

安 阳 师 范 学 院

安阳师范学院本科毕业生毕业论文

基于伽马和对数预处理的人脸和手势识别系统

作 者 马新程
系（院） 计算机与信息工程学院
专 业 计算机科学与技术
年 级 2020 级
学 号 200912126
指导教师 郭 安
论文成绩 中 等
日 期 2024 年 5 月

安 阳 师 范 学 院

学生诚信承诺书

郑重承诺：所呈交的论文是作者个人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写的研究成果，也不包含为获得安阳师范学院或其他教育机构的学位或证书所使用过的材料。与作者一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

作者签名：马新程

日期：2024年5月20日

导师签名：孙华

日期：2024年5月20日

院长签名：孙华

日期：2024年5月20日

论文使用授权说明

本人完全了解安阳师范学院有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。保密论文在解密后遵守此规定。

作者签名：马新程

导师签名：孙华

日期：2024年5月20日

安 阳 师 范 学 院

目 录

摘 要:	1
1 引言	1
1.1 研究背景	1
1.2 研究目的和意义	1
1.3 研究内容及创新点	1
2 相关技术介绍	2
2.1 人脸识别技术介绍	2
2.2 手势识别技术介绍	2
2.3 百度 AI 平台介绍	2
2.4 图像处理技术介绍	3
3 图像预处理算法	3
3.1 图像预处理思路	3
3.2 对数变换方法	3
3.3 伽马变换方法	4
3.4 直方图均衡化处理	5
4 系统的设计与实现	6
4.1 总体设计思路	6
4.2 详细设计	7
5 系统演示和优化	13
5.1 各模块功能展示	13
5.2 系统问题及优化	14
5.2.1 系统存在的问题	14
5.2.2 系统问题的解决和优化	14
6 总结与展望	14
参考文献	15

基于伽马和对数预处理的人脸和手势识别系统

马新程

(安阳师范学院计算机与信息工程学院, 河南 安阳 455000)

摘 要: 人脸识别和手势识别是计算机视觉中的重要分支, 随着人工智能的不断发展这两项技术在智能门禁, 安防监控和人机交互等领域广泛应用, 对人脸识别和手势识别技术的不断研究在提高安全性, 增强便利性, 促进智能化发展以及人机交互等方面都具有重要意义。传统的人脸识别和手势识别会受到图像清晰度和光照等因素的影响, 图像的曝光度过高和过低都会使图像损失部分细节从而影响识别的准确度, 为了改善这一影响, 在进行人脸和手势识别之前, 先使用伽马变换和对数变换对曝光度高于180和低于50的图像进行预处理使图像的曝光度处于一个合理的范围, 再将预处理过的图像上传并进行识别可以提高图像识别的准确度。本系统使用QT进行界面开发并结合百度AI所提供的图像识别接口以及图像预处理方法来实现人脸和手势识别的功能。

关键词: 人脸识别; 手势识别; 百度AI; 伽马变换; 对数变换

1 引言

1.1 研究背景

随着人工智能技术的不断发展, 人脸识别和手势识别作为计算机视觉领域的重要研究方向, 已经在各个领域被广泛应用。人脸识别和手势识别技术可以应用于身份验证, 安防监控, 人机交互和医疗辅助等多个领域。对于提高社会生产力和改善人们生活质量有着重要意义。

在过去的几年时间中, 百度 AI 平台推出了一系列先进的人脸识别和手势识别技术, 为相关的领域研究和应用提供了强大的技术支持。其中, 基于深度学习的人脸识别技术在识别准确率和速度上取得了显著的进展, 可以更加快速和准确的对人脸信息进行识别; 而基于神经网络的手势识别技术可以实现对手势信息更加高效的识别和解析, 为人机交互和虚拟现实等领域提供了更加自然和直观的交互方式^[1]。然而, 当前的人脸识别和手势识别技术在一些特定的背景下仍然存在不足之处, 例如在复杂环境下的光照变化, 姿态变化和关键部位的遮挡等情况会造成识别准确率的下降。以及对不同年龄, 性别的人群会出现识别差异的情况。因此进一步研究来提升人脸和手势识别技术的准确性, 适用性和鲁棒性具有重要的研究意义和实用价值。本文主要基于百度 AI 提供的人脸和手势识别技术, 深入研究人脸识别和手势识别的相关算法原理和方法, 通过图像预处理的方法来优化人脸识别和手势识别的准确度。

1.2 研究目的和意义

人脸识别和手势识别作为当前人工智能领域的研究热点, 它们的目标是提高人脸和手势识别的准确度, 使系统可以更加准确高效的从图像中识别出人脸和手势信息。人脸和手势识别技术的核心目的是通过对识别技术的不断优化和研究使计算机能够达到像现实世界中人们对人脸和手势信息识别的效果。人脸识别技术在安全认证、监控安防和社交娱乐等领域中发挥着重要作用。这项技术可以用于身份验证与权限控制, 比如手机的人脸识别解锁和一些地方的智能门禁等安全领域, 该技术可以提供比传统密码验证更加便捷高效的认证方式。在一些公共场所和金融机构等重要领域, 人脸识别技术可以应用于实时监控和识别, 对违法犯罪行为进行识别和预警从而提高安全性。人脸识别技术在社交娱乐领域也被广泛应用, 随着短视频领域的兴起, AI换脸和人脸动画特效也流行起来, 增加了用户在社交平台上的趣味性和

互动性。手势识别技术的研究目的主要体现在人机交互,辅助工具和虚拟现实等方面。在智能设备操控中,手势识别以其直观、自然的优势,逐渐替代传统按键指令,广泛应用于智能家居控制,显著提升用户操作便捷度与设备使用效率。在虚拟现实手势识别也被广泛应用,用户可以通过一些手势操作完成对虚拟世界的物体操作,增加了用户与虚拟世界的交互性。手势识别技术也可以作为辅助工具来使用,对于存在沟通障碍的残障人士,手势识别技术能够将手语翻译为文字,打破语言壁垒,方便这类人群与外界沟通,切实提升了生活质量。综上所述,人脸识别和手势识别技术的研究对提高生活便利性,增强安全性,改善用户体验和推动技术进步都有着重要意义。

1.3 研究内容及创新点

本文的研究内容是基于百度AI的人脸和手势识别系统,该系统主要是对百度AI提供的人脸识别和手势识别接口进行封装,实现人脸和手势识别功能。通过QT对系统的图形界面进行开发,实现功能交互。本系统人脸识别和手势识别的图像获取主要通过电脑的摄像头来进行获取,通过这一方式获取的图像质量会受到光照,姿态和摄像头像素等因素的影响。为了改善这一因素,通过使用一些图像预处理方法对图像进行处理,从而改善图像的质量来提高图像上传识别的准确度。本系统所采用的图像预处理策略,是基于OpenCV对不同曝光度的图片采取针对性的图像处理策略来优化图像的质量。图像的曝光度主要是根据图像的灰度级来确定的,使用OpenCV将原始图像转化成灰度图,然后对图像的灰度值进行平均值计算,将计算的平均灰度值作为图片的曝光度。对于曝光度低于50的图像,采用对数变换的方式来对低灰度级的图像进行处理,从而提高图像的亮度然后再结合直方图均衡化处理来增强图像的对比度。曝光度高于180的图片,被认定为过曝光的图片,这类图片会损失图像原有的部分细节,所以采用伽马变换的方式对图像的各个像素进行处理,来降低图像的灰度级使图像变暗,再使用直方图均衡化处理来增强图像的对比度。如果图像的曝光度介于50和180之间则直接采用直方图均衡化处理,使图像的灰度级更加均匀从而提高图像的对比度。通过上述的图像预处理算法对摄像头获取的图片进行针对性的处理,再上传到相关的人脸和手势识别接口进行识别可以提高图像识别的质量。

2 相关技术介绍

2.1 人脸识别技术介绍

人脸识别是通过计算机的相关方法对图像中的人脸进行定位和识别的一种技术。该过程涉及人脸检测、对齐、特征提取和比对。人脸检测确定图像中人脸的存在和位置;人脸对齐通过几何变换消除不同图像间人脸位置和姿势的差异;特征提取则将人脸转化为特征集合或向量,常见方法包括局部二值模式、线性判别分析和主成分分析等。最后,将待识别图像的人脸特征与已有数据进行比对,以确定人脸信息。

然而,传统方法受图像清晰度、光照和特征提取质量等因素影响,识别准确度有限。随着深度学习特别是卷积神经网络(CNN)的发展,人脸识别技术取得了显著进步,CNN通过多层卷积和池化操作提取高度抽象且富含语义的图像特征,这些特征可以更全面的反映图像真实内容^[2]。为提升模型稳定性和适应能力,通常采用数据增强策略来丰富训练数据,正则化技术减少过拟合现象,以及迁移学习方法在已有模型上微调。这些策略使得结合深度学习的人脸识别体系展现出更高的准确性和抗干扰能力。

2.2 手势识别技术介绍

手势识别是通过计算机技术识别手部信息的过程,主要包括检测、跟踪和识别三个步骤。检测基于颜色、形状和纹理等特征确定手部位置;跟踪则运用运动分析或模板匹配追踪手部运动轨迹,如光流法和卡尔曼滤波器等方法;识别则是提取手部形状、运动等特征,通过分类器如支持向量机、决策树等方法来识别手势含义。

传统技术受限于图像质量和特征提取质量。而结合深度学习,手势识别准确性得以提升,通过大量数据集训练神经网络提取更深层次的特征表示,这些特征更抽象、更具代表性,能更好地反映手势含义^[3]。数据增强和正则化技术增强模型泛化能力和拟合性,迁移学习则优化已训练模型的性能。因此,深度学习在手势识别中广泛应用,克服了传统技术的局限性。

2.3 百度 AI 平台介绍

百度AI平台依托其核心引擎百度大脑,整合了视觉、自然语言处理、深度学习等热门AI技术。其中,人脸识别技术基于深度学习,通过CNN提取图像特征,RNN分析特征,并利用SVM进行分类预测,实现高精度的人脸及表情识别^[4]。手势识别技术同样基于深度学习,能自然、直观地检测并识别手势,提升用户体验。该技术通过图像获取、特征提取、模型分析和结果转化等步骤,将手势转化为指令或操作。百度AI平台以其强大的技术支持,为开发者提供了极大的便利。

2.4 图像处理技术介绍

图像处理是利用计算机技术改造和优化图像的过程,涵盖增强、分割、滤波等操作。图像增强通过调整亮度、对比度和降噪来优化视觉效果。图像在计算机中表现为像素点的二维矩阵,处理核心是对这些像素值进行计算和变换,以实现不同视觉效果。

本系统采用OpenCV开源计算机视觉库进行图像处理。OpenCV功能强大,包含C/C++编写的丰富函数和模块,支持多种语言接口,便于跨平台开发,它集成了图像处理、计算机视觉算法,如图像增强、滤波、人脸识别、光流分析、特征检测与匹配等^[5]。OpenCV经过优化,确保算法在各类硬件上高效运行,实现快速准确的图像处理和视觉应用。作为广泛应用的计算机视觉资源库,OpenCV在图像处理、视觉研究、医学影像分析等领域发挥重要作用。

3 图像预处理算法

3.1 图像预处理思路

由于摄像头获取的图片受到自身像素,光照等因素的影响所以需要对获取的图片进行预处理,或者采用像素水平更高的摄像头。如果采用其他摄像头通过本系统识别界面的下拉选项框选取外接的摄像头。如果不采用更高像素的摄像头则采用如下图像增强算法对象进行增强处理。

该系统所采用的图像增强策略基于图像的不同曝光等级实施分层优化。针对输入图片,本系统将其划分为三个曝光级别区间:当图像的曝光值低于阈值50时认定为低曝光度图片,首先运用对数变换方法提升图像的整体亮度,随后再配合直方图均衡化技术使图像的灰度级分布更加均匀以进一步强化图像的视觉效果;而当图像曝光度高于上限180时将图片认定为过曝光图片,则采取伽马校正方法适当降低过高的曝光水平使图像的曝光度处于50~180这一区间,在此之后同样应用直方图均衡化来增强图像的对比度;至于曝光度介于50至180之间的常规图像,则直接利用直方图均衡化技术进行图像质量改善。

3.2 对数变换方法

在OpenCV库中,对数变换是一种广泛应用的图像处理技术手段,其目的是为了优化图像的视觉对比度,特别是在凸显图像中暗部区域的细节表现方面效果显著。对数变换的主要思想是对图像的每一个像素进行取对数的操作,从而扩展暗部像素的细节使图像的整体亮度提高。如图1所示。

在图像处理过程中对数变换通常按照公式(1)对各像素点进行数学计算,其中 r 是输入图像的像素值即原始像素值, s 是输出图像的像素值, c 是一个调节参数,用来控制图像对比度的增强程度。

$$s = c * \log(1 + r) \quad (1)$$



图1 对数变换对低曝光度的处理

为了防止输出图像的像素经过对数处理后超过 $0 \sim 255$ 的区间,通常将对数变换和归一化结合使用使其经过处理后的图像像素在 $0 \sim 255$ 区间之内。对数变换之所以可以使图像的暗部细节得到提高主要是因为对数函数的性质,在公式(1)中,当原始像素 r 很小时经过计算该点会是一个负值,由于对数函数的增长速度很快,负值的对数值变化会比正值的对数值变化要大^[6]。因此对数变化后原始图像中像素较暗的像素值经过对数变换后会有更加明显的增加,可以理解为对数变换增强了低灰度级的像素权重,从而将暗部的细节拉伸到更明亮的像素范围使图像看起来更加明亮。

3.3 伽马变换方法

伽马变换是一种常见的图像增强算法,主要是通过调整每个像素的值来改变图片的对比度和亮度。伽马变换的底层逻辑涉及到了幂运算,通过对伽马值的设定来输出想要达到的图像效果,这种变化的目的是通过调整图像的对比度来改善图像的可视化效果。伽马变换通常按照公式(2)对图像的各像素点进行数学运算, c 为常数用来调节输出图像的亮度, r 为原来图像的每个像素的像素值, γ 是该公式的关键用来确定每个像素值进行的幂运算的大小。

$$s = c * (r^\gamma) \quad (2)$$

γ 的取值也有要求通常为大于0的常数,伽马的取值也决定了图像的对比度,在伽马变换过程中,若选取的伽马值小于1,则这一操作将会集中强化原始图像中相对暗淡的灰度层级部分;反之,若伽马值大于1,则该变换会侧重提升图像中较为明亮的灰度级别区域,以此达到调节和优化整体图像对比度的目的^[7]。如图2所示。



图2 伽马变换对高曝光度照片的处理

采用伽马变换处理曝光度过高的图片主要是因为伽马变换可以有效的改善过曝图像的质量。首先，过曝图像的像素动态范围非常大，图像中的部分细节可能会丢失，而伽马变换主要是通过幂运算的形式进行图像的像素调整，可以对图像的像素进行非线性压缩，将图像中高灰度值区域的细节拉回到合适范围内，使得图像的动态范围更加广泛从而适应更广泛的显示设备。伽马变换可以增强图像对比度，过曝光的图像通常对比度都较低，因为大部分像素都处于较高的灰度值，通过伽马变换可以调整图像的对比度，使图像中过曝光地方的细节更突出从而提高图像的识别度。通过伽马变换可以对图像的不同通道的色彩进行调节，使图像的色彩更加均衡，还原场景中的真实色彩信息。所以采用伽马变换处理过曝光的图片可以很好的提高图像的对比度。

3.4 直方图均衡化处理

上述提及的对数变换与伽马变换分别针对曝光不足和过曝光这两种极端情况提供了解决方案。而在处理曝光度介于两者中间状态的图像时，则首选直方图均衡化技术以提升图像质量。直方图均衡化的核心理念在于，通过调整图像像素值的分布格局，使其灰阶分布趋于均衡，由此达到增强整体对比度的效果。这种方法尤其适用于那些灰度分布不均的图像，经过均衡化处理后，能够有效凸显图像原本可能被忽视的细节部分^[8]。

直方图均衡化处理基本公式如公式(3)所示。 n_j 代表当前灰度级的像素个数， k 的范围是 $0 \sim L-1$ ， L 是图像中可能的灰度级总数， S_k 为经过变换后的图像像素值。

$$S_k = \sum_{j=0}^k \frac{n_j}{n} \quad (3)$$

最后根据处理过的像素值重新构建图像，将原始图像中的每一个像素都替换成经过均衡化处理后新像素，从而生成了直方图均衡化处理的图像。如图3所示。

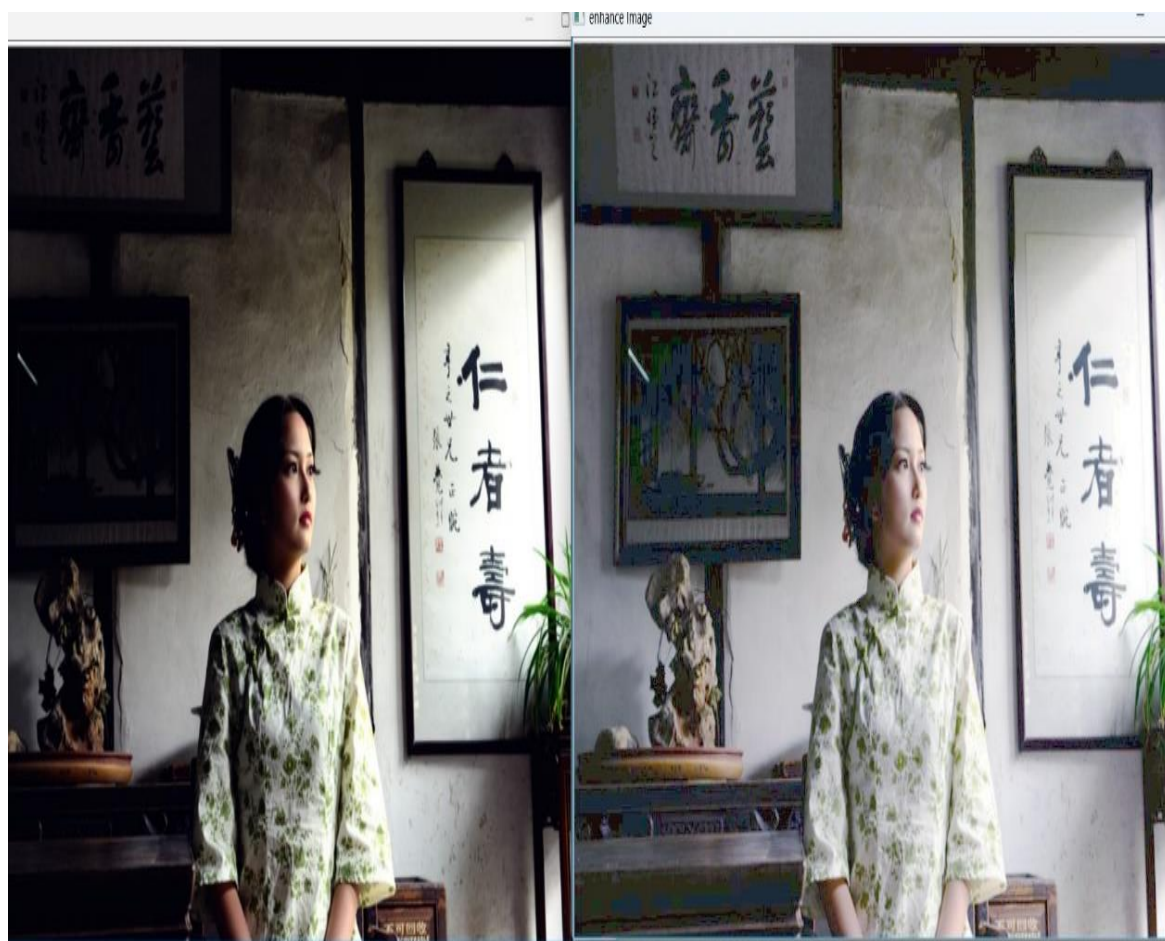


图3 直方图均衡化处理正常曝光度图片

综上所述，本系统采用对数变换，伽马变换和直方图均衡化处理来对摄像头获取的图片进行预处理，通过这几种图像处理算法的结合使用使原图像的曝光度和对比度都处于一个比较合适的范围，然后在进行图像的识别可以进一步提高图像识别准确度，从而达到预期的效果。

4 系统的设计与实现

4.1 总体设计思路

本系统采用了QT开发工具构建用户界面，QT是一款广泛应用于开发图形用户界面应用程序及嵌入式系统的跨平台C++框架编程工具。它包含了一系列的库和工具可以使开发人员创造出可移植，可扩展的高性能程序。系统主体分为三个界面，通过主界面进行交互选择进行人脸还是手势识别。人脸识别和手势识别界面由四个主要部分构成分别是图像显示控件，摄像头显示控件，拍照按钮和信息展示框组成。通过点击拍照按钮获取摄像头中的照片，然后进行图像的预处理，预处理思路为先将图片转化成灰度图，然后计算平均灰度值来作为后续曝光度的判断依据，对与曝光度低于50的图像先采用对数变换处理然后在使用直方图均衡化处理，曝光度高于180的图片采用伽马变换来降低图像的曝光度然后在使用直方图均衡化处理。如果图像的曝光度介于50~180之间则直接采用直方图均衡化处理。将预处理好的图像进行Base64编码格式转换并上传进行识别，对返回的图像识别信息进行格式化处理并从中提取出所需的人脸或手势信息并将其显示到信息展示框中。该系统执行流程如图4所示。

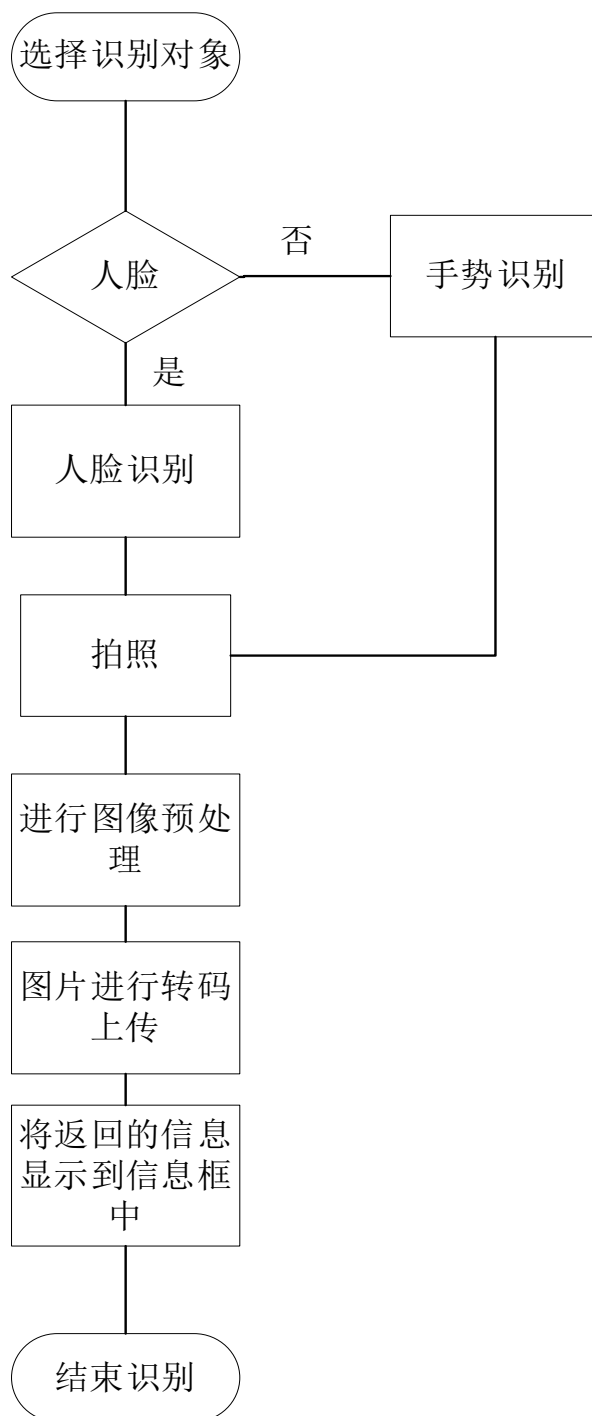


图4 系统流程图

4.2 详细设计

首先是对系统主界面的设计，设置一个主界面用来简单的功能跳转，主界面设置两个按钮分别跳转到两个不同的识别界面，点击按钮触发点击事件进行界面跳转，主界面后台保持运行用来进行多次界面跳转操作，如图5所示。



图5 系统主界面

子界面的功能是用来进行人脸或手势识别，该界面通过四个控件组成，使用标签控件来展示摄像头捕获的照片，设置一个按钮完成拍照操作当点击拍照按钮是触发对应的拍照方法将摄像头的照片进行捕获操作。使用摄像头取景器控件来展示电脑摄像头所获取的图像。使用文本框控件来存放人脸或手势识别的信息。人脸和手势识别界面布局效果如图6所示。

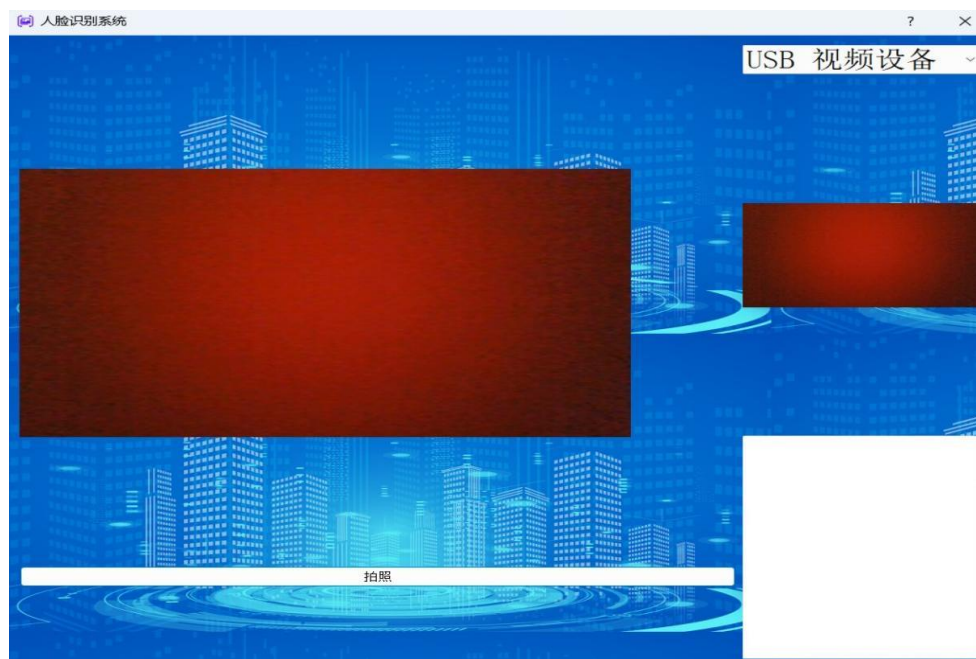


图6 子界面布局

子界面实现的功能是对摄像头获取的照片进行上传识别和显示识别的信息，对拍照按钮设置功能，设计相关的槽函数来响应按钮的点击事件，点击按钮触发拍照函数对摄像头的照片进行获取。由于该系统的图像获取来源主要是电脑的摄像头，在获取图像时会受到光照和

自身像素等因素的影响，可能会影响图像识别的质量。所以采用OpenCV对获取的图像进行预处理，将处理好的图像在进行转码上传识别。图像预处理的思路是针对不同曝光度的图像采取针对性的图像预处理方法，曝光度低于50的图片采用对数变换的方式来处理图像从而提高图像的曝光度，然后使用直方图均衡化处理进一步提高图像的对比度。曝光度高于180则使用伽马变换的方式来降低图像的曝光度，在使用直方图均衡化处理提高图像对比度。对于介于50到180曝光度之间的图片则直接使用直方图均衡化处理来提高图像的对比度。图像预处理的实现流程如图7所示。

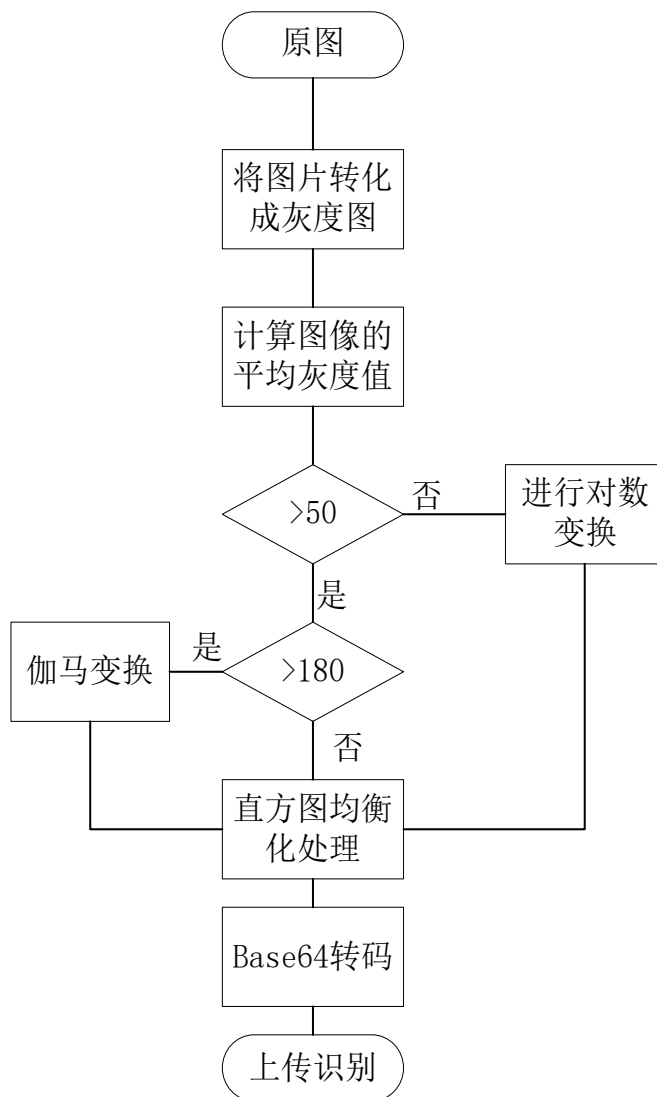


图 7 图像预处理流程图

通过上述图像预处理流程，对预处理模块有了核心方向，下面开始对图像预处理模块进行详细设计。首先需要对摄像头获取的图像进行灰度处理，将转化好的灰度图进行平均灰度值计算，得出的平均灰度值作为图像的曝光度，用于后续分类处理时的判断依据。

通过计算得出图像的平均灰度值，使用条件判断语句来将图片进行分类。如果图像的平均灰度值小于50则进行对数变换，对数变换是对图像上的每一个像素点进行数学运算，将运算结果替代原来的像素值。设计思路如下，首先建立一个与原图大小相同的Mat对象，然后使用循环遍历图像中所有的像素点，将每一个点的像素值通过公式(1)进行计算，并将计算过的像素值赋给创建的结果对象，由于对像素值进行的数学运算可能会导致部分像素点的值超出

0~255的范围，所以使用归一化的操作将新图像的像素值控制在0~255区间之内。对数变换后图像的灰度级分布图如图8所示。对数变换的伪代码如下。

算法 1 对数变换

输入: *src*(原图像)

输出: *dst*(对数变换后的图像)

```
1: dst  $\leftarrow$  src.size()(创建一个与原图大小相同的图片)
2: 遍历图片的所有像素点
3: for i  $\leftarrow$  0 ; i < src.row - 1 do
4:   for j  $\leftarrow$  0 ; j < src.col - 1 do
5:     dst  $\leftarrow$  对每一个像素点进行对数变换
6:   end for
7: end for
8: normalize(dst, 0, 255)对图像进行归一化操作
9: return dst
```

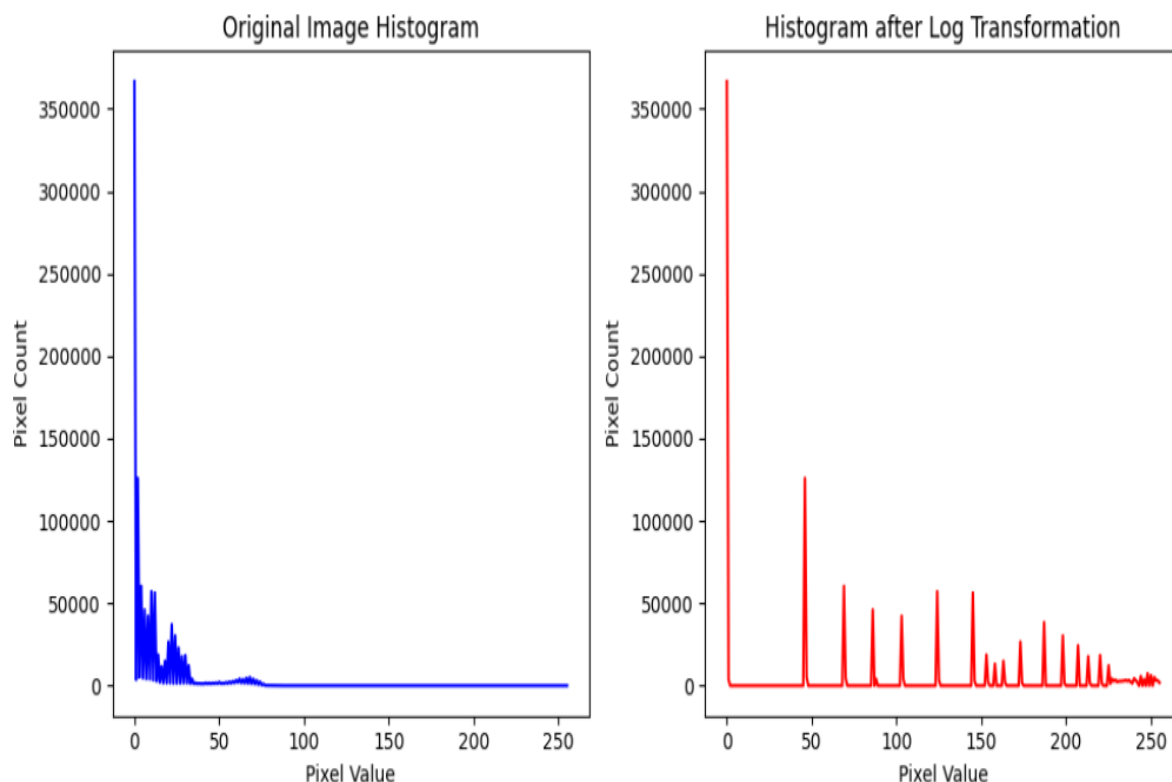


图8 对数变换后图像灰度图

如果图像的Mean值大于180，则需要使用伽马变换来降低图像的曝光度。伽马变主要涉及到幂运算，对图像上的每一个像素进行幂运算，将计算的结果赋给新的图像从而达到降低曝光度的操作。具体实现步骤为，首先创建一个长度为256的浮点数组，用来储存伽马映射函数，将0~255每个数的四次方存入数组中，然后通过循环遍历图像所有的像素，通过查找结果数组中对应的值来进行伽马变换，将变换后的结果赋值给新的图像，由于伽马变换也涉及到数学运算，可能造成像素值的越界问题，采用归一化的方式将像素值缩放到0~255之间。伽马变换后图像灰度级分布图如图9所示。伽马变换伪代码如下。

算法 2 伽马变换

输入: src (原图像)

输出: dst (伽马变换后的图像)

```
1: 初始化查找表
2:  $float\ pixels[256]$ 
3: for  $i \leftarrow 0 ; i < 255$  do
4:   将下标对应的值进行幂运算, 用于后续查找
5:    $pixels[i] \leftarrow powf(i, 4)$ 
6: end for
7: 创建输出图像
8:  $dst \leftarrow Mat(src.size())$ 
9: 遍历图像并执行变换
10: for  $i \leftarrow 0 ; i < src.rows - 1$  do
11:   for  $j \leftarrow 0 ; j < src.cols - 1$  do
12:      $dst \leftarrow pixels[src[i][j]]$ (从查找表找到伽马变换后的值)
13:   end for
14: end for
15:  $normalize(dst, 0, 255)$ 对图像进行归一化处理
16: return  $dst$ 
```

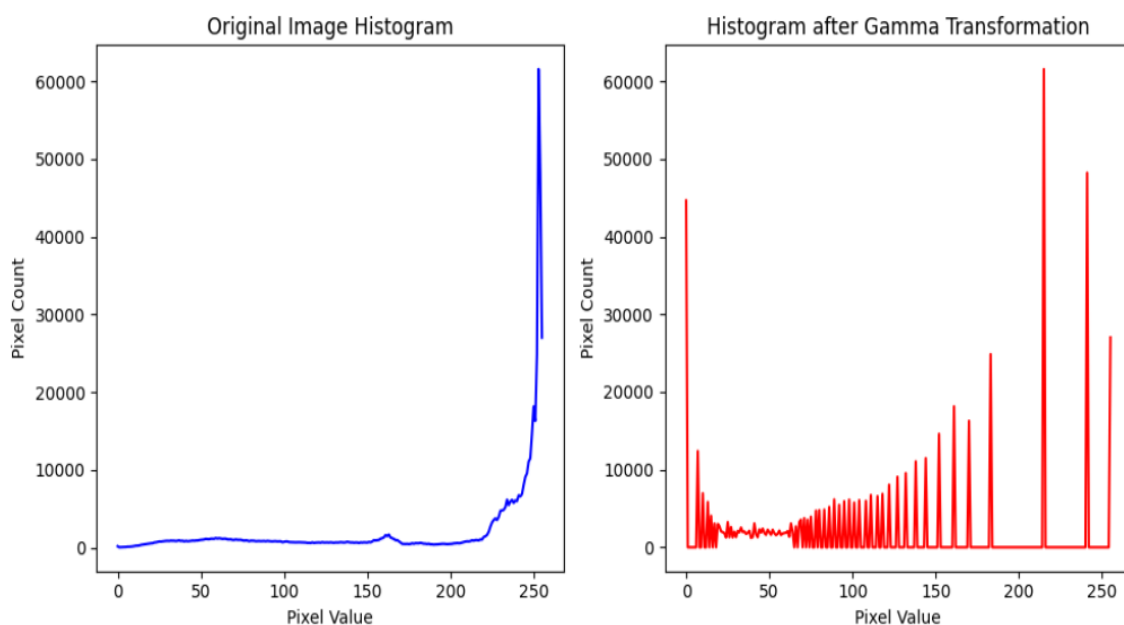


图9 伽马变换后图像灰度级分布图

若图像的曝光度介于50~180之间, 则对图片直接使用直方图均衡化处理。直方图均衡化处理是通过重新分布图像像素的强度值来扩展图像的动态范围, 使图像的像素分布更加均匀, 从而提高图像的对比度和视觉效果。本系统采用的直方图均衡化处理思路是, 首先创建一个与原图像大小相同的Mat对象用来储存结果图, 将原图像的每一个通道分离出来用一个临时对象储存, 通过循环对每一个通道的图像进行直方图均衡化处理, 并将结果赋值给创建的结果

变量。再通过合并操作将分离的通道合并为一个多通道图像。直方图均衡化处理后图像灰度级分布如图10所示。直方图均衡化伪代码如下。

算法 3 直方图均衡化

输入: *src*(原图像)

输出: *dst*(直方图均衡化处理的图像)

- 1: 创建与原图大小相同的输出图像
- 2: 创建*srcRgb*[]存放原图的色彩通道值
- 3: 将原图的通道数分离成独立通道存入*srcRgb*[]中
- 4: **for** $i \leftarrow 0 ; i < \text{src.channels}$ **do**
- 5: 将*srcRgb*[*i*]中的值进行直方图均衡化处理
- 6: **end for**
- 7: *dst* \leftarrow 将均衡化处理的通道进行合并
- 8: **return** *dst*

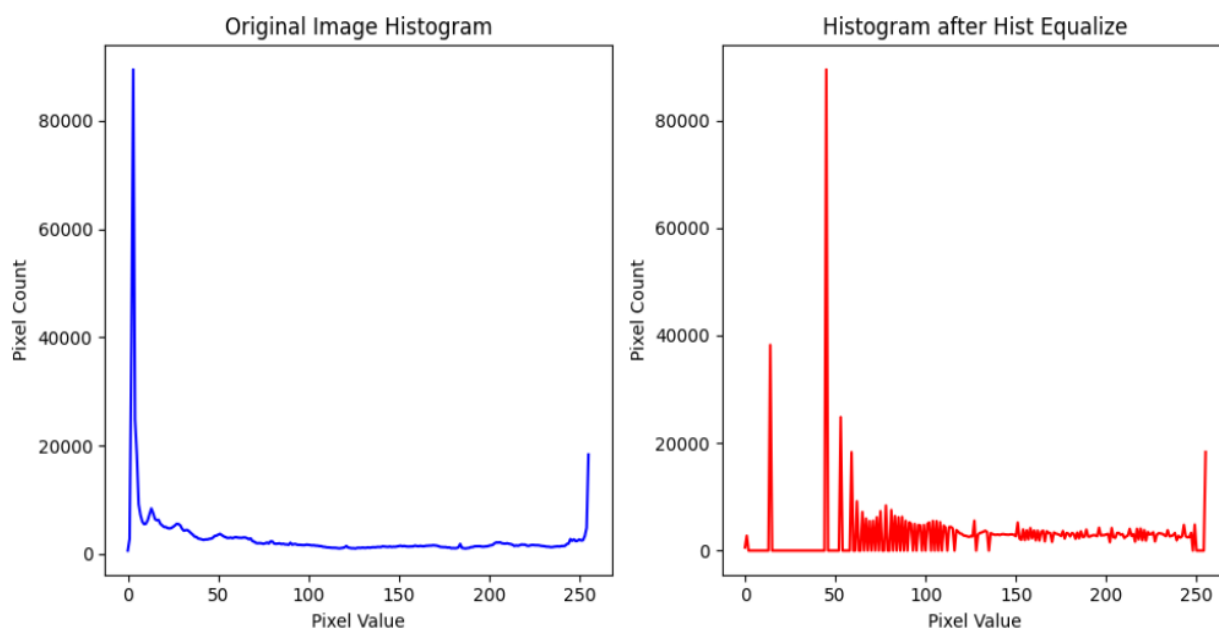


图10 直方图均衡化后图像灰度级分布图

对图像预处理结束后，便可以将图片上传识别了。在进行照片上传前需要对相应的接口进行配置，首先需要通过百度AI官方文档的信息使用HTTPS协议封装向相应接口的URL信息然后采用POST方法上传URL信息，然后解析返回的信息从中获取人脸识别和手势识别接口的Access_token信息，将预处理好的图片进行Base64编码格式转换，将预先获取的token值和处理过后图像的Base64编码按照HTTPS协议进行URL的格式封装通过POST方法进行数据上传^[9]。上述百度AI的接口调用均需要HTTPS协议作为网络传输协议，使用该协议的目的在于对数据安全性进行保护，保证数据传输的完整性同时确保身份安全性。该协议在HTTP协议的基础上结合了SSL/TLS安全协议，HTTPS通过加密的方式在网络上进行数据传输，可以有效的保护数据被非法获取和恶意篡改。特别是对于像百度AI接口这类可能涉及到人脸，语音等敏感信息的传输场景，采用加密传输机制尤为重要，它能切实保证数据在传输过程中的安全性。

将识别返回的信息进行格式化处理，提取所需的关键信息显示到文本展示框中。以上就是本系统的详细设计思路。

5 系统演示和优化

5.1 各模块功能展示

本系统主要有三个界面组成分别是主菜单界面，人脸识别界面和手势识别界面，主界面的功能是点击人脸识别或手势识别按钮可以进行相应界面的跳转，主界面实现效果图如图11所示。



图11 系统主界面

人脸识别界面的功能是将摄像头图像显示到界面的右侧，通过点击拍照按钮可以完成对人脸信息的识别，并将识别的信息显示到文本框中，根据提取的信息定位到人脸所在位置，并画出人脸位置框。效果如图12所示。



图12 人脸信息识别

在手势识别界面通过点击按钮完成对手势信息的识别和定位，并根据识别信息将手势位置在图像中绘制出来。效果如图13所示。



图13 手势信息识别

5.2 系统问题及优化

5.2.1 系统存在的问题

通过对主界面的功能和人脸及手势识别界面功能的测试，虽然该系统可以实现预期的基本功能但同样也存在一些问题，由于系统识别需要调用百度AI的人脸识别以及手势识别接口，该过程涉及到网络数据传输所以在连续进行识别操作时会因为网络的原因出现卡顿现象，为了避免这种现象需要使用比较稳定的网络环境。

排除网络问题的影响系统还有卡顿现象，通过检查发现是由于在进行系统设计时将图像预处理的图像增强算法和图像转码操作都写在了主程序中，这两个操作在程序执行时会占据系统的大量资源在结合网络数据传输会导致系统出现卡顿现象影响使用体验。

该系统进行识别时使用的电脑自带的摄像头，摄像头自身的像素水平会对图像的识别效果产生影响，同时由于电脑摄像头不能自由翻转和移动，用户识别自由度受到了限制需要通过相应的方法改善识别自由度。

5.2.2 系统问题的解决和优化

为了缓解由于资源消耗过高导致的系统运行卡顿问题，计划采用多线程技术。具体来说，是将本系统中耗时较多的操作置于子线程中并行执行。多线程编程作为并发编程的一种实践手段，借助系统支持并发执行的优势，使得程序可以在同一时间内运行多个线程，从而提升了系统的效率和响应速度。这一技术充分利用了现代计算机硬件多核与多处理器的特性，通过让多个任务并行开展，极大地提高了整个程序的运行速率。

本系统采用QT作为开发平台，在此环境下，对传统的线程处理进行了相应的封装处理。在QT中实现多线程操作的基本步骤大体如下所述：首要步骤是定义一个基于QThread类的线程对象，值得注意的是，此类是从QObject类派生而来^[10]。在这个自定义的线程类中，将具体实现子线程所需的业务逻辑，故而，可以将此类视作承载线程执行内容的一个容器。然后在线程类中进行具体功能逻辑的设计，通常放入比较耗时的操作，例如本系统中的图像增强算法相关的方法，以及对预处理过的图像进行Base64转码的操作。将这些耗时的操作放入子线程中是为了使系统执行更加流畅，以免阻塞UI界面的响应。其次对于线程操作的信号和槽函数的设计，信号和槽的机制是QT中线程通信的关键，需要设置两个信号分别代表线程的启动和线程执行结束，还需要设置一个槽函数用于完成图像增强操作和图像转码操作。通过多线程的使用在进行多次人脸和手势识别时系统的流畅度比未使用多线程时有明显提高。

在原系统的基础上,为了改善只能使用电脑自带摄像头的问题,添加了一个摄像头设备选择框。点击选择框可以选择电脑外接的其他摄像头设备。在进行设备切换时,系统会先将原来的摄像头资源进行关闭和释放,然后构建一个新的摄像头类进行图像的显示。通过这一方法可以实现摄像头设备的自由切换,使得进行图像识别时可以对任意方位的图像进行获取,而不局限于电脑屏幕前这一局部区域的图像,从而大大提高了用户使用时的自由度。

6 总结与展望

本文介绍了人工智能在日程生活中的广泛应用,对人脸识别和手势识别技术进行了深入研究。本文依托百度AI提供的深度学习和大数据分析支持的人脸和手势识别技术,设计实现了一个完整的人脸和手势识别系统。

本系统的设计采用QT框架进行用户界面开发实现简单的功能交互,并通过摄像头模块对图像进行获取。为了提高图像识别的质量,本文着重介绍了图像预处理的方法,利用OpenCV库对图像进行了对数变换,伽马变换以及直方图均衡化处理等图像增强算法,分别对低曝光,高曝光以及正常曝光的图像采取针对性的图像预处理操作来增强图像的质量^[1]。通过对图像进行一系列的图像增强算法,进一步加深了对图像处理的认识。在进行系统设计时还采用了多线程的技术来改进系统的性能,将图像预处理操作和图像转码操作放入子线程中执行,防止主线程因为计算密集型任务造成系统的阻塞。针对系统仅使用电脑自带的摄像头造成的系统识别局限性,提出了摄像头设备选择框的优化方案,通过这一方法使用户可以使用外接的摄像头进行识别,提高了识别视角的自由度。经过一系列的功能测试,该系统已经具备对摄像头中人脸和手势进行识别和定位的功能。

尽管该系统已经可以对人脸和手势进行识别,并且做出了一些改良。但是还可以进行进一步的优化,比如研究更加高效的图像预处理方法来适应更加复杂的光照影响和人体姿态变化进一步提高系统识别的抗干扰能力和识别准确性。未来对系统的设计应该考虑更多的应用场景如移动端,嵌入式设备等更为复杂的环境,从而实现更加完善的系统。

参考文献

- [1] 王育阳,古玉锋,肖子叶,等. 基于静态手势识别的智能交互系统研究[J]. 技术与市场, 2024, 31(01):20-24.
- [2] 唐荣辉. 新时代背景下基于深度学习的人脸识别技术开发探究[J]. 信息与电脑(理论版), 2024, 36(03):124-126.
- [3] 赵鸿图,李豪,梁梦华. 复杂背景下多特征结合的深度学习手势识别[J]. 电子测量技术, 2023, 46(23):77-84.
- [4] 张琪立,马民生,薛艳芬. 基于深度学习技术的人脸识别算法优化与应用研究[J]. 信息记录材料, 2023, 24(12):146-148.
- [5] 叶荣,何云,高泉,等. 基于OpenCV和模糊数学的茶叶病害分析方法研究[J]. 北方园艺, 2024, (04):145-153.
- [6] 于佳. 基于深度学习的图像对比度增强算法[D]. 北京邮电大学, 2023.
- [7] 程晶晶. 一种自适应低照度彩色图像增强算法[J]. 南宁师范大学学报(自然科学版), 2023, 40(03):51-56.
- [8] 莫文彬,伊力哈木·亚尔买买提. 通道注意力机制的局部遮挡人脸表情识别[J]. 计算机应用与软件, 2024, 41(02):145-151.

- [9] 张牧琢, 张隼霏. 基于百度AI的物体图像识别软件的设计与实现[J]. 电脑知识与技术, 2023, 19(33): 11-13.
- [10] 吴晓甜. 基于Qt的高性能MWD系统软件开发[D]. 西安石油大学, 2023.
- [11] 陈鹏慧, 张楠珂. 基于OpenCV的仿生机器鱼视觉系统的分析与研究[J]. 电子测试, 2022, 36(14): 31-33+47.

Face and gesture recognition system based on gamma and logarithmic preprocessing

MA Xincheng

(School of Computer & Information Engineering, Anyang Normal University, Anyang, Henan 455000)

Abstract: Face recognition and gesture recognition is an important branch of computer vision, with the continuous development of artificial intelligence, these two technologies are widely used in the fields of intelligent access control, security monitoring and human-computer interaction, and the continuous research on face recognition and gesture recognition technology is of great significance in improving security, enhancing convenience, promoting intelligent development and human-computer interaction. Traditional face recognition and gesture recognition will be affected by factors such as image sharpness and illumination. Too high or too low exposure of images will cause the image to lose some details, thus affecting the accuracy of recognition. In order to improve this impact, before face and gesture recognition, Firstly, gamma transform and logarithm transform are used to preprocess the images with exposure above 180 and below 50 so that the exposure of the images is in a reasonable range, and then the preprocessed images are uploaded and recognized to improve the accuracy of image recognition. This system uses QT interface development and combines the image recognition interface provided by Baidu AI and image preprocessing methods to realize the face and gesture recognition functions.

Key words: Face recognition; Gesture recognition; Baidu AI; Gamma transform; Logarithmic transform