

视觉手势识别综述

易靖国 程江华 库锡树

(国防科技大学电子科学与工程学院 长沙 410073)

摘要 随着计算机视觉的发展和人机交互的需要,手势识别研究取得了蓬勃的发展。然而,很少有对手势识别进行全面分析、总结、评述的相关综述研究。针对这一问题,就基于计算机视觉的手势识别技术的发展现状展开综述,系统地分析和梳理了近三十年来基于计算机视觉的手势识别技术,按照手势识别的流程,将手势识别分为检测分割、分析和识别3个主要步骤,并对每个步骤分别进行评述,总结归纳其适用范围及优缺点。最后对手势识别进行总结展望。

关键词 人机交互,计算机视觉,手势识别,综述

中图法分类号 TP391 文献标识码 A

Review of Gestures Recognition Based on Vision

YI Jing-guo CHENG Jiang-hua KU Xi-shu

(College of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China)

Abstract In recent years, with the development of computer vision and the requirement of human computer interaction, gesture recognition research has made conspicuous progress. However, there are few little review research to gesture recognition with analysis, summarizing and evaluation. In this paper, in connection with summarizing the research and development status of the computer vision gesture recognition, we analyzed several gesture recognition technologies based on compute vision in the past 30 years. The steps of hand gesture recognition were proposed from three main steps: detection and segmentation, analysis and recognition. We summarized their application fields and their advantages and disadvantages, and then made a conclusion and the future scope.

Keywords Human computer interaction, Computer vision, Hand gesture recognition, Review

近年来,随着计算机视觉和机器学习等相关学科的发展,人机交互技术(Human Computer Interaction)已经成为人们日常生活中不可或缺的重要部分。作为计算机科学研究领域中的一个重要组成部分,人机交互的发展历程已历经了半个多世纪,并且取得了很大的进步和提高。相对于传统的二维图形用户界面,人机交互的最终目标是实现人与机器自然地交流,从而为操作者提供更为直观、舒适的交互体验。这些年来开展了大量的相关研究,包括人脸、步态、面部表情、手势以及体态等^[1]交互技术研究。在这些交互信息中,手势具有直观性、自然性和丰富性的特点,其识别研究引起了人们极大的关注。

手势是一种符合人类日常习惯的交互手段,在日常生活中人们之间的交流通常会辅以手势来传达一些信息或表达某种特定的意图,它包括静态手势和动态手势两种类型。最初的手势识别主要是利用机器设备的直接检测来获取人手与各个关节的空间信息,其典型代表设备如数据手套^[2]等。1983年,来自AT&T的Grimes^[3]原创性地发明了最早的数据手套;1984年,VPL公司生产的数据手套能够使用光纤传感器检测出手指的弯曲程度;Liang等^[4]利用数据手套识别出台湾手语中多个基本词条,其识别率达到90.5%;1991年,Vir-

tual Technologies公司推出了Cyber Glove,该手套包含了线形弯曲传感器和电子张力变形测量器,能够获得较高的精度和稳定性;吴江琴等^[5]使用Cyber Glove进行中国手语识别,对孤立词的识别率达到90%,对简单语句的识别率则达到了92%;Assam和Grobel^[6]从带颜色手套的实验者的记录视频中提取特征,采用隐马尔科夫模型(Hidden Markov Model, HMM)识别了261个孤立词汇,其正确率达到91.3%。后来,因数据手套受到了手势的自然性和可识别的手势较少的限制,光学标记法诞生并取代了数据手套。该方法也可提供良好的效果,但仍需较为复杂的设备。机器设备的直接检测虽然使得手势识别的准确度和稳定性得到了提高,但却限制了手势的自然表达方式。基于视觉的手势识别系统,相比于穿戴设备手势识别系统,其能够使操作者徒手以更加自然的方式进行人机交互。它是指对视频采集设备拍摄到的包含手势的图像信息传输给计算机,然后通过计算机视觉技术进行处理,进而对手势进行识别。这种方法不但不需要高昂的设备费用,并且在操作时也更加方便。这也是手势识别未来的发展趋势,为此自然场景下基于视觉的手势识别应运而生。本文主要对基于视觉的手势识别技术进行分析和总结。

视觉手势识别(Vision Gesture Recognition)是人机交互

易靖国(1993—),男,硕士生,主要研究方向为计算机视觉与智能信息处理,E-mail:yijingguo-nudt@126.com;程江华(1979—),男,副教授,主要研究方向为视频图像处理及模式识别,E-mail:jianghua-eheng@nudt.edu.cn;库锡树(1963—),男,教授,主要研究方向为电路与系统,E-mail:kuxishu@126.com。

技术研究的一个分支领域,是机器视觉领域的一个重要的研究内容,其发展状况与人机交互技术紧密相连^[7]。成熟的视觉手势识别产品已在世界范围内被广泛应用,典型的消费类电子产品有:2003年,Sony公司推出了一款名为EyeToy的手势识别设备,这种设备能将玩家的动作传输到游戏画面,使玩家互动^[8];2010年11月,微软公司推出的Kinect体感设备在手势跟踪与识别方面有着出色的表现,它能实时识别用户手势,结合Xbox,使用户完成对游戏的控制指令^[9];2012年,三星推出的智能电视ES8000,结合用户的手势,可以对电视进行换台、搜台以及音量调节等操作。

在众多手势识别文献中,尽管前人已对其方法进行过综述,但主要是对发展现状进行总结,且介绍的是所有类型的手势识别方法,尚无专门针对基于视觉的手势识别进行综述的文献。相对于采用机器设备检测的手势识别方法,视觉手势本身具有多样性、多义性以及差异性等特点,加之人手是复杂的变形体以及复杂的背景环境。所以,视觉手势识别成为多学科交叉的研究课题,也是目前以及将来一段时间内计算机视觉研究的热门方向。本文将30多年来的视觉手势识别方法与涉及的相关技术进行系统总结,指出当前存在的问题,展望该领域的发展趋势。

1 国内外研究现状

在20世纪90年代初,就开始了视觉手势识别的研究。付永刚^[10]采用两个摄像机实现了一个双手交互的Video Desk系统,系统将采集的视频图像与预先定义好的捏取、指点及拖拽等常用手势进行对比匹配来进行手势识别。1991年,日本富士通公司完成了对46个手势符号的识别工作^[11]。HyeSun^[12]建立了手势的六维特征向量,并利用整合的HMM识别出了13种手势,这种方法相比于传统的HMM方法而言,能够通过手势之间的关系来改善识别结果,其不足之处是过程比较复杂。Wilson和Bobick^[13]提出了一种基于状态的手势识别方法,通过将手势看作为空间中的一条轨迹,将轨迹曲线划分成不同的状态,然后用一组连续的状态序列来表示手势,将等待识别的图像序列与原型进行对比,所得的结果作为识别的判据。Korea Polytechnic大学的Lee等^[14]用熵分析法从背景较为复杂的视频流中分割并提取出手势区域以进行手势识别,该系统可识别6种手势,平均识别率达到95%。美国MIT媒体实验室的Darrel等^[15]采用动态时间规整算法对“Hello”手势进行识别,识别率达到了97%。Yang^[16]对40个美国手语进行识别,所采用的方法是时延神经网络,学习的对象是手势的运动轨迹。该方法在在训练集上的识别率高达99%,测试集的识别率也达到了96%。

虽然国内对手势识别的研究较晚,但取得的成果较显著,比较具有代表性的主要有:清华大学祝远新等^[17]提出了一种基于表现的新的手势识别技术。该课题组通过结合手势的运动表现、形状表现和时序信息建立了动态手势的时空表现模型。为抽取时空表现模型的参数,提出了基于运动、形状和颜色等多模式信息分层融合的策略。而且建立的实验系统可对12种手势进行在线识别,识别率超过90%^[18,19]。北方交通大学的王延江等人提取手势轨迹中关键点的运动方向,将之与标准手势中所有可能的特征码进行匹配,从而实现识别手势轨迹。中科院软件所的王西颖等^[20]结合HMM与模糊神

经网络提出了一种基于HMM-FNN模型的结构,能够识别出复杂背景下的动态手势。上海交通大学的刘江华等^[21]通过跟踪双手的运动识别的动态手势,采用光流法和耦合隐马尔科夫模型,所能达到的识别率为96.7%。张习文和王西颖^[23]等利用一组二维手势模型来替代三维模型,其过程是先利用贝叶斯分类器对静态手势进行识别,然后动态跟踪图像中的手指和指尖。由于该方法结合了基于模型和表现方法的特点,因此大大地减少了计算量。中科院自动化所的方亦凯^[22]提出一种快速的尺度空间特征检测方法,通过对手势图像中的Blob和Ridge结构的检测,得到手掌与手指的结构描述,进而完成手势识别。

2 视觉手势识别的关键技术

基于视觉的手势识别通常分为3个步骤,首先,输入图像经手势分割进行分离,定位出动态手势;然后,根据需求选择手势模型进行手势分析,并依据模型提取手势参数;最后,根据模型参数对手势选择合适的算法进行手势识别,具体流程如图1所示。本文以手势识别步骤为主脉来剖析手势识别的关键技术。

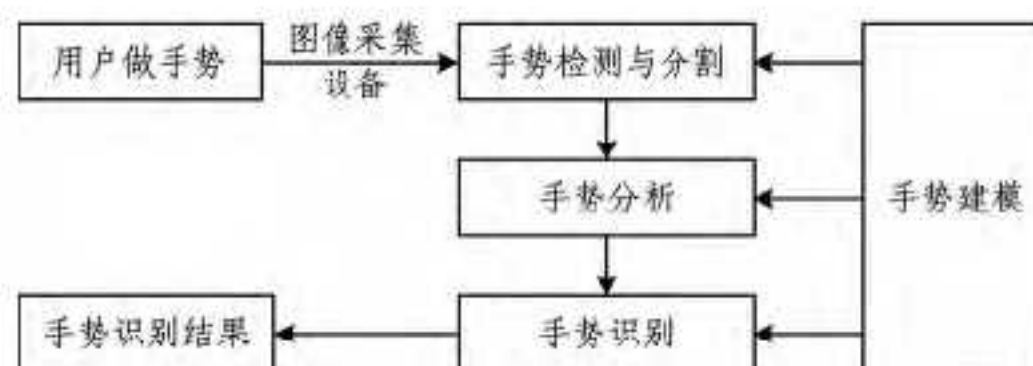


图1 手势识别流程

2.1 手势检测与分割

手势的检测与分割是手势识别的前提和基础,其分割识别的效果直接影响到下一步手势分析及手势识别的结果。手势检测的主要目的就是检测当前图像中是否存在手势以及其具体位置,而手势分割的实质就是将手势与背景分离。从手势分割的技术实现上来看,常见手势的检测分割方法主要有:基于运动信息的检测分割、基于徒手表现特征的检测分割以及多模式信息的检测分割。

2.1.1 基于运动信息的检测分割

手势动作往往与背景存在差异,现有文献资料利用运动信息达到手势分割的目的。这一领域主要有差影和光流两种研究方法。差影法对消除背景图像具有明显的效果,但缺陷是只适用于运动对象的分割。与之相比,光流计算方法不需要预先获取图像背景,在复杂环境下能够清晰地表示手势的运动。Hackenberg等^[24]运用光流法,实现手势的实时追踪;刘蓉等^[25]采用佩戴在手腕的单个加速度传感器获取手势加速度信号,提出一种实时手势加速度动作分割和识别方案。基于运动的分割方法一般需要在一些假设前提下才能发挥作用,例如差影法需要前景图像和背景图像有明显的颜色差别,而光流法要求背景为静止状态和尽量保持光照恒定,且图像里的主运动分量必须是手势运动。

2.1.2 基于表现特征的检测分割

表现特征是指手的肤色、纹理、指尖、手型和手的轮廓等。在众多的手势分割方法中,肤色分割则是最常用的手势分割方法,依据选择色彩空间的不同,研究者在不同的颜色空间下建模对手势进行分割。VAN^[26]研究基于RGB空间的肤色建模方法;刘昌盛^[27]采用YUV颜色空间进行手势分割;黄国

范^[28]等利用人体肤色的聚类特性,在 YCbCr 空间构建皮肤颜色分布模型并进行手势分割。该类分割方法不受手势形状变化的影响,但缺陷是受外界环境的干扰较大,如在不同的光照条件下,肤色变化较大,当光源位置角度或亮度发生改变时,误检出伪肤色的概率显著增加。

2.1.3 基于多模式信息的检测分割

为克服复杂环境下单一手势分割方法的局限,一些学者提出基于多线索融合的分割方法。如 Weng^等^[29]融合颜色、运动和形状定位,提高了手势分割的准确度;曹昕燕等^[30]提出结合肤色模型和运动检测技术的手势分割算法;方奎等^[31]提出基于样本集均匀化的肤色模型,建立了基于统计分析的手势分割方案。为降低复杂环境下手势分割的难度和计算复杂度,很多研究通过在手指或手掌上作特殊标识,或者强制要求使用单色的墙壁、特殊颜色的服装来简化背景。从分割准确度的角度来看,这些方法确实提升了系统分割的性能,但诸多人为的限制不仅影响了人机交互的自然性,而且应用范围也受到了一定的限制。

2.2 手势建模与分析

2.2.1 手势建模

对于基于视觉的手势识别系统而言,手势模型是至关重要的。对于复杂的人机交互,则需要一个精确而高效的模型使得系统能够识别用户的绝大多数手势。

目前,主要的建模方式是通过手势的外在形态来建模,它直接建立在图像的基础之上,通过对手势图像序列的分析来给手势建模。在静态手势识别中,常用的模型一般是基于手势图像属性的模型,如结构、边界、图像特征向量以及区域直方图特征等等。而对于动态手势识别而言,模型主要是建立在图像本身的变化或者运动轨迹的基础上。从目前的文献来看,手势建模的方法大致可归结为两类:1)基于表现的手势建模;2)基于三维模型的手势建模。

基于表现的手势建模是利用手势在图像序列里的表现特征为手势建模。常见的模型策略有灰度图、可变形模板、图像特征属性以及运动参数模型。灰度图主要采取直方图统计方式作为手势模型,通过积累图像序列里的时空信息来建立手势模型^[26]。可变模板是从轮廓中抽取一些特征点,构成具有一定的形变能力并能够描述如平移、旋转等一系列全局运动的点集^[32]。图像特征属性是指从图像序列中抽取独立且具有辨识性的参数,如轮廓、边界、图像矩、区域直方图等^[33]。运动参数表现模型更多地考虑了模型的时空关系,可选择的运动参数包括平移、形变以及方位等。这类表现模型利用光流等手段分割图像中的运动区域,建立手势的时空表现模型。基于表现的手势模型计算复杂度低,仅利用少数的局部特征就可以建立手势模型,因此应用较为广泛,但表现特征通常只适用于交流性手势,其通用性受到一定的限制。

基于三维模型的手势建模又可以分为纹理模型、网络模型、几何模型以及骨架模型。其中最常使用的三维模型是三维骨架模型,其参数是经过简化的关节角度参数和指节长度。考虑到动态手势中存在着大量的关节约束和运动依赖关系,需要处理的模型信息非常巨大,因此从降低计算负担考虑,常常采用几何模型替代或部分恢复的方法对模型简化处理。前者是利用简单的 3D 几何结构对骨架模型进行进一步简化,加快计算机实时处理和渲染的速度。方奎等^[31]以手区域的

重心为圆心构造切割圆,实现手指的计数,并完成手势的分类识别。完全恢复手势模型需要对手势的形状和运动具有先验知识,一般很难用于实时操作,因此实际应用中,经常采取部分恢复手势模型的策略,如只恢复指尖或手掌的信息,就可完成定位和导航等相对复杂的任务。基于骨架的三维模型适合于所有手势,但是从高维空间中实时恢复手势非常困难。针对大容量高维度的手势特征库,往往需要高效的特征搜索算法,但这些算法随着特征维数不断的升高,系统可能陷入“维数灾难”。因此很多系统加入了约束条件以免系统陷入所谓的“维数灾难”,如霍亚松^[44]等从归一直方图入手,提出非重叠式梯度方向直方图,该方法在人手和人体检测上的实验表明,不仅检测速度得到了提高,准确度也得到了改善。

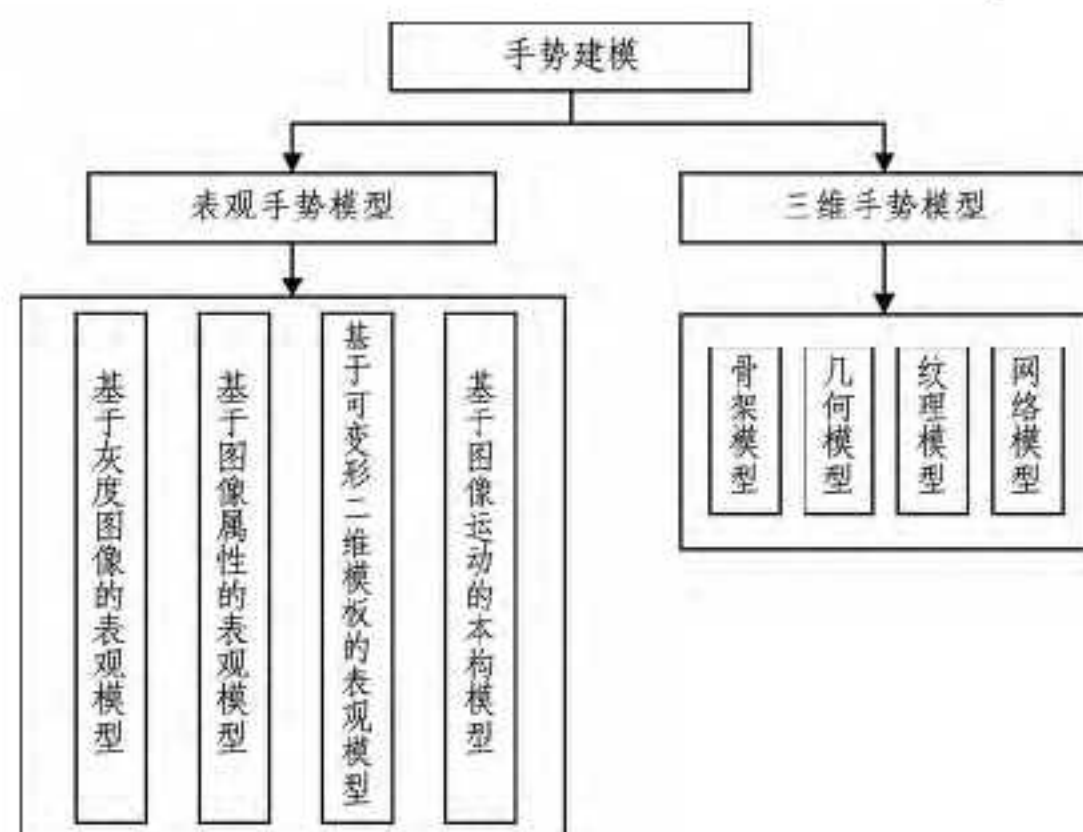


图2 手势建模的分类

2.2.2 手势分析

手势模型确定后,手势分析阶段的目标主要是根据选择的手势模型来估算参数。分析阶段由特征检测和参数估计两部分任务组成。在特征检测的过程中,首先必须定位用户,根据所用线索的不同,定位技术可分为基于颜色、运动以及模式定位等 3 种。对于参数估计而言,又可分为基于表现的手势模型参数估计和对三维手势模型的参数估计。综上所述,基于表现的手势模型共有 4 类,其中基于灰度图像本身的表现模型有许多不同的参数,在最简单的情况下,可以选择模型视图序列作为参数,也可以使用序列里各帧图像关于平均图像的特征分解表示。基于可变形二维模板表现模型的典型参数是模板节点的均值和其方差通过在训练集上进行主成分分析可得到的模型参数。基于运动图像的参数包括平移运动参数、旋转运动参数以及图像变形参数等。对于三维手势模型通常涉及到两组参数:角度参数和直线参数,而从检测出的特征去估计这些运动学参数通常包括初始参数和参数随时间变化的更新这两个环节。

2.3 手势识别

手势识别是对手势区域分割后的图像进行特征值提取和手势模型参数估计,是将参数空间中的点或轨迹分类到该空间里某个子集的过程,其包括静态手势识别和动态手势识别,动态手势识别最终可转化为静态手势识别。常见的研究方法包括以下 3 类。

2.3.1 基于模板匹配的方法

模板匹配是通过对大量手势样本的学习得到的手势模板,其核心思想是将输入目标数据与之前保存好的模板进行匹配,通过计算二者的相似度来完成匹配过程,通常采用的相

似度计算方法有:欧斯距离、Hausdorff 距离、夹角余弦等。模板匹配识别算法是比较简单的识别算法,其计算简单,速度快,不受光照、背景、姿态及标清等外界环境因素的影响,应用范围比较广。张良国等^[45]基于 Hausdorff 距离建立了一个手势识别系统,并首次利用 Hausdorff 距离模板匹配的思想,实现了中国手语手指字母集上的基于单目视觉的 30 个手指字母手势识别,识别率高达 96.7%。但其分类的准确率往往不高,不同手势之间的可分的程度也取决于特征值的选择,所以其可识别的手势数量不多,较为局限。

基于模板的方法不仅是指相互对应的模板匹配算法(Template Matching),还包括对时间、空间进行规整后的模板匹配算法,如动态时空规整算法和动态规划算法。由于动作快慢的不同,图像序列中的每一幅图像很难做到相互对应,因此进行时间上的规整是算法的重要步骤。基于动态时间规整的识别(DTW)算法是代表性的模板匹配法,它是一种基于动态规划思想且对非线性时间进行归一化再模式识别的算法。算法的主要原理是对于两个不同的时间范围,使用非线性规整函数对它们时间轴上的波动进行建模,为消除两个时间范围的差别,DTW 通过变化其中一个时间轴,从而在不同时间轴上消除时空表示模式之间的差别。陈文^[34]提出了适合基于加速度传感器的智能终端的 E-DTW 手势识别算法,通过斜率限定曲线路径,减轻模板匹配的计算量,降低手势识别的成本。该算法无需对手势进行时间上的分割来确定手势的起点和终点,但缺陷是实时性差。

2.3.2 基于人工神经网络的方法

人工神经网络(ANN)也叫神经网络(NN),主要指一种效用人脑神经网络的功能技术。神经网络具有高度的并行性、自适应性及移动的学习能力等特点,一些学者将其应用于手势识别的领域。赵新龙等^[35]采用基于 BP 神经网络的方法,设计出一套符合用户设计习惯的编辑手势对其进行识别,以实现计算机对草绘设计行为的准确理解和对草图的快速编辑与修改。而目前使用最多、应用最广的还是误差反向传播神经网络,也叫 BP 神经网络,它是由 McClelland 于 20 世纪 90 年代提出的。BP 算法包含两个过程:1)正向传播,即信号从入口处进入网络,依次通过各层,最后输出结果;2)误差反向传播,当输出与预估值出现偏差时,计算出两者的误差,将误差值向输入层反向传播依此来修正每层的权值和阈值,使误差最小,从而得到的输出结果更靠近预估值。Tusor 等^[36]利用模糊理论和人工神经网络的结合,建立了模糊神经网络的拓扑结构,用模糊特征值去描述和区分不同的手势;Li 等^[47]采用新型贝斯神经网络方法,在保证识别速率的同时,很好地避免了过拟合的问题。

神经网络的模型种类繁多,对于不同的需求可衍生出不同的形式,但其准确的识别还是建立在大容量的训练样本的基础之上。虽然一些改进的网络训练迭代规则可以提高训练速度,但仍不能完全避免过长的学习训练过程,因此该方法不适合需要实时识别的用户。

2.3.3 基于概率统计模型的方法

为了解决静态姿势识别中出现的分类错误等问题,往往引入基于概率统计的贝叶斯分类器,但贝叶斯分类器所涉及的密度函数的维数和具体表达形式难以确定,所以往往需要进行估计和假设,这一特性限制了贝叶斯分类器在复杂的手势识别过程的应用。

另一方面,HMM 作为基于概率统计的典型方法得到了广泛的应用。HMM 在 1968 年被提出^[37],其过程是一个双重的随机过程,它包括状态转移的随机过程和观察值输出的随机过程,其中状态转移的随机过程是隐式的,提高观察序列的随机过程表现。对于人的手势可表示为一个双随机过程,其内部状态是依赖于各个状态的观察输出信号表现出来的,而不能直接观察到马尔科夫过程,而对于手势信号的各个状态可以通过预定义的方式先行确定。

HMM 早期主要用于语音识别领域,近年来在手势识别方面亦有很大的发展。Elmezain^[38]利用 Baum-Welch 算法(BW)训练 HMM 已完成手势识别;严焰等^[39]利用 HMM 对手势指令进行建模,并运用 K-Means 算法矢量量化手势特征序列来提高手势识别性能;北京理工大学的张静等^[49]利用改进的条件随机场(Conditional Random Field,CRF)算法完成动态手势识别,识别率较传统的 HMM 有了显著地提高;基于向量化的 HMM 也得到了广泛的应用^[40,48]。该方法在训练样本足够时,可以达到很高的识别率,能够处理不同长度的信息。但因其需要为每种手势分别建立 HMM,计算量巨大,影响了系统的实时性能。

3 存在问题及发展趋势

近年来,手势识别技术已经由原来的理论探索阶段逐渐过渡至理论研究与实践产品并行的阶段,其研究取得了巨大的进步。现有成熟的技术基本能对实际中静态或动态手势进行检测识别,适用背景的复杂度和抗干扰能力都能满足一般用户的需求,准确率和识别率也得到了显著的提高。但现有的算法在识别率、运算速度、普适性及准确性等方面还有诸多问题有待解决,归纳起来,现有的视觉手势识别方法存在的主要问题及发展趋势是:

(1)针对复杂背景对手势分割的问题。目前大多数研究者均将注意力集中在手势最终的识别方面,并在单一背景下将手势进行分割,然后采用常用的识别方法将识别结果通过系统表达出来。但在现实应用中,手势通常处于复杂环境下,如光线过亮或过暗、手势距离采集设备不一、较多非手势等。而复杂的背景因素则会导致手势分割的准确性下降,不能精确地提取并识别手势,因此继续增强手势识别系统的鲁棒性,在保证高准确率的前提下提高适用性是更具挑战的课题。

(2)基于新平台的新研究。此前的研究大多数利用普通摄像头所提供的图像信息,虽然可以取得手势的形状、颜色以及运动信息,但却无法直接获得手势的骨架等三维特征。随着具有新功能的硬件的出现,如微软公司推出的深度视觉传感器设备 Kinect,所采集的图像不仅包括传统的 RGB 颜色信息外,同时获取了图像的深度信息,因此基于 Kinect 深度视觉结合二维视觉的研究方法正逐渐引起学者的广泛关注。如谈家谱等^[41]利用 Kinect 改进指尖检测的难点并完成手势识别,识别率高达 97.92%;中国传媒大学的吴晓雨等^[46]在利用 Kinect 获取手心节点的基础上,通过改进的 HMM 完成了动态手势识别;北京交通大学王松林等^[42]基于 Kinect 完成手势识别并实现用手势操控机器人。深度传感设备对复杂背景下手势分割性能带来了显著的提高,拓展了传统视觉手势识别的应用范围,对于如何进一步利用深度摄像头等新硬件已经成为了目前的研究热点。手势识别方法对比如表 1 所列。

表 1 手势识别方法对比

识别方法	代表算法	优点	缺点	应用情况
基于模板匹配的方法	传统的模板匹配法	计算简单,速度快,不受外界环境影响,应用范围广	准确率不高,可识别手势少	适用于静态手势识别
	动态时间规整 ^[34] (DTW)	算法简洁,消除时间上的非线性波动,识别率得到了提高	运算量大,实时性不高	静态手势识别、动态手势识别均适用,更适用于静态手势识别
基于人工神经网络的方法	BP 神经网络 ^[36]	变化丰富,可满足不同应用的需求,具备自动对信息处理的能力,抵抗噪音的干扰,有较好的鲁棒性	训练过程冗长,不适合实时识别	
基于隐马尔科夫模型的方法	隐马尔科夫模型 ^[37] (HMM)	提供了时间尺度的不变性,识别性能好,识别率高	初始化过程复杂,计算量巨大,系统实时性能差	适用于动态手势识别

(3) 综合多学科已有的研究成果。到目前为止,手势识别仍然集中在技术层面实现问题,综合利用其他学科已有的研究方面相对不足。除了需要计算机硬件、算法、软件等学科之外,手势识别还需要心理学、人类工效学的共同努力。冯志全等^[43]从认知心理的角度对人的手势行为进行分析研究,建立了用户手势的行为预测模型,并研究了基于此模型的交互算法。从现状分析来看,综合多学科的手势识别的研究,是值得研究者深入考虑的内容。

结束语 视觉手势识别在人机交互领域具有不可替代的地位。经过 30 多年的广泛研究,基于视觉的手势识别已经取得了很大的进展,但由于手势通常处于复杂的环境下,而复杂的背景因素则会导致手势分割的准确性下降,不能精确地提取并识别手势。本文从手势分割、手势模型建立、手势分析以及手势识别等多个方面对基于视觉的手势识别技术的研究现状和发展前景进行了分析,总结出了各种方法的优缺点,指出当前研究存在的问题以及发展趋势。随着人机交互的发展,视觉手势识别技术研究渐渐趋于成熟,各种算法的推陈出新必将给视觉手势识别带来新的研究途径。

参 考 文 献

- [1] 李勇,高文,姚鸿勋. 基于颜色手套的中国手指语字母的动静态识别[J]. 计算机工程与应用, 2002, 17: 55-58
- [2] Meenakshi P. Hand gesture recognition based on shape parameters[C]// International Conference on Computing, Communication and Applications (ICCC A). 2012: 317-319
- [3] Grimes G J. Digital data entry glove interface device: Technical Report US Patent 4[R]. 1983: 414-537
- [4] Ouhyoung M, Liang R H. A sign language recognition system using hidden markov model and context sensitive search[C]// Proceedings of the ACM Symposium on Virtual Reality Software and Technology. 1996: 59-66
- [5] 吴江琴,高文. 基于视觉手套数输入的汉语手指字母的识别[J]. 模式识别与人工智能, 1999, 12(1): 74-78
- [6] Grobel K, Assan M. Isolated sign language recognition using hidden markov model [C]// IEEE International Conference on Computational Cybernetics and Simulation. 1997, 1: 162-167
- [7] 周航. 基于计算机视觉的手势识别系统研究[D]. 北京: 北京交通大学, 2007
- [8] Marks, Richard. EyeToy-Innovation and beyond. sony compute entertainment america[OL]. <http://blog.us.playstation.com/2010/11/03/eyetoy-innovation-and-beyond>
- [9] 余涛. Kinect 应用开发实战: 用最自然地方式与机器对话[M]. 北京: 机械工业出版社, 2012
- [10] 付永刚,张凤军,戴国忠. 双手交互界面研究进展[J]. 计算机研究与发展, 2005, 42(4): 604-613

- [11] Takahashi T, Kishino F. Hand gesture coding based on experiments using a hand gesture interface device[J]. ACM SIGCHI Bulletin, 1991, 23(2): 67-74
- [12] Park H S, Kim E Y, Jang S S, et al. An HMM based gesture recognition for perceptual user interface[M]// Advances in Multimedia Information Processing. Springer Berlin Heidelberg, 2005: 1027-1034
- [13] Bobick A F, Wilson A D. A state-based approach to the representation and recognition of gesture[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1997, 19(12): 1325-1337
- [14] Lee J S, Lee Y J, Lee E H, et al. Hand region extraction and gesture recognition from video stream with complex background through entropy analysis [C]// Proceedings of the 26th Annual International Conference of the IEEE EMBS. San Francisco, CA, USA, 2004: 490-493
- [15] Darrell T, Pentland A. Space-time gesture [C]// IEEE Conference on Compute Vision and Pattern Recognition. 1993: 335-340
- [16] Yang M H, Ahuja N. Recognizing hand gestures using motion trajectories[M]. Face Detection and Gesture Recognition for Human-Computer Interaction. Springer US, 2001: 53-81
- [17] 祝远新,徐光祐,黄裕. 基于表现的动态孤立手势识别[J]. 软件学报, 2000, 11(1): 54-61
- [18] 任海兵,祝远新,徐光裕,等. 连续动态手势的时空表现建模及识别[J]. 计算机学报, 2000, 23(8): 824-828
- [19] 任海兵,祝远新,徐光裕,等. 复杂背景下的手势分割与识别[J]. 自动化学报, 2002, 28(2): 256-261
- [20] 王西颖,戴国忠,张习文,等. 基于 HMM-FNN 模型的复杂动态手势识别[J]. 软件学报, 2008, 19(9): 2302-2312
- [21] 刘江华,陈佳品,程君实. 基于光流及耦合隐马尔可夫模型的动态手势识别[J]. 上海交通大学学报, 2003, 37(5): 720-723
- [22] 方亦凯,程健,汪孔桥,等. 基于快速尺度空间特征检测的手势识别方法[J]. 中国图象图形学报, 2009, 14(2): 214-220
- [23] 王西颖,张习文,戴国忠. 一种面向实时交互的变形手势跟踪方法[J]. 软件学报, 2007, 18(10): 2423-2433
- [24] Hackenberg G, McCall R, Broll W. Lightweight palm and finger tracking for real-time 3D gesture control[C]// Proceeding on Virtual Reality Conference. Singapore, 2011: 19-26
- [25] 刘蓉,刘明. 实时手势加速度动作分割与识别研究[J]. 小型微型计算机系, 2012, 33(7): 1620-1624
- [26] Van-Den-Bergh M, Van-Gool L. Combining RGB and TOF cameras for real-time 3D hand gesture interaction[C]// IEEE Workshop on Application of Computer Vision. New Jersey, 2011: 66-72
- [27] 刘昌盛,厉树忠,赵姝颖,等. 一种复杂背景下的手部图像分割方法[J]. 河北科技师范学院学报, 2007, 21(3): 46-49

- [28] 黄国范,程小平. 基于历史的动态手势识别[J]. 西南大学学报(自然科学版),2009,31(1):106-110
- [29] Weng Chuan-bo, Li Yang, Zhang Ming-min, et al. Robust hand posture recognition integrating multi-cue hand tracking[M]. Germany:Springer-Verlag,2010:497-508
- [30] 曹昕燕,赵继印,李敏. 基于肤色和运动检测技术的单目视觉分割[J]. 湖南大学学报(自然科学版),2011,1(1):78-83
- [31] 方奎,欧阳宁,莫建文. 基于椭圆模型的手势识别与仿真[J]. 计算机仿真,2011,28(3):267-280
- [32] 关然,徐向民,罗雅愉,等. 基于计算机视觉的手势检测识别技术[J]. 计算机应用与软件,2013,30(1):155-160
- [33] 冯志全,杨波,郑艳伟. 基于特征点分布分析的手势特征检测方法[J]. 计算机集成制造系统,2011,17(11):2333-2340
- [34] 陈文. 基于加速度传感器的智能终端手势识别关键技术研究[D]. 长沙:国防科学技术大学,2011
- [35] 赵新龙,方贵德,沈莉芳. 基于 BP 神经网络的草图编辑手势识别[J]. 浙江水利水电专科学校学报,2006,18(4):31-33
- [36] Tusor B, Varkonyik A R. Circular fuzzy neural network based hand gesture and posture modeling[C]//Proceeding on Instrumentation and Measurement Technology Conference(I2MTC). Texas,2010:815-820
- [37] Poritz A B. Hidden markov models: A guided tour[C]//IEEE Conference on Acoustics, Speech and Signal Processing. 1988:7-13
- [38] Elmezain M, Al-Hamadi A, Michealis B. Real-Time Capable System for Hand Gesture Recognition Using Hidden Markov Models in Stereo Color Image Sequences[J]. WSCG Journal,2008,16(1):65-72
- [39] 严焰,刘蓉,黄璐,等. 基于 HMM 的手势识别研究[J]. 计算机应用,2012,32(增刊 1):128-133
- [40] Glomb P, Romaszewski M, Sochan A, et al. Unsupervised parameter selection for gesture recognition with vector quantization and hidden markov model[J]. Lecture Notes in Computer Science,2011(69):170-177
- [41] 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):1795-1800
- [42] 王松林. 基于 Kinect 的手势识别与机器人控制技术研究[D]. 北京:北京交通大学,2014
- [43] 冯志全,杨波,李毅,等. 基于交互行为分析的手势跟踪方法[J]. 计算机集成制造系统,2012,18(1):31-40
- [44] 霍亚松,张锐. 非重叠低维度梯度方向直方图[J]. 模式识别与人工智能,2014,27(3):242-247
- [45] 张良国,吴江琴,高文,等. 基于 Hausdroff 距离的手势识别[J]. 中国图象图形学报 A 辑,2002,7(11):1144-1150
- [46] 吴晓雨,杨成,冯琦. 基于 kinect 的手势识别算法研究及应用[J]. 计算机应用与软件,2015,32(7):173-176
- [47] Linjun L, Shuling D. Bayesian neural network approach to hand gesture recognition[C]//Proceedings of 2014 IEEE Chinese Guidance, Navigation and Control Conference. 2014:2019-2023
- [48] Wang Liang, Liu Gui-xi, Duan Hong-yan. Dynamic and combined gestures recognition based on multi-feature fusion in a complex environment[J]. The Journal of China Universities of Posts and Telecommunications,2015,22(2):81-88
- [49] Liling M, Jing Z, Junzheng W. Modified CRF algorithm for dynamic hand gesture recognition[C]//Proceedings of the 33rd Chinese Control Conference. 2014:4763-4767

(上接第 89 页)

参 考 文 献

- [1] Park J, Chen Z, Kiliaris L, et al. Intelligent vehicle power control bases on machine learning of optimal control parameters and prediction of road type and traffic congestion[J]. IEEE Transactions on Vehicular Technology,2009,58(9):4741-4756
- [2] Rahman H A, Marti J R, Srivastava K D. Road traffic forecasting through simulation and live GPS-Feed from inter-vehicle networks[C]//Global Humanitarian Technology Conference (GHTC). 2012:36-40
- [3] Chung E H, Sharlaby A. Expected time of arrival model for school bus transit using real-time global positioning system-base automatic vehicle location data[J]. Journal of Intelligent Transportation System: Technology, Planning, and Operations,2007,11(4):157-167
- [4] 王建,邓卫,赵金宝. 基于贝叶斯网络多方法组合的短时交通流量预测[J]. 交通运输系统工程与信息,2011,11(4):147-153
- [5] 李天雷. 基于 GPS 数据的公交行程时间计算与预测系统[D]. 长春:吉林大学,2009:7-25
- [6] 张松林,张昆. 连续随机变量非线性函数的期望和方差的近似求法[J]. 大地测量与地球动力学,2008,4(28):107-110
- [7] 刘伟平. 改进的图像匹配方法在汉字识别中的应用[D]. 广州:暨南大学,2015:18-32
- [8] Hsiao S W, Tsai H C. Applying a hybrid approach based on fuzzy neural network and genetic algorithm to product form design[J]. International Journal of Industrial Ergonomics,2005,35(5):411-428
- [9] 于海燕. 基于模糊神经网络的电力短期负荷预测[D]. 天津:天津大学,2007
- [10] Capi G, Doya K. Evolution of recurrent neural controllers using an extended parallel genetic algorithm[J]. Robotics and Autonomous Systems,2005,52(2):148-159
- [11] 叶菁,张莹,阮一文. 一种改进型交叉算子和自识别高变异算子新型遗传算法的研究[J]. 福州大学学报(自然科学版),2009,37(6):809-817
- [12] 陈长征,王楠. 遗传算法中交叉和变异概率选择的自适应方法及作用机理[J]. 控制理论与应用,2002,19(1):41-43
- [13] 吕淑平,祝捷. 一种改进的自适应混合神经网络盲分离算法[J]. 计算机应用研究,2013,30(4):1055-1057
- [14] 李晓华,丁雍,斯茵. 模糊综合评判模型在宝坻水源地地下水水质评价中的应用[J]. 地下水,2014,36(1):6-8
- [15] Duru O. A fuzzy integrated logical forecasting model for dry bulk shipping index forecasting-An improved fuzzy time series approach[J]. Expert Systems with Applications,2010,37(7):5372-5380
- [16] Raju D K B, Harsha N, Raju V K V. Prediction of tensile Strength of friction stir welded joints Using artificial neural networks[J]. International Journal of Engineering Research & Technology,2012,1(9):1-5