



申请同济大学工学硕士学位论文

关于人脸检测技术的研究

培养单位：电子与信息学院

一级学科：控制科学与工程

二级学科：模式识别与智能系统

研 究 生：庄茵芳

指导教师：武妍 教授

二〇〇六年三月

摘要

随着数字技术和因特网的发展,人脸检测和识别广泛地应用于身份识别、人机交互等领域中。目前对于人脸检测算法的研究已经相当深入,但对于复杂背景下的人脸检测仍难以取得较好的效果。此外,人们将研究的重心从灰度图像转向彩色图像中的人脸检测。在人脸检测中,精度和速度是一个人脸检测系统的两个重要方面,一般都希望一个系统既能有很高的精度又能达到实时的速度。但是实际研究中这两方面又常常存在矛盾。一般是在精度不能满足的条件下,牺牲速度来满足精度。如何能在保证精度的前提下,有效地提高系统的速度,对人脸检测的研究有很重要的意义。

针对以上关键性问题,本文在以下几个方面进行了创新性的探索:

- 1) 对传统的 SVM 检测算法进行改进,即在学习和检测过程中,对人脸和非人脸进行预处理之后,先需经过 ICA 特征提取得到对应模式的特征向量(即新的分类依据),再采用 SMO 算法训练支持向量机。这里创造性地采用核独立分量分析算法,并在算法中运用不完全 Cholesky 分解来降低时间复杂度。对求得的特征采用遗传算法再进行一次特征选取以便有效地减少特征空间的维数,得到独立基的一个最优子集。实验表明,和传统的 SVM 方法相比,检测率提高了将近 15%,检测速度也有明显的提高。
- 2) 对链式 Boosting 方法进行了一些研究工作。首先介绍了链式 Boosting 的结构、训练算法以及对算法的一些优化。然后将链式 Boosting 运用到人脸检测中。在该检测方法中,先利用核独立分量分析算法构造重建图像信噪比阈值函数,再使用 Boosting 算法构造一个基于信噪比阈值的检测函数序列,然后将它们组合成一个总检测函数,形成一个 Boosting 分类器,再使用链式 Boosting 算法用相同的方法构造一个个 Boosting 分类器(结点),并将它们串起来形成链式 Boosting 结构,据此可以判别一幅图像是否为人脸图像。实验表明,链式 Boosting 方法检测率比 1) 方法要有所提高,检测速度也比 1) 方法快;但是虚警数随着检测率的提高也迅速增加,同时它需要较长的训练时间。
- 3) 提出了一种新的基于肤色模型和模板匹配的人脸检测算法。在模板匹配中,提出了一种基于边缘检测的人脸平均模板。该模板相对于一般的灰度平均脸

模板,人脸特征更加突出(五官都以白色突出显示)。对于遮挡和重叠的人脸区域,将基于边缘检测的人脸平均模板匹配之后已检测到的人脸区域置黑,用该模板的一半再次进行模板匹配。通过求得候选区域的倾斜角,把模板按照这个角度进行旋转后再进行检测,一定程度上解决了平面内旋转人脸的检测问题。在使用基于边缘检测的平均脸模板进行模板匹配后所得的人脸候选区域中再进一步使用了基于边缘检测的平均耳模板进行模板匹配,一定程度上解决了深度旋转角度过大的人脸的检测问题。而在进行人脸验证时,先通过大量的实验建立头发模型(该模型可以检测黑发和染过的头发),剔除那些没有头发相邻的候选人脸区域,再使用改进了的三分图模型验证规则进一步对余下的人脸区域进行验证。实验表明,本文提出的模板匹配方法在检测率、虚警数以及检测时间上达到了最好的折衷,并且可以避免训练过程。

关键词: 人脸检测, 肤色模型, 模板匹配, 核独立分量分析, 链式 Boosting

ABSTRACT

With the development of digital technology and the Internet, face detection and face recognition are widely used in the fields such as identification recognition, human-computer interaction, etc.. At present there exist relatively in-depth researches on face detection algorithms. However, as to face detection in complex scenes, relatively good detection results are hard to get. Moreover, there is a tendency for researchers to shift their attention from face detection in gray images to face detection in color images. Besides precision and rate are two important aspects of a face detection system. It is generally hoped that the system has high precision and rate simultaneously. However in reality precision and rate are in conflict with each other. Therefore it is of significance to do researches on how to upgrade the system rate under the predefined precision.

Here we do some creative researches on the following three aspects in order to solve the above problems:

- 1) Make some revisions on the traditional SVM face detection algorithm. In the training and testing processes, first do independent component analysis to get the independent features. Then we adopt the traditional SMO algorithm to train support vector machine. Here we use kernel independent component analysis(KICA) to get those key features and use Incomplete Cholesky Decomposition to decrease the time complexity of KICA. Then we apply Genetic Algorithm to further select features. Finally we get the optimized subset of independent components. Experimental results show that the detection rate upgrades 15% compared with that of the traditional face detection method based on SVM. And the detection is also accelerated greatly.
- 2) Do some researches on Boosting Chain algorithm. First introduce its structure, training algorithm and several optimizations of the original algorithm. Then apply Boosting Chain algorithm to face detection field. The first step of this method is to construct the signal-noise threshold function by means of kernel independent component analysis. The second step is to apply Boosting algorithm to construct a

series of detection functions based on this signal-noise threshold function model and combine them into one detection function. We call it Boosting classifier or Boosting node. We use this method several times to construct several Boosting classifiers. Finally we thread them into the Boosting Chain structure and use it to detect faces. Experimental results show that the face detection method based on Boosting Chain has better detection performance and uses less detection time than the first method. However the false alarm number increases swiftly with the improvement of the face detection rate and the method also requires long time to train the detector.

3) Propose a new face detection algorithm based on face color model and template matching. In the template matching step, propose average face template based on edge detection. This template makes facial features more obvious. As to partly covered face regions, apply left half, right half, upper half and lower half of the above average face template based on edge detection to further detect faces. After getting the tilted angle of the face and rotating the template accordingly, do the face detection. After that, use average ear template based on edge detection to further detect faces with large out-of-plane rotation angles. In the rule validation step, first do a lot of experiments to get the hair model (It can not only detect black hair, but also detect dyed hair). Then use hair model and the priori knowledge of hair lying just beside face regions to remove some of the candidate regions. After that, use revised face San-fen image model to validate the rest candidate regions. Experimental results show that the face detection method based on template matching has gained reasonable detection rate, false alarm number and detection time at the same time and it also avoids the training process.

Key Words: face detection, skin color model, template matching, kernel independent component analysis (KICA), Boosting chain

学位论文版权使用授权书

本人完全了解同济大学关于收集、保存、使用学位论文的规定，同意如下各项内容：按照学校要求提交学位论文的印刷本和电子版；学校有权保存学位论文的印刷本和电子版，并采用影印、缩印、扫描、数字化或其它手段保存论文；学校有权提供目录检索以及提供本学位论文全文或者部分的阅览服务；学校有权按有关规定向国家有关部门或者机构送交论文的复印件和电子版；在不以赢利为目的的前提下，学校可以适当复制论文的部分或全部内容用于学术活动。

学位论文作者签名： 庄尚芳

2006 年 3 月 10 日

经指导教师同意，本学位论文属于保密，在 年解密后适用本授权书。

指导教师签名：

学位论文作者签名：

年 月 日

年 月 日

同济大学学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，进行研究工作所取得的成果。除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人创作的、已公开发表或者没有公开发表的作品的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。本学位论文原创性声明的法律责任由本人承担。

签名：庄南芳

2006 年 3 月 10 日

第1章 绪论

1.1 人脸检测问题的提出

人脸检测 (Face Detection) 是指在输入图像中确定所有人脸 (如果存在) 的位置和大小的过程^[1]。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、研究十分活跃的课题。

人脸检测问题最初来源于人脸识别 (Face Recognition), 人脸识别的研究可以追溯到 20 世纪 60—70 年代, 经过几十年的曲折发展已日趋成熟。人脸检测是自动人脸识别系统中的一个关键环节, 但是早期的人脸识别研究主要针对具有较强约束条件的人脸图像 (如无背景的图像), 往往假设人脸位置已知或很容易获得, 因此人脸检测问题并未受到重视。近几年随着电子商务等应用的发展, 人脸识别成为最有潜力的生物身份验证手段, 这种应用背景要求自动人脸识别系统能够对一般环境图像具有一定的适应能力, 由此所面临的一系列问题使得人脸检测开始作为一个独立课题受到研究者的重视。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^[2]。

1.2 人脸检测的学术和应用价值

人脸检测是智能人机接口技术研究的基本内容。不仅是智能接口系统的重要功能模块, 而且其他智能感知系统的基础和前提。人脸检测技术为其他感知系统 (人脸识别、表情识别、唇读等) 的技术研究提供了前提条件, 即有效、快速的人脸定位。

人脸检测还是其他领域中的一项关键技术, 例如、视频会议、基于对象的编码、应用视觉系统 (如安全检查、保安全管理)、Web 搜索、三维人脸合成、视频检索、基于内容的图像检索等。在数字音频领域有广泛影响的 MPEG 系列国际标准也在征求人脸检测算法。

由于人脸是一类细节变化复杂的自然结构目标, 如果能够找到解决这些问

题的方法，成功构造出人脸检测与跟踪系统，将为解决其他类似的复杂模式检测问题提供重要的启示。

1.3 人脸检测的难点

人脸检测问题的挑战性在于^[2]：

- 1) 人脸由于外貌、表情、肤色等不同而存在很大差异性，具有模式可变性；
- 2) 人脸上可能存在眼镜、胡须等附属物；
- 3) 作为三维物体的人脸的影像不可避免地受到光照的影像，光照的不均匀性特别是偏光的影响极大破坏了人脸特征空间分布的规律性。
- 4) 自然景物中存在大量与人脸相似的图像，如斑驳的树叶、老式建筑物的外表面等，这些复杂背景都极易与人脸图像相混淆。

1.4 目前的人脸检测方法

人脸图像所包含的特征十分丰富，如肤色特征、轮廓特征、镶嵌图特征、直方图特征、器官对称性特征和投影特征、模板特征等。如何充分利用这些特征，是人脸检测要研究的一个关键问题。人脸具有复杂而细致的变化，因此一般需要采用多种特征综合的方法。

通过长期研究，人脸检测方法总的来说可分为基于启发式模型的方法和基于统计学习的方法。前者根据先验知识抽取人脸特征，然后进行建模，是人脸检测的传统方法。后者把人脸检测问题当作一个标准的模式识别问题来处理，通过对某一学习算法的训练，获得想要得到的人脸检测器。统计学习的方法也可以利用人脸的先验知识，所不同的是，这些知识是被融入到学习过程中被间接使用的。

1.4.1 基于启发式模型的方法

对一个典型的复杂背景下的人脸检测问题，基于启发式模型的方法首先要分割出一些初级的视觉特征，例如边缘特征、灰度特征、颜色特征等。这些视觉特征还比较低级，不能包含太多的信息量，需要对它做进一步的分析，即特

题的方法，成功构造出人脸检测与跟踪系统，将为解决其他类似的复杂模式检测问题提供重要的启示。

1.3 人脸检测的难点

人脸检测问题的挑战性在于^[2]：

- 1) 人脸由于外貌、表情、肤色等不同而存在很大差异性，具有模式可变性；
- 2) 人脸上可能存在眼镜、胡须等附属物；
- 3) 作为三维物体的人脸的影像不可避免地受到光照的影像，光照的不均匀性特别是偏光的影响极大破坏了人脸特征空间分布的规律性。
- 4) 自然景物中存在大量与人脸相似的图像，如斑驳的树叶、老式建筑物的外表等，这些复杂背景都极易与人脸图像相混淆。

1.4 目前的人脸检测方法

人脸图像所包含的特征十分丰富，如肤色特征、轮廓特征、镶嵌图特征、直方图特征、器官对称性特征和投影特征、模板特征等。如何充分利用这些特征，是人脸检测要研究的一个关键问题。人脸具有复杂而细致的变化，因此一般需要采用多种特征综合的方法。

通过长期研究，人脸检测方法总的来说可分为基于启发式模型的方法和基于统计学习的方法。前者根据先验知识抽取人脸特征，然后进行建模，是人脸检测的传统方法。后者把人脸检测问题当作一个标准的模式识别问题来处理，通过对某一学习算法的训练，获得想要得到的人脸检测器。统计学习的方法也可以利用人脸的先验知识，所不同的是，这些知识是被融入到学习过程中被间接使用的。

1.4.1 基于启发式模型的方法

对一个典型的复杂背景下的人脸检测问题，基于启发式模型的方法首先要分割出一些初级的视觉特征，例如边缘特征、灰度特征、颜色特征等。这些视觉特征还比较低级，不能包含太多的信息量，需要对它做进一步的分析，即视觉特征还比较低级，不能包含太多的信息量，需要对它做进一步的分析，即特

题的方法，成功构造出人脸检测与跟踪系统，将为解决其他类似的复杂模式检测问题提供重要的启示。

1.3 人脸检测的难点

人脸检测问题的挑战性在于^[2]：

- 1) 人脸由于外貌、表情、肤色等不同而存在很大差异性，具有模式可变性；
- 2) 人脸上可能存在眼镜、胡须等附属物；
- 3) 作为三维物体的人脸的影像不可避免地受到光照的影像，光照的不均匀性特别是偏光的影响极大破坏了人脸特征空间分布的规律性。
- 4) 自然景物中存在大量与人脸相似的图像，如斑驳的树叶、老式建筑物的外表面等，这些复杂背景都极易与人脸图像相混淆。

1.4 目前的人脸检测方法

人脸图像所包含的特征十分丰富，如肤色特征、轮廓特征、镶嵌图特征、直方图特征、器官对称性特征和投影特征、模板特征等。如何充分利用这些特征，是人脸检测要研究的一个关键问题。人脸具有复杂而细致的变化，因此一般需要采用多种特征综合的方法。

通过长期研究，人脸检测方法总的来说可分为基于启发式模型的方法和基于统计学习的方法。前者根据先验知识抽取人脸特征，然后进行建模，是人脸检测的传统方法。后者把人脸检测问题当作一个标准的模式识别问题来处理，通过对某一学习算法的训练，获得想要得到的人脸检测器。统计学习的方法也可以利用人脸的先验知识，所不同的是，这些知识是被融入到学习过程中被间接使用的。

1.4.1 基于启发式模型的方法

对一个典型的复杂背景下的人脸检测问题，基于启发式模型的方法首先要分割出一些初级的视觉特征，例如边缘特征、灰度特征、颜色特征等。这些视觉特征还比较低级，不能包含太多的信息量，需要对它做进一步的分析，即特

征分析。在特征分析阶段，利用人脸的几何信息就可以把这些低级视觉特征组织成高级的、具有全局概念的人脸特征。通过特征分析，人脸区域就可以被确定下来。

1) 基于器官特征的方法

利用人脸五官的空间位置分布等规则，检测图像中是否有满足这些规则的图像块来检测人脸。Ki 和 Cipolla^[3]利用了大量的特征，包括几何、空间、灰度等各种度量，对人脸进行粗定位。人脸的几何关系由一个树型结构表示，并计算出概率属性，作为判断是否为人脸的条件。这是一种自底向上的方法。另外还有自顶向下的方法，即根据一个人脸模型（一般是正面人脸模型）先在一个比较大的范围内寻找人脸后选区，由粗到精地在一个最佳范围内定位人脸候选区，然后检测各种人脸器官特征。最为经典的就是 Yang 和 Huang^[4]提出的镶嵌图（Mosaic Image）的方法。Morimoto 和 Flickner^[5]在一个鲁棒性强的瞳孔定位技术的基础上，结合人脸结构信息，实现了多人脸的检测。

2) 基于彩色信息的方法

人脸的肤色在颜色空间中的分布相对比较集中，利用这个特点可以检测人脸。这种方法的最大优点是对姿态变化不敏感。用彩色信息检测人脸的关键是合理选择色度坐标。常用的方案是将彩色的 R、G、B 分量归一化。目前人们研究最多的是如何提取彩色的色度信息，即将 RGB 彩色空间转化为其他彩色空间，以突出色度信息。

Martin^[6]设计了肤色模型表征人脸颜色，利用一个感光模型对输入图像修正和补偿，并建立人脸颜色分类器，其输出作为神经网络的输入进行检测。Tony 等^[7]采用高斯混合模型（GMM, Gaussian Mixture Model）表示人脸肤色 R、G、B 各分量的统计分布，通过阈值比较判断像素是否为人脸像素。DAI 等^[8]利用彩色信息并结合人脸的纹理特征，设计了 SGLD 共发矩阵方法进行检测。Wu^[9]通过在 XYZ 彩色空间中，对人脸肤色和头发颜色的分布建立基于模糊逻辑理论的描述模型，通过计算隶属度来确定人脸肤色区域。

3) 基于形状分析的方法

从几何的观点和统计的角度来分析人脸的形状特征，主要是人脸的大小、对称性、形状参数等信息，这种方法作为细检测的途径，通常和基于彩色信息的方法结合在一起使用。

4) 基于模板匹配的方法

模板匹配的方法一般是先定义一个标准人脸模板, 计算输入图像与模板的似然度, 然后确定一个似然度阈值, 用以判断该输入图像中是否包含人脸。标准人脸模板可以是固定的样板, 也可以是带参变量的曲线函数, 即变形模板。

梁路宏等^[10]提出“双眼-人脸”模板的方法, 将平均脸的双眼模板剪裁出来, 检测人脸时先用双眼模板再使用人脸模板, 以提高匹配速度并取得更好的性能。

刘翼光^[11]将人脸图像二值化后, 得到人脸轮廓, 然后与模板库中不同大小的人脸模板匹配, 根据图像与模板之间的距离及相关阈值判定取舍。

Yuille 等^[12]提出用可变形模板来描述人脸的形状信息。所谓可变形模板, 是对眼睛、嘴巴等面部器官形状的一种参数化描述, 例如眼睛可以用一个圆外加两条抛物线来表示。同时, 与模板的性质相对应, 定义一个与图像边缘、峰值、谷值等相关的能量函数。将模板动态地作用于图像, 通过修改其参数使能量函数达到最小, 也就是通过模板的变形在图像中找到其最佳匹配。马宇飞等^[13]利用广义 Hough 变换的思想, 设计出了一种匹配人像边缘曲线可变形模板, 提高主持人广播帧检测方法的通用性。山世光等^[14]设计了一种由粗到精的面部特征提取策略, 在一定程度上解决了可变形模板参数初值确定的问题。

1. 4. 2 基于统计模型的方法

此类方法将人脸区域看作一类模式, 即模板特征, 使用大量人脸与非人脸样本训练、构造分类器, 通过判别图像中所有可能区域属于哪类模式的方法实现人脸检测。实际上是把人脸检测问题转化为统计模式识别的二分类问题。

1) 基于特征空间的方法

此类方法将人脸区域图像变换到某一特征空间, 根据其在特征空间中的分布规律划分人脸与非人脸两种模式。

主元分析法 (Principal Component Analysis, PCA) 是一种常用的方法, 它根据图像的统计特性进行正交变换 (K-L 变换), 以消除原有向量各个分量间的相关性, 变换得到对应特征值依次递减的特征向量, 即特征脸^[15]。Moghaddam 等发现人脸在特征脸空间的投影聚集比较紧密, 因此利用前若干张特征脸将人脸向量投影到主元子空间 F 和与其正交的补空间 \bar{F} , 相应的距离度量分别称为 DIFS (Distance in Feature Space) 和 DFFS (Distance from Feature Space)。对于人脸检测问题, 由于没有考虑非人脸, 样本的分布需要同时使用 DIFS 和 DFFS

才能取得较好的效果。属于特征空间方法的还有因子分解方法 (Factor Analysis, FA) 和 Fisher 准则方法 (Fisher Linear Distribution, FLD)。此外, 小波变换也被用于人脸检测, 提取人脸的多分辨率特征作为分类的依据^[16]。

2) 基于人工神经网络的方法

人工神经网络 (ANN) 方法是把模式的统计特性隐含在神经网络的结构和参数之中, 对于人脸这类复杂的、难以显式描述的模式, 基于神经网络的方法具有独特的优势。

Sung 和 Poggio^[17]首先对人脸样本集和非人脸样本集聚类, 以测试样本与人脸样本集和非人脸样本集的子类之间的距离作为识别特征向量, 利用多层感知器 (MLP) 网络作为分类器。

Rowley 等对正面端正人脸和正面旋转人脸的检测单独进行了研究^{[18][19]}。对于正面端正的人脸, 仅使用了三层前向网络: 输入层结点一一对应 20×20 像素的图像区域; 中间层结点分为对应不同人脸区域的若干组, 与输入层部分连接; 输出层一个结点, 输出值从 -1 到 1 区间的值表示这个区域是否为人脸。使用相同的人脸样本和不同自举过程收集的非人脸样本训练了多个正面人脸检测 ANN, 对它们的检测结果进行仲裁, 以进一步减少错误报警。对于正面旋转人脸的检测使用了旋转角度检测器 ANN 以及正面人脸检测器 ANN, 并使用相似的多个 ANN 仲裁方法降低错误报警。

3) 基于隐马尔科夫模型的方法

用于描述信号统计特性的隐马尔科夫模型 (Hidden Markov Models, HMM) 目前也被应用于人脸检测, 一般只使用人脸样本进行训练, 主要针对用于人脸识别的头肩部图像。Nefian 等^[20]根据正面人脸由上到下各个区域 (头发、额头、双眼、鼻子、嘴) 具有自然不变的顺序这一事实, 使用一个包含五个状态的一维连续 HMM 加以表示, 将头部图像按照这五个区域划分为互有重叠的条块, 对各块进行 K-L 变换, 选取前若干个变换系数作为观测向量训练 HMM。后来还提出了基于嵌入式 HMM 的人脸检测方法^[21], 该方法同时考虑到人脸由左到右各个特征的自然顺序, 使用了二维 HMM 并且采用二维 DTC 变换的系数作为观察向量。

4) 基于支撑向量机的方法

支持向量机 (Support Vector Machine, SVM) 是 Vapnik 等提出的基于结构风险最小化原理 (Structural Risk Minimization Principal, SRM) 的统计学习理论

[22], 用于分类和回归的问题。SRM 使 VC (Vapnik Chervonnikis) 维数的上限最小化, 这使得 SVM 方法比基于经验风险最小化 (Empirical Risk Minimization Principal, ERM) 的人工神经网络方法具有更好的泛化能力。

Osuna 等^[23]将 SVM 方法用于人脸检测, 对每一个 19×19 像素的检测窗口使用 SVM 进行分类, 以区分人脸和非人脸窗口。SVM 的训练使用了大量人脸样本和自举方法收集的非人脸样本, 并且使用逼近优化的方法减少支持向量的数量。SVM 的方法用于人脸检测也取得较好的效果。

由于长期以来的训练 SVM 需要求解计算复杂度极高的二次规划问题, 限制了该方法的应用。Platt^[24]提出的 SMC (Sequential Minimal Optimization) 算法解决了 SVM 训练困难的问题。

5) Adaboost 方法

Paul Viola 和 Michael Jones 于 2001 年提出了基于 Adaboost 的人脸检测算法^[25], Adaboost 算法通过一个迭代训练过程将一些弱分类器组合得到一个强分类器。在第一次训练出一个弱分类器后, 训练样本的权重得到调整, 从而使没有被第一次训练出的分类器正确分类的样本的权重增加。如此迭代下去, 最终得到的分类器是对每次训练得到的弱分类器的一个线性组合。由于其中的每个简单分类器使用了计算速度非常快的特征, 因此从根本上解决的检测的速度问题。Li 通过利用 3 个视角的检测器, 实现了一个多视角的人脸检测算法。

1.5 肤色信息在人脸检测中的应用

上面在对人脸检测方法进行综述的过程中, 采取的分类原则主要从方法论的角度出发。当然, 分类原则不同自然会产生不同的分类结果。例如, 若以检测过程中是否利用了色彩或肤色信息, 我们可以分为基于彩色信息的人脸检测算法、基于灰度信息的人脸检测算法以及彩色信息与灰度信息相结合的人脸检测算法。在彩色信息与灰度信息相结合的人脸检测算法中, 我们又可以根据如何应用肤色信息这一角度再进行细分, 例如, 可以分为肤色信息作为前期预处理的方法、肤色信息作为后期验证的方法等。

目前, 色彩信息的利用已经越来越广泛地应用在计算机视觉等研究领域, 肤色信息笼统地说就是通过某种方法判断出彩色图像中哪一点或哪一块区域是肤色。正如前面的分析所示, 将肤色信息应用到人脸检测中, 会起到积极的作

用,但如何利用肤色信息也有许多不同的考察问题的角度。

1) 将肤色判断作为人脸检测的预处理

这是肤色信息在人脸检测算法中比较常见的应用,目前的算法大多采用肤色区域分割与人脸验证相结合进行。

对于彩色图像,在确定肤色模型之后,首先可以进行肤色检测。在检测出肤色像素后,需要根据它们在色度上的相似性和空间上的相关性分割出可能的人脸区域,同时利用区域的几何特征或灰度特征进行是否是人脸的验证,以排除其它色彩类似肤色的物体。区域分割与验证在很多方法中是密切结合、统一考虑的。

在一些情况下,仅根据肤色像素的聚积特征即可完成区域分割。Yoo 等^[26]利用肤色像素的连通性分割出去与,使用椭圆拟合各个区域,根据区域长短轴的比率判断是否是人脸。Terrillon 等^[27]利用颜色信息得到肤色区域,然后利用运动信息来判定人脸和手。Gergesha 等^[28]利用面部几何特征,以及利用肤色概率映射迭代阈值化得到图像中可能的脸部区域,然后再利用一种快速的眼睛定位技术进行验证。

对于较为复杂的情况,则需要考虑两方面的问题:(1)由于光照和脸部器官的影响,人脸可能被分裂为若干互不连通的肤色区域;(2)人脸区域可能与其它类肤色区域连接在一起。

聚类-归并-验证策略是较常用的解决方法:首先将肤色像素按照较为严格的颜色一致性和几何约束条件聚类为取与,然后按照一定规则进行归并,归并过程中利用其它特征进行验证。Rein-Lien Hue 等^{[29][30]}使用一种光线补偿技术对输入图像的 R、G、B 值进行校正,将校正后的 R、G、B 值变换到 YcbCr 色彩空间,然后在 YcbCr 空间中进行非线性变换,使用椭圆肤色模型在变换后的空间中检测肤色像素,利用肤色相似性与空间相邻性合并肤色区域,得到候选人脸区域,最后构建眼睛、嘴的轮廓映射对候选区域进行验证,确定是否为人脸。

另一种策略是先用较弱的条件将肤色像素聚类为区域,再将符合一定条件下的区域分裂开。还有一些方法使用肤色模板进行搜索匹配,等等。

斯坦福大学图像组 EE368 也在这方面进行了大量的研究,于 2002-2003 年已经发表了许多关于这方面的研究文献^[31],主要利用肤色分布特征,结合灰度信息或者统计特性,处理彩色图像中的人脸检测问题。

2) 将肤色判断作为人脸检测的后期验证

既然肤色判断可以作为人脸检测的预处理，那么在没有使用肤色信息的人脸检测算法的后期，我们可以增加一个用肤色判断来进行验证的环节，它不但可以用来衡量前期的检测算法的结果好坏，也完全可以集成到一个完整的人脸检测算法中去。有实际工作证明这样的肤色信息后期验证对于提高人脸检测算法的正确率、降低误报率有很大的帮助。

几种不同的肤色模型：

(1) 基于 RGB 肤色模型

RGB 色彩空间是最常用的色系坐标。由于彩色图像是多光谱图像的一种特殊情况，其对应与人类视觉的三基色，即红、绿、蓝三波段，它是对人眼的光谱量化性质的近似。因此，利用 R、G、B 三基色的这三个分量来表征颜色是很自然的一种格式。所以，一般摄像系统都采用 RGB 色系坐标，通常图像颜色就用这种色系坐标标识。在 RGB 色系坐标中，每一个数字图像中像素点的颜色都可以用三维空间中的第一象限中的一个点来表示。当 RGB 颜色空间的任何颜色分量都为零，表示为黑色；三个分量达到最大值时，表示为白色；而当三个分量相等时，标识为灰度。一般以 RGB 色系坐标系为基础描述其它色系坐标系，将其它色系坐标系的基色描述为 RGB 三色的线性或非线性函数。

定义了基于 RGB 的肤色模型如下：

$R > 95$ ，且 $G > 40$ ，且 $B > 20$ ，且 $\max\{R, G, B\} - \min\{R, G, B\} > 15$ ，且 $\text{abs}(R - G) > 15$ ，且 $R > G$ ，且 $R > B$ 。

从上边的肤色模型定义中可以看出，这种模型是通过分别统计 RGB 三种分量的数值特征取得的，基本是线性关系，也没有考虑光照等因素，所以，针对有复杂光照的图像其检测效果就不理想了。

(2) 归一化 RGB 肤色模型

在一定的光照条件下，人脸的肤色在归一化的 RGB 色彩模型里符合高斯分布，即每个像素的亮度值为 $I = R + G + B$ ，其归一化的三种颜色分量为：

$$r = R/I; \quad g = G/I; \quad b = B/I, \quad \text{且 } r + g + b = 1. \quad (1.1)$$

其中 r , g , b 三者中的两者是独立的。

高斯模型 $N(\mu, \sum^2)$ 可用来表现出不同人的肤色在色度空间的聚类特性，其中：

$$\mu = (\bar{r}, \bar{b}) \quad (1.2)$$

$$\bar{r} = \frac{1}{N} \sum_{i=1}^N r_i \quad (1.3)$$

$$\bar{b} = \frac{1}{N} \sum_{i=1}^N b_i \quad (1.4)$$

其中 N 为用于建模的像素点总数。

$$\Sigma = \begin{pmatrix} \sigma_{rr} & \sigma_{rb} \\ \sigma_{br} & \sigma_{bb} \end{pmatrix} \quad (1.5)$$

对于参数 \bar{r} , \bar{b} , σ_{rr} , σ_{rb} , σ_{br} , σ_{bb} , 这六个参数都可以通过实验确定, 值分别为:

$$\bar{r}=156.5599, \bar{b}=117.4361, \sigma_{rr}=160.1301,$$

$$\sigma_{rb}=12.1430, \sigma_{br}=12.1430, \sigma_{bb}=299.4574.$$

使用归一化 RGB 色彩空间中的 r , g , b 只去除了 R , G , B 中的相对亮度成分, r , g , b 中仍存在亮度信息, 因此这种算法的亮度适应性不好。

(3) 基于 YIQ 的肤色模型

在 YIQ 色系坐标中, Y 分量标识颜色的亮度, I 和 Q 是两个相互正交的分量, 用来标识色度信息。这种颜色空间充分利用了人眼的色分辨力。从 RGB 空间向 YIQ 空间转换的公式如下:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.320 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1.6)$$

对于 YIQ 色系空间, 肤色满足如下条件:

$$(Y \leq 44) \parallel (Y \geq 223) \parallel (I \leq 0) \parallel (I \geq 64)$$

在该颜色空间中, 肤色的色度值和光源的光谱成分密切相关。可以看出, 这种肤色模型提取了亮度信息和色度信息, 其肤色模型比较简单。因此, 肤色聚类的结果只能应用于特定的光照条件下, 即亮度在 $(Y \leq 44) \parallel (Y \geq 223)$ 范围内, 不能有高光和阴影的干扰。

(4) 基于 YCbCr (或称 YUV) 的肤色模型

这是彩色电视体系的色彩表示体系, 它兼容老式黑白电视, 在多媒体技术中广泛应用。它可以从 RGB 格式线性变化得到。与 RGB 相比, 它的最大优点在于只需占用极少的带宽, 其中, Y 表示亮度, 即灰度值, UV 表示的是色度。

这种颜色空间与 YIQ 很相似。

YCbCr 色彩空间具有如下一些优点：

1. YCbCr 色彩空间具有与人类视觉感知过程类似的构成原理。
2. YCbCr 色彩空间被广泛的应用于电视显示等领域中，也是许多视频压缩编码，MPEG、JPEG 等标准中普遍采用的颜色表示格式。
3. YCbCr 色彩空间具有与 HIS 等其他一些色彩格式相类似的将色彩中的亮度分量分离出来的优点。
4. 相比 HIS 等其他一些色彩空间，YCbCr 色彩空间的计算过程和空间坐标表示形式比较简单。

从 RGB 颜色向 YCbCr 颜色转换的公式如下：

$$\begin{aligned} Y &= 0.299 * R + 0.587 * G + 0.114 * B \\ Cb &= (B - Y) * 0.565 + 128 \\ Cr &= (R - Y) * 0.713 + 128 \end{aligned} \quad (1.7)$$

则 YCbCr 肤色模型如下所示：

$$\frac{(x - e_{c_x})^2}{a^2} + \frac{(y - e_{c_y})^2}{b^2} \leq 1 \quad (1.8)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} * \begin{bmatrix} C_b - c_x \\ C_r - c_y \end{bmatrix} \quad (1.9)$$

其中， $c_x = 109.38$ ， $c_y = 152.02$ ， $\theta = 2.53$ ， $e_{c_x} = 1.6$ ， $e_{c_y} = 2.41$ ， $a = 25.39$ ， $b = 14.03$ 。

对于待检测图像中的像素，使用 (1.7) 式得到了 Cb, Cr 值，然后代入 (1.9) 式计算得到的 (x, y) ，最后代入 (1.8) 式是否满足。如果满足，则说明此像素值落在了椭圆肤色模型内部，即该像素为肤色。

这种模型很好地考虑了肤色模型是非线性，其肤色模型是近似于椭圆。所以，这种模型比较理想。在^{[36][37]}中使用了此肤色模型进行人脸检测，取得了较好的结果。

(5) 基于 KL 颜色变化的肤色模型

各种肤色定位方法中，关键问题是寻找到一个合适的色系坐标系，使得这个坐标系的变化集中在某二维上，由此设定阈值以确定肤色，使得在此阈值范围

内的非肤色越少越好。而 KL 变换正是具有这种集中变化的作用，因此我们将 KL 变换应用到肤色定位上，建立 KL 肤色坐标系。

令 $\mathbf{X} = [R, G, B]^T$ 为原始的三基色， $\mathbf{Y} = [K_1, K_2, K_3]^T$ 为变换后的三基色。采样足够的有代表性的肤色样本构成 \mathbf{X} ，计算 \mathbf{X} 的协方差矩阵：

$$\mathbf{C}_X = \mathbf{X}\mathbf{X}^T \quad (1.10)$$

由 \mathbf{C}_X 的特征向量组成的矩阵 Φ （同样按对应的特征值从大到小组合）即为我们所求的变换矩阵。此时：

$$\mathbf{Y} = \Phi^T \mathbf{X} \quad (1.11)$$

并且 \mathbf{Y} 的三个分量相互独立。

为建立 KL 肤色坐标系，首先从人像数据库中选取足够多的有代表性的小块肤色区域，对其进行 KL 变换，得到变换矩阵为：

$$\begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} = \begin{bmatrix} 0.666 & 0.547 & 0.507 \\ -0.709 & 0.255 & 0.657 \\ 0.230 & -0.797 & 0.558 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1.12)$$

这 4000 个肤色样本的值域为：

$$110.2 < K_1 < 376.3, \quad -61.3 < K_2 < 32.9, \quad -18.8 < K_3 < 19.5 \quad (1.13)$$

这一阈值即可作为 KL 肤色过滤器的阈值。对任意待处理彩色图像，先用式 (1.12) 对其进行色彩转换，然后按式 (1.13) 的阈值进行判断：在该范围内的置 1，否则置 0，即可得到过滤肤色后的二值图像。在此二值图像上可实现人脸的初定位。

1.6 人耳生物特征

人耳的结构并非完全没有规律可循，而是像人脸那样具有一定的标准特征。通常，一只标准的耳朵由以下部分构成：外耳轮廓、对耳轮（处于外耳轮内侧，并与其平行）、耳垂以及独特的 U 形凹口（位于耳洞和耳垂之间）。

之所以将人耳作为一种生物特征，是因为其满足了生物特征所应具备的四个基本性质：

1. 普遍性。人人都具有这种生物特征。
2. 唯一性。美国犯罪学专家 A Iannarelli 进行了两组实验。第一组比较

California 随机抽取的 10 000 只人耳样本；第二组比较兄弟和双胞胎的耳型。研究发现所有被测人耳都是唯一的。尽管双胞胎长相相似，但是耳朵的结构尤其是外耳和耳垂部分是不同的。该研究结果表明利用人耳生物特征进行个体识别具有很大的可能性。

3. 稳定性。据医学研究的结果表明：人耳在婴儿出生 4 个月之后，是成比例生长的，整体结构比率基本保持不变，唯有耳垂在 4 个月到 8 岁之间较之耳朵的整体生长有较大的拉伸，大约是其整体生长速度的 4 倍，但是 8 岁到 70 岁耳垂的拉伸速度同人耳整体结构生长速度基本保持一致，70 岁以后耳垂的拉伸速度才继续加快。由此可见在 8 岁到 70 岁之间，每个人耳朵的结构特征是独一无二的。

4. 可采集性。采集人耳图像不涉及接触式指纹扫描中存在的卫生问题，也不会像测量虹膜和视网膜那样被人们带来紧张情绪。虽然不如人脸或姿态容易捕获，但相对虹膜、视网膜和指纹，人耳面积较大，所以容易捕获，并且可以在不打扰个体正常活动的情况下进行采集，用户容易接受。

相对于人脸图像来说，人耳图像不受表情、化妆的影响，也不易损伤，经过处理后耳环、眼镜等影响也可消除。此外，人耳图像具有更一致的颜色分配，在转化为灰度图像时信息丢失少，而对于人脸图像进行同样的操作，眼睛的颜色信息几乎全部丢失。人耳图像表面更小（大约为人脸的 1/20 到 1/25），信息处理量更少。

1.7 主要研究内容

本文就现有的一些人脸检测算法进行改进，以争取获得精度和速度的一个较优的折衷。人脸检测算法可以通过不同方法对系统加速，例如，选择计算量小的特征。由于肤色特征计算量相对较少，但是只使用单一的肤色特征一般情况下误检率比较高，检测精度也有限，因此通常将它用在一个检测系统的初始阶段，对图像中的人脸位置进行初定位，然后再用其它的特征进行判断，去除初检结果中的假脸，这样可以既提高检测的速度又保证了检测的精度。再如，使用 ICA 来降低特征空间的维数，从而提高了检测速度。基于 Boosting 和 Cascade 算法的系统在速度方面具有很大的优越性。但是要进一步提高检测精度，就需要级联更多的强分类器，这样会降低检测速度。另外，在基于模板匹配的人脸

检测算法中使用多种模板进行匹配，在提高检测率的同时却降低了检测时间，因此需要根据实际应用对检测时间的要求来对模板数目进行控制。

1.8 论文创新点

- 1) 使用核独立分量分析算法构造重建函数的信噪比，依据重建图像的信噪比构造阈值函数，将该阈值函数当成弱学习算法。
- 2) 在模板匹配中，提出了一种基于边缘检测的人脸平均模板。对于遮挡和重叠的人脸区域，用基于边缘检测的人脸平均模板的一半再次进行模板匹配。
- 3) 通过求得候选区域的倾斜角，把模板按照这个角度进行旋转后再进行检测，一定程度上解决了平面内旋转人脸的检测问题。
- 4) 使用了基于边缘检测的平均耳模板进行模板匹配，一定程度上解决了深度旋转角度过大的人脸的检测问题。
- 5) 在学习和检测过程中，对人脸和非人脸先经过核独立分量分析算法来提取特征得到对应模式的特征向量，在算法中运用不完全 Cholesky 分解来降低时间复杂度。对求得的特征采用遗传算法再进行了一次特征选取以便有效地减少特征空间的维数。

1.9 论文的安排

第一章，绪论。首先介绍了人脸检测研究的背景、意义、难点以及国内外研究动态几个方面。然后介绍了肤色信息在人脸检测中的意义以及应用。最后介绍了人耳生物特征的性质以及较之人脸生物特征所具有的优势。

第二章，介绍了一种新的基于核独立分量分析和遗传算法的 SVM 人脸检测算法。

第三章，介绍了链式 Boosting 的结构、训练算法以及对算法的一些优化，然后将链式 Boosting 运用到人脸检测中。这里使用核独立分量分析算法构造重建函数的信噪比，依据重建图像的信噪比构造阈值函数，将该阈值函数当成弱学习算法。

第四章，提出了一种新的基于肤色和模板匹配的人脸检测算法。提出了边缘平均脸模板、边缘平均半脸模板、耳模板和变形模板。

第五章，总结和展望。对全文做了总结，提出了存在的一些问题，指出了今后应努力的方向和对未来工作的展望。

第2章 基于 KICA 的人脸检测算法

本章针对复杂背景的彩色静态图像,提出基于肤色和独立分量分析的人脸检测算法。首先使用YCbCr肤色模型进行肤色提取;在噪声滤波和确定候选脸之后,进行独立分量的提取。独立成分分析(Independent Component Analysis, 简称ICA)法是基于信号高阶统计特性的分析方法。ICA不仅可以分割混合的信号,而且还可以有效地提取信号中的低级特征。ICA基强调了图像中的边缘信息。而且不象小波,ICA基是对训练数据自适应的。ICA是一种无监督学习算法,它已经被成功应用到许多领域,如语音处理、人脸识别等。这里创造性地使用了独立分量分析的一种新算法——核独立分量分析算法来提取独立分量;然后使用标准遗传算法进行二次特征的提取从而获得独立基的最优子集。而支持向量机是一种很强大的有监督学习算法。这里将ICA和SVM集成起来,研究基于有监督学习和无监督学习的混合的学习机制,并将其应用到人脸检测问题中。实验结果表明,采用这种ICA/SVM的混合学习方案,单独的SVM方法具有更好的效果,具有较高的检测率。

2.1 肤色提取

这里采用第一章介绍的方法来提取肤色区域,即使用 YCbCr 肤色模型。从 RGB 颜色向 YCbCr 颜色转换的公式如下:

$$\begin{aligned} Y &= 0.299 * R + 0.587 * G + 0.114 * B \\ Cb &= (B - Y) * 0.565 + 128 \\ Cr &= (R - Y) * 0.713 + 128 \end{aligned} \quad (2.1)$$

则 YCbCr 肤色模型如下所示:

$$\frac{(x - e_{c_r})^2}{a^2} + \frac{(y - e_{c_b})^2}{b^2} \leq 1 \quad (2.2)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} * \begin{bmatrix} C_b - c_r \\ C_r - c_b \end{bmatrix} \quad (2.3)$$

其中, $c_x = 109.38$, $c_y = 152.02$, $\theta = 2.53$, $e_{c_x} = 1.6$, $e_{c_y} = 2.41$, $a = 25.39$, $b = 14.03$ 。

对于待检测图像中的像素, 使用 (2.1) 式得到了 Cb, Cr 值, 然后代入 (2.3) 式计算得到的 (x, y) , 最后代入 (2.2) 式是否满足。如果满足, 则说明此像素值落在了椭圆肤色模型内部, 即该像素为肤色。

这种模型很好地考虑了肤色模型是非线性, 其肤色模型是近似于椭圆。所以, 这种模型比较理想。

2.2 噪声滤波

经过颜色空间转化和二值化处理后, 人脸图像不可避免地会出现噪声。此时, 在复杂背景的图像中可能会出现很多与人脸肤色相似的或相同的区域, 也可能是手臂和腿部区域被误检为人脸。为此, 我们采用了中值滤波器去除与肤色相似的噪声点, 同时希望保持图像的细节。中值滤波器的模板领域越大, 消除噪声的能力越强, 图像细节被去掉的越多。综合这两个因素, 我们选用 (3×3) 领域的中值滤波器。

2.3 确定候选脸

为了进一步排除分割、滤波后并不含有人脸的类肤色区域, 需要对获得的二值图像中的目标区域进行形状分析, 以实现算法简化。

通过像素间连通性的分析将每个区域标记出来后, 分别对每个区域进行垂直、水平扫描, 得到每个区域外围矩形的高 (H)、宽 (W), 计算每个区域的高宽比 ($S=H/W$)。将高宽比 S 在 0.5 到 2 范围内的区域粗略检测确定为候选脸区域, 而不满足这种比例关系的区域被去除。

2.4 基于 KICA 算法的独立特征提取

2.4.1 建立训练样本集

全文实验建立了一个非人脸库和一个利用 AR 人脸库构造的小型人脸图像集。非人脸库由 110 张背景中包含类肤色、类人脸轮廓和非人脸肤色区域的图像组成。人脸库由 30 个（男性 15 人，女性 15 人）不同测试者，每人 10 幅共 300 幅人脸图像组成。其中包括了不同光照条件下、不同表情、以及脸旋转（偏转、微低、微扬）、移位、戴眼镜、围巾等人脸图像。图像序号 1-10 分别表示正面端正、面带笑容、夸张表情、右侧光、左侧光、高光、戴墨镜、戴墨镜并有右侧光、戴墨镜并有左侧光、戴围巾等人脸。

ICA 的训练样本集由非人脸库中的前 90 张图像和人脸库中每人具有代表性的序号为 1、4、8、10 的 4 幅图像构成，整个训练样本集共 210 幅图像。人脸库中序号为 3、5、9 的 3 幅图像，共 90 张构成遗传算法训练集。

2.4.2 预处理

独立分量分析提取人脸特征依赖于图像间的空间灰度分布相关性，所以实验前必须先对人脸图像进行预处理。人脸的预处理包括：

（1）人脸图像的几何归一化（位置校准）

目的是解决由于成像距离和人的姿势造成的人脸尺度和头部旋转角度上的差异。具体包括：平面旋转矫正（歪正）、深度旋转矫正（扭脸）、尺度归一化和掩模四个环节。本文人脸库中，没有过度的脸部扭曲图像，所以只需对人脸图像进行了平面旋转矫正和尺度归一。经过几何校准，可以在一定程度上获得人脸表示的几何不变性。步骤是：

Step1 图像旋转：使两眼的连线在一条水平线上，这保证了人脸在图像平面内的旋转不变性。

Step2 图像裁剪：取两眼的距离为 d ，经过裁剪，在 $2d \times 2d$ 的图像内，使两眼中心点固定于 $(d, 1.5d)$ 处，保证了人脸在图像平面内的平移不变性。

Step3 图像缩小放大变换，得到统一大小的校准图像。规定校准图像的大小为 64×64 像素，则缩放倍数为 $B = 2d / 64$ 。这使得两眼距离 d 为定长（32 个像素点），保证了人脸在图像平面内的尺度不变性。

（2）直方图均衡

目的是使转换后的图像在各个灰度值的分布概率相同，这样对于原本过亮或过暗的训练样本，可提高对比度，减小训练样本对明暗变化的敏感度。

(3) 图像的亮度补偿

本文中采用的亮度补偿的思路是：把图片中亮度最大的5%的像素提取出来，然后线性放大，使得这些像素的平均亮度达到255。根据求得的系数把整个图片的亮度进行线性放大。

(4) PCA 降维

在进行独立分量分析前，适当的对观测信号降维，舍弃协方差矩阵中某些较小的特征值，可使得 ICA 的工作量大大减少。使用训练图像的主特征向量而非原始图像来训练得到 ICA 的基，这是因为：在 ICA 混合模型中，假定图像是独立元的线性混合，而主特征向量又是训练样本图像的线性组合，因此用训练图像的主特征向量代替训练图像不会丢失信息。对每张大小为 64×64 像素 ICA 训练图像集的“标准图像”的 PCA 降维过程为：

Step1 建立 ICA 训练集。令 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, $n = 210$, 为样本容量；

Step2 计算平均脸。

$$\mathbf{x}_{avg} = \sum \mathbf{x}_i / n \quad (2.4)$$

Step3 用每张脸减平均脸得到新的训练集：

$$\mathbf{X} = \{(\mathbf{x}_1 - \mathbf{x}_{avg}), (\mathbf{x}_2 - \mathbf{x}_{avg}), \dots, (\mathbf{x}_n - \mathbf{x}_{avg})\} \quad (2.5)$$

Step4 计算矩阵 $\mathbf{X}^T \mathbf{X}$ 的特征向量：

$$\mathbf{p}_i = \frac{1}{\sqrt{\lambda_i}} \mathbf{X}^T \mathbf{v}_i \quad (2.6)$$

及其特征值 λ_i , $i = 1, \dots, n$, 其中 \mathbf{v}_i 是 $\mathbf{X}\mathbf{X}^T$ 的特征向量；

Step5 将特征值从大到小排序： $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, 其对应的特征向量为 \mathbf{p}_i 。

这样，每一幅人脸图像都可以投影到由 \mathbf{p}_i 张成的子空间中；

Step6 选择 120 个特征值最大的特征向量 \mathbf{p}_j ($j=1 \dots 120$), 并定义为矩阵

$$\mathbf{P}^T = \{\mathbf{p}_1^T, \mathbf{p}_2^T, \dots, \mathbf{p}_{120}^T\}.$$

(5) 白化处理

所谓“白化”即对观测数据施加一个线性变换，使得新向量 \mathbf{z} 的各个分量间

互不相关,同时 z 的协方差矩阵为单位阵: $E\{zz^T\} = \mathbf{I}$, 新向量 z 称为空间白色或简称为白色,该过程称为白化。

为减少分离结果幅度的不确定性,事先假设源信号 s 为单位方差,这一假设也同时使得 s 成为空间白色向量。将矩阵 P 的每一行数值减去行中值,结果送入一个零相位白化滤波器,则白化矩阵可表示为:

$$\mathbf{W}_p = 2(P^T P)^{-\frac{1}{2}} \quad (2.7)$$

从而得到空间白色:

$$\mathbf{Z} = \mathbf{W}_p \mathbf{P}^T \quad (2.8)$$

经过白化预处理后,分离矩阵 w 从 120×120 简化为 $120 \times \frac{120-1}{2}$, 矩阵自由度降低一倍。因此,白化处理减少了ICA的计算复杂度,将白化处理后的矩阵 z 作为ICA的输入矩阵进行独立分量分析。

2.4.3 独立特征提取

为找出人脸图像中统计独立的基图像,我们对标准人脸库进行独立分量分析(ICA)^[51]。根据图像合成模型,在 x 中的人脸图像可看成是由相互独立的源图像 s 经过线性混合得到的, A 是未知混合矩阵, $x = AS$ 。基于输出相互独立性原则,ICA可有效估计出分离矩阵 w ,将源图像从观测图像中分离出来,产生统计独立的输出矩阵 u 。

ICA用于人脸描述时,将每一幅二维图像采用行堆叠方式变成一维向量,构成输入矩阵 x 的行向量,则ICA的输出矩阵 $u = wx$ 的行向量也为图像, u 是独立源图像的重现,即人脸图像的独立基图像(或称为一组独立的面部特征)。正是这些独立基图像的不同线性组合描述了人脸库中的每一幅图像。混合矩阵 A 的每一个行向量惟一对应着每幅人脸的线性组合系数, A 可以通过计算分离矩阵 w 得到: $A = W^{-1}$ 。

ICA用于人脸检测时,将每一幅待检测人脸图像在独立基图像构成的人脸图像空间上投影,依据基图像的不同投影系数来描述该待检测人脸,并根据投影系数,在特征空间设计适当的分类器来实现分类识别。设 x_f 为待检测人脸,在人脸空间中有:

$$x_f = a_1 u_1 + a_2 u_2 + \dots + a_n u_n \quad (2.9)$$

(a_1, a_2, \dots, a_n) 为投影系数。

目标函数（优化判据）的建立和寻优算法（优化算法）的设计是 ICA 的两个核心问题^[52]。其中 ICA 方法的稳健性主要取决于目标函数的选择，而算法的收敛速度、占有内存情况等则主要依赖于优化算法。从某种意义上说，同一个目标函数可以有不同的优化算法，而同一个优化算法也可以用来优化不同的目标函数。

核独立分量分析算法^[53]不是独立分量分析算法的核化版本，而是一种新的独立分量分析算法。本文提出的算法不是基于单个非线性函数，而是基于整个函数空间。在再生希尔核空间中进行，并且使用“核窍门”（kernel trick）来高效地搜索整个空间。使用整个函数空间而不是单个函数使得该算法适应于多种源并且对于多种源分布具有很好的鲁棒性。

早期的独立分量分析算法是基于对比函数，该对比函数是根据一个为特殊目的而选择的单一固定的非线性函数的期望而定义的。仔细选择这个非线性函数使得它的期望近似于互信息且使得该近似具有很好的鲁棒性能，从而得到更加复杂的算法。所有这些早期的算法都将注意力集中在对于特定非线性函数的选择上。

这里算法所采用的对比函数是基于再生希尔核空间的正规化相关性度量定义的。其中正规化相关性分析是一种度量两个多维变量之间线性关系的方式，定义为：

$$\rho_F^* = \max_{f_1, f_2 \in F} \frac{\text{cov}(f_1(x_1), f_2(x_2))}{(\text{var } f_1(x_1) + \kappa \|f_1\|_F^2)^{1/2} (\text{var } f_2(x_2) + \kappa \|f_2\|_F^2)^{1/2}} \quad (2.10)$$

则其近似值为：

$$\hat{\rho}_F^* (\mathbf{K}_1, \mathbf{K}_2) = \max_{\alpha_1, \alpha_2 \in \mathbb{R}^N} \frac{\alpha_1^T \mathbf{K}_1 \mathbf{K}_2 \alpha_2}{(\alpha_1^T (\mathbf{K}_1 + \frac{N\kappa}{2} \mathbf{I})^2 \alpha_1)^{1/2} (\alpha_2^T (\mathbf{K}_2 + \frac{N\kappa}{2} \mathbf{I})^2 \alpha_2)^{1/2}} \quad (2.11)$$

则该问题可以转化为广义特征值的求解问题，即：

$$\begin{pmatrix} 0 & \mathbf{K}_1 \mathbf{K}_2 \\ \mathbf{K}_2 \mathbf{K}_1 & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \rho \begin{pmatrix} (\mathbf{K}_1 + \frac{N\kappa}{2} \mathbf{I})^2 & 0 \\ 0 & (\mathbf{K}_2 + \frac{N\kappa}{2} \mathbf{I})^2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \quad (2.12)$$

这里用的对比函数是根据一个这个正规化相关性度量来定义的，该对比函数与相互信息有关。即：

$$C(W) = -0.5 * \log \delta^*(K_1, \dots, K_m) \quad (2.13)$$

其中 $\delta^*(K_1, \dots, K_m) = \det \kappa_x / \det D_x$ 为核泛化方差 (kernel generalized variance)。 κ_x 中对角线上的元素为 $K_i + (N * \kappa / 2) * I$, $\kappa_x(i, j) = K_i * K_j (i \neq j)$ 。 D_x 是对角阵, 其对角线上的元素为 $K_i + (N * \kappa / 2) * I$ 。 K_i 为 $N * N$ 维 Gram 阵。

然而这个算法牵涉到寻找 $m * N$ 维的矩阵的广义特征值, N 数据个数, m 数据的维数。简单地执行该算法, 其时间复杂度为 $O(m^3 N)$, 这样的算法复杂度使得该算法难于用于大型数据集。考虑到 Gram 阵的特征值有大到小变化很快, 所以对于 Gram 阵的低阶近似通常可以成功代替 Gram 阵应用于算法中, 从而降低算法复杂度, 使得其降为与数据点集的个数呈线性关系。即寻找一个矩阵 G , $N * M$ 维, $M \ll N$, 从而 $K - G * G^T$ 的模小于某个给定的值。这可以通过不完全的 Cholesky 分解完成。不完全 Cholesky 分解算法涉及一次从 Gram 阵 K 中选择一列, 选择的标准是使得近似误差的下降程度达到最大。经过 p 步之后, 得到的 G_p 为 $N * p$ 维。对余下的 $N - p$ 列进行排序, 这是通过比较 $K - G_p G_p^T$ 中的对角线上的元素得到的。不完全 Cholesky 分解的时间复杂度只有 $O(M^2 N)$, 而且由于只需要存储 K 的对角线元素, 所以其空间复杂度只有 $O(MN)$ 。

具体算法如下:

输入: $N \times N$ 维的半正定矩阵 K

精度参数 η

1. 初始化: $i = 1$, $K' = K$, $P = I$, for $j \in [1, N]$, $G_{ij} = K_{ij}$

2. While $\sum_{j=1}^N G_{jj} > \eta$

- $j^* = \operatorname{argmax}_{j \in [1, N]} G_{jj}$
- 更新 P : 令 $P_{ii} = 0$, $P_{j^* j^*} = 0$, $P_{ij^*} = 1$, $P_{j^* i} = 1$
- 令 $K'_{1:N, i} \leftrightarrow K'_{1:N, j^*}$, $K'_{i, 1:N} \leftrightarrow K'_{j^*, 1:N}$

- $\mathbf{G}_{i,li} \leftrightarrow \mathbf{G}_{j',li}$
- $\mathbf{G}_{ii} = \sqrt{\mathbf{K}_{ii}^{'}}$
- $\mathbf{G}_{i+1:N,j} = \frac{1}{\mathbf{G}_{ii}} (\mathbf{K}_{i+1:N,j}^{'} - \sum_{j=1}^{i-1} \mathbf{G}_{i+1:N,j} \mathbf{G}_{ji})$
- $\mathbf{G}_{jj} = \mathbf{K}_{jj} - \sum_{k=1}^i \mathbf{G}_{jk}^2$
- $i \leftarrow i+1$

3. 输出 \mathbf{P}, \mathbf{G} 和 $M = i - 1$ 。

输出: $N \times M$ 维的三角阵 \mathbf{G} , 排列矩阵 \mathbf{P} , 从而使得 $\|\mathbf{PKP}^T - \mathbf{GG}^T\| \leq \eta$ 。

我们将广义特征值求解的问题, 即 $\kappa_x \alpha = \lambda \mathbf{D}_x \alpha$, 转化为标准特征值求解问题,

即 $\mathbf{C}^{-T} \kappa_x \mathbf{C}^{-1} \beta = \lambda \beta$, 其中 $\mathbf{D}_x = \mathbf{C}^T \mathbf{C}$, $\beta = \mathbf{C} \alpha$, 从而

$$\kappa_x' = \mathbf{D}_x^{-1/2} \kappa_x \mathbf{D}_x^{-1/2} = \begin{pmatrix} \mathbf{I} & \mathbf{r}_x(\mathbf{K}_1) \mathbf{r}_x(\mathbf{K}_2) & \dots & \mathbf{r}_x(\mathbf{K}_1) \mathbf{r}_x(\mathbf{K}_m) \\ \mathbf{r}_x(\mathbf{K}_2) \mathbf{r}_x(\mathbf{K}_1) & \mathbf{I} & \dots & \mathbf{r}_x(\mathbf{K}_2) \mathbf{r}_x(\mathbf{K}_m) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_x(\mathbf{K}_m) \mathbf{r}_x(\mathbf{K}_1) & \mathbf{r}_x(\mathbf{K}_m) \mathbf{r}_x(\mathbf{K}_2) & \dots & \mathbf{I} \end{pmatrix} \quad (2.14)$$

其中 $\mathbf{r}_x(\mathbf{K}_i) = \mathbf{K}_i (\mathbf{K}_i + \frac{N\kappa}{2} \mathbf{I})^{-1} = (\mathbf{K}_i + \frac{N\kappa}{2} \mathbf{I})^{-1} \mathbf{K}_i$ 。如果我们求得了 κ_x' 的特征向量 α_i' , 可以通过 $\alpha_i = (\mathbf{K}_i + \frac{N\kappa}{2} \mathbf{I})^{-1} \alpha_i'$ 得到最终广义特征向量, 而特征值相同。

对于每个 \mathbf{K}_i 我们使用不完全 Cholesky 分解使得 $\mathbf{K}_i = \mathbf{G}_i \mathbf{G}_i^T$, 其中 \mathbf{G}_i 为 $N \times M_i$ 维, $M_i \ll N$ 。然后对 \mathbf{G}_i 进行奇异值分解获得 $N \times M_i$ 维单位正交阵 \mathbf{U}_i 和 $M_i \times M_i$ 维的对角阵 Λ_i , 满足关系式:

$$\mathbf{K}_i = \mathbf{G}_i \mathbf{G}_i^T = \mathbf{U}_i \Lambda_i \mathbf{U}_i^T \quad (2.15)$$

令 \mathbf{V}_i 表示 \mathbf{U}_i 的正交补阵, 从而使得 $(\mathbf{U}_i, \mathbf{V}_i)$ 是 $N \times N$ 维的正交单位阵, 则有:

$$\mathbf{K}_i = \mathbf{U}_i \Lambda_i \mathbf{U}_i^T = (\mathbf{U}_i, \mathbf{V}_i) \begin{pmatrix} \Lambda_i & 0 \\ 0 & 0 \end{pmatrix} (\mathbf{U}_i, \mathbf{V}_i)^T \quad (2.16)$$

$$\text{而 } \mathbf{r}_x(\mathbf{K}_i) = (\mathbf{K}_i + \frac{N\kappa}{2}\mathbf{I})^{-1}\mathbf{K}_i = (\mathbf{U}_i\mathbf{V}_i)\begin{pmatrix} \mathbf{R}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}(\mathbf{U}_i\mathbf{V}_i)^T = \mathbf{U}_i\mathbf{R}_i\mathbf{U}_i^T \quad (2.17)$$

其中 \mathbf{R}_i 是对角阵, 其对角线上的元素为对 Λ_i 的对角线上的元素施以 $\lambda \rightarrow \frac{\lambda}{\lambda + N\kappa/2}$

的变换而得到的。则前面的 κ_x 变成如下形式:

$$\kappa_x = \mathbf{U}^T\mathbf{R}_x'\mathbf{U} + \mathbf{V}^T\mathbf{V} = (\mathbf{U}^T\mathbf{V}')\begin{pmatrix} \mathbf{R}' & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}(\mathbf{U}^T\mathbf{V}')^T \quad (2.18)$$

其中 \mathbf{U}' 是 $mN \times mM$ 维, \mathbf{V}' 是 $mN \times (mN - mM)$ 维, \mathbf{R}'_x 是 $mM \times mM$ 维, $(\mathbf{U}^T\mathbf{V}')$ 是单位正

$$\text{交。 } \mathbf{U}' = \begin{pmatrix} \mathbf{U}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{U}_m \end{pmatrix}, \quad \mathbf{V}' = \begin{pmatrix} \mathbf{V}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{V}_m \end{pmatrix}$$

$$\mathbf{R}'_x = \begin{pmatrix} \mathbf{I} & \mathbf{R}_1\mathbf{U}_1^T\mathbf{U}_2\mathbf{R}_2 & \dots & \mathbf{R}_1\mathbf{U}_1^T\mathbf{U}_m\mathbf{R}_m \\ \mathbf{R}_2\mathbf{U}_2^T\mathbf{U}_1\mathbf{R}_1 & \mathbf{I} & \dots & \mathbf{R}_2\mathbf{U}_2^T\mathbf{U}_m\mathbf{R}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_m\mathbf{U}_m^T\mathbf{U}_1\mathbf{R}_1 & \mathbf{R}_m\mathbf{U}_m^T\mathbf{U}_2\mathbf{R}_2 & \dots & \mathbf{I} \end{pmatrix}$$

由于 κ_x 和 $\begin{pmatrix} \mathbf{R}'_x & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$ 相似, 所以:

$$\det \kappa_x = \det \mathbf{R}'_x \quad (2.19)$$

特征向量的关系是:

$$\alpha = \mathbf{U}'\beta \Rightarrow \beta = \mathbf{U}'^T\alpha \quad (2.20)$$

这里对我们的对比函数求导的核心子程序是对于 Gram 阵的导数的计算。主要问题是虽然 Gram 阵 \mathbf{K} 是对称半正定的, 但是它的导数是对称但不是半正定或者半负定的, 所以不完全 Cholesky 分解不能直接应用到这些导数上来求取低阶近似。但是对于高斯核, 导数可以表示为两个半正定的 Gram 阵之差或者之和。可以对这两个半正定的 Gram 阵分别施以不完全 Cholesky 分解。最终求解对比函数导数的时间复杂度为 $O(m^2M^2N)$ 。

预处理后, 将得到的行向量即 z_1, \dots, z_N 作为该算法输入, 采用高斯核。经过该算法后输出为 W , 即滤波器 W , 独立的基图像 U 可以通过滤波器组 W 而获得, 即: $W * Z = U$ 。

算法的整体框架如下:

- a. 按以下步骤求出 $C(W)$ (即对比函数)
 - a 1. 首先对估计出来的源 (z_1, \dots, z_N) 求其中心化的 Gram 阵, K_1, \dots, K_m , 其中

$$u_i = W * z_i, \text{ 其中的 } W \text{ 是一个参数矩阵。}$$
 - a 2. 通过不完全 Cholesky 分解来求解 $\det \kappa_x$
 - a 3. $C(W) = -0.5 * \log \delta^x(K_1, \dots, K_m)$, 其中 $\delta^x(K_1, \dots, K_m) = \det \kappa_x / \det D_x$
- b. 最小化 $C(W)$
- c. 将 W 的每个行向量正交归一化, 记为 $[W_1, W_2, \dots, W_N]$.

因此, 独立基图像 U 的不同线性组合重构了人脸库中的每一幅图像:

$$X_{recICA} = C_{train} * U, \quad C_{train} = W^{-1} \quad (2.21)$$

其中, 矩阵 C_{train} 的每一个行向量唯一对应着每幅人脸的线性组合系数。

2.4.4 独立基选择

由于图像数据维数高, 对求得的二次特征有必要再进行一次特征选取以便有效地减少特征空间的维数。我们采用遗传算法来求得独立基的一个最优子集。

遗传算法 (Genetic Algorithm, GA) 是一种模拟生物进行理论提出的仿生类优化搜索算法, 由 Michigan 大学的 Holland 教授于 1975 年首次提出^[54]。它模拟生物进化过程中的选择、交叉、变异操作, 用概率化的寻优方式, 自适应地获取和指导优化方向和空间。遗传算法的优点之一是只要问题空间 (目标函数) 能够被描述为求极值的形式, 就可以加以优化, 而且在遗传算法的搜索过程中也不要求目标函数是连续可微的约束。

遗传算法是具有“生成+检测”功能的迭代算法。它以适应度函数为评估依据, 通过对候选解群体中的个体施以遗传操作, 实现群体内个体结构 (模式) 的重组并进行迭代, 群体中的个体 (问题的解) 一代一代不断优化并逐渐逼近最优解。

下面我们具体给出遗传算法每一步选用的方法以及其参数。

Step1 种群规模。在遗传算法中，候选解与染色体一一对应，染色体个数即解群的大小。群体规模大小对算法优化的最终结果和执行效率有重要影响。标准遗传算法 SGA 是按随机方法在最优解分布范围内产生一定数目的个体组成初始种群。在人脸特征选取问题中，120 个独立基可以有 2^{120} 中不同的解，在可容忍的计算复杂度情况下，我们选择种群规模为 60。

Step2 编码。编码是将问题空间的各参数映射为遗传空间的染色体。各参数映射的染色体的基因呈一维排列称为一维染色体编码。一维染色体编码常采用二进制编码，即染色体由 0 和 1 的位串组成。二进制编码符合最小字符集编码规则，应用广，具有较大的优势。本文对独立基也采用二进制编码。

令染色体 q 的位串长度为 120，则染色体的每一位基因对应一个独立基。若某位取 1，表示保留相应的独立基；若取为 0，表示舍弃该独立基，由于大多数特征分量将被保留，在初始化种群时每一位取 1 的概率设为 0.8。

Step3 适应度函数。适应度函数反映了个体对环境的适应程度，用于父代选择。适应度函数对算法的收敛速度和结果有极大的影响。它可以定义为任意集合，任何非线性、不可微、非连续的正函数。本文给定染色体 q 的适应度函数为：

$$fit(q) = \sum_{f \in \omega} \delta(f, q) \quad (2.22)$$

ω 是用于遗传算法训练的人脸图像库，该适应度函数的意思是：将 ω 中的一张人脸图像 f 放在染色体 q 所选择的独立基空间中，若人脸图像 f 被正确检测，则 $\delta(f, q) = 1$ ，若 f 未被正确检测，则 $\delta(f, q) = 0$ 。 ω 中每一张人脸图像的检测结果之和即是该染色体的适应度值。显然， $fit(q)$ 的大小代表了染色体 q 在遗传算法训练样本集上被正确检测的人脸样本个数，其取值范围是 $0 \leq fit(q) \leq 90$ 。

Step4 选择算子。选择是一个对个体串依其适应度值进行挑选的过程（即亲本挑选过程），选择算子具有将种群集中于高适应度区域的能力。

赌轮法是一种简单的选择策略，但可能会导致最好的解群成员无法在下一代中产生子代成员，从而产生所谓的随机错误。同时，在适应度相差较大时，其它个体串几乎不会被选中，造成种群单一化，缺乏多样性。最优个体保留法

可以减少随机错误。因此我们采用最优个体保留法和赌轮法相结合的混合选择机制。

最优个体保留法是将适应度值最大的一个个体直接保留到下一代，其余 59 个染色体由赌轮法进行选择复制构成。过程为：首先对种群 ψ 中的所有个体求总适应度 f_{sum} ；接着每个个体按对应的适应度值比例占据盘面区域；然后，选择一个值在 0 到 f_{sum} 之间的随机数 n 。假设 ω 在每个染色体选择的独立基构成的人脸图像空间上的平均检测率为 80%，那么 ω 在种群所构成的 60 种不同的人脸图像空间上正确检测人脸图像的个数将是 $90 \times 80\% \times 60$ 张。所以我们选择 $n = 4320/2$ ；最后转动一次轮盘。若得到个体的适应度值加上前面所有个体的适应度之和大于或等于 n ，则选择该个体，它被完整保留到下一代。共转动 59 次，产生 59 个亲本。这种选择机制使染色体 q 被选择的概率是

$$P_q = f(q) / \sum_{q \in \psi} f(q)。$$

Step5 交叉算子。交叉算子的作用是按一定概率 P_c 交换父代染色体的部分基因，形成新的子代染色体。其目的是增加群体中个体的离散程度，在遗传空间产生最优解。我们采用最基本的单点交叉。

首先从被选中的 60 个染色体中随机挑选两个串进行配对，交叉的位置也随机选定，然后这些串在这个交叉点之后的部分被分割交换，取交叉概率 $P_c = 0.8$ 。这个过程产生两个新串。每个新串都在某种程度上与父代双亲相似。这种既有组织又有随机性的信息交换可以促进那些高于平均适应度值的个体更多地出现在下一代中。

Step6 变异算子。变异算子的作用是按一定概率改变子代染色体的基因，维持种群多样性，提高局部随机搜索能力，同时打破由于初始值选择不当造成的模式限制，避免算法早熟。根据二进制的编码方式，我们采用单点的位点变异。随机选择个体的一位取反，对每个染色体取变异概率 $P_m = 0.03$ 。

Step7 循环终止。按照上述迭代过程，直到每代中的最优染色体的适应度值保持一定的代数未改变，此时终止循环。最后那些被选中的独立基构成了新的特征矩阵： $U = [u_1, u_2, \dots, u_l]^T$ ，它们构成一个 1 维的特征子空间。

2.5 基于 SVM 的 ICA 特征分类

获取特征之后,就建立好了一个 SVM 训练样本集 $\{\mathbf{c}_i, d_i\}_{i=1}^N$, 其中 $d_i = \{1, -1\}$ 是 ICA 特征 \mathbf{c}_i 的类别, N 是训练样本集的大小。通常 SVM 是在约束条件:

$$\sum_{i=1}^N y_i \alpha_i = 0$$

$$0 \leq \alpha_i \leq C \quad i=1, \dots, N \quad (2.23)$$

下对 α_i 求解下列函数的最大值:

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \quad (2.24)$$

求解上述问题后得到的最优分类函数是:

$$f(\mathbf{x}) = \text{sgn}\left(\sum_{i=1}^N \alpha_i^* y_i K(\mathbf{x}_i, \mathbf{x}) + \mathbf{b}^*\right) \quad (2.25)$$

$\text{sgn}()$ 为符号函数, $K(\mathbf{x}_i, \mathbf{x}_j)$ 是内积函数, 主要有三类: 多项式内积函数, 核函数内积函数, S 型内积函数。该文采用核函数内积函数。由于非支持向量对应的 α_i^* 均为 0, 式中的求和实际上只对支持向量进行。 \mathbf{b}^* 是分类阈值, 可以用任意一个支持向量求得, 或通过两类中任意一对支持向量取中值求得。 C 为某个指定的常数, 它实际上起到控制错分样本惩罚的程度的作用, 实现在错分样本的比例与算法复杂度之间的折中。

通过求解上面的约束优化问题, SVM 试图最大化两类训练样本的分类间隙或分类间隔, 并同时最小化训练误差 (为 $2/\|\omega\|$) 如图所示:

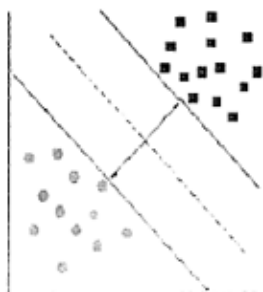


图 2.1 最优分类面

2.6 实验结果与分析

将使用肤色模型提取出来的候选脸区域，先进行预处理。预处理包括：

- (1) 将二值图像中被框出的各个候选脸矩形区域映射到原待检测的彩色图像上，然后尺度归一化为 64×64 。
- (2) 然后进行直方图均衡，在一定程度上纠正不同相机的输入增益以及增强图像的对比度。
- (3) 图像的亮度补偿
- (4) PCA 降维（不一定有，可选步骤）
- (5) 白化处理。使得新向量的各个分量间互不相关，同时其协方差矩阵为单位阵，减少 ICA 计算复杂度。

我们将同样的这组测试集，经过人脸区域粗定位得到候选脸在不同的检测方法上作性能比较。结果如下：

表 2.1 各方法在测试集上的性能比较

检测方法	检测率 (%)	虚警数	平均检测时间 (s)
SVM 方法（不作任何特征提取）	75.1	26	26.8
FastICA ^[56] （不进行 PCA 降维，不作特征选择）	83.9	20	30.8
FastICA（PCA 降维，不作 GA 特征选择）	85.7	12	7.6
FastICA（PCA 降维，GA 特征选择）	89.1	7	5.8
KICA（不进行 PCA 降维，不作特征选择）	86.9	13	50.6
KICA（PCA 降维，不作 GA 特征选择）	88.6	9	14.5
KICA（PCA 降维，GA 特征选择）	90.2	5	11.6

从表 2.1 中数据可以看出：

- FastICA 算法和 KICA 算法在检测率和误检人脸数目上均优于传统的不作任何特征提取的 SVM 算法。
- 在 FastICA 算法和 KICA 算法中是否降维对检测速度有很大的影响。本实验

训练阶段是将 ~~210~~ 维降为 120 维，因此检测时间几乎快了四倍。

- 对降维后的 FastICA 算法和 KICA 算法提取的独立基是否作特征选择以获得最有利于分类的独立基图像，影响着算法的检测率。表中经过 PCA 降维得 FastICA 算法和 KICA 算法进行特征选择和不进行特征选择的检测率并不是提高的特别明显。这是因为，本文特征选择所使用的遗传算法非常适合于在巨量空间内搜索最优值，而经过降维后的独立基只有 120 维，所以选择后得到的最优子集的维数较高，也就是说，大部分的独立基向量会被选中，这使得在最优独立基构成的人脸图像空间上的检测率与一般独立基构成的特征子空间上的检测率相差不大。但是遗传算法对 FastICA 和 KICA 算法进行特征选择，在检测率，误检人脸数目和平均检测时间上还是体现出其优势。
- 无论是从检测率还是从误检数方面，KICA 算法都比 FastICA 算法来得优越，但是 KICA 算法要花费更多的检测时间。

使用 KICA 算法和 GA 算法提取特征的结果图如下所示。

通过对实验结果的分析，基于肤色和核独立分量分析的人脸检测算法在少许的约束条件下具有相当程度的误差容忍性，对部分环境采光变化，不同尺寸大小，各种姿态和表情，轻度侧面旋转，部分遮挡和比较复杂背景等情况有较高的检测率和较低的误检率，同时兼顾了检测速度。但是也存在一些问题，如对于具有大量类肤色区域的复杂背景图像，该算法的漏检和错误报警率会有所提高。





图 2.2 使用 KICA 和 GA 的部分结果图

大部分误检的主要原因是：(1) 在复杂背景的图像中有可能出现颜色和形状都与人脸近似的区域，如正面手掌等；

大部分漏检的主要原因有：(1) 有与肤色相似的区域与人脸区域相连，导致在确定候选脸的步骤中被判决条件否决；(2) 人脸深度旋转角度过大；(3) 图片中人脸处于暗处，彩色信息失效，导致肤色监测器将其滤掉了；(4) 图片中存在多人脸重叠。

2.7 本章小结

这里对传统的 SVM 人脸检测算法进行改进，即除了包含学习过程和检测过程之外，还多了一个对样本进行 ICA 特征提取部分。也就是说在学习和检测过程中，对人脸和非人脸进行预处理之后，还要经过 ICA 特征提取得到对应模式的特征向量（即新的分类依据），再采用 SMO 算法训练支持向量机。这里创造

性地采用核独立分量分析算法来进行独立分量的提取，并在算法中运用不完全 Cholesky 分解来降低时间复杂度。由于图像数据维数高，对求得特征再进行了一次特征选取以便有效地减少特征空间的维数。这里采用遗传算法来求得独立基的一个最优子集。

实验表明，在 FastICA 算法和 KICA 算法中是否降维对检测速度有很大的影响。用遗传算法得到的最优子集的维数虽然较高，但是对 FastICA 和 KICA 算法进行特征选择，在检测率，误检人脸数目和平均检测时间上也体现出一定的优势。此外，KICA 算法都比 FastICA 算法来得优越，但是 KICA 算法要花费更多的检测时间。

第3章 链式 Boosting 算法在人脸检测中的应用

由 Viola 提出的 AdaBoost^[32] 训练算法是序列式前向搜索过程, 它使用贪婪的选择策略。它的前提条件是单调性。而这个前提可能无法满足。而由 Stan Li 提出的 FloatBoost 训练算法^[33]将 Floating 搜索和 AdaBoost 融会贯通。由 FloatBoost 训练出的检测器具有更好的检测性能和较少的特征, 然而 FloatBoost 算法不稳定, 并且当处理比较复杂的问题时比较费时。

而链式 Boosting^[34] 训练算法在传统的 AdaBoost 算法的过程中使用历史信息。并且采用线性递归特征消除方法^[35]来消除 AdaBoost 中的冗余信息, 从而避免了局部最优。除此之外, 采用线性递归特征消除方法还获得了最优阈值, 避免了 AdaBoost 采用的阈值确定方法所带了一些问题。

这里将链式 Boosting 算法应用到了人脸检测器的训练中。首先创造性地使用核独立分量分析算法构造重建函数的信噪比, 依据重建图像的信噪比构造阈值函数, 将该阈值函数当成弱学习算法。再使用 Boosting 算法构造一个基于信噪比阈值的检测函数序列, 然后将它们组合成一个总检测函数, 形成一个 Boosting 分类器, 再使用链式 Boosting 算法用相同的方法构造一个个 Boosting 分类器(结点), 并将它们串起来形成链式 Boosting 结构, 据此可以判别一幅图像是否为人脸图像。

3.1 链式 Boosting 的结构

链式 Boosting 方法的主要思想就是将一个个 Boosting 分类器连成一个链形结构, 后一个 Boosting 分类器利用前一个 Boosting 分类器所产生的历史信息。设第 i 个结点分类器为:

$$\Phi_i = \text{sign} \left[\sum_{t=1}^{m_i} \alpha_{i,t} h_{i,t}(x) - b_i \right]. \quad (3.1)$$

为了使用第 i 层的历史信息, 定义:

$$h_{i+1,0}(x) = \sum_{t=1}^{m_i} \alpha_{i,t} h_{i,t}(x) \quad (3.2)$$

那么，第 $i+1$ 层可用新的特征 $h_{i+1,0}(x)$ 来初始化，即可表示为：

$$\Phi_{i+1} = \text{sign}\left[\sum_{t=0}^{m_{i+1}} \alpha'_{i+1,t} h_{i+1,t}(x) - b'_{i+1}\right] \quad (3.3)$$

设 $\alpha_{i+1,t} = \alpha'_{i+1,t} / \alpha'_{i+1,0}$ ，且 $b_{i+1} = b'_{i+1} / \alpha'_{i+1,0}$ ，则该公式可写成：

$$\phi_{i+1} = \text{sign}\left[\sum_{t=1}^{i+1} \sum_{j=1}^{m_j} \alpha_{t,j} h_{t,j}(x) - b_{i+1}\right] \quad (3.4)$$

因此，第 i 层分类器就被链到第 $i+1$ 层。重复这个过程 for $i=1, \dots, M$ ：

$$\phi_i = \text{sign}\left[\sum_{k=1}^i \sum_{j=1}^{m_j} \alpha_{k,j} h_{k,j}(x) - b_i\right] \quad (3.5)$$

这样层叠式 Boosting 就被连成一个链状结构，见图 3.1：

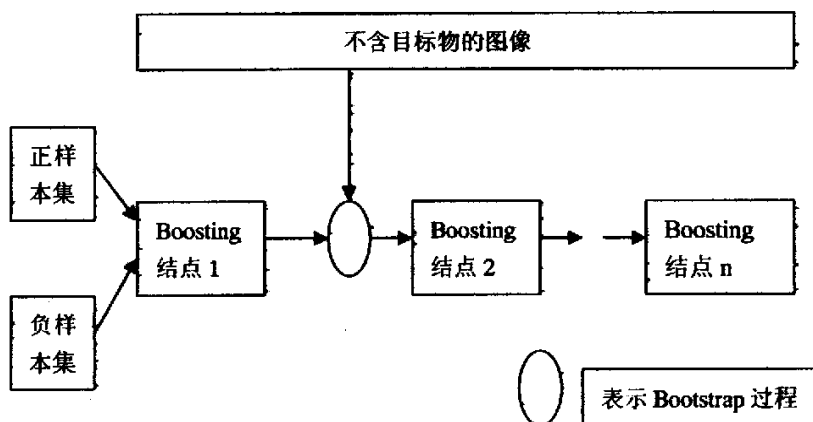


图 3.1 链式 Boosting 学习算法 (Bootstrap)

在链式 Boosting 的结构中，前面的分类器是后面的分类器的前缀，即：

$$\Phi_1 \subset \Phi_2 \subset \dots \subset \Phi_M \quad (3.6)$$

其中最后一个分类器结点包含了所有在链式 Boosting 中用到的特征。

3.2 链式 Boosting 的训练算法

与 AdaBoost 不同的是链式 Boosting 在 M 步的学习过程中, 使用一个正样本集和 M 个不同的负样本集。对于链式 Boosting 学习算法先作如下改进: 首先、正样例的权重可以根据正样例数直接给定, 负样例的权重将根据前一个分类节点的分类错误数进行调整, 在这里, 我们采用^[36]的权重调整方法:

$$w_j^{i+1} \leftarrow w_j^0 \exp[-y_j \sum_{k=1}^i \sum_{l=1}^{m_k} \alpha_{k,l} h_{k,l}(x)] \quad (3.7)$$

其次, 不引入 $h_{i,0}(x)$, 而是根据前面的 Boosting 分类器来进行后续的训练。

由上述改进得出一般的链式 Boosting 算法如下:

- (1) 初始化: $i=0$, $F_0=1$, $\phi=\{\}$, 对于所有的正样例 $x_j: w_j=1/p$; 对于所有的负样例 $x_j: w_j=1/n_i$ 。
- (2) 当 $F_i > F$ 时, 做
 - 1) $i=i+1$
 - 2) 用初始权重 w_j , 正样本集 P 和负样本集 N_i 上训练 ϕ_i , 以使 f_i , d_i 达到要求的值
 - 3) 优化结点训练器 (见 3.3)
 - 4) $F_i = F_{i-1} * f_i$, $\phi = \phi \cup \{\phi_i\}$
 - 5) 在非人脸图像集上评估 ϕ , 将错误的检测放入负样例集 N_{i+1}
 - 6) 对于每一个在 N_{i+1} 中的样例 x_j , 按权值修改公式

$$w_j^{i+1} \leftarrow w_j^0 \exp[-y_j \sum_{k=1}^i \sum_{l=1}^{m_k} \alpha_{k,l} h_{k,l}(x)] \text{ 修改权值。}$$

其中: P 为正的训练集, N_i 为第 i 个负的训练集, f_i 为第 i 层的最大误确定率, d_i 为第 i 层的最小检测率, w_j 为样例 x_j 的权重, F 为总的误确定率, D 为总的检测率, M 为算法中使用的分类器数目, ϕ_i 为第 i 个 Boosting 分类器, $h_{i,j}$ 为第 i 层第 j 个弱分类器, $\alpha_{i,j}$ 为弱分类器 $h_{i,j}$ 的参数, b_i 是分类器 ϕ_i 的阈值, m_i 是 ϕ_i 中弱分类器的数目。

对于每一层次的检测器都有一个检测率 d_i 和误检率 f_i , 那么整个 K 层的检测器的总检测率和误检率为:

$$D = \prod_{i=1}^K d_i \quad F = \prod_{i=1}^K f_i \quad (3.8)$$

3.3 链式 Boosting 训练算法的优化

在链式 Boosting 的每一步中, 其性能是精度和速度的折中, 即使用较多的特征获得较高的检测率, 但是需要花费更多的评估时间。AdaBoost 提出通过简单调整每个分类器的阈值来试图取得检测率和虚警率的折中, 但是该方法却导致虚警率的急剧增加。此外, 由于每个分类器的计算时间与 f_i 和 m_i 有关, 对于 f_i 和 m_i 的优化就显得尤为重要。为了解决上述问题, 提出如下优化方法。

3.3.1 对于 α 的优化

使用以下的简略符号: $T = m_i, h_j(\mathbf{x}), \alpha_j = \alpha_{i,j}, b = b_i$ 和 $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_T\}$ 。那

么 $\Phi_i = \text{sign}[\sum_{j=1}^{m_i} \alpha_{i,j} h_{i,j}(\mathbf{x}) - b_i]$ 可以看成是弱分类器 $\{h_1(\mathbf{x}), h_2(\mathbf{x}), \dots, h_T(\mathbf{x})\}$ 的线性组合。当

$h_i(\mathbf{x})$ 确定之后, 样本 \mathbf{x}_i 将从原来的特征空间 \mathbf{F} 映射到新的空间 \mathbf{F}^* (维数是 T) 的一点, 即 $\mathbf{x}_i^* = h(\mathbf{x}_i) = \{h_1(\mathbf{x}), \dots, h_T(\mathbf{x})\}$ 。对于 α 参数的优化问题变成了在新的空间中寻找最优分割平面的问题。

根据^[37] 通过求解最大化二次形式:

$$L(\beta) = \sum_{i=1}^n \beta_i - \frac{1}{2} \sum_{i,j=1}^n \beta_i \beta_j y_i y_j (h(\mathbf{x}_i) \cdot h(\mathbf{x}_j))$$

$$\text{约束条件是: } \sum_{i=1}^n \beta_i y_i = 0 \text{ 和 } C_i \geq \beta_i \geq 0, i=1, \dots, n \quad (3.9)$$

其中 C_i 是跟据常数 C , 风险参数 ω 和训练集来设置。该方法的解为:

$$\beta^0 = (\beta_1^0, \beta_2^0, \dots, \beta_n^0) \quad (3.10)$$

从而得到优化的 $\alpha_i = \sum_{j=1}^T \beta_j y_j h_j(\mathbf{x}_i)$ 。通过调整 b 和 ω 可以获得最优解。

3.3.2 去除 Boosting 算法中的冗余信息

由于 AdaBoost 使用贪婪选择策略来进行前向序列搜索, 所以不可避免地引入了冗余。而 FloatBoost 采用回溯策略, 在增加新的弱分类器的同时剔除不好的弱分类器。虽然 FloatBoost 在一定程度上降低了冗余, 但是学习过程不稳定。

这里使用以下的算法来消除一定的冗余, 其中将前述的分类器表达成:

$$f(x) = \sum_{i=1}^I \alpha_i h_i(x) + b.$$

1. 在 $\{h_i(x)\}$ (其中 $i=1, \dots, M$) 集合上训练一个线性 SVM 分类器, 得到权值 w 。
2. 将 w (即 α) 依序从小到大排列, 假定序号为: i_1, i_2, \dots, i_M
3. $k=1, \dots, N$, 其中 N 是特征消除常量
 - a) 剔除特征 h_{i_k}
 - b) 计算现在的学习精度 p_k
 - c) 将 h_{i_k} 放回特征集中
4. 将对应 p_k 最大的那个 h_{i_k} 剔除
5. $M=M-1$, 转 1

3.4 链式 Boosting 算法在人脸检测中的应用

该文采用 KICA 构造重建信噪比阈值函数, 并且在此基础上, 使用层叠式 Boosting 算法构造一个基于该信噪比阈值的检测函数序列, 然后将它们组合成一个总检测函数, 形成一个 Boosting 分类器, 再使用相同的方法构造一个 Boosting 分类器 (结点), 并将它们串起来形成链式 Boosting 结构, 据此可以判别一幅图像是否为人脸图像。

3.4.1 构造重建图像信噪比阈值函数

在执行该算法之前, 先对人脸图像进行一系列的预处理, 以达到位置校准和灰度归一的目的。经过预处理, 得到了所谓的“标准图像”, 以归一化的标准图像作为训练样本集。预处理后, 一幅 $l \times h$ 大小的人脸图像将其按列相连构成一个 $M=l \times h$ 维的列向量。将得到的列向量作为该算法输入, 即 y_1, \dots, y_N , 代表 N 幅图像, 采用高斯核, 经过该算法后输出为 W , 即滤波器 W , 独立的基图像 U

可以通过滤波器组 W 而获得, 即: $W * X = U$ 。

则人脸向量 P 在由 w_1, w_2, \dots, w_N 张成的子空间中的投影 $U = W^T P$,
 $W = [w_1, w_2, \dots, w_N]$ 。重构时, 取前 K 个独立分量进行重构, 即 $P_i = WU \approx \sum_{i=1}^K w_i U_i$,
 这样就可以把人脸图像向量从原来的 N 维降到 K 维。定义重建图像的信噪比:

$$R_{SN}(P) = 10 * \log\left(\frac{\|P\|^2}{\|P - P_i\|^2}\right) \quad (3.11)$$

若其小于阈值 θ , 则可判断 P 不是人脸图像。这样就得到了一个个用于人脸检测的阈值检测函数:

$$\text{sign}(R_{SN}(P) - \theta) \quad (3.12)$$

3.4.2 用链式 Boosting 方法进行人脸检测

Boosting 算法是建立在“寻找粗糙的弱分类器比寻找很强的分类器要容易很多”这一发现的基础上的。它把许多弱分类器融合在一起, 从而获得一个分类能力很强的判别准则。Boosting 方法的主要思想是给定一弱学习算法和一训练集 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ 这里 x_i 为第 i 个训练样本的输入, y_i 为分类问题的类别标志。算法开始先给每一个训练样本赋以相等的权值 $1/n$, 然后用该学习算法对训练集训练 T 轮, 每次训练后, 对训练失败的训练样本赋以较大的权值, 也就是让学习算法在后续的学习中集中对比较难的训练样本进行学习, 从而得到一个预测函数序列 h_1, h_2, \dots, h_r , 其中 h_j 也有一定的权值, 预测效果好的预测函数的权值较大, 反之较小。最终的预测函数 H 采用有权值的投票方式对新样本进行判别。

链式 Boosting 方法的主要思想就是将一个个 Boosting 分类器连成一个链形结构, 后一个 Boosting 分离器利用前一个 Boosting 分类器所产生的历史信息。

以上已经依据重建图像的信噪比 $R_{SN}(x)$ 构造了阈值函数 $\text{sign}(R_{SN}(x) - \theta)$ 。现在将该阈值函数当成弱学习算法: $h_i \in \{t(R_{SN}(x) | \theta) = \text{sign}(R_{SN}(x) - \theta)\}$ 。其中 x 是作为输入数据的样本图像, $R_{SN}(x)$ 是得到人脸图像 x 对应的重建图像信噪比, θ 为检测阈值。

Boosting 算法如下:

- (1) 给定 $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$; $y_i \in \{+1, -1\}$

初始赋予每个样本相等的权重 $1/N$; $D_1(i) = 1/N$ 。

- (2) For $t=1, 2, \dots, T$ Do

- 在 θ 的所有可能的取值范围 $[\theta_1, \theta_2]$ 内选取一个能使样本集的检测错误率最小的 θ^* , 构造 $h_t(x) = \text{sign}(R_{SN}(x) - \theta^*)$, 使得:
 $(\theta^*) = \underset{\theta}{\operatorname{argmin}} \sum_i D_t(i) [h_t(x_i) \neq y_i]$ 。
- 计算该预测函数的错误率 E_t (所有被错误分类的样本的权重和):

$$E_t = \sum_i D_t[i, h_t(x_i) \neq y_i] \quad (3.13)$$

得到:
$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - E_t}{E_t} \right) \quad (3.14)$$

- 根据错误率更新样本的权重:

$$D_{t+1}(i) = D_t(i) * \exp(-\alpha_t y_i h_t(x_i)) / Z_t \quad (3.15)$$

其中 Z_t 为使得 $\sum_{i=1}^N D_{t+1}(i) = 1$ 的规范化因子。

则基于重建图像信噪比的链式 Boosting 算法为:

- (1) 初始化: $i=0$, $F_0=1$, $\phi=\{\}$, 对于所有的正样例 x_j ; $w_j = 1/p$; 对于所有的负样例 x_j ; $w_j = 1/n_i$ 。
- (2) 当 $F_i > F$ 时, 做
- 1) $i=i+1$
 - 2) 用初始权重 w_j , 在正样本集 P 和负样本集 N_i 上训练 ϕ_i , 以使 f_i, d_i 达到要求的值 (见上面 Boosting 算法)
 - 3) 优化结点训练器 (见 3.3)
 - 4) $F_i = F_{i-1} * f_i$, $\phi = \phi \cup \{\phi_i\}$
 - 5) 在非人脸图像集上评估 ϕ , 将错误的检测放入负样例集 N_{i+1}

6) 对于每一个在 N_{i+1} 中的样例 x_j ，按权值修改公式

$$w_j^{i+1} \leftarrow w_j^0 \exp[-y_j \sum_{k=1}^i \sum_{l=1}^{m_k} \alpha_{k,l} h_{k,l}(x)] \text{ 修改权值。}$$

OpenCV 组织提供了算法训练的特征分类器参数，应用这些参数可以很好的进行人脸检测。

3.5 实验结果与分析

本文取1000幅包含人脸的图像和1000幅不包含人脸的图像作为训练样本，这些图像来自于AR, Rockfeller, FERET, BioID和Internet网。训练集中的人脸有 $[-30^\circ, 30^\circ]$ 的平面内和深度旋转变换。将从1000幅包含人脸的图像中切割出人脸，进行尺度归一化为 20×20 大小，然后分别进行镜像变换、四个方向上1个像素的位移、在15度以内的平面内旋转以及在20%范围内的尺度缩放，从而产生共8000幅包含人脸的图像。

测试集一部分来自于标准的MIT+CMU人脸库。该库包含了125幅灰度图像，其中含有483个正面人脸；一部分来自于Internet网上的图片。

在MIT+CMU测试集上虚警数和检测率的关系见表3.1：

表3.1 在MIT+CMU测试集上虚警数和检测率的关系

检 测 率 (%) 虚 警 数	方 法	AdaBoost	FloatBoost	本文的方法
10		78.0	78.1	79.5
50		88.2	88.0	90.8
100		90.3	90.1	91.9
150		91.7	93.1	95.2

由表3.1可以看出虚警数和检测率是一对矛盾：当虚警数大时，检测率较高；而当虚警数小时，检测率较低。还可以看出我们的方法在相同虚警数下检测率要优于Adaboost和Floatboost。

使用链式Boosting算法分别在MIT+CMU和数字照片测试集上进行人脸检测的结果如图3.2和图3.3所示。

3.6 本章小结

由Viola提出的AdaBoost是序列式前向搜索过程，它使用贪婪的选择策略。它的前提条件是单调性。而这个前提可能无法满足。而由Stan Li提出的

FloatBoost算法将Floating搜索和AdaBoost融会贯通。由FloatBoost训练出的检测器具有更好的检测性能和较少的特征，然而FloatBoost算法不稳定，并且当

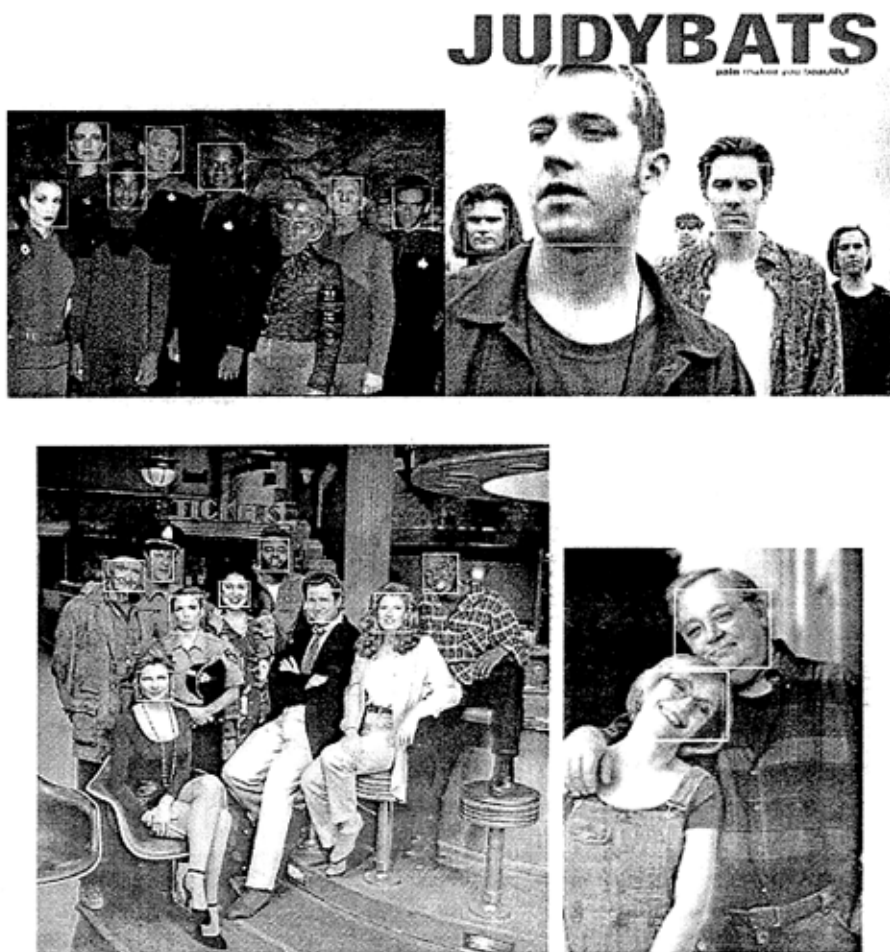




图3.2 在MIT-CMU测试集上使用链式Boosting的部分实验结果图



图3.3 在数字化照片上使用链式Boosting的部分实验结果图

处理比较复杂的问题时比较费时。而链式 Boosting 在传统的 AdaBoost 训练算法的过程中使用历史信息。并且采用线性递归特征消除方法来消除 AdaBoost 中的冗余信息，从而避免了局部最优。除此之外，采用线性递归特征消除方法还获得了最优阈值，避免了 AdaBoost 采用的阈值确定方法所带来的一些问题。

本章首先介绍了链式 Boosting 的结构、训练算法以及对算法的一些优化。

然后将链式 Boosting 运用到人脸检测中。在该检测方法中,先利用核独立分量构造重建图像信噪比阈值函数,再使用 Boosting 算法构造一个基于信噪比阈值的检测函数序列,然后将它们组合成一个总检测函数,形成一个 Boosting 分类器,再使用链式 Boosting 算法用相同的方法构造一个个 Boosting 分类器(结点),并将它们串起来形成链式 Boosting 结构,据此可以判别一幅图像是否为人脸图像。实验表明该方法无论在检测率还是在虚警率方面都要优于传统的 AdaBoost 和 FloatBoost。当然,无论是链式 Boosting,还是传统的 FloatBoost 和 AdaBoost,都需要长时间的训练过程。

第4章 基于模板匹配的人脸检测算法

基于模板匹配的方法是属于人脸检测方法两大类中的基于启发式模型的方法。

这里提出的基于模板匹配的方法如图4.1所示,在“可能肤色区域提取”阶段,这里使用 YCbCr 肤色模型,它是一种比较合理并且省时有有效的肤色提取方案。在“形态学处理”首先检测二值化后的各个图像分块,对分块大小进行阈值过滤,大量非人脸分块可以过滤掉。之后对待检测原图像进行边缘提取,将提取结果和“形态学处理”后的二值化反转图像进行“与”运算,得到保留了人脸细部特征的二值化图像。将可能肤色分块分析放到“第一次模板匹配”时进行(检测各分块区域的连通区域数)。“第一次模板匹配”阶段提出了一种基于边缘检测后的平均脸模板匹配的人脸检测方法。在“第二次模板匹配”阶段,在使用上面方法进行检测的结果的基础上提出了使用基于边缘的半个平均脸模板再来进行人脸检测,以保证重叠和遮挡人脸不会被漏检。在“第三次模板匹配”阶段,使用人耳模板,在第二次模板匹配检测结果的基础上再一次进行检测来防止对于旋转角度过大的人脸的漏检。在规则验证阶段,先使用头发模型,把没有头发相邻的候选人脸区域排除掉,然后使用改进了的三分图验证规则进行再一次地验证。

4.1 可能肤色区域的提取

这里仍旧使用 YCbCr 肤色模型来进行肤色提取。

也有一些人脸检测方法采用计算相似度的方法来提取肤色区域。通过采用大量的肤色标本来训练系统,得到一个肤色分布的中心,之后考虑要考察的像素距离该中心的远近得到一个相似度,一个像素一个相似度,最终得到一个待检测图像的相似度分布图,再按照一定规则和阈值对该分布图二值化,将肤色区域置白,背景区域置黑。

肤色分布可以通过高斯模型 $N(m, C)$ 表示,其中 m 是均值, C 是协方差矩阵。

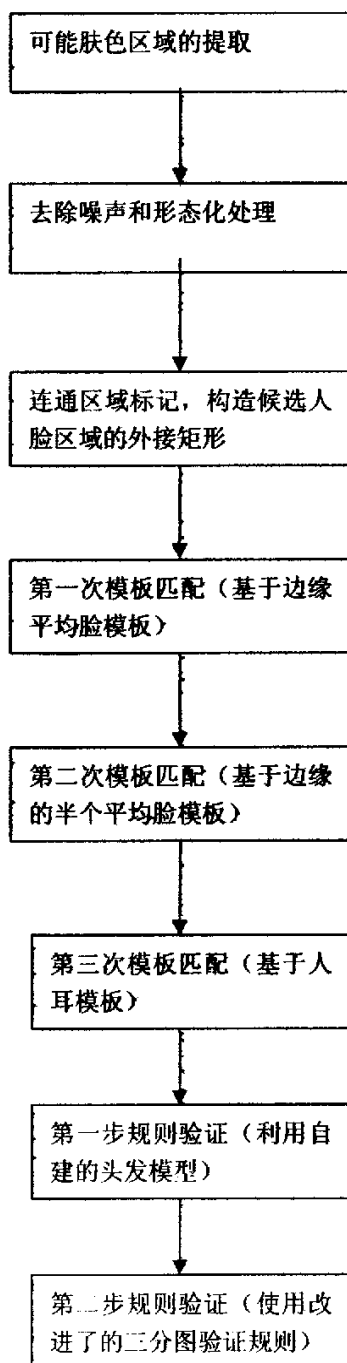


图 4.1 算法流程图

$$\mathbf{m} = E\{\mathbf{x}\}, \quad \mathbf{x} = (Cr, Cb)^T \quad (4.1)$$

$$\mathbf{C} = E\{(\mathbf{x} - \mathbf{m})(\mathbf{x} - \mathbf{m})^T\} \quad (4.2)$$

通过这个肤色高斯分布，可以得到某幅彩色图像中任意一个像素点属于皮肤的概率。对于某像素点 s ，从RGB空间转换到YCbCr彩色空间，得到色度值 (Cr, Cb) ，那么这个像素的肤色概率密度可以通过下面的式子计算得到：

$$P(Cr, Cb) = \exp[-0.5(\mathbf{x} - \mathbf{m})^T \mathbf{C}^{-1}(\mathbf{x} - \mathbf{m})], \quad \mathbf{x} = (Cr, Cb)^T \quad (4.3)$$

但是此方法数学计算比较复杂，花费时间较多，实时性很差。经过对大量人脸和肤色的统计结果进行分析，使用YCbCr肤色模型来提取肤色效果好且满足了实时性的要求（见图4.3）。



图 4.2 处理前的图片



图 4.3 使用 YCb'Cr' 彩色空间提取肤色

4.2 去除噪声

在对肤色区域进行二值化之后,会出现很多孤立的像素,这些像素就是噪声。噪声的出现有很多原因,比如在图像生成的时候由于设备的原因导致有的像素失真,从而产生了噪声。像这样的噪声一般都是少量的、孤立的一个或几个像素,所以,在进一步处理之前,先把这样的坏像素,即噪声,处理掉。

这里采用一种基于像素“密度”的滤波方法来去除噪声。这是一种类似于“卷积”的滤波方法,就是在以每一个肤色像素为中心的 5×5 邻域内统计肤色像素的个数,超过半数时中心点保留为肤色,否则认为是非肤色。

4.3 形态化处理

在对肤色区域进行了上述滤波之后,可以消除很多的噪声。虽然大量孤立的噪声点被消除,但是对于与肤色区域邻接的那些噪声点不会被上述滤波剔除,而这些噪声的存在影响后边的处理,所以使用形态化操作再次扫描图像。开运算可以去除孤立的小点、毛刺和小桥(即连通两块区域的小点)。这里连续使用两次开运算来消除毛刺和小桥。开运算是先膨胀后腐蚀。膨胀采用的是四方向判断的算法,就是说当前点的上下左右有一个点的颜色是白色时,就把当前点的颜色填充成白色。使用膨胀操作是为了连接脸部区域中不连续的块。随后的腐蚀操作采用的也是四方向判断算法,就是说检查一个点的上下左右4个点,如果当前点的颜色为白色,并且有一个点的颜色和当前点的颜色不同,那么就把当前点用别的颜色(黑色)填充掉。腐蚀是为了使非人脸区域变小,恢复到膨胀前的大小。

其次,要为人脸可能大小设定一个阈值。如果阈值设为一个定值,待检测图片换作其他图片,结果可能相差很大。如果阈值设得太高,有些人脸区域可能由于遮挡和原图片本身只有部分人脸等因素而漏判。经过试验发现,对于大部分图片,如果阈值设定为整个图片大小的 $1/100$,大部分单独琐碎的背景区域(\leq 阈值)经过开运算后可以去除。

然后,由于存在着人脸间相连或者人脸和某些背景区域相连的情况,可能会造成人脸的检测定位的失败,因此需要将这些大的图像块进行分离。最理想的状况是一个图像块就是一个单独人脸。将原图像转化为灰度图像。再经过均

值滤波器(3*3)，结果如图 4.4:



图 4.4 经过均值滤波器(3*3)处理后的结果图

滤除掉影响图像强度的一阶和二阶导数的噪声之后，采用 Canny 算子进行边缘检测，结果见图 4.5:



图 4.5 采用 Canny 算子进行边缘检测的结果图

之后将前面处理后的二值图像和边缘检测所得反转图像进行相“与”运算，结果见图 4.8:



图 4.6 前面处理后的二值图像



图 4.7 边缘检测所得反转图像



图 4.8 相“与”的结果图

这样将人脸的细部特征保留下来，并且将大的图像块进行一定程度上的分割，同时尽可能的去除掉背景的影响，为以后的模板匹配打下良好的基础。

4.4 连通区域标记，构造候选人脸区域的外接矩形

经过前边的步骤后，整个图像被分割为一个个连通的肤色区域。在一般情况下，单个人脸肤色区域是一个连通的肤色区域。当然在极少数情况下由于光照角度等等因素，导致单个的人脸区域也有可能被分割到多个肤色区域，但是这种情况在自然情况下很少发生。所以，在通常情况下，在这些肤色连通区域中有的含有人脸，有的没有人脸。我们使用种子填充法将肤色连通区域进行标记，这样就可以得到肤色连通区域的外接矩形。

经过了种子填充法之后，肤色连通区域都被做了标记，并且每一个连通区域（即为候选人脸区域）都被一个外接矩形围住。这样以后的工作主要集中于对这些矩形区域进行处理就可以了。

4.5 可能肤色分块分析

经过以上步骤，可以进行初步的人脸检测定位。

为了确定各个分块是否是人脸，需要进一步进行判断。通常的方法：首先，找出能包含各个分块的矩形区域，依次对各个矩形区域进行判断，流程如图4.9所示。

其次，计算各个分块的欧拉数，反映出各个分块各有多少个洞。因为人脸的眼睛、鼻子和嘴唇经过以上步骤后会以黑色的洞显示出来，因此通过所计算的欧拉数，设定一个阈值（通常设为 ≥ 0 ，因为眼睛至少占两个洞），当分块欧拉数 ≥ 0 时，将该分块视为非人脸区域；反之，该分块进入下一轮的人脸区域候选。

但是考虑到如果分别计算各个分块的宽高，之后进行条件比较并不能有效解决人脸重叠和遮挡的问题（因为人脸遮挡之后，其宽和高有比较大的变化，可能因为不能满足以上条件而被漏判）。同时欧拉数的判断也不够准确，因为非人脸皮肤或者背景分块也有可能满足欧拉数阈值条件而被错判为人脸。所以这里选择比较简洁的方法，就是矩形区域都取为正方形，边长等于分块区域的宽高两者之间的较大值。对分块的判断放到后面步骤进行。

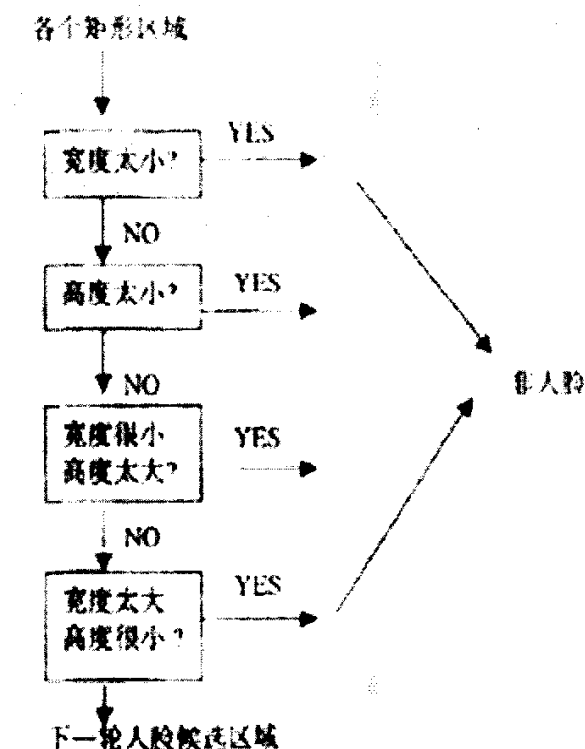


图 4.9 可能肤色分块分析的流程图

4.6 基于模板匹配的人脸检测方法

概括地说，基于模板匹配的方法是在图像灰度的层次上直接比较目标模板和候选图像区域之间的相似性。对于人脸检测而言，就是利用人脸的灰度模板在被检测图像中寻找最近似人脸的区域。与基于特征的方法等其他的人脸检测方法相比，模板匹配方法具有简单、直观的特点，适应性强，对图像质量的依赖度较低，鲁棒性强。本章主要讨论了使用模板匹配在复杂背景彩色图像中多个多姿态人脸的检测问题。

4.6.1 模板匹配的基本原理

在机器识别事物的过程中，常需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校准, 或根据已知模式到另一幅图中寻找相应的模式, 这就叫匹配。如在遥感图像处理中, 把不同波段传感器对同一景物拍的多光谱图像按像点对应套准, 然后根据像点的性质进行地物分类。如果利用在不同时间对同一地面拍摄得到两幅图片, 经套准后找出其中特征有了变化的点, 就可以用来分析图中那些部分发生了变化。在描述三维动态景物、计算物体的空间结构和运动参量等方面匹配技术都发挥着重要的作用, 因此匹配的方法受到广泛重视。

早期的图像匹配技术主要用于几何校正后的多波段遥感图像的套准, 借助于求互相关函数的极值来实现。在三维景物分析中, 由于三维成像中有透视失真、运动遮挡及阴影混入和噪声干扰等不利因素, 三维图像匹配至今仍是公认的技术难题。但在二维图像处理领域, 匹配技术已经成为一个重要技术手段, 众多研究者在这一基础上已发展出大量相关理论和方法。

模板匹配方法简单说就是计算一个固定的目标模式和一个未知图像块的相异程度, 用一个阈值决定输出。模板匹配可以用来确认一幅图中是否存在某种已知模板图像。

设有一模板 T (大小为 $M \times M$) 放在搜索图 S ($N \times N$) 上平移, 模板覆盖下的那块搜索定义为子图 $S^{i,j}$, i, j 为这块子图的左上角像点在 S 图中的坐标 (参考点), 则 i, j 的取值范围为:

$$1 < i, j < N - M + 1 \quad (4.4)$$

比较 T 和 $S^{i,j}$ 的内容, 若两者一致, 则 T 和 $S^{i,j}$ 之差为零, 所以可以用下列两种测度之一来衡量 T 和 $S^{i,j}$ 的相似程度:

$$D(i, j) = \sum_{m=1}^M \sum_{n=1}^M [S^{i,j}(m, n) - T(m, n)]^2 \quad (4.5)$$

$$D(i, j) = \sum_{m=1}^M \sum_{n=1}^M |S^{i,j}(m, n) - T(m, n)| \quad (4.6)$$

如果展开 (4.5), 则有

$$D(i, j) = \sum_{m=1}^M \sum_{n=1}^M [S^{i,j}(m, n)]^2 - 2 \sum_{m=1}^M \sum_{n=1}^M S^{i,j}(m, n) \times T(m, n) + \sum_{m=1}^M \sum_{n=1}^M [T(m, n)]^2 \quad (4.7)$$

式(4.7)右边第一项是模板覆盖下那块子图的能量, 它随 (i, j) 的位置缓慢改变, 第三项表示模板的总能量, 是一个与 (i, j) 无关的常数, 而第二项是子图图像和模板的互相关, 随 (i, j) 而改变。当 T 和 $S^{i,j}$ 匹配时这一项的取值最大, 因此可以用下列相关函数作为相似性测度:

$$R(i, j) = \frac{\sum_{m=1}^M \sum_{n=1}^M S^{i,j}(m, n) \times T(m, n)}{\sum_{m=1}^M \sum_{n=1}^M [S^{i,j}(m, n)]^2} \quad (4.8)$$

或者归一化为:

$$R(i, j) = \frac{\sum_{m=1}^M \sum_{n=1}^M S^{i,j}(m, n) \times T(m, n)}{\sqrt{\sum_{m=1}^M \sum_{n=1}^M [S^{i,j}(m, n)]^2} \sqrt{\sum_{m=1}^M \sum_{n=1}^M [T(m, n)]^2}} \quad (4.9)$$

其中 $0 < R(i, j) < 1$, 仅在比值 $\frac{S^{i,j}(m, n)}{T(m, n)}$ 为常数时 $R(i, j)$ 取得极大值1。利用相关性进行模板匹配的计算量很大, 模板需要在 $N - M + 1$ 个位置做相关计算。

4. 6. 2 基于平均脸模板匹配的人脸检测

使用模板匹配的方法进行人脸检测, 首先需要将待检测的彩色图像转换为灰度图。对图像中每个像素点可做如下变换:

$$Gray(i, j) = 0.11 * R(i, j) + 0.59 * G(i, j) + 0.3 * B(i, j) \quad (4.10)$$

式中 $Gray(i, j)$ 为变换后灰度图像中像素点灰度值, $R(i, j)$ 、 $G(i, j)$ 、 $B(i, j)$ 分别是原图中像素的R、G、B值。

(1) 人脸检测所需模板的制作

一般模板匹配的方法采用平均脸模板(即127幅彩色人脸的平均)。如果平均脸所用人脸图片很少的话, 结果可能不理想。同时可以看到平均脸的计算也

要占用一些时间，且人脸的细部特征不够突出。也有采用包括脖子部分的平均脸作为模板的方法，相对于一般的平均脸模板效果更好^[40]，但同样需要大量人脸图片和花费较多时间。

为了节约时间和减少需要的人脸图片，我们可以采用下面的方法去取得人脸模板：将 127 幅人脸图片全部转化为灰度图。制作人脸模板需要对含有人脸样本的图像按照一定的比例进行剪裁，取出能体现人脸模式的局部人脸图像。剪裁步骤如下：首先在人脸图像中人工定点，标定左、右眼与左、右嘴角的位置；如果双眼连线与水平存在夹角，则利用旋转变换公式保持双眼水平。然后以经过双眼中点的脸部中轴线为对称轴，按照图 4.10 所示的剪裁比例进行剪裁：

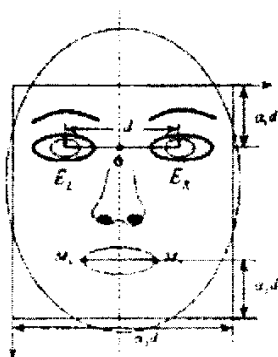


图 4.10 人脸样本裁剪比例示意图

其中以双眼距离 d 为基准，双眼向上的比例系数为 $\alpha_1 = 0.35$ ，嘴部向下的比例系数 $\alpha_2 = 0.2$ ，剪裁图像宽度的比例系数为 $\alpha_3 = 1.4$ 。接着利用尺度伸缩变换公式将尺度归一化为 60×60 。对图像进行灰度分布标准化，将图像灰度均值和方差调整到 128.0 和 64.0。之后对每幅图片进行边缘提取（采用 Canny 算子）以提取脸部的重要特征（见图 4.15）。



图 4.11 裁剪之前的原图



图 4.12 裁剪之后的图



图 4.13 尺度归一化后的图



图 4.14 灰度标准化之后的图



图 4.15 边缘提取之后的图

将边缘提取之后的人脸图片求均值。由于边缘提取之后的图片并不理想，人脸特征并不突出（见图 4.16）。



图 4.16 边缘提取之后的人脸图片求均值之后的结果图

为了使人脸特征突出，将上图减去平均灰度并且进行图像增强。由此得到的图 4.17 中眼、鼻、嘴、耳等特征都比较突出，因此将之作为第一次匹配的模板是合适的。



图 4.17 减去平均灰度并且进行图像增强之后的结果图

同时考虑到分块矩形区域有大有小，模板匹配时，需要缩放模板，这样要占用一定的时间。为了节省时间，我们将模板存储为6个不同的尺度：

设原模板大小为 $P = 60 \times 60$ 像素，

6个模板大小依次设定为 $1.4P, 1.2P, P, 0.8P, 0.6P, 0.4P$ 。

可以看到该模板有以下几个优点：

1) 模板大部分都是黑色，白色部分表明脸部边缘和五官轮廓，这样可以节省模板存储空间并且可以节省模板匹配的时间。

2) 由于大部分是黑色，可以看到除非人脸候选分块和模板有比较精确的匹配，否则相关性都很低，便于除去非人脸分块。

3) 下图是一幅人脸的边缘提取后的图像和边缘提取之后的手臂的图像。由下图可以看出，图4.18比图4.19复杂。因为两幅图像的原始彩色图像大小差不多并且颜色差不多。如果用一般的模板匹配的方法计算相关性，它们和模板的相关性可能差不多，这样就有错判或者误判的可能。但是由于我们现在提取了边缘，可以看出我们只要简单计算两图各有几个连通区域，可以得出图4.18连通区域有37个，图4.19连通区域只有10个，因为人脸比手臂，腿，衣服等大部分的非人脸区域复杂，因此可以轻易地将候选人脸区域中的这些非人脸分块过滤掉。



图 4.18 人脸的边缘提取后的图像



图 4.19 边缘提取之后的手臂的图像

用上述的人脸边缘平均模板匹配，选用的判定条件为：如果相关系数 $\text{corr} \geq 0.2$ 并且连通区域 ≥ 20 为人脸可能区域；否则为非人脸区域。

(2) 模板匹配准则

利用模板匹配方式，就不免要谈及到模板匹配的准则。常用的度量匹配程度的准则主要有欧式距离和相关性系数，其中、相关性系数也可以看作是角度

距离。

设尺度为 $W \times H$ 的人脸模板灰度矩阵为 $T = [t_{i,j}]$ ，灰度均值为 μ_T ，方差为 σ_T ，与模板同样尺度大小的检测窗口内的图像矩阵为 $X = [x_{i,j}]$ ，所对应的均值为 μ_X ，方差为 σ_X 。从而模板与输入图像之间的欧式距离 $d(T, X)$ 和相关系数 $r(T, X)$ 为：

$$d(T, X) = \sqrt{\frac{\sum_{i=0}^{H-1} \sum_{j=0}^{W-1} (t_{i,j} - x_{i,j})^2}{W \cdot H}} \quad (4.11)$$

$$r(T, X) = \frac{\sum_{i=0}^{H-1} \sum_{j=0}^{W-1} (t_{i,j} - \mu_T)(x_{i,j} - \mu_X)}{W \cdot H \cdot \sigma_T \cdot \sigma_X} \quad (4.12)$$

其中相关系数 $r(T, X)$ 越大表明模板与检测图像的匹配程度就提高了；而欧式距离 $d(T, X)$ 恰好相反。

文献^[10]提出利用 $r(T, X)$ 和 $d(T, X)$ 综合起来作为匹配程度的度量，即：

$$q(T, X) = r(T, X) + \frac{\alpha}{1 + d(T, X)} \quad (4.13)$$

其中 α 为权重系数，可以根据实验数据，使前后两部分的量值相当而取定的平均当量。

我们在实验中发现，仅使用相关系数 $r(T, X)$ 就已经表现出了良好的检测效果，而计算 $q(T, X)$ 时多计算了 $d(T, X)$ ，这无疑增加了检测时间的开销，并且检测结果相差不大。故在我们所实现的模板匹配算法中只使用了相关系数作为匹配程度的度量。

(3) 多尺度搜索

对于进行模板匹配时，由于只知道图像中人脸的大致尺寸，而不知道人脸的确切尺寸，则需要对不同分辨率的图像进行检测。一种常用的方法是按一定比例对图像进行多尺度缩放，对每一尺度图像分别用 6 个模板进行检测。

设在一幅待检图像中, 最大可能人脸的尺度为 W_{\max} , 最小可能人脸的尺度为 W_{\min} , 如果模板的尺度为 W_T , 并设比例缩放因子为 $1/\alpha$, 尺度依次缩放次数为 k , 则应满足:

$$\alpha^k = W_{\max} / W_{\min} \quad (4.14)$$

从而:

$$k = [(\log W_{\max} - \log W_{\min}) / \log \alpha] + 1 \quad (4.15)$$

其中 $[\cdot]$ 表示取整。

设第 0 次的图像为原始图像, 则第 i 次缩放 (针对第 $i-1$ 次) 的比例系数为:

当 $i=1$ 时, $\alpha_i = \beta$, 其中 $\beta = W_T / W_{\min}$;

$$\text{当 } 1 < i \leq k, \alpha_i = 1/\alpha。 \quad (4.16)$$

(4) 变形模板和弹性模板的引入

在实际图像中的人脸不可避免的存在一些少许的平面内倾斜, 如果能得到候选区域的倾斜角 θ , 那么把模板按照这个角度进行旋转后再进行检测, 将有利于提高模板匹配的成功率。

Henry Chang 提出通过分析肤色区域的一些基本特征 (质心、倾角、高宽比等) 来优化区域, 他将区域的倾斜角通过如下的公式计算得到:

$$\theta = \frac{1}{2 \cdot \text{atan}(\frac{b}{a-c})} \quad (4.17)$$

其中, $a = \sum_{i=1}^n \sum_{j=1}^m (\dot{x}_{ij})^2 B[i, j]$, $b = 2 \sum_{i=1}^n \sum_{j=1}^m \dot{x}_{ij} \dot{y}_{ij} B[i, j]$

$$c = \sum_{i=1}^n \sum_{j=1}^m (\dot{y}_{ij})^2 B[i, j]$$

$$\dot{x}_{ij} = x_{ij} - \bar{x}, \quad \dot{y}_{ij} = y_{ij} - \bar{y}$$

$B[i, j]$ 表示 $m \times n$ 维候选人脸区域, (x_{ij}, y_{ij}) 表示图像上像素点的坐标。 (\bar{x}, \bar{y})

为区域的质心,

$$\bar{x} = \frac{1}{A} \sum_{i=1}^n \sum_{j=1}^m j B[i, j], \quad \bar{y} = \frac{1}{A} \sum_{i=1}^n \sum_{j=1}^m i B[i, j] \quad (4.18)$$

将旋转 θ 角度后的模板进行匹配检测,有利于提高检测率。实际上弹性和变形模板方法的应用已经广受重视,是模式识别领域的重要研究方向^{[42][43][44][45]}。

(5) 半个人脸模板的引入

上面提到的基于边缘平均脸(完整人脸)模板匹配来检测人脸的方法由于没有考虑到人脸重叠和遮挡问题,因此在人脸重叠和遮挡时效果不是很好,可能会出现漏判,为此我们可以采取取半个人脸模板的方法进行进一步的检测。所用模板如图4.20所示:



图4.20 半人脸模板,从左到右依次编号为1、2、3、4

将原来边缘检测之后所取得的模板一分为二,分别进行匹配。当然,花费时间也是要考虑的因素,因此如果仅仅考虑人脸被遮挡下半部分的情况,就选用模板3;如果仅仅考虑左右遮挡的问题,那么就选用模板1或者2。如果计算机性能优良并且对人脸检测系统时间要求不高的话,几个模板可以同时使用,从而可以取得良好的效果。

其中人脸判定条件为:相关系数 $corr$ 阈值取0.6。如果相关系数 $corr \geq 0.6$,则为人脸可能区域;如果相关系数 $corr \leq 0.6$,则为非人脸区域。之所以阈值取0.6是因为大部分人脸已经在前面被检测出来,并且由于未被检测人脸可能不完整,连通区域不再成为能判定人脸的必要条件,因此阈值取得比较高。

图4.21可以看到没有改进前的算法,可能会出现漏检(人脸重叠和遮挡),但是误判几率很小(因为同时要满足边缘提取后模板的匹配和连通区域多少的判断);而改进后的算法,由于将已检测人脸区域置黑,而对其他可能人脸区域再一次进行检测,一来可以防止重复检测,大幅度减少了系统运行时间,二来如果改进前算法有漏检的话,改进后的算法则将这种漏检的可能进一步减小(图4.22)。

综合改进前算法和改进后的算法,除了系统时间有所延长外,人脸检测的成功率将有很大提高。



图 4.21 改进前的算法的结果图



图 4.22 改进后的算法的结果图

(6) 使用人耳模板

上面提到的基于边缘平均脸（完整人脸）模板匹配和基于边缘半个脸模板匹配来检测人脸的方法对正向人脸和旋转角度较小的人脸可以比较有效的检测定位，但是还存在一些问题需要改进，即对于深度旋转角度较大的人脸会出现漏判。这里提出利用人耳模板进行模板匹配来进一步减少漏检。由于国内外还没有专门的人耳图像库，所以必须自己构建人耳图像库。用数码相机或利用现有的标准人脸库收集 127 幅全侧人脸图片，将它们全部转化为灰度图（部分图片见 3.23）。



图 4.23 部分全侧人脸示意图

接着手工将人耳裁剪出来，如图 4.24 所示：



图 4.24 手工裁剪出的人耳示意图

然后进行尺度归一化（必须满足人耳区域高度与宽度的比值为 1.5，这里取 51×28 ），灰度标准化（将图像灰度均值和方差调整到 128.0 和 64.0），如图 4.25：



图 4.25 尺度归一化和灰度标准化之后的人耳示意图

最后将这些人耳图片取平均，从而得到平均人耳模板。



图 4.26 平均人耳模板示意图

可以看到该耳模板仍然具有细部特征不够突出的缺点。所以我们考虑使用边缘耳模板来代替。对每幅耳图像进行边缘提取（采用 Canny 算子）以提取耳部的重要特征：



图 4.27 每幅耳图像进行边缘提取后的结果图

将边缘提取之后的耳图片求均值：



图 4.28 平均边缘人耳模板示意图

为了使特征更加突出，将其减去平均灰度并且进行图像增强：



图 4.29 减去平均灰度并且进行图像增强后的平均边缘耳模板

由于侧面人脸与人耳面积的比值为 11.3, 根据剩下的各个可能肤色区域块的面积大小，对耳模板进行尺度缩放，然后进行匹配。其中各个可能肤色区域块的面积的求取使用矩特征：

图像 $f(x, y)$ 的 $p+q$ 阶原点矩用下式定义^[41]：

$$m_{pq} = \iint_{\zeta} x^p y^q f(x, y) dx dy \quad (p, q = 0, 1, 2, \dots) \quad (4.19)$$

其中， $f(x, y)$ 是对应输入图像坐标 (x, y) 处的灰度值， ζ 为图像亮度函数定义的像素空间区域。

图像的 $f(x, y)$ 的 $p+q$ 阶中心矩：

$$\mu_{pq} = \iint_{\zeta} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (p, q = 0, 1, 2, \dots) \quad (4.20)$$

其中 (\bar{x}, \bar{y}) 为灰度图像的灰度重心。

对于二值图像来说，零阶几何矩 m_{00} 就是目标区域的几何面积。

在以上检测结果的基础上将已检测人脸区域置黑，而对其他可能人脸区域用人耳模板再一次进行检测，逐像素匹配。经过多次实验确定判定条件为：相关系数 $corr$ 阈值取 0.4。如果相关系数 $corr \geq 0.4$ ，则为人脸可能区域；如果相关系数 $corr \leq 0.4$ ，则为非人脸区域。结果见图 4.30:



图 4.30 使用人耳模板改进后的算法的结果图

4.7 规则验证

由前面的介绍可知，模板匹配只是全局的匹配，而并不是局部的匹配，并不能再现人脸局部的特性，仅靠模板匹配得到的人脸并不是可靠的，还需要进一步的确认。模板匹配出的结果可能是“假脸”，如何去掉“假脸”，做进一步的人脸确认的方法有很多，如神经网络方法^[46]，SVM 方法^[47]等。

4.7.1 第一步验证-----选出头发区域来（不考虑秃顶）

在通常情况下，如果不戴帽子的时候，在人脸上方总是头发存在，通过找到头发的位置，我们就可以大致知道人脸比较准确的位置。对于彩色图像，头发的颜色模型：

$$Y = 0.30 * R + 0.59 * G + 0.11 * B \quad (4.21)$$

其中： R 、 G 、 B 是原始彩色图像像素的值。

通过大量实验定义判定函数如下:

$$Y < 0.3 \quad (4.22)$$

当 Y 的值在该范围内时, 该区域可能是头发。这里由于考虑到存在染发的情况, 因此在划定 Y 值的范围时, 规则定义比较宽松。

当满足以上函数时把这些像素点留下, 过滤掉其他的像素。结果见图 4.31 和图 4.33:



图 4.31 使用头发模型提取头发的结果图 (没有染发)



图 4.32 有染发的原图



图 4.33 使用头发模型提取头发的结果图（有染发）

然后，同上边处理肤色区域的方法一样，对其进行滤波和形态化。然后再求出其外接矩形并且进行灰度化，最后再求出每一个头发区域所在的重心。

求重心的公式如下：

$$\bar{x} = \frac{\sum_{x=m}^n x * g(x,y)}{n-m}, \quad \bar{y} = \frac{\sum_{y=m'}^{n'} y * g(x,y)}{n'-m'} \quad (4.23)$$

其中， $g(x,y)$ 是像素 (x,y) 所在的灰度值， m, n, m', n' 分别为所求区域的开始点的坐标和结束点的坐标。使用上面的公式，即可求出每一个头发区域的重心坐标。一般情况下，由于头发是比较对称的，所以其重心坐标也是位于中间位置，并且由于人脸区域是在头发的下边，所以，人脸也是位于重心的下面，或者头发重心落在人脸区域。人脸区域基本上以头发重心所在垂直线为对称轴。

找到了头发重心之后，我们可以进一步排除那些不含人脸的矩形区域，因为头发所在区域的宽度与实际宽度通常相差不大，它们的宽度之比通常不会超过 $[0.5, 2]$ 这个范围。另外，头发区域一般都是紧邻人脸区域，所以对于没有相邻的区域就可以断定不是人脸区域。

4. 7. 2 第二步验证----改进了的人脸三分图模型验证规则

在这里我们采用了文献^[48]和文献^[49]比较类似的验证方法，采用了人脸三分图模型的验证规则，为了适应不同姿态的变化，我们对验证规则进行了改进：

首先对待确认区域（模板匹配结果，已经标准化）使用垂直梯度算子和典型梯度算子获得垂直梯度图像与典型梯度图像，具体形式如下：

设待确认灰度图像 $[f(i, j)]$ （其中 i, j 分别为图像的行、列坐标），垂直梯度图像 $[V(i, j)]$ ，典型梯度图像 $[D(i, j)]$ ，则有：

$$V(i, j) = |f(i, j) - f(i+1, j)|. \quad (4.24)$$

$$D(i, j) = |f(i, j) - f(i+1, j)| + |f(i, j) - f(i, j+1)|. \quad (4.25)$$

然后对典型梯度图像 $[D(i, j)]$ 利用其均值进行二值化，得到二值图像 $[B(i, j)]$ 。与文献^{[52][53]}中的方法类似，将 $[V(i, j)]$ 和 $[B(i, j)]$ 划分成 3×3 个大小相同的块，统计每个子块的灰度值。若待确认区域是“人脸”，则应满足如下规则：

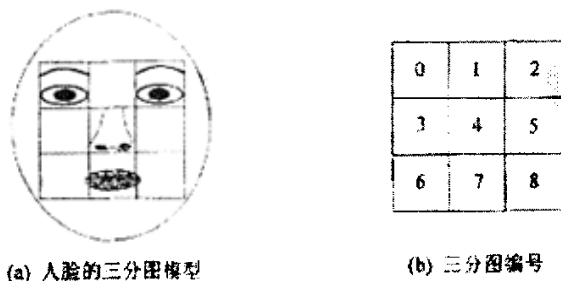


图 4.34 规则验证所采用的人脸三分图模型与编号

- (1) 梯度规则（ $\bar{V}[k]$ 表示垂直梯度图像 $[V(i, j)]$ 三分图中第 k 子块的平均灰度值）

$$\text{Rule1: } \bar{V}[0] > \bar{V}[1], \text{ 且 } \bar{V}[0] > \bar{V}[3], \text{ 且 } \bar{V}[0] > \bar{V}[6]$$

$$\text{Rule2: } \bar{V}[2] > \bar{V}[1], \text{ 且 } \bar{V}[2] > \bar{V}[5], \text{ 且 } \bar{V}[2] > \bar{V}[8]$$

$$\text{Rule3: } \bar{V}[6] < 1.5\bar{V}[7], \text{ 且 } \bar{V}[8] < 1.5\bar{V}[7]$$

$$\text{Rule4: } (\bar{V}[7] + 0.5\bar{V}[6] + 0.5\bar{V}[8]) > 1.5\bar{V}[3], \text{ 且 } (\bar{V}[7] + 0.5\bar{V}[6] + 0.5\bar{V}[8]) > 1.5\bar{V}[5]$$

- (2) 二值规则（ $\bar{B}[k]$ 表示图像 $[B(i, j)]$ 中 k 子块中灰度为1像素点所占的块比

例)

Rule1: $\bar{B}[0] > 0.3$, 且 $\bar{B}[0] > \bar{B}[1]$, 且 $\bar{B}[0] > \bar{B}[3]$

Rule2: $\bar{B}[2] > 0.3$, 且 $\bar{B}[2] > \bar{B}[1]$, 且 $\bar{B}[2] > \bar{B}[5]$

Rule3: $\bar{B}[4] > 0$

Rule4: $(\bar{B}[7] + 0.5\bar{B}[6] + 0.5\bar{B}[8]) > 1.5\bar{B}[3]$, 且 $(\bar{B}[7] + 0.5\bar{B}[6] + 0.5\bar{B}[8]) > 1.5\bar{B}[5]$

之所以采用梯度规则中规则 3、4 与二值规则的规则 4 这种形式, 是考虑到对非正面人脸适应性问题的。

4.8 实验结果和分析

采用本文提出的算法部分结果如下:



图 4.35 检测结果图 1



图 4.36 检测结果图 2



图 4.37 检测结果图 3



图 4.38 检测结果图 4



图 4.39 检测结果图 5

表 4.1 本算法的检测结果

待检测图片	存在人脸数	人脸正确定位数	错判和漏判数
1	3	3	0
2	5	5	0
3	5	3	3
4	5	3	2
5	8	7	1

表 4.2 本算法与一般基于肤色和模板匹配的算法的比较

	一般基于肤色和模板匹配的算法 ^[50]	本文算法
正确率(%)	82.7	92.2
平均检测一幅人脸图片所耗费的时间(s)	2.15	6.28

其中表 4.1 是对图 4.35 到图 4.39 这五张图片的测试结果的小结, 而表 4.2 是对自建的测试集的小结, 其中测试集中的图片一部分来在网上的照片, 一部分为私人照片, 共计 30 张, 均为复杂背景下多人脸多姿态的彩色图片, 图片大小归一化为 550*380。

分析试验结果的图、表和基于前面的算法讨论, 可以看出:

- 1) 本文提出的算法相对于一般基于肤色和模板匹配的算法, 检测的正确率有明显的提高。
- 2) 本文提出的算法没有出现将背景物体错判为人脸的情况, 说明结合边缘检测后的平均模板和连通区域的方法, 特别是连通区域阈值的设定, 可以有效去除背景的影响。
- 3) 如果人脸不是正向, 而是发生平面内旋转, 本文算法可以准确定位。
- 4) 如果人脸不是正向, 而是发生深度旋转, 本文算法也可以比较准确地定位人脸。

4.9 本章小结

这里针对复杂背景下彩色图像中人脸的检测问题提出了一种新的基于肤色模型和模板匹配的人脸检测算法。首先结合人体肤色的色彩特性和各彩色空间

的优劣,同时考虑到肤色提取时间的长短,提出了利用 $YCb'Cr'$ 彩色空间提取肤色的方法。在实际图像中的人脸不可避免的存在一些少许的平面内倾斜或旋转,这里提出了一种根据候选区域的基本特质来计算倾斜角,把模板按照这个角度进行旋转后再进行匹配检测的方法。在模板匹配中,提出了一种基于边缘检测的人脸平均模板,与保留了人脸细节的二值化图像进行模板匹配,结合相关系数阈值和连通区域阈值的过滤。对于遮挡和重叠的人脸区域,将基于边缘检测的人脸平均模板匹配之后已检测到的人脸区域置黑,用该模板的一半再次进行模板匹配。在平均脸模板匹配之后已检测到的人脸区域置黑上,提出了使用平均耳模板进一步进行匹配,从而在一定程度上克服了对深度旋转人脸的漏检。在规则验证部分,提出了在用三分图模型验证规则进行人脸验证之前,先通过大量的试验建立头发模型来进行第一步的验证,该头发模型不仅可以提取黑发而且可以提取染过之后的头发,适应性和鲁棒性都很强。

实验表明,该方法比传统的模板匹配方法在检测率和虚警率上有明显的优势,但是它要花费较长的检测时间。

第 5 章 三种方法之比较

为了对这三种方法进行比较，本文采用统一的测试集进行性能测试和比较。测试集中的所有图片均为彩色图片，并且均在复杂背景下，其中一部分来自互联网上的照片，一部分为作者的私人图片，一共 30 张图片，共 145 张人脸，图片大小归一化为 550×380 。部分图片见图 5.1:





图 5.1 测试集部分图片

先进行肤色提取及噪声滤波，然后分别使用 KICA/GA/SVM 的方法、链式 Boosting 方法和基于多个模板匹配的人脸检测方法来进行检测，结果见表 5.1。其中表中的“虚警数”表示将非人脸检测为人脸的数目。

表 5.1 三种人脸检测方法之性能比较

方法	平均检测一幅人脸图片所用的时间 (s)	检测率 (%)	虚警数
KICA/GA/SVM	38.4	90.3	24
链式 Boosting	0.1	94.1	49
多模板匹配	6.32	92.3	13

由表可以看出，由于 KICA/GA/SVM 的方法中引入了特征提取降低了特征空间的维数，从而使其检测速度要比传统的 SVM 方法来得快，但是由于 KICA 比较耗时，所以它比链式 Boosting 慢了将近 400 倍，比多模板匹配慢了 6 倍；与链式 Boosting 和多模板匹配相比，其检测率也是最低的。链式 Boosting 方法在传统的 AdaBoost 算法的过程中使用历史信息。并且采用线性递归特征消除方法来消除 AdaBoost 中的冗余信息，从而与 AdaBoost 方法相比只需要较少的弱分类器就可以获得相当的检测率，因而提高了检测速度，使其成为三种方法中速度最快的方法；但是我们也可以看到它在提高检测率的同时也在迅速增加虚警数，同时它还需要很长的训练时间。而这里的基于多个模板匹配的人脸检测方法由于引入了耳模板克服了旋转角度过大造成的漏检，也引入了半人脸模板克服了遮挡和重叠所造成的漏检，从而提高了检测率；然而，由于涉及多个模板多个尺度的搜索，因此它要比链式 Boosting 慢了将近 60 倍，但它使用了肤色信息来提取候选肤色区域，避免了全图搜索，因而比 KICA/GA/SVM 的方法快了将近 6 倍；并且该方法可以避免训练过程。所以从总体性能上来说，基于多个模板匹配的人脸检测方法最优，链式 Boosting 方法次之，而 KICA/GA/SVM 的方法最差。

第6章 结束语

6.1 总结

从全文的论述可以知道,人脸检测定位是解决人脸识别,人脸特征提取和其他人脸应用问题的关键技术,对它的研究是目前计算机视觉和计算机图像处理等研究领域的一个热点。此外,人脸检测定位的技术和方法可以很容易地拓展到物体检测定位和识别领域。本文系统研究了人脸检测的几种方法并作了一定的改进。现总结如下:

- 1) 在传统的 SVM 检测算法中,先采用核独立分量分析算法进行特征提取,并在算法中运用不完全 Cholesky 分解来降低时间复杂度,对求得的特征采用遗传算法再进行了一次特征选取以便有效地减少特征空间的维数,得到独立基的一个最优子集。实验表明,和传统的 SVM 方法相比,检测率有了明显的提高,并且由于特征空间维数降低了,检测速度又提高了不少。
- 2) 对链式 Boosting 方法进行了一些研究工作,即将链式 Boosting 运用到人脸检测中。在该检测方法中,先利用核独立分量构造重建图像信噪比阈值函数,再使用 Boosting 算法构造一个基于信噪比阈值的检测函数序列,然后将它们组合成一个总检测函数,形成一个 Boosting 分类器,再使用链式 Boosting 算法用相同的方法构造一个个 Boosting 分类器(结点),并将它们串起来形成链式 Boosting 结构,据此可以判别一幅图像是否为人脸图像。实验表明,链式 Boosting 方法检测率比 1) 方法要有所提高,检测速度也比 1) 方法快;但是虚警数随着检测率的提高也迅速增加,同时它需要长时间的训练时间。
- 3) 提出了一种新的基于肤色模型和模板匹配的人脸检测算法。在模板匹配中,提出了一种基于边缘检测的人脸平均模板。该模板相对于一般的灰度平均脸模板,人脸特征更加突出(五官都以白色突出显示)。对于遮挡和重叠的人脸区域,将基于边缘检测的人脸平均模板匹配之后已检测到的人脸区域置黑,用该模板的一半再次进行模板匹配。通过求得候选区域的倾斜角,把模板按照这个角度进行旋转后再进行检测,一定程度上解决了平面内旋转人脸的检测问题。在使用基于边缘的平均脸模板进行模板匹配后所得的人脸候选

区域中再进一步使用了平均耳模板进行模板匹配，一定程度上解决了深度旋转角度过大的人脸的检测问题。而在进行人脸验证时，先通过大量的实验建立头发模型（该模型可以检测黑发和染过的头发），剔除那些没有头发相邻的候选人脸区域，再使用改进了的三分图模型验证规则进一步对余下的人脸区域进行验证。实验表明，由于使用肤色信息进行人脸候选区域的提取，从而避免了对全图的多模板多尺度搜索，从而加快了检测速度；并且可以避免训练过程；从来说，本文提出的模板匹配方法在检测率、虚警数以及检测时间上达到了最好的折衷。

6.2 展望

模板匹配的方式没有考虑对“反例”（非人脸）的学习问题，适应性并不是很强，而且模板匹配是全局匹配，匹配形式比较粗糙，所得到的结果并不足以反映真正的“人脸”，必须要结合进一步细检（如规则验证）才能得到较满意的结果。然而由于人脸模式的复杂性，姿态各异，很难建立合适的规则来反映不同姿态的人脸，必然会带来一定的漏检和误检。由于进行了多重模板匹配和多尺度搜索，系统运行时间相对于一般的基于模板匹配和肤色的方法有所延长，选取更合适的模板来减少系统运行时间是未来亟待解决的问题，例如，现在女性很多都纹眉，采用女性眉毛的平均作为模板是未来可以尝试的工作。当然其它方面的配合必不可少，人脸的眼睛的检测定位对上述模板的成功与否起着至关重要的作用。

对于 KICA/GA/SVM 算法，可以考虑进一步改进 KICA 理论，从而进一步减少检测和训练所花费的时间。

对于基于多模板匹配的人脸检测方法而言，相对于其他方法其检测时间较长，这是该算法的不足之处，主要指由于多次模板匹配造成的。因此怎样减少模板匹配的时间，怎样选取更好的模板，将是下一步要做的工作。

这里实现的人脸检测算法都是基于静态图像的出发点提出的，并没有考虑对动态图像序列（视频流）的人脸检测问题（人脸跟踪），故进一步的研究重点应放在人脸跟踪问题上。考虑序列的帧间信息并结合基于静态图像的检测算法来完成对视频流的人脸检测问题。

目前存在很多的人脸检测方法。但是对于这些方法，人脸检测的好坏没有

一个公正的测试环境和评估算法。因此需要建立一系列标准来评估人脸检测算法的优劣。同时，由于现在的人脸检测方法都存在着缺点和不足，减少时间和提高检测准确率的矛盾依然存在，将多种检测方法相结合，是未来人脸检测方法的趋势和潮流。

同时，可以看到人脸检测和识别具有非常广阔的应用前景。例如，应用于下一代智能计算机的人机交互界面研发方面；应用于电子商务网络安全认证的推广方面；应用于银行自动安全监控和 ATM 自动提款机储户身份认证方面；应用于案件侦查时，检索犯罪嫌疑人人脸图像方面；应用于游戏设计的虚拟人物互动交流场景方面。人脸检测技术的研究，未来必将和其他新型学科和技术相互交叉，如人工智能、神经网络等，我们相信随着它的发展，必定能推动人脸研究的跨越式发展，进而推动整个多媒体信息技术的飞速发展。

致谢

在硕士期间取得的研究成果，首先要归功于导师武妍教授对我的关心和指导。武老师严谨的治学态度，对学术孜孜不倦的追求为我树立了一个科研工作者的榜样，将使我受益终身。从论文选题、资料收集整理、评估分析、到论文撰写的整个过程中，都得到了武老师的耐心指导。在此谨向武老师致以诚挚的感谢，感谢她对我们的严格要求。

其次要感谢王守觉院士。王院士作为一个科学家的操守、对新知识新技术的勇于探索，都令我非常的敬佩。王院士对于我的谆谆教导给了我很多学习和研究上的启发。

同时还要感谢实验室的朱君波老师，他在日常学习中给我不少有益的建议。感谢谷秧波、宋金晶、杨洋、施凝等的帮助，和他们平时的探讨给了我很多启示。实验室中互相帮助、交流、学习的和谐气氛，使我在实验室度过了难忘的时光。

在此，我向他们表示最衷心的感谢！

2006年3月

参考文献

- [1] Yang M.H., Kriegman D.J. & Ahuja N, Detecting faces in images: a survey. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002, Vol.24 (1):34—58
- [2] 梁路宏, 艾海舟, 徐光佑. 人脸检测研究综述. *计算机学报*, 2002, Vol.25 (5):449—458
- [3] Kin.C.Y. & Cipolla R, Feature-based human face detection. *Image and Vision Computing*, 1997, Vol.15 (9):713—715
- [4] Yang G. & Huang T.S, Human face detection in complex background. *Pattern Recognition*, 1994, Vol.27 (1):53—63
- [5] Morimoto C. & Flickner M, Real-time multiple face detection using active illumination. *Proceedings of the Fourth International Conference on Automatic Face and Gesture Recognition*, 2000: 26—30
- [6] Martin H. & Hunke H, Locating and tracking of human faces with neural networks. *Techn. Rep. CMU-CS*, 1994: 94—155
- [7] Tony S. & Pentland A, Parametrized structure from motion for 3D adaptive feedback tracking of faces. *IEEE International Conference on Computer Vision and Pattern Recognition*, 1997: 144—150
- [8] Dai Y. & Nakano Y, Face-texture model based on SGLD and its application in face detection in a color scene. *Pattern Recognition*, 1996, Vol.29 (6):1007—1017
- [9] Wu H., Chen Q. & Yachida M, Face detection from color images using a fuzzy pattern matching method. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1999, Vol.21 (6):557—563
- [10] 梁路宏, 艾海舟, 何克忠. 基于多模板匹配的单人脸检测. *中国图像图形学报*, 1999, Vol.4A (1):825—830
- [11] 刘翼光, 沈理. 利用 Hausdorff 距离人脸图像定位算法. *计算机研究与发展*, 2001, Vol.38 (4):475—481
- [12] Yuille A., Hallinan P. & Cohen D, Feature extraction from faces using deformable templates. *International Journal of Computer Vision*, 1992, Vol.8 (2):99—111
- [13] 马宇飞, 白雪生, 徐光佑等. 新闻视频中口播帧检测方法的研究. *软件学报*, 2001, Vol.12 (3):371—381
- [14] 山世光, 高文, 陈熙林. 基于纹理分布和变形模板的面部特征提取. *软件学报*, 2001, Vol.12 (4):570—577
- [15] Moghadam B. & Pentland A, Probabilistic visual learning for objects representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1997, Vol.19 (7):696—710
- [16] Zhu Y., Schwartz S. & Orchard M, Fast face detection using subspace discriminate wavelet features. *IEEE Conference on Computer Vision and Pattern Recognition*. Hilton Head Island,

- South Carolina, USA, 2000: 636—642
- [17] Sung K. & Poggio T, Example-based learning for view-based human face detection. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1998, Vol.20(1):39—51
 - [18] Rowley H.A. & Baluja S, Neural network-based face detection. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1998, Vol.20(1):23—38
 - [19] Rowley H.A., Baluja S. & Kanade T, Rotation invariant neural network-based face detection. *IEEE Conference on Computer Vision and Pattern Recognition*, 1998: 38—44
 - [20] Nefian A.V. & Hayes M.H, Face detection and recognition using hidden Markov models. *Proceedings of IEEE Conference on Image Processings*, Chicago, 1998: 141—145
 - [21] Nefian A.V. & Hayes M.H, An embedded HMM based approach for face detection and recognition. *Proceedings of IEEE Conference on Acoustics, Speech and Recognition*, Phoenix, Arizona, 1999: 3553—3556
 - [22] Boser B.E., Guyon I.M. & Vapnik V.N, A training algorithm for optimal margin classifier. *Proceedings of the 5th ACM Workshop on Computational Learning Theory*, Pittsburgh, 1992: 144—152
 - [23] Osuna.E., Freund.R. & Girosi.F, Training support vector machines: an application to face detection. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Pucrot Rico, 1997: 130—136
 - [24] Platt J, Using analytic QP and sparseness to speed training of support vectors. *MIT Press*, 1999.: 112—116
 - [25] Viola P. & Jones M, Robust real time object detection. *8th IEEE International Conference on Computer Vision*, Vancouver, British Columbia, 2001
 - [26] Yoo.T.W. & Oh.I.S, A fast algorithm for tracking human face based on chromatic histograms. *Pattern Recognition Letters*, 1999, Vol.20(10): 967—978
 - [27] Terrillon.T.C., Pilpre.T. & Niwa.Y, Robust face detection and hand posture recognition in color images for human-machine interaction. *Pattern Recognition*, 2002, Vol.1: 11—15
 - [28] Gergesha.M. & Panchanathan.S, Face detection from color images by interactive thresholding on skin probability maps. *IEEE International Symposium*, 2002, Vol.5: 673—676
 - [29] Rein-Lien H., Abdel-Mottaleb.M. & Jain.A.K, Face detection in color images. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2002:696—706
 - [30] Rein-Lien H, Face Detection and Modeling for Recognition[D].PhD thesis, Michigan State University,2002
 - [31] <http://ise.stanford.edu/>
 - [32] Viola P. & Jones M, Robust real time object detection. *IEEE ICCV Workshop on Statistical and Computational Theories of Vision*, Vancouver, Canada, 2001
 - [33] Li S.Z., Zhang Z.Q. & Shum H, FloatBoost learning for classification. *Proceedings of the 16th Annual Conference on Neural Information Processing Systems*. Vancouver, Canada, 2002

- [34] Rong X., Long Z. & Hong-Jiang Z, Boosting chain learning for object detection, Ninth IEEE International Conference on Computer Vision (ICCV'03), Vol.1: 709
- [35] Guyon I., Weston J. & Barnhill S, Gene Selection for Cancer Classification Using Support Vector Machine. BIONWulf Technical Report, 2000
- [36] Freund Y. & Schapire R.E, A decision-theoretic generalization of online learning and an application to Boosting. Journal of Computer and System Sciences, 1997, Vol.55(1):119-139
- [37] Vapnik V.N, Statistical learning theory. John Wiley and Sons, Inc., New York, 1998
- [38] Shinjiro K. & Jun O, Automatic skin-color distribution extraction for face detection and tracking, ICSP 2000. The 5th Int. Conf. On Signal Processing, 2000, Vol.2: 1415—1418
- [39] Rein-Lien H., Mohamed A.M. & Anil K. J, Face detection in color images. IEEE Trans on Pattern Analysis and Machine Intelligence, 2002, Vol.24 (5): 696—706
- [40] Waqar M., Noman A. & Chung-Tse M, Face detection project. Department of Electrical Engineering. Stanford University, 2003: 1—12
- [41] 孙印详等著. 模式识别中的特征提取与计算机视觉不变量. 北京: 国防工业出版社, 2001
- [42] Yuille A.L., Hallinan P.W. & Cohen D.S, Feature extraction from faces using deformable templates. International Journal of Computer Vision. 1992, Vol.8 (2):99—111
- [43] 叶亮. 一种基于变形模板匹配的人脸检测方法. 计算机工程, 2004, Vol.30(11): 115—117
- [44] 袁超, 张长水. 基于多模板匹配的自动人脸检测. 电子学报, 2000, Vol.28 (3): 95—98
- [45] 张春雨, 陈绵书, 陈贺心. 人脸正面图像的机器识别. 计算机工程与应用, 2004, Vol.40 (2): 62—65
- [46] Rowley H A, Baluja S. & Kanade T, Neural network-based face detection. IEEE Trans PAMI, Vol.20(1):23—38
- [47] Edgar O., Robert F. & Federico G, Training support vector machines: an application to face detection. Computer Vision and Pattern Recognition, 1997: 130—136
- [48] 梁路宏等. 基于仿射模板匹配的多角度单人脸定位. 计算机学报, Vol.23 (6): 640—645
- [49] 姜军, 张桂林. 一种基于知识的快速人脸检测方法. 中国图像图形学报, 2002, Vol.7A (1):6—10
- [50] 梁路宏, 艾海舟. 基于模板匹配和人工神经网络确认的人脸检测. 电子学报, 2001, Vol.4A (6):744-747
- [51] Bartlett M.S., Lades H.M. & Sejnowski T.J, Independent component representation for face recognition. Proc. SPIE Conf. on Human Vision and Electronic Imaging III[C], 1998, 3299:528-539
- [52] 杨福生等. 独立分量分析及其在生物医学工程中的应用. 国外医学生物医学工程分册 [M], 2000, Vol.23 (3): 129
- [53] Francis R.B. & Michael I.J, Kernel independent component analysis, Journal of Machine Learning Research, 2003, Vol.3 (1): 1-48
- [54] Holland J.H. Adaptation in natural and artificial systems. MI: University of Michigan Press[M], 1975

- [55] Platt J, Using analytic QP and sparseness to speed training of support vector vectors[M]. MIT Press. 1999
- [56] Hyvarinen A, fast fixed-point algorithm for independent component analysis. Neural Computation, 1997, Vol.9 (7):1483—1492

个人简历 在读期间发表的学术论文与研究成果

个人简历:

庄茵芳, 女, 1979年11月23日生。

2002年7月毕业于同济大学, 计算机科学与技术专业, 获得学士学位。

2003年9月就读于同济大学, 攻读硕士研究生。

已发表的论文:

[1] 武妍, 庄茵芳。基于核独立分量分析的人脸识别方法。计算机科学 2005, Vol. 32 (8.A): 264—267

待发表论文:

[1] 武妍, 庄茵芳。基于 KICA 和链式 Boosting 算法的人脸检测。ISSN 2006

作者: [庄茵芳](#)
学位授予单位: [同济大学](#)

参考文献(56条)

1. [Yang M H, Kriegman D J, Ahuja N](#) [Detecting faces in images:a survey](#) 2002(01)
2. [梁路宏, 艾海舟, 徐光档, 张钊](#) [人脸检测研究综述](#)[期刊论文]-[计算机学报](#) 2002(5)
3. [Kin C Y, Cipolla R](#) [Feature-based human face detection](#)[外文期刊] 1997(09)
4. [Yang G, Huang T S](#) [Human face detection in complex background](#) 1994(01)
5. [Morimoto C, Flickner M](#) [Real-time multiple face detection using active illumination](#) 2000
6. [Martin H, Hunke H](#) [Locating and tracking of human faces with neural networks](#) 1994
7. [Tony S, Pentland A](#) [Parametrized structure from motion for 3D adaptive feedback tracking of faces](#) 1997
8. [Dai Y, Nakano Y](#) [Face-texture model based on SGLD and its application in face detection in a color scene](#)[外文期刊] 1996(06)
9. [Wu H, Chen Q, Yachida M](#) [Face detection from color images using a fuzzy pattern matching method](#)[外文期刊] 1999(06)
10. [梁路宏, 艾海舟, 何克忠](#) [基于多模板匹配的单人脸检测](#)[期刊论文]-[中国图象图形学报](#) 1999(10)
11. [刘翼光, 沈理](#) [利用Hausdorff距离人脸图像定位算法](#)[期刊论文]-[计算机研究与发展](#) 2001(4)
12. [Yuille A, Hallinan P, Cohen D](#) [Feature extraction from faces using deformable templates](#) 1992(02)
13. [马宇飞, 白雪生, 徐光右, 史元春](#) [新闻视频中口播帧检测方法的研究](#)[期刊论文]-[软件学报](#) 2001(3)
14. [山世光, 高文, 陈熙霖](#) [基于纹理分布和变形模板的面部特征提取](#)[期刊论文]-[软件学报](#) 2001(4)
15. [Moghadam B, Pentland A](#) [Probabilistic visual learning for objects representation](#) 1997(07)
16. [Zhu Y, Schwartz S, Orchard M](#) [Fast face detection using subspace discriminate wavelet features](#) 2000
17. [Sung K, Poggio T](#) [Example-based learning for view-based human face detection](#)[外文期刊] 1998(01)
18. [Rowley H A, Baluja S](#) [Neural network-based face detection](#)[外文期刊] 1998(01)
19. [Rowley H A, Baluja S, Kanade T](#) [Rotation invariant neural network-based face detection](#) 1998
20. [Nefian A V, Hayes M H](#) [Face detection and recognition using hidden Markov models](#) 1998
21. [Nefian A V, Hayes M H](#) [An embedded HMM based approach for face detection and recognition](#) 1999
22. [Boser B E, Guyon I M, Vapnik V N](#) [A training algorithm for optimal margin classifier](#) 1992
23. [Osuna E, Freund R, Girosi F](#) [Training support vector machines:an application to face detection](#) 1997
24. [Platt J](#) [Using analytic QP and sparseness to speed training of support vectors](#) 1999
25. [Viola P, Jones M](#) [Robust real time object detection](#) 2001
26. [Yoo T W, Oh I S](#) [A fast algorithm for tracking human face based on chromatic histograms](#) 1999(10)
27. [Terrillon T C, Pilpre T, Niwa Y](#) [Robust face detection and hand posture recognition in color images for human-machine interaction](#) 2002
28. [Gergesha M, Panchanathan S](#) [Face detection from color images by interactive thresholding on skin probability maps](#) 2002
29. [Rein-Lien H, Abdel-Mottaleb M, Jain A K](#) [Face detection in color images](#) 2002

30. [Rein-Lien H Face Detection and Modeling for Recognition](#) 2002
31. [查看详情](#)
32. [Viola P. Jones M Robust real time object detection](#) 2001
33. [Li S Z. Zhang Z Q. Shum H FloatBoost learning for classification](#) 2002
34. [Rong X. Long Z. Hong-Jiang Z Boosting chain learning for object detection](#)
35. [Guyon I. Weston J. Barnhill S Gene Selection for Cancer Classification Using Support Vector Machine](#) 2000
36. [Freund Y. Schapire R E A decision-theoretic generalization of online learning and an application to Boosting](#) 1997(01)
37. [Vapnik V N Statistical learning theory](#) 1998
38. [Shinjiro K. Jun O Automatic skin-color distribution extraction for face detection and tracking](#) 2000
39. [Rein-Lien H. Mohamed A M. Anil K J Face detection in color images](#) 2002(05)
40. [Waqar M. Noman A. Chung-Tse M Face detection project](#) 2003
41. [孙即详 模式识别中的特征提取与计算机视觉不变量](#) 2001
42. [Yuille A L. Hallinan P W. Cohen D S Feature extraction from faces using deformable templates](#) 1992(02)
43. [叶亮 一种基于变形模板匹配的人脸检测方法\[期刊论文\]-计算机工程](#) 2004(11)
44. [袁超. 张长水 基于多模板匹配的自动人脸检测\[期刊论文\]-电子学报](#) 2000(3)
45. [张春雨. 陈绵书. 陈贺新 人脸正面图像的机器识别\[期刊论文\]-计算机工程与应用](#) 2004(2)
46. [Rowley H A. Baluja S. Kanade T Neural network-based face detection\[外文期刊\]](#)
47. [Edgar O. Robert F. Federico G Training support vector machines:an application to face detection](#) 1997
48. [梁路宏 基于仿射模板匹配的多角度单人脸定位](#)
49. [姜军. 张桂林 一种基于知识的快速人脸检测方法\[期刊论文\]-中国图象图形学报A辑](#) 2002(1)
50. [梁路宏. 艾海舟. 徐光佑. 张钺 基于模板匹配与人工神经网络确认的人脸检测\[期刊论文\]-电子学报](#) 2001(6)
51. [Bartlett M S. Lades H M. Sejnowski T J Independent component representation for face recognition](#) 1998
52. [杨福生 独立分量分析及其在生物医学工程中的应用\[期刊论文\]-国外医学\(生物医学工程分册\)](#) 2000(3)
53. [Francis R B. Michael I J Kernel independent component analysis](#) 2003(01)
54. [Holland J H Adaptation in natural and artificial systems](#) 1975
55. [Platt J Using analytic QP and sparseness to speed training of support vector vectors](#) 1999
56. [Hyvarinen A fast fixed-point algorithm for independent component analysis](#) 1997(07)

本文读者也读过(9条)

1. [高鹏东 人脸检测和造型方法的研究\[学位论文\]](#)2003
2. [杨杰. 吴鹏. YANG Jie. WU Peng 人脸检测系统在DSP上的实现\[期刊论文\]-武汉理工大学学报\(信息与管理工程版\)](#) 2007, 29(10)
3. [杨海城 面向复杂场景的人脸检测\[学位论文\]](#)2006
4. [孙海峰 基于AdaBoost人脸检测算法在复杂背景中的研究与应用\[学位论文\]](#)2010
5. [刘佳. 韩鹏 浅谈人脸检测技术\[期刊论文\]-黑龙江科技信息](#)2009(14)
6. [张琳 基于多种信息综合的动态人脸检测\[学位论文\]](#)2005

7. [韩相军](#), [关永](#), [王雪立](#), [万玛宁](#) [基于DSP的人脸检测系统研究](#)[会议论文]-2005
8. [杨林](#) [基于Web的人脸检测跟踪系统的设计与实现](#)[学位论文]2005
9. [多姿态小目标人脸检测算法研究](#)[期刊论文]-[计算机工程与设计](#)2009, 30 (20)

本文链接: http://d.g.wanfangdata.com.cn/Thesis_Y855313.aspx