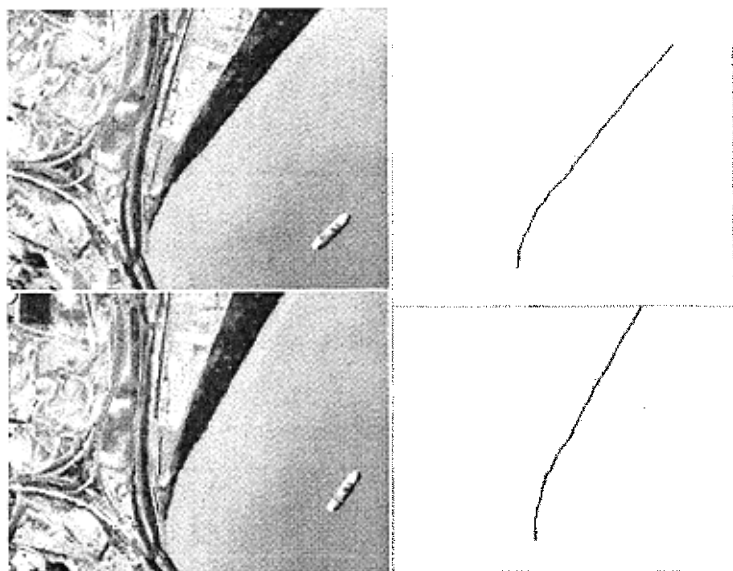
设 $F(P)$ 、 $F'(\hat{P})$ 分别为该点的几何描述算子,我们的目标是设计合适的描述算子,以极小化如下范函:

$$\arg \min_T \| \Pr(F) - \Pr(F') \| \quad (3.19)$$

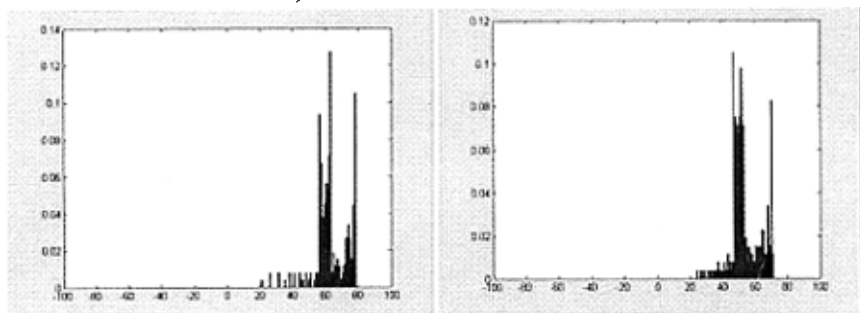
这里 $\Pr(\cdot)$ 是概率密度函数,表示此描述算子的分布。

比如下图所示,两副图像的特征曲线存在旋转关系。对其中一条曲线中的任意一点,记该点斜率为 ψ , 且在另外一条曲线上的对应点的斜率为 $\hat{\psi}$, 显然有 $\hat{\psi} = \psi + \theta$ 。既然以上所述关系符合曲线上每一个点,那么也应当符合斜率 \Pr 的分布关系。在计算出曲线上

点的斜率分布后，利用互相关就可以得到对应的旋转角度 θ 。



a) 曲线特征提取结果



b) 曲线上点斜率直方图（斜率分布）

图 3.5 曲线上点斜率分布示意图

3.3.1 平移和旋转

$$\hat{P} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} P + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (3.18)$$

有三个变换参数待求： θ, t_x, t_y

为求角度 θ ，选择 F 形式为曲线上点的斜率 $slop$ ，统计出斜率分布，得：

$$\Pr(slop') - \Pr(slop) = \theta$$

通过对斜率分布求互相关便可得出两条曲线间的角度差；

在做出角度补偿后，选择 F 形式为曲线上点坐标 Xor 、 Yor ，同样统计坐标分布，得：

$$\Pr(Xor') - \Pr(Xor) = t_x$$

$$\Pr(Yor') - \Pr(Yor) = t_y$$

3.3.2 尺度缩放

$$\hat{P} = sP \quad (3.19)$$

有一个变换参数待求： s

由于曲线上任意一点的曲率半径 R 正比与尺度变换参数，即 $R(\hat{P}) = s * R(P)$ ，所以选择 F 形式为曲率半径的对数 $\ln |R(\cdot)|$ ，得：

$$\Pr(\ln |R(\hat{P})|) = \Pr(\ln |R(P)|) + \ln s$$

3.3.3 拟仿射变换

$$\hat{P} = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} P + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (3.20)$$

有 5 个参数待求： s_x 、 s_y 、 θ 、 t_x 、 t_y

由于此时曲率和斜率不再相互独立，所以不能直接套用前两种变换中介绍的方法，需要先重写变换模型形式：

$$\hat{P} = \sqrt{|s_x s_y|} \begin{pmatrix} \rho & 0 \\ 0 & \frac{1}{\rho} \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} P + \begin{pmatrix} t_x \\ t_y \end{pmatrix}, \quad \text{其中 } \rho^2 = \frac{s_x}{s_y} \quad (3.21)$$

令曲线 C, \hat{C} 分别由参数 t, \hat{t} 来表示，定义 $M = [P'(t) \ P''(t)]$ ， $\hat{M} = [\hat{P}'(\hat{t}) \ \hat{P}''(\hat{t})]$ 为关于参数的导数形式，则：

$$M = [P'(t) \ P''(t)] = \begin{bmatrix} x' & x'' \\ y' & y'' \end{bmatrix}, \quad \hat{M} = [\hat{P}'(\hat{t}) \ \hat{P}''(\hat{t})] = \begin{bmatrix} \hat{x}' & \hat{x}'' \\ \hat{y}' & \hat{y}'' \end{bmatrix}$$

可推出 $\hat{M} = TM(\frac{ds}{d\hat{s}})$ (详细参见^[106])

如果取 $(\frac{ds}{d\hat{s}}) = 1$ ，便可以得到：

$$\ln |\hat{M}| = \ln |s_x s_y| + \ln |M|$$

因此选择 F 形式为 $\ln |[P'(t) P''(t)]|$ ，得：

$$\Pr(\ln |\hat{M}|) = \ln |s_x s_y| + \Pr(\ln |M|)$$

以上式分布关系求得尺度参数，并修正曲线的尺度变化。

再根据关系式 $\hat{x}' \hat{y}' = \frac{\sin(2\theta)}{2} (y'^2 - x'^2) + \cos(2\theta) x' y'$ 求解角度 θ

同样由关系式 $|\frac{\hat{x}'}{\hat{y}'}| = |\rho^2| \frac{|\cos \theta \cdot x' + \sin \theta \cdot y'|}{|-\sin \theta \cdot x' + \cos \theta \cdot y'|}$ ，选择 F 、 F' 满足：

$$\Pr(\ln |\frac{\hat{x}'}{\hat{y}'}|) = 2 \ln |\rho| + \Pr(\ln |\frac{\cos \theta \cdot x' + \sin \theta \cdot y'}{-\sin \theta \cdot x' + \cos \theta \cdot y'}|)$$

以上式求得参数 ρ ，剩下的平移参数可按照前文所述求得。

3.3.4 仿射变换

$$\hat{P} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} P + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (3.22)$$

有 6 个参数待求： a, b, c, d, t_x, t_y

首先仍然求解 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 的行列式值 M ，显然我们有如下关系：

$$\frac{\hat{x}''}{\hat{x}'} = \frac{ax'' + by''}{ax' + by'}$$

做一个简单的参数变换 $(a, b) = \sqrt{a^2 + b^2} (\sin \phi, \cos \phi)$ ，上式转化为：

$$\frac{\hat{x}''}{\hat{x}'} = \frac{\sin(\phi)x'' + \cos(\phi)y''}{\sin(\phi)x' + \cos(\phi)y'}$$

由上式可以求解出 $\frac{a}{b} (= \tan \theta)$ ；用同样的方法可以求出 $\frac{c}{d}$ 。

同样还有关系式： $\frac{\hat{x}'}{\hat{y}'} = \left(\frac{b}{d}\right) \frac{\frac{a}{b}x' + y'}{\frac{c}{d}x' + y'}$ ，求出 $\frac{b}{d}$

至此，我们求得 $\alpha = \frac{a}{b}, \beta = \frac{b}{d}, \gamma = \frac{c}{d}$ ，带入行列式 M ，得到 $M = (ad - bc) = a^2 \frac{\alpha - \gamma}{\alpha^2 \beta}$

由上式可以求解出 a ，继而得出 b, c, d 。

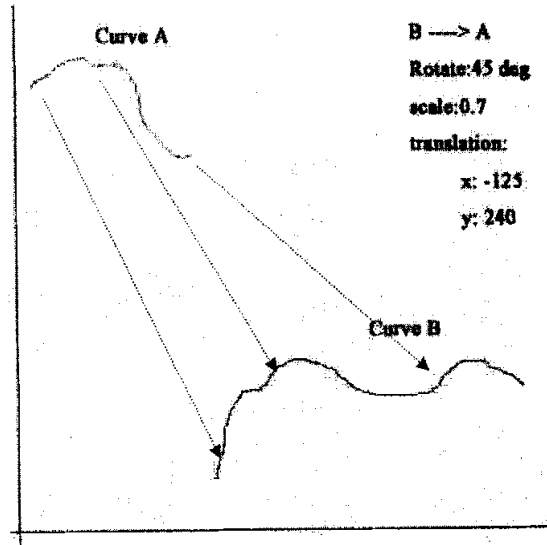
3.3.5 算法设计和实验结果

图像曲线提取出的是粗糙的离散点，首先要拟合出更加光滑的曲线才可以计算曲线上各点导数。在文献^[109]中给出了较为准确的导数计算方法，可以抵抗噪音点的干扰。本文采用三次样条函数拟合，保持曲线的光滑性且方便计算各点导数。另外，由于我们将考察的是局部特征的分布，通过建立直方图来近似得代替分布函数，所以 Bin 的宽度必须足够细以保证所有采样点能落入尽可能多的 bin 上。对于那些长度太小的曲线，由于采样点太少很容易引起误差，所有这样的曲线都先不予考虑参与计算。最后，由于即使匹配的曲线之间本身可能有差异，所以多次迭代的效果会更好。

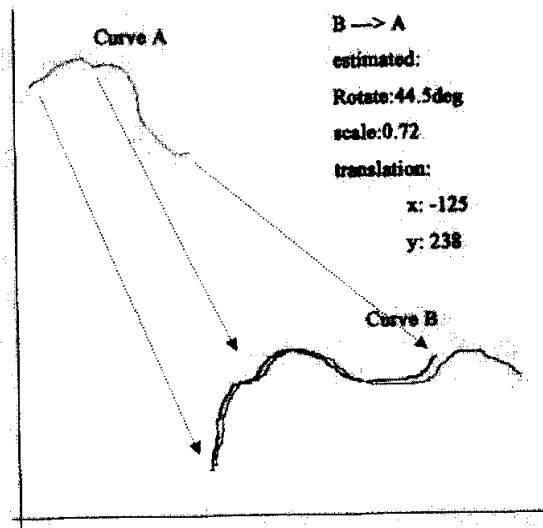
下表给出了如图 3.5 的曲线对根据曲线局部几何特征分布估计变换参数的试验结果：

表 3.1 曲线配准试验结果

拟仿射变换参数估计	旋转角度	尺度	平移 X	平移 Y
Case1 真实值	0	1.0	50	50
Case1 估计值	0	1.0	50	50
Case2 真实值	5	1.0	50	50
Case2 估计值	5	1.0	50	50
Case3 真实值	10	1.0	50	50
Case3 估计值	9.0	1.0	50	50
Case4 真实值	45	1.0	50	50
Case4 估计值	44.5	1.0	48	51
Case5 真实值	0	0.8	100	100
Case5 估计值	0	0.8	100	102
Case6 真实值	45	0.8	100	100
Case6 估计值	43	0.78	95	103
Case7 真实值	45	1.25	100	100
Case7 估计值	44.5	1.20	95	93
Case8 真实值	45	0.7	-125	240
Case8 估计值	44.5	0.72	-125	238



a) 待配准曲线 A、B



b) 曲线 A、B 配准结果

图 3.5 基于曲线局部几何特征的配准示例

3.4 RANSAC 算法修正曲线对应关系

本文采用的基于局部几何特征分布的曲线配准优于一般点特征配准的最大贡献在于，它直接略过了建立点对应关系的困难步骤。然而，曲线与曲线的对应关系还是需要的。显然，只有曲线 A 与它的变形 A' 之间相互匹配才是合理的；又比如，如果一幅图像提取出 3 条完整曲线，另外一幅图像提取出 5 条曲线，那么配对关系又当如何确立，这同样需要进一步设计算法解决。

RANSAC(Random Sample Consensus)^[108]算法是一种在强噪音干扰下仍能鲁棒得将模型与数据匹配的有效算法。它的基本思路如下：

给出一个匹配问题和待估计的参数 \mathcal{E} ，假设：

1. 参数可以通过 N 项数据估计，总共有 M 个数据。
2. 随机选取的符合模型的数据的概率为 P_g 。
3. 算法找不到任何符合模型的数据的概率 P_{fail} 。

则算法执行步骤：

1. 随机选取 N 个数据
2. 估计参数 \mathcal{E}
3. 找出在此参数下 M 个数据中多少数据在允许的误差范围内符合模型，记为 K
4. 如果 K 足够大，则算法退出
5. 重复步骤 1—4 共 L 次
6. 如果执行至此，算法失败

其中 K 取值通常取决与用户认为数据中大致合理的数据所占的比例。

L 取值方法如下：

$$\begin{aligned}
 P_{fail} &= L \text{ 次连续失败的概率} \\
 &= (\text{一次测试失败的概率})^L \\
 &= (1 - \text{一次测试成功的概率})^L \\
 &= (1 - (\text{随机取的数据符合模型的概率})^N)^L \\
 &= (1 - (p_g)^N)^L
 \end{aligned}$$

$$\text{所以 } L = \frac{\log(p_{fail})}{\log(1 - (p_g)^N)} \quad (3.33)$$

RANSAC 本身的效率比较低，但幸运的是前文所述的 Graph Cut 算法保证了不会产生大量曲线，所以采用基于 RANSAC 的随机匹配—验证的策略代价并不高。

设输入从图像 A 和图像 B 提取的曲线集分别为 $\Lambda_A : \{C_i\}$, $\Lambda_B : \{C_j\}$, 曲线 RANSAC 匹配算法伪代码如下:

```

sort( $\Lambda_A$ ), sort( $\Lambda_B$ ); /* based on curve length or area enclosed */

for i = 1 : | $\Lambda_A$ |

    caculate_geometry_descriptor( $C_i$ ); /* based on formula introduced in section 3.3 */

    hist(A, i) = create_descriptor_histogram( $C_i$ , descriptor_type);

end

/* same for  $\Lambda_B$  */

for i = 1 : | $\Lambda_A$ |

    for j = 1 : | $\Lambda_B$ |

        transform_params(i, j) = estimator(hist(A, i), hist(B, j)); /* correlation the distributions */
        image(C) = transform(image(A), transform_param(i, j)); /* transform the original image */
        error = similarity(image(B), image(C)); /* here similarity function can be various type */
        if error < threshold
            output(transform_params(i, j)); exit();
        end
    end
end
end

```

在验证当前估计参数正确性时,验证方法主要是通过比较其他曲线在此参数表示的变换下与目标曲线集重合的程度。曲线与曲线的重合程度可以用第二章介绍的 HD 距离来度量。如果已知图像灰度变化不剧烈的情况下,也可以按照伪代码所示的将原始图像变形,再与目标图像做相似性度量。相似性度量函数也可以按照第二章选择多种方法。

当输入的曲线集合元素数量较多时,就需要考虑提前结束算法,此时就要利用到公式(3.33)实现计算 L 值。当出现连续 L 次失败时便主动退出程序。

由于在一般仿射变换下不会更改同一图像中曲线间的相对关系。所以还可以进一步设计一些启发式规则指导每次抽取样本的策略而不是 RANSAC 的随机抽选。其可靠性和实用性仍然有待研究。

3.5 本章小结

本章介绍了基于曲线局部几何特征分布的配准方法,避开了直接在可能出现大畸变情况下寻找点对对应关系的困难。首先引入基于图分割的区域边界检测模型,试验证明由于图分割综合考虑了图像局部和全局的信息,在遇到复杂背景图像时效果远好于一般边缘检测算子。在合理提取特征曲线后,利用寻求曲线上几何特征分布之间的相关性来求解变换模型参数。由于特征曲线可能不止一条,便利用 RANSAC 方法寻找曲线与曲线之间的对应关系,并同时求解最优变换参数。因为特征曲线数量不多,且曲线在畸变下的稳定性好,所以此算法比普通的基于点匹配的算法更鲁棒、更有效。

然而,由于曲线为离散表示,几何特征计算时用到多处近似计算(导数等),所以最终求解得到的变换参数只是粗略值。若需要精确定位配准参数,通常高级特征的效率并不如低级特征,也就是说在配准问题上基于线特征的算法精确度不如基于点特征的算法。所以下一章将介绍在粗配准之后进一步求精的点特征配准方法。

第四章 基于特征点匹配的配准模型

4.1 算法描述

当利用上一章节的算法初步修正图像后（变换原始图像或者反变换畸变图像），此时图像间只存在微小变换，如果需要进一步修正求精，此时可以利用点匹配的配准方法。本文采用的特征点匹配算法流程如图所示。

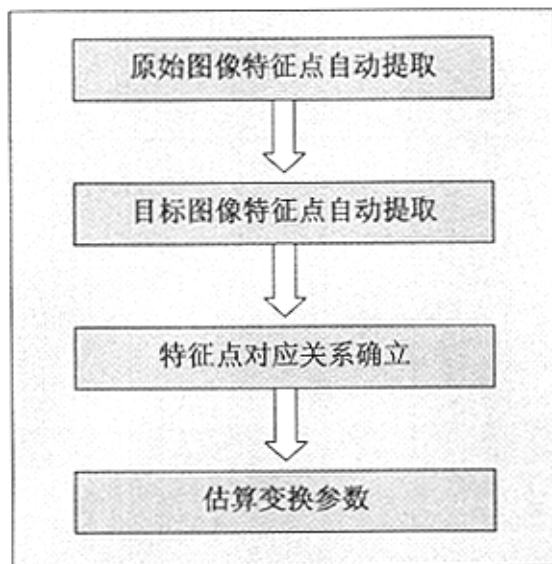


图 4.1 基于特征点匹配的配准算法

4.2 原始图像特征点自动提取

好的特征点提取算法应当满足以下条件：

1、特征点的重复度高

在不同条件下拍摄的两幅图像中抽取的特征点具有高的一致性与稳定性；

2、特征点所包含的信息量大

也就是说被提取的点很突出，能够有效地反映图像的主要结构信息。

为了有效地配准两幅图像，特征点的抽取算法还应该具有旋转、平移不变性，并且在发生小的尺度变化和透视形变时，能够检测出相同位置特征点。

本文选用基于 Gabor 滤波器的特征点提取方法。

Gabor 滤波器为带权值的有方向性的正弦波与高斯窗口内的 Dirac 函数卷积的结果，

它定义如下：

$$G(x, y, \theta, f) = e^{i(f_x x + f_y y)} e^{-(f_x^2 + f_y^2)(x^2 + y^2)/2\sigma^2} \quad (4.1)$$

其中, x, y 为空间坐标, θ 为滤波器方向角, f 为中心频率, f_x 为 $f \cos \theta$, f_y 为 $f \sin \theta$, σ 为高斯方差。图 4.2 显示了 4 个方向的 Gabor 滤波器波形, 图 4.3 显示了频域中的一组 Gabor 滤波器, 从中可以清晰得看出 Gabor 滤波器的多尺度和多方向特性。

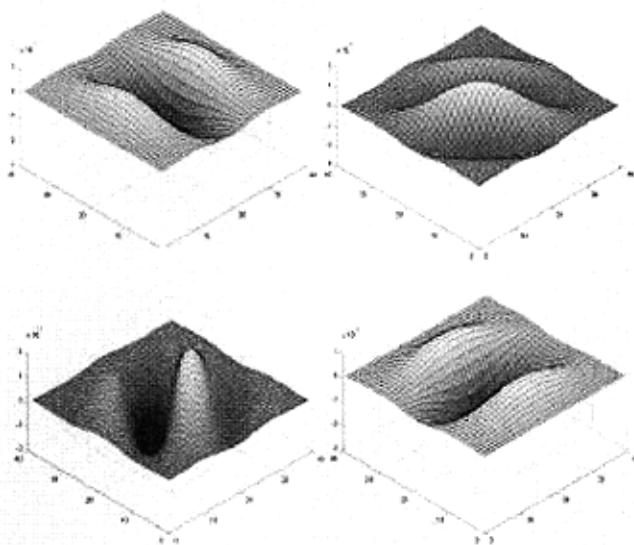


图 4.2 一组 4 个方向的 Gabor 滤波器

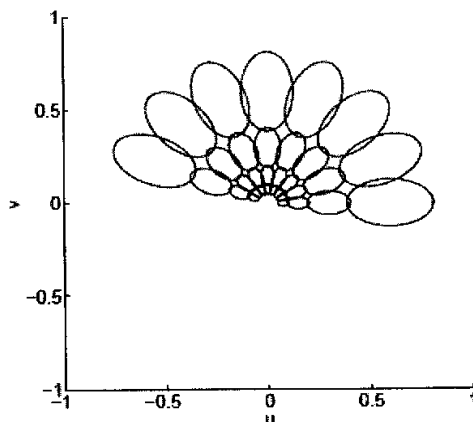


图 4.3 频域里的一组 Gabor 滤波器

由图像与 Gabor 滤波器卷积后, 设计如下响应能量度量函数:

$$F(x^*, y^*) = \max_{x, y \in N_{xy}} F(x, y) \quad (4.2)$$

$$F(x, y) = \max_{\theta} F(x, y, \theta) \quad (4.3)$$

$$F(x, y, \theta) = \|E_s(x, y, \theta) - \alpha E_r(x, y, \theta)\| \quad (4.4)$$

$$E_s(x, y, \theta) = I * G(x, y, \theta, f_s) \quad (4.5)$$

由上可知，在一个邻域内存在某个方向上能量响应的极大值的点就可以作为粗选的特征点。为了使得选择出来的特征点分布不要过于集中，尽量分布在图像的主要结构位置上，在选择特征点时可以在粗选之后用自适应的参数修正参考邻域大小，使得高密度的特征点区域降低特征点数目。图 4.4 给出试验结果。

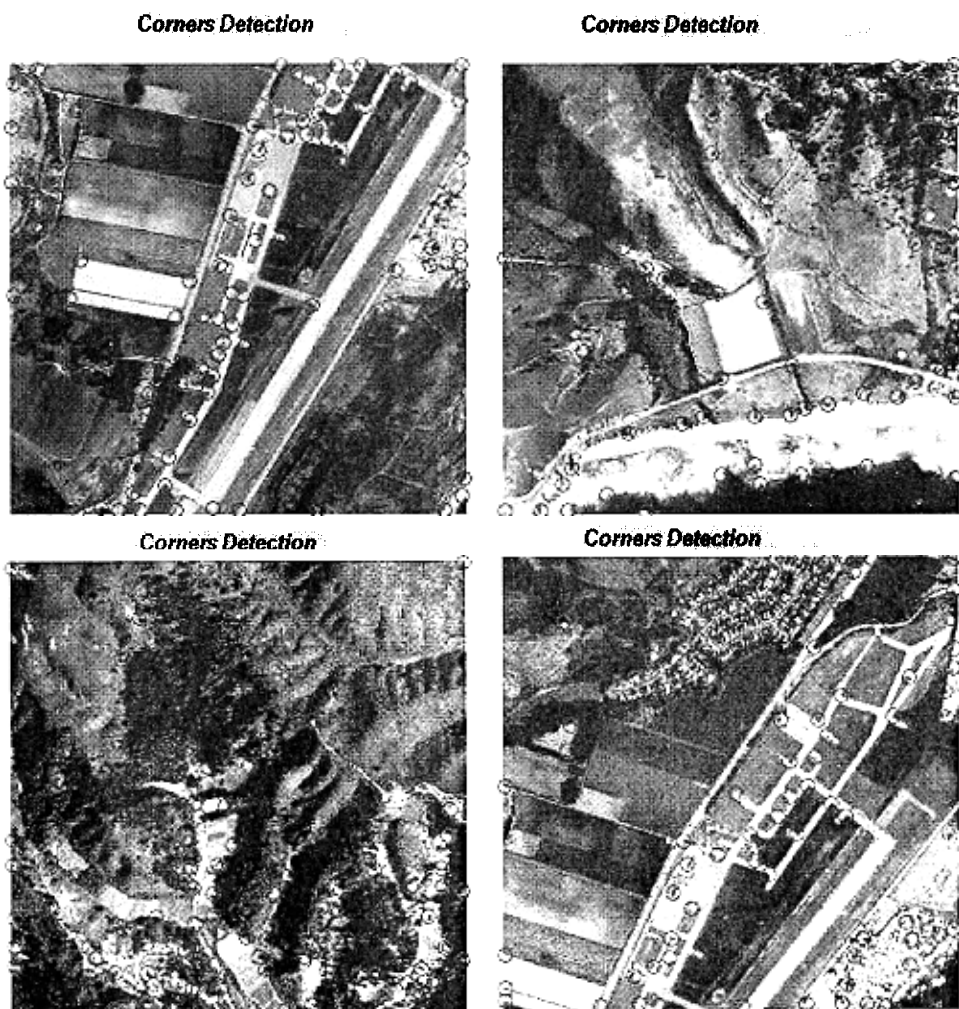


图 4.4 特征点自动提取示例

4.3 目标图像特征点自动提取

设原始图提取的特征点集为 P , $P = \{p_i(x_i, y_i)\}$, 对每个特征点 p_i 选取周围 9×9 的小块作为特征块, 在对应目标图位置的一个窗口领域 (30×30) 内做 NCC 计算:

$$\rho = \frac{\left| \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (A_{ij} - \bar{A})(B_{ij} - \bar{B}) \right|}{\sqrt{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (A_{ij} - \bar{A})^2 \cdot \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (B_{ij} - \bar{B})^2}}, \quad 0 \leq \rho \leq 1 \quad (4.6)$$

其中 \bar{A} 、 \bar{B} 分别是 A, B 两个子块内的灰度均值。

求得相关系数最大的位置为初选的对特征点 q_i , (p_i, q_i) 为一组对应匹配点对。

由于已经经过第三章介绍的粗配准, 此时 NCC 计算已经是在较小几何畸变的情况下, 旋转和尺度的干扰比较小, 所以这种寻找对应点的方式是合理的。但是由于特征点数相对较多, 必须采用快速 NCC 计算方法(详见第 5 章介绍)。

4.4 错误匹配点的消除

由于图像变换前后可能存在内容上的差异(如遮挡物和景观变化等), 相关计算不能保证在最大相关峰值的点为对应的匹配点, 因此必须去除一些明显不符合变换准则的匹配点对。准则为两个方面:

- 1) 归一化相关系数未超过阈值(如 0.7);
- 2) 假设 $MP = \{p_i \leftrightarrow q_i, i=1, 2, \dots, N\}$ 为初选的匹配点对集合, 且假设匹配点对之间的关系可以表示成仿射变换: $q_i = s \cdot R \cdot p_i + T$ 。当尺度和旋转参数较小时 $q_i \cong p_i + T$, 因而大多数匹配点对之间的平移量基本接近。对所有匹配点对之间的平移量差 T_x , T_y 做统计, 去除在 3σ 区间外的点对。

4.5 变换参数估计

设最终选定的匹配点对集为 S ，并假设匹配点存在仿射变换关系，问题转化为基于最小平方误差的优化：

设 $e = \sum \|sRp_i + T - q_i\|^2$ ，最小化该函数，分别令 $\partial e / \partial T = 0$ ， $\partial e / \partial s = 0$

$$\text{得到： } T = \bar{q} - sR\bar{p}, \text{ 其中 } \bar{p} = \frac{1}{N} \sum_{i=1}^N p_i, \quad \bar{q} = \frac{1}{N} \sum_{i=1}^N q_i, \quad (4.7)$$

$$s = \frac{\sum_{i=1}^N q_i'^T R p_i'}{\sum_{i=1}^N p_i'^T p_i'}, \text{ 其中 } p_i' = p_i - \bar{p}, q_i' = q_i - \bar{q} \quad (4.8)$$

$$\text{于是 } e = \sum_{i=1}^N q_i'^T q_i' - \frac{[\sum_{i=1}^N q_i'^T R p_i']^2}{\sum_{i=1}^N p_i'^T p_i'} \quad (4.9)$$

关于旋转参数 R 最小化(4.9)式，可以按照以下步骤求得（详细证明文献^[110]）：

$$1) \quad K = \sum_{i=1}^N q_i' p_i'^T \quad (4.10)$$

$$2) \quad \text{将矩阵 } K \text{ 做奇异值分解 (SVD), 得到 } K = U \Lambda V^T \quad (4.11)$$

$$3) \quad R = U S V^T, \quad S = \begin{pmatrix} 1 \\ \det(UV^T) \end{pmatrix} \quad (4.12)$$

4.6 实验结果

本实验针对较小几何畸变的图像配准，用本章节的特征点匹配算法估计变换参数。

图 4.5 中的两幅图像不仅存在几何畸变，是不同时不同设备采集的，所以在图像内容上有差异。图 4.6 是本文算法给出的配准结果。表 4.1 是一组测试结果，可以看出在小变换存在的情况下，使用特征点的配准方法可以克服传统的相关算法对旋转、光照变化的影响。

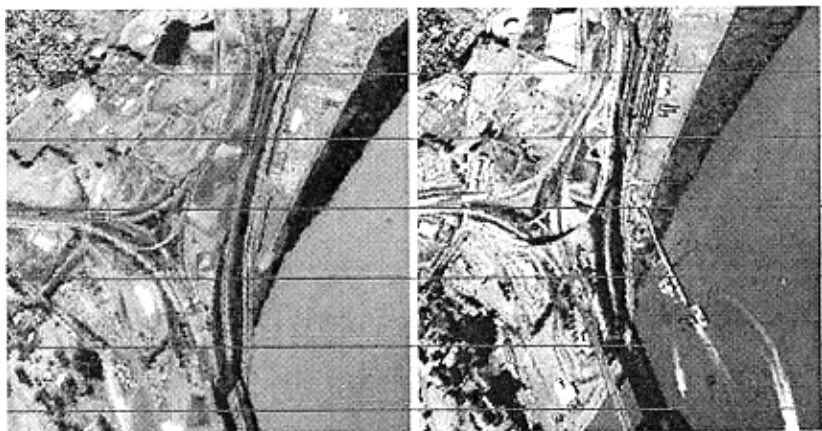


图 4.5 配准前两幅图像间的比较。
存在几何畸变和内容上的改变，清晰度也不同。

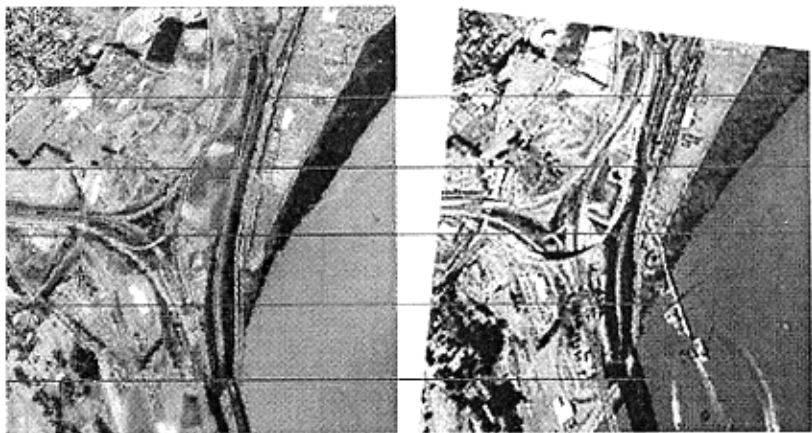


图 5.5 配准后两幅图像间的比较

表 5.1 曲线配准试验结果

拟仿射变换参数估计	旋转角度	尺度	平移 X	平移 Y
Case1 真实值	0	1.0	15	15
Case1 估计值	0	1.0	15	15
Case2 真实值	5	1.0	15	15
Case2 估计值	4.8	1.0	13	18
Case3 真实值	10	1.0	10	20
Case3 估计值	8.5	1.0	9.8	14.5
Case4 真实值	0	0.8	15	15
Case4 估计值	0	0.8	14	15
Case5 真实值	5	0.9	15	15
Case5 估计值	4.8	0.95	12	18

4.7 本章小结

本章介绍了基于特征点匹配的配准方法。特征点为 Gabor 滤波器卷积后的局部响应极值点，点对点关系通过小窗口邻域内的快速 NCC 计算获得，而最终变换参数通过 SVD 方法求解获得。特别需要强调本算法只适合微小几何畸变，如果发生了较大畸变（旋转超过 7 度，尺度变化超过 10%，平移超过了 NCC 计算的窗口大小）都会使得点对映射关系的建立时出现错误匹配点，且 SVD 算法本身也只适应在小旋转和尺度变换。所以基于特征点匹配的配准算法虽然精度超过高级特征匹配方法（比如第三章的曲线配准），但是能容忍的畸变程度很小。很自然的，可以考虑将前两章的算法统一起来，贯序执行，互相弥补不足。合并的整体算法在下一章介绍。

第五章 混合配准算法及其优化

5.1 优化需求与混合算法描述

当面对复杂背景的图像配准时,有效的配准算法的计算效率难免会比较低。当需要配准的图像尺寸增大时,优化的需求就必须得到重视。回顾本文前两章的算法,在计算上有多处瓶颈:

- 基于图分割的区域边界提取中,需要构造连接强度矩阵,矩阵大小正比与图像尺寸的平方。当图像尺寸为 1024×1024 时,连接强度的矩阵就要构造为 $1024 \times 1024 \times 1024 \times 1024$,普通计算机难以处理这样大小的矩阵操作。即使强度矩阵是稀疏矩阵,尺寸大小仍然需要减小,因为求解广义特征值问题时仍然需要复杂的大型矩阵为中间变量。所以,图分割适用的图像越小越好。
- 图分割输出的区域个数 K 通常不能取得很大,否则过分割输出多个曲线会增加曲线配准的计算量,且过分割考虑了过多的细节,而这恰恰是边缘检测算子的弊端。因此,图分割适用的图像细节越少越好。
- 基于 Gabor 提取特征点的点配准算法中,点对应关系是利用 NCC 计算获得的。尽管互相关在频域上有快速算法,但是 FFT 变换首先就要做延拓至 2 的指数的尺寸大小,且归一化的互相关系数计算也无法直接套用 FFT。然而,此算法的鲁棒性依赖有效特征点对的数量,而当特征点增多时又将面临大量 NCC 计算,因此必须降低 NCC 计算的复杂度。
- 通常配准算法都遵循循环迭代,逐步求精的思路,因此必须减少每次迭代的算法代价,同时还需要保证迭代的收敛性。

为此,本章介绍针对以上问题的优化办法,并综合给出如下配准混合算法的描述,见图 5.1 所示算法框架。

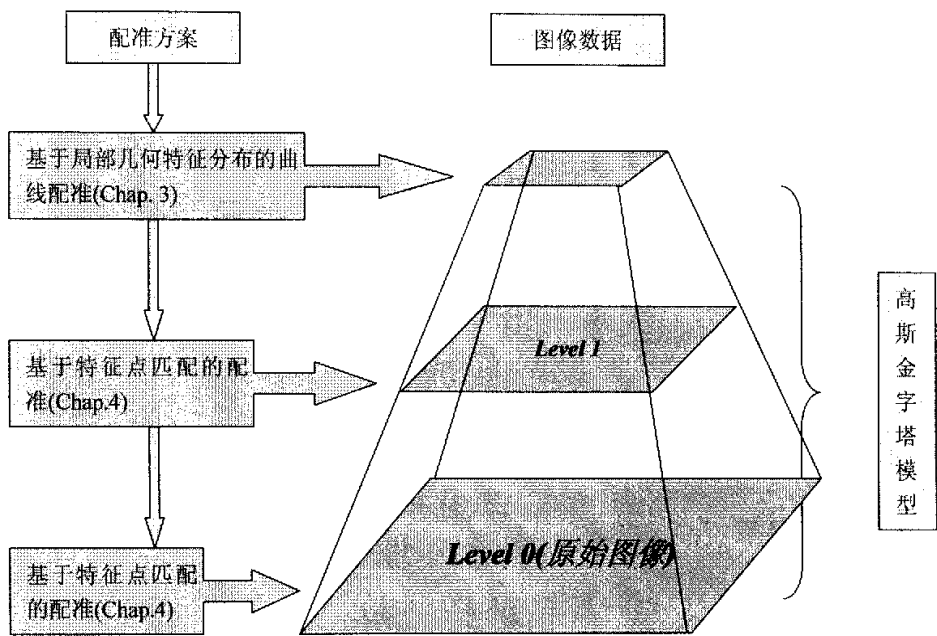


图 5.1 Hybrid 配准算法框架图

5.2 优化一：多分辨率图像结构

图像的金字塔结构建立了图像在多个分辨率下的表示形式。建立金字塔结构的目的在于产生一系列连续的参数用来表达图像对一系列不同尺度或者不同方向的滤波器的响应。在这里为了减少数据量，我们将采用高斯金字塔模型，它表示原始图像的一系列分层低通滤波分量表示，上层对应着下层的低通分量，如图 5.2:

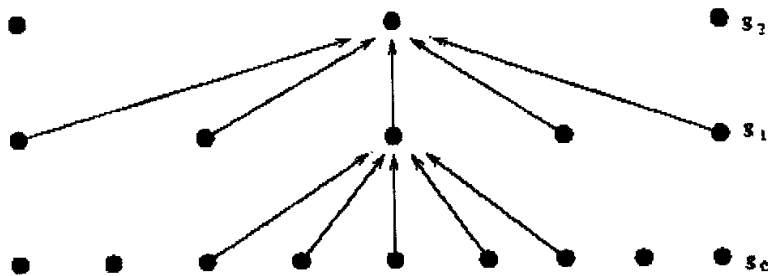


图 5.2 高斯金字塔结构

图像的低通滤波由图像与高斯核卷积完成，将结果降采样便得到上层数据。以 5×5 大小的离散滤波核和 4 倍降采样系数为例，定义如下算子：

$$REDUCE(I)[i, j] = \sum_{m=1}^5 \sum_{n=1}^5 w[m, n] I[2i + m, 2j + n] \quad (5.1)$$

由上式，高斯金字塔如下定义：

$$G_0 = I \quad (5.2)$$

$$G_{i+1} = REDUCE(G_i) \quad (5.3)$$

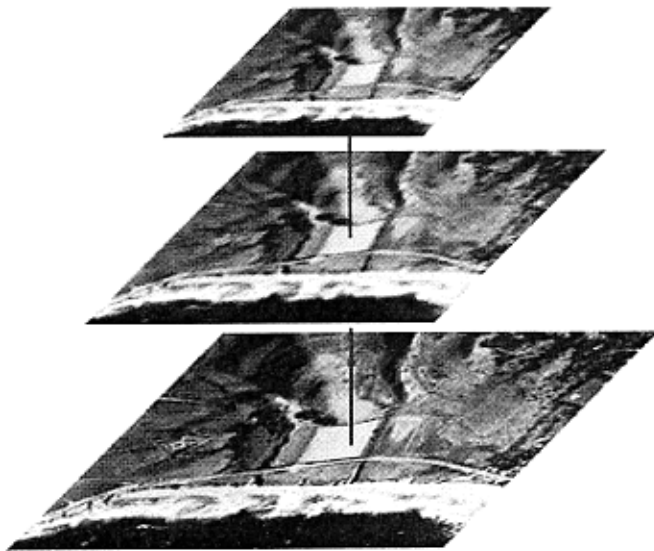


图 5.3 图像三层高斯金字塔结构示例

由于图分割方法计算量大，所以我们只在最高层图像上做曲线配准，这样选择的另外一个好处是，在最高层的低频信号图像上保留了图像整体结构而忽略了细节部分，更加适合图像分割。

5.3 优化二：快速 NCC 计算

两个信号的相关计算是特征检测的一个标注算法，经典介绍的快速计算方法皆是 FFT 变换法。但是，归一化相关系数的计算公式并没有特别简单和有效的频域计算公式，所以通常都是在空间域按照公式(4.6)直接计算的，计算量相对较大。为便于讨论，我们重写公式如下：

$$r(u, v) = \frac{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}] [t(x-u, y-v) - \bar{t}]}{\sqrt{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x-u, y-v) - \bar{t}]^2}} \quad (5.4)$$

这里设 \bar{t} 为图像 $t(N \times N)$ 的均值, $\bar{f}_{u,v}$ 为图像 $f(M \times M)$ 在当前窗口 (u, v) 下的均值。

再设 $t' = t(x - u, y - v) - \bar{t}$;

分子部分:

由于图像 t 的均值可以事先算好, 所以可以改写为:

$$\sum_{x,y} f(x, y) t'(x - u, y - v) - \bar{f}_{u,v} \sum_{x,y} t'(x - u, y - v)$$

又因为 t' 的均值为 0, 所以上式剩下第一个求和项, 可以用频域相关计算。

分母部分:

第二项只与图像 t 有关, 也可以只算一次, 大致需要 N^2 次运算。

第一项在每一次移动窗口下都需要计算一次方差, 总共大致需要 $3N^2(M - N + 1)^2$ 次运算, 运算量已经超过了频域相关。为此, 为图像 f 计算以下累计量:

$$s(u, v) = f(u, v) + s(u - 1, v) + s(u, v - 1) - s(u - 1, v - 1) \quad (5.5)$$

$$s^2(u, v) = f^2(u, v) + s^2(u - 1, v) + s^2(u, v - 1) - s^2(u - 1, v - 1) \quad (5.6)$$

且 $s(u, v) = s^2(u, v) = 0$ 当 $u, v < 0$

此时定义在窗口位置为 (u, v) 下, 图像 f 的能量为:

$$e_f(u, v) = s^2(u + N - 1, v + N - 1) - s^2(u - 1, v + N - 1) - s^2(u + N - 1, v - 1) + s^2(u - 1, v - 1) \quad (5.7)$$

将 $\sum_{x,y} [f(x, y) - \bar{f}_{u,v}]^2$ 中展开二次项, 便可以直接利用之前计算好的累计项 S 带入

计算, 共需 $3M^2$ 次运算, 比直接计算明显减少了计算量。

5.4 优化三：并行计算探讨初步

由于配准算法计算比较复杂, 很多学者也着手考虑配准的并行处理。Ozkan^[111]使用了 5 个处理器来执行一种基于遗传算法的图像配准技术; Turton^[112]设计出并行配准算法的硬件实现; 数据在多个处理器上分配处理后, 计算效率被明显提高。

配准问题的并行处理需要考虑算法和数据的分配、负载平衡、任务申请调度等方面。在算法和数据的分配上最常见的是考虑同一算法需要执行多个不同数据的要求, 例如第四章中介绍的 NCC 计算。负载平衡也是一个重要环节, 尤其当分配算法和数据得到的益处与处理器之间相互通信的时间具有同等数量级时, 如果通信与数据传递过于频繁, 很可能将降低并行算法的效率。并行算法的效率可以简单用如下公式评价:

$$Speed_N(Alg.) = \frac{t_{seq}(Alg.)}{t_N(Alg.)} \quad (5.8)$$

$$Efficiency_N = \frac{Speed_N(Alg.)}{N} \quad (5.9)$$

$$L = \frac{1}{N} \sum_{i=1}^N (t_i - t_N) \quad (5.10)$$

(5.8)式表示算法 ALG 在 N 个处理器上提高的速度, (5.9)式表示使用每个处理器的效率。 (5.10)表示平均负载非平衡系数。其中 t_i 为每个处理器一次执行的总消耗时间, t_N 为执行配准算法的时间, t_{seq} 为顺序执行配准算法的总时间。

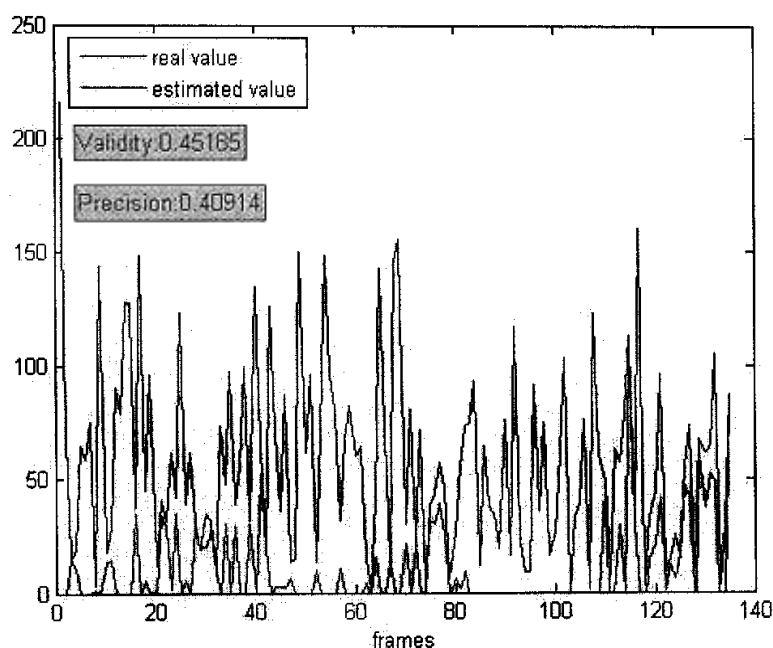
在并行算法执行的平台方面, 按照执行指令和数据的方式可以分为: 1、单指令单数据(SISD); 2、单指令多数据(SIMD); 3、多指令单数据(MISD); 4、多指令多数据(MIMD)。现代计算机都采用 MIMD 的体系, 并称为大规模并行处理器结构(Massive Parallel Processors, MPP)。以下两种 MPP 解决方案可用于图像配准的并行算法:

表 5.1 高性能计算平台^[113]

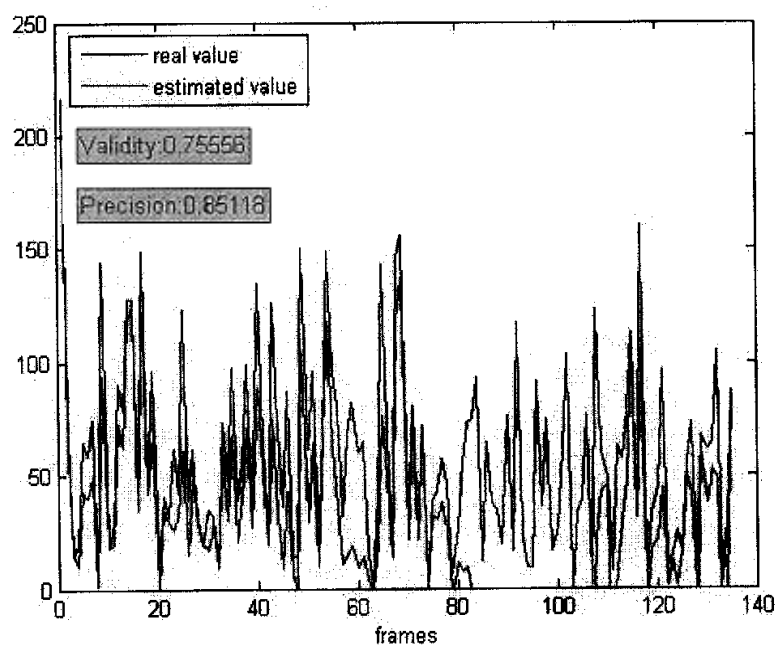
Features	Beowulf (usrhive.gfsc.nasa.gov)	Cray T3E-900 (yukon.arsc.edu)
Processors	128 Pentium Pro 200 MHz	DEC Alpha 26124 450MHz
No. of Nodes	64	256
Memory/PE	384 Mbytes	256 Mbytes
Interconnection Network	FastEthernet (100 Mbits/sec)	3D Torus (122Gbits/sec)
Communication Library used	Message Passing Interface (MPI)	Message Passing Interface (MPI)
Operating System	Linux	Unicos
Peak Performance	25.6 Gflops	234 Gflops

5.5 混合配准算法实验

如图 5.4 为普通相关算法与本文提出的混合配准算法的对比实验。实验所用 2 组图像中参考图像序列为卫星图像, 畸变图像序列为实际航拍图像, 畸变类型包括平移、旋转、轻微尺度、光照非线性变化、景物内容的改变。共有 135 副航拍图片与卫星图片之间的配准。其中蓝色曲线代表实际变换参数, 红色曲线代表算法估计的变换参数。可以看到普通基于相关的算法在很多此类图像的配准上都失败了, 但是本文算法的有效性提高了很多。



A) 互相关算法配准结果



B) 混合算法配准结果

图 5.4 混合配准算法 vs 互相关配准算法。

5.6 本章小结

本章综合前两章的算法, 提出由粗到精的配准混合算法。为了提高效率, 分别提出了在多分辨率模型、快速相关计算和并行计算的优化方案。Gauss 金字塔多分辨率模型减少了数据量并增加了特征提取的鲁棒性, 快速 NCC 计算提高了搜索效率, 并行计算可以进一步加快计算速度。以较为复杂的卫星图片为实验图库, 通过实验对比体现了此配准混合算法的有效性。

第六章 总结和展望

6.1 本文主要工作总结

本文针对 2D 图像存在几何畸变和灰度变化情况下,提出一种基于特征的分层配准模型,由粗到精逐步求解。

图像中的区域边界由于在通常的灰度变换、光照变化下保持稳定,所以以区域边界为代表的曲线适合作为初步配准时图像特征。提取曲线的方法采用基于图分割的方法,综合考虑了图像中的灰度、纹理和空间信息,曲线输出效果明显优于一般的边缘检测算子。曲线之间的对应关系和变换参数的最优估计同时在 RANSAC 算法下完成,参数估计通过基于曲线局部几何特征分布的互相关求得,不需要建立点与点的对应关系,因而也绕开了由于灰度信息变化大造成的点对应关系建立的难题。在经过粗配准后,进一步利用基于 Gabor 滤波提取的特征点做 SVD 求解点对的精确配准。Gabor 滤波器保证了特征点一般处于图像中的拐点等明显标志性位置,图像畸变后这些点一般保持相对稳定的信息,所以在小范围窗口经过 NCC 计算即可得到对应点关系。最后利用 SVD 方法在最小平方误差意义上求得变换参数,作为配准算法的最后输出。

为了提高运算速度,本文实现了多分辨率的图像模型,对最高一层低频图像采用曲线配准,在其余层次采用特征点配准,以上一层的输出为基准限制下一层的搜索范围,进一步减少计算量。并利用快速 NCC 算法减少特征点匹配中的大量相关计算负担。最后提出并行计算的初步构想,将像 NCC 这样的同一种算法执行不同数据的情况分配到多个处理器上同时计算。

6.2 进一步研究方向

将来进一步研究方向包括以下几个尚未解决或实现的目标:

- 基于图分割的区域曲线提取算法仍然需要快速算法,当图像尺寸增大时,求解大型矩阵的特征值会非常耗时。
- 在计算曲线的局部几何特征分布的相关运算时,如果出现多个峰值,则必须进一步设计算法判断正确的峰值位置。
- 提高曲线的局部几何特征分布在计算上的精确性,包括各阶导数运算、曲线拟合方法。
- 提高点配准模型抗噪音点的能力,目前的算法在出现强噪音情况下会出现很多错误的误匹配点对。
- 实现完整的图像配准并行计算机制和测试环境。在国内外都很少有在这方面的成果,主要问题是如何设计更好的可分离算法模块并行执行的方法。

附录 A 基于 K 平均聚类的纹理特征提取

K-平均聚类是基于混合高斯概率密度模型的参数估计方法，本文将训练图像与一组滤波器卷积^[105]，其滤波响应结果构成特定纹理特征向量。把特征向量看成多元随机变量，并假设特征向量集合服从多元混合高斯概率密度模型分布，利用 K-平均聚类算法自动得到各种纹理分类的混合高斯概率密度参数，从而可以方便的构建基本纹理特征库。这里主要介绍文献^[114]中阐述的聚类方法。

● 混合高斯模型的参数估计

混合高斯模型的参数估计，是对由不同子类构成的概率密度分布进行建模。为了估计混合高斯概率密度模型分布的参数，需要决定子类的数量和每个子类的参数。可以通过使用训练数据的采样和估计子类的数目及其参数来实现。

设Y为混合高斯分布的m维随机向量。我们假设这个模型有K个子类。下面这些参数用来描述第k个子类。

π_k - 子类k的概率。

μ_k - 子类k的m维均值。

R_k - 子类k的m×m协方差矩阵。

我们使用 π , μ 和 R 表示参数集合 $\{\pi_k\}_{k=1}^K$, $\{\mu_k\}_{k=1}^K$ 和 $\{R_k\}_{k=1}^K$ 。分类的完整的参数集合由K和 $\theta = (\pi, \mu, R)$ 给出。K为正整数， $\pi \geq 0$ ，并且 $\sum_k \pi_k = 1, \det(R) \geq \varepsilon$ ， ε 大小依赖与具体的应用。

令 Y_1, Y_2, \dots, Y_N 为N个从分类采样的多维特征向量。设对于每个向量 Y_i ，其从属子类是由随机变量 X_n 给出，当然，一般 X_n 我们开始并不知道其具体值。进一步设每个子类服从多元高斯分布，向量 Y_n 的概率密度函数在 X_n 是第k个子类的情况下可写为：

$$P_{y_n|x_n}(y_n | k, \theta) = \frac{1}{(2\pi)^{m/2}} |R_k|^{-1/2} \exp\left\{-\frac{1}{2}(y_n - \mu_k)' R_k^{-1} (y_n - \mu_k)\right\} \quad (A.1)$$

因为每个采样的子类 X_n 的概率密度分布参数未知，所以为了计算 Y_n 的概率密度函数，给定参数 θ ，由条件概率定义得：

$$P_{y_n}(y_n | \theta) = \sum_{k=1}^K P_{y_n|x_n}(y_n | k, \theta) \pi_k \quad (A.2)$$

这个采样序列 $Y = \{Y_n\}_{n=1}^N$ 属于类别k的总概率的对数为：

$$\log P_y(y|k, \theta) = \sum_{n=1}^N \log \left(\sum_{k=1}^K P_{y_n|x_n}(y_n|k, \theta) \pi_k \right) \quad (\text{A.3})$$

根据上式可估计参数 K 和 $\theta \in \Omega^{(K)}$ 。利用最大似然(ML)估计, 得到:

$$\hat{\theta}_{ML} = \arg \max_{\theta \in \Omega^{(K)}} \log P_y(y|k, \theta) \quad (\text{A.4})$$

在上式中关于 K 的 ML 估计并没有很好的定义, 因为随着 K 的增长 \log 函数值总是在增长的, 换句话说, 分类数量越多, 模型当然越准确。解决这个问题的一种办法是 Rissanen^[115]提出的最小描述长度准则法(MDL), 约束使用的模型阶数, 防止过学习问题。本文采用的目标函数为:

$$L = K(1 + m + \frac{(m+1)m}{2}) - 1 \quad (\text{A.5})$$

$$MDL(K, \theta) = -\log P_y(y|K, \theta) + \frac{1}{2} L \log(Nm) \quad (\text{A.6})$$

其中 N 是指采样个数。因此在当前的采样个数下, 选择的 K 值应当同时使 MDL 函数最小化。

由于在实际的应用过程中, 采样的数据是不完全的, 因此用期望值-最大化(EM)算法可以准确地解决这类“不完全”的数据问题。EM 算法是将向量集合 Y_n 大致分类, 并求取其分类参数 $\theta^{(i)}$, 再重新根据概率最大化的式子求取每个向量的分类, 如此反复迭代得到最终的分类结果。首先假定有 k 个分类, 在计算出每个类别的分类参数 $\theta^{(i)}$ 后, 属于某个子类 k 的向量 y_n 的概率密度分布就可以由贝叶斯规则来求得, 公式如下:

$$P_{x_n|y_n}(k|y_n, \theta^{(i)}) = \frac{P_{x_n|y_n}(y_n|k, \theta^{(i)}) \pi_k}{\sum_{l=1}^K P_{x_n|y_n}(y_n|l, \theta^{(i)}) \pi_l} \quad (\text{A.7})$$

重新计算每个分类新的预测均值和协方差, 记为 $\bar{\pi}_k$, $\bar{\mu}_k$ 和 \bar{R}_k 。计算公式如下:

$$\bar{N}_k = \sum_{n=1}^N p_{x_n|y_n}(k|y_n, \theta^{(i)}) \quad (\text{A.8})$$

$$\bar{\pi}_k = \frac{\bar{N}_k}{N} \quad (\text{A.9})$$

$$\bar{\mu}_k = \frac{1}{N_k} \sum_{n=1}^N y_n p_{x_n|y_n}(k | y_n, \theta^{(i)}) \quad (\text{A.10})$$

$$\bar{R}_k = \frac{1}{N_k} \sum_{n=1}^N (y_n - \bar{\mu}_k)(y_n - \bar{\mu}_k)^t p_{x_n|y_n}(k | y_n, \theta^{(i)}) \quad (\text{A.11})$$

以上参数可由最大化下式得到的：

$$(\pi^{(i+1)}, \mu^{(i+1)}, R^{(i+1)}) = \arg \max_{(\pi, \mu, R) \in \Omega^{(K)}} Q(\theta; \theta^{(i)}) = (\bar{\pi}, \bar{\mu}, \bar{R}) \quad (\text{A.12})$$

初始化参数 $\theta^{(1)}$ 的选择将决定 EM 算法是否能够有效地收敛到局部极值。初始的聚类数目 K_0 ，由用户根据应用来决定，其他参数由下式决定：

$$\pi_k^{(1)} = \frac{1}{K_0} \quad (\text{A.13})$$

$$\mu_k^{(1)} = y_n \text{ where } n = [(K-1)(N-1)/(K_0-1)] + 1 \quad (\text{A.14})$$

$$R_k^{(1)} = \frac{1}{N} \sum_{n=1}^N y_n y_n^t \quad (\text{A.15})$$

● K-平均聚类算法步骤

1. 初始化分类数目为 K_0 。
2. 使用(A.13)，(A.14)和(A.15)初始化 $\theta^{(1)}$ 。
3. 应用迭代EM算法(A.12)直到MDL(K, θ)中的变化小于 ε 。
4. 记录参数 $\theta^{(K, \text{ifinal})}$ ，和MDL(K, $\theta^{(K, \text{ifinal})}$)的值。
5. 如果子类数量大于1，减少聚类的数量，令 $K \leftarrow K-1$ ，并且返回到第三步。
6. 选择 K^* 的值和最小化MDL值得最终的参数 $\theta^{(K, \text{ifinal})}$ 。

参考文献

- [1] L. G. Brown, A survey of image registration techniques, *ACM Computing Surveys* 24 (1992) 326 - 376.
- [2] B. K. Ghaffary, A. A. Sawchuk, A survey of new techniques for image registration and mapping, *Proc. SPIE: Applications of Digital Image Processing* 432 (1983) 222 - 239.
- [3] P. A. van den Elsen, E.-J. D. Pol, M. A. Viergever, Medical image matching - a review with classification, *IEEE Engineering in Medicine and Biology* 12 (1993) 26 - 39.
- [4] J. B. A. Maintz, M. A. Viergever, A survey of medical image registration, *Medical Image Analysis* 2 (1998) 1 - 36.
- [5] H. Lester, S. R. Arridge, A survey of hierarchical non-linear medical image registration, *Pattern Recognition* 32 (1999) 129 - 149.
- [6] L. M. G. Fonseca, B. S. Manjunath, Registration techniques for multisensor remotely sensed imagery, *Photogrammetric Engineering and Remote Sensing* 62 (1996) 1049 - 1056.
- [7] E. Gulch, Results of test on image matching of ISPRS WG, *ISPRS Journal of Photogrammetry and Remote Sensing* 46 (1991) 1 - 18.
- [8] J. le Moigne, First evaluation of automatic image registration methods, *Proc. Int'l. Geoscience and Remote Sensing Symposium IGARSS'98* (Seattle, Washington, 1998) 315 - 317.
- [9] W. K. Pratt, Correlation techniques of image registration, *IEEE Trans. Aerospace and Electronic Systems* 10 (1974) 353 - 358.
- [10] W. K. Pratt, *Digital Image Processing*, John Wiley, New York, 2nd ed., 1991.
- [11] D. I. Barnea, H. F. Silverman, A class of algorithms for fast digital image registration, *IEEE Trans. Computing* 21 (1972) 179 - 186.
- [12] A. Goshtasby, Template matching in rotated images, *IEEE Trans. Pattern Analysis and Machine Intelligence* 7 (1985) 338 - 344.
- [13] R. Berthilsson, Affine correlation, *Proc. Int'l. Conf. on Pattern Recognition ICPR'98* (Brisbane), Australia, (1998) 1458 - 1461.
- [14] A. Simper, Correcting general band-to-band misregistrations, *Proc. IEEE Int'l. Conf. on Image Processing ICIP'96*, 2 (Lausanne, Switzerland, 1996) 597 - 600.
- [15] P. vanWie, M. Stein, A Landsat digital image rectification system, *IEEE Trans. Geoscience Electronics* 15 (1977) 130 - 136.
- [16] P. E. Anuta, Spatial registration of multispectral and multitemporal digital imagery using Fast Fourier Transform, *IEEE Trans. Geoscience Electronics* 8 (1970) 353 - 368.
- [17] D. P. Huttenlocher, G. A. Klanderman, W. J. Rucklidge, Comparing images using the Hausdorff distance, *IEEE Trans. Pattern Analysis and Machine Intelligence* 15 (1993) 850 - 863.
- [18] Dubuisson, A modified Hausdorff distance for object matching, *Proc 12th Int'l Conf on Pattern Recognition*, 1994

- [19] Azencott, Multiscale identification of buildings in compressed large aerial scenes, Proc. 13th International Conference on Pattern Recognition (ICPR'96) ,Vol 3, p. 974, 1996.
- [20] Sim, Pyramidal robust Hausdorff distance for object matching, Proc. IEEE Int' Conf. on Image Processing ICIP'99, 1999.
- [21] Jiang, "Phase correlation-based matching method with sub-pixel accuracy for translated and rotated images", ICSP'02 Proceedings. p752-755, 2002
- [22] E. de Castro, C. Morandi, Registration of translated and rotated images using finite Fourier transform, IEEE Trans. Pattern Analysis and Machine Intelligence 9 (1987) 700 - 703.
- [23] B. S. Reddy, B. N. Chatterji, An FFT-based technique for translation, rotation and scale-invariant image registration, IEEE Trans. Image Processing 5 (1996) 1266 - 1271.
- [24] Q. Chen, M. Defrise, F. Deconinck, Symmetric phase-only matched filtering of Fourier-Mellin transform for image registration and recognition, IEEE Trans. Pattern Analysis and Machine Intelligence 16(1994) 1156 - 1168.
- [25] P. Viola, W. M. Wells, Alignment by maximization of mutual information, Int'l. Journal of Computer Vision 24 (1997) 137 - 154.
- [26] R. K. Sharma, M. Pavel, Multisensor image registration, Proc. Society for Information Display XXVIII (1997) 951 - 954.
- [27] H. S. Sawhney, R. Kumar, True multi-image alignment and its applications to mosaicing and lens distortion correction, IEEE Trans. Pattern Analysis and Machine Intelligence 21 (1999) 235 - 243.
- [28] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, P. Suetens, Multimodality image registration by maximization of mutual information, IEEE Trans. Medical Imaging 16 (1997) 187 - 198.
- [29] P. Th' evenaz, M. Unser, A pyramid approach to sub-pixel image fusion based on mutual information, Proc. IEEE Int' l. Conf. on Image Processing ICIP' 96 (Lausanne, Switzerland, 1996) 265 - 268.
- [30] P. Th' evenaz, M. Unser, An efficient mutual information optimizer for multiresolution image registration, Proc. IEEE Int' l. Conf. on Image Processing ICIP' 98 (Chicago, Illinois, 1998) 833 - 837.
- [31] P. Th' evenaz, M. Unser, Spline pyramids for inter-modal image registration using mutual information, Proc. SPIE: Wavelet Applications in Signal and Image Processing (San Diego, California, 1997) 236 - 247.
- [32] N. Ritter, R. Owens, J. Cooper, R. H. Eikelboom, P. P. van Saarloos, Registration of stereo and temporal images of the retina, IEEE Trans. Medical Imaging 18 (1999) 404 - 418.
- [33] C. Studholme, D. L. G. Hill, D. J. Hawkes, An overlap invariant entropy measure of 3D medical image alignment, Pattern Recognition 32 (1999) 71 - 86.
- [34] D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach, D. J. Hawkes, Non-rigid registration of breast MR images using mutual information, Proc. Medical Image Computing and Computer-Assisted Intervention MICCAI' 98 (Cambridge, Massachusetts, 1998) 1144 - 1152.
- [35] B. Likar, F. Pernus, Elastic registration of muscle fiber images, Proc. Int' l Workshop on Biomedical Image Registration WBIR' 99 (Bled, Slovenia, 1999) 24 - 45.
- [36] J. P. W. Pluim, J. B. A. Maintz, M. A. Viergever, Mutual information in multiresolution contexts, Proc.

- Int' l Workshop on Biomedical Image Registration WBIR' 99 (Bled, Slovenia, 1999) 46 - 60.
- [37] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, D. J. Hawkes, A comparison of similarity measures for use in 2D - 3D medical image registration, *IEEE Trans. Medical Imaging* 17(1998) 586 - 595.
- [38] A. Goshtasby, G. C. Stockman, C. V. Page, A region-based approach to digital image registration with-subpixel accuracy, *IEEE Trans. Geoscience and Remote Sensing* 24 (1986) 390 - 399.
- [39] J. Flusser, T. Suk, A moment-based approach to registration of images with affine geometric distortion, *IEEE Trans. Geoscience and Remote Sensing* 32 (1994) 382 - 387.
- [40] A. Goshtasby, G. C. Stockman, Point pattern matching using convex hull edges, *IEEE Trans. Systems, Man and Cybernetics* 15 (1985) 631 - 637.
- [41] M. Holm, Towards automatic rectification of satellite images using feature based matching, *Proc. Int' l. Geoscience and Remote Sensing Symposium IGARSS' 91* (Espoo, Finland, 1991) 2439 - 2442.
- [42] Y. C. Hsieh, D. M. McKeown, F. P. Perlant, Performance evaluation of scene registration and stereo-matching for cartographic feature extraction, *IEEE Trans. Pattern Analysis and Machine Intelligence* 14 (1992) 214 - 237.
- [43] M. Sester, H. Hild, D. Fritsch, Definition of ground control features for image registration using GIS data, *Proc. Symposium on Object Recognition and Scene Classification from Multispectral and Multisensor Pixels*.
- [44] M. Roux, Automatic registration of SPOT images and digitized maps, *Proc. IEEE Int'l. Conf. on Image Processing ICIP' 96* (Lausanne, Switzerland, 1996) 625 - 628.
- [45] P. A. Brivio, A. D. Ventura, A. Rampini, R. Schettini, Automatic selection of control points from shadow structures, *Int' l. Journal of Remote Sensing* 13 (1992) 1853 - 1860.
- [46] N. R. Pal, S. K. Pal, A review on image segmentation techniques, *Pattern Recognition* 26 (1993) 1277 - 1294.
- [47] W. H. Wang, Y. C. Chen, Image registration by control points pairing using the invariant properties of line segments, *Pattern Recognition Letters* 18 (1997) 269 - 281.
- [48] S. Moss, E. R. Hancock, Multiple line-template matching with EM algorithm, *Pattern Recognition Letters* 18 (1997) 1283 - 1292.
- [49] H. Li, B. S. Manjunath, S. K. Mitra, A contour-based approach to multisensor image registration, *IEEE Trans. Image Processing* 4 (1995) 320 - 334.
- [50] X. Dai, S. Khorram, Development of a feature-based approach to automated image registration for-multitemporal and multisensor remotely sensed imagery, *Int' l. Geoscience and Remote Sensing Symposium IGARSS' 97* (Singapore, 1997) 243 - 245.
- [51] V. Govindu, C. Shekhar, R. Chellapa, Using geometric properties for correspondence-less image alignment, *Proc. Int' l. Conf. on Pattern Recognition ICPR' 98* (Brisbane, Australia, 1998) 37 - 41.
- [52] D. Shin, J. K. Pollard, J. P. Muller, Accurate geometric correction of ATSR images, *IEEE Trans. Geoscience and Remote Sensing* 35 (1997) 997 - 1006.
- [53] H. Maitre, Y. Wu, Improving dynamic programming to solve image registration, *Pattern Recognition* 20

(1987) 443 - 462.

- [54] S. Z. Li, J. Kittler, M. Petrou, Matching and recognition of road networks from aerial images, Proc.2nd European Conf. on Computer Vision ECCV' 92 (St. Margherita, Italy, 1992) 857 - 861.
- [55] N. Vujovic, D. Brzakovic, Establishing the correspondence between control points in pairs of mammographic images, IEEE Trans. Image Processing 6 (1997) 1388 - 1399.
- [56] J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Analysis and Machine Intelligence 8 (1986) 679 - 698.
- [57] D. Marr, E. Hildreth, Theory of edge detection, Proc. R. Soc. London B 207 (1980) 187 - 217.
- [58] J. B. A. Maintz, P. A. van den Elsen, M. A. Viergever, Evaluation on ridge seeking operators for multi-modality medical image matching, IEEE Trans. Pattern Analysis and Machine Intelligence 18(1996) 353 - 365.
- [59] J. B. A. Maintz, P. A. van den Elsen, M. A. Viergever, Comparison of edge-based and ridge-based registration of CT and MR brain images, Medical Image Analysis 1 (1996) 151 - 161.
- [60] G. Stockman, S. Kopstein, S. Benett, Matching images to models for registration and object detection via clustering, IEEE Trans. Pattern Analysis and Machine Intelligence 4 (1982) 229 - 241.
- [61] A. S. Vasileisky, B. Zhukov, M. Berger, Automated image co-registration based on linear feature recognition, Proc. 2nd Conf. Fusion of Earth Data (Sophia Antipolis, France, 1998) 59 - 66.
- [62] S. Growe, R. Tonjes, A knowledge based approach to automatic image registration, Proc. IEEE Int'l. Conf. on Image Processing ICIP' 97 (Santa Barbara, California, 1997) 228 - 231.
- [63] J. Ton, A. K.Jain, Registering Landsat images by point matching, IEEE Trans. Geoscience and Remote Sensing 27 (1989) 642 - 651.
- [64] M. Ehlers, Region-based matching for image registration in remote sensing databases, Proc. Int'l. Geoscience and Remote Sensing Symposium IGARSS' 91 (Espoo, Finland, 1991) 2231 - 2234.
- [65] Q. Zheng, R. Chellapa, A computational vision approach to image registration, IEEE Trans. Image Processing 2 (1993) 311 - 325.
- [66] B. S. Manjunath, C. Shekhar, R. Chellapa, A new approach to image feature detection with applications, Pattern Recognition 29 (1996) 627 - 640.
- [67] W. S. I. Ali, F. S. Cohen, Registering coronal histological 2-D sections of a rat brain with coronal sections of a 3-D brain atlas using geometric curve invariants and B-spline representation, IEEE Trans. Medical Imaging 17 (1998) 957 - 966.
- [68] S. Banerjee, D. P. Mukherjee, D. D. Majumdar, Point landmarks for registration of CT and NMR images, Pattern Recognition Letters 16 (1995) 1033 - 1042.
- [69] L. M. G. Fonseca, M. H. M. Costa, Automatic registration of satellite images, Proc. Brazilian Symposium on Computer Graphic and Image Processing (Brazil, 1997) 219 - 226.
- [70] J. W. Hsieh, H. Y. M. Liao, K. C. Fan, A fast algorithm for image registration without predetermining correspondence, Proc. Int'l. Conf. on Pattern Recognition ICPR' 96 (Vienna, Austria, 1996) 765 - 769.
- [71] L. Kitchen, A. Rosenfeld, Gray-level corner detection, Pattern Recognition Letters 1(1982) 95 - 102.

- [72] L. Dreschler, H. Nagel, Volumetric model and 3-D trajectory of a moving car derived from monocular TV-frame sequence of a street scene, Proc. Int' l. Joint Conf. Artificial Intelligence (Vancouver, Canada, 1981) 692 - 697.
- [73] P. R. Beaudet, Rotationally invariant image operators, Proc. Int' l. Joint Conference on Pattern Recognition(Kyoto, Japan, 1978) 579 - 583.
- [74] W. Forstner, E. Gulch, A fast operator for detection and precise location of distinct points, corners and centers of circular features, Proc. ISPRS Workshop on Fast Processing of Photogrammetric Data(Interlaken, Switzerland, 1986) 281 - 305.
- [75] A. Noble, Finding corners, Image and Vision Computing 6 (1988) 121 - 128.
- [76] T. M. Lehman, A two stage algorithm for model-based registration of medical images, Proc. Int'l. Conf. on Pattern Recognition ICPR' 98 (Brisbane, Australia, 1998) 344 - 352.
- [77] S. M. Smith, J. M. Brady, SUSAN - A new approach to low level image processing, Int'l. Journal of Computer Vision 23 (1997) 45 - 78.
- [78] D. Lavine, B. Lambird, L. Kanal, Recognition of spatial point patterns, Pattern Recognition 16 (1983) 289 - 295.
- [79] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, H. C. Wolf, Parametric correspondence and chamfer-matching: Two new techniques for image matching, Proc. 5th Int'l. Joint Conf. Art. Intell. (Cambridge, Massachusetts, 1977) 659 - 663.
- [80] J. Flusser, Object matching by means of matching likelihood coefficients, Pattern Recognition Letters16 (1995) 893 - 900.
- [81] S. Abdelsayed, D. Ionescu, D. Goodenough, Matching and registration method for remote sensing images, Proc. Int'l. Geoscience and Remote Sensing Symposium IGARSS'95 (Florence, Italy, 1995)1029 - 1031.
- [82] D. Skea, I. Barrodale, R. Kuwahara, R. Poeckert, A control point matching algorithm, Pattern Recognition26 (1993) 269 - 276.
- [83] T. Suk, J. Flusser, Vertex-based features for recognition of projectively deformed polygons, Pattern Recognition 29 (3) (1996) 361 - 367.
- [84] T. Suk, J. Flusser, Point-based projective invariants, Pattern Recognition 33 (2000) 251 - 261.
- [85] M. K. Hu, Visual pattern recognition by moment invariants, IRE Trans. Information Theory8,179 - 187.
- [86] J. Flusser, T. Suk, Pattern recognition by affine moment invariants, Pattern Recognition 26(1993) 167-174.
- [87] X. Dai, S. Khorram, A feature-based image registration algorithm using improved chain-code representation combined with invariant moments,IEEE Trans. Geoscience and Remote Sensing37(1999)2351 - 2362.
- [88] J. Sato, R. Cipolla, Image registration using multi-scale texture moments, Image and Vision Computing 13 (1995) 341 - 353.
- [89] J. Flusser, T. Suk, Degraded image analysis: An invariant approach, IEEE Trans. Pattern Analysis and Machine Intelligence 20 (1998) 590 - 603.

- [90] J. Flusser, B. Zitov' a, Combined invariants to linear filtering and rotation, *Int'l. Journal of Pattern Recognition and Artificial Intelligence* 13 (8) (1999) 1123 - 1136.
- [91] J. Flusser, J. Boldys, B. Zitov' a, Invariants to convolution in arbitrary dimensions, *Journal of Mathematical Imaging and Vision* 13 (2000) 101 - 113.
- [92] R. S. Mitra, N. N. Murthy, Elastic maximal matching, *Pattern Recognition* 24 (1991) 747 - 753.
- [93] S. Ranade, A. Rosenfeld, Point pattern matching by relaxation, *Pattern Recognition* 12 (1980) 269 - 275.
- [94] G. Medioni, R. Nevatia, Matching images using linear features, *IEEE Trans. Pattern Analysis and Machine Intelligence* 6 (1984) 675 - 685.
- [95] J. K. Cheng, T. S. Huang, Image registration by matching relational structures, *Pattern Recognition* 17 (1984) 149 - 159.
- [96] F. H. Cheng, Point pattern matching algorithm invariant to geometrical transformation and distortion, *Pattern Recognition Letters* 17 (1996) 1429 - 1435.
- [97] H. Ogawa, Labeled point pattern matching by fuzzy relaxation, *Pattern Recognition* 17 (1984) 569 - 573.
- [98] S. Z. Li, Matching: Invariant to translations, rotations and scale changes, *Pattern Recognition* 25 (1992) 583 - 594.
- [99] K. E. Price, Relaxation matching techniques: a comparison, *IEEE Trans. Pattern Analysis and Machine Intelligence* 7 (1985) 617 - 623.
- [100] C. Davatzikos, J. L. Prince, R. N. Bryan, Image registration based on boundary mapping, *IEEE Trans. Medical Imaging* 15 (1996) 112 - 115.
- [101] W. Peckar, C. Schnorr, K. Rohr, H. S. Stiehl, Two step parameter-free elastic image registration with prescribed point displacements, *Journal of Mathematical Imaging and Vision* 10 (1999) 143 - 162.
- [102] J. Shi and J. Malik. Normalized cuts and image segmentation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'97)*, p731--737, 1997
- [103] S. Yu and J. Shi. Multiclass spectral clustering. In *Proc. of ICCV*, 2003
- [104] XY Stella. Segmentation using multiscale cues. In *Proc. CVPR*, vol 1, p247 - 254, 2004
- [105] P. Perona and J. Malik. Detecting and localizing edges composed of steps, peaks and roofs. In *Proc. Int. Conf. Computer Vision*, pages 52 - 7, Osaka, Japan, Dec 1990.
- [106] A. Bruckstein, Invariant signatures for planar shape recognition under partial occlusion, In *Proc. CVGIP: Image Understanding*, Vol 58, p49 - 65, 1993.
- [107] J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. *J. Optical Society of America*, 7(2):923 - 32, May 1990.
- [108] M. A. Fischler, R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Comm. of the ACM*, Vol 24, pp 381-395, 1981.
- [109] P. Meer and I. Weiss, Smoothed differentiation filters for images, *Journal of Visual Communication and Image Representation*, Vol. 3, pp. 58--72, 1992.
- [110] Shinji Umeyama. Least-Squares Estimation of Transformation Parameters Between Two Point Patterns, *IEEE Trans. PAMI*, 1991, vol. 13: pp. 376-380

- [111] M. Ozkan, J. M. Fitzpatrick, and K. Kawamura, "Image Registration for a Transputer-Based Distributed System," in Proc. the 2nd International Conference on Industrial & Engineering Applications of AI & Expert Systems(IEA/AIE-89), pp. 908-915, June 6-9, 1989.
- [112] B. Turton, T. Arslan, and D. Horrocks, "A hardware architecture for a parallel genetic algorithm for image registration," in Proc. of IEE Colloquium on Genetic Algorithms in Image Processing and Vision, pp. 111-116, Oct. 1994.
- [113] Prachya Chalermwat, High Performance Automatic Image Registration For Remote Sensing, Phd Thesis, 1999
- [114] Bouman, "CLUSTER: An Unsupervised Algorithm for Modeling Gaussian Mixtures", technique report, <http://www.ece.purdue.edu/~bouman>
- [115] J. Rissanen, "A Universal Prior for Integers and Estimation by Minimum Description Length", Annals of Statistics, vol. 11, no. 2, pp. 417-431, 1983.

攻读硕士期间发表和完成的学术论文

- [1] Donghui Wang, Chenyang Cui, **Zheng Wu**, "Matching 3D models with global geometric feature map", in *Proc. IEEE 12th Int. Conf. on Multimedia Modelling*, Beijing, Jan. 2006
- [2] **Zheng Wu**, Hui Qian, Miaoliang Zhu, "Geo-registration of aerial images by robust feature matching", in *Proc. SPIE Visual Communications and Image Processing 2006*, San Jose, California USA, Jan. 2006
- [3] 吴铮, 钱徽, 朱淼良, "一种新的基于特征的航空图像匹配算法", 第十二届中国图形图像学术会议, p. 381-386, 北京, Oct. 2005
- [4] **Zheng Wu**, Donghui Wang, Miaoliang Zhu, "Video stream interception and erotic content filtering", in *Proc. 16th Int. Conf. on Computer Communication*, vol.2, p. 1601-5, Beijing, China, 2004
- [5] 龚建军, 朱淼良, 吴铮, 王东辉, "多媒体短消息实时阻断与分类系统", 计算机工程与设计, p356-359, 第 26 卷(第二期), 2005
- [6] Donghui Wang, Miaoliang Zhu, **Zheng Wu**, "Real-time blocking system for MMS", in *Proc. SPIE Asia-Pacific Optical and Wireless Communications*, vol. 5283, p. 67-71, China, Nov., 2003
- [7] 吴铮, 朱淼良, 龚建军, 王东辉, "基于视频分割与皮肤检测的视频敏感内容检测技术", 第十届中国人工智能学术年会, 北京, Nov. 2003

致谢

本文是在导师朱淼良教授、项目负责人钱徽副教授的悉心指导和严格要求下完成的。朱老师崇高的敬业精神、严谨求实的治学态度、开阔的思路以及深厚的人情味都给我留下了深刻的印象，并将使我终生受益。自我本科二年级认识朱老师以来，朱老师给了我许多宝贵的建议和启发，使我克服了课题研究中的重重困难，从一个对图像处理领域完全生疏的门外汉成长起来，让我更加的自信。在此，衷心地感谢朱老师给予我的帮助、鼓励和关心。同时也感谢钱徽师兄在本论文撰写过程中的帮助和指点，我们的合作一直非常愉快。课题研究中遇到的各种困难，无论是理论上还是实际程序设计上的，我们都一起努力解决。所以，本论文能够得以完成，也有钱师兄的很多贡献。如果本文有任何疏漏和不足的地方，是我个人的责任。

感谢王东辉副教授、袁昕副教授、张其前博士、张涣成博士、以及已经毕业的师兄龚建军、陈峰、桂本桦、徐少君和实验室同事陈沈轶在学习和生活中给予我的热心帮助和大力支持。没有你们的工作，我们的实验室不可能取得今天的成就，也无法形成如今团结、进取的良好学术氛围。

衷心感谢我的父母和家人多年来对我的关心、支持和鼓励，完全解除了我读书期间经济上的后顾之忧。

感谢当年同届混合班的同学，和你们一起奋斗的日子至今难以忘怀。

感谢所有曾帮助过我的老师、同学和朋友。

光阴似箭，在浙江大学六年多的学习生活马上就要结束了。在这些年中，我学到了很多的东西，丰富了知识，提高了技能，拓宽了思路，综合素质得到了显著提高。这一切都得益于母校浙大的培养。感谢学校各级领导、学院各位老师对我的亲切关怀，我以曾为浙大学子而自豪。

2006-2-15 于求是园