

南开大学

本科生毕业论文（设计）

中文题目：一种挥手识别系统的设计与实现

外文题目：Design and Realization of A Hand Waving
Recognition System

学 号：1310650

姓 名：赵泽坤

年 级：2014 级

学 院：计算机与控制工程学院

系 别：自动化与智能科学系

专 业：智能科学与技术

完成日期：2018 年 5 月

指导教师：杨巨峰 副教授

关于南开大学本科毕业论文（设计） 的声明

本人郑重声明：所呈交的学位论文，是本人在指导教师指导下，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、已公开发表或没有公开发表的作品内容。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。

学位论文作者签名：

年 月 日

本人声明：该学位论文是本人指导学生完成的研究成果，已经审阅过论文的全部内容，并能够保证题目、关键词、摘要部分中英文内容的一致性和准确性。

学位论文指导教师签名：

年 月 日

摘 要

文章介绍了一种挥手动作识别系统的设计与实现方法。该系统通过人脸检测结果提取肤色特征，并结合人手和脸部的空间关系，采用基于区域的图像分割方法对手区域进行分割；之后，通过对手轮廓曲率的变化来定位指尖位置，利用指尖在轮廓凸包间的几何关系排除误检点；最后，在以上工作的基础上，利用长短滤波器对手位置周期性变化特征进行监测，用于识别挥手动作。综上，系统主要分为三个主要步骤：检测分割、分析定位、姿态识别。

该系统基于机器人操作系统(Robot Operating System, ROS)平台，通过主题-消息机制接收视频信号和发布挥手识别信号。为了证明系统的可靠性，在实际环境中对系统中三个步骤分别进行了检测速度和实现效果的分析。实验结果表明该系统可以完成实时、稳定的挥手识别任务。

关键词： 人机交互；肤色模型；人手检测；指尖定位；挥手识别

Abstract

This thesis describes a design and implementation method of a hand-waving recognition system. The system extracts skin color features based on face detection results, combines the spatial relationship between hands and face, and utilizes region-based image segmentation methods to segment the hand region; then, the fingertip position is located using the curvature change of the hand contours, and the fingers are positioned. The geometric relationship between the tip and the convex hull eliminates the misdetection point. Finally, based on the above work, the long-short filter is used to monitor the periodic variation of the position of the hand during waving, and is used to identify the waving movement. In summary, the system is mainly divided into three main steps: detection and segmentation, analysis and positioning, posture recognition.

In this paper, the system is applied to the robot operating system ROS platform to receive video signals and issue wave recognition signals through the topic-message mechanism. Through the experimental verification, the three steps of the system were analyzed respectively for the detection speed and the realization effect. The system can achieve real-time, stable hand-waving recognition task.

Key Words: Human-Computer Interaction; Skin Color Model; Hand Detection; Fingertip Positioning; Hand-Waving Recognition

目 录

摘 要	I
Abstract.....	II
目 录	III
第一章 绪论	1
第一节 研究背景与意义.....	1
第二节 相关研究工作.....	2
第三节 论文主要内容.....	3
第二章 挥手检测算法的设计	6
第一节 基于脸部信息的人手检测方法	6
2.1.1 人脸检测.....	6
2.1.2 颜色采样与区域估计.....	7
第二节 人手区域显著性分割	9
2.2.1 人手整体分割.....	9
2.2.2 指尖检测.....	11
第三节 挥手动作识别.....	13
第四节 本章小结.....	15
第三章 实验结果与分析	17
第一节 实验平台介绍.....	17
第二节 人手检测结果.....	18
3.2.1 评估标准.....	18
3.2.2 肤色分割效果.....	19

3.2.3 人手分割效果.....	21
第三节 指尖定位结果.....	24
第三节 挥手识别结果.....	25
第四节 本章小结.....	27
第四章 总结与展望	28
第一节 论文总结.....	28
第二节 研究方向展望.....	29
参考文献	30
致谢	32

第一章 绪论

第一节 研究背景与意义

智能计算设备在日常生活中的日益普及与发展,引发了人们对人与计算机之间更直观,更有效的交互方式的需求。随着计算机技术和计算机视觉理论的不断发展与完善,利用数字图像和视频对感兴趣目标进行检测和识别的研究工作也变得越来越热门。近些年,基于视觉的人机交互方式吸引了许多研究人员的兴趣,已经成为主流的人机交互研究方向。

在一个基于视觉的人机交互系统中,首先要进行的便是对感兴趣目标进行检测,也称之为目标提取技术。通常,目标检测技术会将感兴趣目标的分割和识别进行结合,并且有较好的准确性和实时性。手势是人类非语言交流的一种主要交流形式,分为静态手势和动态手势。作为日常生活中人与人之间相互交流的重要途径,手势能够增强演讲的说服力,用来强调和传递情感。例如,挥手或者指向这些手势都是某种无声的演讲。新一代人机交互技术的研究^[1]表明,由于大量的手势姿态可以完成人与人之间日常信息的交流与传递,基于手势的人机交互方式可以有效地替代当前基于键盘,鼠标以及各种触摸设备等传统的交互模式。同时,基于视觉的手势识别使得创建全新的,更加直观和强大的人机交互界面成为可能,使得未来人机交互不再局限于特殊接触式硬件。手势识别在游戏,虚拟现实,机器人技术,医疗保健环境中的非接触式界面,增强障碍物的日常生活以及手语识别等方面有很多不同的应用。

对于人手的准确检测有利于高效地提取图像和视频信号中的手部信息,是手势自动识别系统的首要环节^[2]。然而,由于复杂背景、形变以及运动造成的模糊等因素,导致自然场景下的人手检测仍然是一个具有挑战性的问题。人脸检测作为一个已经比较成熟的研究课题,并且可以提供对于人手检测的有效信息。因此,通过人脸检测实现对手部特征的提取是一个不错的方法。另外,挥手动作作为一种简单的动态手势,有传递信息简单,容易识别,具有周期性,不易产生歧义等优点。据此,作者在本文中介绍了一种实时、稳定的挥手识别系统,并应用该系统在实际环境中,并在多个视频序列上进行了实验。

第二节 相关研究工作

不论哪种手势的识别方法，都需要首先进行人手区域检测，其次是对人手特征的提取，最终实现对某种手势姿态的识别。为了实现稳定，高效的手势识别系统^[3]，三个难点也必须被克服：

1. 场景的多样性：即使在多变的场景条件下，针对不同的肤色、变化的光照条件，一个优秀的手势识别系统也应该正常识别到目标手势。
2. 处理的实时性：作为一种自然的人机交互方式，系统应该做到对目标手势的实时识别。简单、稳定的人手特征可以更好地被应用于机器视觉系统中。
3. 系统的灵活性：系统必须考虑到手势的变化范围。因为很少有人可以每次完成一模一样的动作，因此系统必须具有灵活性，并允许某一个手势在某个范围内的变动。

Mo 和 Zhao 等人提出的一些尝试解决这些问题的方法^[4,5]都对用户或应用场景进行了一定条件的限制。比如，假设背景的单一性、人手是场景中唯一的具有肤色颜色的物体、光照条件稳定等。然而，这些限制使得手势识别方法在不同场景下应用的可能性降低，使得算法缺乏普遍性。

为了解决复杂背景下基于视觉的人手检测问题，许多研究人员提出了通过提取人手的视觉显著特征，如颜色，运动信息^[6]，形状^[7]以及多特征融合^[8]来进行人手检测的方法。首先，人手颜色的提取必须选取一个合适的色彩空间来实施。在多个色彩空间中，肤色颜色都存在聚类特性，如标准化的 RGB 色彩空间、HSV 色彩空间、 $YCbCr$ 色彩空间以及 YUV 色彩空间等。优秀的色彩空间可以有效地将色度与颜色的亮度分量分开。这是由于通过仅使用色彩的色度相关分量，可以实现对在一定程度上光照变化下，人手颜色特征检测的稳健性。2012 年，Yang 等人^[6]利用人手的运动特征，采用背景减除来检测快速运动的人手区域，但是这种方法要求背景环境静止不动。基于形状特征的方法是依据人手轮廓来进行模板匹配来对人手区域进行检测，但在挥手动作中的人手存在一定的形变导致其效果并不理想。为了更好地检测到正确的人手区域，Sridhar 等人^[9]提出了基于包含深度信息的手势识别方法。Zhou 等人^[10]利用深度摄像头得到的深度信息，提出通

过深度阈值来区分人手区域的方法。其中，深度阈值通过预设人手质心相对摄像头的最远和最近距离来确定，在这个距离内的像素点作为人手像素点。文献^[11]中指出对检测的人手区域加上尺寸限制可以有效减少噪声干扰。另一方面，一些研究者提出通过身体其他部分位置来估计人手深度阈值的方法^[12,13]，避免了假设人手作为距离深度摄像机最近的物体，如利用人脸检测，和躯体检测等相关技术。当然，类似的方法也可以应用在彩色图像来估计人手的空间位置^[14]。

提取到图像中人手区域后，由于遮挡或者特殊光照条件等原因，人手肤色区域可能会出现不连续的情况。通过结合肤色信息和深度阈值是一个更好的解决人手分割问题的方法。

解决人手分割问题的另一个方式是通过聚类和区域生长^[15]的办法。通过聚类可以连接类似相邻点来获得最大连通区域，如肤色区域或者拥有同样深度信息的区域。区域生长方法^[15]则是通过在所检测的目标区域内部加入“种子点”，通过寻找相连接的点来生长，进而覆盖目标区域。1991年，Vincent^[16]提出了分水岭算法，用类似于盆地积水的过程处理图像分割和轮廓检测问题^[17]。2004年，Bieniecki^[18]指出在针对彩色图像利用分水岭算法存在一定的过分割问题。2011年，Chen等人^[19]通过人脸部分的“种子点”在深度图像中进行区域生长来得到整个身体区域。在通过预计的躯干半径和距离头部的距离信息来确定人手区域。

通过对近几年的手势识别领域的工作进行分析^[20,21]，大部分研究工作基于微软 Kinect 深度摄像机或者其他深度摄像机，利用骨骼数据进行关节点^[22]跟踪进行姿态的分类和识别，或通过 SVM、隐马尔科夫模型、神经网络等方法对各种手势姿态进行分类。本文立足于实践，并且只关心识别一类手势，因此采用了一种简单的基于算法的挥手动作识别方法。

第三节 论文主要内容

随着日益增长的社会需求，对现实生活中随处可见的各种视觉场景进行快速地、有效地分析显得越来越重要。本文的主要研究目标包括：人手检测方法研究和基于算法设计的挥手姿态识别的应用研究。

由于这些目标的实现都离不开对图像场景信息的理解和分析，并考虑到现实生活中复杂多变的环境信息，本文以人手检测为出发点，探索，并应用其技术在

机器人系统 ROS 中以进行模块化，尝试解决现实生活中一些实际应用问题，如机器人餐厅服务，智能机器唤醒等。本文在基于人脸检测的结果上提出了一种实时、稳定的挥手识别方法。本论文的内容分为以下几个章节：

第一章介绍本文的研究背景和意义，总结了对人手检测和人手姿态识别国内外研究的发展历程和现状，以研究问题和难点出发，从肤色检测模型，人手区域提取，图像分割以及对手势进行识别，分类方面进行了阐述，并对论文的主要内容进行了简单介绍。

第二章提出了一种实时、稳定的挥手检测的方法。首先，在第一部分通过结合主流的人脸检测技术，对基于 Haar 特征级联分类器方法和基于标准化像素差异的 Normalized Pixel Difference (NPD)人脸检测方法进行测试，以提取肤色特征和空间关系约束条件，来估计人手空间分布和肤色特点，为对后续进行肤色分割处理准备；其次，利用图像分割中的基于区域的分割方法-分水岭算法，提取大面积连通的肤色区域，之后通过 8 向种子生长算法对候选区域进行填充，以提取和标记更加准确的分割结果，同时在准确分割结果的基础上，通过几何法结合基于人手轮廓采样点的曲率变化，筛选、并去掉仅利用曲率方法和径向方法产生的误检点，以准确提取到各个指尖点的位置；最后通过对挥手动作的具体分析，在上述工作的基础上，提出了一种简单，高效的挥手动作识别算法，实现了在实际环境中对挥手动作的识别工作。本文提出了解决挥手检测系统中关键问题的一系列方法：

1. 基于脸部的颜色特征和空间特征对手区域特征进行预测的方法。即使肤色在一些颜色空间比较容易通过阈值进行分割，但对于不同光照和人种，这种阈值分割法通常会产生错误的结果。即使通过在大量肤色样本进行机器学习的方法，在面对个例特定情况时，也会存在错误的分割结果。然而，由于人脸特征相对稳定，主流研究算法已经实现较好的稳定和实时性，本文提出一种基于人脸的颜色和空间特征对手特征进行预测。
2. 一种基于分水岭算法的人手区域分割方法。首先，在对人手区域有较好预测的基础上，可以得到一个粗糙的人手二值分割图。之后，通过对该二值图进行一系列的形态学变化，得到标记了前景（人手）和背景（非

人手) 的标记图。该标记图可以使得分水岭算法有效避免过分割效果。通过 8 向种子生长法对分割出的前景区域进行填充, 得到尺寸和肤色像素占比等信息, 用以判定该区域是否为人手区域。

3. 基于曲率和指尖几何关系的指尖检测方法和简单, 稳定的挥手检测方法。在对人手轮廓描述较好的情况, 系统通过曲率变化和各个指尖点与曲率变化极值点之间的几何关系的分析, 有效地筛选和预测出正确的指尖位置。挥手检测方法通过人手位置的周期变化特点, 利用类似于背景更新的滤波算法, 对是否存在挥手动作进行判定。

第三章则是介绍了系统的实现和对真实环境中实验数据的分析。首先对系统工作平台 Ubuntu 下的 ROS 系统进行了简单介绍, 对图片格式信息转化和信息通讯机制进行了说明; 并对实验设计中的数据采集, 以及量化评估方法进行了介绍, 在此基础上, 分别针对系统中肤色分割、人手区域提取、指尖检测以及挥手检测每部分进行了效果和运行速度的分析, 并在采集的数据集上进行了正确率、召回率、各种特定错误检测情况和误判数据的统计。

第四章对本文进行了总结, 并对今后进一步的研究方向进行了讨论。

第二章 挥手检测算法的设计

第一节 基于脸部信息的人手检测方法

本章基于图像中的检测物体的区域分布和颜色特征,并结合统计的方法在多帧的图片中实现自顶向下,任务驱动的挥手动作识别。在现实生活中,挥手动作发生时,每个人人手的形状和挥动频率都会有所不同,而且在不同光照条件下以及复杂的背景环境(含有与肤色相似颜色的物体),这些因素都会对实验结果造成很大影响。然而,人类进行挥手动作时,其人手位置和人脸存在一定的几何关系,在尺寸和颜色特征上也存在相关性。基于以下观察,本章提出了一种基于区域和统计特性的人手显著性检测方法:

1. 人脸检测技术已相当成熟,在不同光照和肤色条件下,也能实现实时,鲁棒性高的检测效果。
2. 人类在进行挥手动作时,人手与脸部在区域分布,颜色和尺度等多种特征上存在相关性,利用这些先验知识,可以大大减少算法复杂性,提高检测速度和效率。
3. 人类在进行挥手示意时,通常会正面朝向示意对象。
4. 为了实现实际场景应用,检测算法应该具有简单快速的特点。

2.1.1 人脸检测

由于本文的方法是以人脸与挥手动作在图像中的空间关系为基础设计的,因此,鲁棒、快速的人脸检测方法是至关重要的。在实际实验过程中,对基于 Haar 特征的级联分类器的经典方法^[23]和基于标准化像素差异的 NPD 人脸检测方法^[24]进行了实验。在非约束环境下,基于 Haar 特征的级联分类器效果很差。所谓非约束环境,是相对于约束情况下的人脸数目单一,背景简单,正面直立的情况而言的,这种情况下,人脸尺度多变,数量多变,姿势多样,遮挡口鼻,表情夸张,分辨率低下。在开放的人脸检测非约束评测集 FDDB 上,基于 Haar 特征的级联分类器方法达到了真正例率 65.93%的误检率,而 NPD 人脸检测方法达到了真正

例率 82.79%的误检率。

通过在实际环境中进行测试，两种方法都能满足实时检测的要求，基于 Haar 的级联分类器方法在 OpenCV 中可以直接调用，通过调整参数实现很好的检测效果；基于 NPD 的人脸检测方法在非正面的人脸图像上拥有更好的效果。

2.1.2 颜色采样与区域估计

1) 基于人脸检测的颜色采样

通过肤色模型进行人手检测是一种常用的普遍方法。这种方法通过对色彩空间的变换，使得人类肤色特征在某一色彩空间如内呈现良好的聚类特性，使得可以利用简单的代数解析式来对肤色区域进行分割处理。然而，由于各种不同的光照条件和人种肤色的多样性，基于特定阈值的分割方法不仅可能无法正确检测到肤色区域，甚至会将人物背景中具有相似颜色的区域也作为分割结果。

但是，即使在不同肤色和光照的条件下，现有的人脸检测方法依然可以实现鲁棒的检测效果。并且注意到在同一场景下，如图 2.1，人物脸部与人手的显著性颜色特征基本吻合；并且，一张图片中，人手与人脸必然在图片中存在一定的空间关系。因此，基于人脸的颜色和空间特征来进行人手的检测应该会实现更好的效果。

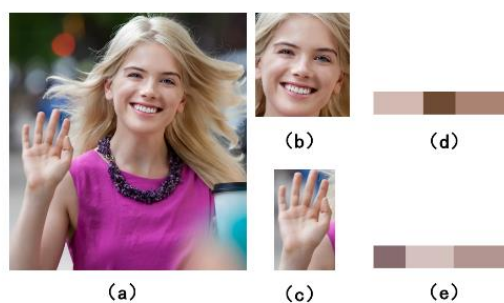


图 2.1 正常背景下挥手动作脸部与人手主要颜色提取结果

为了在人脸区域提取出更好表征颜色特征，通过对人脸的颜色采样，提取出显著的 RGB 颜色特征，并将其转化到 YC_bC_r 色彩空间，通过分割通道，提取到基于 C_b 和 C_r 两个通道的特征。 YC_bC_r 色彩空间中，Y 代表的是亮度分量， C_b 代表蓝色色度分量， C_r 代表红色色度分量。在数字域中，RGB 色彩空间到 YC_bC_r

色彩空间的转换关系可以表示为:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.418 & 0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.1)$$

基于多次测试实验,若 C_r 分量初始化设为阈值在 136~170 之间, C_b 分量初始化设为阈值在 70-127 之间时,如式 2.2,将符合条件区域的像素值置为 1,其他范围置为 0,检测效果较好。

$$P(x, y) = \begin{cases} 1, & Y(x, y) \in [40, 140] \cap C_r(x, y) \in [136, 170] \cap C_b(x, y) \in [70, 127] \\ 0, & otherwise \end{cases} \quad (2.2)$$

这种分割得到的图像中白色区域为肤色区域,黑色为背景。之后,通过实际采样的颜色信息,对初始化阈值进行不断更正,始终使得检测到人脸部肤色的均值位于阈值范围的中值位置。

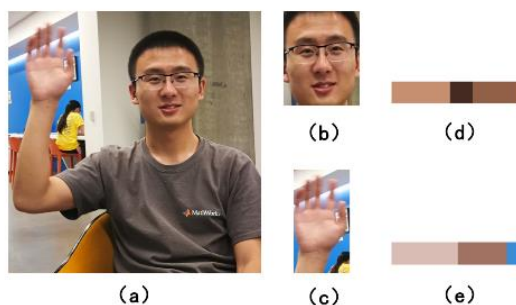


图 2.2 特殊背景下挥手动作脸部与人手主要颜色提取结果

由于如图 2.2 在特殊背景下,脸部和人手的主要颜色都可能受到背景颜色的影响。因此,在对初始阈值的调整上,只允许对阈值范围进行小比例的缩小,避免了将背景颜色作为显著性颜色判定依据。并且,在前几帧中对脸部显著性颜色进行提取采样,用于进行后续的阈值分割处理。

2) 人手位置的区域估计

另外,考虑到在招手动作中,人手和人脸在图像中必定存在一定的空间关系。通过对两者相对空间关系的估计,可以实现更好的检测效果。在对图像显著性研究的文献^[25]中,Cheng 提出了一种基于区域的空间关系的显著性物体检测方法,该方法利用区域间的空间距离使得某一区域的显著性得到整体的突出。相对于基于像素直方图检测的方法,效果得到了明显的提升。同时,文献^[26]中指出对于基

于区域的显著性物体检测方法可以有效提高检测精度和正确性。

基于上述观察,本文通过采集大量含有挥手动作的图片和视频数据进行颜色分割处理,得到基于肤色分割出的二值图。通过对所有图像中的人脸区域和人手相关空间位置的统计,得到其二者之间的相互关系。通过引入空间关系信息,可以有效增加区域的空间位置对人手检测结果的影响。特别的,对于任意轮廓区域 r_k , 仅基于空间关系,可能作为人手区域的显著性定义为:

$$S(r_k) = \log\left(\frac{L - D_s(r_k, r_1)}{\sigma_1}\right) + \beta_1 \sin \theta_1 + \log\left(\frac{L - D_s(r_k, r_2)}{\sigma_2}\right) + \beta_2 \sin \theta_2 \quad (2.3)$$

其中, r_1 和 r_2 分别是脸部的中心点在横轴方向左, 右各 2 倍脸宽的像素点;

$D_s(r_k, r_1)$ 和 $D_s(r_k, r_2)$ 分别表示区域 r_k 的中心点与 r_1 和 r_2 的空间距离; L 表示 2.5 倍的脸部宽度; σ_1 和 σ_2 用来控制空间关系的权值响度, 实验中取值均为 10; θ_1 和 θ_2 分别表示区域 r_k 中心点与 r_1 和 r_2 的顺时针夹角, β_1 和 β_2 用来控制空间关系的权值响度。

此关系式可以用于在视频或者单个图像中,对于挥手动作的检测。具体来说,在分割的二值图像 I 中,越靠近统计人手与人脸相对位置关系的轮廓区域, $S(r_k)$ 值越大, r_k 越有可能是人手目标区域。在实际实验中,对于归一化后的 $S(r_k)$ 值,我们只考虑 $S(r_k) > 0.4$ 的区域进行检测。

第二节 人手区域显著性分割

2.2.1 人手整体分割

通过修正后的颜色特征进行阈值分割,得到不同的分割结果。之后,对多个阈值分割结果进行求和操作,再通过中值滤波操作,以及一定的腐蚀和膨胀操作进行组合形成开运算,用来消除噪音、在纤细处分离不同区域、平滑较大区域的边界的同时并不明显改变其面积。

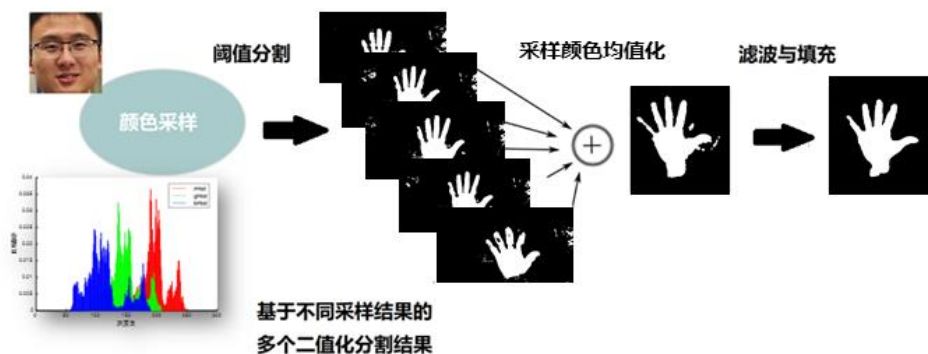


图 2.3 人手显著性分割步骤

由于在检测过程中存在强光或佩戴物遮挡等不利条件，检测结果可能被割裂成为多个区域。为了改善分割效果，Vincent^[16]等人提出的基于区域的形态学图像分割方法“快速分水岭分割算法”，但直接应用该方法会存在过分割现象，使用滤波对图像平滑预处理可以在一定程度上减少细节和噪声对分割结果的干扰。但是通过滤波平滑处理会使得图像变得模糊，导致图像细节缺失，分割的准确性降低。文献^[18]指出通过控制标记种子点数量的方法可以有效防止过分割。如图 2.4 中所示，本文通过以肤色分割结果为基础，作为人手区域检测的先验知识，来实现控制标记种子点的数量。应用分水岭算法求出大块肤色区域的边缘轮廓部分，再利用 8 向种子算法对边缘轮廓内部进行填充处理。

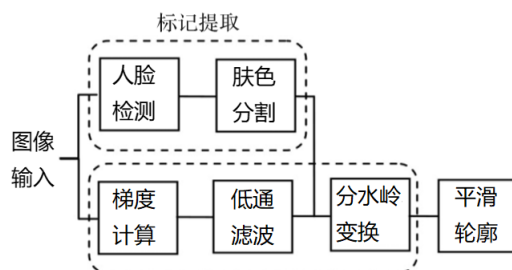


图 2.4 基于标记的分水岭算法

分水岭算法作为一种基于拓扑理论的形态学分割方法，其通过将图像作为测地学上的拓扑地貌，以图像中各点的像素值作为海拔高度的表示，通过从不同区域的低谷涌出的泉水淹没相邻区域，最终泉水所波及的范围为集水盆地，不同水流相遇的界限为分水岭。然而一般的分水岭算法会对微弱边缘、图像中的噪声、物体表面细微的灰度变化造成过度的分割。如图 2.5 所示，本文使用了改进的标

记分水岭算法，它使用预定义的标记来引导对图像的分割。标记利用先前人手二值分割的结果生成。具体来讲，对二值图像通过腐蚀和滤波操作移除噪声和微小物体得到前景，通过膨胀和阈值处理得到背景。标记图像则是对前景和背景和融合。通过这种基于肤色二值分割标记的分水岭分割算法，可以有效刻画出图像中人手轮廓。

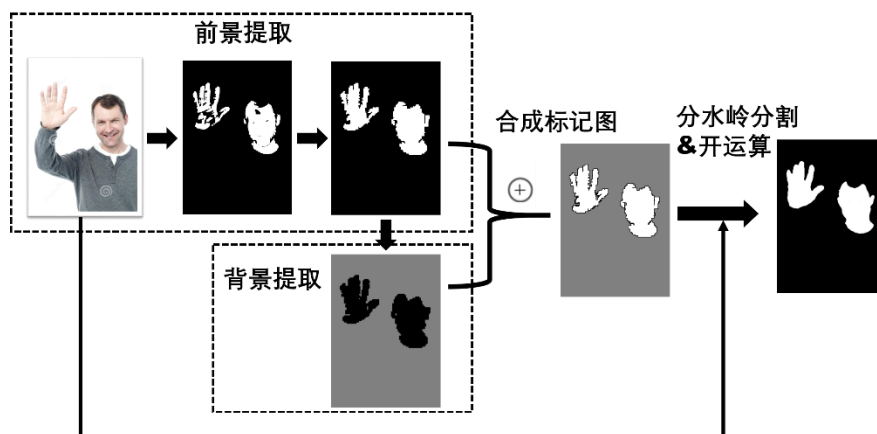


图 2.5 实际运行步骤

为了正确提取到人手区域，通过 8 向种子填充算法，对轮廓内部区域进行标记，以得到该区域像素数目和其中具有肤色特点像素的占比信息。

同时，为了正确检测到人手区域，排除其他皮肤区域如脸部，采用以下约束：

1. 当存在多个候选区域时，优先考虑最大面积区域；
2. 若候选区域内与脸部区域重合，则抛弃该区域；
3. 对在候选矩形区域内肤色面积的比例进行统计，当其小于 65% 时，认为此轮廓包围区域为人手。

通过对人手轮廓的凸点检测，提取包围轮廓的凸包，完成人手的显著性检测。

2.2.2 指尖检测

为了进一步细化分割工作，并为精细化的动作进行识别和分析，本文也进行了多指尖检测的工作。早期的指尖检测系统，主要依赖于特定的输入设备，如电子手套或者特定颜色的指套，但这些方法需要借助外界设备来进行检测，不符合自然条件下的交互场景。近些年，基于计算机视觉理论的发展，对裸手指尖检测的方法愈来愈流行。Lee 等人^[27]提出通过轮廓曲率和距离相结合的方法进行指尖

的检测和筛选，但最终曲率检测结果仍然存在误判。李博男和林凡结合物理学中“场”和“势”的概念，通过计算指尖点的势来识别其位置。一些研究者通过结合深度图像信息，计算掌心于轮廓边缘距离的方法来进行指尖检测。文献^[28]中提出利用手指的径向对称特征来检测，但该方法在各个手指间距较大时，会对指根节点及逆行误判。通常的方法有基于人手的重心距离法和基于轮廓的曲率几何分析法。本文提出了一种结合基于人手轮廓的曲率和几何特征的指尖检测方法。

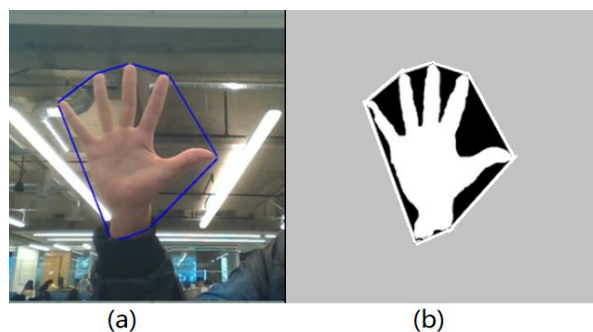


图 2.6 人手凸包图示

如图 2.6 中所示，(a)为 RGB 图像凸包，(b)为将肤色分割处理后的示意图。在图 2.6(a)中，蓝色的轮廓线为人手区域的凸包；在图 2.6 (b)中，凸包与手掌之间的黑色区域均为凸包缺陷区域。每个凸包缺陷区域有四个特征量：起始点，结束点，距离凸包最远点，最远点到凸包的距离。

```

struct CvConvexityDefect
{
    CvPoint* start;           // 凸包缺陷区域起始点
    CvPoint* end;           // 凸包缺陷区域终止点
    CvPoint* depth_point;    // 凸包缺陷区域距离凸包边缘最远点
    float depth;            // 最远点与凸包边缘的距离
};

```

如果可以正确地检测到五指指尖内四个凸包缺陷区域的最远点，就相当于确定了五个指尖的位置。然而，如图 2.8(c)中所示，由于噪声和人手的腕关节结构使得存在多个凸包缺陷区域，使得最远点检测结果并不理想。

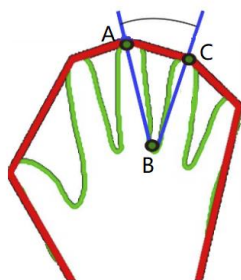


图 2.7 人手指尖检测理论示意图

为了过滤掉不相关的凸包缺陷区域最远点，本文介绍的方法利用正确指间点与正确凸包缺陷区域最远点间的几何关系进行筛选。如图 2.7 所示，点 A、点 B 与点 C 之间构成了一个锐角 α ，并且点 A 或点 C 与点 B 之间的距离 l 也是有一定的长度限制的。设人手检测区域高度为 H ，则锐角 α 与距离 l 应满足以下关系：

$$\alpha < 80^\circ, l > 0.4H \quad (2.4)$$

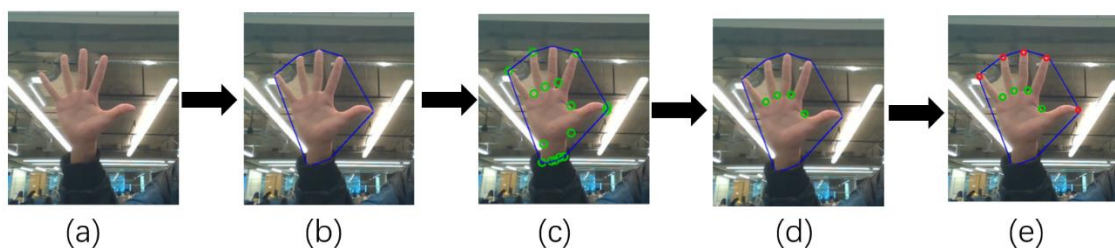


图 2.8 人手指尖检测步骤

第三节 挥手动作识别

挥手这一手势是不论年龄大小的人都能够做的一个简单动作。这是一个友好的、快乐的手势。人们通常挥手来打招呼和道别，或者在繁忙的餐厅里来引起服务员注意，需要他们提供服务。作为一种简单的动态手势，挥手动作具有强烈的周期性特征，算法在已经确定了人手区域的条件下，再去验证是否存在周期变化的运动来确定挥动手势是否存在。

在确定对挥手动作识别的算法之前，我们首先要对挥手动作进行一定的规律总结。微软公司在深度摄像机 Kinect 中实现利用人体骨骼检测对人体姿态的识别，其中就包含对挥手手势的识别。在 XBOX 体感游戏系统中将挥手动作定义

为：从胳膊开始到肘部弯曲。用户以胳膊肘为焦点来回移动前臂，移动平面和肩部在一个平面上，并且胳膊和地面保持平行，在手势的中部，前臂垂直于后臂和地面。然而，以这样的定义来进行检测，不仅要求算法对身体其他部分，如肘关节点，进行检测，也使得手势姿态检测系统的灵活性降低。本文作者认为，在对某一手势进行识别时，应该采取尽可能多的样本，然后试图标准化这一手势，更关注于手势的核心部分，而不是外在的细枝末节。因此，在本文中，将挥手动作定义为：人手在五指朝上状态下，与头部的距离在单位时间内发生周期变化的次数大于某一阈值，且能维持一定时间。算法通过对计算人手挥动离头部距离在时间序列上变化的特征来判断视频中是否存在挥手动作。

使用长滤波器(2.5)和短滤波器(2.6)是检测人手运动区域的关键，定义为

$$L(x, y) = (1 - w_L)L(x, y) + w_L F^t(x, y) \quad (2.5)$$

$$S(x, y) = (1 - w_S)S(x, y) + w_S F^t(x, y) \quad (2.6)$$

式中， $F^t(x, y)$ 为当前帧， $L(x, y)$ 为长滤波器输出结果， $S(x, y)$ 为短滤波器输出结果， w_L 和 w_S 分别为长、短滤波器对应权值。在本文实际实验环境下， w_L 取值为0.1，这一取值取决于测试者的挥手频率，频率越快， w_L 应取值越大，反之则较小。 w_S 取值为0.9，这一取值主要用于去除噪声并防止图片虚化。在合适的权值条件下，长滤波器的输出结果会与手本身自身挥动频率像协调，使得背景图片像素与人手像素相互混合。

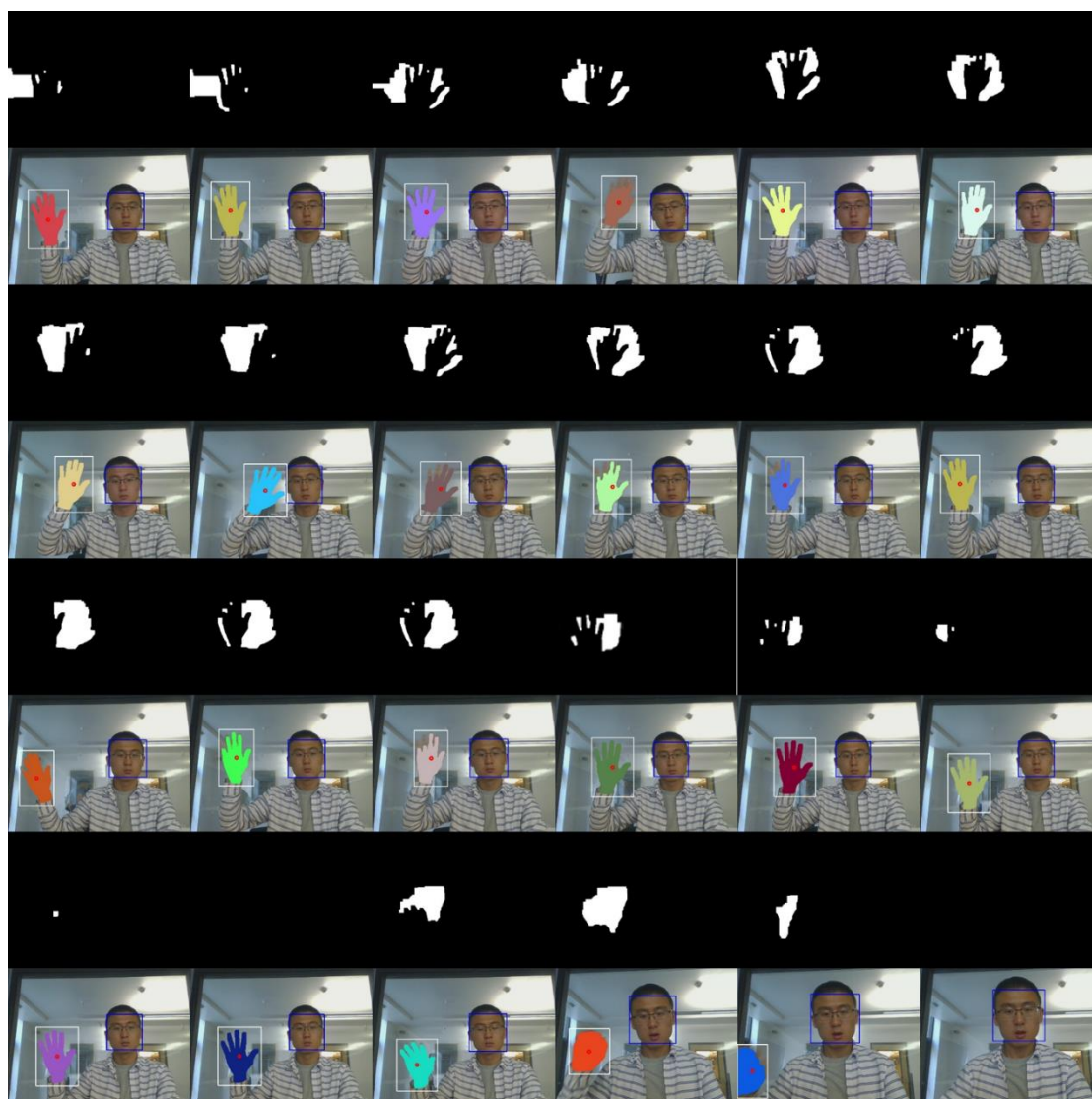


图 2.9 挥手动作发生的视频序列处理结果

如图 2.9，这一方法应用于灰度分割图时，长滤波器结果则是对所有人手经过的区域进行了分割后的结果，短滤波器则是基于当前帧的人手分割结果。通过求取长滤波器结果中区域的质点，并统计在后续视频帧的人手区域经过该点的次数来判断是否存在周期性挥手动作。在本文实验中，认为周期性运动时间大于 2 秒且次数大于 3 次，挥手动作形成。

第四节 本章小结

在本章第一节中，针对人脸检测，介绍了基于 Haar 特征的级联分类器和基于 NDP 特征的检测方法，由于挥手状态下人脸为正面朝向，考虑到稳定，快速，方便参数调节的特点，实际系统中我们采用了基于 Haar 特征的级联分类器的经

典方法。在图像中检测到人脸的基础上，通过对人脸部分的颜色直方图采样来进行初始化颜色描述。之后，通过脸部和挥手动作用在图像中的空间关系，确定挥手动作用在图片中可能出现的区域。在这一区域内，通过多个视频帧的采样颜色来进行阈值分割生成多个二值图像。再对这些二值图像进行均值融合，减少噪声影响。在本章第二节中，讨论了如何对二值肤色分割图像进行处理，以得到平滑和无噪的二值检测结果。首先，利用中值滤波与一系列膨胀和腐蚀操作来进行预处理，再通过图像分割分水岭算法和 8 向种子填充法进一步对分割结果进行优化。得到了更好的人手分割结果。在这一基础上，我们通过基于轮廓采样点的曲率变化关系和指尖与手掌的几何关系，判断出了指尖的位置点，有利于后续人手姿态的识别。第三节中，在之前工作基础上，通过对挥手手势的核心部分进行分析和总结，使用了一种基于长短滤波器的方法来确定挥手动作的运动范围，并对次数，频率进行统计来确定挥手动作是否存在。

第三章 实验结果与分析

第一节 实验平台介绍

为验证本文的方法，并与其它机器人系统模块进行通信，系统采用基于 Ubuntu14.04 系统下机器人操作系统 ROS Indigo 平台。实验中，PC 配置为 Intel® Xeon® E3-1505M v5 Processor，内存为 4G。ROS 是一个开源操作系统，应用于机器人控制的一种次级操作系统。通过提供了类似传统操作系统的功能，其中包括底层设备控制、进程间消息的传递和程序包管理等。

在 ROS Indigo 平台中使用 Intel 公司的开源计算机视觉库 OpenCV 的一个关键问题是：ROS 环境下图像格式为 sensor_msgs 类型，而 OpenCV 处理的图像格式是 Mat 类型。

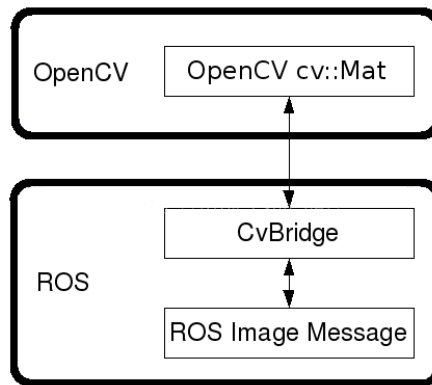


图 3.1 ROS 机器人操作系统图像格式转化示意图

如图 3.1 所示，ROS 提供了这两种图像格式相互转换的软件包 cv_bridge，使得我们可以将 ROS 系统下读取的图像转换成 Mat 格式，再使用 OpenCV 进行图像处理。

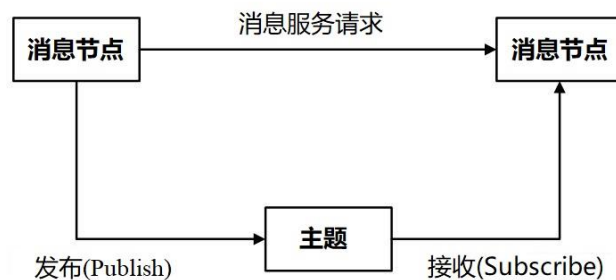


图 3.2 ROS 机器人操作系统消息传递机制示意图

如图 3.2, 在 ROS 系统中, 消息信号的识别以及传递依靠主题(topic)来进行订阅 (subscribe) 和发布 (publish)。对于本文应用系统的挥手信号, 程序首先要对摄像头信号进行订阅, 以收到 sensor_msgs/image 类型的图像信息; 当检测出存在挥手动作时, 发出 std_msgs::String 类型的信号:

```
# ros::Publisher waveman;  
#ros::NodeHandle nh;  
#image_transport::ImageTransport it(nh);  
#image_transport::Subscriber sub=it.subscribe("camera/rgb/image_raw",1,imageCallback);  
#ros::Subscriber waveInfo_sub = nh.subscribe("detectWave", 1, waveInfoCallback);  
# waveman = nh.advertise<std_msgs::String>("wave_detect", 1);
```

第二节 人手检测结果

3.2.1 评估标准

在本节中, 分别针对系统对肤色和人手的分割效果进行了分析, 量化评估。针对图像的分割效果的评估, Haralick 和 Shapiro 在文献中提出了四个标准:

1. 对于某些特征 (如灰度、纹理), 同一个区域的图像应该一致和均匀;
2. 区域内部应该简单, 没有很多空洞;
3. 相邻的区域在满足区域内部一致性的特征上应该有显著的区别;
4. 每个区域的边界应该简单而不粗糙, 并且空间位置准确;

这种主观上针对图像分割效果的判定虽然针对少量图片在一定情况下有效, 但针对大量数据集的分割结果来说, 并无法做到准确的判断。因此, 为了量化的比较和分析图像分割的结果, 本文在分割结果与真值结果比较的基础上, 利用正确率和召回率进行量化分析。

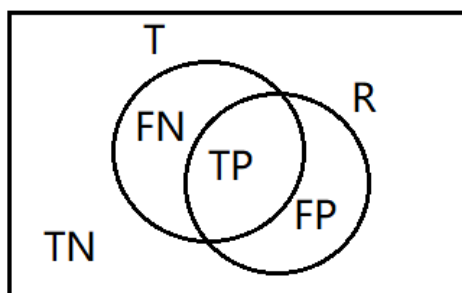


图 3.3 图像分割结果与真值比较示意图

图 3.3 中，T 表示真实值分割结果，R 表示系统分割结果，TN 为背景区域，TP 为 T 与 R 重叠区域，FN 和 FP 分别为 T 和 R 中排除重叠部分的区域。准确率（Precision）是指检测结果中的真值(TP)占所有检测结果(TP+FP)的比例；查全率（Recall）是指检测结果中的真值(TP)占全部真值(TP+FN)的比例。

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3.1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.2)$$

另外，由于在检测过程中，当所有图片区域都判定为目标区域时，召回率便可以达到 100%；并且在进行图片检索中，出现错误的结果比从海量图片中漏掉几个正确的结果要更糟糕，所以对于算法的评估，准确率比召回率的要求更高。因此，F-measure 方法通过调节权重，综合衡量了图像分割结果的准确率和查全率，定义如下：

$$F_{\beta} = \frac{(1 + \beta^2)\text{Precision} \times \text{Recall}}{\beta^2\text{Precision} + \text{Recall}} \quad (3.3)$$

其中，通常被设为 0.3 来强调准确率的重要性。

3.2.2 肤色分割效果

为了较为客观的实际验证算法的鲁棒性，通过从采集互联网图片和现实生活中采集到的图片数据，建立了一个包含 200 张图片的测试数据集。整个数据集涵盖了不同肤色，光照条件，位置，姿态的人物形象。换句话说，这些人物形象没有任何先验知识或者约束。在整个实验过程中，没有进行参数的调整。

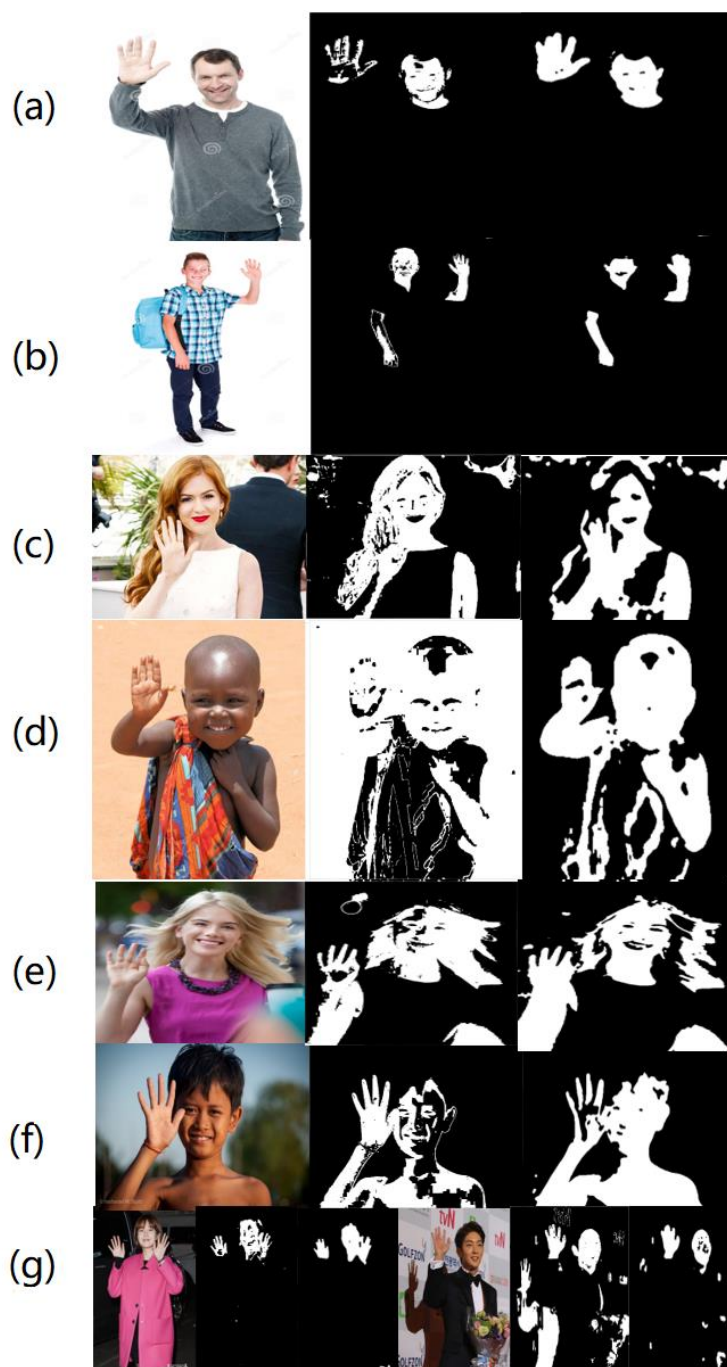


图 3.4 部分肤色检测结果实例

图 3.4 中，(a)、(b)是简单背景下的肤色检测结果，(c)、(d)对比不同肤色的检测结果，(e)、(f)和(g)表示复杂场景下的肤色检测结果。从左到右分别为输入图像，固定阈值检测，基于人脸颜色特征的检测结果



图 3.5 部分肤色检测结果实例

图 3.5 中, (a)是阈值处理下的检测结果、(b)是基于脸部颜色特征的肤色检测结果, (c)是通过在(b)结果基础上利用空间关系和形态学去除噪声的结果。图 3.4 与图 3.5 表明, 在相同条件下, 基于人脸颜色采样的肤色分割方法可以有效提高肤色检测效果, 在不同光照和肤色条件下也可以得到较好的输出结果。

3.2.3 人手分割效果

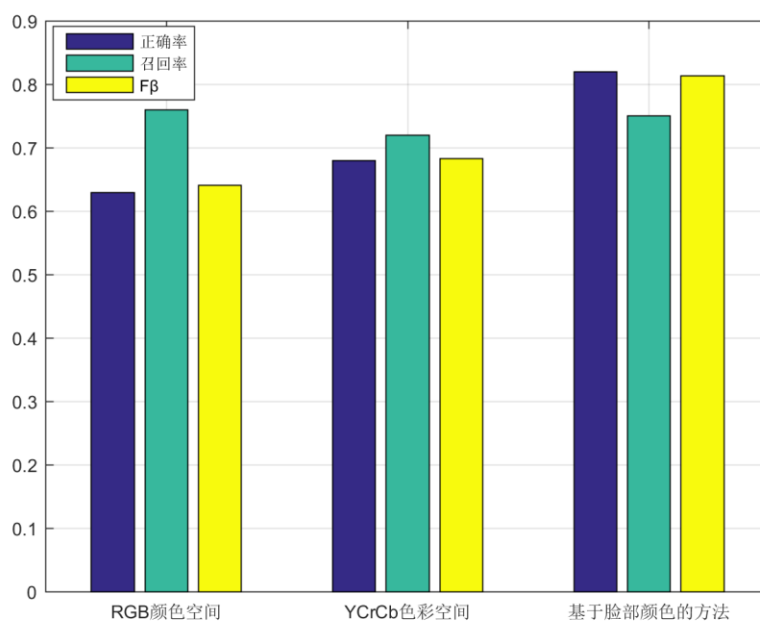


图 3.6 网络图片中的人手检测结果实例

同时, 为了对本章的算法进行更为客观的评价, 作者以 Mittal 等人建立的人手检测数据集为基础, 对其中 500 张图像进行了测试, 原始图像大小约为 512×375 , 24 位色彩位数编码, 并对测试速度进行了计算, 如表 3.1 所示。

表 3.1 不同像素大小图像

图像像素大小(宽×高)	检测耗时 (ms)
1280×720	1200
1152×648	912
1024×576	685
896×503	540
768×432	392
512×375	168
256×144	32

图 3.7 各种方法的正确率、召回率以及 F_{β} 度量的比较

为了更好得评估对视频格式中挥手动作人手的检测，并且由于对视频格式数据进行人手真实值标定的困难，本文采用了一种任务导向的评判方法。评判数据集包含 40 个 avi 格式视频文件，每个视频包含约 30,000 帧的不同背景下的挥手动作，视频分辨率为 640×480 像素。评判方法通过分析系统分割结果来对评判效果。一个好的分割结果应该是有利于人手轮廓和细节特征提取的。换句话说，一个好的人手检测方法应该有利于后续的手势识别工作。



图 3.8 部分肤色检测结果实例

具体来说，在每个视频中，对于分割结果的评判采用抽取样本帧的方法，随机抽取每个视频中 100 帧，若其中手掌和手指的形状被较好地描述，则认为系统输出成功。相反的，系统输入失败，当出现下列情况：不完整地检测出手掌区域（错误 1）；不完整的检测到手指（错误 2），错误检测到背景物体为手指（错误 3）。同时，由于在检测过程中，即使没有完整检测到人手，但依然准确识别了人手区域的情况，本文引入另外一个评估指标：对人手区域检测判定的准确性。

表 3.2 算法性能比较

视频序列	检测成功率	错误 1 发生率	错误 2 发生率	错误 3 发生率	区域判定准确率
1	97%	2%	3%	2%	98%
2	95%	3%	5%	3%	97%
3	85%	5%	9%	3%	92%

第三节 指尖定位结果

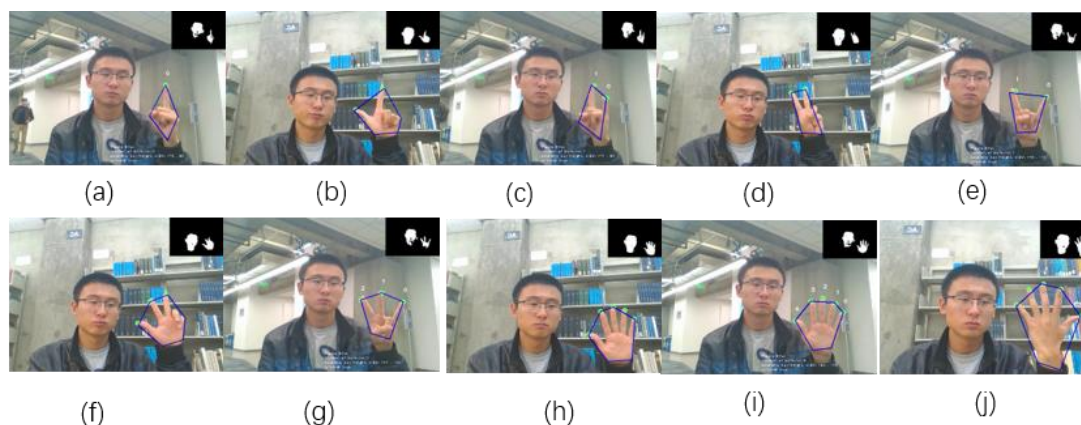


图 3.9 指尖定位检测结果

为了验证本文的指尖检测方法，针对常见的数目为 0 至 5 的六类指尖手势进行了测试，得到了一系列的检测结果，如图 3.6 所示。在实验过程中，手在镜头前左右晃动和旋转，但对每类手势，指尖数目不发生变化。针对每一类手势的视频序列，记录了 1000 帧，每隔 5 帧对画面进行提取，则可以提取 200 帧。

表 3.3 不同类别下的指尖检测结果

指尖数目	帧数	成功次数	误触发次数
0	200	198	2
1	200	199	1
2	200	200	0
3	200	197	3
4	200	199	1
5	200	197	3

另外，本文方法也同国内外一些主流的指尖检测算法^[27,28]进行了比较和分析。文献^[27]利用曲率与与手心质点距离的特性来进行指尖检测，但该方法容易在较短手指上发生误检。文献^[28]利用手指的径向特征对每个手指的进行检测，但容易将指根节点进行误判。而本文算法采用了基于曲率的人手指尖几何关系的方法，对检测结果进行了筛选和判定，在复杂条件下也能稳定地检测出各个指尖的位置。实验结果表明了该方法的鲁棒性。

表 3.4 算法性能比较

指尖检测方法	检测成功率	误检率	平均耗时
本文算法	98%	2%	0.072
基于曲率和距离	97%	2.5%	0.070
基于径向特征	95%	3%	0.08

第三节 挥手识别结果

为了评价本文挥手识别方法的性能，采用了不同环境，和对象对系统进行了测试，表 2 为实验结果。测试过程为：与采样摄像头分别距 0.5 米、1 米、1.5 米、2 米、3 米内进行测试，每次测试时间为 1 分钟，在挥手动作发生 4 秒内发出识别到挥手动作信号为成功，超出 4 秒时间为失败，若实际无挥手动作发生，但系统发出识别到挥手动作信号为误触发。

表 3.5 不同距离与场景下的挥手识别结果

场景条件	距采样镜头 距离(米)	识别成 功率	误触发 次数	响应耗时 (秒)
室内-单人	0.5	100%	0	2
	1	95%	0	2
	1.5	95%	0	3
	2	90%	0	3
	3	80%	0	3
室内-多人	0.5	100%	0	2
	1	90%	2	2
	1.5	85%	4	2
	2	80%	6	3
	3	70%	6	3

室外-单人	0.5	100%	0	2
	1	100%	0	2
	1.5	95%	0	3
	2	85%	1	4
	3	80%	2	5
室外-多人	0.5	100%	2	2
	1	90%	2	2
	1.5	85%	2	3
	2	70%	4	4
	3	70%	6	5

从测试统计结果表 3.5 中可以看出，系统的实际检测效果与距离，和环境背景有关。在距离小于 2.5 米的情况下，检测成功率和准确率较好。因为在系统中通过肤色检测和一系列的形态学处理后，对于近距离的脸部信息可以提取到更准确的信息，有利于肤色分割和空间关系的确定。而在复杂背景的条件下，容易产生误触发，是因为复杂环境下，可能检测到背景中的多个肤色区域，难以对其与目标区域进行区分。

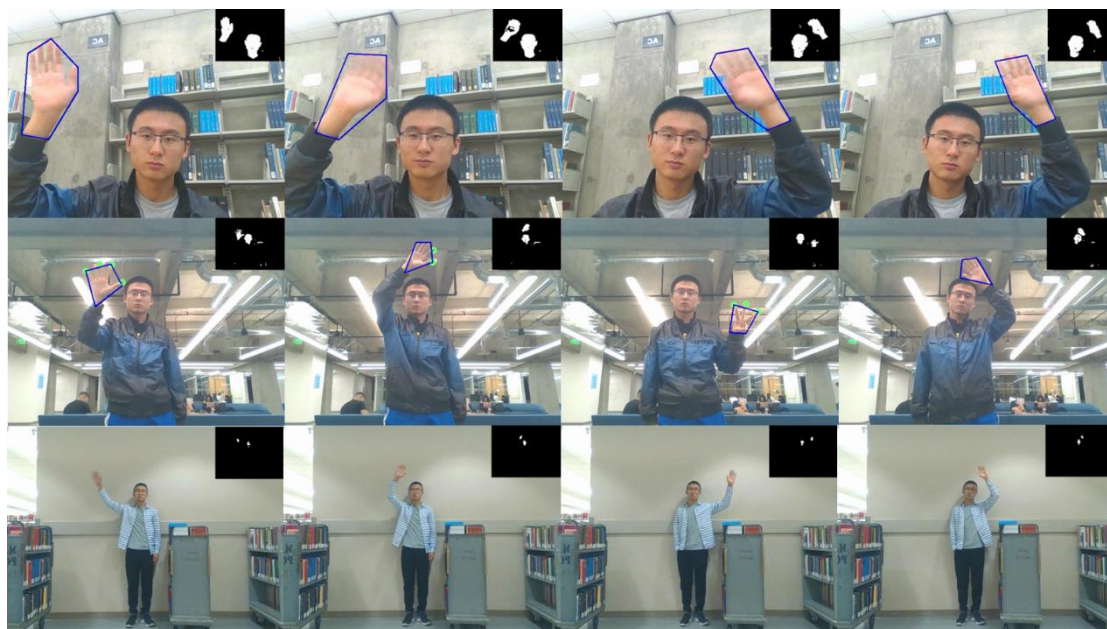


图 3.10 距离为 1 米(第一列)、2 米(第二列)、3 米(第三列)的挥手动作识别实例

第四节 本章小结

本章对实时的肤色分割方法,以及人手检测算法进行了检测速度和准确性的评估,并且在权威检测数据集上进行了测试实验。并将挥手检测算法在不同实际环境中进行了测试,进行了量化分析。实验结果表明,本文算法可以实现实时检测,整个系统的检测速度已经达到每帧 20 毫秒左右。对于应用于实际环境的挥手检测系统,也可以考虑对目标场景进行多帧分时检测,如每隔 5 帧检测人手位置变化情况。

第四章 总结与展望

第一节 论文总结

本文对基于肤色分割的挥手动作识别技术进行了相关研究。在本次实验中，实验主要以下几个阶段的工作：

第一阶段是通过人脸颜色和空间关系特征，实现人手的显著性区域提取。本次实验使用了基于使用了基于 Intel 公司的开源计算机视觉库 OpenCV 实现的色彩空间转化和形态学处理，对于图像中肤色部分进行了快速，高效地分割。但由于仅仅通过分割方法无法确定人手具体位置，本文通过空间关系和像素点占比等限制条件，确定候选人手区域。进一步，通过利用人手轮廓上各个采样点曲率的变化关系，以及几何特征，对人手的指尖区域进行了定位，由此完成了对人手的精细化分割，并实现了手指数目以及位置的确定，为后续手势识别工作打下了坚实基础。

第二阶段则是对挥手动作的识别和判定。本阶段分为两个部分。第一部分是对视频中前后帧进行长滤波和短滤波处理，为了使滤波结果更符合实际系统需求，需要对挥手动作的频率进行估计，已完成滤波参数的调节。在取得第一部分滤波结果的基础上，第二部分则是进行人手移动区域的周期性检测。首先通过上一步长滤波器得到的数据确定距头部最远和最近位置，以确定中间点。用短滤波结果与中间点位置进行处理分析，确定人手运动状态，最后对是否发生挥手动作进行判定。

第三阶段则是实验平台介绍、数据的采集与分析。本阶段分为三个部分。第一部分是简单介绍了机器人操作系统中图像格式转化，以及系统中消息传递机制。第二部分则是进行人手区域检测方法的分析。首先通过采集互联网图像数据，在各个场景和肤色情况下于传统算法结果进行对比。然后，在标准人手检测数据集上进行测试，并对算法处理速度进行统计。第三部分则是通过采集日常生活中各类场景下，在不同距离的条件下，对系统进行成功率、误触发率和识别耗时特性的分析，检验挥手识别系统的实际应用效果。

第二节 研究方向展望

在本次实验中，在对人手特征的提取上依赖于鲁棒、快速的人脸检测算法，相对于传统阈值肤色分割，这种方法提供了较好的肤色检测结果；但这种方法也对系统性能和应用场景进行了限制。对于下一步的工作，可以考虑利用卷积神经网络对手掌进行三维姿态估计，建立人手骨架，通过仿射投影结果确定手指状态。

除此之外，也可以结合目标场景的深度信息来进行人手的检测和手势识别工作。深度传感器，如微软 Kinect 深度传感器、双目摄像头、华硕 Xtion Pro 深度传感器等，可以产生连续的深度图像。深度信息的获取进一步增加定位的准确性，增强了算法在不同场景中的适应能力。对于 2D 图像中难以解决的重叠、光线变化、快速移动以及分割类肤色物体等问题上，深度信息都可以提供更方便和直观的解决方案。进一步来讲，将系统搭载在机器人系统上，与其他传感器的信息进行交流和融合，得到更好的人机交互体验。

参考文献

- [1] Ebert A, Gershon ND, van der Veer GC. Human-computer interaction[J]. KI-Künstliche Intelligenz, 2012, 26(2):121-126.
- [2] Rautaray SS, Agrawal A. Vision based hand gesture recognition for human computer interaction: a survey[J]. Artificial Intelligence Review, 2015, 43(1):1-54.
- [3] Stergiopoulou E, Sgouropoulos K, Nikolaou N, Papamarkos N, Mitianoudis N. Real time hand detection in a complex background[J]. Engineering Applications of Artificial Intelligence, 2014, 1(35):54-70.
- [4] Mo Z, Neumann U. Real-time hand pose recognition using low-resolution depth images[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, New York, USA, 2006, 1499-1505.
- [5] Zhao J, Chen T. An approach to dynamic gesture recognition for real-time interaction[C]. The Sixth International Symposium on Neural Networks, Berlin, May 26-29, 2009, 369-377.
- [6] Yang C, Jang Y, Beh J, Han D, Ko H. Gesture recognition using depth-based hand tracking for contactless controller application[C]. IEEE International Conference on Consumer Electronics (ICCE), Jan 13, 2012, 297-298.
- [7] Zhu X, Jia X, Wong KY. Pixel-level hand detection with shape-aware structured forests[C]. Asian Conference on Computer Vision, Singapore, Nov 1-5, 2014, 64-78.
- [8] 李逢, 桑农, 王洪智, 颜轶, 高常鑫, 刘乐元. 多特征融合的实时人手跟踪算法[J]. 控制理论与应用, 2017, 34(6):797-802.
- [9] Sridhar S, Mueller F, Oulasvirta A, Theobalt C. Fast and robust hand tracking using detection-guided optimization[C]. IEEE Computer Vision and Pattern Recognition (CVPR), Boston, USA, June 8-10, 2015, 3213-3221.
- [10] Ren Z, Yuan J, Zhang Z. Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera[C]. Proceedings of the 19th ACM international conference on Multimedia, Scottsdale, Arizona, USA, Nov 28-Dec 1, 2011, 1093-1096.
- [11] Li Z, Jarvis R. Real time hand gesture recognition using a range camera[C]. Australasian Conference on Robotics and Automation, Sydney, Australia, Dec 2-4, 2009, 21-27.
- [12] Lahamy H. Real-time hand posture recognition using a range camera[D]. University of Calgary, 2013.
- [13] Van den Bergh M, Van Gool L. Combining RGB and ToF cameras for real-time 3D hand gesture interaction[C]. IEEE Workshop on Applications of Computer Vision (WACV), Kona, Hawaii, Jan 5,-7, 2011, 66-70.
- [14] Do NH, Yanai K. Hand detection and tracking in videos for fine-grained action recognition[C]. Asian Conference on Computer Vision, Singapore, Nov 1-5, 2014, 19-34.

-
- [15] Droschel D, Stückler J, Behnke S. Learning to interpret pointing gestures with a time-of-flight camera[C]. The 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Lausanne, Switzerland, Mar 6-9, 2011, 481-488.
- [16] Vincent L, Pierre S. Watershed in digital spaces: An efficient algorithm based on immersion simulations[J]. IEEE Transactions of Pattern Analysis and Machine Intelligence, 1991, 13(6):583-598.
- [17] 余辉, 曲昌盛, 李金航. 复杂环境下的人手轮廓提取方法[J]. 计算机工程与应用, 2015, 0(14):170-174.
- [18] Bieniecki W. Over segmentation avoidance in watershed based algorithms for color images[C]. Proceedings of IEEE International Conference on Modern Problems of Radio Engineering, Telecommunication and Computer Science, Lviv, Ukraine, Feb 24, 2004, 169-172.
- [19] Chen CP, Chen YT, Lee PH, Tsai YP, Lei S. Real-time hand tracking on depth images[C]. Visual Communications and Image Processing (VCIP), Tainan, Taiwan, Nov 6-9, 2011, 1-4.
- [20] 易靖国, 程江华, 库锡树. 视觉手势识别综述[J]. 计算机科学, 2016, 43(6A):103-108.
- [21] Hassan M, Ahmad T, Liaqat N, Farooq A, Ali SA, Hassan SR. A review on human actions recognition using vision based techniques[J]. Journal of Image and Graphics, 2014, 2(1):28-32.
- [22] 谈家谱, 徐文胜. 基于 Kinect 的指尖检测与手势识别方法[J]. 计算机应用, 2015, 35(6): 1795-800.
- [23] Viola P, Jones MJ. Robust real-time face detection[J]. International journal of computer vision, 2004, 57(2):137-154.
- [24] Liao S, Jain AK, Li SZ. A fast and accurate unconstrained face detector[J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 38(2):211-23.
- [25] Cheng MM, Zhang GX, Mitra NJ, Huang X, Hu SM, Global contrast based salient region detection[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, USA, June 20-25, 2011, 409-416
- [26] Jiang H, Wang J, Yuan Z, Wu Y, Zheng N, Li S. Salient object detection: A discriminative regional feature integration approach[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun 23-27, 2013, 2083-2090
- [27] Lee D, Lee S G, Vision-based finger action recognition by angle detection and contour analysis[J]. ETRI Journal, 2011, 33(3): 415-422.
- [28] 梅萍华, 李斌, 朱中的, 等. 基于径向对称变换的实时指尖检测算法[J]. 中国科学技术大学学报, 2011, 41(2): 101-107.

致谢

从学期开始到此刻完成毕业论文的撰写，五个月的时间一闪而过，我的本科生生涯也即将画上句号。回首往昔，感慨顿生，有遗憾，有失落，也有满足和自豪。五个月的毕业设计在我的本科生生涯里画上了浓墨重彩的一笔。在此，衷心感谢杨老师在毕业设计方面给我的悉心指导，杨老师总是能够一针见血地指出我实验中的问题，指引出我实验的正确方向。论文的撰写已接近尾声，但是杨老师严谨开放的治学态度却始终让我铭记于心。

同时，也要感谢大学四年中教给我知识，给予我帮助的老师，是你们见证引导了我的成长，让我不断进步。

感谢我的同学、朋友们，感谢你们陪我度过了这美好四年时光，虽聚散天涯，感谢我们曾经相遇。

感谢我的家人们一直以来对我的支持，在我失落时给我鼓励，在我伤心时给我安慰，你们永远是我最强后盾，给予了我无限前进的力量。

感谢我生命中遇见的所有人，因为有你们的出现使得我的生命如此丰富多彩。