

SHANGHAI JIAO TONG UNIVERSITY

**学士学位论文**

## BACHELOR’S THESIS



论文题目： 数控冲床的排料算法优化设计

学生姓名: 王 昕 宇

学生学号: 5140219395

专 业: 电气工程及其自动化

指导教师: 唐 厚 君

学院(系): 电子信息与电气工程学院

**数控冲床的排料算法优化设计**

摘要

在数控冲床排料过程中，图像匹配技术在某些特殊板材的排料中有重要意义。但是长期以来，由于缺乏一种效率与准确性兼具的图像匹配算法，该技术一直没能大范围应用。本文提出了一种基于积分图和T检验法的倒金字塔图像匹配算法。该算法能在不降低匹配准确性的情况下提高效率。整体算法是一个“先粗后精”的搜索过程，核心思想如下：

首先，使用倒金字塔算法“粗”搜索。该算法利用了积分图能够快速计算任意矩形面积的特点，将面积特征作为一种搜索过滤标准，在庞大的搜索空间中，快速筛选出与模板相近的子图。

接着，由于面积特征在多次使用后会出现独立性下降、筛选效果不佳的情况，本文又提出了T检验法筛选方案。该方案根据已知的面积特征结果，通过假设检验的方式，进一步减少了搜索子图的数量。

最后，由于“粗”搜索筛选掉大部分子图，剩余的搜索子图紧紧围绕在目标附近，具有良好的梯度相关特性，所以采用了基于梯度上升的LSS算法，精准定位搜索目标。

本文提出的图像匹配方法在速度上有了较大提升，与常用的FFT算法相比，在只有一个目标的情况下，有近一倍的效率提升。该算法另一个巨大优势是可以更换相似性度量方式，有利于算法的推广。

关键词：图像匹配，倒金字塔搜索，积分图，T检验法，LSS

**CNC PUNCH NESTING ALGORITHM OPTIMIZATION**

**ABSTRACT**

In the process of CNC punching, image matching technology has important significance in the nesting of some special plates. However, for a long time, due to the lack of an image matching algorithm with both efficiency and accuracy, the technology has not been widely applied. This paper presents an inverted pyramid image matching algorithm based on integral graph and T test. This algorithm can improve efficiency without reducing the accuracy of matching. The overall algorithm has a "coarse and fine" search process. The core ideas are as follows:

First of all, we use the inverted Pyramid algorithm for "coarse" search. The algorithm uses the integral graph to quickly calculate the characteristics of arbitrary rectangle area. The area feature is used as a search filtering standard. In the large search space, the sub graph which is close to the template is quickly screened.

Then, because of the decrease in independence and poor screening effect after the use of area features, this paper proposes a T test method. According to the known area characteristic results, the scheme further reduces the number of search subgraphs by means of hypothesis testing.

Finally, as the "rough" search filters out most of the subgraphs, the remaining search subgraphs are closely around the target and have good gradient correlation characteristics, so the LSS algorithm based on gradient rise is adopted to accurately locate the search target.

The image matching method proposed in this paper has greatly improved the matching speed. Compared with the common FFT algorithm, it has nearly twice the efficiency improvement in the case of only one target. Another great advantage of the algorithm is that the similarity measure can be replaced, which is conducive to the generalization of the algorithm.

**Key words:** image matching, Inverted pyramid search, Integration graph, T test, LSS

目 录

[第一章 绪论 1](#_Toc516099935)

[1.1 课题背景及选题意义 1](#_Toc516099936)

[1.1.1 数控冲床简介 1](#_Toc516099937)

[1.1.2 课题的研究内容及意义 3](#_Toc516099938)

[1.2国内外研究现状 4](#_Toc516099939)

[1.2.1 基于灰度信息的匹配算法 4](#_Toc516099940)

[1.2.2 基于特征的匹配算法 4](#_Toc516099941)

[1.3 算法性能度量及研究目标 5](#_Toc516099942)

[1.4 各章节内容安排 5](#_Toc516099943)

[第二章 传统图像匹配算法综述 6](#_Toc516099944)

[2.1基于灰度的匹配算法 6](#_Toc516099945)

[2.1.1 经典模板匹配 6](#_Toc516099946)

[2.1.2 基于FFT的模板匹配 9](#_Toc516099947)

[2.1.3 SSDA算法及改进 12](#_Toc516099948)

[2.1.4 图像金字塔 13](#_Toc516099949)

[2.1.5 启发式搜索 14](#_Toc516099950)

[2.2基于特征的匹配算法 15](#_Toc516099951)

[2.2.1 区域特征 16](#_Toc516099952)

[2.2.2 边缘特征 17](#_Toc516099953)

[2.2.3 点特征 18](#_Toc516099954)

[2.3本章小结 19](#_Toc516099955)

[第三章 基于积分图的倒金字塔搜索 20](#_Toc516099956)

[3.1基于面积特征的搜索区域筛选 20](#_Toc516099957)

[3.1.1 SSDA算法回顾及新算法提出 20](#_Toc516099958)

[3.1.2 基于面积特征的筛选 20](#_Toc516099959)

[3.1.3 积分图的计算 22](#_Toc516099960)

[3.1.4 采样矩形的选取 23](#_Toc516099961)

[3.1.5 灰度图像的处理 23](#_Toc516099962)

[3.2 T检验法 24](#_Toc516099963)

[3.2.1 假设检验 24](#_Toc516099964)

[3.2.2 T检验法原理 25](#_Toc516099965)

[3.2.3 T检验法应用 25](#_Toc516099966)

[3.3 精确定位 26](#_Toc516099967)

[3.3.1 聚类 26](#_Toc516099968)

[3.3.2 相似性度量 27](#_Toc516099969)

[3.3.3 搜索方法 27](#_Toc516099970)

[3.3.4 匹配修正 28](#_Toc516099971)

[3.4 算法总结 28](#_Toc516099972)

[3.4.1 算法流程图 28](#_Toc516099973)

[3.4.2 算法复杂度分析 29](#_Toc516099974)

[3.5 本章小结 29](#_Toc516099975)

[第四章 算法应用效果及分析 31](#_Toc516099976)

[4.1 算法分步结果 31](#_Toc516099977)

[4.1.1 倒金字塔搜索 31](#_Toc516099978)

[4.1.2 精准定位 32](#_Toc516099979)

[4.2 与FFT算法对比 33](#_Toc516099980)

[4.3 本章小结 34](#_Toc516099981)

[第五章 结论与展望 35](#_Toc516099982)

[5.1 全文工作总结 35](#_Toc516099983)

[5.2 后续工作展望 35](#_Toc516099984)

[参考文献 36](#_Toc516099985)

[谢辞 38](#_Toc516099986)

第一章 **绪论**

近年来，随着计算机视觉技术的迅速发展，其在各个领域的应用逐渐增加。此外，工业对生产质量和效率的要求也越来越高。在此背景下，计算机视觉技术开始在工业领域大范围应用。计算机视觉就是用机器来代替人眼进行观察和判断，常用于大批量生产中的过程控制和质量检测等。相比于人眼观察，其检测速度更快，误差更小，因此能显著提高生产效率。数控冲床就是一个计算机视觉的重要应用场景。本文深入研究了计算机视觉中的图像匹配技术，并应用该技术优化数控冲床的排料，提高生产效率。

1.1 课题背景及选题意义

1.1.1 数控冲床简介

随着计算机技术的发展和自动控制理论的成熟，自动化设备在生产中占有了越来越多的比重，其中数控冲床就是一种自动化的生产设备。它将冲头产生巨大的压力作用在金属板材上，使金属发生形变直至边缘完全断裂，进而得到各种形状的工件。图1-1为待加工的金属板材，图1-2为冲压出来的工件。

**图 1-1 金属板材 图1-2 冲压工件**

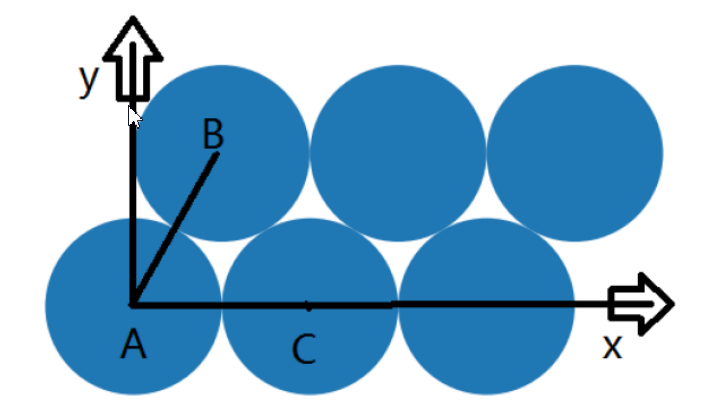
冲床的核心原理为将电机的圆周运动转化为滑块的直线运动，通过滑块带动模具向下冲压板材，使其形变。模具的形状和期望加工的工件形状相同，上下两个为一组模具，将板材置于中间，挤压出指定形状。图1-3为冲床外观图。

  
**图1-3 冲床**

冲床的发展经历了从手动到自动的巨大飞跃。在此过程中，伺服传动系统的成熟和自动排料算法的发展功不可没。在手动冲床的时代，需要工人用手把持板材，将其对准模具，然后操作冲床的冲头带动模具下压，冲压出工件。伺服系统的功能就是替代人工送料，实现送料的自动化。通过伺服电机、传送带、螺杆等一系列传动机构，数控冲床可以控制板材在水平面x, y方向自由移动。伺服系统的应用能够明显提高板材移动的效率和稳定性，而且节省了大量人力成本，也避免了在冲床这种大型机械附近操作，可能带来的人身危险。

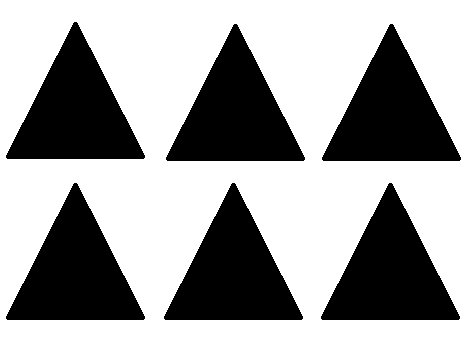
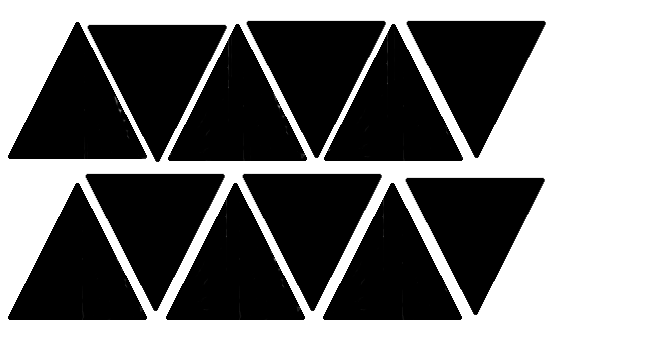
但是只能移动板材并不能实现完全的自动化生产，在人工操作中，首先人眼寻找板材上可以冲压工件的区域，然后大脑分析应该怎么排布工件才能在有限的板材上冲压出最多的工件，最后是手动操作板材进行冲压。为了实现这些功能，数控冲床首先需要一个摄像头来观察板材，其次需要一个工业处理器来最优化的排布工件，也称为排料。接着，处理器将排料的结果输出给伺服系统，最后伺服系统控制板料和冲头完成工件的冲压。其中，伺服系统较为成熟，相机和工控机硬件也没有瓶颈，所以数控冲床的关键点在于软件的开发。

软件实际上替代了人脑的分析过程，即观察摄像头得到的板材图片，确定排料方案，最后得到可以排放工件位置的坐标并发送给伺服系统。这个过程涉及两个技术，一个是图像处理，即从摄像头拍到照片中得到板材的位置和形状，并确定可以排料的有效区域。另一个是自动排料，在这个过程中，我们将已知的工件尽可能多的排布在有效区域内。目前大部分的排料使用的是套料模板，所谓套料模板就是将工件组成一个小的集合，这个集合是我们已知的最优方案，然后用这种方案逐步的放到更大的有效区域内，再次追求最优解。也就是说这是一种由局部最优扩展到全局最优的方案。下图1-4就是一种圆形的套料模板。按照这种模板排布圆形工件能够实现同样面积摆放最多工件。



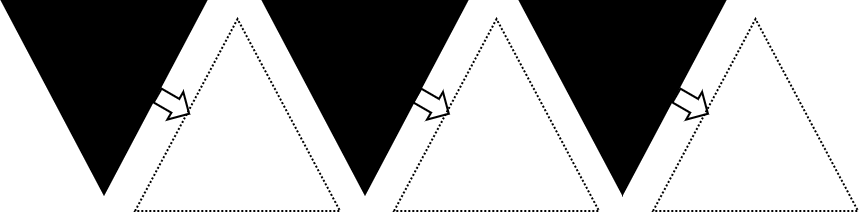
**图 1-4 圆形工件套料模板**

上述的自动排料方法能够解决大部分简单工件的优化排料，但是对于较为复杂或者形态特殊的工件效果却不好，例如图1-5三角形工件的排布。对比图1-6，显然后者排料方案更佳。

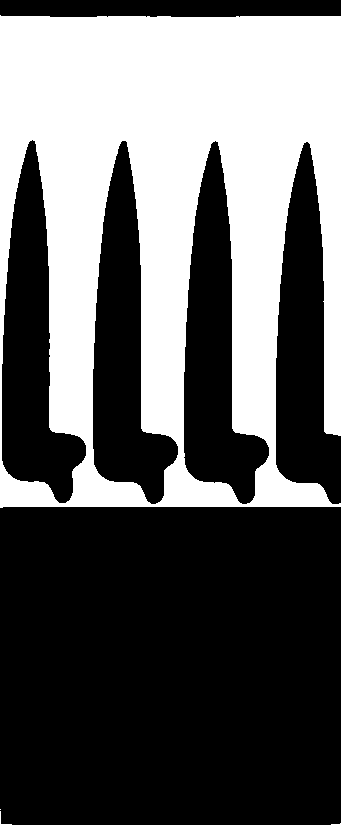
**图 1-5 普通三角排料 图 1-6 合理的三角排料**

但是要冲压出图1-6的排料效果并不容易。首先，模具的方向是不变的，即不能通过旋转模具来达到冲压相反方向工件的目的，只能将板材冲压两遍。第一遍可以冲压出图1-5的效果，第二遍为了冲压出相反方向的工件，则要翻转板材送入数控冲床。此时为了刚好在相邻三角形空隙中冲压出下一个工件，在排料算法中一般采取的策略是，先找到原来已经被冲压过工件的孔洞，根据孔洞确定空隙上的排料位置，如图1-7所示。



**图 1-7 翻转后的板材上确定排料位置**

要想通过孔洞确定排料位置，分为两个步骤。首先要得到板材上孔洞的位置，然后才可以通过模板形状大小等特征确定空隙中的排料位置。后者较为简单，技术成熟，而前者精确获取孔洞位置的算法还有待研究，本论文的主要研究内容就是如何得到孔洞位置。采取的策略为图像匹配，即用已知的工件模板在搜索图像中寻找孔洞位置。其中工件模板就是一个工件图如图1-8，因为孔洞是由板材冲压过工件形成的，所以工件模板的形状大小和孔洞完全相同，以下统称为模板。而搜索图像是摄像头拍到的经过第一遍冲压，包含了多个孔洞的图片如图1-9。

**图1-8 工件模板**   **图1-9 待匹配的搜索图像**

1.1.2 课题的研究内容及意义

为了优化上一小节提到的工件排料，本文将主要研究图像匹配技术。实际上图像匹配是机器视觉的基础算法之一，在多个领域有深远影响。上世纪70年代，美军首次在潘兴II地对地中程导弹[1]中使用了图像匹配技术。在末端制导段，通过将已知目标图片与导弹当前拍摄图片进行匹配，精准定位目标位置，实现了30m级的命中精度。此后，各国相继对图像匹配技术投入了大量人力物力，而图像匹配也因此得到了快速发展，并在数字图像处理中广泛应用。

所谓图像匹配技术，就是识别不同场景，不同时间，不同传感器，不同视角等获得的

两张或多张图片的共同物体，实现在空间位置上的对应。或者是用已知模板在另一幅图片中寻找相应的模式，已知模式一般称为模板图像，如图1-8。而待检测的图像一般称为搜索图像，如图1-9。图像匹配在导弹制导[1]、医疗图像处理、三维场景重构以及目标的识别与跟踪方面都有着广泛应用。虽然不同的应用场景有具体的应用方案，但是他们的理论算法基本相同，所以图像匹配的理论突破很容易带来多种技术的共同进步。而图像匹配的理论算法要想有所突破，需要各个基础领域的共同努力，如数值计算、数字信号处理、人工智能等，另外基础数学领域更是图像处理算法的重要支撑。

随着计算机技术以摩尔定律的快速发展，计算机运算速度得到了巨大的提升，目前的已经能够满足绝大多数行业的需求。但在图像处理中，由于计算数据量大，实时性要求高，所以减少匹配算法复杂度依然是主要的研究方向。另外，由于拍摄时间、角度、光照等影响，模板图像和搜索图像存在一定程度的灰度失真及几何畸变，这些都给匹配的准确度带来困难。所以合理的模板和目标之间相似性度量方法也是研究人员重要的研究方向。如果能够提出一套匹配效率高、鲁棒性强的匹配方案，那么图像匹配技术将在导航制导、模式识别等领域发挥更大更积极的作用。因此，图像匹配技术的研究具有十分重要的应用价值和理论意义。

1.2国内外研究现状

图像匹配技术第一次应用在美国军方，并取得了不错的效果，但当时匹配技术还处在起步阶段，算法较为粗糙，效率也比较低。但随着数十年的发展，图像匹配算法得到了极大地丰富，各种高效准确的方法被陆续提出，根据算法匹配的原理可以大致分为如下两类：

1.2.1 基于灰度信息的匹配算法

基于灰度的匹配算法[2]也被称为相关法、模板匹配法。此方法原理易懂，实现简单，是人们最早开始研究的匹配算法。它以模板和搜索图像的灰度信息作为相似性的度量，对搜索图像每个像素点进行相似性的最优搜索，来确定搜索图像中是否存在和模板相匹配的目标并确定目标位置。其中“归一化互相关[3]”是一种典型的基于灰度的匹配算法。该算法适应能力强且易于掌握，使用方便。但这种方法需要处理的数据量大，并对图像的畸变较为敏感。

所以近年来，人们的研究方向主要集中在更快的搜索算法和更稳定的相似性度量方案。搜索算法中性能较好的有图形金字塔算法[4]、惯性相似性检测算法[5]（SSDA）、快速傅里叶变换[6]等算法。其中快速傅里叶变换算法凭借速度优势得到了广大工程人员的认可，著名开源软件OpenCV的图像匹配方案就是使用了该算法。不过上述算法主要以“归一化互相关”作为相似性的度量，其他的相似性度量方式也在不断发展，如平方绝对差匹配[4]、投影匹配[7]和信息熵等，它们也都各自所长。

随着人工智能的发展，各种启发式搜索算法呈现井喷式发展，如遗传算法[8]、帝国竞争算法[9]、灰狼算法[10]等。这些方法通过启发式搜索，极大地缩短了匹配的时间，但是容易陷入局部最优解，稳定性还有待提高。

1.2.2 基于特征的匹配算法

特征匹配是指通过提取模板图像和搜索图像的特征，并且对提取出来的图像特征进行参数化描述，然后对比两图像的参数进行匹配。所以使用特征匹配算法分为两个步骤，首先提取特征使其参数化，然后对特征参数进行对比匹配。常用的特征有区域特征、边缘特征和点特征。其中区域特征包括面积、矩等，边缘特征包括轮廓、几何基元，点特征一般是角点。特征提取的方法往往需要大量的数学知识及图像形态学知识，运算量较大，而且没有统一的范式，需要根据应用场景具体情况具体分析。但是特征提取算法也有它无法比拟的优势，那就是仿射不变性和尺度不变性，对旋转、平移、缩放都有很好的适应性，抗

干扰能力极强。

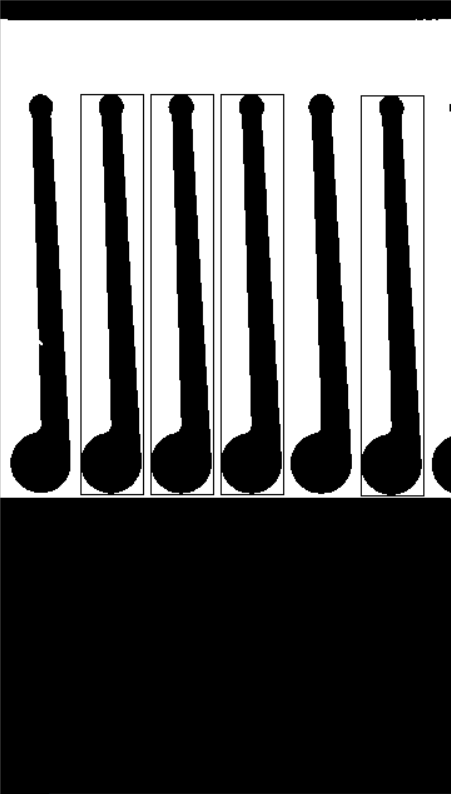
目前特征提取与匹配的常用方法有：小波变换法[11]、边界特征法[12]、形状不变矩法[13]等。但是应用最广泛的还是基于特征点的匹配，近年来，学者们关于特征点提取也提出了好多笋子，如Harris算子[14]、SUSAN算子[15]、SIFT算子[16]、SURF算子[17]。其中以SIFT和其改进算法SURF应用最为广泛。

1.3 算法性能度量及研究目标

各种各样的图形匹配算法，需要一些标准来判断其性能的优劣。常见的评价指标一般有以下三个：

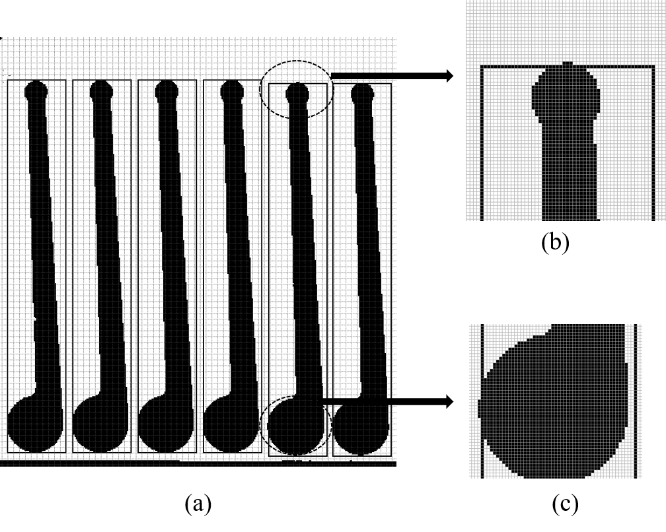
（1）匹配稳定性，可以用能够正确匹配的图像数占总目标数的比例来度量

其中*m* 为正确匹配的次数，*N*为总的目标数。如图1-10所示，正确匹配数为4个，总目标6个，显然稳定性不佳。



**图1-10 匹配稳定性**

（2）匹配准确性，如图1-11中(b)(c)所示，算法在搜索图像中检测的目标位置和目标的实际位置之间可能存在误差，该误差越小，则匹配准确性越高。



**图1-11 匹配准确性**

（3）匹配实时性，算法从读取图片数据到目标坐标输出之间所需要的时间定义为匹配时间，该时间越短则匹配速度越快。

当然在不同的环境下会有不同的具体要求，但是核心的准则在于能否满足应用场景的需求，能否优于之前的算法。在本课题数控冲床排料优化中，显然匹配成功率要保证达到100%，这样的算法才能够应用到生产中。匹配准确性也要尽可能的高，误差尽量不超过2个像素，毕竟排料过程中工件之间的安全距离只有2个像素。匹配速度也要尽可能的快，一方面是冲床使用的是控制器性能不高，可能无法在指定时间完成计算；另一方面，匹配所用时间越短，其他排料处理的时间就可以越长，可以实现整体性能的改良。

1.4 各章节内容安排

第一章，绪论。首先介绍了选题背景及图像匹配技术的研究意义，接着简单的介绍了当前图像匹配技术在国内外发展情况，最后明确了课题的任务及要达到的性能指标。

第二章，传统图像匹配算法综述。详细介绍了目前图像匹配的主流算法。其中基于灰度匹配的算法有FFT相关匹配、序贯相似性检测、图像金字塔和启发式算法等。在基于特征的匹配中，本章根据特征的分类介绍了区域特征、边缘特征和点特征及其各自对应的匹配算法。

第三章，基于积分图的倒金字塔搜索。本章介绍了新算法“由粗到细”的搜索过程。首先，通过面积特征快速筛选搜索子图，接着使用T检验法对面积特征进行补充。最后，在得到与目标十分接近的子图后，利用LSS算法精细定位最佳匹配位置。

第四章，算法应用效果。本章首先罗列了倒金字塔算法的每个步骤得到的阶段性结果，并将结果与之前的理论分析相比较。接着测试整个算法的时间性能，并与FFT算法相对比。

第五章，结论与展望。本章总结了整个课题的研究工作，并对得到的结果进行分析，提出了下一阶段的研究方向。

第二章 传统图像匹配算法综述

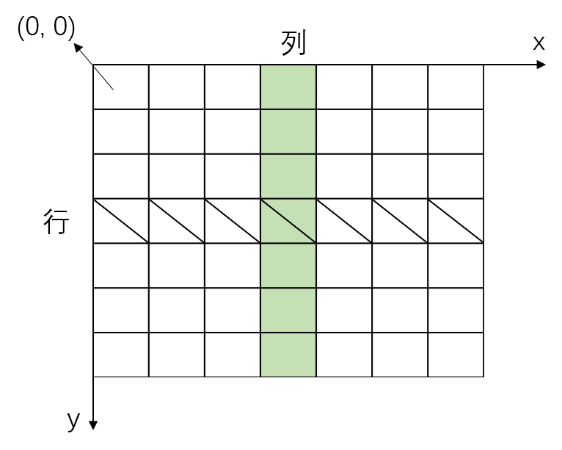
经过几十年的发展，专家学者针对图像匹配提出了种类繁多的算法，大致可以分为两类，即基于灰度的匹配算法和基于特征的匹配算法。前者原理易懂，实现简单，但对图像畸变敏感，而后者算法复杂，计算量大，但往往拥有尺度不变性和仿射不变性，适用范围更广。下面，将详细介绍这两种策略对应的算法原理及实现方式。

2.1基于灰度的匹配算法

基于灰度的匹配算法也可以称为基于统计的算法，他通过统计模板与搜索图像之间的差异大小来判断匹配的相似性。

2.1.1 经典模板匹配

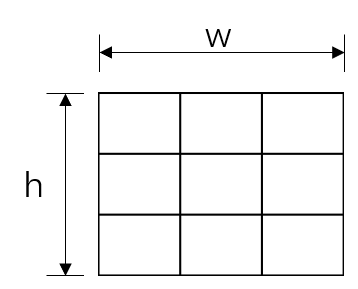
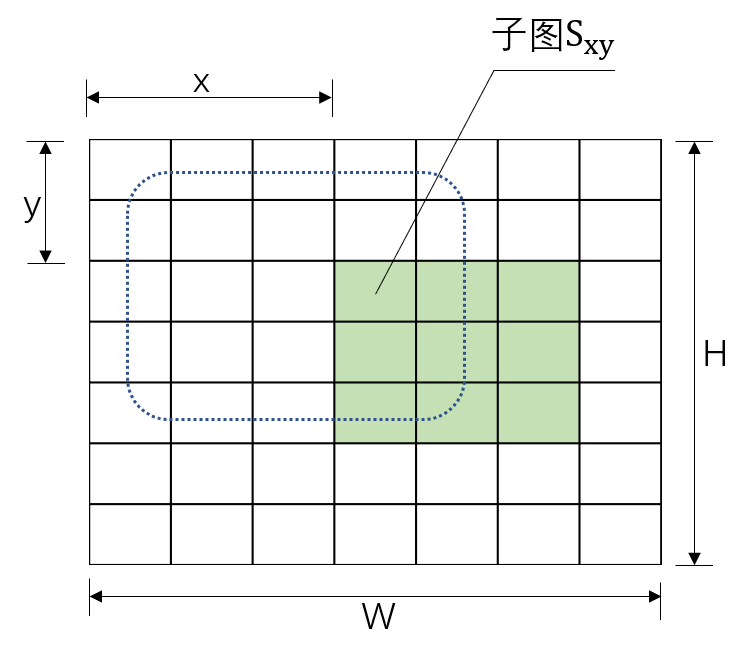
在介绍模板匹配算法之前，本文先介绍算法中可能用到的术语及标识。如图2-1所示，在计算机中，图像是以点阵的形式存储，左上角点坐标一般定义为（0, 0），向右为*X*轴正方向，向下为*Y*轴正方向。



**图2-1 图像基本标识**

2.1.1.1 模板匹配算法的流程

模板匹配算法中，模板是已知的，为一幅较小的图片，而搜索图像是一幅较大的图片，它可能包含着和模板相似的待匹配目标。模板匹配就是确定搜索图像中是否存在目标并确定其位置。本文中假设模板*T*的大小为，如图2-2。搜索图像的*S*的大小为，如图2-3。在匹配的过程中，模板*T*在搜索图像*S*上平移，搜索窗口和模板大小相同，该窗口覆盖的子图记为，其中*x*，*y*是子图左上角顶点在搜索图像*S*中的坐标。显然*x*，*y*的搜索范围是：，，即图2-3虚线框内的区域。

**图2-2模板图*T*** **图2-3 搜索图像*S***

模板匹配的具体过程如下：

（1）模板在搜索图像的搜索范围内依次从左到右，从上到下的移动

（2）计算模板*T*在每个位置与该位置对应的子图之间的相似性大小。最简单的相似性度量为平均绝对差算法（Mean Absolute Differences，简称MAD算法）。MAD算法的度量公式为式(2-1)。

若越小，说明该子图与模板越相似。

（3）在遍历过搜索区域全部个点后，取其中最小的所对应的位置坐标（*x, y*）作为最终匹配结果。

从上面模板匹配的过程中，我们可以得知算法的核心分为两部分，一是搜索方法，上述过程中使用的搜索方案是完全遍历，即将有效搜索区域内所有子图的相似性大小都计算了一遍，然后从中找出相似性最大点。这样的搜索方式所需的计算量是巨大的，其算法复杂度为。显然我们也可以每隔一个点算一次，这样效率提高一倍，当然也随之产生了误差，在后面的章节中本文将详细介绍各种精巧高效的搜索策略。第二个核心就是相似性度量方法，上面用的是MAD算法，该方法思路简单容易理解，并且易于编程，匹配精度也较高，但是计算量很大，容易受噪声影响，尤其是模板与搜索图像整体亮度不同的时候，计算出来的结果会较大，不能准确反应相似性。接下来本文将详细介绍几种常用的相似性度量算法。

2.1.1.2 几种常见相似性度量算法及其对比

(1)平均绝对差算法（Mean Absolute Differences，简称MAD算法）,如式(2-1)所示。

(2)零均值绝对差值和(Zero mean absolute difference sum,简称ZAD)，其相似性度量公式如式(2-2)：

其中和分别为子图和模板图的平均灰度值，如式(2-3)。该算法消除了图像灰度值变化的影响，在搜索图像和模板平均亮度不同时也会有较好的匹配效果。

(3)误差平方和算法(Sum of Squared Differences，简称SSD算法)，其相关性度量公式如式(2-4)：

此算法与MAD算法较为相似，MAD使用L1距离来度量差异，而SSD算法使用L2的欧氏距离来度量子图和模板的差异。

(4)零均值差的平方和算法(Zero mean Squared Difference sum，简称ZSD)，其相关性度量公式如式(2-5)：

该方法与ZAD原理相似，都可以消除整体灰度变化带来的匹配误差。

(5)归一化互相关算法(Normalized Cross Correlation，简称NCC算法)，在实际应用中，NCC算法最为常见，但它的度量公式并不像上述几种算法直观，本文将通过SSD算法推导出NCC算法。首先将式(2-4)展开得：

|  |  |  |
| --- | --- | --- |
|  |  |  |

展开后分为三个部分，分别定义为：

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

我们称为搜索图像中子图的能量，它的大小与匹配点的位置有关，随(*x*, *y*)的变化而逐渐变化。为模板图像的能量，它与搜索位置*(x, y*)无关，是一个常数。是搜索子图与模板图之间的互相关，它随搜索位置(*x, y*)的变化而变化，当模板T与子图相匹配的时候，互相关的值最大。

在近似不变的情况下，可以用来度量相似性，即取得最大值的时候，我们认为匹配效果最好。但如此近似会有一些误差，所以一般采用归一化的形式来定义相似性，如式(2-10)

进一步推广ZAD和ZSD算法的经验，对式(2-10)进行零均值处理，来消除整体灰度变化的影响。最终得到NCC算法的度量公式(2-11)。

对于图像质量较好，失真较小的情况，算法MAD和SSD都会有较好的效果，并且实现简单，运算量较小。而对于灰度变化较大的模板和搜索图像，算法ZAD和ZSD效果较好，但由于计算平均灰度时需要提前遍历一遍图像，所以相应计算复杂度上升。而NCC算法抗干扰能力最强，匹配的效果也最好，但公式复杂，计算量大，效率较低。

2.1.2 基于FFT的模板匹配

上一小节的经典匹配算法，虽然大多数情况下准确性较好，但是时间复杂度却高达。这意味着对于的搜索图像和的模板，想要在普通PC上运行该算法，一次匹配的时间大概需要十几分钟。而实际应用中，一般视频是25帧/s，冲床排料中拍照至少也要1帧/s。所以提高运算效率是模板匹配能否应用的关键。本小节将要介绍一种数学方法来提高计算效率。

经典的模板匹配算法计算复杂的主要原因在于需要对每个搜索位置都计算一次公式(2-11)，忽略了相邻搜索点之间相似性的内在联系。为了降低算法复杂度，将式(2-11)变换为公式(2-12)的形式。

式(2-12)中与坐标位置有关的三项分别记为：

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

2.1.2.1 快速傅里叶变换

式(2-13)中是子图与模板之间的互相关，也可以看作模板与子图之间的二维卷积，根据傅里叶变换的性质：时域的卷积可以由对应频域的乘积经过傅里叶反变换得到，即如式(2-16)所示。

其中是搜索图像的傅里叶变换，是模板图像的傅里叶变换共轭，由于模板图像相较搜索图像较小，所以一般将模板图像*T*从用0扩充至。为了进一步简化计算，可以构造式(2-17)的复图像。

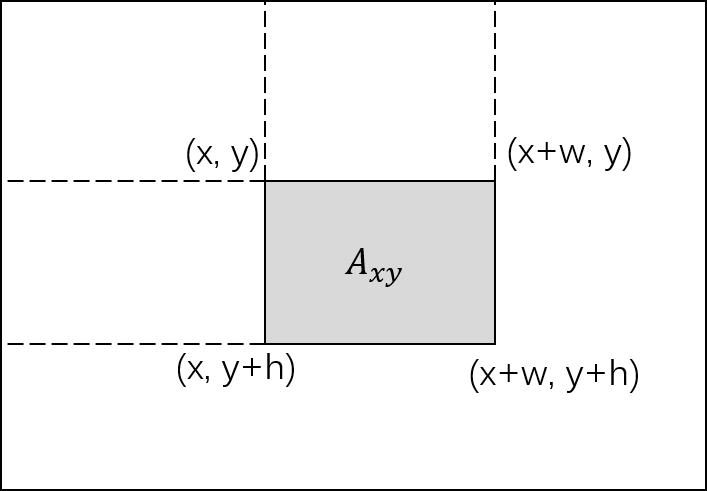
根据二维傅里叶变换的性质可以得知：

通过上式（2-17）、（2-18）、（2-19），计算一次*Z(x, y)*的傅里叶变换*Z(u, v)*就可以得到*S(u, v)*和*T(u, v)*。然后再利用式（2-16），就可以一次性将每个搜索位置的都计算出来。

2.1.2.2 积分图

剩余两项，分别是搜索子图的灰度值和灰度平方的积分，利用积分图[18]可以快速算出。积分图的概念最早由Viola提出，并应用在Harr特征的计算中。对任意图像*A*，其积分图*I*在 处的灰度值为图像*A*中该点左上方所有像素点灰度值之和，如式(2-20)所示。

积分图最大的好处在于能够快速计算出原图像*A*中任意方形子图的像素灰度的积分。如图2-4和式(2-21)所示。



**图2-4 积分图**

所以计算搜索图像S的灰度积分图及灰度平方积分图后，，即可利用式（2-21）快速算出。

最终，使用二维FFT和积分图的模板匹配算法的具体流程为：

（1）使用FFT计算出；

（2）计算搜索图像的灰度积分图和灰度平方积分图，据此计算，；

（3）根据式（2-12）计算每个搜索点的相似性；

（4）在所有搜索点中寻找相似性最大的点作为最终匹配结果。

2.1.2.3 复杂度分析

当以遍历的方式计算式（2-11）的时候，主要的计算量在乘法运算上。对每个位置使用一次式（2-11），计算量大概为，总共需要遍历的搜索点数为，所以经典匹配的运算量大概为：

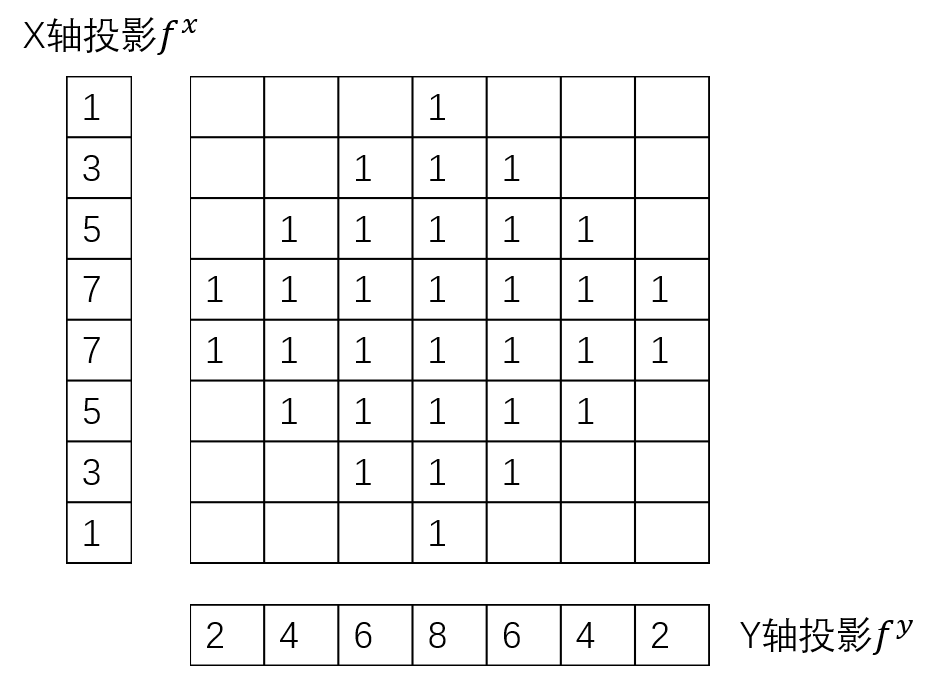
从式（2-22）来看，经典算法的运算量大致与搜索图像和模板图像大小的乘积成正比。

对于本节介绍的FFT算法，积分图主要涉及的是加分运算，基本可以忽略，乘法运算集中在FFT和IFFT，典型的快速傅里叶变换或其反变换需要乘法次数为。所以在本算法中乘法运算量大致为：

由上式可以看出，基于FFT的方法计算复杂度与模板的大小无关。对于1000\*1000的搜索图像来说，根据经验，当模板大小超过13\*13的时候，基于FFT的方法的速度就会更快。所以相比于经典方法，模板图像越大，本方法的优势越大。

2.1.2.4 基于一维FFT匹配算法

上面的相似性度量方案都是基于二维的，子图中每个点都参与计算，为了简化计算，我们也可以将二维信息压缩为一维[19]。具体方法如图2-5所示，将图像横向或纵向投影成一维的数据，这里的投影指的是把一列数据或者一行数据进行累加，用累加和来表示该列或行的信息。



**图2-5 投影图**

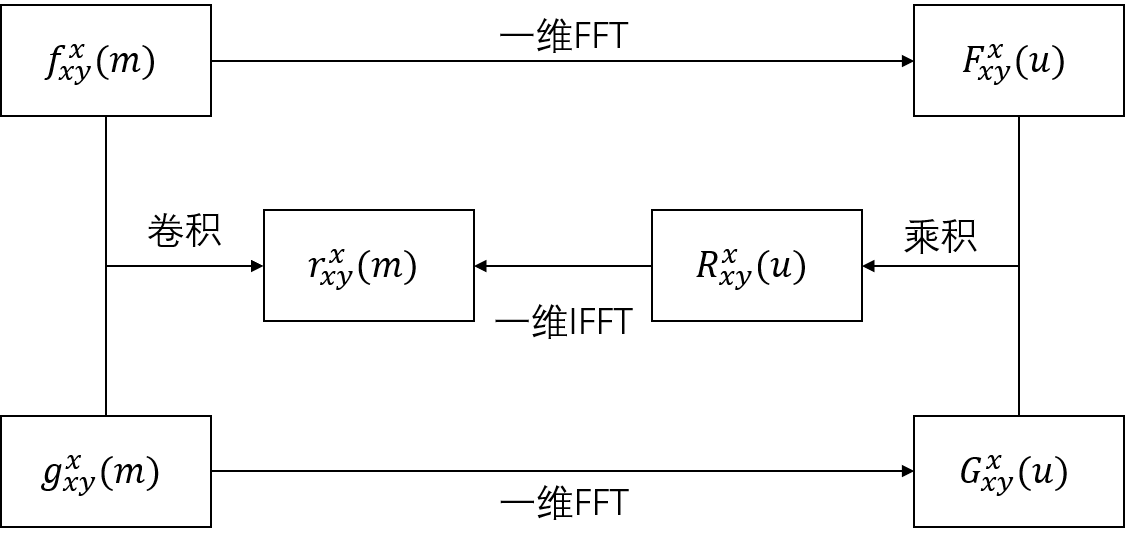
利用投影将数据压缩成一维后，我们用子图和模板*T*的一维投影进行匹配相关性计算*，*计算公式如式（2-24）。

其中表示子图在*X*轴方向的投影，为模板在*X*轴方向的投影。显然，上式使用了*X*轴方向投影进行匹配，同理，也可以使用*Y*轴投影数据，这里不再赘述。为了能够使用FFT快速计算式（2-24），我们将其变换成如下形式：

对比式（2-24），我们可以得到：

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

其中可以迭代得到，为常数，可以由一维的FFT计算得到。计算过程示意图如下2-6。



**图2-6 一维FFT示意图**

从图2-6可以看出，一维FFT算法与二维相同，都是将时域卷积变换成频域乘积来简化计算。相比于二维FFT算法，该方法显然具有计算量小速度快的特点，但是由于二维信息压缩至一维，信息丢失严重，匹配的准确度不高。所以此算法适用于，对精确度要求不高但实时性十分重要的场合。

2.1.3 SSDA算法及改进

序贯相似性检测（Sequential Similarity Detection Algorithms，简称SSDA）最早在1972年由Barnea 和 Silverman共同提出。实践表明，SSDA算法的计算效率较FFT相关算法要高一至两个数量级。其算法的精髓在于虽然也要遍历全部搜索点，但是不同搜索子图并不充分计算相似性，在相似性低的子图快速略过，对相似性高的子图详细计算，使得计算过程对匹配目标更有针对性。

SSDA算法流程如下：

(1)定义绝对误差：

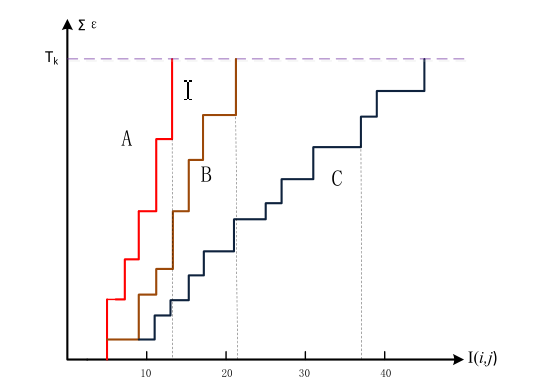
该定义与ZAD的定义式（2-2）较相似，但是去掉了两次积分，表示单个像素点间的误差，式中和可以参照式(2-3)；

(2)取固定阈值；

(3)在子图中随机选取像素点，计算它与模板中相对应点之间的绝对误差，继续选取不重复的像素点，计算误差，将这些误差累加，直至累加和超过阈值，记录此时已经选取的像素点数量*r*，定义相似性度量：

由式（2-29）可知，代表着绝对误差和大于阈值的最小累加次数；

(4)把最大的点作为最佳匹配点，因为该点需要选取最多的像素点，误差才能达到阈值，说明该点对应的子图与模板相差很小。图2-7给出A,B,C三个点的累计误差增长曲线。其中A,B两点累计误差增长很快，说明该点对应的子图与模板偏差较大，每次选取的像素点绝对误差都很大，所以很快就累加超过了阈值。而点C的增长缓慢，说明和模板相近，有可能是匹配点。



**图2-7 累计误差增长曲线**

在SSDA算法的计算过程中，和模板图相似性很差的子图选取较少的像素点就可以停止计算直接排除，而和模板图相近的子图，则需要选取更多的像素点，详细计算匹配相似性，因此计算更具有针对性，运算速度显著提高。

在此基础上，人们也提出了一种自适应阈值算法。此算法特点在于能够更新阈值的大小。首先将阈值设置为一个较大的值，此时会出现，与模板相似性较高的子图遍历其全部个像素点后，累计误差依然小于阈值。此时 将赋值给即：

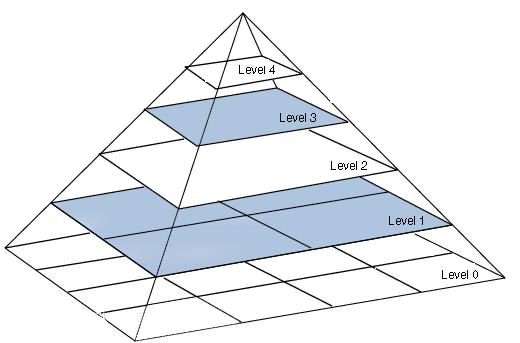
记录此时坐标。遍历所有子图过程中遇到上述情况都更新阈值，并记录下坐标。遍历完毕即可获得最佳匹配点和相应子图位置。

该方法十分依赖最佳匹配点出现的位置，如果最佳匹配点很早出现，那么阈值就会提前变得很小，后面的搜索子图也会很快被淘汰，但是若其出现很晚，那么大部分搜索子图按照较大阈值来淘汰，反而速度变慢。

2.1.4 图像金字塔

无论是FFT算法还是SSDA算法，都需要计算每个搜索子图的匹配相关性，而本节介绍的图像金字塔方案则是一个由粗到精的细化过程，不需要计算每一个搜索子图。

图像金字塔的提出是基于经典的遍历搜索方案，该方案的复杂度为，当能够减小搜索图像或模板图像的大小的时候，复杂度会快速下降。所以自然可以想到对原图像和模板进行抽样。当横纵坐标均二次抽样的时候，搜索图像和模板大小均变为1/4，算法复杂度变成原来1/16。但是直接二次抽样，可能导致锯齿效应，并丢失重要信息。所以一般会使用的均值滤波器来平滑图像，也就是说将范围内像素的平均灰度值作为二次抽样后一个像素点值。另外，一般来说图像多次缩小两倍比一次缩小超过两倍的效果更好。而多次缩小两倍的效果就是每次图像大小变为原来1/4，将得到的图像从大到小的依次向上堆叠，就形成了如图2-8的图像金字塔。



**图2-8 图像金字塔**

应用图像金字塔匹配的流程如下：

(1)将搜索图像和模板图像分别平滑滤波，然后二次抽样，多次执行这一步，得到图2-8所示图像金字塔

(2)对最顶层的图像进行经典的模板匹配，得到最佳匹配点位置

(3)将高层的匹配结果映射到下一层，即将匹配结果坐标乘2。考虑到映射得到的位置不准确，将下一层的搜索范围设定为匹配点周围小区域，记为ROI。

(4)遍历每层的ROI得到该层最佳匹配点，直至最底层结束。

在图像金字塔算法中，显然抽样次数越多，最顶层的图像越小，计算速度就会越快，但是，过多的抽样会导致图像分辨率降低，我们感兴趣的标识可能被忽略，导致无法识别。所以合适的层数至关重要，一般我们会取4层，但要具体情况具体分析。另外，被平滑抽样后，最上层的待识别物体的边缘会严重失真，层数越多这种情况越明显，所以在高层的时候匹配阈值就要相应改变，如NCC算法，在匹配相似性较低时，我们也可以认为找到了目标。

图像金字塔算法最大的优点在于运算速度快，以四层金字塔为例，经过三次平滑抽样，总的计算复杂度变为了原来的1/4096。但实际应用时也会有一些问题，如层数的选择，层数太少计算速度慢，而层数太多可能导致找不到目标。而且算法需要一个阈值来判断目标是否存在，顶层的阈值也较难选取。总之，图像金字塔算法是牺牲了稳定性来尽可能提升效率。

2.1.5 启发式搜索

随着人工智能的发展，各种启发式搜索算法也逐渐应用到图像匹配技术当中。所谓启发式搜索就是无需遍历全部搜索范围内子图，而是随机选取部分子图，计算其相似性后，推测最佳匹配点可能的方向，然后向该方向移动，直至到达最佳匹配位置。典型的启发式搜索算法有遗传算法、灰狼算法、帝国竞争算法等。本小节以较为简单的帝国竞争算法为例，详细介绍启发式搜索。

帝国竞争算法（imperialist competitive algorithm，简称ICA）是Atashpaz-Gargari和Lucas于2007年提出的一种基于帝国殖民竞争机制的迭代算法。ICA算法主要包括如下四个部分：

(1)初始帝国产生。我们在搜索区域内随机选取部分子图，每个子图作为一个国家，而每个子图和模板之间的相似性代表着国力的强弱。首先，设随机产生国家，选取其中相似性靠前的个国家视为帝国主义国家，剩下的个国家视为殖民地。其次，根据每个帝国相似性的大小分配殖民地的个数，确定个数后在个国家中随机选取相应个数分配给帝国主义国家。所以一个帝国中包含一个帝国主义国家和众多殖民地如图2-9所示。

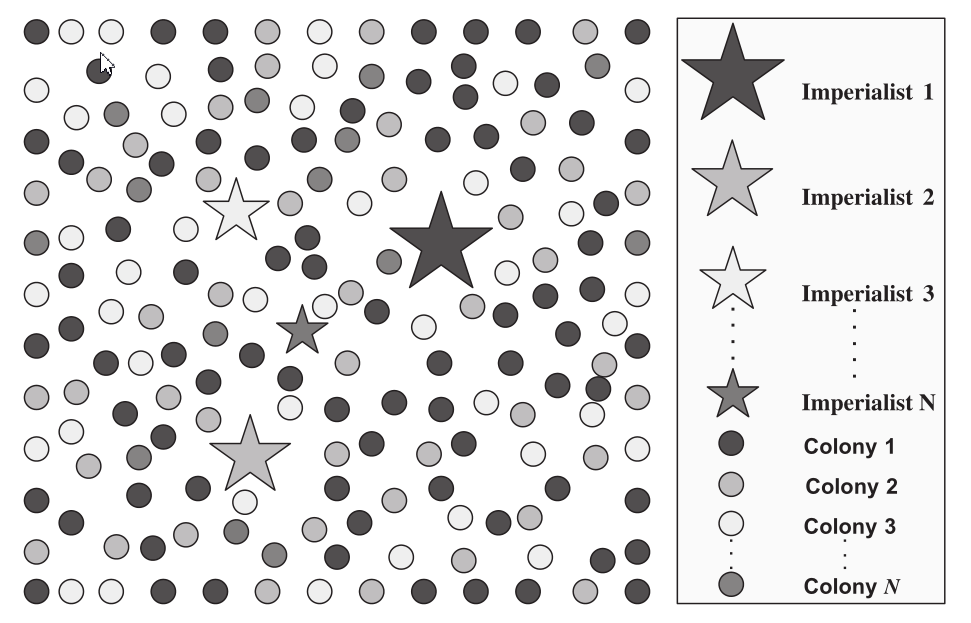
(2)同化机制。帝国主义国家向其殖民地推广文化思想以方便控制的过程称之为同化，ICA中使用殖民地向帝国主义国家移动来模拟同化过程。在一个殖民地移动到新的位置后，它的相似性可能会大于其帝国主义国家的相似性，这时，将调换帝国主义国家和殖民地的角色。

(3)竞争机制。不同帝国之间会有竞争蚕食。首先计算不同帝国的总体实力大小，总体实力包括帝国主义国家的实力及其附属殖民地的实力。选择最弱的帝国最弱的殖民地作为各个帝国的竞争对象，实力高的帝国得之。

(4)帝国灭亡。在多次的同化和竞争的迭代后，会出现一个帝国失去其所有殖民地的情况，此时意味着该帝国灭亡。随着帝国逐个灭亡，最终剩下的帝国即为匹配目标。

该方法有许多值得注意的地方，首先是参数的选取，包括总的国家数、初始帝国数等，这些都会影响匹配效率和准确率。其次，第二部分同化机制中，殖民地会向帝国主义国家移动，其中移动步长的选取至关重要，影响着算法收敛速度和结果准确度，一般是确定移动方向，步长在一定范围内随机取值，以防止错过最佳匹配点。

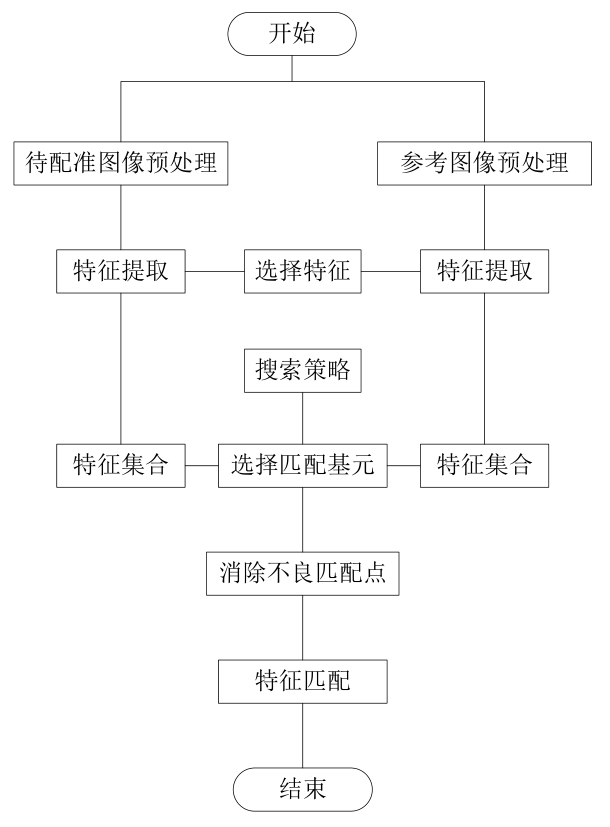
启发式搜索无需遍历全图，提高了匹配速度，但是容易陷入局部最优解，匹配稳定性相对较低。



**图2-9 帝国竞争算法**

2.2基于特征的匹配算法

上一节介绍了基于灰度的图像匹配算法，该算法易于理解便于实现，但是对于旋转、缩放、严重畸变的目标很难准确识别甚至无法识别。而本节的基于特征的匹配算法则能够解决这些问题。所谓图像特征，是一幅图像中几何特征或结构特征，是其区别于其他图片的本质属性。基于特征的匹配算法一般先提取搜索图像和模板图像中的特征，然后根据两者特征的对应关系进行匹配。匹配的具体流程图如下图2-10。



**图2-10 特征匹配流程**

从图2-10的流程中，我们可以知道，基于特征的匹配最为关键的三个步骤为：

(1)特征提取，首先选取合适的特征，然后从图像中应用各种形态学方法，将特征提取出来；

(2)特征描述，提取出特征后，为了匹配的方便，使用特定的数据结构描述，如轮廓特征可以用链表来描述；

(3)特征匹配，在搜索图像和模板得到的特征集合中寻找最为相近的特征参数进行匹配。

图像中一般包含着各种各样的特征信息，如纹理特征、颜色特征、空间关系特征、形状特征等。而基于特征的图像匹配对于特征的选取有着严格的要求，一方面要求特征有着同类聚集而异类分散的特性，另一方面则希望在光照、视角等变化的情况下，特征能够保持稳定。几种常见的特征可以分为三类：区域特征、边缘特征和点特征。

2.2.1 区域特征

当目标区域内灰度或几何等整体特征与其他区域有明显不同时，就可以利用区域特征进行匹配。最简单的区域特征就是面积。以二值化后的图像举例，一个大小为的子图，其面积可以由式(2-31)求出。

由于经过二值化，所以其中的值为0或1。在2.1.2.2节我们知道，面积特征可以通过积分图快速求得。

面积是区域矩广义特征中的特例，对，（p, q）阶矩可以定义为式(2-32)。

由上式可知，就是区域的面积。式（2-32）中的矩会随区域的大小变化而变化，我们通常需要一个具有尺度不变性的特征，所以将矩除以面积得到归一化矩：

从归一化矩中，我们可以得到一个感兴趣的特征即重心（）。归一化矩的大小依然受位置（*x, y*）的影响，为了得到平移不变性，我们可以计算其相对于重心的归一化矩，即归一化中心矩：

Hu在1962年提出的Hu不变矩，其将归一化中心矩定义为：

Hu根据提出了7个与目标的旋转、平移、伸缩不变的量值，本文将其中4个常用量值列出：

|  |  |  |
| --- | --- | --- |
|  |  |  |

本文专门对Hu不变矩做了一些测试，结果如表2-1。

**表2-1 不同完整性的目标Hu矩**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 图片标号 |  |  |  |  |
| a | 0.00157504 | 2.69387e-07 | 2.8192e-10 | 6.06076e-11 |
| b | 0.00137267 | 5.76834e-07 | 1.59913e-10 | 2.20844e-10 |
| c | 0.00165901 | 3.30829e-07 | 6.41433e-10 | 1.30315e-10 |
| d | 0.00164375 | 2.40754e-07 | 4.66723e-10 | 7.155e-11 |
| e | 0.00178099 | 5.3472e-07 | 8.03918e-10 | 1.22032e-10 |
| f | 0.00168862 | 6.63757e-07 | 5.56248e-10 | 9.89794e-11 |
| g | 0.00157452 | 2.27367e-07 | 3.31143e-10 | 9.56789e-11 |
| h | 0.00157022 | 1.50989e-07 | 5.06323e-10 | 1.59089e-10 |



**图2-11 Hu矩测试图片**

在图2-11的a-h的图片中，a为标准模板图像，b-d为经过平移、旋转、缩放后的图像，e-f分别是模板图像在其上、下、左、右有部分缺失的图像。将这8幅图像计算其Hu不变矩得到数据如表2-1所示。

分析数据表2-1可知，Hu矩确实对平移旋转及缩放具有不变性，数据基本在相同数量级。分析e-f的数据，我们发现，当目标有小的缺失的时候，即当前位置与最佳匹配位置有偏差的时候，Hu矩并不能有效判别出来，所以Hu矩更适合粗匹配，而不适合精确定位目标。

2.2.2 边缘特征

边缘普遍存在于目标与目标、目标与背景之间，是一幅图像中最基本的特征之一。边缘蕴含了图像大量的内在信息，如形状、阶越性质和方向等，是人们描述和识别目标的重要特征参数。倘若能够成功地提取出图像的边缘，那么图像的识别及其分析将会变得容易很多，准确度也会大幅提高。

图像边缘的提取一般分为四个步骤：

(1)滤波，由于一般的边缘检测都是基于梯度的变化，所以对噪声十分敏感。因此首先要对原图像进行滤波，过滤掉干扰点。

(2)增强，上一步滤波的时候，在过滤掉噪声点的时候，也将实际的边缘模糊化了，所以我们需要增强步骤来使邻域前强度变化明显的点凸显出来。

(3)检测，经过增强后，会有许多点的梯度较大，但这些点并不全是边缘，这需要根据实际情况确定，人们提出了多种方法来确定哪些点是真正的边缘点。

(4)定位，经过检测后，有些时候需要定位边缘的位置，例如，匹配的时候需要确定匹配点坐标，所以需要最后定位。

Roberts[20]、Sobel[21]、Prewitt[22]、Canny[23]算子等都是图像边缘提取的常用算法。其效果如下图2-12。



**图2-12 不同算子的边缘提取效果**

在提取到图像边缘后，我们一般先进行边缘的描述如使用链表、样条逼近或多边形等。描述是为了更好的匹配，所以描述的方式一般根据匹配的方法确定。本文中，以Hausdorff距离【】作为匹配的标准，使用点集的形式描述边缘。

我们将模板的点集记为T，将搜索图像的点集记为E。两个点集之间的Hausdorff距离用下式表示：

上式中

式(2-37)表示，对于模板点集*T*中任意一点*t*，遍历点集E找到与t最相近的点e，取两者欧式距离作为t的距离，然后取点集*T*中距离最大的*t*作为。的定义正好与之对称。所以式(2-36)表示的Hausdorff距离如果希望是一个较小的值，就要求模板边缘每个点与至少一个搜索图像边缘点相近，同时搜索图像上每个边缘点也至少和一个模板边缘点相近。

由上面分析可知，Hausdorff距离对于理想边缘的匹配效果会很好，但是如果目标发生畸变或者发生了旋转、缩放后，匹配效果会较差，并且匹配的过程运算量也比较大。

2.2.3 点特征

特征点是指在其周围灰度变化剧烈的点，通常为拐点、角点、交叉点等。特征点往往具有尺度不变性的特征，另外特征点一般为图像的局部特征，所以对于遮挡、灰度变化、图像形变等会有较好的匹配效果。常用的特征点匹配算法有SUSAN算法、SIFT算法及其改进算法SURF。下面简单介绍SIFT算法。

SIFT特征点检测算法最早由David G Lowe于1999年提出，并于2004年在总结了各种基于不变特征检测方法的基础上，深入的发展和完善了这个基于尺度空间的特征匹配算法。SIFT特征匹配算法主要分为两个部分，一个是SIFT特征的生成，即从多幅图像中提取对旋转、亮度变化、尺度缩放无关的特征向量；第二阶段是运用SIFT特征向量进行匹配。

SIFT特征的提取分为如下四个步骤：

(1)尺度空间极值点检测：首先生产各种尺度空间的图像，使用高斯微分函数在不同尺度空间图像中搜索潜在的对于旋转和尺度不变的兴趣点。

(2)关键点定位：在上一步得到的候选位置上，通过精细的模型的拟合来确定尺度和位置。算法一般依据他们的稳定程度来选择关键点。

(3)方向确定：为了达到旋转不变性，算法首先要标注关键点的方向，然后经过变换得到不变性的特征。

(4)关键点描述：为了详细描述关键点，在每个关键点的邻域内，计算图像局部的梯度。这些梯度被转换成高维向量的形式，以方便匹配。

匹配方法：特征点的匹配是通过计算两组特征点的高维向量的欧式距离实现的。欧式距离越小，则相似性越高，当欧式距离小于阈值时，可以认为匹配成功。

SIFT及其改进算法SURF在特征点匹配领域有着统治性地位，因为该算法对图像旋转、缩放、甚至仿射变换都能保持不变性。但特征点匹配算法也存在一些共通的缺点，首先计算量大实时性不高，其次对于光滑目标，如圆，基本无法提取特征点。但是特征点匹配也有其大量的适用场景如人脸识别、图像拼接等。

2.3本章小结

本章主要介绍了近年来常用的图像匹配算法，分为两大类即基于灰度的匹配和基于特征的匹配。前者的主要研究方向为相似性度量方式的改进以及搜索方法的创新，后者则主要为寻找一种同类聚集而异类分散的特征而努力。两类方法各有优缺点，需要根据具体情况具体分析。

第三章 基于积分图的倒金字塔搜索

在上一章的介绍中可以发现，基于灰度的匹配和基于特征的匹配是泾渭分明的，各自有一套完整的匹配方案。而本章中，我们创新性的将最简单的面积特征融入到灰度匹配的搜索过程当中，赋予了传统匹配方法新的可能。

3.1基于面积特征的搜索区域筛选

3.1.1 SSDA算法回顾及新算法提出

序贯相似性检测（简称SSDA）是一种基于灰度的匹配算法，也可以说是基于统计的算法，在遍历搜索区域的过程中，统计该子图与模板之间误差大小，当总的误差大于阈值的时候就结束该子图的匹配，接着统计下一个子图的误差，具体算法流程可以参照2.1.3小结。

我们也可以这样认为，该搜索区域共个待搜索子图，先在模板中任选一点，然后计算所有子图中对应点与模板中该点的误差大小，此时如果有的子图误差大小达到了阈值，则直接被筛选掉，不再参与接下来的计算。接着对剩下的子图，再次在模板中取一不重复的点，计算剩下子图中对应点与之误差，将两次误差累加，如果有的子图的误差和达到阈值，则该子图也不再参与后面计算，即在第二轮筛选中被淘汰了。以此类推，不断地选取不重复的点来计算误差，不断的淘汰累计误差达到阈值的点，最终剩下来的，就是选取了最多误差计算点，但是误差和还没有达到阈值的子图，说明该子图和模板之间相似性极高，为最佳匹配子图。另外一种情况就是选取了模板中全部的个点，有些子图的累计误差和还没有达到阈值，这时就可以将经过轮淘汰，剩下的子图中，累计误差和最小的子图作为最佳匹配子图。

在上一段的描述中，我们用了一种全新的角度来看待SSDA算法，即层层过滤直至筛选最佳匹配子图。在这种描述中，可以更清楚的看出，很多待匹配子图并没有较多的参与SSDA算法的计算过程，而是在淘汰的前期就早早的退出了计算。这也就是SSDA算法提高效率的核心。

借鉴这种层层过滤的策略，本文提出了一种倒金字塔搜索，即将全部搜索子图看作金字塔最上面一层，拥有最多待搜索子图，接着定义一种过滤标准，只有通过该过滤标准的子图才能下沉到金字塔下一层。使用不同的过滤方法重复该过程，我们就可以得到唯一一个或很少数量的子图，这些子图都经历了多次过滤，证明了它们和模板有着高度相似性。

3.1.2 基于面积特征的筛选

在上一节的分析中可以得知，倒金字塔算法的核心在于层与层之间的过滤方法的选择，我们可以将SSDA算法作为倒金字塔方法的特例，先分析它的过滤方案。

SSDA算法中，它的不同层的过滤相互关联，以总的误差和作为过滤标准。每多选取一个误差计算点，它就会通过累计误差和过滤一次，也就多出了一层。以一个大小为的二值化模板为例，其中可以选择的误差计算点有10000个，以1%错误率作为淘汰的阈值即累计误差和达到100即淘汰，因为二值化图像中只有0,1，绝对误差也只有0,1，所以要想达到淘汰的阈值，至少要选择100个误差计算点，但这是最好的情况。而在模板0,1各占50%的情况下，模板与全为0的搜索子图进行比对时，平均需要200个误差计算点才能将其淘汰。这样的像素值全为0、与模板没有半点相似性的子图也能够通过200层的过滤，更不要提那些和模板稍有相似的子图，虽不至于达到封顶的10000层，但也要上千层才能过滤掉。

从上面分析中可以发现，作为倒金字塔搜索的一个特例，SSDA算法并不十分优越，只是比需要对每个搜索子图都进行10000次计算的方法好一些而已。究其原因，主要在于低效的过滤方案。毕竟每个误差计算点的信息量很少，只有随着过滤层数的增多，每层的累计误差和的信息量才会逐渐增加，过滤效率才会随之加快。所以在倒金字塔搜索算法中，一个包含信息量大、效率高的过滤方案就是我们追求的目标。

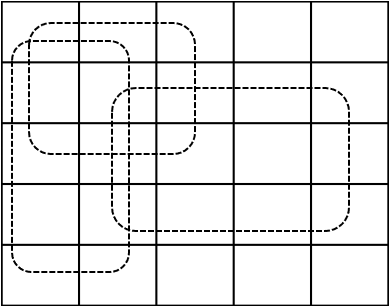
当我们将目光投向基于特征匹配的时候，我们发现面积这个特征就是我们想要找的目标。我们这里说的面积就是将待测面积的矩形内所有像素点的值求和，即

该矩形宽高为*w*和*h*，矩形左上角点坐标为(x, y)。在二值化图像中，该式求出的A就是像素值为1 的区域面积，而在灰度图像（像素值0-255）中*A*则是灰度和，我们将在3.1.5中详细介绍灰度图像的处理。目前可以将图像都视为二值化图像。面积特征的优势在于：

(1)信息量大，面积中基本包含着整个矩形的信息，虽然信息有所损失，但是相比于单个像素点，信息量不可同日而语。

(2)易于计算，如果信息量大但计算不方便，那就退化成了经典的全图遍历型匹配了。而面积特征的计算由于积分图的存在而变得十分容易。在小节3.1.3将详细介绍积分图的计算。

(3)独立性，这是由倒金字塔搜索的本质决定的，在每一层的过滤中，过滤方法必须彼此独立，这样才能保证每一层都可以过滤掉和模板某一方面不相似的子图。如果不独立，两层过滤标准相同，就只能发挥一层的功效。面积特征可以满足这一点，一般取模板中不同大小不同位置的矩形作为每一层的过滤标准，如图3-1，每个虚线框都可以计算面积，由于框内的点不同，所以这些面积特征彼此独立。



**图3-1 模板中矩形的选取**

当选定面积特征来筛选子图后，接着要讨论的就是如何用面积度量子图和模板的相似性。理论上可以用子图矩形面积减模板矩形面积，这个差值越接近0则相似度越高，但是，由于选取矩形大小的不一，差值还需要归一化才能更好的描述相似性。所以为了简单起见，直接用子图矩形面积除以模板矩形面积，即式(3-2)所示的比值*R*作为相似性度量。

其中为子图中选取的以(*x, y*)为左上角点，*s, t*为宽高的矩形面积，同理。当*R*越接近1，则匹配效果越好。这种相似性定义的方式在3.1.5中还会看到它的好处。

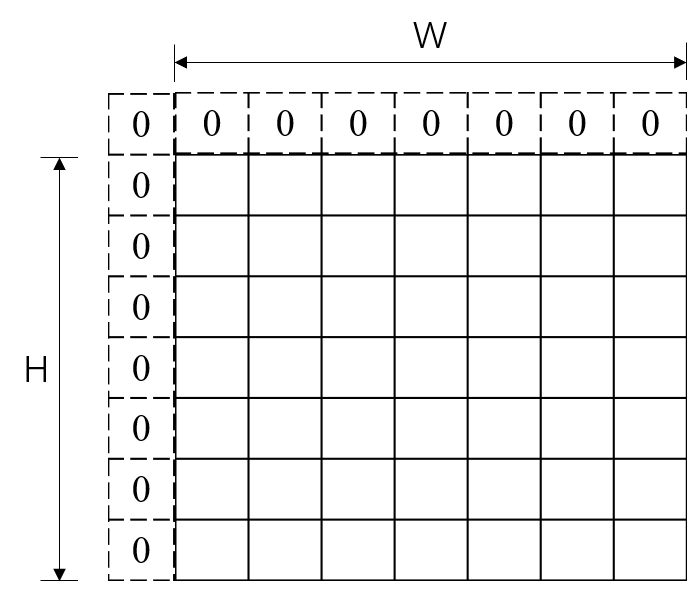
得到相似性度量的方法后，具体的过滤公式就可以给出。在考虑了搜索图像中目标的长宽可能发生畸变，本文以最大10%畸变计算，面积比的变化范围为，最终过滤标准的公式为式(3-3)。

在层层的过滤中，凡是矩形面积比不满足式(3-3)的都认为不是最佳匹配子图，不会传递到下一层。

3.1.3 积分图的计算

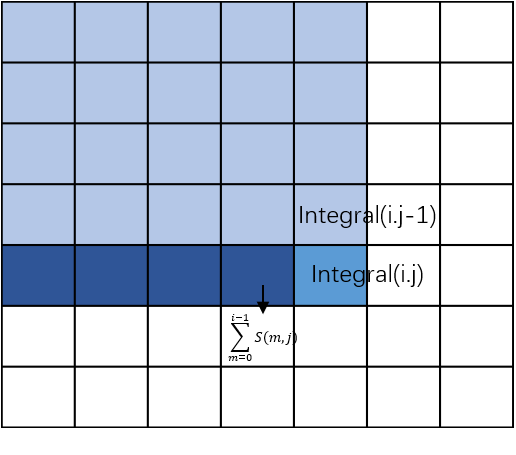
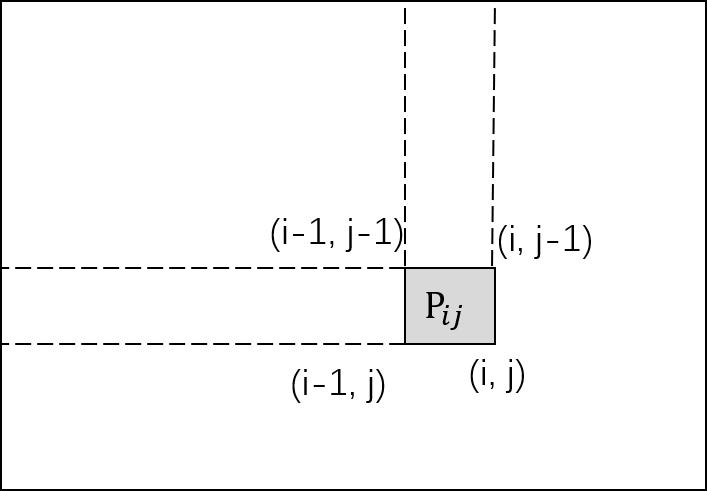
在基于面积特征的倒金字塔搜索中，面积的简单快捷计算至关重要。在2.1.2.2小节，图2-4和公式(2-21)说明了如何由积分图计算面积。本小节将简单介绍如何获得积分图。

设待计算积分图的图像P的大小为，为了各行各列计算的统一性，将P扩充出一个-1行和一个-1列，以0填满扩充像素。总的图像大小变为，如图3-2。



**图3-2 以0扩充后图像**

扩充图像主要是为了原图像第一行和第一列的积分运算方便，以0扩充也不会影响到积分值的大小。积分图的计算有如下图两种方法。

**图3-3 方法一 图3-4 方法二**

两种方法均是先从左到右、再从上到下逐个计算每个像素点的左上方所有像素积分，并存储在对应的大小为的积分图I中。积分图I中每个点的值都可以通过之前计算出的点迭代出来。具体的迭代方法为：

(1)图3-3中：

其中，和分别代表积分图I中当前点和其上一行点的存储值，代表原图像P中深蓝色长条所在像素总和，最后代表原图像P中该点的像素值。

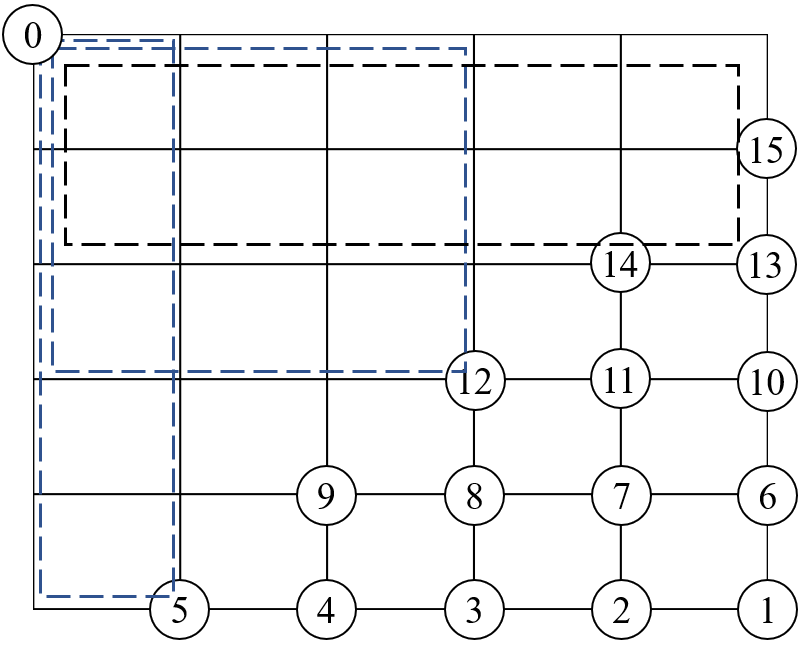
(2)图3-4应用的则是由积分图求面积公式(3-5)。

在求积分图的过程中，已知的是P中(i, j)点像素值，未知的是该点积分图，所以公式可以变形为式(3-6)。

上述两种计算方式，对原图像遍历一次，就能得到其全部积分图。而且计算过程均是加法，效率较高。

3.1.4 采样矩形的选取

图3-1是在模板上随机选取矩形计算面积特征，以此来筛选子图。但实际情况中，矩形的选取对算法的应用效果有着至关重要的影响，选取的标准依然是信息量大、容易计算、独立性。



**图3-5 矩形的选取**

图3-5是本文的矩形的选取情况，其中共选出15个矩形，以标号为0的点为矩形左上角点，以标号为1-15的点为矩形右下角点，图中的虚线框就分别框选出了5号、12号、13号矩形包含的像素点。图3-5为模板大小的示意图，一般情况下，点1-15的横纵坐标都在边框的5等分点处。

下面从矩形选取标准上来说明这样选取矩形的优势所在：

(1)信息量大，以0为左上角点，而右下角点1-15都集中在这个模板的右下方，这样矩形的面积尽可能的大，包含的信息就多。避免了子图和模板选取矩形的面积很小，有少量的误差时，两者面积比值就会严重偏离1，造成误淘汰。

(2)容易计算，根据由积分图计算面积的公式(2-21)，想要得到矩形面积，首先需要获取矩形四个顶点在积分图中的像素值，然后进行加减运算。根据图3-5选取矩形，图像上边界和左边界上的点会多次重复使用，提高了效率。

(3)独立性，图3-5采用5等分点的形式来定义矩形位置，在保证每个矩形都包含较多像素点的同时，也确保了每个矩形之间都有较多的特有点。

另外，本文采用5等分点取矩形，得到了5+4+3+2+1共15个矩形。当然，面对不同大小的模板，可以改变取点方式如4等分点、6等分点等，都会取得不同数量的矩形。而每个矩形都对应一个层与层之间的过滤标准，矩形数量多则过滤层数多，效果好，但计算量随之增加。而且层数过多，过滤效果也会不再明显，所以层数的选择要根据实际情况而定。

3.1.5 灰度图像的处理

这里的灰度图像指的是像素点的值可以取0-255内的任意整数或者0-1内的任意浮点数的图像。对于这种图像，我们也可以用式(3-1)来计算像素点的和*A*。但是此时A并不能代表真实意义上的面积了，因为可能矩形中每个像素点值都不为0，面积就是整个矩形。但是在和模板进行对比时，由于两者选取的矩形像素点个数相同，所以，其中是矩形内像素点的平均灰度，也包含了矩形的大量信息，同样可以用来过滤子图。为了逻辑上的连贯，我们依然将*A*称为面积。

此时，灰度图像的匹配和二值化图像将完全相同。首先选取矩形，计算面积A，接着得到子图与模板面积比*R*，最后根据*R*的大小判断是否淘汰。但是，灰度图像有一个自己独特的问题，即整体灰度变化。也就是说虽然搜索图像中存在和模板大小形状相同的目标，但是由于光照等条件不同，目标灰度比模板图像的灰度整体小一些，如此一来根据上面流程计算出来的*R*全部小于1，即使包含目标的子图也无法通过过滤。

这时就体现出用面积比作为相似性度量的优点了，我们可以先算一个矩形的面积比，以此矩形作为标定，其他矩形计算面积比全部使用标定修正，如式(3-7)。

其中为标定系数，一般将整个模板作为标定矩形，即：

其中(x, y)为子图的位置，*w*, *h*为模板的宽和高。经过标定后会减少一个矩形（即图3-5中的1位置矩形）的过滤作用，但是由于一般过滤层数有余量，所以算法的整体效果基本保持不变。

3.2 T检验法

T检验法[24]是统计推断中的一种常用方法。本文提出的倒金字塔搜索将T检验法作为子图过滤的一个重要淘汰标准，下面将详细介绍T检验法的原理及其在本算法中的应用。

3.2.1 假设检验

在统计学中，由于个体的客观差异，以致抽样误差的不可避免，当遇到几个样本的均值或方差与已知总体的均值或方差有偏差时候，我们要分析造成这个结果的原因是抽样的偶然误差，还是总体的本质误差。假设检验就是判断是偶然误差还是本质误差的方法。

假设检验的基本步骤为：

(1)做出假设，面对一组抽样数据，当我们想要检验总体的均值是否为理想的总体均值。则我们做出原假设，备选假设。

(2)构造统计量，构造一个合适的统计量，要求在原假设成立的情况下，统计量T的分布完全已知。

(3)选定拒绝域，确定一个显著性水平，构造一个拒绝域W，使得，即构造一个小概率事件。

(4)检验假设，将抽样数据代入，看统计量T的观察值是否落在区域W中:

有上面步骤可以看出，假设检验的核心原理是“小概率事件一般不会发生”。

3.2.2 T检验法原理

首先，我们认为抽样的样本来自正态分布的总体，也是用假设来检验均值。由基本统计学知识我们知道，总体均值的无偏估计在假设的情况下，满足：

也就是说此时在已知的情况下，可以构造出一个已知的正态分布。但是在好多问题中，总体的方差并不知道，此时需要用的无偏估计样本方差来代替。于是得到估计量：

当原假设为真的时候，上式统计量，即T满足自由度为n-1的t分布，这种利用t分布的检验法被称为T检验法。当给定显著性水平时，若备选假设，则拒绝域W为：

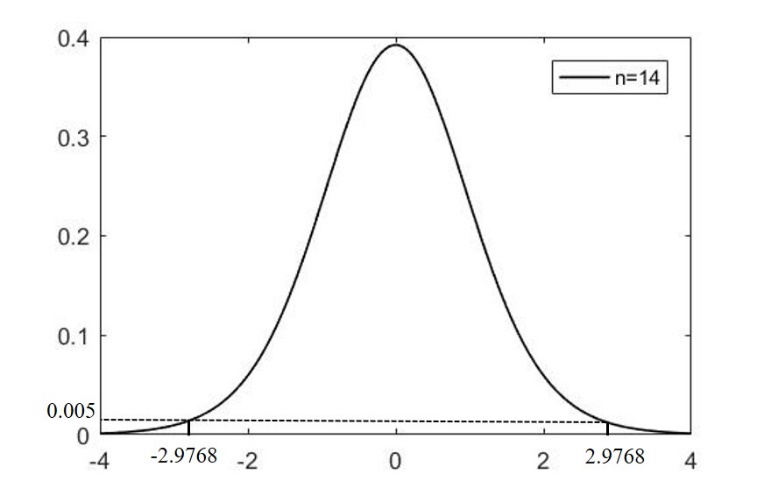
3.2.3 T检验法应用

在3.1节介绍的倒金字塔搜索方案中，选取了15个矩形面积特征，我们认为子图和模板面积比为1，以此作为判断标准对子图层层过滤。实际应用发现，式(3-3)的判断标准较为宽松，有一些似是而非的子图也能够通过，此问题可以通过增加层数来解决，但是当层数增加过多后，矩形面积特征之间的独立性减弱，过滤效果不再增强。所以提出了应用T检验法来做最后的过滤，即把之前每层得到的面积比R作为总体特征的抽样值，应用T检验法来判断这些面积比是否满足假设。

在一般的倒金字塔搜索中，面积比*R*共15个，

其中：

画出自由度为14的t分布图：



**图3-6 t分布**

本文中选取的显著性水平，则拒绝域W：

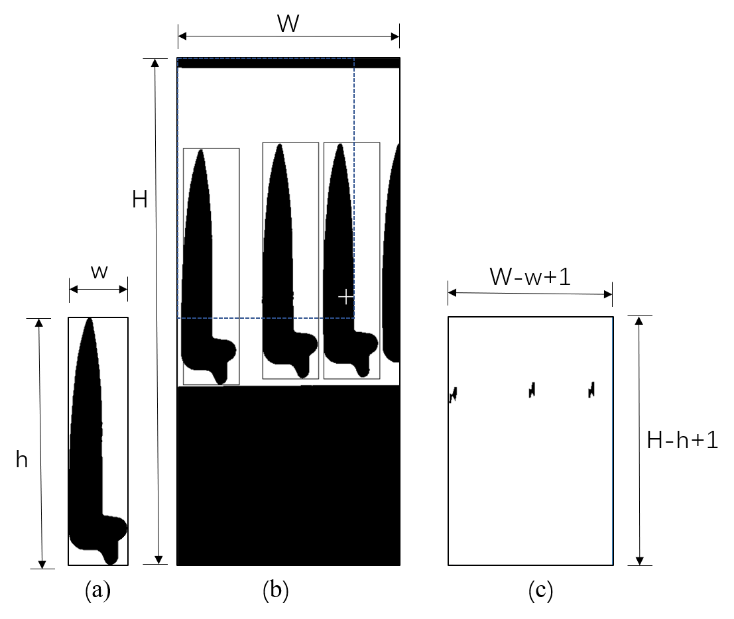
对于T检验法的应用，实际上是先验性的假设了面积比R服从正态分布。另外假设检验的应用整合了每层滤波的参数，得到了一个对搜索子图更强的约束，所以能够进一步过滤。

3.3 精确定位

无论是基于面积特征的过滤还是T检验法的进一步筛选，众多的搜索子图总会剩余下一些与模板十分相近的子图，但是往往一个搜索图像中目标，对应了多个子图，即多个子图都和模板十分相似，所以针对剩下的为数不多的子图，我们需要从中选出一个与模板最相似的子图，作为检测出的目标位置。这个过程我们称之为精确定位。由此也可以看出，倒金字塔搜索是一个先粗后精的搜索算法，而精确定位就是“精”的过程。

3.3.1 聚类

经过了筛选后，剩余的子图往往较少，并且聚集在实际目标附近。如图3-7所示。



**图3-7 一个匹配实例**

上图中，(a)为模板图像，(b)为搜索图像，实线方框为精准定位后的最终选定的子图，较大的虚线框为搜索子图左上角点的范围即搜索区域，(c)为搜索区域，其中黑色点表示经过多轮筛选后剩下的子图左上角点的位置，也就是倒金字塔的最底层。

由图3-7(c)可以看出，最底层的子图有明显的聚类特征，即这些子图都分别聚集在目标附近，形成一个个子图集合。这也证明了倒金字塔搜索的良好特性即将无关子图淘汰，只剩下相似性高的子图。为了得到一个最佳匹配位置，我们需要将每个目标对应的子图集和分离开。理论上，这是一个聚类问题，可能需要类似于K-mean算法才能解决，而实际上，在本图像匹配的问题中，有一个前提条件，即目标不会大面积重叠，而子图又十分接近目标。这就造成了聚集在同一目标附近的子图是大量重叠的，而不同目标对应子图却很少重叠。根据这一先验条件可以很容易将子图分类。

具体的流程如下：

(1)将倒金字塔最底层的子图放入链表中；

(2)取链表第一个子图为新的一类，将该子图移出链表；

(3)遍历链表其他子图，其中若有子图与(2)中子图范围上大量重叠则归为一类，并将该子图从链表移出；

(4)遍历完成后，可以得到一个目标的完整子图集合。此时若链表中还有子图，则跳转至步骤(2)，否则结束。

3.3.2 相似性度量

在得到每个目标的子图集合后，需要从集合中找出和模板最相近的子图。这时最关键在于相似性的度量。

首先，最简单可靠的是经典匹配中的归一化相关性度量，如式(2-11)所示。该方法得到了广泛的检验，匹配准确性和稳定性都很高。

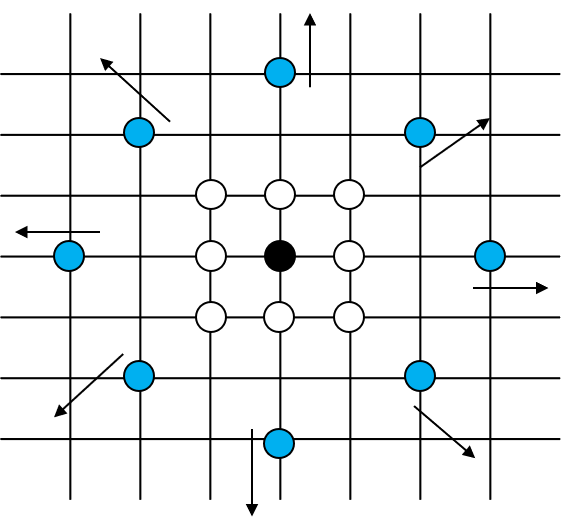
不过，由于本算法使用了T检验法，据此，我们提出了另一种简单的度量方式即使用式(3-12)计算得到的统计量T来度量相似性。在图3-6 t分布的曲线上，我们可以看出，统计量T的值越接近0，则其概率越大。我们有理由认为，大概率事件更容易发生，也就是说统计量T对应概率大的子图更有可能是目标位置。所以说可以认为T越接近0，则相似度越大。

但是，显然这种方法是基于15个采样矩形的，信息量远小于归一化相关性方法，但是其计算速度却是后者难以企及的。

3.3.3 搜索方法

由于归一化相关性方法更加成熟，所以以下以该方法为主。在目标的子图集合较小的时候，可以直接算出所有子图与模板之间的归一化相关系数，取其中最大值即为最佳匹配位置，但是对于子图集合较大的情况，将归一化相关系数全部算出较为费时，本节将介绍一种可以提高效率的搜索方法。

线性正方形搜索算法[25](Line Square Search，简称LSS)是一种基于梯度的搜索方法。该方法分为线性搜索和正方形搜索两部分，如图3-8。



**图3-8 LSS搜索**

图3-8中黑色点为搜索起始点，周围8个白色点为正方形搜索，外围8个蓝色点是线性搜索，代表了8个搜索方向，所以8个蓝色点最多只会计算一个，这取决于正方形搜索中得到的梯度方向。具体算法流程如下：

(1)计算搜索起始点及其周围8个点的相关系数，若起始点相关系数最大则跳转至(3)，否则执行(2)；

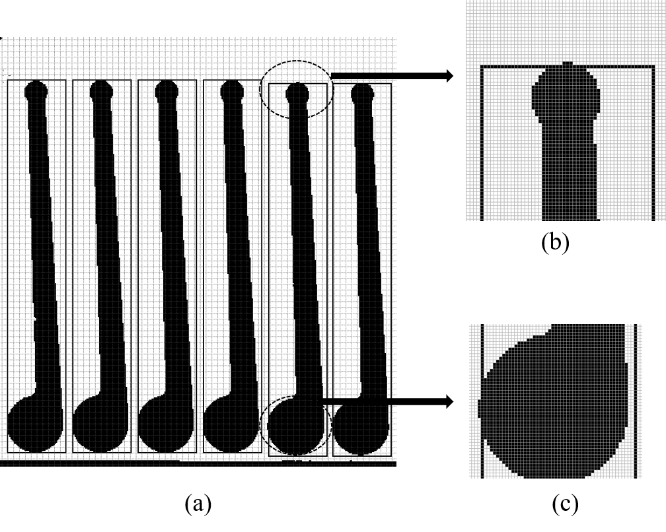
(2)由正方形搜索的结果计算相关系数梯度增大的方向，计算该方向线性搜索点即图3-8中蓝点，如果相关系数继续增加则继续在该方向上取点计算相关系数，直至相关系数开始变小为止。此时将搜索方向上最大相关系数的点作为新的起始点，跳转回(1)。

(3)结束搜索，将此时的起始点作为搜索结果。

从流程中我们可以看出，LSS算法中正方形搜索和线性搜索的交替进行，正方形搜索准确判断否是为最优点，线性搜索则帮助算法快速向最优点靠近。所以LSS算法是一个收敛速度和准确性俱佳的搜索算法。但我们也应该注意到，本算法是在子图十分靠近最佳匹配位置的情况下才能发挥作用的，因为只有这种情况下，梯度方向才能指向最佳匹配位置，否则，LSS算法将失效。

3.3.4 匹配修正

经过前面章节中介绍的倒金字塔搜索和进一步定位后，我们就可以得到在指定相似性度量下的最优匹配位置。但是对于实际工程的问题，我们还可以做一些小的修正，本节就是对数控冲床排料中的图像匹配的结果做出修正。



**图3-9 结果偏差示意图**

图3-9(a)中，黑色方框代表了最佳匹配子图的位置。而从图3-9(b)(c)中可以看出，有些匹配的结果并不完美，会有一些偏差。由于偏差很小，只是在目标附近1-2个像素的偏移，这种偏移往往是因为搜索图像中目标比模板偏小导致的。为了后续操作的便利，需要将目标偏向稳定在搜索子图的左上角，即上边界和左边界和目标完全相切，下边界和右边界可能与目标稍有分离。以上边界为例，具体策略如下：

(1)计算子图上边界与目标之间的交点个数，若为0则执行(2)，否则执行(3)；

(2)计算子图中第二行(即上边界下一行)与目标交点数量，若数量不为0，则跳转至(4)，否则将子图整体向下移动一行，执行(1)；

(3)将子图整体向上移动一行，然后执行(1)；

(4)结束，将此时子图所在位置定义为修正后位置。

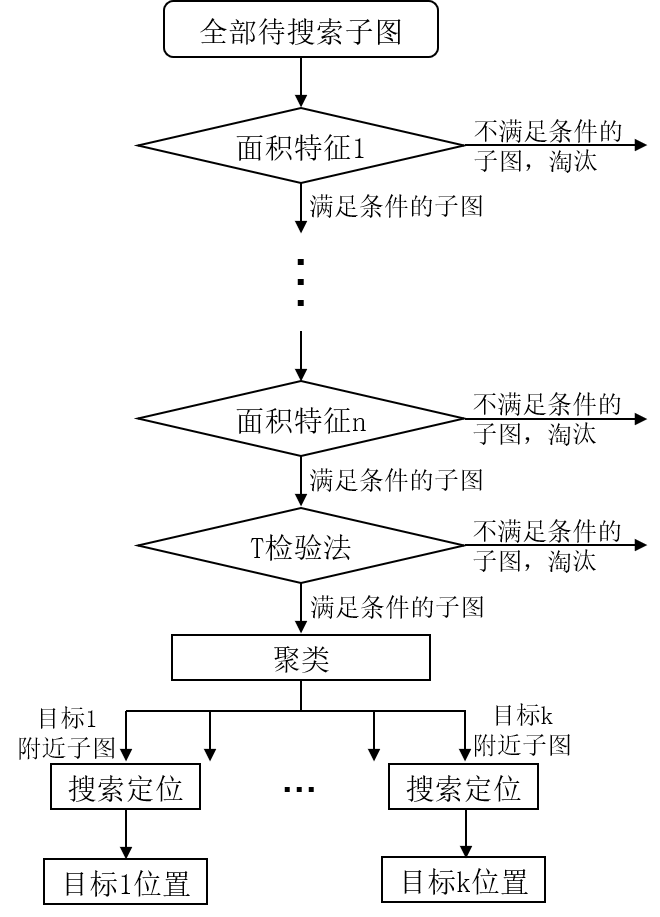
右边界同理。目标位置偏向的稳定，有利于在冲床排料优化中根据目标位置确定其对应的排料位置。在其他图像匹配的应用中，也可以做类似针对性修正。

3.4 算法总结

前面的章节详细介绍了基于积分图的倒金字塔搜索的原理，本小节将对整个算法流程进行总结，并分析该算法的复杂度。

3.4.1 算法流程图

基于积分图的倒金字塔搜索算法主要分为两个部分，一个是前期的倒金字塔搜索，在这一部分中，分别使用面积特征及与面积相关的T检验法层层筛选全部的搜索子图，在金字塔底可以得到少量剩余的子图。二是精确定位目标，首先对剩余的子图按目标聚类，接着定义合适的相似性度量方式，最后针对每个目标对应子图集合，或是遍历搜索或是LSS搜索，得到相似性最佳的子图。具体流程如图3-10。



**图3-10 算法流程**

3.4.2 算法复杂度分析

复杂度的分析也是为两部分，一是倒金字塔搜索，二是精确定位。同样设搜索图像大小，模板图像大小为。

在倒金字塔搜索过程中，首先计算了积分图，由于只遍历了一遍搜索图像而且是加法，时间复杂度为。在倒金字塔过滤的过程中，计算最多的子图需要计算15次面积比和一次T检验法，而最少的只需要计算一次面积比。而且整个搜索过程也只需要遍历搜索图像一次，因此时间复杂度也是。

精确定位所用的方法分为完全遍历和LSS搜索，这两种方法的选择取决于倒金字塔搜索后还剩余的子图数量，数量较少则选择完全遍历，否则选择LSS搜索。倒金字塔最后一层子图数量与目标数和模板的大小有关，大致正比于，即剩余的子图数与目标数*k*、模板的宽高的对数分别成正比。当采用“归一化相关”度量相似性的时候，每个子图与模板相似性系数的计算复杂度为，所以精确定位的大致复杂度为：。

上述的两个过程为并列关系，即时间复杂度是加和运算，所以总复杂度如式3-12所示。

相比于FFT算法的复杂度，如式(2-32)，在目标不多，模板较小的情况下，本算法有较大优势。

3.5 本章小结

本章从SSDA算法的回顾开始，深入分析了该算法能够提高效率的原理，并在此基础上提出了本文的倒金字塔过滤算法。接着介绍了算法的核心是基于面积特征的筛选，由于面积特征有着包含信息量大易于计算的特点，使得倒金字塔搜索拥有了极高的效率优势。然后介绍了T检验法，该方法能够在面积特征面临瓶颈时进一步淘汰不相关子图。由于这是一个先粗后精的搜索算法，所以接着介绍了精确定位策略。最后，根据算法流程，详细分析了算法的复杂度，相较于常用的FFT算法有明显优势。

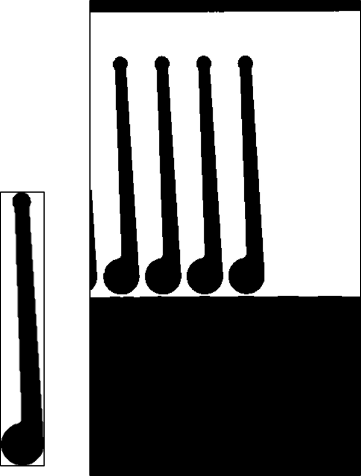
第四章 算法应用效果及分析

4.1 算法分步结果

本文提出的算法分为倒金字塔搜索的“粗匹配”和精准定位的“精匹配”，下面分别给出算法每一步的效果。

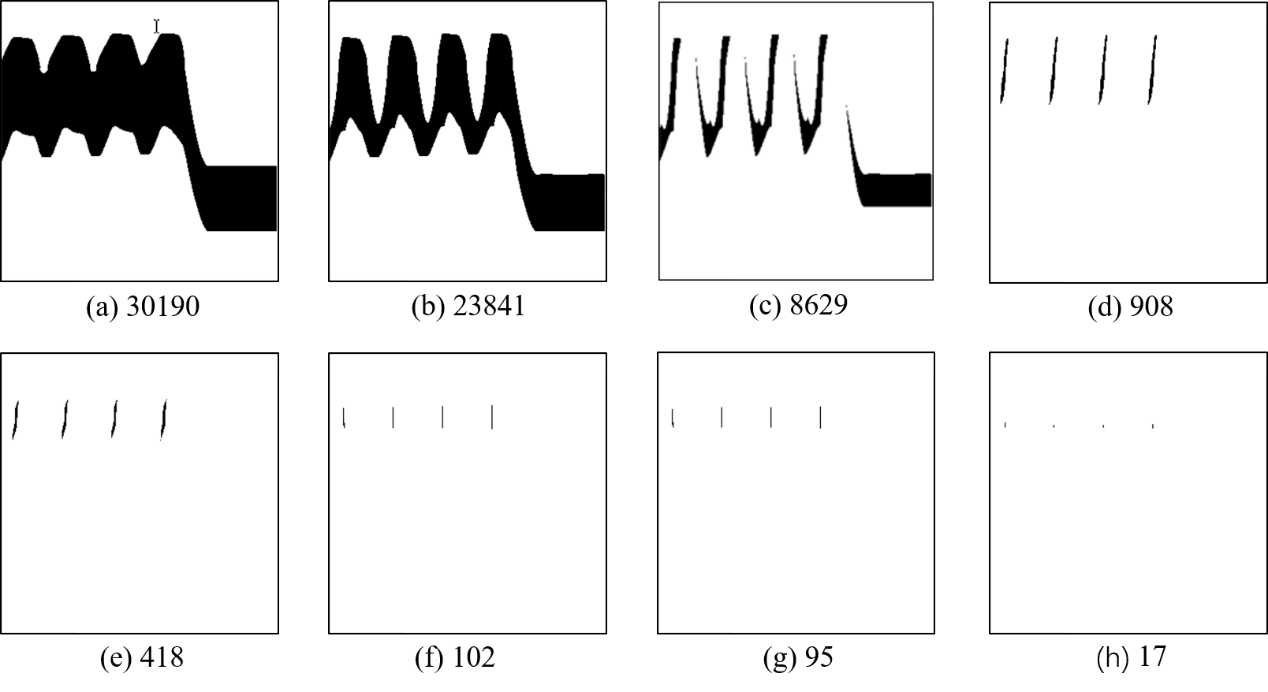
4.1.1 倒金字塔搜索

倒金字塔搜索是一个层层过滤的过程，分为面积特征过滤和T检验法过滤。



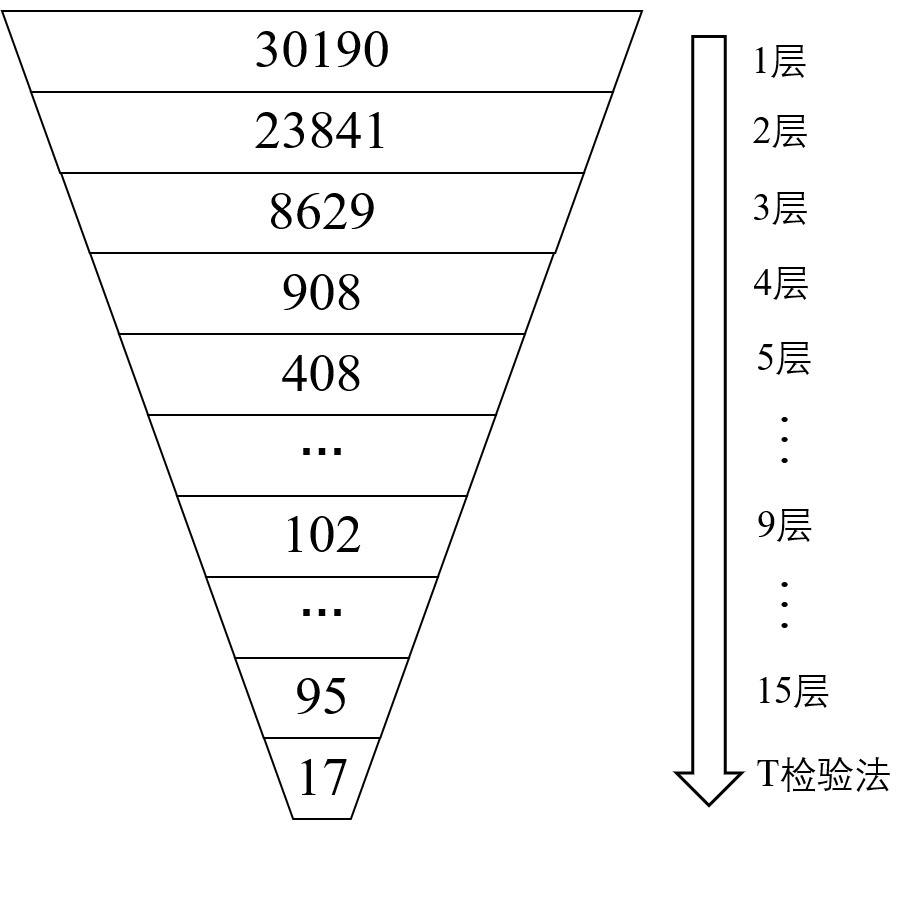
**图4-1 模板与搜索图像**

以图4-1的模板与搜索图像举例，其中模板大小为，搜索图像大小为。所以搜索子图的数量为。图4-2则为层层筛选的效果。



**图4-2过滤效果**

通过筛选的搜索子图在图4-2中用其左上角点表示，即图中黑色点。(a)-(h)标号旁边的数字代表了图中黑色点的数量。其中(a)-(e)分别为经历了前1-5层筛选后的结果，(f)表示第9层结果，(g)表第15层结果，最后一张图(h)，则表示经过T检验法后的结果。图4-3更加形象的表示了倒金字塔搜索效果。



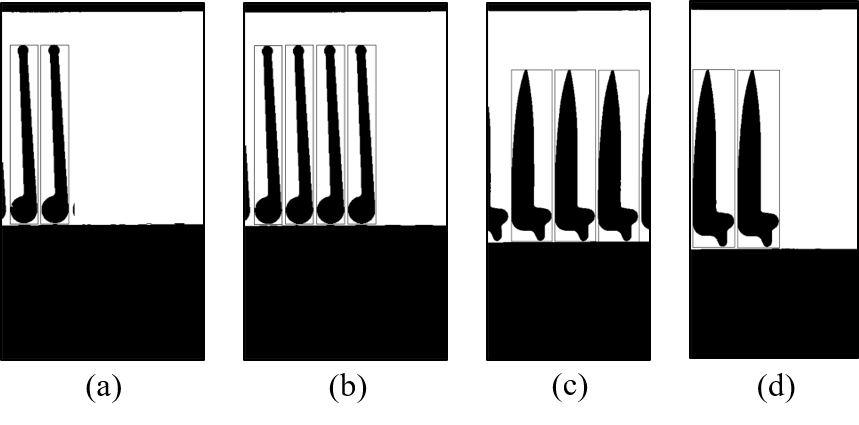
**图4-3 倒金字塔搜索效果**

图4-3清楚的反应了倒金字塔搜索的优势，由于面积特征的优越性，前几层金字塔都能够过滤大量的搜索子图，但同时也可以看到，面积过滤是有瓶颈的，如在9层到15层之间，只过滤了不足10%的子图，也意味着此时再增加金字塔层数已经没有意义了。不过使用了T检验法后，过滤效果又再次明显，说明理论分析中，假设的T检验法成立条件是基本正确的。

4.1.2 精准定位

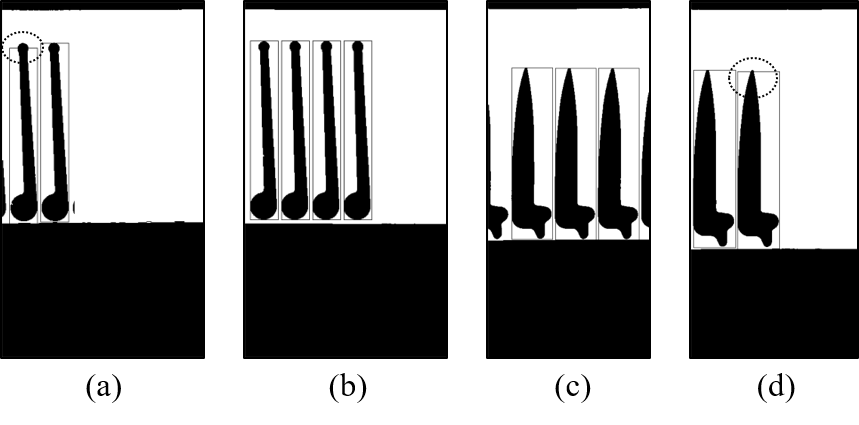
4.1.2.1 LSS与遍历搜索

在精准定位中，本文提出了两种搜索方式即LSS与遍历搜索，图4-4和图4-5为这两种搜索方式的实现效果，表4-1是两种方法匹配所需要的时间。



**图4-4 LSS搜索**

由图4-4和图4-5可以看出，两种方法均能够找到最佳匹配位置，但是明显LSS算法的准确性更高，从图4-5(a)(d)虚线圈内可以看到，遍历搜索时会有一点偏差。究其原因，主要在于，搜索图像中目标发生了畸变，导致在倒金字塔筛选的过程中将最佳匹配位置淘汰掉了，那么再遍历剩余子图自然找不到最佳匹配位置。而LSS搜索却不会有这个问题，因为该方法只需要倒金字塔搜索提供一个目标附近的搜索起始点，并不是在剩余的子图中搜索，而是按梯度方向直接逼近目标，所以总是能够找到最佳匹配点。



**图4-5 遍历搜索**

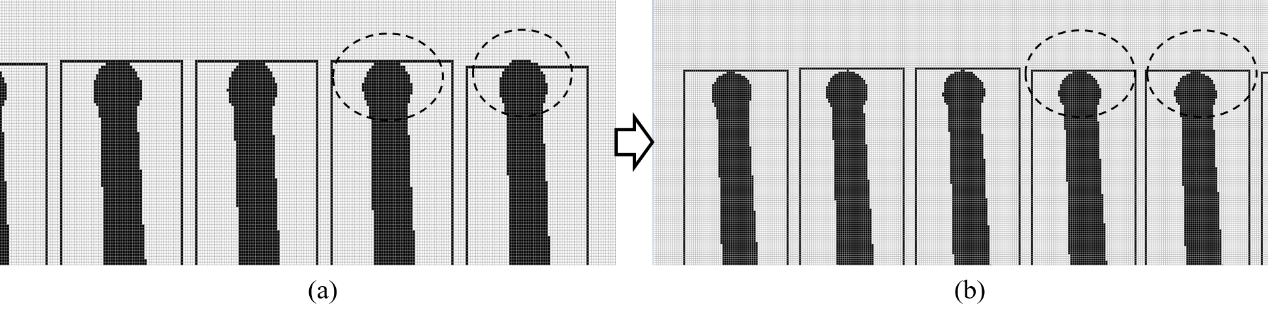
表4-1 列出了两种方法在不同情况下的匹配所需时间，在表中(a)与(b)、(c)与(d)的对比中可以看出，匹配时间和目标的数量几乎成正比，根据之前分析的算法复杂度公式(3-12)，说明精准搜索占据了大量时间，而倒金字塔筛选只占较少时间。从表LSS与遍历两行的数据来看，LSS搜素时间长短更稳定，但由于遍历搜索和倒金字塔筛选剩余的子图数量直接相关，所以对于不同模板与搜索子图，其匹配时间相差较大。

**表4-1 LSS与遍历所需时间(ms)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 图像编号 | a | b | c | d |
| LSS搜索 | 69 | 166 | 250 | 181 |
| 遍历搜索 | 39 | 74 | 550 | 377 |

4.1.2.2 匹配修正

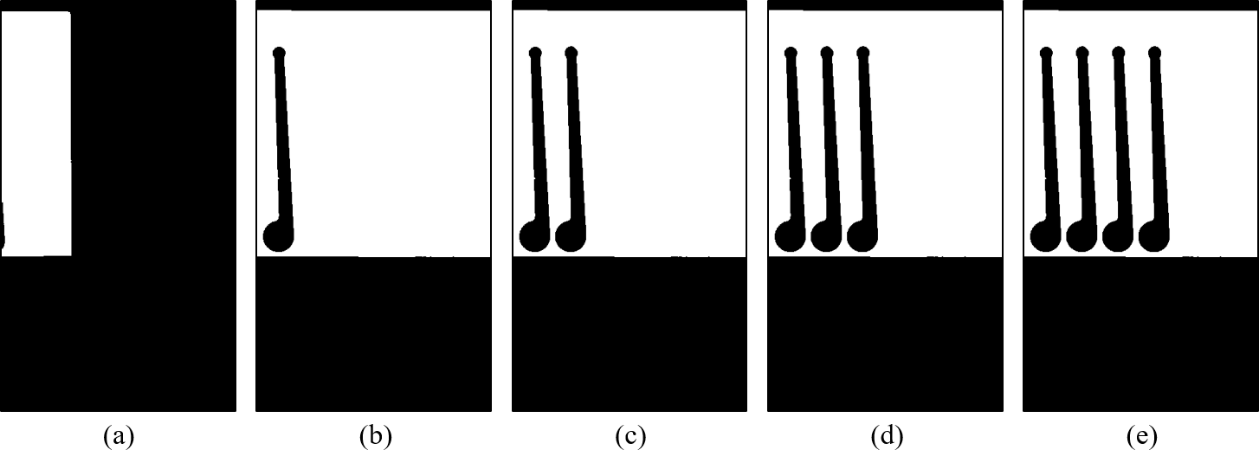
匹配修正是在特殊情况下的一种针对性算法，目的是让目标偏向搜索子图的左上角，图4-6表示了图(a)经过修正后达到图(b)的情况。由于该方法只是针对局部范围的操作，所以基本不会对整体时间复杂度造成影响。



**图4-6 匹配修正**

4.2 与FFT算法对比

在众多基于灰度的算法中，SSDA算法效率较高，但是由于其相关性度量类似于零均值绝对差值法，所以相比于应用归一化互相关法的匹配方式，它的精度较低。而另一种图像金字塔方法，稳定性不佳，所以最终本文使用了OpenCV提供的FFT算法作为倒金字塔算法的对比参照。由于FFT算法和本文算法都是使用归一化互相关作为相似性度量，所以匹配的准确性完全相同，所以主要对比匹配速度。



**图4-7 对比测试图**

将图4-7中(a)-(e)作为搜索图像，分别使用FFT算法和倒金字塔算法进行图像匹配，将匹配时间记录在表4-2中。

**表4-2 FFT与倒金字塔算法时间对比(ms)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 图像编号 | a | b | c | d | e |
| FFT | 40 | 28 | 46 | 33 | 39 |
| 倒金字塔 | 4 | 20 | 31 | 42 | 64 |

表4-2中分别列出了FFT算法与倒金字塔方法匹配时间，该数据由型号为i5 8250U的英特尔CPU测出。为了更清楚了表示表4-2中的内容，将数据画成折线图4-8。

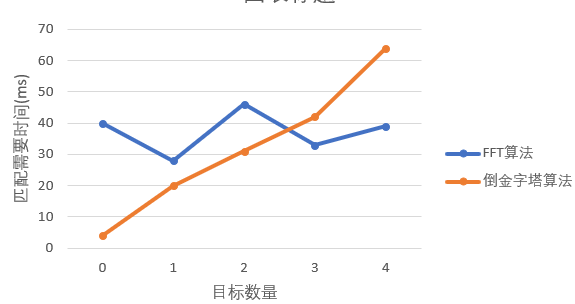


图4-8 匹配速度对比

在图4-8中，我们可以看出，FFT算法面对不同数量的目标运算时间相对稳定，而倒金字塔方法则对目标数量极为敏感，近似成线性增加。在目标较少是，倒金字塔方法有一定优势，而随着目标的增加，FFT算法速度则会更快。但值得注意的是，FFT算法使用的是OpenCV库函数，经过了专业的优化，而倒金字塔方法为直接编程实现，没有经过专门的优化，所以图4-8的数据并不能表现出倒金字塔算法的全部性能。

4.3 本章小结

本章首先罗列了金字塔搜索每步的中间结果，该结果表明，面积特征和*T*检验法都能够很好的筛选匹配子图，证明上一章的理论分析是基本正确的。接着，将倒金字塔方法与FFT算法进行对比，发现目标数量对倒金字塔效率有很大影响，在目标较少时匹配速度优于FFT算法，而目标较多时则慢于FFT算法。同时这也说明倒金字塔的粗略搜索效率很高，而针对每个目标的细化搜索效率较低，是改进的重点。

第五章 结论与展望

5.1 全文工作总结

在数控冲床排料过程中，图像匹配技术是其中重要一环，而一个效率高准确性好的图像匹配算法一直是研究人员努力的目标。本文在前人研究的基础上提出了一种全新的图像匹配方案，并完善了方案的细节使其得以编程实现。在此对全文的工作做如下总结：

首先，详细的分析了各种基于灰度和基于特征的图像匹配算法，其中基于灰度的算法容易理解、方便实现，但是在目标旋转、缩放、遮挡的情况下，无法识别，由于传统灰度匹配算法复杂度较高，所以近年来的的主要研究方向是降低算法复杂度。而基于特征的匹配算法受目标形变的影响较小，但是较难理解，特征提取困难且不稳定，所以目前特征匹配的主要研究方向为特征的选择和提取。

接着，在前人工作的基础上，本文提出了一种“先粗后精”的搜索策略，该策略创新性的将特征匹配应用于“粗”搜索而将灰度匹配应用于精细搜索，兼顾了效率和精度。“粗”搜索中，本文提出了倒金字塔搜索算法。由于应用积分图可以快速计算任意矩形的面积特征，所以倒金字塔算法多次使用面积特征层层筛选搜索子图，并应用T检验法来加强筛选效果，最终筛选剩余的搜索子图都紧紧围绕在目标附近，有较好的梯度特性。在精细搜索中，利用剩余子图良好的梯度特性，本文又提出了LSS梯度上升算法，使用该算法，能够快速找到最佳匹配位置。

最后，将上述图像匹配算法编程实现并与常用的FFT相关算法进行了对比。对比测试中，使用的图片为数控冲床现场采集而来，有较高参考价值。由于两种算法采用了相同的相似性度量方式，所以匹配的准确性完全一致。在效率方面，本文提出的算法匹配时间与目标数量相关，实际测量结果也表明匹配时间和目标数量基本呈线性关系。在目标数量较少时，算法效率要优于FFT相关算法，而目标较多时则FFT算法效率更高，目标数量为3时，两种算法效率近似相同。

从倒金字塔算法匹配的结果来看，基本符合第三章中的理论分析，能够稳定的找到目标，并且准确性也得到了保证。另外式(3-12)分析得到的算法复杂度公式也基本成立，只是从时间结果看，明显，精准搜索需要的时间更长，导致了整个算法时间近似与目标数目成正比而基本与倒金字塔搜索无关，不过这也从另一方面证明了倒金字塔搜索在效率上的优越性。

另外考虑到本文在对比中使用是OpenCV专业工程师编写的FFT算法，而本文提出的算法只经过简单优化，所以本算法效率还有较大的提升空间。不仅如此，与FFT算法和NCC度量方式完全捆绑不同，本文提出的倒金字塔搜索中，精细搜索部分可以任意更换相似性度量方式，增强了算法的灵活性。

5.2 后续工作展望

本文提出了基于积分图的倒金字塔搜索算法，虽然其匹配效果较传统算法优势明显，但依然有较大的改进空间。

(1)改进精准定位算法。本文提出的倒金字塔粗略搜索有极高的筛选效率，但是精确匹配中LSS算法，效率较低，单目标情况下后者占到总搜索时间的80%，多目标时更能达到90%以上，所以迫切需要一种更加优良的精准定位算法。

(2)改进相似性度量方式。在精准定位中，本文依然将“归一化互相关”作为相似性度量，而该方法广为接受主要原因是其可以通过FFT快速计算。所以在倒金字塔算法能够快速“粗”搜索的情况下，完全可以尝试其他度量方式来进行精细定位，甚至可以使用计算量较大的特征匹配算法。

参考文献

[1] 王红梅, 张科, 李言俊. 图像匹配研究进展[J]. 计算机工程与应用, 2004, 40(19):42-44.

[2] Zhang S, Zhou Y. Template matching using grey wolf optimizer with lateral inhibition[J]. Optik - International Journal for Light and Electron Optics, 2017, 130:1229-1243.

[3] Krattenthaler W, Mayer I J, Zeiller M. Point correlation: a reduced-cost template matching technique[C]// Image Processing, 1994. Proceedings. ICIP-94. IEEE International Conference. IEEE Xplore, 1994:208-212 vol.1.

[4] Carsten Steger, Markus Ulrich, Christian Wiedemann. 机器视觉算法与应用[M]. 清华大学出版社, 2008.

[5] Barnea D I, Silverman H F. A Class of Algorithms for Fast Digital Image Registration[M]. IEEE Computer Society, 1972.

[6] 殷松峰, 王一程, 曹良才,等. 基于FFT和积分图的快速相关匹配[C]// 中国光学学会2010年光学大会论文集. 2010.

[7] 张娅丽, 刘波, 解周凯,等. 基于投影匹配的目标姿态测量方法研究[J]. 传感技术学报, 2010, 23(6):825-828.

[8] 赵启. 图像匹配算法研究[D]. 西安电子科技大学, 2013.

[9] Duan H, Xu C, Liu S, et al. Template matching using chaotic imperialist competitive algorithm[J]. Pattern Recognition Letters, 2010, 31(13):1868-1875.

[10] Zhang S, Zhou Y. Template matching using grey wolf optimizer with lateral inhibition[J]. Optik - International Journal for Light and Electron Optics, 2017, 130:1229-1243.

[11] 李峰, 周源华. 基于小波变换的图像匹配算法[J]. 上海交通大学学报, 1999, 33(9):105-107.

[12] 张志佳, 黄莎白, 史泽林. 新的基于边缘特征的图像相关匹配方法[C]// 全国光电技术学术交流会暨全国红外科学技术交流会. 2003:635-638.

[13] MingKuei Hu. Visual pattern recognition by moment invariants[J]. Information Theory Ire Transactions on, 1962, 8(2):179-187.

[14] 谢东海, 詹总谦, 江万寿. 改进Harris算子用于点特征的精确定位[J]. 测绘地理信息, 2003, 28(2):22-23.

[15] Mallat S, Zhong S. Characterization of Signals from Multiscale Edges[J]. Trans IEEE, 1992, 14(7):710-732.

[16] Lowe D G. Object recognition from local scale-invariant features[M]. IEEE, 1999.

[17] Sayago Benito M. Speeded up robust features (SURF) : performance test[J]. 2013.

[18] Crow, Franklin C. Summed-area tables for texture mapping[J]. Proc. SIGGRAPH, 1984, 1984, 18(3):207-212.

[19] 胡竞, 林喜荣, 黄新营,等. 实时图像快速定位算法及其应用[J]. 计算机工程与应用, 2005, 41(35):59-62.

[20] 王冰. 用Roberts算子进行边缘处理[J]. 甘肃科技, 2008, 24(10):18-20.

[21] 袁春兰, 熊宗龙, 周雪花,等. 基于Sobel算子的图像边缘检测研究[J]. 激光与红外, 2009, 39(1):85-87.

[22] 杨道普, 马秋禾, 石磊. 边缘检测Prewitt算子的改进算法[J]. 测绘科学, 2008(s1):100-101.

[23] 陈强, 朱立新, 夏德深. 结合Canny算子的图像二值化[J]. 计算机辅助设计与图形学学报, 2005, 17(6):1302-1306.

[24] 武爱文. 概率论与数理统计[M]. 上海交通大学出版社, 2014.

[25] 丁贵广, 郭宝龙. 基于线性搜索的快速运动估计算法[J]. 西安交通大学学报, 2004, 38(2):136-139.

谢辞

本次毕业设计能够顺利完成，首先我要感谢的指导老师唐厚君教授，唐老师不仅在课题进行过程中不断给予我理论上的帮助，常常提出切实可行的意见，还在学习习惯、学习态度上对我严加要求，让我在学习中时时保持着充沛的精力。其次，我要感谢王堃学长，在整个毕设过程中，学长从选题到理论研究再到编程实现，每一步都给予了我巨大的关心和帮助。最后，我要感谢在我二十几年的生命里，一直陪伴我、关心我的父母，是他们无私的爱支撑着我迎接一个又一个挑战。

**CNC PUNCH NESTING ALGORITHM OPTIMIZATION**

With the development of computer technology and the maturity of automatic control theory, automation equipment occupies more and more proportion in production, among which CNC punch press is an automated production equipment. It applies a huge pressure to the punch and acts on the metal plate to deform the metal until the edge is completely broken, resulting in various shapes of the workpiece.

In the application of the CNC punch press, the stamping of some special plates needs both positive and negative punching two times. The second stamping needs to determine the new stamping position according to the result of the first stamping, so the image matching technique is needed to locate the first stamping position at this time accurately. But for a long time, due to the lack of a matching algorithm with both efficiency and accuracy, the technology has not been widely applied.

With the rapid development of computer technology with Moore's law, computer computing speed has been greatly improved, and now it is able to meet the needs of the vast majority of industries. But in image matching, because of the large amount of computation data and the high real-time requirement, it is still the main research direction to reduce the complexity of matching algorithm. In addition, due to the change of time, angle and illumination, there is a certain degree of gray distortion and geometric distortion in the template image and the search image, which all bring difficulties to the accuracy of the matching. Therefore, a reasonable method of similarity measurement between templates and targets is also an very important research direction for researchers. If a matching algorithm with high matching efficiency and strong robustness is proposed, the image matching technique will play a more active role in navigation guidance, pattern recognition and many other fields. Therefore, the research of image matching technology has very important application value and theoretical significance.

In this paper, an inverted Pyramid image matching algorithm based on integral graph and T test is proposed. The algorithm can improve the efficiency without reducing the accuracy of matching. The following is the main work of this article.

Firstly, intensity-based method and feature-based image matching algorithms are two of the main algorithm, in which the intensity-based matching algorithms includes FFT correlation, sequential similarity detection, image Pyramid and heuristic algorithm and so on. The feature-base matching algorithm can be divided into regional features, edge features and point features according to feature categories. The intensity-based method is easy to understand and can be realized easily, but it can’t be recognized under the circumstances of rotation, zooming and occlusion. Because of the high complexity of the traditional intensity-based method, the main research direction in recent years is to reduce the complexity of the algorithm. The feature-based matching algorithm is less affected by the target deformation, but it is difficult to understand, the feature extraction is difficult and unstable as well, so the main research direction of the feature matching is to select and extract the features.

Then, on the basis of predecessors' work, this paper proposes a search strategy that is “crude and refined”,

(1) Use the inverted pyramid algorithm for a rough search. The inspiration of this method comes from the filtering effect of the SSDA algorithm. The SSDA algorithm relies on the difference of each pixel point to filter the subgraph, but each pixel contains a very low amount of information, which leads to the low efficiency of the algorithm. In this paper, the area features are used to filter the subgraphs. The area features include most of the information in the region. So the area features contain a very large amount of information .As a result, most of the search subgraphs can be eliminated by a few screening times. Moreover, because of the existence of integral graph, the area feature can be calculated only by several addition operations, which has a great advantage of efficiency.

(2) However, after repeated use of the area feature, the pixels will be more and more repeated between the selected area rectangles, therefore the area characteristics will gradually lose its independence. At this time, if we continue choosing more rectangular areas as the screening standard, the screening effect will become worse. At this time, this paper proposes a screening scheme based on T test. The scheme takes advantage of the selection of a lot of area feature data and integrates it to make full use of the overall data distribution effect, assuming that it satisfies the t distribution. Subgraphs satisfying the hypothesis test are left behind and unsatisfied are continuously eliminated. Furthermore, the number of search subgraphs is further reduced by means of hypothesis testing.

(3) Since the "coarse" search filters out most of the sub-graphs, the remaining search sub-graphs will be tightly around the target. In this case, the different sub-graphs corresponding to different targets can be separated easily by clustering. Because the targets in the subgraph are very close, when calculating the similarity, there will be a distinct peak of the best matching position, and the similarity of the other positions is relatively low, that is, it has a good gradient-related characteristic. Therefore, a linear square search algorithm based on gradient ascent is used in this paper. The method gradually approaches the target along the direction of the gradient rise until the search target is accurately located.

Finally, the above image matching algorithm is programmed and compared with commonly used FFT correlation algorithms. In the comparison test, the pictures used were collected from the site of CNC punch press so this test has a very high reference value. Since the two algorithms use the same similarity measurement method, theoretically the matching accuracy of the two algorithms should be exactly the same, and the factual result also proves this point. In terms of efficiency, the algorithm matching time proposed in this paper is related to the number of targets. The actual measurement results also show that the matching time and the target number are basically linear. When the number of targets is small, the efficiency of the algorithm is better than that of the FFT correlation algorithm. When the number of targets is large, the FFT algorithm is more efficient. When the number of targets is 3, the efficiency of the two algorithms is approximately the same.

From the result of the inverse pyramid algorithm matching, it basically conforms to the theoretical analysis in Chapter 3 and can also find that the target stably and the accuracy are guaranteed. In addition, the algorithm complexity formula obtained by the analysis of (3-12) is also basically established. But, from the time results, obviously, the precise search takes much more time, resulting in the entire algorithm time to be approximately proportional to the number of targets and is essentially independent of the inverted pyramid search. On the other hand, this also proves the superiority of the inverted pyramid search in terms of efficiency.

In addition, considering that this article is used in the comparison is an FFT algorithm written by a professional engineer of OpenCV, but the algorithm proposed in this paper is simply optimized, so the efficiency of this algorithm has a huge potential. Not only that, it is different from FFT algorithm and NCC metric completely bounded. In the inverted pyramid search proposed in this paper, the fine search part can replace the similarity measure arbitrarily and enhance the flexibility of the algorithm.

In the theoretical analysis of this paper, inverted pyramid search algorithm should have greater efficiency advantages than FFT algorithm, but the actual results are similar, indicating that the algorithm still has great potential. In addition, this paper still uses NCC as a measure of similarity and does not exert the advantage of this algorithm for better similarity measurement. Therefore, the following two improvements are proposed for future work.

(1) Improve precision positioning algorithm. The inverted pyramid rough search proposed in this paper has a very high screening efficiency, but the LSS algorithm in the exact matching is less efficient. In the case of single target, the latter accounts for 80% of the total search time, and it can reach more than 90% in multiple targets. Therefore, there is an urgent need for a more precise positioning algorithm.

(2) Improve the similarity measure. In the precise positioning, this paper still regards "normalized cross-correlation" as the similarity measure, and the main reason why this method is widely accepted is that it can be quickly calculated by FFT. Therefore, in the case of an inverted “pyramid” algorithm that can quickly search roughly, it is entirely possible to try other metrics for fine-grained positioning, and it is even possible to use a feature-matching algorithm with a large amount of computation.