

## 摘 要

人脸检测是指通过对输入图像进行分析,确定其中人脸的数目、大小、位置、姿态等信息的过程。人脸检测作为人脸识别的先期工作,是计算机模式识别领域研究热点,涉及模式识别、图像处理、生理学、心理学、认知科学等许多领域。由于人机交互、安全监测等应用的日益广泛,这使得人脸检测技术的研究越来越重要。

目前人脸检测方法可归类基于图像块、基于模版匹配、特征不变和基于知识四种。其中基于特征不变的算法,首先用各种数字图像处理方法对输入图像进行处理,根据结果提取某些特征与人脸某些共性特征进行比较,继而判断某一区域是否为人脸。其中眼睛这一器官有着举足轻重的作用,只要眼睛被精确定位,再辅助用脸部其他器官,如眉、鼻、嘴等信息,就可以比较准确的定位出人脸位置。

利用地形学标定的相关理论,本文介绍了一种新的眼睛定位算法。该算法是基于图像成像过程中,瞳孔对光源会有不同程度的反射,会有明显的亮斑效应的光学原理而提出来的。用地形学方法对灰度图像进行标定,眼睛部位会呈现地形学标定中山谷点的典型特征。文章首先利用地形学标定算法找到灰度图像中眼睛的所有候选点位置集合;然后利用眼睛本身固有的灰度、几何信息构造了若干条筛选规则对候选眼睛对集合进行精简操作;通过研究各种投影函数在眼睛对窗口的特点,作者发现眼睛区域积分投影和方差投影的极大值和极小值位置会呈现一定特征,而构造出了积分投影的极大值点、方差投影的极小值点和候选眼睛点三者位置在一定容忍区域之内的眼睛对验证规则。最后运用该验证规则对精简后的眼睛对进行了验证,达到了精确定位眼睛的目的。

实验表明,本文提出的眼睛定位算法,在现在的实验数据下具有较高的正确定位率和鲁棒性,而且可以较好地、快速地定位出单一背景或者复杂背景中单一人脸图像中的眼睛。对于多脸孔的图像,也有较好的定位效果。

**关键词:** 人脸检测, 眼睛定位, 地形学标定

## ABSTRACT

Face detection is the process of searching human faces in an image. As the previous step of face recognition, face detection is related to the fields of pattern recognition, image processing, physiology, cognitive science etc. Over the past ten years face detection has been thoroughly studied in computer vision research for its interesting applications, such as face recognition system, surveillance system and machine interface.

Many algorithms have been got through decade's years of study. In generally, there are four typical face detection methods included based on multiple templates, based on appearance, based on knowledge and based on invariant. Using digital image processing techniques to enhance the quality of the original image, searching some facial features in the image and comparing those features with the features of standard facial model, and locating the existed facial position in the original image are three key steps in the invariant based methods. The eyes hold the balance in all facial features. If precise located the eyes, we can compute the positions of other facial features, such as eyebrows, eyes, nose and mouth, with the help of the priori knowledge of geometrical facial features. So eye locating is thus a very crucial step in the face detection.

As we know, because of the specific reflection characteristic of pupil, the eye center exhibits fuscous while the iris appears bright. If a gray-scale image is treated as a 3D topographic surface with the height of each location being denoted by the intensity of the corresponding pixel, the eye region will show a certain terrain pattern. In particular, the center consists of pit-pixels, surrounded by hillside-like regions. In this paper a new method for automatically locating human eyes in grayscale still images is presented. First, by using topographic feature of eyes, the eye's candidates in the gray image are labeled. Then dilate the smoothed pictures and use some regulations based on the eyes geometrical and grayscale features to screen out more refined eyes pairs. After analyzed the results of experiment we found the positions of the three points (the maximal point of MPF, minimal point of VPF and the eyes candidate point) bounded in a fixed range. So we formed it to a regulation which we employed to perform eye validation in our algorithm.

It is proved that the correct rate of eyes location is very high by new method

based on experimental data. What's more, the new method is easy to realize and is quicker in velocity. And experimental data are proposed in the paper.

**Keywords:** Face detection, eyes location, topographic labeling

## 原创性声明

本人声明：所呈交的论文是本人在导师指导下进行的研究工作。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已发表或撰写过的研究成果。参与同一工作的其他同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签 名：王菊玲 日期：2007.3.23

## 本论文使用授权说明

本人完全了解上海大学有关保留、使用学位论文的规定，即：学校有权保留论文及送交论文复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容。

（保密的论文在解密后应遵守此规定）

签 名：王菊玲 导师签名：丁友东 日期：2007.3.23

## 第一章 绪论

本章前两节通过回顾人脸检测的背景知识及研究现状，简要阐述了眼睛定位在人脸检测中的研究价值，即本论文选题背景和研究意义。第三节介绍现有的一些人脸库，最后在第四节中介绍本文工作内容及后续章节的安排。

### 1.1 课题意义和背景

生物识别技术 (Biometric Identification Technology) 是利用人体生物特征进行身份认证的一种技术。用于生物识别的生物特征有手形、指纹、脸型、虹膜、视网膜、脉搏、耳廓等，行为特征有签字、声音、按键力度等。基于这些特征，人们已经发展了手形识别、指纹识别、人脸识别、发音识别、虹膜识别、签名识别等多种生物识别技术。

其中，人脸识别技术 (Face Recognition Technology) 是一种依据人的面部特征来自动进行身份鉴别的先进生物识别技术。与其它较成熟的人体生物特征识别方法(如指纹、DAN 检测等) 相比,这种技术具有使用方便、不易假冒、识别精度高、直观性突出、基础资料容易获得和成本低等特点，备受学术界关注。尤其是进入20世纪90年代，由于高速度高性能的计算机的出现，人脸识别的方法有了重大突破，人脸识别研究得到了前所未有的重视。

人脸识别包括三个技术环节：人脸检测、人脸特征提取和人脸识别与验证。

人脸检测(Face Detection)是指在给定图像中，确定人脸（是否含有）位置、大小和姿态等信息的过程。人脸检测问题最初起源于人脸识别，是自动人脸识别系统中的一个关键环节。

早期人脸识别主要针对约束较强的图像，人脸位置容易获得，因此人脸检测没有得到重视。但随着人脸识别应用背景的发展，所识别的图像背景越来越复杂，使得人脸检测单独成为一个独立的课题，备受研究者的重视。

今天，人脸检测的应用背景已经远远超出了人脸识别系统的范畴，在基于内容的检索、数字视频处理、视觉监测等方面有着重要的应用价值。

目前,国内外对人脸检测问题研究很多,例如美国的MIT, CMU, Yale等。国内的有清华大学,北京工业大学,中科院计算所和中科院自动化所等研究机构。随着人脸检测研究的深入,国内外发表的有关论文数量也大幅度增长。许多重要的国际会议和期刊都也都涉及到人脸检测问题的研究论题。所涉及的人脸检测算法已经有很多种,即便如此,由于人脸本身固有的特性,很难建立一个评价各种算法的公认体系。

由于人脸非刚性特性与成像环境多样性,使得人脸检测成为一个具有挑战性的复杂的模式检测问题。其复杂性主要体现在下面几个方面:

- 1) 人脸由于外貌、肤色、表情等不同,表现出来的特征各不相同,不可能用同一模式去表达;
- 2) 存在眼镜、头发和头部饰物以及其他外部物体的附属物的遮挡;
- 3) 成像角度的不同造成拍摄出来的人脸的姿态也各不相同;
- 4) 光照的影响,如图像中的亮度、对比度的变化和阴影等。
- 5) 图像的成像条件,如摄像设备的焦距、成像距离,图像获得的途径等。

这些都为解决人脸检测问题造成了难度。如果能够找到一些相关的算法并能在应用过程保证实时性,将为成功构造出具有实际应用价值的人脸检测与跟踪系统提供保证。

目前人脸检测方法可归类基于图像块、基于模版匹配、特征不变和基于知识四种。眼睛在人脸各个特征中处于关键地位,只要眼睛被精确定位,再辅助用脸部其他器官,如眉、鼻、嘴等信息,就可以比较准确的定位出人脸位置。这种基于特征不变的人脸检测算法,具有姿势、观察点、光照条件改变的情况下保持不变的优点。为本文提出的基于地形学标定的人脸眼睛的精确定位算法,提供了一个很好的应用环境,这也是本论文进一步研究的方向。鉴于此,本章下面将对人脸检测研究现状做了进一步介绍。

## 1.2 人脸检测研究现状

人脸检测按照采集图像方式分为静态检测和动态检测。静态检测是从静态图片或者视频序列中找出人脸,如果存在人脸,则输出人脸的数目、每个人脸的位置

及其大小。动态人脸跟踪则是在已检测到的人脸信息基础上,在后续的视频图像中继续捕获人脸的位置及其大小等性质,以达到跟踪识别的目的。

按照采集到的图像又可将人脸检测大致分为如下三个不断深入的阶段: 1) 单一背景单人脸的人脸检测; 2) 复杂背景单人脸的人脸检测; 3) 复杂多人脸的人脸检测。

一般而言人脸检测的框架如下:

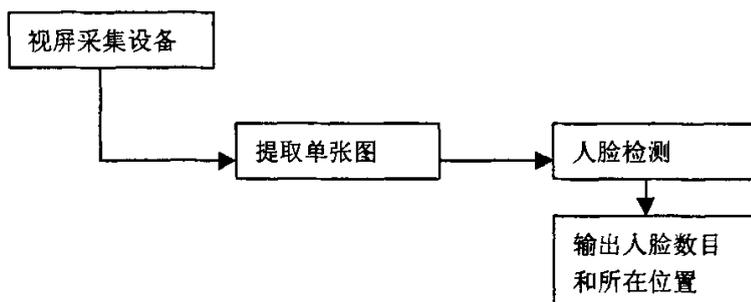


图 1-1 人脸检测的基本框架

纵观人类检测研究历史,按照时间次序可以将人脸检测分为萌芽期,发展及应用期。

早在 19 世纪,英国科学家、探险家和人类测量学家 Galton 就曾经用一组数字代表不同的人脸侧面特征<sup>[1]</sup>。在这一时期,学者致力于寻求人脸的显著特征,从图像处理、统计学等角度进行探讨,试图用尽可能少的人脸特征表示方法来表达人脸特征。这一阶段具有代表性的研究有: MIT 的 Sung 等提出的基于样本学习的方法<sup>[2]</sup>。M H Yang 提出了一种基于 SNoW (Spare Network of Winnows)的学习体系<sup>[3]</sup>,用于检测具有不同特征、表情、姿态和光照条件下的人脸。

90 年代以后,研究者开始结合多种分类算法与人脸特征进行人脸检测。概括起来主要有以下方法:

1) 基于知识的人脸检测算法。这种方法中人脸检测是基于研究者对人脸的

先验知识。研究者把这些先验知识归结成一些规则。

例如人脸中包括两只几何对称的眼睛，眉毛在眼睛的上方，嘴巴在鼻子下面的面部生物特征等。通常，先验知识包含了这些人脸特征之间的相互关系。比较著名的文献有, Yang 等<sup>[4]</sup>首先提出采用由粗到细的三级结构的方法检测人脸。原始图像作为系统的输入，在第一层中，通过一些规则，在整幅图像的范围检测可能的人脸区域。在第二层，在第一层的检测地候选区域基础上运用更加详细的面部特征信息，进行进一步检测。第三层，在第二层检测的人脸候选区域内部进行直方图均衡化，如果被测出的结果法和面部特征，那么人脸就被检测出来了。对于正面人脸情况，该方法可以有效降低误检率，并可以提供一个有效的验证机制，但对于其它视角的人脸，该方法就没有较好的解决办法了。

2) 特征不变的人脸检测算法。这些算法首先用各种数字图像处理方法对输入图像进行处理，根据结果提取某些特征与人脸某些共性特征进行比较，继而判断某一区域是否为人脸。

例如双眼是人脸的突出特征，它们在人脸中占据比较固定的位置，双眼间的距离刻画了人脸的大小。如果我们根据眼睛特征，精确定位出了眼睛位置，这样也就找到了脸相检测的目的。Jeng等<sup>[5]</sup>根据可能的眼睛对获得人脸候选区域，用一个评估函数表示每个区域属于某个特征的可信度，根据评估函数的加权和确定人脸区域。

基于特征不变的人脸检测算法方法具有姿势、观察点、光照条件改变的情况下保持不变的优点。

3) 基于模版匹配的人脸检测算法。这种方法首先是建立人脸的一些标准模版，可以是人脸轮廓，单独眼睛、嘴巴、鼻子等模版，用来描述整张人脸或人脸的部分特性。然后利用一些算法来计算输入图像待测区域与已经存储的标准模版之间的相似程度来进行检测。Bruneli等<sup>[6]</sup>建立了一个标准眼睛模板，利用它来定位眼睛的位置，他们采用五个不同尺度的眼睛模板，在输入图像中搜索眼睛。Govindaraju等<sup>[7]</sup>，使用变形模板(deformable template)匹配头顶轮廓线和左右两条面颊轮廓线，实现人脸定位。

基于模板匹配的方法具有直观，简单的优点，但是搜索计算时间复杂度比较

高。

4) 基于图像块的人脸检测算法。这种方法主要是通过大量学习来制作模版。注意模版匹配中的模板主要是由手工制作出来的, 这里的模版是从一组训练图像经过学习而得来的。通常来说, 基于图像块的人脸检测方法的效率来自于统计数据分析及机器学习。这些图像应该包括人脸外观的具有代表性的变化因素。这些方法主要用于人脸检测。

最近属于这类方法范畴研究比较活跃, 包括主成分分析方法, 支持向量机, 神经网络等等, 其大体思路是先利用主成分分析方法、支持向量机、神经网络方法等将待测图像映射低维空间, 然后通过决策面将人脸与非人脸分开。比较著名例子有, 2001年, P.Viola 发表的文章<sup>[8]</sup>。在 Viola 的文章中, 他不再寻求强的, 复杂的人脸特征表示方法, 分类器由一组简单的快速的弱分类器结合而成, 最后再综合 AdaBoost 和 Cascade (层叠) 算法<sup>[9]</sup>实现了实时的人脸检测系统, 从而使得人脸检测真正地走向实用。目前, 在 Intel 开发的 OpenCV 函数库中, 实现了这一算法, 已在许多领域得到应用, 包括视频序列的处理。

上述各种算法都有自身的优缺点和适用范围, 由于学者的在各自领域的潜心钻研, 也发明了许多卓有成效的检测算法。

总得来说, 现有的人脸检测方法可以分为两大类, 基于规则的方法和基于统计的方法。下面我们对这两类方法做个总结:

#### 1) 基于规则的方法

基于知识的方法一般先提取各种基本特征, 然后根据一些知识规则确认图像中是否包含人脸。这些规则可以通过人的先验知识获得, 也可以通过大量样本学习获得。

这类方法具有直观、直接、简单的优点, 并且不需要训练样本, 对任意情况下的人脸只要建立起相应的规则, 就可以按照这些规则来检测人脸。比较适合于人脸图像分辨率比较高, 人脸比例较大的情况。

#### 2) 基于统计的方法

基于统计的方法一般是通过对大量人脸样本和非人脸样本的学习得到一组模型参, 然后基于模型检测人脸。为了检测大小不同的人脸, 通常采用由粗到细

的分级策略进行处理。对每个分级图像，采用一个固定大小的窗口，通过滑动窗口抽取图像上的不同区域，进行光照校正、直方图修正等预处理后再进行判决。最后需要合并同一位置上多个不同分级检测到的人脸区域。

这类算法具有较强的鲁棒性，适应于复杂背景图像的人脸检测。但这一方法也存在着缺陷，统计模型的建立需要较为一致的训练样本。获取有效的“非人脸”样本有很大的困难。另外，检测速度很慢，计算复杂度很高，对于有视角的人脸，一般没有有效的手段予以解决等等问题的存在。

基于规则的检测方法和基于统计的检测方法通常需要结合起来使用。通过采用简单的特征滤除大量的候选区域，然后采用基于统计特征的方法进行人脸检测，检测速度会提高很多。也可以先采用统计的方法检测出人脸候选区域，然后基于人脸特征进行确认，降低误检率。

我们不得不承认，各种算法都存在着效率与性能上的矛盾。效率与性能本质上是对立的，不可调和的一对矛盾体。在这个对立的前提下尽可能寻找性能与效率的最佳结合点，使检测系统从整体上达到最优，一直是人脸检测科研工作者的追求目标。

### 1.3 人脸图像数据库

在人脸检测与识别的时候，有一系列用于算法比较的人脸库。国际上许多人脸识别小组都收集了自己的人脸数据库，一般人脸库容量不大，最常见的人脸库，包括不到 500 幅人脸图像，如 MIT, Yale, CMU, ORL 等人脸库均为小型库。因为这些数据库是为了特定领域的人脸识别算法而建立的，因此不同的识别程序之间很难进行比较。所以，我们首先来看一看这些数据库的特点和它在人脸检测方面的应用。

FERET 数据库<sup>[10]</sup>是由美国国防部建立起来的通用人脸库，该 FERET 库可以用来评估不同人脸检测的强度和缺点。由于 FERET 库中包括军人的图片，不能在美国以外获得，因此其他国家的研究只能采用本地的人脸库。

Yale 数据库<sup>[11]</sup>是由 P.N.Belhumeur 等人于 1997 年用来进行不同光照条件下人脸识别而创建的，该数据库包括 15 人总共 165 幅灰度图像，每人在不同光照

模型下(正光, 左侧光, 右侧光), 不同表情(高兴, 正常, 悲伤, 疲倦, 惊异和眨眼)和不同脸部妆饰(戴眼镜和不戴眼镜)十一张图像。

UMIST 数据库<sup>[12]</sup>由 20 个在不同姿势下的 564 幅图片组成, 每一组图片都有从右到正面的姿势, 这 20 个图像涵盖了不同种族, 性别, 面貌特征的情况。

BioID 人脸库<sup>[13]</sup>共包含 23 个不同人物的 1521 张正面人脸灰度照片, 照片中的人物大部分为欧美人, 每张照片的解析度为 384\*268。人脸图像在不同季节, 地点采集。光线, 背景, 人脸大小均有较大变化。

国内现在有南京理工大学的彩色人脸库。

本文选用的是 BioID 数据库作为单目标多背景的检测数据库。另外, 单目标单背景数据库选用的是自建的人脸库, 而多目标多背景的样本来自于互联网上的图片与数码相机采集的照片。

## 1.4 论文的主要研究内容与组织框架

人脸检测作为人脸信息处理中的一项关键技术, 近年来成为模式识别与计算机视觉领域内一项受到普遍重视十分活跃的研究课题。

在人脸检测的许多算法中眼睛定位是关键的第一步。因为与人脸其他部件如鼻子、嘴巴相比, 眼睛区域不仅包含了丰富的, 区别于其它部位的重要信息, 而且是对图像进行旋转校正、归一化和均衡化的前提, 同时也是人脸其他部位检测和提取的基础。

对于基于统计特征的人脸检测和识别方法, 首先必须定位眼睛后才能对图像进行旋转校正、归一化和均衡化; 对于基于模板的人脸检测和识别方法, 也必须首先粗定位眼睛, 在此基础上才能定位其它特征点, 而且定位精度要尽量高。能否准确地定位眼睛将直接影响特征点的整体定位质量和识别的准确率。

本文的研究集中在探讨眼睛的精确定位算法上。通过对地形学的灰度图像标定方法研究发现, 眼睛部位的灰度图像呈现典型的地形学标定中的谷点的特征。基于此, 本文提出了一种基于地形学标定的新的人眼定位算法, 其基本思想是, 先利用地形学标定算法找到眼睛的候选点位置, 然后利用眼睛本身的固有灰度, 几何信息排除掉冗余的干扰点, 紧接着利用投影函数对剩余的眼睛结点, 进行验

证并精确定位。

实验表明,本文提出算法有较高的识别率,且利用人脸的轮廓、对称性脸和五官分布特征还能够在一定程度上适用于复杂背景图像中的人脸检测,同时达到较高的检测速度。

本文组织框架如下:第一章,介绍本课题的选题背景、目前研究状况,并概述了本文研究内容。在第二章中简单阐述了与人脸检测相关的基础理论和技术。第三章介绍了眼睛定位的一些常用方法,如用模板匹配方法、神经网络方法和图像灰度信息方法等。第四章介绍了地形学标定的相关理论和算法。在第五章详细表叙了一种基于地形学标定的眼睛定位的新算法。第六章介绍了基于地形学标定的眼睛定位系统的实验设计和结果分析。最后在第七章,我们对论文进行了总结,并对下一步的研究作了一个展望。

## 第二章 人脸检测相关技术

人脸检测与眼睛定位的研究涉及模式识别、图像处理、智能控制、生理学、认知科学，与基于其它生物特征的身份鉴别方法以及计算机人机感知交互领域都有密切联系。本章将回顾与人脸检测和眼睛定位相关的一些理论和技术。

由于摄制条件不同，如光照强弱以及设备优劣的影响，往往存在噪声或者干扰，输入的图像（如一张人脸图像）需要运用一定的技术来提高图像质量。这些提高图像质量的操作，属于图像处理知识范畴，本章第一节对相关技术作了简要介绍。对于人脸检测和眼睛定位来说，对初始图像进行平滑处理，可以大大降低噪声对图像的影响，进而提高特征提取的精度。另外，图像分割也是人脸检测的一项关键技术，目前学术界研究成果很多。如何选择适当有效的平滑、分割算法，是我们必须要面对的问题。所以本章第3、4节中将对这两项技术进行着重介绍。

人脸特征的描述，分类器的选择等等都涉及到模式识别领域的知识。如果把图像中每个子图都看作一个模式，则人脸检测也是一个模式识别问题。近年来人工神经网络已经成功地应用在许多模式识别问题上，比如字符识别、目标识别和自动机器人驾驶。基于神经网络的方法是最近几年人脸检测领域研究的一个比较活跃的方向。在本章的最后一节我们将对神经网络技术进行一个简单的介绍。

### 2.1 图像分析与处理技术

数字图像处理（Digital Image Processing）又称为计算机图像处理，它是指将图像信号转换成数字信号并利用计算机对其进行处理的过程。

数字图像处理作为一门学科大约形成于20世纪60年代初期。早期的图像处理的目的是改善图像的质量，它以人为对象，以改善人的视觉效果为目的。图像处理中，输入的是质量低的图像，输出的是改善质量后的图像，常用的图像处理方法有图像增强、复原、编码、压缩等。

数字图像处理主要研究的内容有以下几个方面：1) 图像变换；2) 图像编

码压缩；3) 图像增强和复原；4) 图像分割；5) 图像描述；6) 图像分类。

1) 图像变换 (image Transform) 主要是指几何变换或空间变换, 是一种建立一幅图像与其变形后的图像中所有各点之间映射关系的函数, 可表示为:

$$[x, y] = [X(u, v), Y(u, v)] \quad (2-1)$$

或

$$[u, v] = [U(x, y), V(x, y)] \quad (2-2)$$

其中,  $[u, v]$  表示输出图像中像素的坐标,  $[x, y]$  表示输入图像中像素的坐标。X、Y、U、V 为惟一确定空间变换的映射函数, 即它们唯一地定义了输入图像和输出图像中所有点之间的几何对应关系。常见的变换包括, 线形变换, 傅里叶变换 (Fourier Transform), 余弦变换, 沃尔什变换, 哈达玛变换和 K-L 变换等变换。

2) 图像编码压缩 (image compression) 是指尽可能少的比特数代表图像或图像中所包含信息的技术。压缩可以在不失真的前提下获得, 也可以在允许的失真条件下进行。有损压缩应用于一般图像, 如风景, 人物照片, 部分医疗图像等, 如大家接触的 JPEG 图像格式一般都是有损压缩。无损压缩应用于认证签名图像处理和档案图像领域, 医疗图像也逐步采用无损压缩方法。常有的压缩方法有: 图像变换、预测压缩技术、自适应压缩技术、彩色影像压缩、二值影像压缩、游程编码、“高频分量”和“低频分量”等。编码是压缩技术中最重要的方法, 它在图像处理技术中是发展最早且比较成熟的技术。

3) 图像增强 (image enhancement) 是数字图像处理过程中经常采用的一种方法。根据图像的特点或存在的问题采取的改善方法或者加强特征的措施称为图像增强。它的目的主要是提高图像的可懂度。图像增强的方法分为空域法和频域法两类, 空域法主要是对图像中的各个像素点进行操作; 而频域法是在图像的某个变换域内, 对图像进行操作, 修改变换后的系数, 例如傅里叶变换、DCT 变换等的系数, 然后再进行反变换得到处理后的图像。

4) 图像复原(image restoration) 是指将图像已知降低因素去掉或减少。图像复原技术与增强技术不同, 它需要了解图像质量下降的原因, 首先要建立"降质模型", 再利用该模型, 恢复原始图像。

5) 图像分割 (image segmentation) 是指根据灰度、彩色、空间纹理、几何形状等特征把图像划分成若干个互不相交的区域, 使得这些特征在同一区域内, 表现出一致性或相似性, 而在不同区域间表现出明显的不同。目前已研究出不少边缘提取、区域分割的方法, 但还没有一种普遍适用于各种图像的有效方法。因此, 对图像分割的研究还在不断深入之中, 是目前图像处理中研究的热点之一。我们将在本章第三节中对图像分割方法进行一个简单的回顾。

6) 图像描述 (image classification) 是图像识别和理解的必要前提。作为最简单的二值图像可采用其几何特性描述物体的特性, 一般图像的描述方法采用二维形状描述, 它有边界描述和区域描述两类方法。对于特殊的纹理图像可采用二维纹理特征描述。随着图像处理研究的深入发展, 已经开始进行三维物体描述的研究, 提出了体积描述、表面描述、广义圆柱体描述等方法。

7) 图像分类 (image classification) 根据各自在图像信息中所反映的不同特征, 把不同类别的目标区分开来的图像处理方法。图像分类(识别)属于模式识别的范畴, 其主要内容是图像经过某些预处理(增强、复原、压缩)后, 进行图像分割和特征提取, 从而进行判决分类。图像分类常采用经典的模式识别方法, 有统计模式分类和句法(结构)模式分类, 近年来新发展起来的模糊模式识别和人工神经网络模式分类在图像识别中也越来越受到重视。<sup>[14][15]</sup>

## 2.2 模式识别相关技术

模式识别(Pattern Recognition)是指按模式抽象对事物或现象进行分类, 辨识类的特征而做出判断的过程。模式识别是研究人类识别能力的数学模型, 并借助于计算机技术实现对其模拟的科学, 其识别系统如图 2-1 所示。

如果模式是张脸, 这就是人脸识别。它是一个具体的模式识别问题, 要求识别出每个人。当然广义上的人脸识别系统还要分析人脸表情等。如果把图像

中每个子图都看作一个模式，则人脸检测也是一个模式识别问题，要求正确得把每个子图模式归入人脸区域类或非人脸区域类。

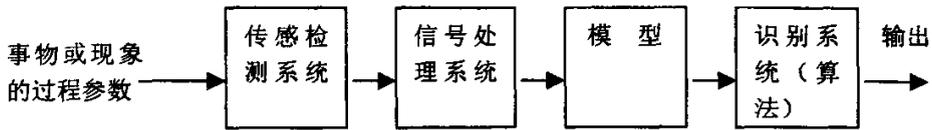


图 2-1 模式识别系统

模式识别诞生于 20 世纪 20 年代，随着计算机与人工智能的兴起，模式识别在 60 年代初迅速发展成为一门科学。几十年来，模式识别研究取得了大量成果，在很多地方得到了成功的应用<sup>[16]</sup>。

模式识别研究主要集中在两方面，一是研究生物体(包括人)是如何感知对象的，属于认识科学的范畴，二是在给定的任务下，如何用计算机实现模式识别的理论和方法。前者是生理学家、心理学家、生物学家和神经生理学家的研究内容，后者通过数学家、信息学专家和计算机科学工作者近几十年来的努力，已经取得了系统的研究成果。

模式识别的方法可以分为以下几类<sup>[17]</sup>：

- 1) 统计模式识别。如果模式类实几何可分的，则几何可分类法按照某种距离度量进行分类；若是几何不可分的，则用概率分类法，如贝叶斯决策理论。
- 2) 结构模式识别。利用语法的递归性，用小而简单的基元和语法规则来描述大而复杂的模式。
- 3) 模糊模式识别。对特征进行模糊化，可以更好地反映问题的本质；对分类结果模糊化，会更有利于下一级分类，因模糊化的分类结果比明确的分类结果包含更多信息。另外，将模糊技术引入其它模式识别算法中，可对它们进行改进，比如优化的模糊 C 均值分类法、模糊 k 近邻分类器。
- 4) 智能模式识别。用神经网络（主要）或逻辑推理。

5) 聚类分析。当没有样本来训练分类器时,要求分类器能根据样本间的相似程度自动分类。

## 2.3 图象平滑相关技术

图象平滑 (image Smoothing) 主要是为了消除被污染图像中的噪声,这是遥感图像处理研究的最基本内容之一。

数字图像的噪声主要来源于图像的获取和传输过程。高斯噪声和椒盐噪声是目前文献中常常提到的两种,已经可以用数学建模的两大噪音。

图像平滑包括空域法和频域法两大类。空间域的图象平滑技术有邻域平均法、空间低通滤波、多图象平均、中值滤波等。在频率域,由于噪声频谱通常在高频部分,可以采用各种形式的低通滤波器的方法减少噪声<sup>[18]</sup>。

下面我们就简单地介绍几种常见的滤波方法:

- 1) 邻域平均法。这是一种简单的空间域处理方法,基本思想是用相邻的几个像素灰度平均值来代替每个像素的灰度。这种方法可以有效地去掉高斯噪声和强度不大的椒盐噪声,但是算法过程中缺少对边缘的考虑,会造成边缘模糊和细节的湮灭。
- 2) 多幅图像平均法。这种方法基本思想就是利用对同一景物的多幅图像相加取平均来消除噪声产生的高频成分。这种方法的难点在于怎样把多幅图像配准,使得对应像素能正确地排序。
- 3) 中值滤波法。基本思想就是以图像中某点为中心将小窗口内的所有像素的灰度排序,将中间值作为该点的灰度值方法。它能很好地去除椒盐噪声,但对高斯噪声无能为力。
- 4) 梯度倒数加权法。基本思想是把中心点像素与它邻域各点的像素取梯度的绝对值,对所得绝对值取倒数。这种方法能较好地保持图像边缘和细节信息,然而对椒盐噪声和脉冲噪声效果并不明显。
- 5) 空间域低通滤波法。图像来说边缘以及噪声干扰的频率分量都处于频率较高的部分,因此可以用低通滤波方法去除噪声。主要借助模板对图像进行邻域操作。空间域滤波器有很多种。它可以有效地处理高斯分布、泊松分布的

噪声,而对较密集的椒盐噪声处理效果较差。

6) 频率域低通滤波法。在分析图像信号的频率特性时,一幅图像中的边缘、跳跃部分以及颗粒噪声代表图像信号的高频分量,而大面积的背景区则代表图像信号的低频分量。用滤波的方法滤除其高频部分就能去除噪声,使图像得到平滑。常见的低通滤波器有:理想低通滤波器、巴特沃思低通滤波器、指数低通滤波器。用频域低通滤波实现平滑,最大的缺点就是处理后的图像有模糊和振铃的效应<sup>[19]</sup>。

另外,在以上几种方法的基础上改进的平滑算法以及基于数学形态学平滑算法等等,研究比较多。图像受噪声污染存在着差异,其中包含的噪声也不尽相同,实际应用种图像往往包含多种不同类型的噪声。因此,好的图像平滑算法应该具备对不同类型噪声的处理能力。我们应该根据自己的实际需要,选择合适自己的平滑算法。

在人脸的局部特征的检测中对细线和小区域的要求都很高,选择滤波器尺寸太大,会造成眼睛区域细节的丢失,为进一步眼睛精确定位造成困难。所以本文选用的高斯平滑滤波器的大小的选择是通过大量实验验证而确定下来的。

## 2.4 图象分割相关技术

图像分割是图像处理中的主要问题,也是计算机领域多层次视觉中的主要问题<sup>[20]</sup>。从二十世纪七十年代起图像分割问题就吸引了很多研究员的高度重视,至今没有一个通用有效的方法,从而被认为是计算机视觉中的一个瓶颈问题。

如果我们感兴趣的目标是人脸区域,比如要进行人脸识别和分析,那就要求首先把人脸区域从背景中分割出来,这就是人脸检测。人脸的检测与分割是图像分割技术的具体应用。

目前学术界已经提出了上千种类型的分割算法,而且近年来每年都有上百篇有关研究报道发表<sup>[21]</sup>。下面我们按照经典得分类方法将其分为:灰度阈值分割法、边缘检测分割法、区域跟踪分割法三种。

1) 灰度阈值分割法:就是先选定一个处于图像灰度取值范围之中的灰度,以确定图像中每个像素点应该属于目标还是背景区域,从而得到相应的二值图

像。阈值分割法可以分为：

全局单阈值，该方法的基本思想是：当灰度级直方图具有双峰特性是，选取两峰之间的谷对应的灰度级作为阈值。

这种方法优点是简单，但是缺点是需要一定的先验只是，而且处理的对象只能是目标和背景对比度相似的图像。

自适应阈值，就是给原始图像分块，对每一块区域选取各自的局部阈值进行分割。这种方法在分割完后，需要采用适当的平滑技术消除边界的不连续性。这类算法的时间和空间复杂度都较大，但具有较好的抗噪性。

多阈值分割，根据不同区域的特点得到几个目标对象，提取的每一个目标采用不同的阈值才能将它们分开。

常见的阈值选取方法有：双峰法、迭代法、大津法、灰度拉伸法和Kirsh算子法<sup>[22]</sup>经典的基于各像素的阈值分割方法有：最大类间方差法（Otsu法），最大熵法、最小误差法等。

2) 边缘检测分割法：边缘是图像最基本特征，是图像分割所依赖的重要基础。可以定义为图像局部特性的不连续性，如灰度突变、纹理结构突变等。图像边缘包含了物体形状的重要信息，它不仅在分析图像时大幅度地减少了要处理的信息量，而且还保护了目标的边界结构。对于边缘的检测常常借助于空域微分算子进行，通过将其模板与图像卷积完成。

在现存的局部技术边缘检测方法中，主要有一次微分（Sobel算子、Roberts算子等）、二次微分（拉普拉斯算子等）和模板操作（Prewit算子、Kirsch算子、Robinson算子），这些算子对边缘过渡比较尖锐且噪音较少等不太复杂的图像，大多提取算法都取得较好效果。

Marr算子、递归滤波器和Canny算子等都是先对图像进行适当平滑，抑制噪音，然后求导，或者先对图像进行局部拟合，然后再用拟合的光滑函数的导数来代替直接的数值导数，这类边缘检测适合于噪声较大的情况。

Snake模型算法是从另一个角度探讨了边缘检测问题，给出一个初始轮廓，然后用图像信息够做内外张力，逼近边缘。这种方法省去了对边缘繁琐的后处理，但提供初始轮廓是它的代价。

3) 区域跟踪法有两种方式: 一种是先给定图像中要分割的目标物体内的一小块作为种子区域。在再种子区域基础上不断将其周围的像素以一定的规则加入其中, 达到最终将代表该物体的所有像素点结合成一个区域的目的。

另外一种先将图像分割成很多一致性较强。如区域内像素灰度相同的小区域。再按一定的规则将小区域融合成大区域, 达到分割图像的目的<sup>[23]</sup>。

分水岭法是一种较新的基于区域增长的图像分割方法。该方法, 先求梯度图像, 然后将梯度图像看成地形图, 对地形图注水进行合并分割。分水岭法有较好的鲁棒性, 但往往会造成过分割。

## 2.5 神经网络相关技术

由于很多复杂的非线性问题常没有算法可循, 人工神经网络(ANN)以其优越的非线性模式分类性能和很强的自组织、自学习能力受到重视。人工神经网络是在探索人脑的计算过程中形成的一种计算模型。对于某些类型的问题, 如学习解释复杂的现实世界中的传感器数据, 人工神经网络是目前知道的最有效学习方法之一<sup>[24]</sup>。

例如, 反向传播算法已在很多实际的问题中取得了惊人的成功, 比如学习识别手写字符, 学习识别口语等。在有些人脸检测算法中, 神经网络已经被用来进行最后的判断工作。在这些算法中, 一般将待检测图像分割为若干个子图, 然后将每个子图归一化后送入神经网络, 它将把每个子图分为两类, 即判断每个子图是人脸还是非人脸,

ANN 学习非常适合于这样的问题: 训练集合为含有噪声的复杂传感器数据, 例如来自摄像机和麦克风的数据。它适合具有以下特征的问题:

- 实例是用很多“属性-值”对表示的。要学习的目标函数是定义在可以用向量描述的实例之上的, 向量由预先定义的特征组成。这些输入属性之间可以高度相关, 也可以相互独立。输入值可以是任何实数。

- 目标函数的输出可能是离散值、实数值或者由若干实数属性或离散属性组成的向量。每个输出值是 0 和 1 之间的某个实数, 对应于在预测分类的置信度 (confidence)。

■ 训练数据可能包含错误。ANN 学习算法对于训练数据中的错误有非常好的鲁棒性。

■ 可容忍长时间的训练。网络训练算法通常比决策树学习这样的算法需要更长的训练时间。训练时间可能从几秒钟到几小时，依赖于网络中权值的数量、要考虑的训练实例的数量、以及不同学习算法参数的设置等因素。

■ 可能需要快速求出目标函数值。尽管 ANN 的学习时间相对较长，但将训练好网络应用于后续的分类或检测，通常是非常快速的。

在许多的人脸检测中神经网络已经被用来进行最后的判断工作。常用的人工神经网络模型有 BP、RBF、Hopfield、Kohonen、ARTMAP 等。其中用得最多的就是 BP 神经网络，下面作个简单介绍：

BP (Back Propagations) 算法，即反向传播算法<sup>[24]</sup>，是基于人工神经网络的模式识别学习算法。反向传播算法是最常用的 ANN 学习技术。基于反向传播算法的神经网络模型，一般被称为 BP 网（如 2-2）。

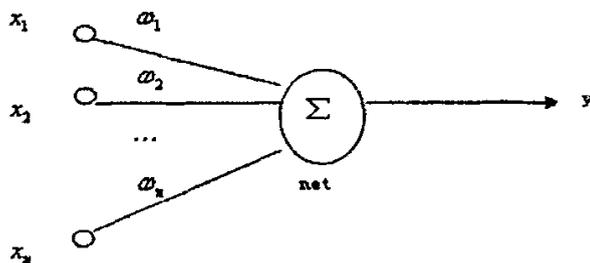


图 2-2 人工神经网络的计算模型

图中  $x_1, x_2, \dots, x_n$  代表  $n$  个输入， $\omega_1, \omega_2, \dots, \omega_n$  代表  $n$  个输入在总的计算模型中的权重， $y$  代表输出值（决策值），其中：

$$net = \sum \omega_i * x_i \quad (2-3)$$

$$y = f(net) \quad (2-4)$$

其中， $f$  是阈值函数。

一个神经网络是由多个这种神经元组成的，两层（如感知器是一种双层神经网络模型，一层是输入层，一层具有计算单元）或者多层，比较复杂的网络由三层以上构成，上一层的输出作为下一层的输入（如图 2-3 所示）。

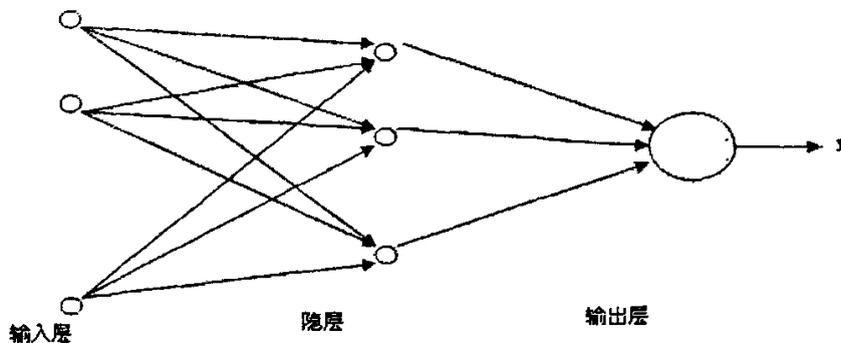


图 2-3 具有隐含层的神经网络

隐层是多个计算单元，不与外界节点相连，可以有 multiple 层。学习的过程就是对权值的修改过程，以使得实际输出接近理想输出。一般的神经网络，因为隐层不直接与外界相连，所以无法计算隐层误差。由此问题引出的 BP 网模型，就是在多层神经网络模型基础上，引入了反向传播算法，由输出层的误差逐层反向传播来间接计算隐层误差。反向传播算法通常采用梯度法来修正权值。

主要算法过程是：首先选定权系数的初始值；然后经过正向过程，输入信息从输入层经过隐层逐层计算各个单元的输出值；再经反向传播过程，通过输出误差逐层向前算出隐层各单元的误差；最后用此误差修正前层权值。

多层前向 BP 网络具有自学习能力，具有一定的推广概括能力，而且特别适合求解内部机制复杂的问题，是目前应用，使得它成为目前最多的一种神经网络形式。但其缺点也是存在的，如学习速度慢，容易陷入局部极值等。

## 第三章 人眼定位相关算法

人脸由眼睛、鼻子、嘴巴、下巴等器官构成。正是因为这些部件的形状、大小和结构上的各种差异才使得世界上人脸千差万别，因此对这些器官形状和结构关系的集合描述，可以作为人脸识别的重要特征。

在人脸检测领域，人脸检测大体可分基于整体或局部的方面考虑可以分为两类：一类是基于人脸各部分特征的检测方法；另一类是把人脸作为一个整体来考虑的检测方法。其中基于人脸分布特征的检测方法，首先是识别出人脸上的有一些有代表性的特征，如眼、鼻、口、面颊、下巴等，然后根据这些特征的存在信任度和其相互之间的集合关系综合起来确定该位置人脸的存在。眼睛因为其固有的几何不变形，使得它在面部特征定位过程中具有举足轻重性的作用。只要眼睛被精确定位，则脸部其他器官，如眉、鼻、嘴等，可由潜在的分布关系比较准确的定位，进而可以检测出人脸的大致位置。

另外，在人脸方位调整中对眼睛的精确定位也是必不可少的步骤，这是因为两眼中心间距受光照或表情变化的影响最小，双眼中心连线方向随人像的偏转而偏转，可以作为图像旋转的依据。而且若以双眼之间的距离对抽取的各个特征值进行归一化，则这些特征值具有平移、旋转和尺度上的不变性<sup>[25]</sup>。

正因为眼睛定位在人脸识别领域中具有如此重要的地位，于是人们研究各种算法来实现眼睛定位，常用的眼睛定位算法有以下几类：霍夫变换法、变形模板法、对称变换法和边缘特征分析法等。

### 3.1 霍夫变换法

霍夫变换是图像处理中从图像中识别几何形状的基本方法之一，良好的抗干扰性，使得应用很广。最基本的霍夫变换是从黑白图像中检测直线，在图像处理领域取得了很多比较满意的成果。针对应用，在检测直线的基础上出现了很多改进算法，目的用来检测数字图像中圆形、抛物线、椭圆等二次曲线。

由于人脸轮廓与特征部件呈明显的几何特性，使得霍夫变换成为人脸识别领域一个常用的方法。眼睛因其不变的几何特征成为人脸部件之中一个重要组成部

在处理好的灰度图像上采用  $a = 0.5$ ，高低阈值分别为30和20的canny分割；然后在眼睛块0-1化图像上运用霍夫变换计算出眼睛的半径值R，并用一个累加器记录所有R值。而累计值最大半径往往并不是我们需要寻找到眼球位置，为了排除干扰，作者巧妙得运用了轮廓线技术。计算霍夫变换检测出来的圆形与图像轮廓的相关度。最后选择那些相关度大且满足一定潜在规则（例如，眼球周围是眼白）的位置最为最终定位的结果。

李粉兰等人<sup>[29]</sup>采用一种基于方向性圆形霍夫变换(Circle Hough Transform)的检测算子来定位眼球中心。文章中作者首先对人脸图像进行边缘提取,并对边缘上的点进行椭圆拟合,由于眼睛宽度和高度的比例是在一定范围之内,所以可以对椭圆的长轴和短轴设定一个阈值,以确定眼睛的候选位置;再运用支持向量集分类器粗略定位眼睛区域,最后在该区域内运用基于方向性圆形霍夫变换检测算子来定位眼球中心。文章采用ORL及自建的数据库进行测试,取得了很好的结果,且该方法对图像平面内旋转、尺度、表情的变化具有较强鲁棒性。

速度慢是霍夫变换方法在人眼定位领域应用的一个制约因素;另外,在图像中要求眼球与周围环境要有反差这个条件也是该方法的一个不利方面。这些是这类方法必须要考虑的问题

## 3.2 变形模板法

把不同传感器或同一传感器在不同时间、不同成像条件下对同一景物获取的两幅或多幅图像在空间上对准,或根据已知模式到另一幅图中寻找相应模式的处理方法就叫做模板匹配。模板匹配是数字图像处理的重要组成部分之一。

模板匹配法主要是模板的选择,根据所选模板的维数可分为二维可变形模板和三维可变形模板。

图 3-2 为简单的二维眼睛模板,由于人脸的旋转角度是任意的,当旋转角度比较大时,其中一部分眼白就会看不见,所以为了适应人脸向两侧作较大角度的旋转,就增加了旋转 30 度和-30 度的两个单眼白的简单眼睛模板。

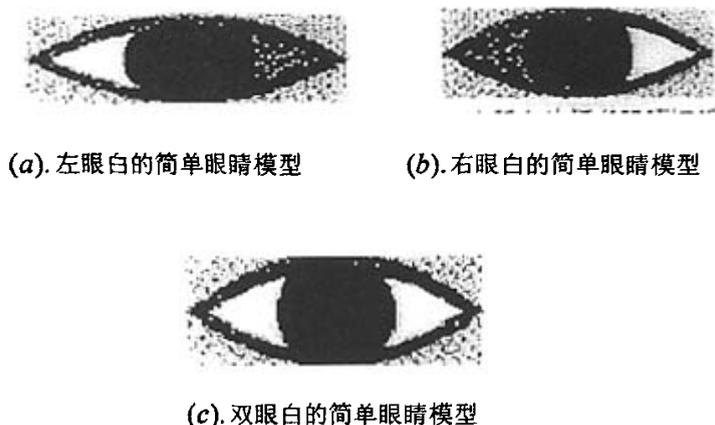


图 3-2 眼睛的变形模板

简单二维眼睛模板可以用如下数学形式表示：

$$SimpleEye = SimpleEye(S, x, y) \quad (3-2)$$

在进行眼睛定位时，不断改变简单眼睛模板的尺寸参数和眼球中心位置参数  $(x, y)$ ，当模板与人脸图像取得最佳匹配时对应的尺度  $S$  就是眼睛的大小，中心位置  $(x, y)$  就是眼睛在人脸图像中的位置。匹配时采用的能量函数与三维可变形模板中的能量函数类似。

三维可变形眼睛模板的生成源于人脸合成时用到的一般三维人脸模型。在一般三维人脸模型中，把眼睛特征点分截取出来，作为可变形眼睛模板的基本点。三维可变形眼睛模板是由10个基本三维点构成的，如图3-3 所示，



图 3-3 三维眼睛模板

三维可变形眼睛模板可以用以下的数学形式来表示：

$$Eye3D = EyeModel(X, S, \theta, x_{eyeball}, y_{eyeball}, z_{eyeball}) \quad (3-3)$$

其中，参数  $X = (x_1, x_2, x_3, \dots, y_1, y_2, y_3, \dots, z_1, z_2, z_3, \dots)$  是由10个三维基本点的坐

标组成,  $S$  表示三维可变形眼睛模板的尺度大小,  $\theta$  表示眼睛模板的垂直旋转角度,  $(x_{eyeball}, y_{eyeball}, z_{eyeball})$  表示眼球的中心, 也是整个三维眼睛模板的中心坐标。眼睛模板的可变形性就体现在不仅模型的尺度  $S$ 、角度  $\theta$  和眼球的中心是可变化的, 而且眼睛模板的形状参数  $X$  也是可变化的。当所有参数都进行变化时, 将产生各种各样的三维眼睛模板。

在模板匹配时, 需要把这些三维模板影射成二维眼睛模板, 然后针对这些二维眼睛模板定义相应的能量函数, 用最小值搜索算法求得最佳匹配。此时所对应的眼睛模板坐标  $(x_{eyeball}, y_{eyeball}, z_{eyeball})$  就是需要求的眼球中心<sup>[30]</sup>。

山世光等<sup>[31]</sup>首先运用肤色模型、模板匹配和特征子脸, 三级人脸检测模型进行人脸检测。在检测到人脸大致区域, 先利用人脸器官构造的先验知识、面部图像灰度分布的峰谷和频率特性粗略地估计眼睛、鼻子、嘴、下巴的大致区域和关键的特征点, 再借助于这些关键点给出模板的初始参数, 这样可以大幅度地提高算法的速度和精度。文中作者用一个圆和两条抛物线来描述眼睛, 能量函数采用的是多边缘检测响应值加权累加算法。

Tsuyoshi Moriyama等<sup>[26]</sup>将眼睛区域分为虹膜、上下眼帘、左右眼角、巩膜、下眼袋、眼帘与眼袋之间的亮处、以及眼袋下面的眼纹几块组成。在文章中, 作者综合考虑眼睛区域的纹理、颜色、几何结构等信息, 用结构参数量化这些信息, 从而模拟眼睛各组成块, 然后用运动参数模拟眼睛运动时各组成块的运动轨迹。在实验中用大量样本训练这些参数, 而通过记录这些参数从而得到眼睛运动的轨迹, 达到跟踪眼睛运动的目的。

尹方平等<sup>[32]</sup>选取80幅光线较好的人脸图像, 对它们做光线规范化后提取 Gabor高频特征来构造模板。文中利用统计模式识别的原理, 在眼睛的周边区域进行模板匹配, 突出眉毛与眼睛这一整体的大致位置, 然后进行二值投影确定人眼的最终准确位置。文章在模板的制作过程中, 是将眉毛与眼睛作为一个整体来制作模板的。首先截下眉毛和眼睛部位统一大小并求平均值来得到左、右眼的模板。在模板匹配的过程中, 首先对待检测的人脸进行预处理与Gabor滤波, 再进行模板匹配, 粗略定位眉毛和眼睛的大致位置, 最后在此小区域内进行二值投影, 可得到眼睛的准确位置。

模板法的两个主要缺点是：一个是对参数初值的依赖程度高，容易陷入局部最小；另一个是计算时间较长。

### 3.3 对称变换法

对称是自然界最常见的现象之一。如果经过某种对称算子的作用后,物体保持原样不变,则称该物体是对称的。例如,字母A是左右对称,Z是对角对称,H既是左右对称又是对角对称,如图3-4所示。

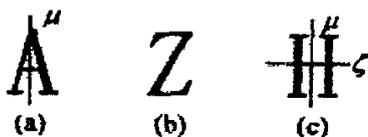


图 3-4 对称的例子

对称可定义为,给定一个二维平面上的物体,如果存在一条将它分成两个呈镜像部分的轴  $x$ , 则称它关于  $x$  对称。如图3-4 (a)所示,字母A关于  $u$  轴对称,图3-4(c)所示中的H关于  $u$ 、 $\zeta$  对称。如果把面也考虑进来,这个对称的定义可以延伸到三维空间。由于对称的存在具有普遍性,所以许多学者提出了不同的计算物体对称程度的方法<sup>[33]</sup>。

目前在人脸器官的定位和特征提取应用比较引人注目的广义对称变换GST (Generalized Symmetry Transform) 对人脸而言,眼睛、鼻子、嘴巴等都具有很强的点对称性,广义对称变换正是用来描述物体点对称的。

Daniel Reissfeld等<sup>[34]</sup>人利用广义对称变换定位人眼进而提取脸部特征。文章中作者首先运用广义对称方法检测出眼睛和嘴巴位置,再利用面部特征的几何性质进行验证,然后利用检测到的特征对人脸图像进行归一化。文章实验表明该方法与其它人脸归一化方法相比,有很好的鲁棒性。

利用广义对称变换进行人脸定位,只考察人眼中心点的强对称性和脸部特征的几何分布,对人脸偏转、表情变化、光照变化等条件不敏感,因而具有很好的鲁棒性。但由于广义对称变换只描述了各点的对称性,当它用于人脸定位及其他一些定位的应用时,会导致待选点较多而使得精确定位比较困难。

针对广义对称变换及其在人脸定位应用上存在的问题,周杰等<sup>[35]</sup>定义了一种新的对称变换——方向对称变换(Directional Symmetry Transform)。文章首先通过预处理对尺度因子进行自动估计,缩小人眼中心点的存在范围,然后进行简化的方向对称变换,再应用简单知识对待选点(即强对称点)进行筛选来得到人眼中心点的。

与广义对称性不同之处在于,方向对称变换是按不同的方向来求取对称性,能够得到对称性在不同方向上的分布情况,而对称性在各个方向上的分布能够在很大程度上反映出物体的基本形状特征,对于不同形状的物体,其中心点的对称性分布表现不同。但是,以上对称变换的计算需要在大范围的尺度上进行,计算量很大。而且,由于只是描述了各点的局部对称性,当它用于脸部特征的提取时,产生的候选特征点较多,不太利于特征点的精确定位。

结合广义对称变换和方向对称变换的优点,刘文予等<sup>[36]</sup>提出一种新的对称变换——离散对称变换,简记为DST (Generalized Symmetry Transform),它不仅具有广义对称变换描述物体对称大小的特点,而且通过对各点领域的考察,去除那些处于规则区域外的点,可大大降低计算量,实现眼睛定位的快速算法。在文章中作者首先定位出双眼中心点,然后利用脸部器官固有的几何比例、相对分布关系来定位嘴和鼻子的位置。通过图像的离散对称变换后,眼睛中心点附近像素的对称值一般都处在对称值最大的前四、五位,对这些具有强对称值的候选点,采用两条简单规则进行筛选。最后取基本符合眼睛分布规律的两点作为最终结果,从而达到了定位的目的。

### 3.4 边缘特征分析法

该算法的基本思想是利用预测法先确定人眼在眉眼区域的大概位置,通过检测眉眼部位的边缘及边缘分组,确定双眼的坐标位置。一般此算法所处理的图像为单人灰度图像。

经典预测策略是利用图像的垂直灰度投影曲线,根据投影曲线凸峰宽度确定人脸的左右边界,然后利用人脸区域的水平灰度投影曲线确定头顶及鼻子中部形成的上下边界。观察不同单人人脸图像的垂直灰度投影曲线,可以发现人脸所在区

域将使垂直灰度投影曲线形成一个具有一定宽度的凸峰。这个凸峰的左右边界大致代表了人脸的左右边界。

当人脸的左右边界确定之后,取左右边界之间的人脸区域作为研究对象,作该图像的水平灰度投影曲线,观察人脸区域图像的水平投影曲线可以发现,此曲线的第一个极小值点对应人的头顶,因为头发的低灰度产生了水平投影曲线的低谷,而曲线的最大值点和次最大值点分别对应人的额头部位和人的鼻中部,这样就可以轻易地求得人眼所处的大致水平位置。

陶亮等人<sup>[37]</sup>为了寻找最佳分割阈值使人眼从复杂图像背景和人脸中分离出来。取分割阈值的初始值 $T_0$ 为直方图低灰度区的第一个峰点对应的灰度值。然后在 $T_0$ 的基础上,增加步长,不断调整阈值大小,直到某个阈值下检测到双眼黑块出现,此时的阈值为最优分割阈值,检测到的人眼块的几何中心也为瞳孔的中心。最后对检测到的眼睛对进行相似性分析来验证检测结果。

由于颜色是人脸最为显著的特征之一,利用肤色模型检测预测法来确定人脸是近几年研究的热潮。根据肤色信息定位人脸区域,其优点是速度快且对姿态不敏感。对背景简单的图像,仅利用肤色分割就可以定位人脸;而对于复杂背景的彩色图像,利用肤色分割的方法可以有效地减小进一步检测的搜索空间,提高检测速度。

Rein-Lien Hsu等人<sup>[38]</sup>文章中为了除去背景对人脸区域的干扰,采用了光线补偿技术,再在YCbCr彩色空间中,根据眼睛处灰度信息,构造非线性公式来定位眼睛位置。文章中还利用嘴巴处灰度信息构造了嘴巴位置的,并且在文章最后利用眼睛和嘴巴特征检测出了脸部轮廓,实现根据脸部特征来检测人脸的目的。

另外,基于虹膜灰度特征标定方法,也是一种眼睛定位的方法。根据眼睛区域灰度信息比较低和虹膜成像特点,眼睛区域在灰度地形学描述中上容易成山谷点的地貌特征。在灰度图像上,利用灰度信息找出候选眼睛点位置。然后通过一定方法排除干扰点,来找到眼睛对的最终位置。

Jun Wang等人<sup>[39]</sup>利用灰度图像标定方法,对眼睛区域的山谷特性进行了分析,且运用切比雪夫微分滤波器对连续的数学定义进行了离散化,来找到离散灰度图像的山谷点位置。最后文章利用分类器方法对这些人眼候选队进行筛选,从

而达到了定位眼睛的目的。实验表明该方法具有很好的检测效果。

不管是投影方法，肤色检测方法还是地形学方法，在边缘检测前都必须先对眉眼区域进行去噪、增强处理，使图像平滑，这是保证下面边缘检测获得较好效果的关键步骤。

然后利用初步检测的结果，用分割算子求出眉眼区域的边缘图像，求解时阈值取得较高，使得边缘点仅由眉毛和眼睛的强边缘以及瞳孔产生的强边缘组成，而不会包含由鼻子的轮廓产生的弱边缘。随后将边缘点进行分组，可以得到几组分离的边缘，其中位于下面的两组分别为左右两眼产生的边缘，取这两个边缘组的中心作为人眼的中心。边缘分组算法中，将分别包含每个边缘分组的最小矩形作为人眼检出，最先检出的两个分组分别对应左眼和右眼。当出现两个分组后算法立即结束，提高了执行效率。

该算法的优势在于首先通过预测算法确定眉眼区域，使数据量大大减少，然后在有效的边缘检测之后，使用了高效的边缘分组策略，在保证正确率的基础上，提高了算法速度。此算法的定位错误出现在头部倾斜度过大，以及侧光太强，还有头部在图像中所占区域过小的情况。

本文讨论的眼睛定位算法就是基于这种边缘特征法的，通过大量试验表明，该方法具有很好的定位效果。

## 第四章 地形学标定算法

人眼可以通过脑中存储经验值,来判定二维图像上各种景物在实际场景中的位置信息。随着计算能力成本的大幅度降低,用计算机代替人眼来识别、跟踪和测量数字图像上物体变成了可能。其中运用地形学知识体系来识别物体是其中一种常见的方法。该方法根据各像素的在数字图像上表现的地貌特征,来对物体进行归类,从而达到了目标物体识别的目的。

众所周知,获取数字化灰度图像的过程中会有信息量的大量流失,会使所获得的地形学知识与实际景物有一定的误差。另外,很多经典的物体识别、提取算法都是基于 0-1 图像提出来的。灰度图像转化成 0-1 图像过程中又不可避免的有更多信息量的流失。因此,如何直接从灰度图像中获得我们所需要的地形学知识成为提高识别效果的一个研究方向。在文献<sup>[40]</sup>中证明利用灰度图像地形学知识要比直接运用 0-1 知识进行自动邮政搜索效果要好。

在地形学知识领域,将一幅灰度图像看成三维的地形学的形面,来对灰度图像进行分析,最早是由 Haralick 等人<sup>[41]</sup>于 1993 年提出来的。在文章中他们运用该方法对手写字体识别中部分连接字体进行了分割,取得了比较好的效果。另外,1995 年, Trier 等也进一步运用了这种基于地形学灰度图像标定方法<sup>[42]</sup>,在文中根据地形学特征,对灰度图像中的每个像素点进行标定成为山顶、山谷、坡状面、平面、背脊面、沟壑面等十二种状态中的一种。然后再对这些特征进行归类,就能获得需要的灰度信息,在文章的最后作者用该方法对写满字母和直线的图像进行了识别,该方法能够获取了很高的识别率。

### 4.1 地形学描述

一般来说,正向地貌包括山顶、鞍部、分水线及漫长的斜坡。这些重要地貌大部分是构造骨骼的主要单元,占着地貌描述中的首要地位。地貌描述的清晰性与易读性,在很大程度上取决于分水岭,合水线,鞍部等骨骼地貌正确与否。选择骨骼地貌时,应细心分辨地貌类型,否则将会造成图形质量不高,模糊不清,使人难以看懂。

地貌可以分成山顶、山谷、马鞍面、坡状面、平面、背脊面、沟壑面七大大种。其中，坡状面又可细分为直倾斜山坡、凸形斜坡、凹形斜坡、台阶状山坡和起伏状山坡五种，而马鞍面又可分为脊背马鞍面和沟豁马鞍面两种。这样合起来有十二种常见地貌特点，见图 4-1。

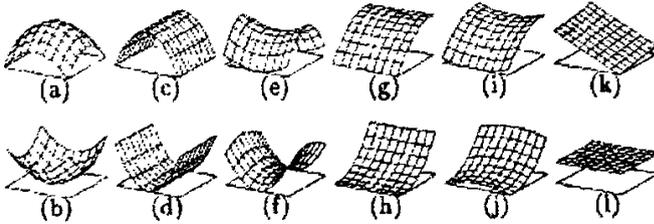


图 4-1 地形学常见的十二种地貌

山顶点描述时会遇到各种不同形状，如尖顶、平顶、圆顶等。它们共性是这些点都是表现为局部最大的那些点，山谷点与山顶点，刚好相反，表现为局部区域内最小的那些点。

马鞍面的特征是一个方向为局部最小值点，另一个方向上表现为局部最大值点的面。背脊面就是一个方向上都是极大值点的而其它方向成坡状的面，沟豁面，刚好相反，是一个方向上成最小其它方向是坡状的面。

平面情形最简单，就是各点相等组成的面。

山坡的表达是构成图的主体，有助于骨骼地貌的突出性与活泼性。山坡地貌大致分为五种类型：

1) 直倾斜山坡，一般在山坡中比较少。只有在冲刷均匀，岩石形同情况下才能遇到，但往往在整个斜坡中只有一段是等斜度变化的。

2) 凸形斜坡，经常能碰到，一般都具有圆滑特征。其中又可分为两种情况：一种是由山顶平缓向下凸出，到中间变陡然后下降；另一种是以较大倾斜向山麓延伸。

3) 凹形斜坡，此种斜坡最为常见。一般特征是由山顶向下倾斜平缓，到一定距离时会急骤下降，在山麓部分以平缓坡度延伸。

4) 台阶状山坡，这种情况占着很重要的地位。多由岩性不同的岩层经冲刷

作用形成,如采用等斜线法描述,会使地貌质量显著降低。

5) 起伏状山坡,也是一种常见地貌,也由岩性不同造成。一般由山顶点,沿着分水线方向连续起伏。可适当采用间曲线来表示,保持地貌的真实性和完美性。

## 4.2 数字地形模型的建立

数字地形模型是把对象区域的平面坐标 $(x, y)$ 和相应点的高度 $(z)$ 的数值用一群地面点描述地表形状的一种方式。这些地面点的投影位置一般按照适当密度的格网式分布,其平面位置和高度可以用摄影测量方法测定或用现有地形采用数字化已输入。

例如,采用方格网法(也有采用三角形法)。这种方法事先选择的对象区域划分成大小相等的方格网,把平面坐标点 $P(x, y)$ 等间隔的规则布置,坐标点(节点)上标高 $(z)$ 从地形图等高线上按一定次序读取,即在三次曲面 $z = f(x, y)$ 的数值中, $(x, y)$ 按方格网有规则的变化,而 $z$ 值作为资料存储,则成为数字地形模型(DTM)。方格网任意节点 $Z_{ij}$ 的标高度为:

$$z_{ij} = f(X_{ij}, Y_{ij}) = Z(I, J) \quad (4-1)$$

其中:  $I, J \geq 0$  的正整数。

而针对我们研究的灰度数字图像没有这么复杂,只是将该图像的灰度值看作该平面坐标 $(x, y)$ 点的高度即可。

这样高度 $z$ 可以表示为:

$$z = f(x, y) = Z(I, J) \quad (4-2)$$

其中,  $I, J \geq 0$  为离散图像的坐标值。

### 4.3 数字地形的数学定义

假设  $z = f(x, y)$  表示地形学表面高度信息, 假设二元函数  $f(x, y)$  在点  $(x, y)$  处的偏导数都存在, 则函数在该点的梯度如下定义:

$$\text{grad}f(x, y) = \nabla f(x, y) = \frac{\partial f(x, y)}{\partial x} \vec{i} + \frac{\partial f(x, y)}{\partial y} \vec{j} \quad (4-3)$$

其中  $\vec{i}$  和  $\vec{j}$  分别是在 X 轴和 Y 轴方向上的单位向量。梯度值可表示为:

$$\|\nabla f(x, y)\| = \sqrt{\left[\frac{\partial f(x, y)}{\partial x}\right]^2 + \left[\frac{\partial f(x, y)}{\partial y}\right]^2} \quad (4-4)$$

若  $\|\nabla f(x, y)\| = 0$ , 则  $(x, y)$  为该函数的临界点, 这就是极值点的一阶必要条件。

设函数  $f(x, y)$  在其临界点  $(x, y)$  的某个邻域内存在二阶的连续偏导数, 其 Hessian 矩阵定义为:

$$H(x, y) = \begin{bmatrix} \frac{\partial^2 f(x, y)}{\partial x^2} & \frac{\partial^2 f(x, y)}{\partial x \partial y} \\ \frac{\partial^2 f(x, y)}{\partial x \partial y} & \frac{\partial^2 f(x, y)}{\partial y^2} \end{bmatrix} \quad (4-5)$$

假设  $w_1, w_2$  是  $H(x, y)$  的两个特征向量,  $\lambda_1, \lambda_2$  分别为它们对应的特征值, 那么有如下判别标准:

1) 若  $\lambda_1 > 0$  且  $\lambda_2 > 0$ , 则临界点  $(x, y)$  为函数  $f(x, y)$  在该领域的极小值点。

山谷点可严格数学定义为: 设函数  $f(x, y)$  在点  $(x, y)$  处存在连续的一阶与二阶偏导数且满足:

$$\|\nabla f(x, y)\| = 0, \quad \lambda_1 > 0 \text{ 且 } \lambda_2 > 0 \quad (4-6)$$

2) 若  $\lambda_1 < 0$  且  $\lambda_2 < 0$ , 则临界点  $(x, y)$  为函数  $f(x, y)$  在该领域的极大值点。

山顶点可严格数学定义为: 设函数  $f(x, y)$  在点  $(x, y)$  处存在连续的一阶与二阶偏导数且满足:

$$\|\nabla f(x, y)\| = 0, \quad \lambda_1 < 0 \text{ 且 } \lambda_2 < 0 \quad (4-7)$$

3) 若  $\lambda_1, \lambda_2 < 0$ , 则临界点  $(x, y)$  为函数  $f(x, y)$  在该领域的鞍点。

鞍点可严格数学定义为: 设函数  $f(x, y)$  在点  $(x, y)$  处存在连续的一阶与二阶偏导数且满足:

$$\|\nabla f(x, y)\| = 0, \lambda_1, \lambda_2 < 0 \quad (4-8)$$

4) 对于山脊的数学表述要复杂得多, 山脊点一个方向上都为极大值, 这个方向上的由极值构成的曲线, 可以是一条水平直线, 也可以是一条向上、向下的直线, 或者向上、向下的曲线。如果这条山脊线不是水平的, 那么它的梯度值就不能为 0, 且在方向  $w_1$  或  $w_2$  有一个局部极大值点。如果这条山脊线是水平的, 这是梯度为 0 了, 其中一个特征值小于 0, 而另外一个等于 0, 严格数学定义如下:

$$\|\nabla f(x, y)\| \neq 0, \lambda_1 < 0, \nabla f \cdot \omega_1 = 0$$

或者

$$\|\nabla f(x, y)\| \neq 0, \lambda_2 < 0, \nabla f \cdot \omega_2 = 0 \quad (4-9)$$

或者

$$\|\nabla f(x, y)\| = 0, \lambda_1 < 0, \lambda_2 = 0$$

5) 沟豁点的数学定义与山脊点相似:

$$\|\nabla f(x, y)\| \neq 0, \lambda_1 > 0, \nabla f \cdot \omega_1 = 0$$

或者

$$\|\nabla f(x, y)\| \neq 0, \lambda_2 > 0, \nabla f \cdot \omega_2 = 0 \quad (4-10)$$

或者

$$\|\nabla f(x, y)\| = 0, \lambda_1 > 0, \lambda_2 = 0$$

6) 平面点的数学定义最简单:

$$\|\nabla f(x, y)\| = 0, \lambda_1 = 0, \lambda_2 = 0 \quad (4-11)$$

坡面的分类比较仔细, 除了上面 6 种的情况, 我们都可以称之为坡面了。本文并不关心具体坡面分类, 详细阐述, 读者可以参考文献<sup>[43]</sup>

## 第五章 基于地形学的人眼定位算法

眼睛定位是人脸识别的关键环节。这是与鼻子，嘴巴等人脸部位相比较，眼睛区域包含了更为丰富的可用于个体区分的重要信息<sup>[44]</sup>。另外，眼睛定位又是面部表情分析和面部位置确定的前提。

眼睛定位分为眼睛位置定位与眼睛轮廓定位两个部分。而大部分眼睛轮廓定位算法都用到了 Yuille 等提出的可变模板法<sup>[45]</sup>。由于模板法的基础就是眼睛位置的定位，所以眼睛位置定位是眼睛定位的关键一步。本文讨论的眼睛定位专指眼睛位置的定位。

一般文献中的眼睛定位算法基于两个前提<sup>[46]</sup>：其一，假设人脸图像已知，通过手工标定或者人脸几何关系标定出包含眼睛的窗口，严格在该窗口上进行眼睛定位。然而，在现实操作中，眼睛窗口的自动标定是非常困难的。其二，运用人脸检测算法粗略提取出人脸的候选位置，眼睛的定位就在确定人脸候选区域上进行的。该方法不足之处也非常明显。首先，眼睛定位的精度依赖于脸部区域检测的鲁棒性；其次，就算检测到了脸相位置，如果不知道脸相的角度，获取眼睛对也是比较困难的事情。

近几年，有学者提出了不需要手工标定脸相位置，或者用某种算法检测出脸相的位置的眼睛定位方法<sup>[47][48]</sup>。在人脸图像中，由于虹膜的像素比较低的灰度特征和成圆形的几何特征，所以人的眼睛有在灰度图像中呈山谷点的地形学特征。

### 5.1 离散化山谷点的数学定义

在人脸图像中，由于虹膜所在位置的像素比较低的灰度特征和成圆形的几何特征，所以人的眼睛的灰度图像中呈现山谷点的地形学特征。

1998年 Kin-Man Lan 等<sup>[49]</sup>利用边界信息和形态学操作构造模板来搜索灰度图像中的山谷点，再对这些山谷点进行聚类来找到眼睛对的位置。

然而由于实际图像存在大量噪音，图像显著极值往往会淹没在大量噪声极值中，所以选择一种抗干扰的离散化算法很有必要。

对离散山谷点数学定义，常见的有采取差分法或者采用切比雪夫多项式方法。但是实践中发现利用这些方法来计算灰度图像的梯度和 Hessian 阵，误差很大，结果也往往不够理想。

对一个标量场  $T(X,Y,Z)$ ，以三维直角坐标系为例，在点  $(X,Y,Z)$  处的梯度定义为：

$$\text{Grad}(x, y, z) = \nabla T := \left[ \frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z} \right] \quad (5-1)$$

对一个向量场  $u(x, y, z) = (u, v, w)$ ，在点  $(x, y, z)$  处的散度定义为：

$$\text{Div}u = \nabla \cdot u = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \quad (5-2)$$

通过奥-高公式可以构造出统一的计算表达式：

$$\nabla F \Big|_{P_0 := (x_0, y_0, z_0)} = \lim_{V \rightarrow P_0} \int_{\partial V} n \cdot F d\delta \quad (5-3)$$

在上式中， $V$  为单连通积分区域； $\partial V$  为  $V$  的边界， $n$  为边界上的外向法矢， $d\delta$  为面积元。

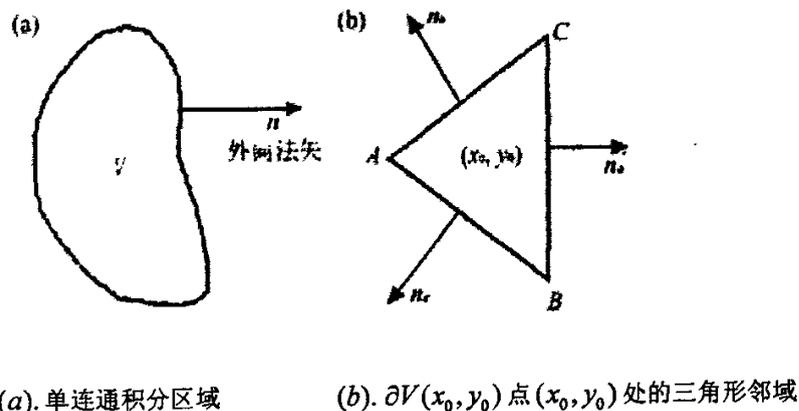


图 5-1 连通区域

上式中的积分计算公式为：

$$\nabla F = \text{Grad}F, \quad n \cdot F = nF \quad (5-4)$$

数字图像是由二维离散的像素点构成，对某一点  $(x_0, y_0)$  的梯度数值计算，一般在该点附近选择一个小邻域。如选取包含  $(x_0, y_0)$  的三角形  $\Delta ABC$ ， $n_a, n_b, n_c$  分别表示三角形外向法矢。根据奥-高公式，可得梯度近似计算公式：

$$\nabla F|_{(x_0, y_0)} \approx \frac{1}{S_{\Delta ABC}} \int_L n \cdot f ds \quad (5-5)$$

其中,  $S_{\Delta ABC}$  表示三角形的面积,  $L: AB \cup BC \cup CA$ 。

而每条边上的积分可以用梯形公式来计算。这样就能够求出梯形的近似表达式:

$$\nabla f(x, y)|_{(x_0, y_0)} \approx -f(x_a, y_a) \cdot \frac{n_a}{h_a} - f(x_b, y_b) \cdot \frac{n_b}{h_b} - f(x_c, y_c) \cdot \frac{n_c}{h_c} \quad (5-6)$$

其中,  $h_a$ ,  $h_b$  和  $h_c$  表示顶点 A, B 和 C 到对应边的距离。

这样计算数字图像的某个像素位置的梯度, 只需要选择包含该像素的小三角形邻域, 该三角形的顶点是由邻近像素组成的。离散化 Hessian 的过程, 可以参考散度离散化的过程<sup>[50]</sup>。

## 5.2 眼睛对的筛选

从标定后的结果可以看出 (如图 4-3 (b)), 基于地形学标定的山谷点来获取候选眼睛对的干扰主要来自嘴角、鼻翼两侧和轮廓比较深的眼角等部位, 而背景干扰主要来自头发, 因为它们有类似山谷点地型特征。我们可以利用眼睛的几何特征及眼睛周围的灰度等信息来消除这些干扰。

### 5.2.1 形态学简述

数学形态学 (Mathematical Morphology) 诞生于 1964 年, 数学形态学的基本思想使用具有一定形态的结构元素去量度和提取图像中的对应形状以达到对图像分析和识别的目的。数学形态学的数学基础和所用语言是集合论, 因此具有完备的数学基础, 这为形态学用于图像分析和处理、形态滤波器的特性分析和系统设计奠定了坚实的基础。数学形态学的应用可以简化图像数据, 保持它们基本的形状特性, 并除去不相干结构。数学形态学的算法具有天然的并行实现的结构, 实现了形态学分析和处理算法的并行, 大大提高了图像分析和处理的速度。

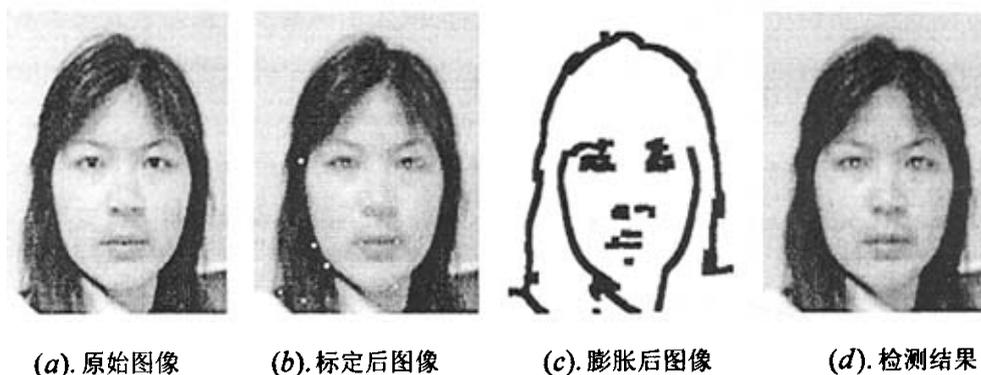


图 5-2 人脸图像定位过程

数学形态学是由一组形态学的代数运算组成。最基本的形态学算子有：膨胀 (expand)、腐蚀 (erosion)、开启(opening)和闭合(closing)，它们在二值图像和灰度图像中各有特点。基于这些基本运算还可推导和组合成各种数学形态学使用算法，用它们可以进行图像形状和结构的分析及处理，包括图像分割、特征提取、边界检测、图像滤波、图像增强和恢复等。

形态学的理论基础是集合论。在图像处理中形态学的集合 (set) 代表着黑白和灰度图像的形状。如黑白二值图像中所有黑色象素点 (pixel) 的集合组成了图像的完全描述。在一个集合中，将被进行形态变换的象素点是被选择的集合  $C$ ，而此集合的补  $C^c$  是没有被选择的集合。通常被选择的集合是图像前景 (foreground)，而未被选择的集合是图像的背景 (background)。

在通常情况下，形态学图像处理以在图像中移动一个结构元素并进行一种类似于卷积操作的方式进行。像卷积核一样，机构元素可以具有任意的大小，也可以包含任意的 0 与 1 的组合。在每个象素位置，结构元素核与在它下面的二值图像之间进行一种特定的逻辑运算。逻辑运算的二进制结果存在输出图像中对应象素的位置上。产生的效果取决于结构元素的大小、内容以及逻辑运算的性质。

腐蚀具有使目标缩小、目标内孔增大以及外部孤立噪声消除的效果；而膨胀是将图像中与目标物体接触的所有背景点合并到物体中的过程，结果是使目标增大、孔径缩小，可以增补目标中的空间，使其形成连通域。数学形态学中另一

对基本运算方法是开运算和闭运算。开运算具有消除图像是细小物体，并在物体影响纤细处分离物体和平滑较大物体边界的作用；闭运算具有填充物体影像内细小空间，连接邻近物体和平滑边界的作用。

### 5.2.2 二值腐蚀与膨胀

数学形态学首先被用来处理二值图像，其基本理论为二值形态学。二值图像是指那些灰度值只能取两个可能值的图像，这两个值通常取为 0 和 1。考虑所有 1 值点的集合 X，则 X 对图像是一一对应的。

如何对集合 X 进行分析？数学形态学认为分析，即是对集合进行变换以突出所需要的信息。其采用的主观“探针”与客观物体相互作用的方法。“探针”也是一个集合，它由我们根据分析的目的来确定。术语上，这个“探针”称为结构元素。选取的结构元素大小及形状不同都会影响图像的处理结果。剩下的问题是如何选取适当的结构元素以及如何利用结构元素对物体进行变换。

为此，数学形态学定义了两个最基本的运算，膨胀和腐蚀。下面我们就对这两个基本运算进行简单地介绍。

#### 1, 腐蚀

腐蚀的数学符号  $\ominus$ ，A 用 B 来腐蚀写作  $A \ominus B$ ，其定义为

$$A \ominus B = \{x | (B)_x \subseteq A\} \quad \text{或} \quad A \ominus B = \{x | B+x \subseteq A\} \quad (5-7)$$

上式中  $(B)_x$  或  $B+x$  均表示将 B 沿矢量 x 平移了一段距离，腐蚀的结果是所有 x 的集合，其中 B 平移后仍在 A 中。如果将 B 看作模板，那么  $A \ominus B$  则由在平移模板过程中，所有可以填入 A 内部的模板的原点组成。图 4-4 和图 4-5 是用结构元素 B 腐蚀 A 的示例。

可以看出，当原点在结构元素内部时，腐蚀具有收缩图像的作用，其输出是输入的一个子集。而当原点不在结构元素的内部时，腐蚀的输出不是输入的子集。

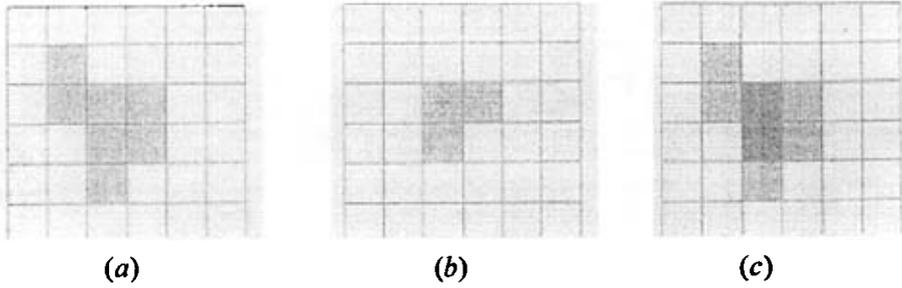


图 5-3 腐蚀运算 (原点包含在结构元素 B 中)

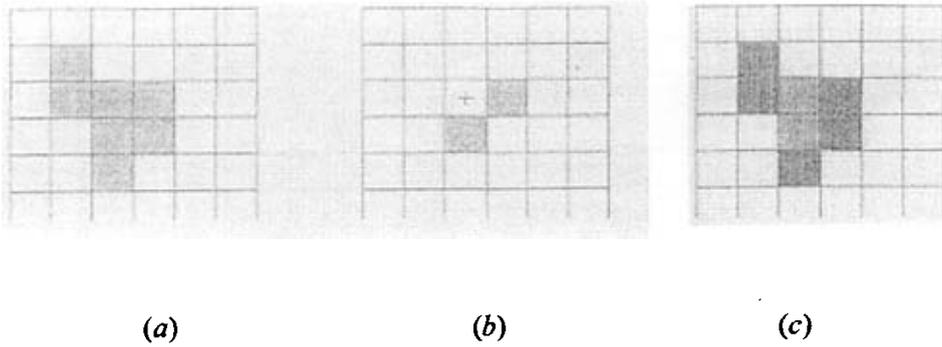


图 5-4 腐蚀运算 (原点不包含在结构元素 B 中)

## 2. 膨胀

膨胀的算符  $\oplus$ , A 用 B 来膨胀写作, 其定义为:

$$A \oplus B = \{x | [(B)_x \cap A] \neq \emptyset\} \quad \text{或} \quad A \oplus B = [A^c - (B)]^c \quad (5-8)$$

式中,  $\hat{B}$  为 B 的映像,  $\hat{B} = \{x | x = -b, b \in B\}$ ,

$A^c$  为 A 的补集,  $A^c = \{x | x \notin A\}$

上式表明用 B 膨胀 A 的过程是先对 B 做关于原点的映射 (即将 B 相对原点旋转 180 度得到  $\hat{B}$ ), 再将其映像平移 X, A 与  $\hat{B}$  的交集非零。即用 B 来膨胀 A 得到的集合是  $\hat{B}$  的位移与 A 至少有一个非零元素相交时 B 的原点位置的集合, 上式表明, A 被 B 膨胀, 即是用  $\hat{B}$  对 A 进行腐蚀。下面图 4-6 和图 4-7 给出了膨胀运算的示例。

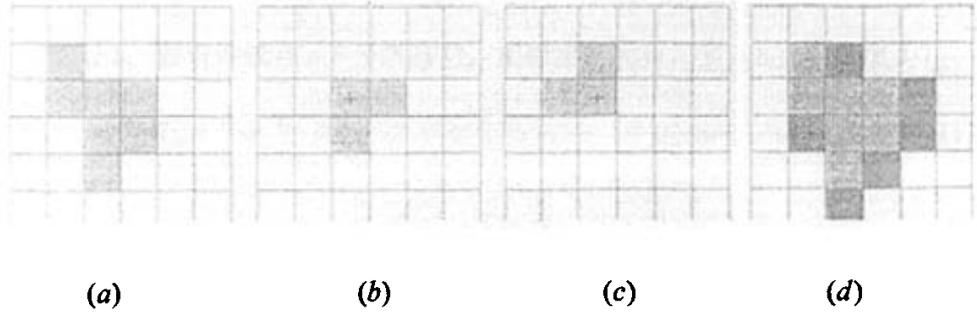


图 5-5 膨胀运算（原点包含在结构元素 B 中）

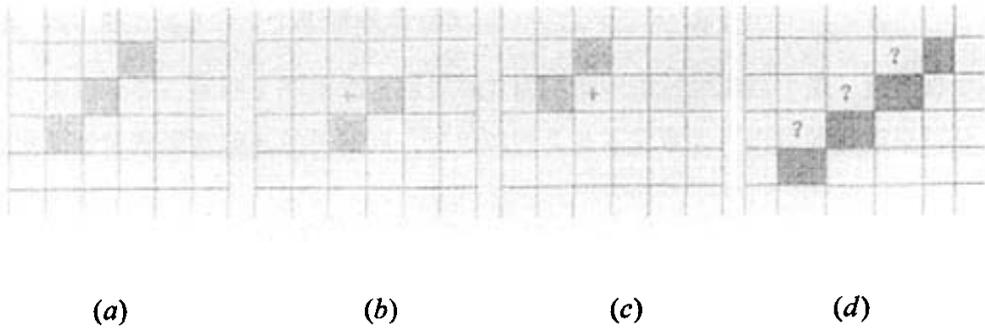


图 5-6 膨胀运算（原点不包含在结构元素 B 中）

### 3. 膨胀和腐蚀的对偶性

膨胀和腐蚀这两种运算是紧密联系在一起，一个运算对图像目标的操作相当于另一个运算对图像背景的操作。其对偶性可表示为：

$$(A \oplus B)^c = A^c \ominus \hat{B} \quad (5-9)$$

$$(A \ominus B)^c = A^c \oplus \hat{B} \quad (5-10)$$

对灰度图像进行处理的灰度形态学算子由二值形态学理论推广而来。

#### 5.2.3 数学形态学眼睛对的筛选

根据脸部生物特征，两只眼睛不可能离得太远或者太近，一般眼睛的距离大约为脸部宽度的 3/5，而鼻梁部位的宽度大约又是眼睛对宽度的 3/5。根据这些几何特征，在包含人脸的图像中，靠得太近或者太远的两候选眼睛点不可能是眼睛对的候选点。另外，若两候选点（如，两候选点相连）距离很近我们将其融合为一个候选点。具体描述如下：

规则 1: 假设候选点  $a, b$  为眼睛对, 图像宽度为  $w$ , 且  $a, b$  点的长度为  $l_{ab}$ ,

$$1) \quad l_{ab} > \frac{3}{5}w \quad (5-11)$$

$$2) \quad l_{ab} < \theta \quad (5-12)$$

其中  $\theta$  为某一个常数, 本文取  $0.1 * w$ ;

若满足 (5-11) 与 (5-12) 中任何一个的候选眼睛对点为假。

规则 2: 假设候选点  $a, b$  为两个眼睛候选点, 若  $a$  为  $b$  的一个 8-连通区域内的, 则, 融合  $a, b$  为一个眼睛候选点  $b$ 。

实验表明, 规则 1 和规则 2 可以很好的降低眼睛对候选点个数。然而对于头发或者背景像素较低的干扰点, 可以通过头发等背景大片出现的特点进行分割来区别。

对分割后的图像, 再用数学形态学知识进行膨胀操作, 显见的好处就是大大突出眼睛区域的灰度信息特征。根据眼睛大小相似性的特点, 膨胀后两眼区域的灰度值为 0 的像素个数应该比较靠近。具体描述如下:

规则 3: 假设候选点  $a, b$  为眼睛对,  $a, b$  点长度  $l_{ab}$ , 设立分别以点  $a, b$  为中心点的  $\frac{l_{ab}}{5} * \frac{l_{ab}}{5}$  的模板, 计算各自模板内灰度为 0 的点的个数  $L_{num}$  和  $R_{num}$ , 若满足,

$$|L_{num} - R_{num}| < \delta \quad (5-13)$$

其中  $\delta$  为某一常数, 本文取  $0.2 * l_{ab}$ , 则  $a, b$  不是眼睛对。

通过规则 2 可以排除眼睛对的背景点干扰。根据眼睛区域的几何特征, 大部分人眼睛外形接近于圆形。在膨胀后的图像中, 分别计算出剩余候选点在各自模板内灰度为 0 的点的个数占模板面积的比重。眼睛区域的比重一般比较大。

规则 4: 假设候选点  $a$  为眼睛对的一支, 设立以点  $a$  为中心点的  $\frac{w}{10} * \frac{w}{10}$  的模板, 计算模板内灰度为 0 的点的个数  $a_{num}$ , 若满足:

$$|a_{num} / (\frac{w}{10} * \frac{w}{10})| < \rho \quad (5-14)$$

其中  $\rho$  为一常数 (本文取了 0.75), 则眼睛对点为假。

对于鼻翼两侧及眼角干扰点, 运用规则 3 可以取得比较满意的实验效果。

经过上述四个筛选规则对候选眼睛对点进行精简, 剩下的眼睛对点一般就是我们需要寻找的那些节点。为了保险起见, 下面一节中将构造积分、方差投影规则对这些候选节点做进一步判断验证。

## 5.3 筛选后眼睛对的验证

由于眼睛区域具有两个显著的特征<sup>[51]</sup>: 其一, 眼睛区域与周围区域相比较暗, 即灰度值较低; 其二, 眼睛区域的灰度变化率较大。所以, 本文将综合考虑这两个特征来精确验证眼睛位置的。

### 5.3.1 投影函数知识简述

投影是一种有效提取图像特征的方法, 通常一幅二维图像可以由两个正交的一维投影函数来分析。维数的降低便于分析图像的特征, 并且减少计算量, 所以投影成为一种重要的图像分析方法。

到目前为止, 已有很多学者将投影函数成功运用于定位面部特征。Kande<sup>[52]</sup>最早将积分投影函数成功地运用于人脸识别, 他首先用拉普拉斯算子对原始灰度图像进行二值化, 然后用积分投影函数对 0-1 图像进行分析。Brunelli 及 Poggio<sup>[53]</sup>对 Kande 的算法作了改进, 他们将积分投影函数应用于边界图分析, 从而确定出面部各个特征的位置。

方差投影概念最早是由 Feng 和 Yuan<sup>[54]</sup>提出, 并同时提出了一种利用方差投影函数定位眼睛的简单方法。后来, 他们又提出了一种多线索定位眼睛的方法, 其中应用了一种眼睛方差过滤器, 而这种过滤器正是利用方差投影函数产生的。由此可以看出, 投影实际上是人脸识别中经常采用的一种定位技术。

### 5.3.2 积分投影函数

最为常用的投影函数是积分投影函数。假设  $I(x, y)$  表示点  $(x, y)$  处的像素灰

度值，在区间 $[x_1, x_2]$  和  $[y_1, y_2]$ 内的垂直积分投影和水平积分投影函数分别表示为 $S_v(x)$  和  $S_h(y)$ ，则

$$S_v(x) = \int_{y_1}^{y_2} I(x, y) dy, \quad (5-15)$$

$$S_h(y) = \int_{x_1}^{x_2} I(x, y) dx \quad (5-16)$$

平均积分投影函数 $M_v(x)$  和  $M_h(y)$ 分别表示为

$$M_v(x) = \frac{1}{y_2 - y_1} \int_{y_1}^{y_2} I(x, y) dy, \quad (5-17)$$

$$M_h(y) = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} I(x, y) dx \quad (5-18)$$

当图像某一列像素灰度均值发生变化时，这种变化会在 $S_v(x)$ 和 $M_v(x)$ 的值上反映出来。同理，当图像某一行像素灰度均值发生变化时，这种变化会在 $S_h(y)$ 和 $M_h(y)$ 的值上反映出来。

因此，可以通过对积分投影函数值得分析提取图像中的特征。例如下图 4-8 中，在水平方向上 3 个不同区域的铰接处，积分投影函数对应的曲线发生了跃变，从而准确反映出了图像中的区域特征。

### 5.3.3 方差投影函数

积分投影函数只考虑了均值的变化，却不能反映均值相同情况下灰度值得变化。如果用方差代替均值，就可以避免这种情况发生，这就是方差投影函数。设在区间 $[x_1, x_2]$  和  $[y_1, y_2]$ 内的垂直方差投影和水平方差投影函数分别表示为 $\delta_v^2(x)$  和  $\delta_h^2(y)$ ，则

$$\delta_v^2(x) = \frac{1}{y_2 - y_1} \sum_{y_1}^{y_2} [I(x, y_i) - M_v(x)]^2, \quad (5-19)$$

$$\delta_h^2(y) = \frac{1}{x_2 - x_1} \sum_{x_1}^{x_2} [I(x, y) - M_h(y)]^2 \quad (5-20)$$

其中,  $M_v(x)$  和  $M_h(y)$  有 4.42 节中公式得到。

当图像某一列像素灰度的方差发生变化时, 这种变化会在  $\delta_v^2(x)$  的值上反映出来。同理, 当图像某一行像素灰度的方差发生变化时, 这种变化会在  $\delta_h^2(y)$  的值上反映出来。

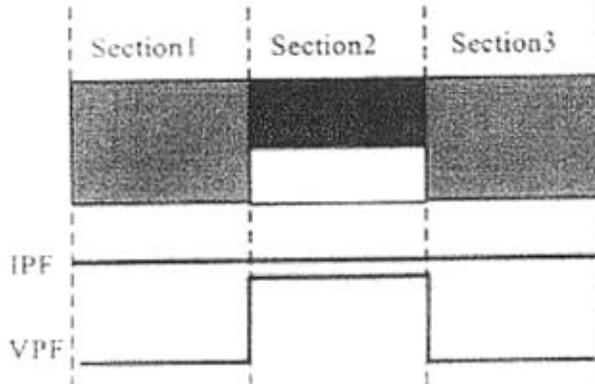


图 5-7 积分投影在跃变处曲线

因此, Feng 和 Yuan 还证明了 VPF 对随机噪音并不敏感。因此, 同样可以利用 VPF 对图像特征进行分析。在上图中, 垂直 VPF 对应在不同区域边界处发生了较大的跃变, 从而反映出了图像在水平方向上的区域变化。

#### 5.3.4 眼睛对的验证

经过上述处理后, 剩下的干扰点只可能存在于眉毛、嘴角和鼻尖处了。我们可以充分利用眼睛部位的对称特征来最后确定眼睛对。

首先必须要得到一幅仅包含眼睛的图像, 即候选眼睛对窗口。假设  $a$  和  $b$  是候选的眼睛对的两个点,  $a, b$  中点为  $o$ , 且距离为  $l_{ab}$ 。以  $o$  为中心点, 沿垂直于  $l_{ab}$  的方向, 截取  $1.6l_{ab} \times 0.3l_{ab}$  大小的眼睛对窗口, 这样取出来的窗口, 包含眼睛的全部信息, 且眼角到窗口左右边缘的距离约为眼睛长度的一半。

在每一个候选眼睛对窗口, 进行垂直方向的投影, 如果  $a, b$  周围灰度信息相差不大, 而且出现必要的谷峰值点, 保留该候选对。这样就可以排除眼睛点与嘴角, 鼻尖等处配对以及它们内部配对造成的干扰。

当然上述过程都是在平滑后的图像上进行操作的。眼睛窗口典型的标准积分，方差垂直方向上的投影函数如图二所示。我们可以清楚地看出，IPF 曲线上的有成对出现的峰值点，则对应到 VPF 曲线上就为一对称的谷值点，而 IPF 谷值点与 VPF 峰值点应该与鼻梁位置相吻合<sup>[53]</sup>。

将窗口划分为，左眼 $[0, 1.6 * l_{ab} / 3)$ ，鼻梁 $[1.6 * l_{ab} / 3, 1.6 * l_{ab} / 2)$ ，右眼 $[1.6 * l_{ab} / 2, 1.6 * l_{ab}]$ 三个区间。因为眼睛对的像数一般比较少，在各自小区间里用逐点比较法，分别计算出 IPF, VPF 函数的最值点横坐标  $L_l, M_l, R_l$  和  $L_v, M_v, R_v$ 。这些点坐标信息与我们地形学标定后的候选眼睛点  $a, b$ ，及眼睛对中心点  $o$  位置相比较，如果三点在一定邻域内，则眼睛对得到验证。

验证规则描述如下：

假设  $a, b$  为候选眼睛对的左右点， $o$  为  $a, b$  的中点， $L_l, M_l, R_l$  和  $L_v, M_v, R_v$  分别为 IPF 和 VPF 函数的三个最值点的横坐标，若满足，

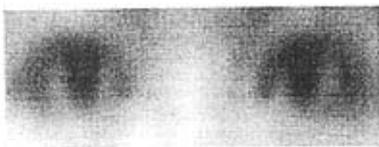
$$|L_l - a| + |L_v - a| < \varepsilon \quad (5-21)$$

$$|R_l - o| + |R_v - o| < \varepsilon \quad (5-22)$$

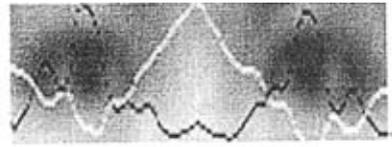
$$|M_l - b| + |M_v - b| < \varepsilon \quad (5-23)$$

其中  $\varepsilon$  为某一常数，本文取  $0.2 * l_{ab}$ ，则眼睛对为真。

这样就确认了被分析的窗口为眼睛对窗口，而标定的候选眼睛点就是要找眼睛点，从而达到了验证定位的目的。



(a).  $1.6l_{ab} \times 0.3l_{ab}$  眼睛窗口



(b). IPF 曲线 (白) 与 VPF 曲线 (黑)

图 5-8 候选眼睛对的验证

## 第六章 眼睛定位系统设计与实现

根据文章提出的基于地形学标定的眼睛定位算法，作者简单设计了一个基于 CxImage 实验平台的眼睛定位系统。该系统由三个模块构成：第一模块为灰度图像的读入与显示过程；第二个模块为眼睛对的标定、精简和验证过程；第三个模块为最终结果输出过程。其中第二模块为整个系统的关键部分，下面各章节都对模块二进行重点性描述。

另外，本章对系统的设计思路、实验环境等作了简单的介绍。在最后一节中作者还详细描述了系统验证的整个过程。所采用的人脸数据集分为：自建的单面孔、统一背景的人脸数据库、BioID 人脸库和网络上下载的多脸孔数据库。并对统计出来的三种不同实验结果，作了简要分析。

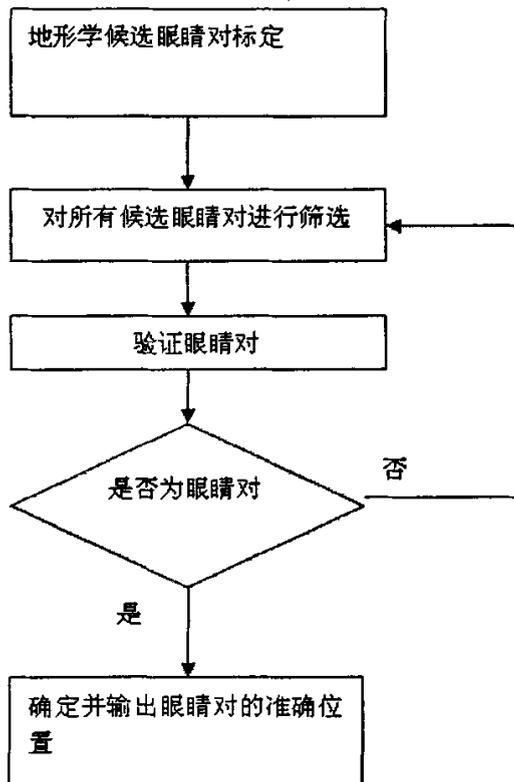


图 6-1 眼睛精确定位整体条件流程图

## 6.1 系统设计

眼睛对的定位过程分为三个步骤，

1) 基于地形学候选眼睛对定位过程。这个过程根据地形学山谷点的特性，在灰度图像上，找到所有呈现该特征的点集。由于眼睛点明显的山谷点特性，如果该灰度图像中有人脸，眼睛点一定就在这个候选点集合当中。

2) 基于形态学的眼睛对筛选过程，该过程主要目的是删除出冗余点，进一步精确检测出眼睛的位置。用到了灰度图像的膨胀操作，构造四个筛选规则来排除掉干扰点。

3) 眼睛对的精确定位的验证过程，该过程主要目的是进一步精确定位及验证出眼睛对。由于眼睛对的明显几何特性，用投影学知识来验证前两步的找到的候选眼睛对精确位置。

检测后得到的图像如图 5-2 (d) 所示。

实验开始需要对灰度图像进行归一化操作，对单一背景的人脸我们全部归一化成解析度为 283\*378 的标准图像。另外，在离散化山谷点定义之前，我们对灰度图像进行平滑处理来降低了噪声的干扰。文中选择了大小为 17\*17， $\delta=3.0$  的高斯滤波器对归一化的图像进行了平滑处理。所有候选眼睛对的位置，我们是用二维数组 A 来存储的。

在眼睛对筛选的步骤当中，我们用形态学的膨胀操作对二值图象进行了处理，选择膨胀算子为 10\*10 的模板，筛选结果用二维数组 B 进行存储的。最后经过验证的最后眼睛对精确位置存储在二维数组 C 中。

本文设计和验证上面算法的实验环境为：操作系统：WinXp，处理器：P4 2.0G，双 CPU，内存：504M。

对单幅图像的具体操作过程（如图 6-2 所示）。

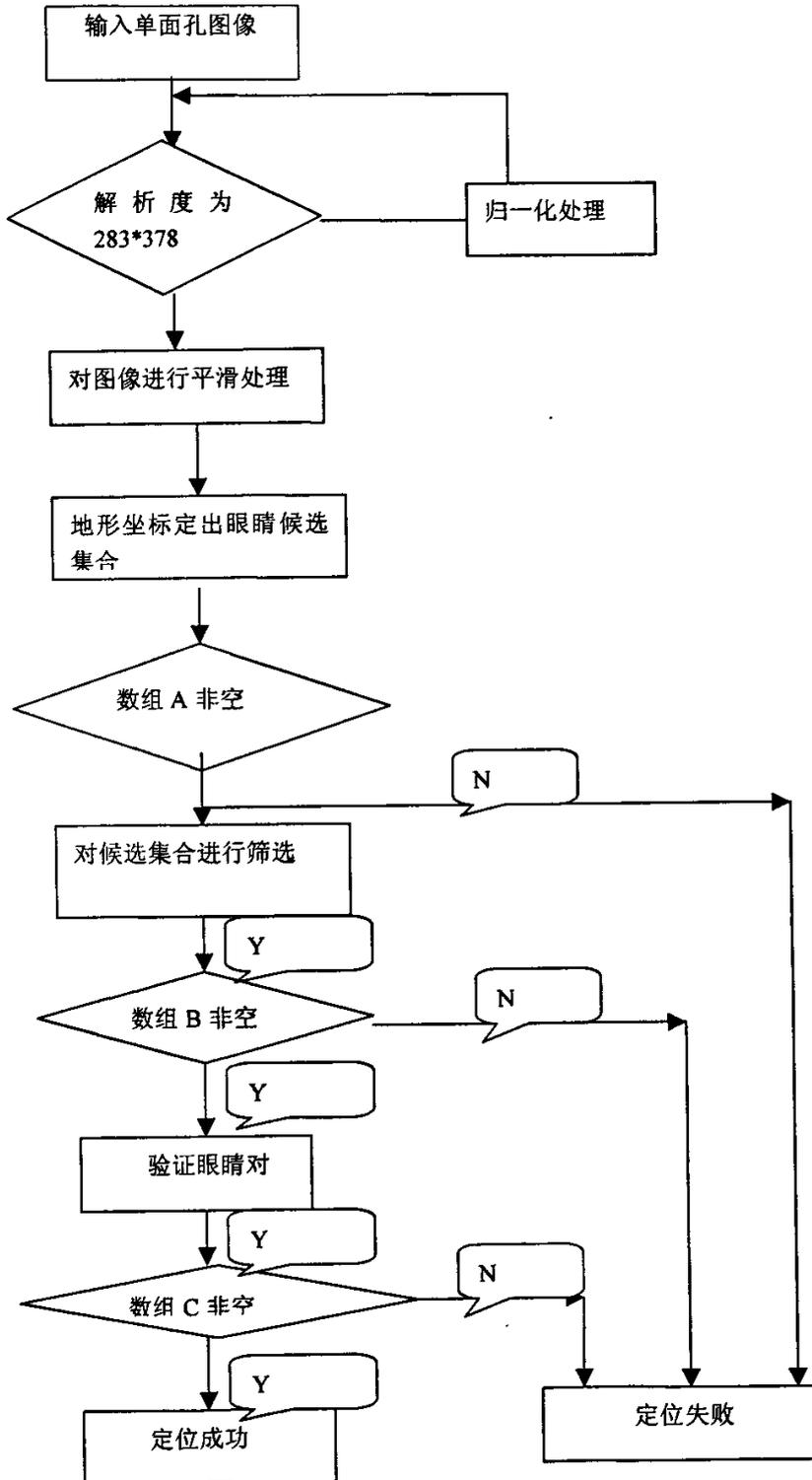


图 6-2 单一备选区域检测流程

## 6.2 VC++实现关键代码

本文试验是基于 CxImage 平台的，用 VC 编码实现了所有功能。部分函数的编写，我们还调用了 ITK 通用函数库中的部分经典函数，例如图像高斯平滑、Sobel 算子分割和形态学膨胀等。另外，实验数据集处理的一些前期工作，我们是直接调用了 MATLAB 中相关函数来完成的，如图像尺寸的归一化等操作。

为了读者更好的了解本文设计的眼睛定位系统，下面我们将对 CxImage 和 ITK 函数库做个简单介绍。其中，CxImage 是一个可以用于 MFC 的 C++类，可以打开，保存，显示，转换各种格式的图像文件。可以实现 BMP<->JPG，PNG<->TIFF 格式等等的转换。其中，CxImage 类库包括了很多优秀的图像操作类，通过它自身所带的 DEMO 界面，可以直接调用这些函数对打开的图像进行处理。与 OpenIL、FreeImage 和 PaintLab 等功能强大的图像库不同，这个类库是完全公开了源代码的免费实验平台。相关内容，读者可以参考文献<sup>[55]</sup>。

另外，ITK 是关于图像处理，分割，配准的开发包，和三维重建方面的功能的函数库。完全开源的代码，类似于 Snake 分割算法等许多优秀的函数，我们都可以直接调来用。但是 ITK 没有显示界面，不能满足许多图像处理研究人员所见即所得的实验需要。在综合考虑 CxImage 显示的优势和 ITK 强大的函数功能后，本文作者改写了 CxImage 函数库中的部分代码，使其可以直接调用 ITK 函数库中的部分函数。对 ITK 函数库感兴趣的读者可参考文献<sup>[56]</sup>。

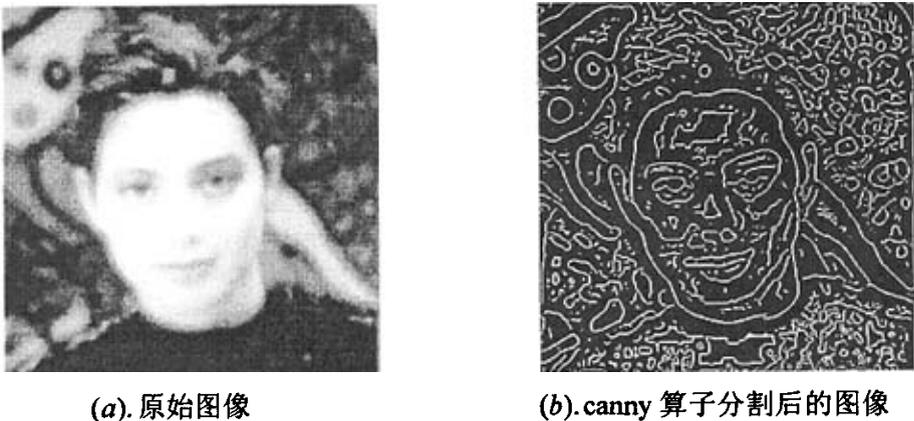


图 6-3 调用 ITK 函数示例

## ■ 图像处理部分

1) 用 MATLAB 函数中 `imresize` 对所有图像进行规一化操作。

2) 图像读入函数:

打开 `CxImage` ->`File`->`Open` 直接打开你需要的图片;

读入的图像存放于一个全局二维数组中: `unsigned int ** dimage`。

3) 图像显示函数:

打开后图像展示在 `CxImage` 平台上, 可以通过指针 `dimage` 进行相关操作。

4) 图像平滑函数:

输入为: 二维数组 `unsigned int ** dimage`;

`bool CDemoDoc::DiscreteGaussianImageFilter();` //对图像进行高斯平滑操作;

输出为: 二维数组 `unsigned int ** dimage`。

## ■ 眼睛定位部分:

1) 候选眼睛对的标定函数:

输入为: 二维数组 `unsigned int ** dimage`;

`bool CDemoDoc::TestTopographic();`//逐一对 `dimage` 中的像素点进行地形学山谷点标定;

输出为: 二维数组 `A`。

2) 候选眼睛对的精简函数:

输入为: 二维数组 `A`, 以及二维数组 `unsigned int ** dimage`;

`bool CDemoDoc::RemoveImageNoise();` //逐一对二维数组 `A` 中的点进行排除;

输出为: 二维数组 `B`。

3) 眼睛对的精确定位:

输入: 二维数组 `B`, 以及二维数组 `unsigned int ** dimage`;

`bool CDemoDoc::DrawMeanAndVarProject();` //逐一对二维数组 `B` 中的点进行眼睛对验证;

输出: 二维数组 `C`。

## ■ 图像输出部分:

`bool CDemoDoc::TestResult();` //用十字架标示出眼睛部位并将结果图像保存在文件夹中的函数。

## 6.3 程序界面

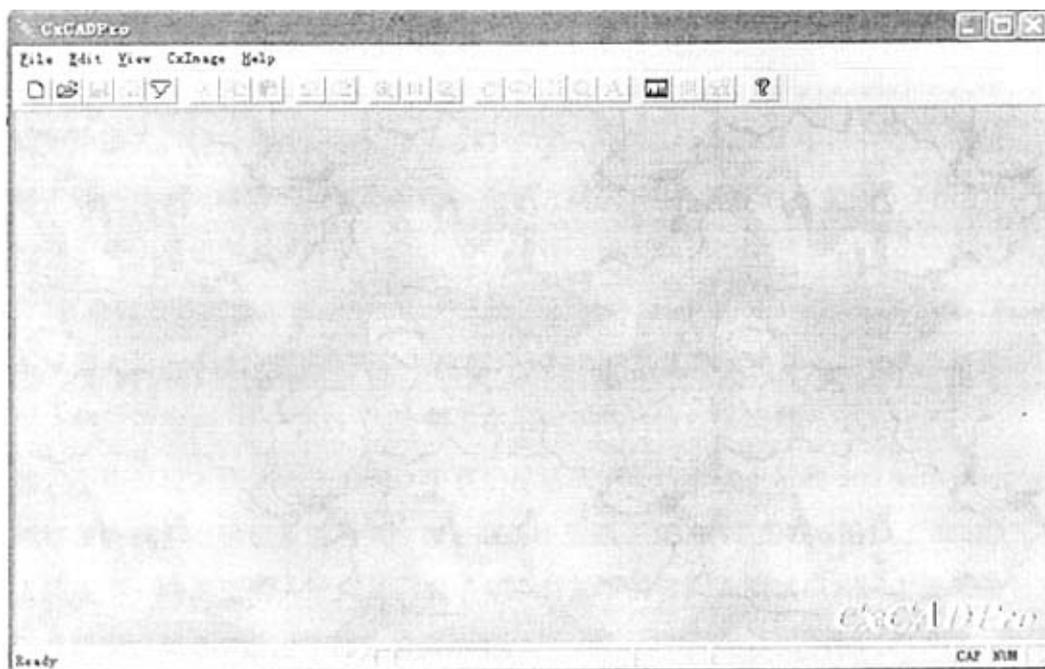


图 6-4 CxImage 平台程序界面



图 6-5 打开图片后的 CxImage 程序界面

## 6.4 实验结果及分析

实验采用单一背景和复杂背景两种不同的数据集。其中单一背景数据库为自建数据库，共包含 240 张不同人物的背景单一的正面灰度照片，照片中人物均为中国人，每张照片的解析度为  $283 \times 378$ 。在这 240 幅图像中，有 229 幅能够正确地找到眼睛对的位置，准确率能够达到 95.4%，操作过程如图 5-2 所示，部分实验结果如图 6-5 所示。

复杂背景数据库来源于 BioID<sup>[15]</sup>人脸库及网络。其中 BioID 人脸库共包含 23 个不同人物的 1521 张正面人脸灰度照片，照片中的人物大部分为欧美人，每张照片的解析度为  $384 \times 268$ 。人脸图像在不同季节，地点采集。光线，背景，人脸大小均有较大变化。在 BioID 中我们挑选其中眼睛形状正常的 300 幅图像，能够正确找到眼睛位置的有 279 幅，准确率达到了 93.0%。图 6-6 给出了 BioID 人脸库中一些典型的人脸照片。对于复杂背景下的多面孔测试照片，主要采集于互联网，图 6-7 给出了部分检测结果。



图 6-6 部分自建数据库检测结果



图 6-7 部分 BioID 数据库检测结果



图 6-8 部分多脸孔图像检测结果

## 第七章 总结及后续工作

### 7.1 总结

针对基于地形学标定在人眼定位过程中的应用，本文的主要研究工作和成果如下：

1) 通过研究图像成像过程，虹膜对光源敏感性的光学原理，表现在灰度图像上眼睛中心位置的灰度值要比其邻域灰度值要高。在地形学上表现为明显的山谷点的地貌特征。用这种方法提高了眼睛对提取的精度和鲁棒性。在离散化山谷点数学定义过程中，文章巧妙地运用奥斯特罗格拉茨基—高斯公式来快速获取离散图像山谷点，缓和了直接从灰度图像中提取物体特征的速度问题。

2) 在精简候选眼睛对集合的过程中，我们采用了经典的数学形态学膨胀操作。在文章中我们根据眼睛区域几何与灰度特征，通过大量的实验提出了一些筛选规则，从而提高了精简候选眼睛对的效率和精度。

3) 通过仔细研究眼睛对的特征，我们发现眼睛区域积分投影和方差投影呈现一定的位置特征。在验证最后眼睛对的过程中，构造了积分投影的极大值点、方差投影的极小值点和候选眼睛点三者位置在一定容忍区域之内的验证规则，巧妙地达到了眼睛对验证的目的。

由于算法并不是基于单幅正面图像进行设计的，所以该算法对图像的偏移以及光照有较好的容忍性。实验表明该算法有较高的定位成功率。

### 7.2 后继工作

人脸检测问题的内涵十分广泛，已有的方法一般都是针对某一类问题提出的。由于人脸检测问题的复杂性，实现通用的人脸检测方法还不现实，因此解决特定约束条件下或某种应用背景下的人脸检测问题仍将是该领域研究的主要课题。

以二维图像作为输入的人脸识别方法的最大不足是方法较为脆弱，易受姿态、光照、表情的影响：如光照的方向、强弱等的影响对识别方法干扰通常较大；人的自然姿态在空间中的偏斜将导致所获得的人脸图像差别很大；人的面

部会表现出各种不同的表情,让现有的二维人脸识别算法在不同的表情下进行正确的人脸识别相当困难。产生这些问题的主要原因是人脸本身是三维的,而照片是对三维面貌进行平面投影的结果,在此过程中必然丢掉一部分重要信息。近年来,在3D人脸建模领域的快速进展,使得获取三维人脸数据变得更容易,这为3D人脸检测与特征定位提供了基础。3D数据真实反映了对象在三维空间中的形状,若是3D深度数据,数据还表征了对象的实际尺寸。通过基于3D数据的人脸检测与定位,有望克服目前基于图像的人脸检测与定位方法所遇到的困难。

人眼定位由于在人脸检测与人脸识别等面向分析领域的关键作用,越来越多的学者致力于该方向的研究。但实际中由于受到人的行为习惯,生理因素等的影响,有时候采集到的人脸图像并不是完全正立的,而是在竖直平面内存在一定的倾角,这就增加了眼睛正确定位难度。

基于地形学标定的方法,在人脸领域应用比较少。根据眼睛周围的特殊灰度信息,提出的山谷点数学定义,如何运用到离散情况,另外,如何平衡平滑模板的大小,才能保证眼睛处的山谷点特征不至于消失等等,以及能不能运用到人脸除眼睛以外的其它特征定位,都是一个值得研究的课题,也是以后的研究工作的重点。

## 参考文献

- 【1】. William A. Barrett, A Survey of Face Recognition Algorithms and Testing Results, National Biometrics Test Center San Jose State University One Washington Square San Jose, CA 95192-0180 IEEE,1998,301~305.
- 【2】. Sung K, Poggio T. Example-based learning for view based human face detection. IEEE Trans Pattern Analysis and Machine Intelligence, 1998, 20(1):39~51.
- 【3】. Yang MH, Roth D, Ahuja N. A SNoW-Based Face Detector. Advances in Neural Information Processing Systems 12, MIT Press, 2000:855~861.
- 【4】. Yang . G and Huang. T. S. Human Face Detection in Complex Background. Pattern Recognition, 1994, 27(1):53~63.
- 【5】. Jeng S H, Liao H Y M, Han C C. Facial Feature Detection Using Geometrical Face Model: An Efficient Approach[J]. Pattern Recognition, 1998:31:273~282.
- 【6】. Bruneli R, Poggio T. Face recognition: Features versus templates. IEEE Transactions on pattern analysis and machine intelligence, 1993, 15(10): 1042~1052.
- 【7】. Govindaraju V , Srihari S N , SherD B. A computational model for face location. P roc IEEE Conference on Computer Vision, Osaka, Japan, 1990, 718~721.
- 【8】. Viola P, Jones M. Rapid Object Detection Using a Boosted Cascade of Simple Features[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, 1:511~518.
- 【9】. Viola P, Jones M. Robust real time object detection Technical Report. CRL 2001/01, Compaq Cambridge Research Laboratory, February 2001.
- 【10】. FERET 人脸库,<http://www.nist.gov/srd/>.
- 【11】. Yale 人脸库,[http://cobweb.ecn.purdue.edu/~aleix/aleix\\_face\\_DB.html](http://cobweb.ecn.purdue.edu/~aleix/aleix_face_DB.html).
- 【12】. UMIST 人脸库. <http://images.ee.umist.ac.uk/danny/database.html>.
- 【13】. BioID 人脸库,<http://www.bioid.com/downloads/facedb/>.
- 【14】. 章毓晋.图像分割.北京:科技出版社,2001.
- 【15】. 冈萨雷斯 (美)等.数字图像处理(英文版).北京:电子工业出版社,2002.
- 【16】. 边肇祺,张学工等.模式识别.北京:清华大学出版社,1999.

- 【17】. 中科永联模式识别培训文章.<http://www.itisedu.com/phrase/200603021030465.html>.
- 【18】. 秦志远, 吴冰, 王艳, 山海涛. 图像平滑算法比较研究及改进策略[M]. 测绘学院学报, 2005, 22(2): 103~106.
- 【19】. 黄涛. 数字图像的平滑. 韶关学院学报, 2004, 25(6): 25~29.
- 【20】. 刘红霞. 图像分割算法的研究和实现. 华东师范大学, 2004 届硕士学位论文.
- 【21】. 张超. 智能图象分割方法研究. 西南交通大学, 2005 届硕士学位论文.
- 【22】. 高秀娟. 图象分割的理论方法及应用. 吉林大学, 2005 届硕士学位论文.
- 【23】. 丁莉, 张琦. 简捷图像分割研究. 现代电视技术, 2004, (12): 140~144.
- 【24】. McGraw Hill. 机器学习. 机械工业出版社. 2003 年.
- 【25】. 朱夏军. 人脸识别的研究及眼睛定位算法. 浙江工业大学, 2003 届硕士学位论文.
- 【26】. Tsuyoshi Moriyama, Takeo Kanade, Jing Xiao, Jeffrey F. Cohn: Meticulously Detailed Eye Region Model and Its Application to Analysis of Facial Images. IEEE Trans. Pattern Anal. Mach. Intell. 2006, 28(5): 738~752.
- 【27】. 王磊, 莫玉龙, 戚飞虎. 基于霍夫变换和眼睑弹性模板的眼睛特征提取[J]. 红外与毫米波学报, 1999, 18(1): 53~60.
- 【28】. Marini R. Subpixellic Eyes Detection. IEEE International Conference on Image Analysis and Processing, 1999, 496~501.
- 【29】. 李粉兰, 徐可欣. 一种应用于人脸正面图像的眼睛自动定位算法. 光学精密工程, 2006, (14): 320~326.
- 【30】. 徐艳, 陈孝威. 人脸检测中的眼睛定位算法研究. 计算机与信息技术, 2006, 24-28.
- 【31】. 山世光, 高文, 陈熙霖. 基于纹理分布和可变形模板的面部特征提取[J]. 软件学报, 2001, 12(4): 570~577.
- 【32】. 尹方平, 赖剑煌, 阮邦志. 非均匀光照下人脸眼睛的定位方法. 中山大学学报, 2003, 42(3): 12~15.
- 【33】. 解明, 马泳. 基于对称变换的人脸图像眼睛定位方法[J]. 光学技术, 2004, 30(2): 237-239.
- 【34】. Reisfeld D, Yeshurun Y. Preprocessing of face images: detection of features and pose normalization[J]. IEEE Trans. Computer Vision and Image Understanding, 1998,

- 71(3):413~420.
- 【35】. 周杰, 卢春雨, 张长水, 李衍达. 基于方向对称变换的人脸定位方法. 电子学报, 1999, 27(8):12~15.
- 【36】. 刘文予, 潘峰. 离散对称变换在人脸图像眼睛定位中的应用[J]. 红外与毫米波学报, 2001, 20(5):375~380.
- 【37】. 陶亮, 庄镇泉. 复杂背景下人眼自动定位[J]. 计算机辅助设计与图形学学报, 2003, 15(1):38~42.
- 【38】. Hsu R.L, Abdel-Mottaleb M, Jain K.L. Face detection in color images. Pattern Analysis and Machine Intelligence, IEEE Transactions, 2002, 24(5):696~706.
- 【39】. Wang Jun, Yin Lijun. Eye detection under unconstrained background by the terrain feature. Proceeding of IEEE International Conference on Multimedia and Expo, 2005, 1528~1531.
- 【40】. Yanowitz S.D, Bruckstein A.M, "A new method for image segmentation", CVGIP, 1989, 46(4):82~95.
- 【41】. Westall J. M, Narasimha M. S. Vertex directed segmentation of handwritten numerals. Pattern Recognition, 1993, 26(10):1473~1486.
- 【42】. Trier O.D, Taxt Torfinn and Jain A. K. Data capture from maps based on gray scale topographic analysis. Proc.of Third Int. Conf. on Document Analysis and Recognition, 1995, 923~926.
- 【43】. Lam K.M. A fast approach for detecting human faces in a complex background. Proc Symposium on Circuits and Systems, Monterey, 1998, 4:85~88.
- 【44】. 宋加涛, 刘济林, 池哲儒, 王 蔚. 人脸正面图像中眼睛的精确定位. 计算机辅助设计与图形学学报, 2005, 17(3):540~545.
- 【45】. Yuille A.L, Hallinan P.W, Cohen D.S, Feature extraction from faces using deformable templates, Int .J. Computer Vision, 1992, 8(2):99~111.
- 【46】. Peng Wang, Green M.B, Qiang Ji, Wayman J. Automatic Eye Detection and Its Validation. Proc.of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 3:164~164.

- 【47】. Cristinacce D and Cootes T. Facial feature detection using AdaBoost with shape constraints. Proc.of 14<sup>th</sup> British Machine Vision Conference, 2003,231~240.
- 【48】. Kothari R and Mitchell J. Detection of eye location in unconstrained visual images, in IEEE International Conference on Image Processing, 1996,519~522.
- 【49】. Lam K.M. A fast approach for detecting human faces in a complex background. In: Proc Symposium on Circuits and Systems, Monterey, 1998, 4:85~88.
- 【50】. 杨朝霞, 逯峰, 李岳升. 图像梯度与散度计算及在边缘提取中的应用. 中山大学学报, 2002,41(6):6~9.
- 【51】. 耿新, 周志华, 陈世福. 基于混合投影函数的眼睛定位. 软件学报, 2003, 14(8):1394~1400.
- 【52】. Kanade T. Picture processing by computer complex and recognition of human faces [Ph.D. Thesis]. Kyoto: Kyoto University, 1973.
- 【53】. Brunelli R, Poggio T. Face recognition: Features versus templates. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1993,15(10):1042~1052.
- 【54】. Feng GC, Yuen PC. Variance projection function and its application to eye detection for human face recognition. Pattern Recognition Letters, 1998,19(9):899~906.
- 【55】. CxImage 函数库代码下载地址. <http://www.codeproject.com>.
- 【56】. ITK 函数库. <http://www.itk.org>.

## 致 谢

谨以此文献给所有关怀和帮助过我的人。

本文是在导师丁友东教授的悉心指导下完成的。承蒙丁老师的亲切关怀和精心指导，并给我提供了良好的学习环境，使我从中获益不浅。丁老师对学生认真负责的态度、严谨的科学研究方法、敏锐的学术洞察力、勤勉的工作作风以及勇于创新、勇于开拓的精神是我永远学习的榜样。在此，谨向丁老师致以深深的敬意和由衷的感谢。

感谢同实验室的其他博士生和硕士生对我研究工作的支持和帮助，以及同寝室同学在生活中对我的关心和帮助。

感谢我的父母，他们在生活上给予我很大的支持和鼓励，是他们给予我努力学习的信心和力量。

最后，感谢所有关心我、支持我和帮助过我的同学、朋友、老师和亲人。在这里，我仅用一句话来表明我无法言语的心情：感谢你们！

## 作者攻读学位期间公开发表的论文

- [1].基于地形学标定的人脸眼睛定位[J]. 计算机应用与软件. 2007, (已录用, 待发表)

## 附录 A 图表索引

图 1-1 人脸检测的基本框架 .....	3
图 2-1 模式识别系统 .....	12
图 2-2 人工神经网络的计算模型.....	17
图 2-3 具有隐含层的神经网络.....	18
图 3-1 眼睛区域二维的多层模型.....	20
图 3-2 眼睛的变形模板 .....	22
图 3-3 三维眼睛模板 .....	22
图 3-4 对称的例子 .....	24
图 4-1 地形学常见的十二种地貌.....	29
图 5-1 连同区域 .....	34
图 5-2 人脸图像定位过程 .....	36
图 5-3 腐蚀运算（原点包含在结构元素 B 中） .....	38
图 5-4 腐蚀运算（原点不包含在结构元素 B 中） .....	38
图 5-5 膨胀运算（原点包含在结构元素 B 中） .....	39
图 5-6 膨胀运算（原点不包含在结构元素 B 中） .....	39
图 5-7 积分投影在跃变处曲线.....	43
图 5-8 候选眼睛对的验证 .....	44
图 6-1 眼睛精确定位流程图 .....	45
图 6-2 单一备选区域检测流程.....	47
图 6-1 调用 ITK 函数示例.....	48
图 6-4 CxImage 平台程序界面.....	50
图 6-5 打开图片后的 CxImage 程序界面 .....	50
图 6-6 部分自建数据库检测结果.....	51
图 6-7 部分 BioID 数据库检测结果.....	52
图 6-8 部分多脸孔图像检测结果.....	52