班　级 130911

学　号 13091062

西安电子科技大学

本科毕业设计论文

题　　目 基于掌纹和摄像头的身份认证技术研究

学　　院 软件学院

专　　业 软件工程

学生姓名 张金辉

导师姓名 高海昌

摘　要

随着信息技术的发展，人们对信息安全的需求越来越高。与传统方法相比，生物认证技术具有安全、稳定的优势，逐渐成为了保障信息安全的主流方法。而非接触式的掌纹识别技术以其使用方便、掌纹特征明显的优势逐渐成为了研究的热点，目前许多学者已经提出了各种行之有效的掌纹特征匹配方法。之前大部分研究集中在后期的特征匹配，而背景去除作为掌纹识别过程中重要的一环并没有得到太多的关注，受复杂背景的限制，掌纹识别系统在生活中的应用还不是很广泛。本文提出了一种对不同光照强度及复杂背景具有鲁棒性的背景去除方法。本文的手掌图像数据库通过普通手机拍摄收集建立，拍摄环境尽可能多样。在YCbCr空间用Laws' masks提取图像的纹理及模糊特征，并加入了邻域特征。将这些特征和目标来训练支持向量机，用得到的训练模型对测试样本进行分类。实验证明，与其他方法相比，本文的方法能在复杂环境下分割出更加清晰的掌形图像，最佳分类正确率达到了94.69%。

关键词： **掌纹识别**　**复杂背景去除**　Laws' masks　**支持向量机**

ABSTRACT

With the rapid development of information techniques, the security of users' information is increasingly desired. Compared with traditional security mechanisms, biometric identification has a great advantage of security and consistency and is becoming the mainstream of protection of information. Since people mainly focused on the palmprint matching phase before, the background removal, which is an important part of palmprint recognition system, in preprocessing phase didn't draw much attention. Therefore, the use of the technique in our daily life is limited by complex background. In this paper, a novel method is proposed that is robust to the varying illumination and changing environment. The image database used in this paper is comprised of photos taken by a mobile phone with as many kinds of backgrounds as possible. Then we extract texture features and haze features in YCbCr color space and neighbor features are added to increase the accuracy. And we can get support vector machine trained and predict the test set using the trained model. According to the experiment result, our algorithm is frequently able to segment the hand out of complex environment with a fairly nice hand shape which reached the accuracy of 94.69%.

Keywords: Palmprint recognition Complex background removal Laws' masks Support vector machine

目　录

第一章　绪论 ····· **1**
1.1　课题研究背景 ····· 1
1.2　掌纹识别技术简介 ····· 3
1.2.1　掌纹的特征 ····· 3
1.2.2　掌纹识别的研究现状 ····· 4
1.3　本文主要研究内容和结构 ····· 5
1.3.1　主要研究内容 ····· 5
1.3.2　文章结构 ····· 6

第二章　手掌图像背景去除 ····· **7**
2.1　固定阈值法 ····· 7
2.2　动态阈值法 ····· 7
2.3　基于机器学习的方法 ····· 9
2.4　本章小结 ····· 10

第三章　基于支持向量机的背景去除方法 ····· **13**
3.1　颜色空间的选取 ····· 13
3.2　图像特征的提取 ····· 14
3.2.1　纹理特征和模糊特征 ····· 14
3.2.2　邻域信息 ····· 16
3.2.3　特征小结 ····· 17
3.3　分类算法：支持向量机 ····· 17
3.3.1　支持向量机简介 ····· 18
3.3.2　支持向量机的应用 ····· 21
3.4　本文方法的具体实现 ····· 21
3.4.1　滤波部分 ····· 22

3.4.2 邻域信息的整合 ····· 23
3.4.3 训练数据的处理及支持向量机的应用 ····· 23
3.5 本章小结 ····· 25

第四章 实验结果和性能评估 ····· 27
4.1 特征分析 ····· 27
4.2 区域分块实验 ····· 30
4.3 有无邻域的实验比较 ····· 30
4.4 本文方法评估 ····· 32
4.5 本章小结 ····· 33

第五章 总结与展望 ····· 35
5.1 主要工作总结 ····· 35
5.2 对未来研究的展望 ····· 36

致 谢 ····· 37

参考文献 ····· 39

第一章　绪论

1.1　课题研究背景

随着信息技术的飞速发展，全球趋于信息化和网络化，固然这带来了经济的高速发展和人们生活上的便利，但同时信息化和网络化的社会也面临着安全方面的挑战。在全球化的今天，人们在更加快捷便利获取信息的同时，也容易泄露一些机密信息。身份认证是人们保障信息安全性的基本措施之一。

传统上[1]，一种方式是通过用户拥有的物品（what I have）来达到身份认证的目的，比如说人们进门开锁的过程，这就是使用钥匙进行认证，还有进入图书馆时使用校园卡进行认证等。一种方式是通过用户掌握的知识（what I know）来进行身份认证，比如说登录计算机需要输入正确的密码，去银行取钱时用户也需要提供密码等。但是可以明显看出，传统的身份认证方式有着它固有的缺陷：

- **认证较麻烦**　很多情况下，为了增强安全性，系统会使用多种认证方式组合的形式，例如，人们在ATM 机上取钱时，需要银行卡和密码；在网上购物时，不仅需要银行卡的密码，还会有手机短信密码。
- **记忆成本高**　在信息时代，人们在网络上需要保密的私人信息越来越多，如各种网站的登录密码，邮箱密码等，如果考虑安全的因素，这些密码需要各不相同，这就导致了记忆成本大大增加，很多时候容易被遗忘，造成生活中的麻烦。
- **容易被伪造**　传统上的身份认证方式，无论是卡，密码，还是钥匙，都是“身外之物”，也就是说，系统对用户身份的认证不是直接对人进行的，而是对其他的物品，系统是无法区分在进行认证的是得到授权的本人还是欺诈者。

由于存在上述的问题，传统的身份认证方式在信息爆炸的今天变得越来越不适用，因此，必须寻找一种能够直接对人本身进行识别(what I am) 的身份认证技术，才能满足社会的需要。

在上世纪末，Jain等人提出可以利用人类的生理特征、化学特征或者行为特征来进行生物识别[2]，例如指纹、气息、虹膜、人脸、掌纹、声音、步态等。人

与人之间的生物特征或多或少地存在差异，这就为通过人的生物特征进行身份识别提供了可能。同时这些生物特征具有终身不变的特点，也很难被不法分子盗取或者伪造，更不可能出现丢失或者遗忘的情况，因此基于生物特征的身份认证技术能够很好的克服上述传统方法存在的缺陷。另外，生物识别技术除了可以用于身份验证（如门禁）之外，还可以用于身份的辨别，即从多个人中辨别出某个特定的人，这大大拓展了生物识别技术可应用的领域，如死者身份辨认、罪犯的抓捕和认定等，而传统的身份认证方式很难用于这样的场合。

近年来，随着各国对信息安全等方面的重视，生物识别技术被越来越多的国家或部门采用。根据Alan Gelb和Julia Clark的调查结果[3]，在大部分的发达国家，早已出现了行之有效的身份认证系统，如在美国的携带有生物信息的护照，同时也有越来越多的发展中国家将生物认证技术用于身份的识别。在世界范围内，从2005年到2010年，基于生物特征的认证技术的市场份额平均每年增长28%，在发展中国家，这个数字达到了34%。仅发展中国家而言，生物识别技术的应用就覆盖了10亿人左右，如图1.1所示。

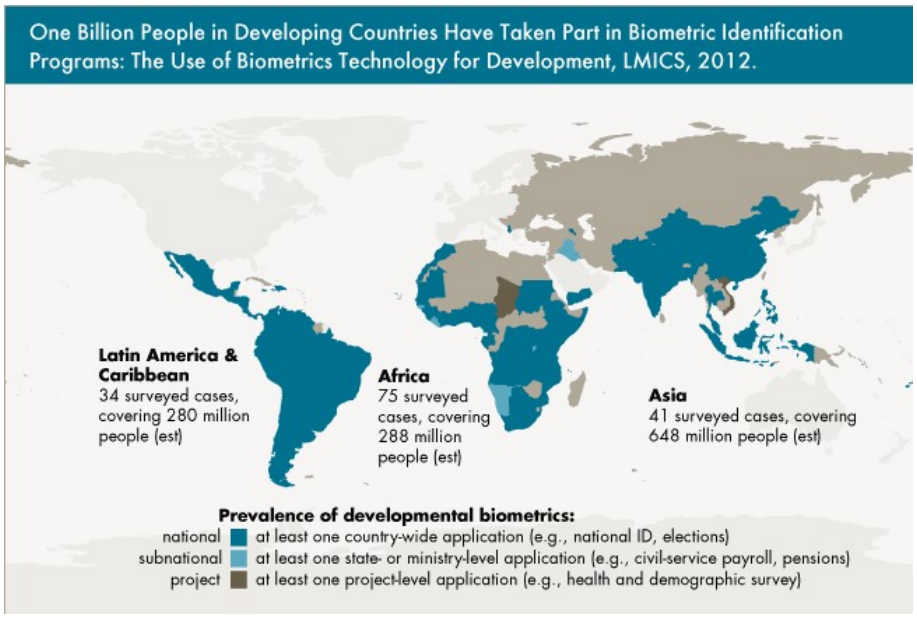

图 1.1 生物识别技术在发展中国家的应用情况[3]

由此可以看出，生物识别技术已经成为保障信息安全的一种重要手段，并且吸引了越来越多的人的关注。从安全角度来讲，生物识别技术克服了传统身份认证技术的三大缺点；从心理学角度来讲，这种技术比较容易被人们接受，比如说系统登录时使用指纹信息的方式已经被大部分人所接受，毕竟生物特征是人自身的信息，用于身份识别的时候不会有太强的排斥感。同时，生物特征的应用省去

了人们携带钥匙、记忆各种密码的麻烦，可以预见，今后生物认证技术将会逐渐渗透到人们的生活中，在保障人们信息安全中起到重要作用。

1.2　掌纹识别技术简介

掌纹识别是张大鹏教授于1998年提出的一种基于生物特征的身份认证方式[4]。它使用人的手掌纹路来对用户的身份进行判别。掌纹是一种非常稳定可靠的生物特征，并且特征明显，即使在低分辨率的图像中也可以轻易辨别，同时它易于提取，在获取的过程中不需要任何的个人信息。因此掌纹识别逐渐成为了生物认证技术的研究热点之一，并已经逐渐应用于实际生活中[5]。

1.2.1　掌纹的特征

与早期的指纹识别相比，掌纹包含有更为丰富的信息，如纹线、乳突纹、细节点、褶皱等，见图1.2。

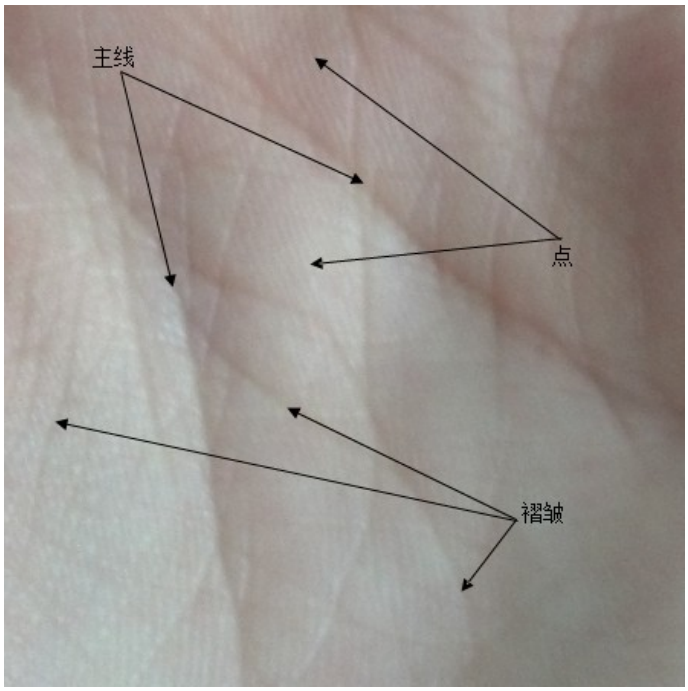

图 1.2 掌纹的不同特征

具体来说，可以用于区别身份的掌纹特征主要包含有以下几类：

1. 主线特征

   手掌上最粗最明显的几条线，即使是在低分辨率图像中也可以轻易分辨出来，一般识别不同的人的时候也常常使用这几条线。

2. 褶皱线特征

比主线更细、更短的一些线，在手掌上的分布毫无规律。不同人之间褶皱线的纹路也各不相同。

3. 点特征

在手掌上形成的和指纹类似的表面特征，如乳突纹在局部形成的有三角形区域的点或小区域。

一般来说，掌纹识别技术采用主线特征或褶皱线特征来进行身份认证的研究。因为这两种纹理不仅在个体间存在差异，也有很好的稳定性，即使在低分辨率、低质量的图像中也可以轻易提取出来。而点特征较为不明显，只有在高质量的图像中才可以被提取用作掌纹特征。随着技术的发展及摄像头精度的不断提高，这些点特征也将被作为特征研究的一部分使得人与人之间的区分度更大，从而得到更好的身份认证效果。

1.2.2 掌纹识别的研究现状

掌纹识别的过程一般分为四个阶段：掌纹采集、掌纹图像预处理、掌纹特征提取以及掌纹特征和数据库信息进行匹配，见图1.3。

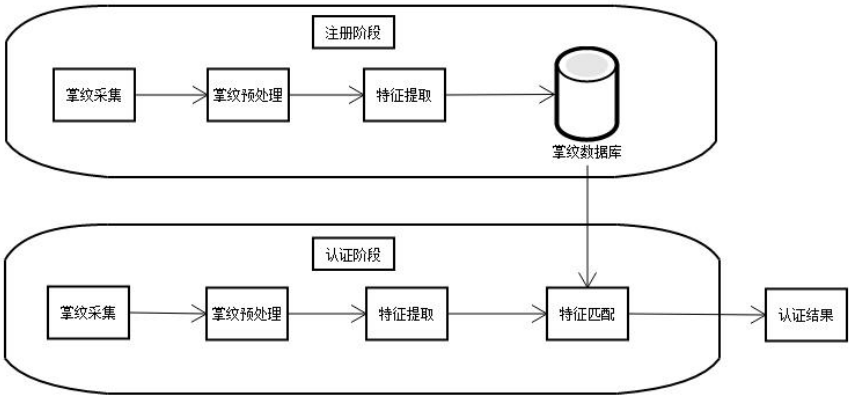

图 1.3 掌纹识别的一般过程

掌纹的采集主要分为两种方式[5]：脱机采集和联机采集。脱机的采集方式是在手掌涂上油墨，然后印在纸上，再去扫描纸上的掌纹；联机采集的方式是利用基于CCD的数字摄像头[6,7]来获取高质量的掌纹图像。显然，脱机采集的方式非常麻烦，难以用在实时系统中，并且人们从心理上对这种掌纹获取的方式有抵触感。在联机采集的机制中，将手掌置于固定的设备上，通过摄像头获取手掌图像，

因此每次都能准确获取到掌纹图像，这也是当今比较流行的掌纹获取方式。随着手机、平板电脑等移动设备的普及和摄像头技术的发展，非接触式的图像采集将会是掌纹识别发展的新方向。

为了准确定位手掌区域，获得掌纹特征进行匹配，我们需要对掌纹图像进行预处理。掌纹的预处理一般包含三个阶段，1) 图像的二值化，2) 手掌方向校正，3) 感兴趣区域（Region of Interest，ROI）的获取。对于CCD设备获取的手掌图像，一般使用大津法[8]进行二值化。进行二值化的算法目前都是一样的，ROI的获取是根据手掌图像的关键点来获取的，而关键点的提取则分为基于切线[6]、基于平分[9]和基于手指[7,10]的方式。大多数的方法都是采用食指、中指、无名指及小指间的三个缝隙作为关键点，利用三个或者其中的两个点来定位手掌。对于使用非接触的方式获得的手掌，Liambas[11]提出了在手掌中作一系列互不相交的最大内切圆来获取掌心区域的方法，同时这个方法还可以克服手掌图像中的噪点、阴影、光照变化等影响ROI提取的因素。Choras等人[12]提出了一种基于手掌边缘和边缘之间的距离的方法来提取关键点，这些关键点既可以将手掌方向进行校正，也可以提取出ROI区域进行手掌特征的提取，但存在的问题是要求手掌区域被严格的从背景中分离出来，但对于稍微复杂的背景该方法并不适用，这也说明了目前非接触式的掌纹识别技术尚不成熟。

掌纹特征的提取和匹配是掌纹识别过程中的重要部分，前人在这一阶段已经做出了很多工作，提出了很多行之有效的算法。归纳起来，可以分为基于线条[13–15]、基于子空间[16,17]、基于统计[18]三大类，还有其他的一些组合了多种方法的机制，例如Hennings-Yeamans等人提出了Log-Gabor滤波的方法[19]，Koichi提出了一种基于相关度的方法[20]等。

1.3　本文主要研究内容和结构

1.3.1　主要研究内容

本文针对掌纹识别技术中的预处理阶段进行了深入的研究，提出了一种能够有效去除手掌图像中复杂背景的方法。本文的方法充分利用手掌和背景在纹理、模糊度、颜色等信息上的差异性，作为特征输入训练支持向量机，达到对图像信息进行分类（手掌或者背景）的目的，以去除背景。经过大量实验，本文提出的分割方法在不同光照条件、不同场景下具有很好的鲁棒性，可以为掌纹识别的后续工作提供一个较好的手掌形状，这就为掌纹识别技术在日常生活中的广泛应用

提供了可能，如手机、平板电脑等。

1.3.2 文章结构

全文共分为五章。第一章介绍了生物认证技术的发展，概括了掌纹识别的优势和发展现状。最后介绍了文章的研究内容和结构安排。

第二章介绍前人在掌纹预处理阶段采用的相关技术，概述人们在复杂背景去除方面做出的工作，并就其方法进行优劣性的分析。

第三章详细介绍本文提出的方法。包括颜色空间的选取及分析，图像特征的选取及分析，支持向量机算法的简介，以及支持向量机在本文方法中的运用。

第四章介绍本文实验的结果。包括使用本文中特征组合比较的实验结果及分析，有邻域信息和无邻域信息两种情况下的性能比较，以及本文方法和现有的背景分割方法的性能对比。实验证明，本文中采取的特征组合和机器学习算法是优于已知的背景分割算法的。

第五章总结本文的研究工作，并对其中的不足提出改进方案。

第二章　手掌图像背景去除

在掌纹图像预处理阶段需要从手掌图像中区分出手掌区域，继而进行关键点的确定，定位ROI区域，之后进行特征的提取和匹配。而预处理的第一步就是图像中背景的去除，背景去除的效果对后期掌纹特征的匹配效果起到重要作用。关于背景去除的方法，人们进行了一定的研究，作出了一些较有成效的工作。

2.1　固定阈值法

前人对手掌图像中背景的去除做了大量研究，提出了许多方法。最早手掌图像的获取通过油墨法，即将手掌涂满油墨在白纸上印下手掌的纹路，这种方法只需要界定白色部分为背景，其余部分为手掌即可。由于油墨法的不便，人们更多使用的是联机获取方式。之后[6–9]手掌图像都是通过基于CCD的摄像设备采集的，如图2.1，这种情况下背景单一固定，使用固定阈值法就可以将手部区域从背景中分离出来。计算公式见式(2-1)，其中$I(x,y)$表示图像坐标为(x,y)像素的灰度值，θ表示区分手掌区域和背景区域的阈值。

$$y = \begin{cases} 1, & if\ I(x,y) \geq \theta \\ 0, & if\ I(x,y) < \theta \end{cases} \qquad \text{式(2-1)}$$

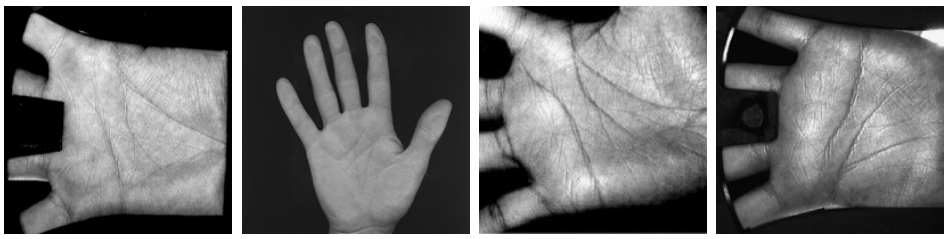

图 2.1 基于CCD的摄像设备采集的手掌图像

2.2　动态阈值法

由于背景往往不是固定的，而是包含有很多无规则的物体，其中的光照条件也各不相同，这时固定阈值法就不再适用。人们提出了基于颜色分布直方图的动态阈值方法，即动态地在颜色分布直方图中找到一个点，能够正确地将手掌和背

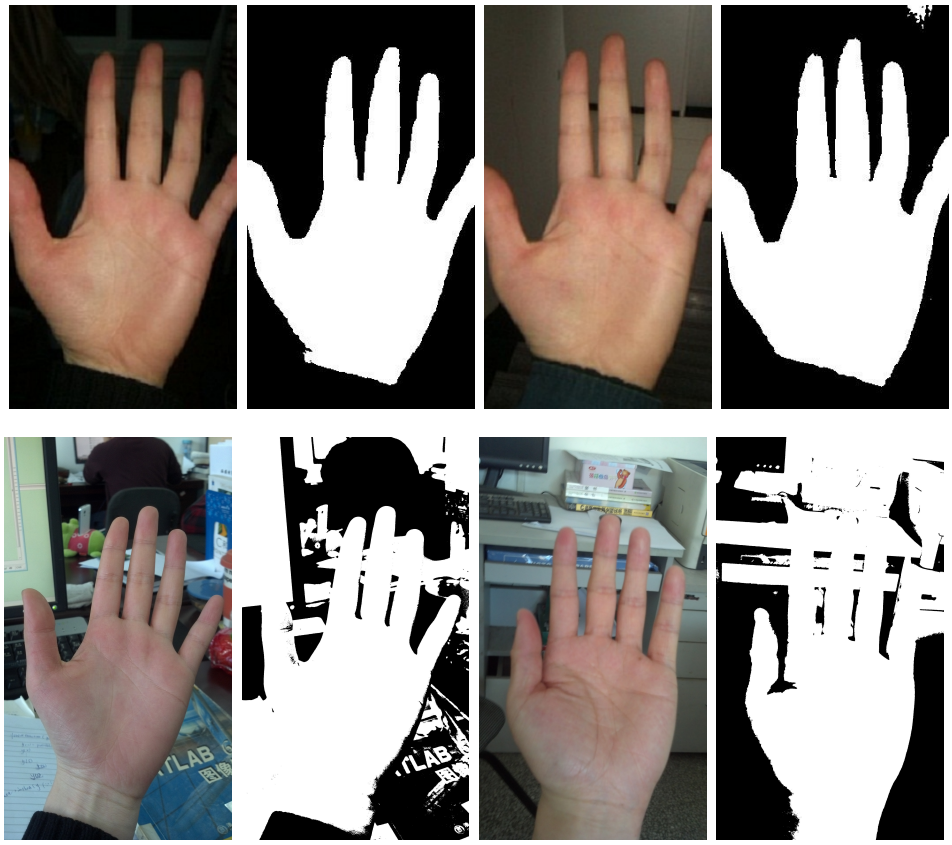

图 2.2 用大津法对图像进行背景去除的结果。第一行为较简单背景的情况；第二行为复杂背景的情况，照片来自本文自建数据库

景分开。Nobuyuki Otsu于1975年提出了大津法[21]，现在许多文献中[22,23]使用这种方法进行手掌与背景的区分。大津法基于这样一种假设：在图像的灰度空间上，目标物体和背景的灰度值服从正态分布。具体来讲，在灰度值分布图上，这个方法可以在分别代表目标物体和背景的两个灰度峰值之间找到一个最低点作为阈值，以这一点为分界点将目标和物体分离开，使得物体和背景两类之间类内距离最小，类间距离最大。Zhu[22]先对图像进行中值滤波消除噪点，然后应用大津法来找到分割手与背景的阈值，这样可以有效去除图像噪声带来干扰，使得分割更为准确。Poon[24]使用手掌和背景的统计信息并应用动态阈值法来进行背景的去除，在他的方法中，由于背景的光照强度较弱，背景较单一，因此分割效果较好。

本文实现了大津法并用数据库中的图片进行实验。图2.2给出通过大津法进行背景去除的结果。从图中可以看出，在背景简单的情况下，大津法非常适于背景去除，如第一行。但如果背景过于复杂，如第二行所示，经过大津法处理的结果就很不理想。因为背景复杂时图像中包含的灰度值范围很大，大津法就不容易确

定准确的阈值。而在我们的日常生活中，通常都是较复杂的背景，经过上述分析及实验结果可以看出，大津法难以适应日常场景的变化。

2.3　基于机器学习的方法

机器学习是上世纪末兴起的一门交叉学科，涉及线性代数、概率论、统计学、凸优化等多门学科，是人工智能的一个分支领域。对机器学习算法的研究目的是使得计算机像人类一样拥有自动“学习”的能力[25]，无论是监督学习还是非监督学习，机器学习算法都是通过分析已有的数据，进而得出规律推理未知的数据。因此，许多研究者提出了基于机器学习算法的手掌图像前景背景分割的方法。

作为机器学习的算法之一，人工神经网络以其自我学习、能够处理输入输出复杂关系、根据输入特征将样本准确分类的能力著称。在这里，输入为从手掌和背景像素中提取的特征，输出为判断该像素点属于哪一类（手掌/背景）。2006年，Doublet[26]提出了在RGB空间对手掌图像进行分割的方法。即将图像中的每个像素的RGB值作为特征，该像素点的分类作为目标来对反向传播（Backpropagation，BP）神经网络进行训练，之后用训练得到的模型对未知的像素点进行分类，见图2.3。

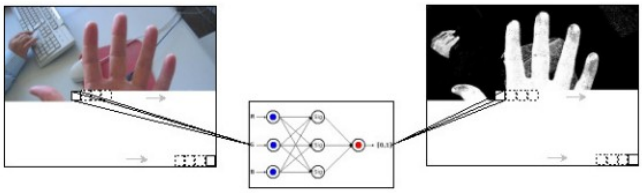

图 2.3 神经网络映射计算[26]

经过实验，该方法确实能够通过每个像素点的RGB值来判断该点为手掌或者背景的概率，但就分割结果来看，手掌图像的细节处理并不是很好，比如说手指缝间的背景去除效果就不是那么理想，导致关键点定位的困难。Yoruk[27]提出了使用Kmeans聚类算法在灰度空间和RGB空间对手掌图像进行分割的方法。作为一种非监督学习算法，Kmeans算法的思想是将样本集分为K类，通过不断迭代聚类中心的方式使得类内距离最小，类间距离最大。由于Yoruk的实验环境较为简单，背景总是比手掌区域稍暗一些，并且场景比较单一，所以严格来说，该方法并不适用于背景复杂、光照强度变化较大的情况。

为了减少RGB颜色空间中的信息冗余和光照敏感性，人们提出了YCbCr颜色空间。由于在YCbCr空间上光照强度和色度是相互独立的，所以它更适于图像像素特征的提取。Y分量是RGB值的加权和，Cb和Cr分量由分别从B和R值中减去Y分量得出，代表蓝色色度和红色色度。RGB和YCbCr的转换关系可表示为式(2-2)。

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad \text{式(2-2)}$$

M.O.Rotinwa等[28]提出了在YCbCr空间上的掌形分割学习方法。他们用Cb和Cr作为像素特征，去掉Y分量以尽量消除光照强度对图像的影响，用三层反向传播（Backpropagation，BP）神经网络进行训练，之后用训练得到的模型参数和Kmeans聚类算法对测试集图片进行二类分类，从而得到背景去除后的手掌部分。但由于去掉光照强度分量之后会降低肤色检测的性能[29]，因此该方法的应用范围十分有限，在背景稍复杂的情况下就无法得出很好的掌形。

另一种基于统计信息的机器学习方法——AdaBoost算法[30]，因为其优秀的分类性能，在很多领域得到了应用，如判断音乐的流派[31]、车辆检测[32]、人脸识别[33]等。AdaBoost算法基于这样一种思想：多个分类器共同对一个样本进行分类，要比一个分类器来进行这个任务的性能好。即分别训练N个弱分类器，再由这些弱分类器组成一个强分类器。Y Han[34]将AdaBoost算法用于复杂背景下的掌形分割。在该方法中，Han提取了图像的Haar特征，用训练集中的手掌形状来对AdaBoost进行训练，以便选出最有效的Haar特征用于级联分类器[35]。但在该方法中仍然需要专门的图像采集设备将手掌部分照亮，以保证分割效果，并不适用于普通手持设备。

2.4 本章小结

随着移动设备的逐渐流行，出于对信息安全的考虑，在移动设备上应用掌纹识别技术是大势所趋，例如在今年的CeBIT展上，富士通公司展示了可以进行掌纹识别的平板电脑原型机，但该平板进行掌纹识别仍需要特定的采集设备，这无疑大大增加了使用成本，并且限制了掌纹设别技术在日常生活中的普及。考虑普通手机或者平板电脑的情况，图像的采集是通过摄像头完成的，而摄像头拍摄的手掌图像不可避免的包含各种不同的背景、不同的物体、不同的光照条件，所以

背景的去除效果直接影响着掌纹特征的匹配，间接决定了这项生物识别技术的普及。

本章介绍了关于掌纹预处理阶段的掌形分割基础，从最初的油墨法到动态阈值法，再到可以应对较复杂背景的机器学习方法。尽管分类方法越来越智能，但在日常生活的场景中仍然不适用，限于此，掌纹识别这种安全方便的身份认证技术没有能够真正普及起来，我们仍然需要一种对光照、背景具有鲁棒性的背景去除方法。

第三章　基于支持向量机的背景去除方法

现今移动设备在人们的生活中迅速普及，但对于手掌图像中复杂背景去除的方法并不是尽善尽美，这就导致了手掌二值化的效果并不理想，这极大影响了掌纹识别技术的易用性，制约了掌纹识别系统在日常生活中的应用。因此迫切需要一种对各种光照、复杂背景具有鲁棒性的技术，来对图像中的前景（手掌）和背景进行分割。本文基于支持向量机算法，在图像的YCbCr空间中提取不同物体的纹理和模糊信息，同时考虑图像中每一区域的邻域信息，据此对手掌和背景进行二类分类。经大量实验证明，本方法具有较好的分类效果，同时对不同光照条件、不同的复杂背景具有鲁棒性。

3.1　颜色空间的选取

我们需要根据移动设备，如手机、平板电脑，拍摄的手掌图像分别提取手掌和背景的特征。人在分辨手掌和背景时的依据之一是它们之间颜色的不同，所以本文的第一步工作就是进行颜色空间的选取。RGB颜色空间是最直观，也是可以从图像中直接获得的颜色信息，由于它的使用简便，在很多基于肤色检测的掌形分割[23,26]的文献中都使用了这个颜色空间，其中Poinsot[23]假设红色分量在手掌区域更加集中而在背景中很少出现，这样两者的颜色就有了很大的对比。为了减少RGB 空间对光照强度的依赖，后来又提出了标准rgb空间[29] （$r+g+b=1$）。实验表明[36,37]，肤色在标准rgb空间上的聚合度高于在RGB空间上，即肤色在标准rgb空间上有更低的方差，因此较之RGB空间，标准rgb空间对肤色检测来说是一个更好的选择。但是在实际应用中，无论是RGB空间还是标准rgb空间，由于三个颜色分量都包含着亮度信息，所以不可避免的会受到外界光照的影响，导致分类的性能并不理想。

为了应对RGB空间的光照敏感性，由RGB进行变换而得到的YCbCr空间随之被提出。YCbCr颜色空间由RGB通过式(2-2)变换得来，Y代表光照强度，Cb代表蓝色色度，Cr代表红色色度，见图3.1。从图中可以看出，图像的Y分量的呈现形式即为图像的灰度空间，表示图像中的光照强度，而Cb和Cr两个分量单纯的表示蓝色分量和红色分量，这样就避免了光照在肤色检测过程中的干扰。因为该空间中光照和颜色之间具有很好的分离度，所以本文的方法采用YCbCr颜色空间。虽

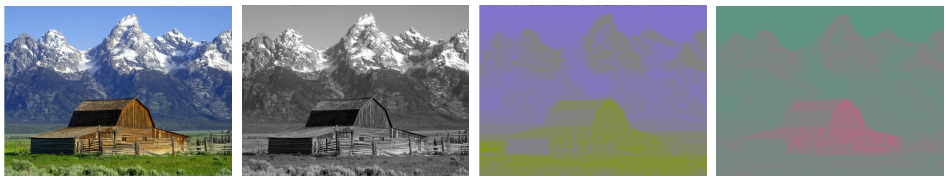

图 3.1 某图片及其Y、Cb、Cr分量图像

然Y作为光照分量，是不同环境、不同背景下的图像中变化最大的一种信息，但经过实验表明[29]，盲目的去除Y分量将会导致检测性能的大幅下降，而收到的对光照具有鲁棒性的效果却不是很理想，因此在本文的方法中保留了Y分量。

3.2 图像特征的提取

3.2.1 纹理特征和模糊特征

接下来的工作是在YCbCr空间中提取出最有利于进行手掌/背景分类的特征。从某种角度上来说，判断前景与背景，实质上是对图像中物体的距离远近进行判断。在掌纹图像中，最近的部分便是手掌，其他部分皆为背景，我们只需要找出距离最近的物体，即手掌，就可以完成对背景的去除。

对人类来说，在一幅照片中判断不同物体的距离是很简单的一件事情[38]。通常我们判断单张图片中物体距离远近的依据是物体的纹理、遮挡、模糊程度、已知物体的尺寸等信息[39,40]，比如说，同一个物体的纹理在不同的距离上观察到的是不一样的，远近不同的物体的纹理也不相同。而模糊是由大气散射造成的一种现象，我们可以通过判断一个物体在图像中的模糊程度来判断其距离的远近。根据我们的经验，手掌的纹理一般情况下和周围的背景是不相同的，因此可以通过纹理来判断手掌的位置。对计算机来说，谈不上已知物体的尺寸，故而舍弃不用。在本文的工作中，一般手和摄像头的距离不会超过30厘米，而背景物体的距离一般大于50厘米，因此可以利用两者间纹理特征和模糊特征的不同来对它们作出区分，从而构造凡是手掌，我们认为是“近”；凡是背景，我们认为是“远”的二类分类问题。

物体的纹理特征一般集中在光照强度分量上（即Y分量），模糊信息一般由颜色分量（即Cb和Cr分量）中的低频信息反映出来。所以我们利用Laws masks滤波模板[41]对图像的Y分量进行滤波以获取纹理特征，用局部均值滤波模板对Cb和Cr分量进行滤波以获取模糊特征[39]。在这里我们使用由Laws’ masks三个一维向量通过卷积计算得到的九个3×3模板对图像进行滤波。三个一维向量

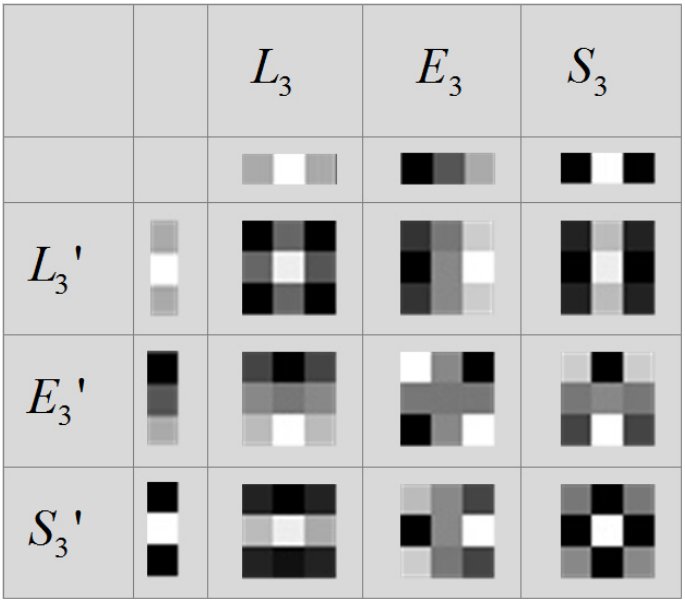

图 3.2 九个Laws’masks滤波模板

如式(3-1)所示。

$$\begin{cases} L_3 = [\ 1 \quad 2 \quad 1] \\ E_3 = [\text{-}1 \quad 0 \quad 1] \\ S_3 = [\text{-}1 \quad 2 \quad \text{-}1] \end{cases} \tag{式(3-1)}$$

其中L_3、E_3、S_3分别用于描述灰度级、边缘和斑点。通过这三个1×3向量生成的九个滤波模板如图3.2所示。 我们用这些模板对图像进行滤波，以获得代表着图像纹理信息和模糊信息的特征。为了说明方便，这里用I_Y、I_{Cb}和I_{Cr}分别代表某一像素的光照强度值、蓝色色度值和红色色度值。具体来讲，在本文中应用九个滤波模板分别和I_Y进行卷积计算得到该点的纹理特征，应用第一个滤波模板（即$L_3 * L_3'$）分别和I_{Cb}、I_{Cr}分量进行卷积计算得到该点的模糊特征，这样我们就能获得对于一个像素点的11个滤波结果，计算过程见式(3-2)。

$$\begin{cases} I_n = I_Y * M_n, \quad n = 1, 2, ..., 9 \\ I_{10} = I_{Cb} * M_1 \\ I_{11} = I_{Cr} * M_1 \end{cases} \tag{式(3-2)}$$

其中M_n表示第n个滤波模板。为了提高计算上的效率，同时为了尽可能包含更多的特征，使得学习算法的分类有更强的置信度，本文将一张图片分成一个个的小区域来进行特征的提取。所以本方法在计算出每个像素的11个纹理特征及模糊特征之后，对每一个小区域i进行特征计算，特征向量生成的具体过程由式(3-3)给

出。

$$F_i = \sum_{(x,y)\in patch(i)} I(x,y) \qquad \text{式(3-3)}$$

其中$I(x,y)$表示坐标为(x,y)的像素点的特征向量（1×11维）。现在给出通过Laws' masks模板滤波后的特征图像（见图3.3），直观上来看，第二张图片和最后两张图片中的手掌部分相对于背景更加明显。但由于特征间不可避免的存在重复信息，造成特征信息冗余，从而导致分类性能下降，同时在机器学习问题中最为关键的一步就是对特征组合的调整，因此最终特征的选取还需要大量的特征组合实验来确定。

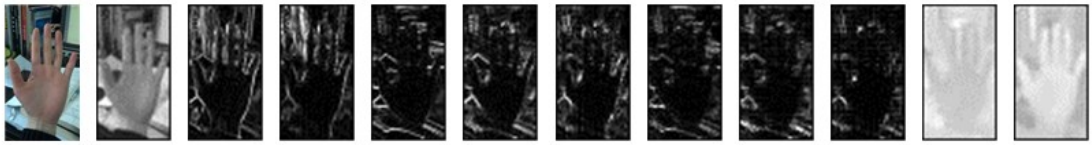

图 3.3 示例图像及其滤波结果，其中第一张为原图，后面11张分别对应通过Laws' masks所得的11个滤波结果

3.2.2 邻域信息

生活经验告诉我们，图像中的某一个小区域是不可能独立存在的，也就是说，某一特定区域中的内容是和它的邻域内容密切相关的（见图3.4）。 具体来说，在

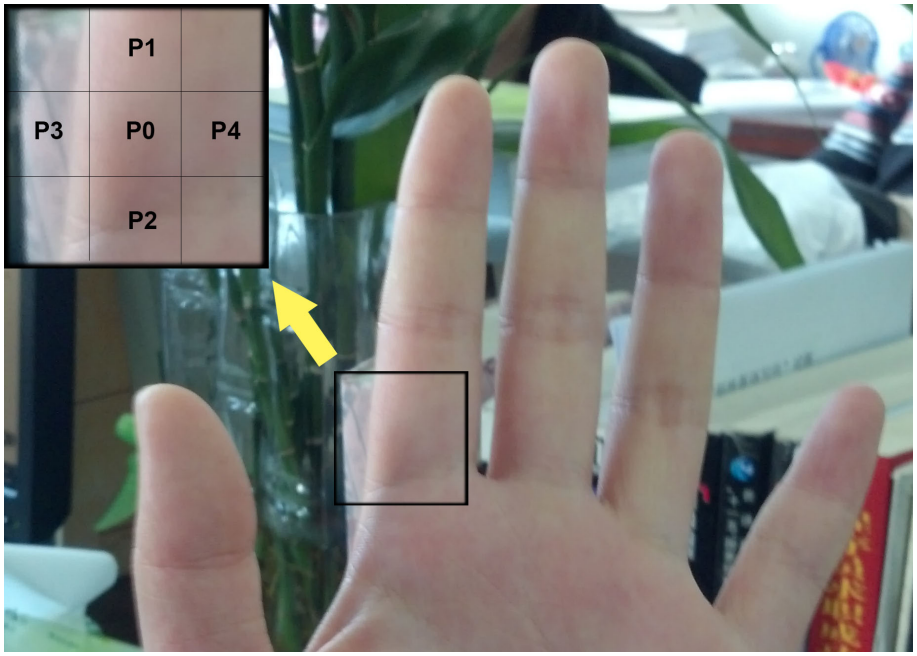

图 3.4 区域P0及其四个邻域

手掌背景去除问题中，某一区域i可以由它的四个邻域的信息来判断它的分类。举例来说，如果一个区域的四周邻域全部是手掌，那么毫无疑问，我们可以确认这个区域也属于手掌。如果它的大部分邻域在背景中，那么它有极大的可能也属于背景一类。在图3.4中，P0的邻域中，P2、P4属于手掌部分，P1的大部分属于手掌，P3属于背景，所以我们有更大的把握判断P0属于手掌区域，尽管实际上有小部分是背景。因此我们考虑将某一区域的邻域纹理特征及模糊特征加入该区域的特征向量，以增大对该区域正确分类的概率。

3.2.3　特征小结

通过以上论述，我们在YCbCr空间提取了图像中每一区域的纹理特征和模糊特征，并与该区域的邻域特征共同组成该区域的特征向量。由于在YCbCr空间上肤色集中在某一特定的区域，同时YCbCr空间对光照的鲁棒性，以及在本文的问题中前景与背景在纹理、模糊上的差别，可以确信本文对特征的提取是合理的，我们的实验也证明了这一点。由于11个滤波特征存在信息重复，从而导致特征信息上的冗余问题，还需要进行特征组合实验，来得到性能最优的特征组合。

3.3　分类算法：支持向量机

根据第二部分提到的，目前的基于机器学习的背景去除方法大部分使用的是人工神经网络，虽然神经网络已经被用于很多领域并获得了成功，但它仍不可避免的存在一些缺陷：

- 神经网络的网络结构需要预先指定，即网络的层数，每层的节点数的确定需要人工参与。一般来说，都是通过经验来判定这些数值，稍有不同就会影响神经网络的性能。
- 神经网络本身是基于Logistic Regression的一种寻优算法，这就使得神经网络很容易陷入局部最优，甚至无法得到最优解。
- 神经网络模型的性能依赖于学习样本，即如果训练样本很小或者样本在样本空间中很集中，那么它就只能对这一部分作出预测，在样本空间其他位置的预测样本则不能被很好的分类。

考虑问题的实际情况，本文使用支持向量机算法作为分类器。支持向量机是基于统计理论的一种机器学习算法[42]，无需经验的参与，这样就保证了支持向量

机的性能不受人们的经验所左右。经过严格的数学证明，支持向量机最终可以转化为一个凸优化问题，这样就保证了算法得到的最优解一定是全局最优。由于支持向量机的理论基础是结构风险最优化，这就使得它具有很好的泛化能力，即使是很小的训练样本，它也能对未知数据作出比较好的预测，在我们的实验中也证明了这一事实，而传统的统计方法只有在训练数据趋近于无穷的情况下才能得到理想的结果。综上所述，在本文中选择支持向量机作为分类算法。

3.3.1 支持向量机简介

在本节给出支持向量机（Support Vector Machine，SVM）学习算法的一个简介。SVM是由Vapnik提出的一种由统计理论发展而来的学习算法，由于它在分类方面出色的性能，应对高维数据空间的能力及可以避免陷入局部最优解，它被广泛用于各种分类问题，比如说人脸识别[43]、图像类型分类[44]及垃圾邮件分类[45]等。

作为一种监督学习算法，首先要用大量的样本对SVM进行训练，之后用得到的模型对测试样本集进行测试来对模型的性能进行评价。SVM的输入为训练集$(x_i, y_i)_{1\leq i\leq m}$，其中$x_i \in R^n$，为训练样本中的特征，$n$代表问题空间的维度，$m$代表训练样本的个数；相对应的，$y_i$表示标记为+1或者−1的目标。算法的目的是找到一个超平面（也叫决策平面），可以把训练样本集中的样本根据y_i值分为两部分，同时也可以对未知分类的样本进行正确分类。这样一个超平面可以表示为式(3-4)的形式。

$$w^T x + b = 0 \qquad \text{式(3-4)}$$

式中w为一个与超平面互相垂直的向量。当训练集线性可分时，我们总可以找到一个(w, b)值使得式(3-5)成立，即找到可将所有样本正确分类的超平面。

$$y_i(w \cdot x_i + b) > 0, \quad i = 1, ..., m \qquad \text{式(3-5)}$$

下面介绍间隔的概念。间隔在SVM中扮演着重要的角色，在这里，为了说明方便，间隔只表示几何间隔，即样本空间中的某一点到分类超平面的距离。对某一特定样本来说，它到超平面的间隔越大，那么对该样本正确分类的置信度越大。为了得到使得所有样本分类置信度最大的超平面，我们需要使得两个类别中的样本到超平面的最小间隔最大。为了获得这个超平面，我们需要求解式(3-6)的最优化问题。

$$\max_{1\leq i\leq m} \frac{y_i(w \cdot x_i + b)}{\|w\|} \qquad \text{式(3-6)}$$

其中$\|w\|$为w的欧几里德范数。为了将问题简化，将函数距离$y_i(w \cdot x_i + b)$固定为1，式(3-6)可以转化为式(3-7)。

$$\begin{aligned} & min\frac{1}{2}\|w\|^2 \\ & s.t. \quad y_i(w \cdot x_i + b) \geq 1, \quad i = 1, ..., m \end{aligned} \tag{式(3-7)}$$

这样，我们就把要求解的问题转化为了一个凸优化问题，而凸优化问题的解一定是全局最优解，这就是支持向量机能够避免陷入局部最优的原因所在。对于该有约束的凸优化问题，可以用拉格朗日乘子法来求解（见式(3-8)）：

$$\min_{w,b} L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{m} \alpha_i \left[y_i \left(w \cdot x_i + b\right) - 1\right] \tag{式(3-8)}$$

其中$\alpha_i \geq 0$是拉格朗日乘子。因为需要求极值，所以使得式(3-8)中w和b的梯度等于0，如式(3-9)和式(3-10) 所示。

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{m} \alpha_i y_i x_i = 0 \Rightarrow w = \sum_{i=1}^{m} \alpha_i y_i x_i \tag{式(3-9)}$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{m} \alpha_i y_i = 0 \qquad \Rightarrow \sum_{i=1}^{m} \alpha_i y_i = 0 \tag{式(3-10)}$$

将式(3-9)与式(3-10)代入式(3-8)，就可以得到一个典型的凸二次优化问题（见式(3-11)）。

$$\begin{aligned} & max \quad W(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \\ & s.t. \quad \sum_{i}^{m} \alpha_i y_i = 0, \quad \alpha_i \geq 0, i = 1, 2, ..., m \end{aligned} \tag{式(3-11)}$$

这样，只需要求解拉格朗日乘子α_i即可，减少了计算量。

当训练集线性不可分时，可以通过核函数技术将训练样本扩展到高维空间，使得训练集在高维空间中线性可分，再进行求解。核函数（Kernel $K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$满足Mercer条件。这样扩展后的优化问题如式(3-12) 所示。

$$\begin{aligned} & max \quad W(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ & s.t. \quad \sum_{i}^{m} \alpha_i y_i = 0, \quad \alpha_i \geq 0, i = 1, 2, ..., m \end{aligned} \tag{式(3-12)}$$

比较常用的核函数有：

1. 线性核：$K(x, y) = x \cdot y$

2. 多项式核：$K(x,y)=[(x\cdot y)+c]^d, c\geq 0$

3. 高斯（径向基函数）核：$K(x,y)=exp\left\{\frac{-\|x-y\|^2}{2\sigma^2}\right\}$

4. 两层神经网络核：$K(x,y)=\tanh[K(x\cdot y)-\delta], k>0$

经过上述推导，可以得到非线性支持向量机学习算法：

算法 3.1: 非线性支持向量机学习算法

输入：训练集$T=\{(x_1,y_1),(x_2,y_2),...,(x_N,y_N)\}$，其中$x_i\in R^n$，$y_i\in\{-1,+1\}$，$i=1,2,...,N$

输出：分类决策函数

（1）选取适当的核函数$K(x_i,x_j)$和参数C，构造支持向量机的对偶形式，并求解式(3-13)

$$\begin{aligned} max \quad & W(\alpha)=\sum_{i=1}^{m}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{m}\alpha_i\alpha_j y_i y_j(x_i\cdot x_j) \\ s.t. \quad & \sum_{i}^{m}\alpha_i y_i=0, \quad 0\leq\alpha_i\leq C, i=1,2,...,m \end{aligned} \qquad \text{式(3-13)}$$

求得最优解$\alpha^*=(\alpha_1^*,\alpha_2^*,...,\alpha_N^*)^T$。

（2）选择α^*的一个正分量$0<\alpha_j^*<C$，计算$b^*=y_i-\sum_{i=1}^{N}\alpha_i^* y_i K(x_i\cdot x_j)$

（3）从而得到决策函数：$f(x)=sign(\sum_{i=1}^{N}\alpha_i^* y_i K(x_i\cdot x_j)+b^*)$

从算法3.1可以看出，式(3-13)是凸二次规划问题，存在全局最优解。

支持向量机算法对样本进行二类分类的例子由图3.5的实验结果给出。可以看出，即使是在比较复杂的非线性情况下，支持向量机通过高斯核（可扩展到无限维）也能够对样本进行有效分类。

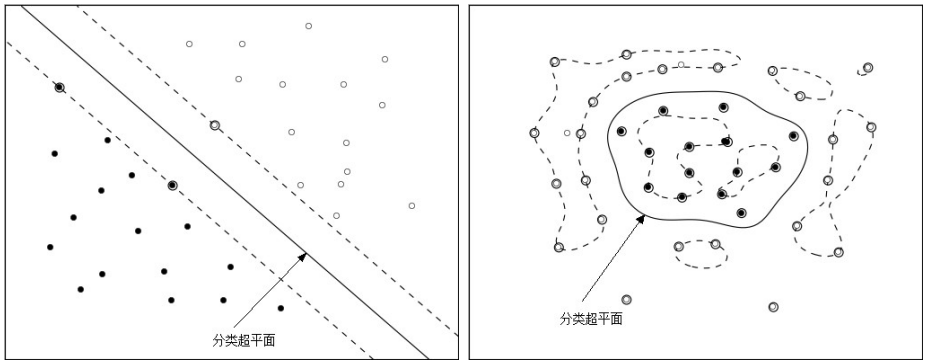

图 3.5 用支持向量机进行二类分类的例子。左图为线性可分的情况，右图为线性不可分的情况，使用了高斯核将样本扩展到了高维空间

3.3.2　支持向量机的应用

在手掌图像背景去除的问题中，由于本文选取的特征个数较多以及分类面（即手掌轮廓）较复杂，本文选择可以将支持向量机扩展到无限维的高斯核函数$K(x,y)=exp\left\{\frac{-\|x-y\|^2}{2\sigma^2}\right\}$。具体的训练阶段和测试阶段的流程如图3.6所示[1]。

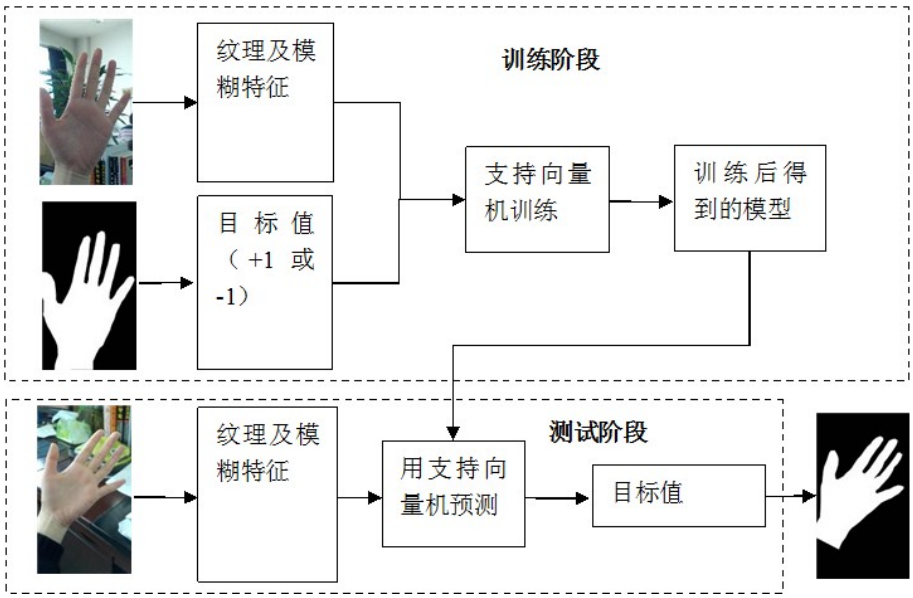

图 3.6 本文方法的流程

我们把图像中的每一个小区域作为一个样本，特征向量根据式(3-3)得出，目标值$y_i \in \{+1,-1\}$ 由手动标记，不同的值表示不同的分类。由于在样本数据库中的背景背景可以分为几个类别，比如说，以电脑为主的背景，以桌面为主的背景等，本文在实验中确保每一种类别的背景都得到了训练。同时，为了避免因为训练最后几张同种背景类别的图片导致的高偏移，训练样本的输入顺序是随机的。

由于在样本量很大时，求解凸二次规划问题就会变得非常低效，在本文的问题中训练样本空间大小在百万级别，因此本文使用序列最小最优化（sequential minimal optimization，SMO）的方法来加快训练速度。

3.4　本文方法的具体实现

据上文的分析过程，本文的实现分为Laws' masks滤波、整合邻域信息、支持向量机分类三部分，由于支持向量机算法本文使用的是Python的Scikit Learn 库

[1]在特征分析实验中（4.1部分），本文使用的是Python的Scikit Learn库[46]，在性能评估实验中（4.3部分及4.4部分），本文使用的是LIBSVM 库[47]。

及LIBSVM库，本文在分类算法的实现上主要集中在特征数据与目标数据的处理上。在Laws' masks滤波及整合邻域信息部分本文使用matlab语言实现，在数据处理部分使用Python语言实现。

3.4.1 滤波部分

在图像的特征提取及处理部分使用Matlab语言实现。对图像进行背景去除的第一步为Laws' masks滤波[39,40,48,49]：

```
MatrixConditioner = .2;
L3 = [1 2 1]/128/MatrixConditioner;
E3 = [-1 0 1];
S3 = [-1 2 -1];
```

代码 3.1 滤波模板定义

代码3.1为本文对Laws' masks的实现，滤波操作由代码3.2给出：

```
temp = conv2(imgY, L3'*L3, 'valid');
temp = [ temp(1*ones(1,round((size(L3,2)-1)/2)),:); temp;...
            temp(end*ones(1,(size(L3,2)-1-round((size(L3,2)-1)/2))
                ),:)];
H(:,:,1) = [temp(:,1*ones(1,round((size(L3,2)-1)/2))) temp...
            temp(:,end*ones(1,(size(E3,2)-1-round((size(E3,2)-1)/2)
                )))];
```

代码 3.2 对图像的Y分量用第一个模板进行滤波

其中H(:, :, 1)表示图像通过Laws' masks模板滤波所得的第一个特征，对应于式(3-2)中$n = 1$的情况，其他10个特征的提取过程与之类似。

```
patch_H1 = patchimg(H(:,:,1),m,n);
patch_H2 = patchimg(H(:,:,2),m,n);
            .
            .
            .
patch_H11 = patchimg(H(:,:,11),m,n);
```

代码 3.3 特征组合

代码3.3将11个特征分开来获取，便于进行之后的特征组合实验。其中patchimg为本文实现的将图像分块的方法，如前文所述，这样可以加快训练速度，同时尽可

能在样本中包含尽可能多的信息，增大分类置信度。m和n表示分块的长宽像素值，一般取m=3，n=3。

3.4.2　邻域信息的整合

加入邻域信息的前提是将每个小区域的特征通过滤波得出，之后对每一区域进行邻域信息的添加，这里的邻域信息即为该区域上下左右四个区域的特征。具体实现时根据某一区域的坐标即可得到其邻域信息。实现过程如下：

```
function [y] = featureVectornew(x,posX,posY)
    y(1)=x(posX,posY);
    y(2)=x(posX,posY-1);
    y(3)=x(posX,posY+1);
    y(4)=x(posX-1,posY);
    y(5)=x(posX+1,posY);
end
```

代码 3.4 邻域信息的整合

其中x为某区域的特征向量，posX和posY表示该区域的坐标值，返回值y即为加入邻域信息后的特征向量。

3.4.3　训练数据的处理及支持向量机的应用

为了方便运行支持向量机算法，本部分代码用Python语言实现。根据本文实现得到的特征文件为“height×width×特征数”的矩阵，而支持向量机支持的数据格式为“样本数×特征数”的矩阵，因此需要对特征数据进行处理，其中用到了scipy包，以读取matlab生成的.mat格式文件。

```
# load feature data
f_data = loadmat(f_filename, mat_dtype = True, matlab_compatible =
    True)
feature = []
feature_mat = list(f_data['feature'])
for i in range(1, height - 1):
    for j in range(1, width - 1):
        feature.append(feature_mat[i][j])
```

代码 3.5 训练数据的处理

程序读取特征文件，得到特征矩阵，根据图片的长宽（以块为单位）来逐块读取特征向量，从而可以得到“样本数×特征数”的矩阵。在for循环中，由于之前加入了邻域信息，导致在图像边缘部分的分块的邻域是残缺不全的（只有两个邻域或三个邻域），因此在构建特征矩阵的过程中忽略了边缘部分。与之相对应的，在目标值向量的构造中也去除了边缘的目标信息：

```
# load target data
t_data = loadmat(t_filename, mat_dtype = True, matlab_compatible =
    True)
target_mat = t_data['target']
t_temp = target_mat[1:height-1, 1:width-1]
# reshape the target matrix to a 1*m vector
t_temp = np.reshape(t_temp, (1, -1))
target = t_temp[0].tolist()
```

代码 3.6 目标数据的处理

由于目标矩阵为“height×width”的矩阵，只需要对其进行reshape操作即可。这样就得到了feature和target两个list，之后就是对支持向量机的训练：

```
clf = svm.SVC(probability = True)
clf.fit(feature, target)
```

代码 3.7 训练支持向量机

在这里使用了Python的Scikit Learn库中的SVC（support vector classification）类，设置为可以输出分类置信概率，便于对结果进行分析和调试，这里的clf就是训练得到的模型，在进行预测时可通过给定的feature给出label值：

```
# deal with the feature data of a test image
for i in range(height):
    for j in range(width):
        # make prediction according to feature data
        target_pre.append(clf.predict(feature_mat[i][j]))
```

代码 3.8 用训练得到的模型进行预测

其中clf.predict()方法可以根据特征数据得到预测的目标值，最终得到的target_pre为预测目标值的list，可保存为.mat文件，方便性能的评估和结果的查看。

3.5 本章小结

在本章中首先介绍了本文选取颜色空间的标准，然后是提取图像的纹理及模糊特征，最后介绍了支持向量机算法以及和神经网络算法的比较。特征选取是机器学习问题中最重要的一个环节，而从图像中直接提取出的特征之间不可避免的存在重复信息，从而造成信息的冗余，所以针对Laws' masks 模板滤波得出的11个特征还需要进行大量的特征组合实验，以得到最好的分类性能。支持向量机算法能够避免陷入局部最优，从而保证了得到的结果一定是唯一的，并且是有最佳性能的。同时使用支持向量机不需要像神经网络那样依赖于经验来构造网络，不会因为人的经验不同而导致性能上的差别。最后给出了对本文方法的具体实现，使用了matlab和Python语言，可以提高原型开发的效率，便于进行实验。

第四章　实验结果和性能评估

在实验阶段，我们收集了来自7个人的总共120张手掌图片构成实验数据库。这些图片是由MOTO ME865型手机在不同光照及随机背景下拍摄的。手机使用的是8MP的摄像头，照片尺寸为1840×3264像素，分辨率为72dpi。数据库中的图像背景包括桌面、电脑、书籍、地板、墙面等。在我们的实验中，我们用70%的图片作为训练样本，用剩下的30%作测试样本。本文的实验是在Intel Xeon E5440 @ 2.83GHz双核、8G内存的环境下进行的。

图 4.1 数据库中的几张示例图片

在本文中，我们用式(4-1)来评估算法的性能：

$$Accuracy = \frac{TP + TN}{P + N} \qquad \text{式(4-1)}$$

对正确率的计算是以图片中的小区域为单位来进行的，每次计算一张图片上分类的正确率。其中TP表示我们将样本正确划分为手掌的数量，TN 表示我们将样本正确划分为背景的数量。P 代表目标为手掌的样本数量，即$y_i = +1$；N代表目标为背景的样本数量，即$y_i = -1$。这样，我们就可以得出每张图片的分类正确率。

4.1　特征分析

在特征提取阶段，我们已经选择Laws' masks模板对图像进行滤波以得到前景和背景的纹理及模糊特征。而直接用滤波模板从图像中提取的特征不可避免的存在信息的重复，基于此，本节将对11个滤波特征进行组合分析。

同时，作为机器学习问题，本文的背景去除方法需要在特征的选择和训练样本数量之间找到一个平衡，使得训练模型在过拟合和欠拟合之间找到一个折中，使得分类性能达到最佳。一般来说，样本较少，尤其是少于特征数量的时候，机器学习容易陷入过拟合，即对训练数据拟合度很好，但对测试数据拟合度很差；同样的，样本过多而特征数量过少的情况下，训练模型容易陷入欠拟合，即对训练数据和测试数据的拟合度都不是太理想。所以从这个角度分析，我们也需要对这11个特征进行调整组合实验。

在本阶段，我们只使用数据库中的16张图片进行特征的组合实验，来达到一个正确率和效率上的平衡。表4.1给出了这11个特征之间的相关度，图4.2给出了这11个特征和目标之间的相关度的绝对值。相关度的计算公式见式(4-2)。

$$R(x,y)=\frac{C(x,y)}{\sqrt{C(x,x)C(y,y)}} \qquad \text{式(4-2)}$$

其中$C(x,y)$表示值x、y的协方差。 结合表4.1和图4.2可以看出，特征1、10和11与

表 4.1 不同特征值之间的相关度参数

	1	2	3	4	5	6	7	8	9	10	11
1	1.00										
2	0.00	1.00									
3	-0.02	0.70	1.00								
4	-0.02	0.33	0.21	1.00							
5	-0.03	0.38	0.41	0.45	1.00						
6	-0.02	0.50	0.44	0.39	0.50	1.00					
7	-0.03	0.18	0.18	0.59	0.42	0.30	1.00				
8	-0.02	0.29	0.26	0.48	0.50	0.41	0.46	1.00			
9	-0.02	0.33	0.36	0.35	0.44	0.53	0.39	0.52	1.00		
10	-0.07	0.02	0.07	0.13	0.12	0.11	0.15	0.13	0.10	1.00	
11	-0.04	-0.01	-0.07	-0.13	-0.12	-0.10	-0.15	-0.13	-0.10	-0.84	1.00

目标的相关度要明显高于其他特征。虽然特征10和特征11之间的相关度参数很高，达到了0.84，说明两者之间存在信息的重复，但由于两者和目标的相关度参数并不算是太高，分别只有0.659和0.732，因此考虑进行实验时两者一起使用。同时观察到特征1与目标的相关度参数要高于剩余的八特征，并且和特征10和特征11的相关度很低，说明特征1中存在对预测目标很有帮助的、但与特征10和11互补的信息，因此可以将特征1用作特征10和11的补充特征。表4.1显示第2个特征到第8个

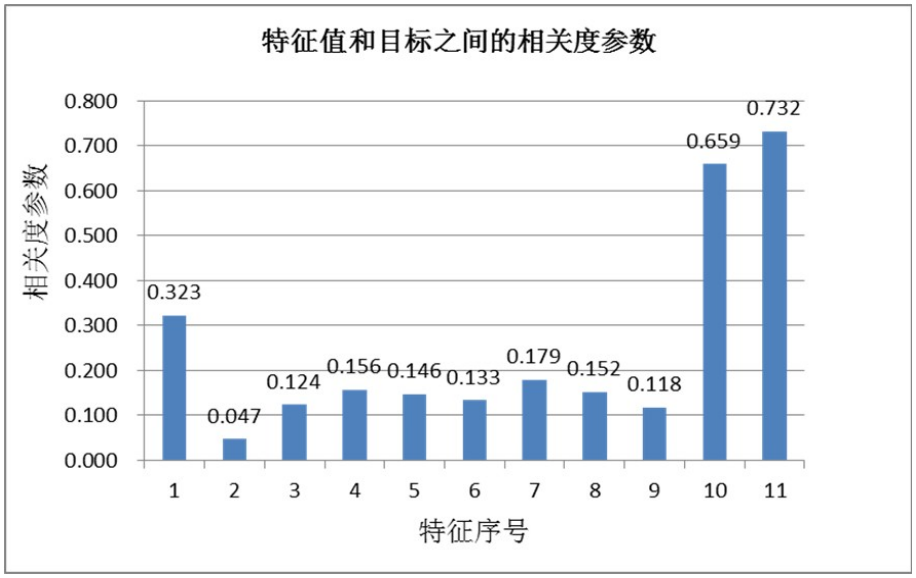

图 4.2 特征值和目标之间的相关度的绝对值

特征与目标的相关度参数都比较低，并且特征之间的相关度参数比较大，为了使得分类性能达到最好，这八个特征不可能同时使用。为了提高实验效率，首先我们使用最好的1、10和11特征作为基本特征，再依次和剩下的特征进行组合实验。我们选取特征的原则是采用这样一种贪心的策略：每次从剩下的特征中选取与目标的相关度参数最大的特征加入。之后将每次选择的特征进行组合，与对应的目标一起输入支持向量机进行训练，然后对测试集的图片进行预测，用式(4-1)得出该特征组合的性能。图4.3给出了我们不同特征组合实验的结果。从实验结果图中可以看出，特征组合（1、4、7、10、11）训练得到的模型有最佳的性能，因此我们选择该特征组合来进行手掌的背景去除工作。

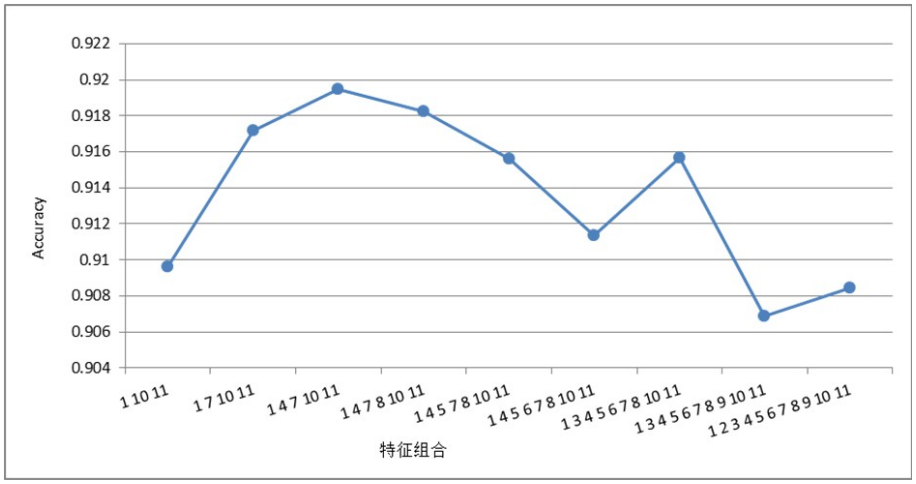

图 4.3 不同特征组合训练得到的模型性能

4.2 区域分块实验

在3.2.1部分中，我们分析从像素级来理解图像是低效的，并且只有多个像素组成的一个小区域在图像内容的展现上才有意义。同时，为了加快支持向量机算法的分类速度，以及使得一个样本中包含的信息尽可能的丰富，以增大分类置信度，我们将图像进行分块，每一块的特征向量由式(3-3)计算得出。

为了证明对区域进行分类的性能及效率，我们进行了分块与不分块的对比实验（见表4.2）。为了简化问题，同时为了便于说明情况，这里使用1、10、11 三个特征组合来进行实验，训练样本使用手掌图像数据库中的43张图片，测试集为16张图片。

表 4.2 分块与不分块的实验对比

	分类正确率	预测时间
分块	0.913	5s
不分块	0.919	12min

从表4.2中可以看出，分块的情况下的预测时间只有5s，而不分块则达到了12min，效率上的对比是显而易见的。但与此同时分块后的分类性能略有下降，这是因为分块导致每张图片上的总样本数减少，一个样本的误分类就会导致正确率出现比较大的浮动（见式(4-1)，总样本数为分母，正确分类样本数为分子）。虽然最终平均正确率下降了0.6%，但速度上是不分块的144倍，因此在本文方法中采用分块的策略。

4.3 有无邻域的实验比较

在第三章讲到了我们加入区域的邻域信息是因为图像中的一个区域是不可能独立存在的，也就是说某一区域属于手还是属于背景可以根据其邻域的特征来判断。比如说我们假设一种极端情况，某一区域的特征值由于某种意外是混乱的，导致支持向量机模型无法正确判断其分类，或者是由于训练模型的问题，导致支持向量机模型无法对该区域正确分类，而其邻域的特征很明显并且能被正确分类，这时我们就可以根据该区域的邻域特征来预测该区域的正确分类。为了验证其性能，我们进行了相同训练样本有邻域特征和无邻域特征的对比实验。图4.4给出了在有邻域信息和无邻域信息的情况下的模型性能对比（16个训练样本）。

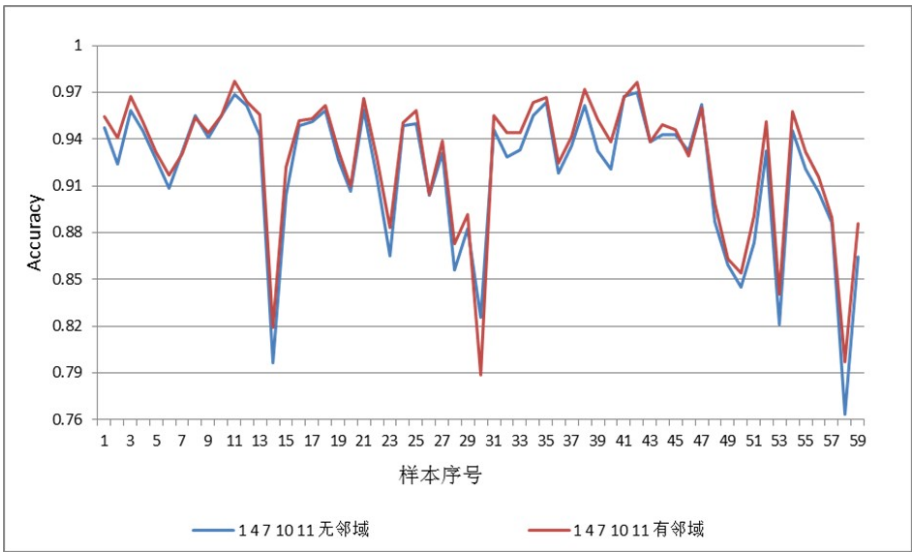

图 4.4 有邻域信息和无邻域信息的情况下的性能对比

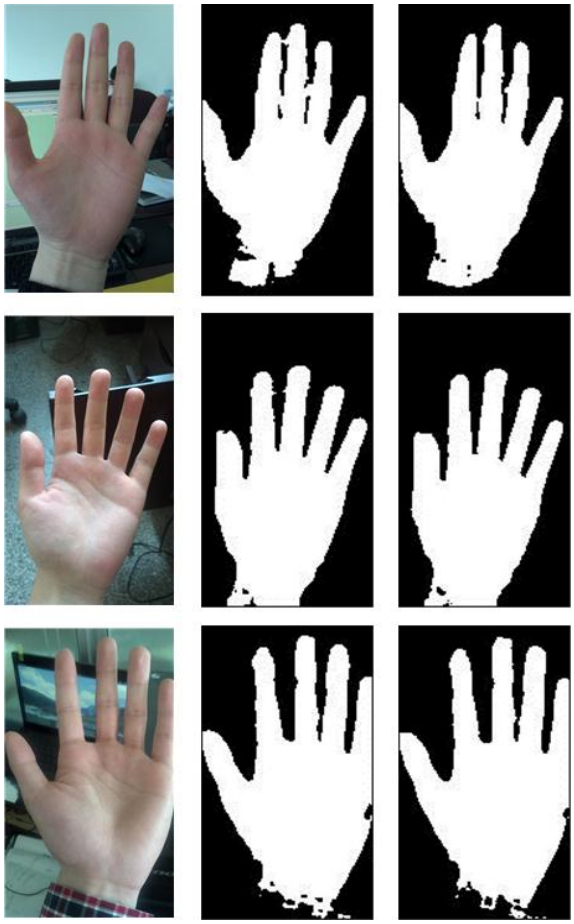

图 4.5 有无邻域特征的结果对比。第一列为原图，第二列为无邻域信息的实验结果，第三列为加入邻域信息的实验结果

根据本文的实验结果，特征向量在加入邻域的情况下的性能整体上要优于无邻域的情况。本文为了提高实验效率，选取了小样本作为支持向量机的训练输入，因此对个别图片的分类效果并不算理想，但足以说明有邻域的情况要好于无邻域的情况。在我们的测试样本中，两种情况下的分类正确率分别为91.87%和90.37%，看起来邻域特征在我们的问题中作用不大，但根据我们的实验结果（见图4.5），加入邻域信息的分类器预测的结果要更加精细，比如说手指缝之间的区域比较清晰，手掌边缘部分比较平滑等。而这些优势对掌纹识别的后续工作，如关键点定位，ROI的提取等非常有利，因此我们加入图片中小区域的邻域信息作为特征向量的一部分。

4.4 本文方法评估

为了对本文提出的方法进行评估，我们还实现了在相关工作中提到过的M.O.Rotinwa的方法[28]和Doublet的方法[26]作为对比实验。M.O.Rotinwa的方法是将每个像素点的CbCr值作为特征，输入三层含有五个隐藏节点的神经网络进行训练，再根据训练得到的模型的参数进行Kmeans分类，从而将背景去除；Doublet的方法是以每个像素的RGB值作为特征，输入三层含有三个隐藏节点的神经网络进行训练，最后用神经网络模型直接对新图像进行预测。同样地，我们使用式(4-1)来评估三种方法的性能。为了比较不同数量的训练样本对分类结果的影响，我们用这三种方法对不同数量的训练集进行训练，然后分别用训练得到的样本对测试集进行预测。图4.6显示了这三种方法在训练样本数分别为14、28、42、56、70、84时的正确率比较。表4.3给出了每种方法在不同数量的训练样本时的正确率。 由图4.6可以看出，M.O.Rotinwa去掉了YCbCr空间中的Y分量以做到

表 4.3 三种方法在不同训练集下的正确率

训练样本数	M.O.Rotinwa的方法	Doublet的方法	本文方法
14	0.822	0.904	0.920
28	0.880	0.925	0.931
42	0.885	0.933	0.944
56	0.887	0.939	0.944
70	0.891	0.939	0.945
84	0.874	0.943	0.947

分类结果不受光照影响，但这样导致了性能的大幅下降，最后样本增加时性能反

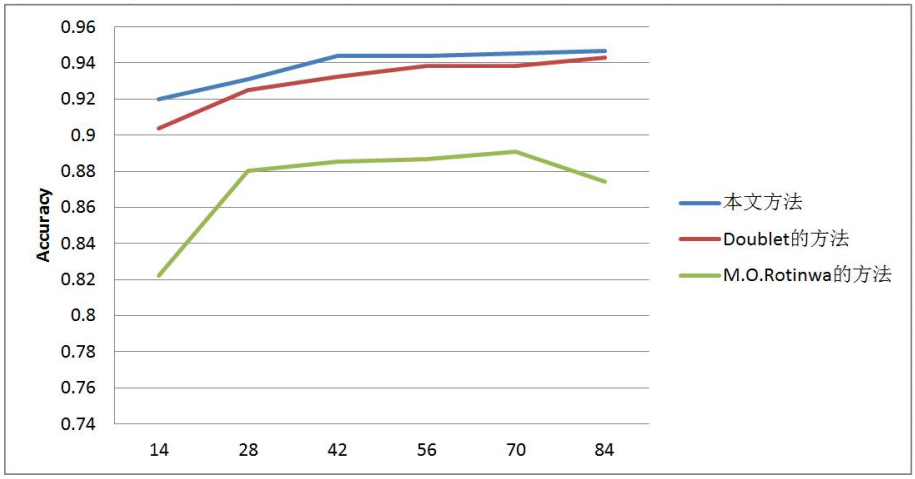

图 4.6 三种方法在不同训练集下的性能对比

而下降，这是因为训练样本过多而特征过少，导致模型出现了高偏移。Doublet的方法在较大样本时和本文方法差距不大，但在小样本的情况下正确率却不是很高。而本文提出的方法无论训练样本的多少，正确率均在92%以上，因此即使在小样本的情况下，本文的方法依然适用。同时根据我们得出的实验结果（见图4.7），用本文的方法处理的图像有更加清晰的手掌轮廓，同时对手掌之外噪点的去除也要明显优于其他两种方法。这样，即使拥有相近的正确率，本文的方法也比Doublet的方法更适于后续的ROI提取、特征匹配等，并且，由于支持向量机算法自身的特性，即使训练样本较少时，本文的方法依然保持良好的性能，因此在训练样本数量有限的情况下，本文的方法同样适用。

4.5　本章小结

本章列出并总结了本文研究过程中的实验结果：特征组合实验、有无邻域特征的对比实验和本文方法与其他方法的对比实验。从目前研究的结果来看，本文的方法成功解决了复杂背景对掌纹识别的干扰，可以清晰的从变化的外界环境中将掌形分离出来，同时本文方法对手掌图像的获取方式没有特殊要求，使用普通的手机或者平板电脑即可做到，这对掌纹识别技术在人们日常生活中普及、更有力的保障人们的信息安全具有重大意义。

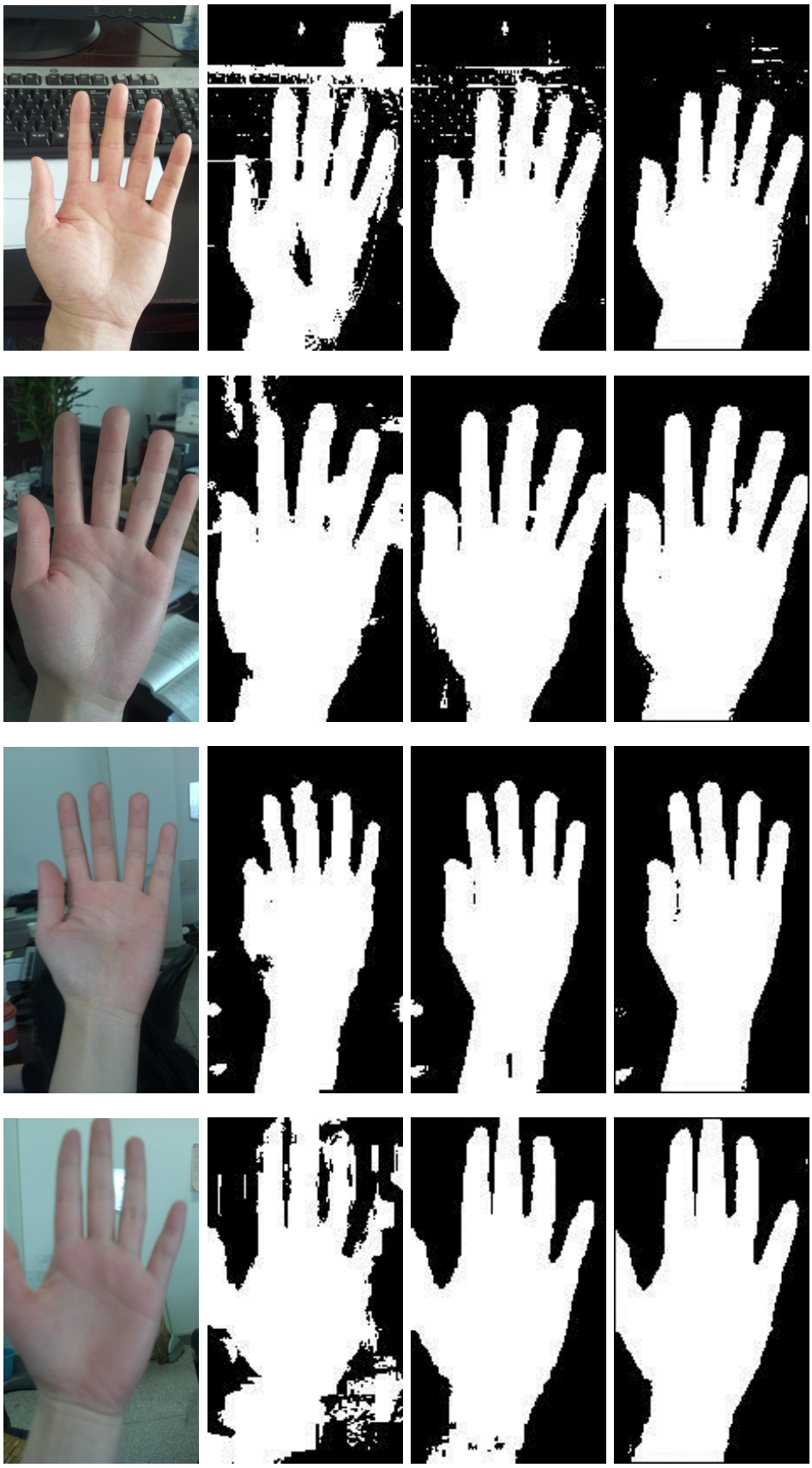

图 4.7 背景去除的结果。其中第一列为原图，第二列为M.O.Rotinwa的方法的结果，第三列为Doublet的方法的结果，第四列为本文方法的结果

第五章　总结与展望

5.1　主要工作总结

随着人们对信息安全的逐渐重视，掌纹认证系统因为它的种种优势，必将成为将来相关领域研究的热点，也终将用于人们的日常生活。但平时人们不可能有专门的仪器设备来获取掌纹而进行身份认证，因此以前的阈值法不再适用，所以在过去手掌图像的复杂背景去除成为了限制这门技术发展的瓶颈。后来人们又基于人工智能领域的一个分支——机器学习，提出了各种方法，并收到了一定的成效，但仍然存在种种限制，比如需要特殊设备、受光照影响大等等。本文就是在这样的背景下提出了一种有效的手掌图像复杂背景去除的方法。

本文的方法是基于支持向量机分类算法和图像的纹理、模糊特征提出的。支持向量机以其不会陷入局部最优解、不受设计者经验制约的优势，在生活中复杂问题的解决上可以获得比神经网络更好的分类结果。本文在特征的提取上不同于前人所做的工作，不是直接根据肤色来进行分类，而是转换了一种思维，将手掌与背景的分类问题转化为将图片中最近的手掌和较远的其他物体进行分类，以立体的思想来看平面的手掌图像。因为两者距离的不同，因此我们可以通过距离来将手掌与背景分类。而我们判断距离远近的时候，除去经验性的知识，依据的是不同物体间的纹理、模糊程度，而Laws' masks模板能够从图像的YCbCr中提取出这些信息。又因为一般直接用模板提取的图像信息不可避免的存在信息重复和信息互补，本文又进行了大量的特征组合实验以找到最优的特征组合。另外，根据我们的生活经验，我们观察图像是不可能将某一区域独立于整体来看的，也就是说，图像中的某一区域的内容、信息，是与它的邻域紧密相关的，很多时候我们可以通过图像中某一区域的邻域来推知该区域的内容，本文中也进行了有无邻域特征的实验，结果证明，有邻域信息时，分类的效果整体上要好于无邻域的情况，更适于进行掌纹识别的下一步工作。

通过本文的分析以及实验，证明了本文提出的方法对变化的光照和复杂背景具有鲁棒性，通过本文方法进行二值化后的手掌较之其他的方法更有利于后续的ROI提取和掌纹特征提取，这极大的促进了掌纹识别技术在我们日常生活中的普及，保障了人们隐私信息的安全。

5.2 对未来研究的展望

由于前人在本文研究的领域所做工作较少，没有成熟的理论基础的支撑，同时限于个人的时间和水平，所以本文提出的方法仍存在很多不足，需要在以后的研究实践中不断进行完善：

- 支持向量机算法是基于二次规划的优化方法，设样本数量为m，求解二次规划问题需要涉及m阶矩阵。当训练样本很大时，会导致占用大量内存和耗费比较长的计算时间，这是支持向量机本身固有的缺陷，虽然人们提出了序列最小最优化等方法来优化支持向量机的训练过程，但效果有限。因此可以考虑使用GPU并行计算来实现支持向量机算法，这样能大大加快计算速度，以节省训练时间。
- 本文的方法已经足以应付较复杂背景，但当背景较单一时，比如说背景主要为白色墙壁时，通过本文的方法进行背景去除的结果反而略差，这主要是因为本文考虑了Y分量，而白色墙壁的光照强度和手掌相近。而背景单一时，使用阈值法可以很容易进行判断，因此，可以考虑本文方法和阈值法的结合，根据颜色直方图分布来判断背景的复杂度，从而采用不同的方法。
- 由于计算及内存空间占用的问题，本文的方法难以直接移植到手机、平板等移动设备上使用。由于机器学习问题的核心是特征、数据，而不在于分类算法，考虑到支持向量机的时空消耗都比较大的缺陷，可以考虑新近兴起的深度学习方法或者其他更加节省时间空间的分类方法。

致　谢

毕业论文即告收尾，能在916实验室度过大学的最后一年，将是我一生难忘的经历。我和高海昌老师于大三相识，那时他严谨的学术态度、渊博的知识就给了我很深的印象。在毕业设计期间，高老师对我的工作进行了细心的指导，同时传授了我很多学术研究的方法，使得我能够在这么短的时间内顺利完成毕设。另外，他平易近人的性格也给予了我生活上很大帮助，使得我能够迅速融入实验室的大家庭中。

感谢刘迪珊师姐、戴菲师姐，在毕设期间能够与她们一起讨论问题、研究数据，使我领略到了进行学术研究的无穷乐趣。在她们的帮助下，我才能够轻松克服在这期间的种种困难。

感谢606宿舍全体舍友，对于我的早出晚归给予充分的理解。这四年有了你们的陪伴，我的大学生活才如此充实。

最后再次感谢我的老师、朋友们！……

谨把本文献给我最敬爱的父母亲以及所有关心我的人！

参考文献

[1] David Zhang. *Automated biometrics: Technologies and systems.* Kluwer Academic Pub, 2000.

[2] Anil K Jain, Ruud Bolle, and Sharath Pankanti. *Biometrics: personal identification in networked society.* kluwer academic publishers, 1999.

[3] Alan Gelb and Julia Clark. Identification for development: The biometrics revolution. Technical report, 2013.

[4] Wei Shu and David Zhang. Automated personal identification by palmprint. *Optical Engineering*, 37(8):2359–2362, 1998.

[5] Priyanka Somvanshi and Milind Rane. Survey of palmprint recognition. *International Journal of Scientific & Engineering Research*, 3(2):1, 2012.

[6] David Zhang, Wai-Kin Kong, and Jane You. Online palmprint identification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(9):1041–1050, 2003.

[7] Chin-Chuan Han. A hand-based personal authentication using a coarse-to-fine strategy. *Image and Vision Computing*, 22(11):909–918, 2004.

[8] Wai Kin Kong, David Zhang, and Wenxin Li. Palmprint feature extraction using 2-d gabor filters. *Pattern recognition*, 36(10):2339–2347, 2003.

[9] Wenxin Li, David Zhang, and Zhuoqun Xu. Palmprint identification by fourier transform. *International Journal of Pattern Recognition and Artificial Intelligence*, 16(04):417–432, 2002.

[10] Chin-Chuan Han, Hsu-Liang Cheng, and Chih-Lung Lin. Personal authentication using palm-print features. *Pattern Recognition*, 36(2):371–381, 2003.

[11] Christos Liambas and Constantine Tsouros. An algorithm for detecting hand orientation and palmprint location from a highly noisy image. In *Intelligent Signal Processing, 2007. WISP 2007. IEEE International Symposium on*, pages 1–6. IEEE, 2007.

[12] Michał Chorań and Rafał Kozik. Contactless palmprint and knuckle biometrics for mobile devices. *Pattern Analysis and Applications*, 15(1):73–85, 2012.

[13] Fang Li, Maylor KH Leung, and Xiaozhou Yu. Palmprint identification using hausdorff distance. In *Biomedical Circuits and Systems, 2004 IEEE International Workshop on*, pages S3–3. IEEE, 2004.

[14] Xiangqian Wu, Kuanquan Wang, and David Zhang. Line feature extraction and matching in palmprint. In *Second International Conference on Image and Graphics*, pages 583–590. International Society for Optics and Photonics, 2002.

[15] Xiangqian Wu, Kuanquan Wang, and David Zhang. Fuzzy directional element energy feature (fdeef) based palmprint identification. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 1, pages 95–98. IEEE, 2002.

[16] Xiangqian Wu, David Zhang, and Kuanquan Wang. Fisherpalms based palmprint recognition. *Pattern recognition letters*, 24(15):2829–2838, 2003.

[17] Guangming Lu, David Zhang, and Kuanquan Wang. Palmprint recognition using eigenpalms features. *Pattern Recognition Letters*, 24(9):1463–1467, 2003.

[18] Xianji Wang, Haifeng Gong, and Hao Zhang. Palmprint identification using boosting local binary pattern. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 3, pages 503–506. IEEE, 2006.

[19] Pablo H Hennings-Yeomans, BVK Vijaya Kumar, and Marios Savvides. Palmprint classification using multiple advanced correlation filters and palm-specific segmentation. *Information Forensics and Security, IEEE Transactions on*, 2(3):613–622, 2007.

[20] Koichi Ito, Takafumi Aoki, and Hiroshi Nakajima. A palmprint recognition algorithm using phase-based image matching. In *Image Processing, 2006 IEEE International Conference on*, pages 2669–2672. IEEE, 2006.

[21] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.

[22] Xijun Zhu, Dazhuan Liu, and Qiulin Zhang. Research of technology for thenar palmprint locating and segmentation. In *Proceedings of 2009 International Workshop on Information Security and Application*, pages 601–604, 2009.

[23] Audrey Poinsot, Fan Yang, and Michel Paindavoine. Small sample biometric recognition based on palmprint and face fusion. In *Computing in the Global Information Technology, 2009. ICCGI'09. Fourth International Multi-Conference on*, pages 118–122. IEEE, 2009.

[24] Carmen Poon, David C. M. Wong, and Helen C. Shen. A new method in locating and segmenting palmprint into region-of-interest. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 4, pages 533–536. IEEE, 2004.

[25] Avrim Blum. Machine learning theory. *Carnegie Melon Universit, School of Computer Science*, 2007.

[26] Julien Doublet, Olivier Lepetit, and Marinette Revenu. Contact less hand recognition using shape and texture features. In *Signal Processing, 2006 8th International Conference on*, volume 3. IEEE, 2006.

[27] Erdem Yoruk, Ender Konukoglu, and Bülent Sankur. Shape-based hand recognition. *Image Processing, IEEE Transactions on*, 15(7):1803–1815, 2006.

[28] MO Rotinwa-Akinbile, Abiodun Musa Aibinu, and Momoh Jimoh Emiyoka Salami. A novel palmprint segmentation technique. In *Informatics and Computational Intelligence (ICI), 2011 First International Conference on*, pages 235–239. IEEE, 2011.

[29] Praveen Kakumanu, Sokratis Makrogiannis, and Nikolaos Bourbakis. A survey of skin-color modeling and detection methods. *Pattern recognition*, 40(3):1106–1122, 2007.

[30] Yoav Freund and Robert E Schapire. Experiments with a new boosting algorithm. In *MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-*, pages 148–156. MORGAN KAUFMANN PUBLISHERS, INC., 1996.

[31] James Bergstra, Norman Casagrande, and Dumitru Erhan. Aggregate features and adaboost for music classification. *Machine Learning*, 65(2-3):473–484, 2006.

[32] Ayoub Khammari, Fawzi Nashashibi, and Yotam Abramson. Vehicle detection combining gradient analysis and adaboost classification. In *Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE*, pages 66–71. IEEE, 2005.

[33] Mian Zhou and Hong Wei. Face verification using gaborwavelets and adaboost. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 1, pages 404–407. IEEE, 2006.

[34] Yufei Han, Zhenan Sun, and Fei Wang. Palmprint recognition under unconstrained scenes. In *Computer Vision–ACCV 2007*, pages 1–11. Springer, 2007.

[35] Eng-Jon Ong and Richard Bowden. A boosted classifier tree for hand shape detection. In *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, pages 889–894. IEEE, 2004.

[36] Jie Yang, Weier Lu, and Alex Waibel. *Skin-color modeling and adaptation.* Springer, 1997.

[37] Ming-Hsuan Yang and Narendra Ahuja. Gaussian mixture model for human skin color and its application in image and video databases. In *Proc. SPIE: Storage and Retrieval for Image and Video Databases VII*, volume 3656, pages 458–466, 1999.

[38] Jack M Loomis. Looking down is looking up. *Nature*, 414(6860):155–156, 2001.

[39] Jeff Michels, Ashutosh Saxena, and Andrew Y Ng. High speed obstacle avoidance using monocular vision and reinforcement learning. In *Proceedings of the 22nd international conference on Machine learning*, pages 593–600. ACM, 2005.

[40] Ashutosh Saxena, Sung H Chung, and Andrew Ng. Learning depth from single monocular images. *Advances in Neural Information Processing Systems*, 18:1161, 2006.

[41] Kenneth Ivan Laws. Textured image segmentation. Technical report, DTIC Document, 1980.

[42] Vladimir Vapnik. *The nature of statistical learning theory.* springer, 1999.

[43] P Jonathon Phillips et al. *Support vector machines applied to face recognition.* Citeseer, 1998.

[44] Olivier Chapelle, Patrick Haffner, and Vladimir N Vapnik. Support vector machines for histogram-based image classification. *Neural Networks, IEEE Transactions on*, 10(5):1055–1064, 1999.

[45] Harris Drucker, Donghui Wu, and Vladimir N Vapnik. Support vector machines for spam categorization. *Neural Networks, IEEE Transactions on*, 10(5):1048–1054, 1999.

[46] Fabian Pedregosa, Gaël Varoquaux, and Alexandre Gramfort. Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research*, 12:2825–2830, 2011.

[47] Chih-Chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[48] Ashutosh Saxena, Min Sun, and Andrew Y Ng. Learning 3-d scene structure from a single still image. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007.

[49] Ashutosh Saxena and Andrew Y Ng. Make3d. http://make3d.cs.cornell.edu/, 2007.