

基于深度信息的动态手势识别综述

陈甜甜 姚 璜 左明章 田 元 杨梦婷

(华中师范大学教育信息技术学院 武汉 430079)

摘 要 随着计算机技术的飞速发展,自然、简单、非接触式的手势识别在人机交互方面备受青睐。动态的手势识别一直是人机交互领域研究的热点与难点,深度传感器的出现为手势识别的研究提供了更加鲁棒的数据。为了解动态手势的发展现状,在广泛调研现有文献和最新成果的基础上,对基于深度信息的动态手势从手势分割、手势建模、特征提取、手势识别 4 个方面进行阐述,介绍动态手势识别相关的应用领域,并对其中存在的难点与问题进行讨论。

关键词 人机交互,动态手势识别,深度信息,手势分割,特征提取

中图分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.12.006

Review of Dynamic Gesture Recognition Based on Depth Information

CHEN Tian-tian YAO Huang ZUO Ming-zhang TIAN Yuan YANG Meng-ting

(School of Educational Information Technology, Central China Normal University, Wuhan 430079, China)

Abstract With the rapid development of computer technology, natural, simple and non-contact gesture recognition is favored in human-computer interaction. Dynamic gesture recognition has always been a hot and difficult issue in the field of human-computer interaction. In order to understand the development status of dynamic gestures, this paper described the dynamic gestures based on depth information from four aspects of gesture segmentation, gesture modeling, feature extraction and gesture recognition based on the extensive investigation of the existing literature and the latest achievements, introduced the applications of dynamic gesture recognition, and discussed the existing difficulties and problems.

Keywords Human-computer interaction, Dynamic gesture recognition, Depth information, Gesture segmentation, Feature extraction

1 引言

随着计算机技术的发展,动态手势识别已经成为一种新兴的人机交互方式,同时也是一个重要的研究方向^[1]。传统的动态手势识别手段主要是利用智能传感设备(鼠标和笔)、可穿戴设备(数据手套^[2])以及单个或多个摄像头进行数据采集的视觉方法。数据手套和手指跟踪器等外部设备定位人体的手的位置和角度,对人的舒适性和自然度有较大影响,并且这种方式的成本较高。基于视觉的手势识别易受到光照和复杂背景的影响,从而导致识别率较低。随着深度传感器的出现,基于深度信息的手势识别^[3-5]通过提取手势的深度数据,有效地消除了光照和复杂背景等干扰,使手势识别技术的研究迈上了一个新的台阶。Jojic 等^[6]最早提出基于视差的指向手势的检测和估计。2002 年,文献^[7]通过使用 coded lighting 技术来获取 3D 数据用于静态手势识别。基于深度信息的动态手势识别, Liu 等^[8]较早提出了一种通过使用主动感测硬件获取实时深度图像数据的序列来识别手势的方

法。Rautaray 等^[9]总结了 2005 年到 2012 年基于视觉的手势识别系统的进展,为后来 HCI 手势领域的研究提供了参考价值。2012 年, Suarez^[10]对基于深度图像的手势跟踪和手势识别进行了简单综述,重点叙述了手势识别的应用程序与识别系统。近几年动态手势识别的研究在特征提取中有所创新,并对识别算法进行了优化,使得手势识别率大大提高。

本文对近几年基于深度信息的动态手势研究进行了归纳梳理,重点阐述了基于深度信息的动态手势识别的研究思路与主流方法。本文第 2 节主要介绍深度传感器与深度数据;第 3 节介绍手势分割的 3 种方法:深度阈值分割、像素点聚类、结合彩色图像和深度图像的分割;第 4 节介绍手势建模方法;第 5 节给出动态手势特征的分类,主要包括动态手势的全局特征、局部特征以及融合特征;第 6 节介绍动态手势识别的主流方法;第 7 节阐述现阶段手势识别研究的具体应用;最后总结全文。

2 深度传感器与深度数据

传统深度相机主要是基于双目视觉^[11]的方法,其原理是

到稿日期:2017-12-29 返修日期:2018-04-08 本文受“十二五”国家科技支撑计划项目(2015BAK33B02,2015BAK27B02)资助。

陈甜甜(1993-),女,硕士生,主要研究方向为计算机视觉、人机交互;姚璜(1984-),男,博士,讲师,主要研究方向为模式识别、计算机视觉, E-mail: yaohuang@mail.ccnu.edu.cn(通信作者);左明章(1969-),男,博士,教授,主要研究方向为教育信息化、数字媒体技术;田元(1982-),女,博士,讲师,主要研究方向为计算机视觉、增强现实;杨梦婷(1995-),女,硕士生,主要研究方向为人机交互。

通过两个相隔一定距离的摄像机同时获取同一场景的两幅图像,利用立体匹配算法找到两幅图像中对应的像素点,随后根据三角原理计算出时差信息,再转换成可以用于表征场景中物体的深度信息。该方法复杂度高,实时性差。目前深度传感器技术^[12]主要有 TOF(Time of Flight)、结构光和激光雷达等,使用得较多的是 TOF 相机。TOF 相机获取深度图像的原理是对目标场景发射连续的光脉冲,然后用传感器接收从物体返回的光,通过探测光脉冲的飞行时间得到目标物体的距离。TOF 相机所获得图像的分辨率远不及彩色图像的分辨率,其深度值易受噪声干扰,此外 TOF 相机的价格昂贵。结构光是具有特定模式的光,基于结构光的深度传感器的原理是将结构光投射至场景,并由图像传感器捕获相应的带有结构光的图案,由于结构光的模式图案会因为物体形状发生形变,因此根据模式图像在捕捉到的图像中的位置以及形变程度利用三角原理计算场景中各点的深度值。结构光测量技术提供了高精度且快速的三维信息,在游戏、医疗等领域已经得到了广泛应用。激光雷达测距技术通过扫描的方式得到场景的三维信息,其原理是按照一定时间间隔向空间发射激光并记录各个扫描点的信号从激光雷达到被测物体,然后又经过物体反射回来的相隔时间,据此计算出物体表面与激光雷达的距离。

利用深度摄像机捕获到的 RGB-D 图像较以往的二维图像具有更加复杂的数据结构,为手势识别领域提供了新的研究方向^[13]。深度传感器设备(如 Kinect、Xtion Pro、TOF 相机等)的出现,为动态手势识别的研究带来了更多的可能。现今在手势识别系统的研究中更多使用微软的 Kinect 设备, Kinect 传感器的原理是采用光编码(Light Coding)技术^[14]。不同于传统的 TOF 或者结构光测量技术,Light Coding 技术利用红外线发射器发射的连续光对 Kinect 视野范围内的测量空间进行编码,红外线接收器读取经物体反射回来的光线并进行解码运算,从而形成深度图像。

2014 年,最新的 Kinect 2.0 版本提供了更高的清晰度和更宽的视角,通过红外技术去除图像中的环境光线,从而恢复物体的真实结构信息。与上一个版本相比,它配备了更高的分辨率的传感器,以每秒 30 帧的速度录制 1920×1080 全高清分辨率视频,16 位深度传感器可捕捉 512×424 像素的深度图像,深度影像的可侦测范围提升到 0.5~8m,进行人体追踪、骨架识别的范围为 0.5~4.5m(该数据表示从深度摄像头到被识别物体的距离)。

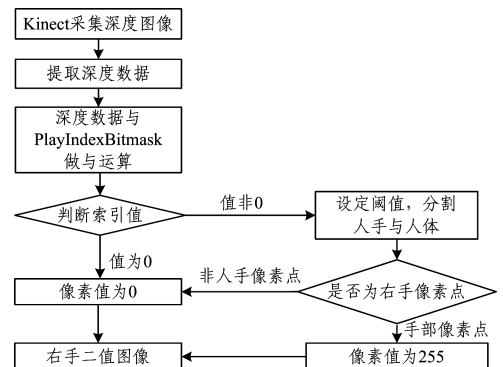
深度数据^[3]是深度传感器的精髓,它提供了第三维的特征,使数据的处理更具有代表性和真实性。深度传感器获取的深度数据,是指其视野范围内的物体到传感器设备的空间距离。研究者利用深度数据可以设计比较简单的手势识别系统,相比传统的手势识别系统,这种技术的优点有:1)对光照强度和纹理变化较不敏感;2)通过单个摄像机提供 3D 信息,而在 RGB 域中需要立体系统来实现相同的目标;3)维护隐私,不可能识别深度摄像机拍摄的人物的面部细节,此功能有助于保持身份机密^[15]。深度信息又可以被进一步处理为一系列关节构成的骨架模型来获取无标记的身体姿势估计,这也为后续的研究工作提供了更广阔的思路。

3 手势分割

手势分割是指把图像中属于手势的像素点与不属于手势的像素点区分开来,得到手势在图像中的区域。手势分割的结果会对需要手部细节特征提取和识别的系统产生一定的影响,手势分割的技术难点在于环境因素的复杂性,以及手动动作在空间上的灵活性导致的手部外形特征复杂多变。目前不存在一种在所有应用系统和背景条件下都能取得良好的分割效果^[16]的方法。常用的手势分割方法可分为基于深度阈值分割的方法、像素点聚类的方法以及彩色图像和深度图像结合的分割方法 3 种。

3.1 深度阈值的分割方法

传统基于计算机视觉的手势分割方法有增加限制法、肤色检测法等,但前者降低了手势的自由性,后者容易受到复杂背景以及光照变化的影响。随着深度传感器的普及,研究人员对基于深度信息的手势分割进行了广泛的研究^[17-19]。文献^[20]利用 Kinect 获取深度图中的每个像素点的深度数据,根据用户和背景索引值不同的原理对人与背景进行分割,然后设置手势与人体深度的阈值再分割出完整的手势,具体流程如图 1 所示。但这种方式只能进行简单的分割,最后的效果并不理想,手腕数据以及手部旁边的衣物都会对分割结果产生影响。



注: PlayerIndexBitmask 的默认值为 7

图 1 基于深度阈值的手势分割流程图

Fig. 1 Flowchart of gesture segmentation based on depth threshold

为解决上述问题,林海波等^[4]通过引入可变阈值的方法去除手腕对手势的影响,进而进行手势分割,然后使用形态学处理使手势分割效果更加准确,并且提出联合手势重心坐标和最大内切圆圆心坐标的手势定位法以准确地定位出手势,同时采用球形阈值限定法处理异常手势动作轨迹,为后续的手势特征提取以及跟踪阶段提供了有力的保证。文献^[21]为避免同样的问题,提出手势的端点检测和可变阈值算法,定位算法与文献^[4]相同,该手势分割的方法更精确、稳定。

基于深度阈值的手势分割方法易受到场景的限制,并且人手需在摄像头最前面。文献^[22]结合深度阈值分割方法的本质,提出基于深度图像的随机决策森林框架改进的手部分割策略。该方法从手部结构的中心点导出新的深度特征,增强了以往深度特征的泛化能力,尽可能多地描述手部像素的深度不变性,从而确保了手部分割的准确性。文献^[23]提出了一种逐帧判断手势标签分布概率进而分割连续手势的实时算法。

3.2 像素点聚类

采用特征空间聚类法进行图像分割是将图像空间中的像素用对应的特征空间点表示,然后根据它们在特征空间的聚类对特征空间进行分割,最后将它们映射到原图像空间得到分割结果。人体各个部分都有较为接近的深度值,根据这个特点,像素点聚类方法对图像中的每个像素点进行聚类,利用人手形状等先验知识提取出包含人手的聚类,从而实现人手分割。

基于像素点聚类方法的手势分割多采用 K-means 算法, K-means 聚类算法最早由 MacQueen 于 1967 年提出,是一种无监督分类方法,也称作 K-均值算法。其基本思想是:通过迭代的方法,逐次更新各聚类中心的值,直到得到最好的聚类结果。文献[24]通过对不同阈值区间的像素进行 K-means 聚类来分割手势区域,当手势输入的数据源发生变化时,可连续应用 K 均值聚类,随机选取聚类的质心作为初始值, K-means 收敛之后,每只手的点被聚类。若两个质心之间的距离小于预定义的值,则将这两个簇合并,从而得到分割后的手势。Wen 等[25]在深度图像中同样利用 K-means 算法进行手部聚类,然后进行手部分割,再检测指尖位置,通过手部数据特征进行手势分类。这种方法的缺点是计算量较大,并且初始质心的位置选择会影响分割效果,如文献[25]中如果两只手的质心之间的距离在设定的阈值之内,则合并两个簇只能检测到一只手,这种方法对单手分割则不会有影响。

3.3 结合彩色图像和深度图像的分割

基于彩色图像的人手分割方法不会受到场景过多的限制,这种方法比较成熟,如帧差法、肤色法。基于深度图像的手势分割不受光照条件变化的影响。利用两种方法的优势,文献[26]和文献[27]引入肤色模型和深度信息提取精确的手势区域数据,前者提出一种椭圆模型的高效活性色区选择方法,后者对手势二值图利用 Border-Following 算法获取手势轮廓,连接相邻视频帧中的手心点坐标作为一条手势的运动轨迹。

深度图像与彩色图像的对齐需要很高的深度精度,而深度数据的噪声与孔洞使得基于深度信息的图像分割有许多不匹配的像素。文献[28]为解决这一问题,采用遗传算法将深度图与 RGB 图像进行匹配,提出一种基于深度图、RGB 图像和像素位置调整分割边的期望最大化(EM)算法。该算法提高了手势分割的精度。文献[29]为消除手势成像中的干扰以及手势目标边界模糊的情况,提出了改进的最大类间方差法以有效地对手势图像进行分割。彩色数据与深度数据融合的方式为手势分割提供了新的研究思路。

除了上述介绍的 3 种常用方法,基于深度信息的动态手势分割也会用到光流场[30]的方法。光流技术的思想是利用图像序列中像素在时间域上的变化以及相邻帧之间的相关性找到上一帧与当前帧之间存在的对应关系,从而计算出相邻帧之间对象运动的信息。基于光流场的方法可以反映出运动不一致的区域,具有较好的鲁棒性,但其存在的缺点是当运动物体有遮挡或存在运动偏差时分割效果不好,并且计算过程复杂。

另外,文献[31]运用深度学习技术在手势分割模块中通过使用两流卷积神经网络(CNN)将帧分类为手势帧和过渡

帧,将连续手势序列分割成独立的手势序列。深度学习在手势分割中具有一定的应用价值。

4 手势建模

在传统手势识别系统中,手势建模是一个最基本的步骤,不同的手势模型方法,其特征提取的方法也不同。对动态手势识别而言,模型主要建立在图像本身有变化或运动轨迹的基础上。总体上,手势建模方法主要分为基于表观的手势模型和基于三维的手势模型。

4.1 基于表观的手势模型

基于表观的手势模型利用手势在图像序列中的表观特征进行建模。常采用灰度图或历史图、可变形模板、图像特征属性以及运动参数模型。灰度图和历史图采用直方图等统计方式统计图像序列中的时空信息来建立手势模型[32];可变形模板是从轮廓影像中抽取一些特征点,构成具有一定的形变,能够描述诸如平移、旋转等全局运动的点集[33];图像特征属性指从图像序列中抽取具有辨识度的属性参数,如轮廓、角点、图像矩等[34];运动参数表观模型更多地考虑模型的时空关系,可选择的运动参数包括平移、旋转、形变以及方位等,这种表观模型可利用光流等手段分割图像中的运动区域,建立手势的时空表观特征[35]。

4.2 基于三维的手势模型

动态手势中存在大量的关节约束和运动依赖关系,骨架模型[36]和几何模型[37]是其常用的方法。三维模型应具有足够的自由度以适应图像中手的尺寸。利用简单的几何结构对骨架模型进行进一步简化,可加快计算机实时处理和渲染的速度。完全恢复手势模型要求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,实际应用中常采取部分恢复手势模型的策略,提取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。

基于表观的手势模型的计算复杂度低,利用少量的局部特征就可以建立手势模型,应用较为广泛但通用性受到限制。基于骨架的三维模型的适应范围广,所需特征维数高,存在的缺点是系统可能会陷入“维数灾难”。因此,采用 PCA 算法[38]等来降低特征空间的维度。在建立手势模型的过程中,深度学习方法的应用越来越广泛,如文献[39]利用两流递归神经网络(2S-RNN)有效地融合多模态特征并对手势序列进行建模。

5 特征提取

动态手势特征主要分为全局特征、局部特征和融合特征,它们不但包括空间信息,还包括时间信息。动态手势的全局时空特征可以通过运动历史图[32]、运动能量图[40]、时空形状[39]等加以描述。局部动态手势特征主要提取图像序列中变化较为明显的局部特征点,再对这些特征点的各种属性进行统计建模,实现对动态手势的描述。常用的局部特征有运动速度、运动方向、梯度方向直方图[41]、尺度不变特征[42]等。融合特征主要包含全局特征和局部特征两种,如轮廓特征与骨骼特征的融合[43]。基于深度信息的动态手势特征分类如图 2 所示。

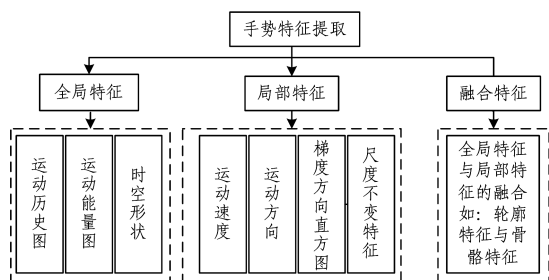


图 2 动态手势特征分类

Fig. 2 Classification of dynamic gesture feature

5.1 动态手势的全局特征

动态手势全局特征利用深度图的信息进行提取。文献[40]对动态手势特征的获取方法是将深度序列的每一帧投影到 3 个正交平面;然后通过计算和阈值化连续帧投影图的差异得到运动能量的二值图,并累计整个深度序列的运动能量二值图以产生深度运动图;最后提取深度运动图的方向梯度直方图特征(DMM-HOG),作为动态手势的描述符。

文献[44]对文献[34]提出的只考虑了手指的外轮廓,而忽略了可以为复杂手势提供重要形状和结构信息的手掌区域的系统进行了改进,在进行手势特征提取时利用最小数近似凸分解的方法检测出手的各个手指部分,利用新的相异性距离度量手指-陆地移动距离(Finger Earth Mover Distance, FEMD),将被识别的手形状与预定义的模板进行匹配。该方法可以更好地分辨略有差异的手势。在此基础上,文献[45]采用改进的 FEMD 套索算法定位手掌图像,提取指尖特征。

针对只提供某种全局特征存在局限性的问题,文献[46]提出以超像素的 EMD(Superpixel Earth Mover's Distance, SP-EMD)作为衡量不同手势相异性的距离度量。不同于 FEMD^[33]只用了轮廓信息,SP-EMD 充分利用了形状(深度)和纹理(颜色)信息,并且这些信息都是以超像素的形式来描述的。联合颜色-深度的 SP-EMD 在数据集 The NTU Hand Digits dataset^[34]上不仅获得了高平均识别准确率,而且显示了优于 FEMD 方法的识别性能。文献[47]针对动态手势识别方法受手势旋转、平移、缩放以及区分度较小的手势的影响,提出了一种基于手势二进制编码和类-Hausdorff 距离模板匹配的手势识别方法。

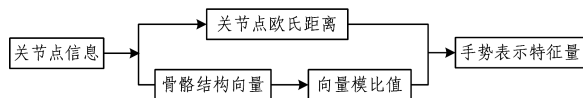
5.2 动态手势的局部特征

动态手势的局部特征主要提取图像序列中变化较为明显的局部特征关键点,对属性进行统计建模,实现对手势的描述。利用静态手势的连续特征可以更加有效地表示手势的三维形状和结构,Zhang 等^[48]设计了一种基于三维物体表面法向量的直方图特征(Histogram of 3D facets, H3DF),在 ASL(American Sign Language)数据集上获得了更好的效果。

文献[49]基于深度信息对手势进行分割之后,提出了 HON4D(Histogram of Oriented 4D Normals)描述子,对手部的兴趣点 (x, y, z, t) 在 4D 空间构建一个面,再量化这个 4D 曲面,得到 HON4D 描述子。在 MSR Gesture 3D 数据集上该特征的效果很好,但该算法是在整个视频序列上进行的,不能进行实时的识别。

深度信息可被进一步处理为骨骼数据,基于关键骨骼点

信息的特征提取是另一种常用方法。动态手势研究通过对人手的追踪定位获得手部(包括手部、肘部和肩部等)关节的三维坐标信息,然后对获取的手部关节信息进行归一化处理,提取动态手势特征构建手势特征描述向量来表示动态手势。如文献[50]根据关节点信息实现对动态手势轨迹的追踪,在获取的关节点信息的基础上,研究了动态手势特征提取及动态手势表示方法。该方法运用人手关节点间的欧氏距离,并将人体骨骼结构特征向量模比值作为主要手势特征,其主要流程如图 3 所示。

图 3 特征提取流程图^[50]Fig. 3 Flowchart of feature extraction^[50]

文献[51]提出了一种基于三维局部稀疏运动的尺度不变特征变换(3D SMO-SIFT)描述子。首先用灰色金字塔、深度金字塔和光流金字塔为每个灰度帧和深度帧建立尺度空间;然后根据光流方差提取感兴趣区域,计算水平方向和垂直方向的方差;在每个感兴趣的区域提取角点作为兴趣点,同时使用灰度和深度光流的信息检测尺度空间中的鲁棒关键点;最后在三维梯度空间和三维运动空间计算 SIFT 描述符。

手势的局部特征很大程度上有利于动态手势的识别,同时有大量文献资料^[52-53]对全局特征与局部特征的融合进行了研究。

5.3 动态手势的融合特征

设计单一的特征提取对于某些特定手势的分类效果较好,但是单一的特征可能不具备稳健性和普适性,在通用的手势识别处理上会出现一定的问题,因此需考虑特征融合。文献[52]提出了一种多重空间特征融合的方法来识别手势深度图像,与文献[48]中的编码法向量的方式相比,其通过主成分直方图与梯度方向直方图在多个尺度上更好地编码了三维物体的形状,通过深度分布直方图避免了形状描述子对深度改变敏感的问题,并补充了深度分布信息。

除了上述将两种形状描述子融合的方式,文献[43]采用骨骼特征(局部特征)与剪影特征(全局特征)融合的方式在 MSR Action3D 数据集上进行测试。MSR Action3D 数据集中部分是人体上部分肢体的运动,如挥手、拍手等,实验证明了这种特征融合方式比单一特征提取的效果更好,表现了特征融合的优势。文献[53]提出了一种在深度序列中融合全局特征运动能量图和局部特征梯度直方图来识别手势的方法。首先将归一后的运动能量图向量细化为一组分段,其对应的帧索引用于分割视频。每个手势动作由 3 个深度运动图(DMM)表示,分别在 3 个正交投影视图中捕捉运动特征,根据 DMM 计算多尺度的定向梯度直方图(HOG)描述符来捕获手势表观的变化特征。目前在 RGBD 视频的行为识别中,为提高识别率,许多方法采用了特征融合的方式。但获取多种特征时,特征融合并不一定是最佳的方式,特征选择的方法值得考虑。如文献[54]选择了关节点表观信息的 HON4D 信息和运动变化的关节点相对距离特征,采用基于熵的自适应特征选择方法进行人体行为识别,将其该方法与特征融合的

算法进行对比,证明了提出的特征选择的方式提高了识别率。近年来,随着基于学习的特征提取方法性能不断提升,深度学习方法逐渐成为最流行的基于学习的特征表示方法^[31]。

综上所述,手势特征提取常常需要对手势特征进行一定的设计,特征设计的合理性在一定程度上会影响手势识别的准确率。对于特定的动态手势,如何更有效地进行融合处理是值得考虑的问题。

6 手势识别

动态手势的识别方法主要分为基于模板匹配的方法和基于模式识别的方法。模板匹配中最常用的是动态时间规整(Dynamic Time Warping, DTW)算法^[55];模式识别中使用较多的是支持向量机^[56](Support Vector Machine, SVM)、隐马尔科夫模型^[57](Hidden Markov Model, HMM)和神经网络^[58](Neural Networks)等。除了主流的几种方法,基于粒子滤波(Particle Filter, PF)的 condensation 算法^[59]、有限状态机^[60](Finite-state machine, FSM)等方法也较为常用。

6.1 基于 DTW 的方法

DTW 算法最早应用于语音识别中,主要用于孤立词的识别,在手势识别中可以解决不同玩家完成动作的时间长度不一的问题。对于动态手势中的模板匹配方法,研究者多采用 DTW 算法^[61]进行处理。最早将 DTW 算法应用于手势的是 Corradini^[62]提出的离线识别小手势词汇。使用 DTW 算法进行手势识别时,首先需要对模板进行训练,对每个手势进行预处理,然后提取特征,将其归一化处理成一个序列作为模板,再对待测手势进行与模板序列相同的处理,最后与模板手势库中的每一个模板进行匹配分析,将距离最小的模板作为识别结果返回,具体流程如图 4 所示。文献^[63]在原始 DTW 算法的基础上提出了一种改进的 N 维动态时间规整方法,将 Kinect 获取的骨骼点进行特征提取与模板训练,最后进行实时手势识别。该算法在效率和实时识别精度上都有所提高。

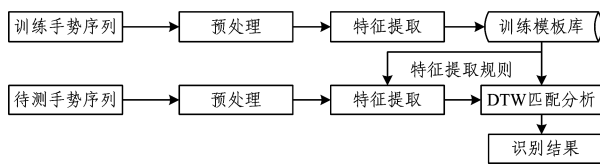


图 4 基于 DTW 算法的动态手势识别流程图

Fig. 4 Flowchart of DTW algorithm based on dynamic gesture recognition

然而,DTW 方法没有采用统计模型框架进行训练,也不容易将上下文的各种知识用于图像识别算法中,因此它在解决大数据量、复杂手势、组合手势识别等问题时较模式识别处于劣势。很多文献针对原始的 DTW 算法计算量大的问题进行优化,比如文献^[50]提出了一种改进的动态时间规整(IDTW)动态手势识别算法,对传统的 DTW 算法的搜索路径进行约束与特征加权,在公开的手势数据集上进行测试。结果表明,所提算法减少了运算量,同时也提高了整体识别率。

6.2 支持向量机

支持向量机在小样本、非线性,以及高维的模式识别的问

题方面的应用较广。而 SVM 关键的步骤为在一个高维的空间内,追求最优化的超平面,此高维的空间通过样本或者特征的空间非线性的变换获得。文献^[64]利用典型相关分析(CCA),通过 HOG3D 描述器对时空兴趣点提取的特征和人体行为构成的 4D 法向量直方图进行降维、融合,通过支持向量机分类器对新特征进行测试分类,最后设计了一个骨骼关节信息的交通手势识别系统。文献^[65]提出了一种 C3D 模型以提取时空特征,通过线性 SVM 分类器在 Chalearn LAP IsoGD 数据集上进行了验证,获得了最先进的效果。在静态手势的识别中,支持向量机的分类方法的应用较多^[66]。

6.3 隐马尔科夫模型

1995 年,Starmer 等^[67]提出了一种基于隐马尔科夫模型的美手语识别方法。隐马尔科夫是比较常用的一种基于状态空间的方法,也是一种成熟的匹配时变数据的数学方法,其主要涉及训练和分类两个阶段。用于解决训练问题的基本方法是 Baum-Welch 算法,而用于分类的基本方法是 Viterbi 算法。在动态手势识别中,每种手势对应一个隐马尔科夫模型。提取代表某手势的特征向量作为训练 HMM 模型的观察序列,使用前向、后向算法训练出 HMM 模型。在进行手势识别时,先对待测手势提取特征向量,使用前向算法分别求出每个 HMM 模型产生该手势的序列的概率,概率最大的 HMM 模型对应的手势即为识别出的结果。

HMM 可以应用于监督学习模式识别任务。HMM 的训练过程由来自特定系统的输出序列(训练序列)组成。训练算法可以通过调整 HMM 的参数,然后从被建模的系统中得到新的观察序列作为 HMM 的输入,输出的结果则取决于模型产生的概率。文献^[68]选择一个主动的基于差分特征的特征描述符,该描述符在传统的几何特征中包含一个序列中的时间扭曲信息。其提出一种基于 HMM 的手势识别方案,可用于开发能够提供增强性能的改进的 HCI 系统。并且对于所有的手势,所提出的方案已经达到了接近 90% 的识别率。文献^[69]对于自闭症典型的刻板姿势,研究了两种机器学习算法(HMM 和 SVM)之间的表现,对 RGB-D 相机获得的某些关节的取向数据顺序进行分析。将这两种机器学习算法的结果与最先进的技术相比较,结果表明其提出的 HMM 方法的平均识别率为 98.89%,该值比 SVM 方法识别的准确率高。HMM 需要大量的训练样本才能得到事物的统计特征,训练的样本越多,HMM 的性能越好。

6.4 基于粒子滤波的方法

1993 年,Gordon 等^[70]提出了一种新的基于顺序重要性采样(SIS)方法的 Bootstrap 非线性滤波方法,奠定了粒子滤波算法的基础。粒子滤波^[71](Particle Filter, PF)是一种基于蒙特卡洛模拟的非线性滤波方法,其核心思想是用随机采样的粒子表达概率密度分布。第一个 PF 在计算机视觉、可视化跟踪领域被称为凝聚算法(Condensation)。粒子滤波的 3 个重要步骤为:1)粒子采样,从建议分布中抽取一组粒子;2)粒子加权,根据观测概率分布、重要性分布以及贝叶斯公式计算每个粒子的权值;3)估计输出,输出系统状态的均值协方

差等。此外,为了应对粒子退化现象,粒子滤波算法还采用了重采样等策略。基于粒子滤波的方法在动态手势系统中可以预测位置,实现指尖跟踪^[72],并且可以进行有效的手势识别^[73]。文献^[74]基于 Icondensation 算法提出融合颜色与深度信息进行重要采样,结合轮廓信息并通过 ASM(Action Shape Model)模型的能量函数进行观察,根据最大后验概率对预定义的手势进行识别。粒子滤波算法对于复杂背景或噪音具有较好的鲁棒性,但是它需要使用大量的样本才能很好地近似系统的后验概率密度。

6.5 有限状态机

有限状态机^[75]是表示有限多个状态以及这些状态之间转移和动作的数学模型。它既可以用作接近和解决问题的开发工具,也可用作开发人员和系统维护人员描述解决方案的正式方法。通常,模型的训练是离线完成的,其使用每个手势的许多可能的例子作为训练数据,并且导出 FSM 中每个状态的参数(标准或特征)。手势的识别可以使用训练的 FSM 在线执行。当输入数据(诸如轨迹的特征向量)被提供给手势识别器时,手势识别器基于输入数据的参数决定停留在 FSM 的当前状态还是跳转到下一个状态。如果它达到最终状态,则一个手势已被识别。文献^[76]提出一种基于 FFMSM 和遗传算法(GAs)的混合学习的自动方法来实现用户姿势的检测。模糊有限状态机(FFMSM)是经典有限状态机的扩展,其中状态和输入是通过模糊推理系统来定义和计算的,使得它们能够处理不精确和不确定的数据。

为解决动态手势识别过程中不同人对于同一动作的关键帧在时间序列上分布不一致的问题,张好^[77]利用有限状态机进行动态手势识别时添加了状态转换的时间约束,其采用自底向上的分层定义策略将手势应用中的各类功能性手势视为动态手势基本单元、运动轨迹与其他动态手势的组合,并利用有限状态机原理和场景深度数据对动态手势基本单元进行识别,其手势识别率达到 92% 左右。

6.6 神经网络

神经网络(即 NNs)又称人工神经网络,它的使用范围比

较广泛。其基本原理是,在计算机系统中模拟生物的神经体系,让计算机可自动进行学习和推理。其优势比较明显,即具有优良的自适应性以及容错性。深度学习的概念源自神经网络的研究,其目的在于建立模拟人脑进行分析学习的神经网络,模仿人脑的机制用于解释数据。深度学习技术在动态手势识别中具有应用^[78-79]。Heickal 等^[80]利用 Kinect 捕捉人体关键部位关节的运动特征作为特征矢量,通过 BP 神经网络和朴素贝叶斯两种方法进行分类,正确率分别达到 98.11% 和 88.84%。Sorice 等^[81]利用 Kinect 获取手势的颜色、深度和关节坐标,并通过神经网络进行识别,识别率达到 95%。中国科技大学的 Tang 等^[82]联合开发了一套利用深度神经网络的实时手势识别系统,对于特定的手语手势利用深度信念网络(DBN)进行识别,对于部分手势可达到 98.12% 的准确率。目前已经提出多种 3D CNN 应用于手势识别^[83-85]。

然而神经网络方法也有一些缺陷,即其训练过程冗长且不适合实时识别^[86]。文献^[87]描述了一种深度动态神经网络的多模式手势识别的新方法,提出了一种基于隐马尔科夫模型的半监督分层动态框架,同时进行手势分割和识别,其中骨架关节信息、深度信息和 RGB 图像是多模态的输入观测数据,这种纯数据驱动的方法在 ChaLean LAP 手势识别挑战中的 Jaccard 指数分数达到了 0.81。该方法的性能与各种先进的基于手势调整的特征的方法以及其他学习方法的表现相当,为深度学习技术以及进一步探索多模式时间序列数据提供了思路。

近两年,深度学习的方法在手势识别中的应用备受青睐,很多研究者对此进行了深入研究。文献^[88]提出了一种深度卷积神经网络(CNN),直接对图像中的手势进行分类,不需要任何分割或检测阶段就可以丢弃不相关的非手区域,设计的手势识别网络能够以独立于用户的方式实时分类 7 种手势。

表 1 总结了近几年主流的动态手势选取的特征以及识别方法,并且列举了对应情况下的手势平均识别率。

表 1 动态手势特征提取、常用方法以及识别率情况

Table 1 Situation of dynamic gesture feature extraction, common methods and recognition rate

识别方法	作者	年份	采用特征	应用数据集	识别率/%	备注说明
DTW	Chaaraoui ^[43]	2013	Skeletal data, contour features	MSR Action 3D	96.4	骨骼与轮廓特征的融合,实验证明融合特征比单一特征更有效
	Cai ^[89]	2017	Skeleton	/	96.5	采用加权距离和有限搜索路径来避免传统 DTW 中的大量计算,提高了识别性能
SVM	Zhang ^[48]	2013	H3DF	ASL	98.9	提出 H3DF 描述符可以有效地表示各种手势的三维形状和结构
	Oreifej ^[49]	2013	HON4D	MSR Gesture 3D	92.45	HON4D 描述符相比 HOG 描述符可以捕捉更复杂的关节形状-像素级的运动线索
	Ohn-Bar ^[90]	2014	HOG+HOG2	APPLICATION-GESTURE SETS	92.8	该研究的手势识别系统被证明可用于车载实时系统
SVM	Yan ^[91]	2015	手部关键特征点运动轨迹	DHA	91.75	结合肤色分割出手部范围,并寻找对应的关键特征点,最终获取手势关键特征点的轨迹。利用支持向量机对数据集中有关手势的数据进行识别和评估
	Marin ^[92]	2016	Fingertip angles, distances, elevations, 3D positions	ASL	96.5	展示了深度相机和 Leap Motion 传感器如何进行手势识别,并引入描述符进行两种不同分类器的实验测试

(续表)

识别方法	作者	年份	采用特征	应用数据集	识别率/%	备注说明
HMM	Camada ^[69]	2017	Orientations data of joints	/	98.89	通过 RGBD 相机获取关键关节的方位数据序列用于分析学习,验证了 HMM 方法比 SVM 以及其他学习方法的识别率更高
	Saha ^[68]	2017	difference signatures	/	90	所提框架已经针对由 60 个不同主题体现的 12 种不同手势进行了测试,对于所有手势已经获得了近 90% 的识别率,证明了框架具有实时应用的价值
	Yao ^[93]	2016	Direction angle, speed, position	/	95.67	数字 0-9 的动态手势识别
	Zhang ^[94]	2017	finger shape, palm normal vector, palm ball radius and palm displacement vector	/	95.6	提出的 HMM-CART 模型的动态手势识别方法的精度要高于传统的单通道 HMM
NNs	Sorce ^[81]	2013	information of color, depth and skeleton	/	95	提出一种基于神经网络的方法以实时检测手的姿态,识别手是否闭合
	Heickal ^[80]	2013	3D Joint position	/	98.11	提出一个实时 3D 全身运动手势系统,通过验证证明基于神经网络的分类器比朴素贝叶斯分类器的准确率更高
	Gasparrini ^[86]	2015	3D Joint position	/	95	从了解测试者健康状况的目的出发对摄入行为进行分析,并根据距离误差评估算法的性能
	Molchanov ^[85]	2016	color+depth+optical flow	Chalearn 2014 dataset	98.2	采用三维卷积神经网络根据多模式数据对动态手势进行同步检测和分类

7 动态手势识别的应用

到目前为止已经出现了很多手势识别相关的应用系统。近年来,手势识别技术被广泛应用于医学^[95]、军事装备训练^[96]、虚拟实验^[97]等领域。

Tuntakurn 等^[95]利用 Kinect 骨骼追踪技术实现了指尖跟踪算法,通过设计手势控制医学图像查看软件。文献^[98]提出基于深度信息特征及空间测量最短路径的手指识别算法,对于检测和识别到的手指,设置粒子滤波追踪手指的运动,并基于手势中手指的个数及名称确定相应的手势。其通过实验验证了手指识别算法和手势识别方法的有效性,并把手势识别应用于三维虚拟图书馆的场景中,驱动场景信息实现体感交互。文献^[3]提出了一种传感器融合方案,将 Kinect 设备的数据与高精度传感器相结合,这种方式也使手势识别系统更加精准、可靠。

手势识别的应用并不适用于所有的场景,更多是一些特定的场景。其在游戏方面的应用不断扩展,比如切水果、哈士奇等游戏已经引入了手势识别版本。

结束语 针对基于视觉的手势识别易受光照和复杂背景的影响,深度图像提供了一种替代的解决方案。深度图像可以打破普通光学摄像头的很多限制,如摆脱光照影响和忽略场景中肤色相近的物体。对于深度信息的手势识别,现阶段已有大量研究,有些研究的实验结果也达到了不错的效果,但更复杂、更重要的时空动态手势识别仍然是一个非常具有挑战性的课题。

1) 动态手势的特征选择

对于动态手势,有运动变化较大的手势,也有局部细节变化的手势,因此选取的特征并不一定全都适用于设定的手势。对于动态手势的特征选择问题,我们可以研究更好的算法或者多特征实时选择的方法,来提高动态手势识别率。

2) 复杂手势的识别

很多静态手势和简单的动态手势识别研究都已经取得了

不错的成果,但对于一些复杂的动态手势仍存在一些困难,原因有:①多手势的相似性。不同的人做同一个手势会有很大差别,而且不同动作之间具有相似性,这种识别具有挑战性。②手势数据的冗余性。采集到手势的特征数据序列中所有的信息不一定都是有贡献的。③人的位置。在进行手势识别时,基本上要求人以及人手正在对深度传感器进行测试,这在很大程度上限制了手势的多样性。当手垂直摄像头或者偏离一定角度时,手势信息将有一定的丢失,这在很大程度上影响了识别率。

深度学习在动态手势识别中具有很大的发展前景,使基于深度信息的动态手势有着更多的可能性。要在动态手势的手势分割、特征提取与手势识别等阶段提升效果,需要人们不断地尝试、创新和完善技术。

参考文献

- [1] RAUTARAY S S, AGRAWAL A. Vision based hand gesture recognition for human computer interaction: a survey [M]. Kluwer Academic Publishers, 2015.
- [2] BHOWMICK S, TALUKDAR A K, SARMA K K. Continuous hand gesture recognition for English alphabets [C] // International Conference on Signal Processing and Integrated Networks. IEEE, 2015: 443-446.
- [3] PALYS T, ZORSKI W. Enhanced movement tracking with Kinect supported by high-precision sensors [C] // Computer Science and Information Systems. IEEE, 2015: 883-888.
- [4] LIN H B, WANG S B, ZHANG Y, et al. Gesture segmentation and positioning based on improved depth information [J]. Journal of Computer Applications, 2017, 37(1): 251-254. (in Chinese)
林海波, 王圣彬, 张毅, 等. 基于改进深度信息的手势分割与定位 [J]. 计算机应用, 2017, 37(1): 251-254.
- [5] LEITE D A T Q, DUARTE J C, OLIVEIRA J C, et al. A System to Interact with CAVE Applications Using Hand Gesture

- Recognition from Depth Data[C]// Virtual and Augmented Reality. IEEE,2014:246-253.
- [6] JOJIC N,HUANG T,BRUMITT B, et al. Detection and estimation of pointing gestures in dense disparity maps[C]// IEEE International Conference on Automatic Face and Gesture Recognition. IEEE,2000:468-475.
- [7] DAWOOD M,GIGENGACK F,JIANG X, et al. First International Symposium on 3D Data Processing Visualization and Transmission[C]// International Symposium on 3d Data Processing Visualization and Transmission,2002. IEEE,2002.
- [8] LIU X,FUJIMURA K. Hand gesture recognition using depth data[C]// IEEE International Conference on Automatic Face and Gesture Recognition. IEEE,2004:529-534.
- [9] RAUTARAY S S,AGRAWAL A. Vision based hand gesture recognition for human computer interaction;a survey [M]. Kluwer Academic Publishers,2015.
- [10] SUAREZ J,MURPHY R R. Hand gesture recognition with depth images:A review[C]// RoMan. IEEE,2012:411-417.
- [11] ZHENG Y J,YANG J. Application of binocular vision system based on one camera in computer vision [J]. Infrared & Laser Engineering,2004,33(4):392-396.
- [12] XIANG X,PAN Z. Depth Camera in Computer Vision and Computer Graphics; An Overview[J]. Journal of Frontiers of Computer Science & Technology,2011,5(6):481-492.
- [13] CUI Z L. Real-time gesture recognition based on RGBD depth images[D]. Suzhou:University of Suzhou,2016. (in Chinese)
崔子璐. 基于 RGBD 深度图像的实时手势识别研究[D]. 苏州:苏州大学,2016.
- [14] MATYUNIN S,VATOLIN D,BERDNIKOV Y, et al. Temporal filtering for depth maps generated by Kinect depth camera[C]// 3DTV Conference: the True Vision-Capture, Transmission and Display of 3d Video. IEEE,2011:1-4.
- [15] GASPARRINI S,CIPPITELLI E,SPINSANTE S, et al. A depth-based fall detection system using a Kinect® sensor [J]. Sensors,2014,14(2):2756-2775.
- [16] LI M,SUN L,HUO Q. Precise hand segmentation from a single depth image[C]// International Conference on Pattern Recognition. IEEE,2017:2398-2403.
- [17] CHEN D,LI G,SUN Y, et al. Fusion hand gesture segmentation and extraction based on CMOS sensor and 3D sensor [J]. International Journal of Wireless & Mobile Computing,2017,12(3):305-313.
- [18] HAN J,SHAO L,XU D, et al. Enhanced computer vision with Microsoft Kinect sensor;a review [J]. IEEE Transactions on Cybernetics,2013,43(5):1318-1334.
- [19] WU X,YANG C,WANG Y, et al. An Intelligent Interactive System Based on Hand Gesture Recognition Algorithm and Kinect[C]// Fifth International Symposium on Computational Intelligence and Design. IEEE,2013:294-298.
- [20] CAO H B. Based on Kinect depth information of static gesture recognition method [D]. Jinan:Shandong University,2016. (in Chinese)
曹海波. 基于 Kinect 深度信息的静态手势识别方法研究[D]. 济南:山东大学,2016.
- [21] LIN H,WANG S,ZHANG Y. Gesture segmentation and positioning based on improved depth information [J]. International Journal of Science,2017,4(3):176-184.
- [22] ZHAO M,JIA Q. Hand Segmentation Using Randomized Decision Forest Based on Depth Images[C]// International Conference on Virtual Reality and Visualization. IEEE,2017:110-113.
- [23] RONG T. Continuous gesture segmentation and identification methods and applications [D]. Nanjing: Nanjing University,2017. (in Chinese)
荣涛. 连续手势分割与识别方法的研究和应用[D]. 南京:南京大学,2017.
- [24] LI Y. Hand gesture recognition using Kinect [C]// Proceedings of 2012 IEEE 3rd International Conference on Software Engineering and Service Science. 2012:196-199.
- [25] WEN Y,HU C,YU G, et al. A robust method of detecting hand gestures using depth sensors [C]// IEEE International Workshop on Haptic Audio Visual Environments and Games. IEEE,2012:72-77.
- [26] HUANG-CHIA S,CHANG-HSIAN M. Hand segmentation with skin color fine-tuning using kinect depth sensor [C]// IEEE 6th Global Conference on Consumer Electronics. 2017:1-2.
- [27] WANG Y. Research on gesture recognition based on RGB-D depth information [D]. Chengdu: Xihua University,2015. (in Chinese)
王艺. 基于 RGB-D 深度信息的手势识别研究[D]. 成都:西华大学,2015.
- [28] JU Z,WANG Y,ZENG W, et al. A modified EM algorithm for hand gesture segmentation in RGB-D data[C]// IEEE International Conference on Fuzzy Systems. IEEE,2017:1736-1742.
- [29] QING L I,TANG H,CHI J N, et al. Gesture Segmentation with Improved Maximum Between-cluster Variance Algorithm [J]. Acta Automatica Sinica,2017,43(4):528-536.
- [30] SONI U,TRIVEDI A,ROBERTS N. Real-time hand tracking using integrated optical flow and CAMshift algorithm[C]// Second International Conference on Research in Computational Intelligence and Communication Networks. IEEE,2017:135-140.
- [31] ZHANG Z,WEI S,SONG Y, et al. Gesture Recognition Using Enhanced Depth Motion Map and Static Pose Map[C]// IEEE International Conference on Automatic Face & Gesture Recognition. IEEE,2017.
- [32] TRAN T H,VO T H,TRAN D T, et al. Dynamic hand gesture recognition using RGB-D motion history and kernel descriptor [C]// International Conference on Advanced Technologies for Communications. IEEE,2014:268-273.
- [33] QIN S,YANG Y,JIANG Y. Gesture recognition from depth images using motion and shape features[C]// International Symposium on Instrumentation and Measurement, Sensor Network and Automation. IEEE,2014:172-175.
- [34] REN Z,YUAN J,ZHANG Z. Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth-camera [C]// ACM International Conference on Multimedia. ACM,2011:1093-1096.
- [35] SINGHA J,LASKAR R H. Hand gesture recognition using

- two-level speed normalization, feature selection and classifier fusion[J]. *Multimedia Systems*, 2017, 23: 1-16.
- [36] BERGH M V D, GOOL L V. Combining RGB and ToF cameras for real-time 3D hand gesture interaction[C]// *IEEE Workshop on Applications of Computer Vision*. IEEE Computer Society, 2011: 66-72.
- [37] CHONG Y, HUANG J, PAN S. Hand Gesture Recognition Using Appearance Features Based on 3D Point Cloud[J]. *Journal of Software Engineering & Applications*, 2016, 9(4): 103-111.
- [38] WU L Y, WEI S N, ZHOU B B, et al. Hierarchical extreme learning machine gesture recognition method based on PCA dimension reduction [J]. *Electronic Measurement Technology*, 2017, 40(3): 82-88. (in Chinese)
吴良圆, 魏书宁, 周棒棒, 等. 基于 PCA 降维的分层超限学习机手势识别方法[J]. *电子测量技术*, 2017, 40(3): 82-88.
- [39] CHAI X, LIU Z, YIN F, et al. Two streams Recurrent Neural Networks for Large-Scale Continuous Gesture Recognition[C]// *International Conference on Pattern Recognition*. IEEE, 2017: 31-36.
- [40] HE F Z. Research on gesture recognition based on Kinect depth information [D]. Wuhan: Huazhong University of Science and Technology, 2015. (in Chinese)
贺芳姿. 基于 Kinect 深度信息的手势识别研究[D]. 武汉: 华中科技大学, 2015.
- [41] JMAA A B, MAHDI W, JEMAA Y B, et al. Arabic sign language recognition based on HOG descriptor[J]. *Proceedings of the Spie*, 2017, 225: 102250H.
- [42] YASIR F, PRASAD P W C, ALSADOON A, et al. SIFT based approach on Bangla sign language recognition[C]// *IEEE, International Workshop on Computational Intelligence and Applications*. IEEE, 2016: 35-39.
- [43] CHAARAOUI A A, PADILLALOPEZ J R, FLOREZRE-VUELTA F. Fusion of Skeletal and Silhouette-Based Features for Human Action Recognition with RGB-D Devices[C]// *The Workshop on Consumer Depth Cameras for Computer Vision*. IEEE, 2013: 91-97.
- [44] REN Z, YUAN J, MENG J, et al. Robust Part-Based Hand Gesture Recognition Using Kinect Sensor [J]. *IEEE Transactions on Multimedia*, 2013, 15(5): 1110-1120.
- [45] JU Z, GAO D, CAO J, et al. A novel approach to extract hand gesture feature in depth images [J]. *Multimedia Tools & Applications*, 2016, 75(19): 11929-11943.
- [46] WANG C, CHAN S C. A new hand gesture recognition algorithm based on joint color-depth Superpixel Earth Mover's Distance[C]// *International Workshop on Cognitive Information Processing*. IEEE, 2014: 1-6.
- [47] FENG Z Q, LIU H. Hand Gesture Recognition Using Binomial Coding and Class-Hausdorff Distance[J]. *Acta Tribune of Electronics*, 2017, 45(9): 2281-2291. (in Chinese)
冯志全, 刘弘. 结合手势二进制编码和类-Hausdorff 距离的手势识别[J]. *电子学报*, 2017, 45(9): 2281-2291.
- [48] ZHANG C, YANG X, TIAN Y L. Histogram of 3D facets: a characteristic descriptor for hand gesture recognition [C]// *Proceeding of 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition(F-G)*. IEEE, 2013: 1-8.
- [49] OREIFEJ O, LIU Z. HON4D: Histogram of Oriented 4D Normals for Activity Recognition from Depth Sequences[C]// *Computer Vision and Pattern Recognition*. IEEE, 2013: 716-723.
- [50] ZHANG J R. Study on Somatosensory Interaction Technology of Virtual Environment Based on Kinect Gesture Recognition [D]. Chongqing: Chongqing University of Posts and Telecommunications, 2016. (in Chinese)
张建荣. 基于 Kinect 手势识别的虚拟环境体感交互技术研究[D]. 重庆: 重庆邮电大学, 2016.
- [51] LIN J, RUAN X, YU N, et al. One-shot learning gesture recognition based on improved 3D SMOsIFT feature descriptor from RGB-D videos [C]// *Control and Decision Conference (CCDC)*, 2015 27th Chinese. IEEE, 2015: 4911-4916.
- [52] GAO Z. Gesture recognition based on multi-spatial feature fusion [J]. *Microsystems*, 2016, 37(7): 1577-1582. (in Chinese)
高喆. 多重空间特征融合的手势识别[J]. *小型微型计算机系统*, 2016, 37(7): 1577-1582.
- [53] LIANG C, QI L, CHEN E, et al. Depth-based action recognition using multiscale sub-actions depth motion maps and local autocorrelation of space-time gradients [C] // *IEEE, International Conference on Biometrics Theory, Applications and Systems*. IEEE, 2016: 1-7.
- [54] GONG D Y, HUANG M, ZHANG H B, et al. An adaptive feature selection method for human behavior recognition in RGBD [J]. *Journal of Intelligent Systems*, 2017, 12(1): 1-7. (in Chinese)
龚冬颖, 黄敏, 张洪博, 等. RGBD 人体行为识别中的自适应特征选择方法[J]. *智能系统学报*, 2017, 12(1): 1-7.
- [55] BERNFT D, CLIHORD J. Using Dynamic Time Warping to Find Patterns in Sequences [C] // *International Conference on Knowledge Discovery and Data Mining*. AAAI Press, 1994: 359-370.
- [56] FUREY T S, CRISTIANINI N, DUFFY N, et al. Support vector machine classification and validation of cancer tissuesamples using microarray expression data [J]. *Bioinformatics*, 2000, 16(10): 906.
- [57] FINE S, SINGER Y, TISHBY N. The Hierarchical Hidden Markov Model: Analysis and Applications [M]. *Machine Learning*, 1998, 32 (1): 41-62
- [58] HAYKIN S. *Neural Networks: A Comprehensive Foundation* [J]. *Neural Networks*, 2004, 2(2004): 41.
- [59] WU T, DING X. Video object tracking using improved chamfer matching and condensation particle filter [C] // *Proceedings of SPIE-The International Society for Optical Engineering*. 2008.
- [60] HONG P, HUANG T S, TURK M. Gesture modeling and recognition using finite state machines [C] // *IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE, 2002: 410-415.
- [61] PLOUFFE G, CRETU A M. Static and Dynamic Hand Gesture

- Recognition in Depth Data Using Dynamic Time Warping [J]. *IEEE Transactions on Instrumentation & Measurement*, 2016, 65(2):305-316.
- [62] CORRADINI A. Dynamic Time Warping for Off-Line Recognition of a Small Gesture Vocabulary[C]//*IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems*. IEEE, 2001:82-89.
- [63] HUANG D, YANG J. Temporal Gesture Recognition Based on Improved N-Dimensional Dynamic Time Warping[D]. *Electronic Science & Technology*, 2017.
- [64] ZHANG Y, ZHANG S, LUO Y, et al. Gesture track recognition based on Kinect depth image information and its applications [J]. *Application Research of Computers*, 2012, 29(9): 3547-3550.
- [65] LI Y, MIAO Q, TIAN K, et al. Large-scale gesture recognition with a fusion of RGB-D data based on the C3D model[C]//*International Conference on Pattern Recognition*. IEEE, 2017: 25-30.
- [66] GHOSH D K, ARI S. Static Hand Gesture Recognition Using Mixture of Features and SVM Classifier [C] // *International Conference on Communication Systems & Network Technologies*. IEEE, 2015: 1094-1099.
- [67] STARNER T, PENTL A. Visual Recognition of American Sign Language Using Hidden Markov Models [C] // *International Workshop on Automatic Face & Gesture Recognition*. 1995: 189-194.
- [68] SAHA S, LAHIRI R, KONAR A, et al. HMM-based gesture recognition system using kinect sensor for improvised human-computer interaction [C] // *International Joint Conference on Neural Networks*. IEEE, 2017: 2776-2783.
- [69] CAMADA M Y O, CERQUEIRA J J F, LIMA A M N. Stereotyped gesture recognition: An analysis between HMM and SVM [C]//*IEEE International Conference on Innovations in Intelligent Systems and Applications*. IEEE, 2017: 328-333.
- [70] GORDON N J, SALMOND D J, SMITH F M A. Novel approach to non-linear and non-gaussian Bayesian state estimation [J]. *IEE Proceedings F-Radar and Signal Processing*, 1993, 140(2): 107-113.
- [71] ISARD M, BLAKE A. CONDENSATION—Conditional Density Propagation for Visual Tracking [J]. *International Journal of Computer Vision*, 1998, 29(1): 5-28.
- [72] HUANG L. Particle filter algorithm and its application in fingertip tracking [D]. Xiangtan: Xiangtan University, 2015. (in Chinese)
黄龙. 粒子滤波算法及其在指尖跟踪中的应用研究[D]. 湘潭: 湘潭大学, 2015.
- [73] SUN R B. Research on gesture recognition technology based on particle filter [D]. Nanjing: Nanjing University of Science and Technology, 2009. (in Chinese)
孙荣斌. 基于粒子滤波的手势识别技术研究[D]. 南京: 南京理工大学, 2009.
- [74] LIU F W, DING G Y, LI S Q, et al. An Algorithm of Human Hand Tracking and Gesture Recognition Based on ICONDENSATION Algorithm [J]. *Journal of Beijing Institute of Technology*, 2007, 27(12): 1069-1072. (in Chinese)
刘法旺, 丁刚毅, 李善青, 等. 基于 ICONDENSATION 算法的人手跟踪与手势识别算法 [J]. *北京理工大学学报*, 2007, 27(12): 1069-1072.
- [75] RAMEY A, GONZALEZ-PACHECO V, SALICHS M A. Integration of a low-cost RGB-D sensor in a social robot for gesture recognition [C] // *ACM/IEEE International Conference on Human-Robot Interaction*. IEEE, 2012: 229-230.
- [76] ALVAREZ-ALVAREZ A, TRIVINO G, CORDÓN O. Body posture recognition by means of a genetic fuzzy finite state machine [C] // *IEEE, International Workshop on Genetic and Evolutionary Fuzzy Systems*. IEEE, 2011: 60-65.
- [77] ZHANG H. Kinect-based gesture recognition and its application [D]. Dalian: Dalian University of Technology, 2014. (in Chinese)
张好. 基于 Kinect 的手势识别及其应用 [D]. 大连: 大连理工大学, 2014.
- [78] OTBERDOUT N, BALLIHI L, ABOUTAJDINE D. Hand pose estimation based on deep learning depth map for hand gesture recognition [C] // *Intelligent Systems and Computer Vision (ISCV)*. IEEE, 2017: 1-8.
- [79] ZHANG L, LIANG B, JIANG A. Recent Advances of Deep Learning for Sign Language Recognition [C] // *International Conference on Digital Image Computing: Techniques and Applications*. 2017: 1-7.
- [80] HEICKAL H, ZHANG T, HASANUZZAMAN M. Real-time 3D full body motion gesture recognition [C] // *IEEE International Conference on Robotics and Biomimetics*. IEEE, 2013: 798-803.
- [81] SORCE S, GENTILE V, GENTILE A. Real-Time Hand Pose Recognition Based on a Neural Network Using Microsoft Kinect [C] // *Eighth International Conference on Broadband and Wireless Computing, Communication and Applications*. IEEE, 2013: 344-350.
- [82] TANG A, LU K, WANG Y, et al. A Real-Time Hand Posture Recognition System Using Deep Neural Networks [J]. *ACM Transactions on Intelligent Systems & Technology*, 2015, 6(2): 1-23.
- [83] HUANG J, ZHOU W, LI H, et al. Sign Language Recognition using 3D convolutional neural networks [C] // *IEEE International Conference on Multimedia and Expo*. IEEE, 2015: 1-6.
- [84] MOLCHANOV P, GUPTA S, KIM K, et al. Hand gesture recognition with 3D convolutional neural networks [C] // *Computer Vision and Pattern Recognition Workshops*. IEEE, 2015: 1-7.
- [85] MOLCHANOV P, YANG X, GUPTA S, et al. Online Detection and Classification of Dynamic Hand Gestures with Recurrent 3D Convolutional Neural Networks [C] // *Computer Vision and Pattern Recognition*. IEEE, 2016: 4207-4215.
- [86] GASPARRINI S, CIPPITELLI E, GAMBI E, et al. Performance analysis of self-organising neural networks tracking algorithms for intake monitoring using Kinect [C] // *Iet International Conference on Technologies for Active and Assisted Living*. IET, 2015: 1-6.

- [5] HUANG X Q, SHAO S, WANG H, et al. The robustness of interdependent clustered networks[J]. *EPL*, 2012, 101(1): 18002-18007.
- [6] SHAO S, HUANG X Q, STANLEY H E. Robustness of partially interdependent network formed of clustered networks[J]. *Physical Review E*, 2014, 89(3): 032812.
- [7] LIU H, ZHOU G G, FU P H. Local Evolving Model Research of Layered Supply Chains Complex Networks [J]. *Computer Science*, 2013, 40(2): 270-273. (in Chinese)
柳虹, 周根贵, 傅培华. 分层供应链复杂网络局域演化模型研究[J]. *计算机科学*, 2013, 40(2): 270-273.
- [8] MENICHETTI G, ASTA L D, BIANCONI G. Network Controllability Is Determined by the Density of Low In-Degree and Out-Degree Nodes[J]. *Physical Review Letters*, 2014, 113(7): 078701.
- [9] CAO W B, XIONG X. Local-world Evolution Model of Complex Supply Chain Network Under Edge Benefit [J]. *Application Research of Computer*, 2016, 33(1): 75-81. (in Chinese)
曹文彬, 熊曦. 边效益因素下复杂供应链网络局域演化机制[J]. *计算机应用研究*, 2016, 33(1): 75-81.
- [10] XIE F J, CUI W T. Complex Structural Properties and Evolution Model of the Ground Express Network [J]. *Journal of Systems Management*, 2016, 25(2): 364-369. (in Chinese)
谢逢洁, 崔文田. 陆运快递网络的复杂结构特性及演化模型[J]. *系统管理学报*, 2016, 25(2): 364-369.
- [11] FU J Y, ZHANG J, XIONG J, et al. Research on the Weighted Local Evolution Model of Urban Logistics Network's Spatial Structure [J]. *Compound Systems and Complexity Science*, 2015, 12(3): 38-44. (in Chinese)
付江月, 张锦, 熊杰, 等. 城市物流网络空间结构加权局域世界演化模型研究[J]. *复杂系统与复杂性科学*, 2015, 12(3): 38-44.
- [12] WATTS J, STROGATZ S H. Collective Dynamics of Small-World' Networks[J]. *Nature*, 1998, 393(6684): 440-442.
- [13] 徐野. 复杂互联系统与网络鲁棒性研究[M]. 北京: 电子工业出版社, 2015: 57-58.
- [14] FU P H, LI J, LIU Y C. Cluster Supply Chain network Evolving Model Based on Degree and Path Preferential Attachment [J]. *Operations Research and Management*, 2013, 22(1): 120-125. (in Chinese)
傅培华, 李进, 刘燕楚. 基于度与路径优先连接的集聚型供应链网络演化模型[J]. *运筹与管理*, 2013, 22(1): 120-125.
- [15] GE W, QIN X H. Evolution Model of Supply Chain Network Based on Complex Network [J]. *Business Age*, 2012(32): 18-19. (in Chinese)
葛伟, 秦效宏. 基于复杂网络的供应链网络演化模型[J]. *商业时代*, 2012(32): 18-19.
- [16] 孙玺菁, 司守奎. 复杂网络算法与应用[M]. 北京: 国防工业出版社, 2015: 67-69.

(上接第 51 页)

- [87] WU D, PIGOU L, KINDERMANS P J, et al. Deep Dynamic Neural Networks for Multimodal Gesture Segmentation and Recognition [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2016, 38(8): 1583-1597.
- [88] BAO P, MAQUEDA A I, DEL-BLANCO C R, et al. Tiny hand gesture recognition without localization via a deep convolutional network[J]. *IEEE Transactions on Consumer Electronics*, 2017, 63(3): 251-257.
- [89] CAI L, CUI S, MIN X, et al. Dynamic hand gesture recognition using RGB-D data for natural human-computer interaction[J]. *Journal of Intelligent & Fuzzy Systems*, 2017, 32(5): 3495-3507.
- [90] OHN-BAR E, TRIVEDI M M. Hand Gesture Recognition in Real Time for Automotive Interfaces: A Multimodal Vision-Based Approach and Evaluations[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2014, 15(6): 2368-2377.
- [91] YAN L M, DU B, LI Y, et al. An extraction of key feature points and recognition of dynamic gesture trajectory based on depth data [J]. *Industrial Control Computer*, 2015, 28(11): 86-88. (in Chinese)
严利民, 杜斌, 李跃, 等. 基于深度数据的关键特征点提取及动态手势轨迹识别[J]. *工业控制计算机*, 2015, 28(11): 86-88.
- [92] MARIN G, DOMINIO F, ZANUTTIGH P. Hand gesture recognition with jointly calibrated Leap Motion and depth sensor [J]. *Multimedia Tools & Applications*, 2016, 75(22): 1-25.
- [93] YAO Y Y. Research and implementation of key technologies of dynamic gesture recognition[D]. Chongqing: Chongqing University of Posts and Telecommunications, 2016. (in Chinese)
姚圆圆. 动态手势识别的关键技术研究及实现[D]. 重庆: 重庆邮电大学, 2016.
- [94] ZHANG Q, DENG F. Dynamic Gesture Recognition based on LeapMotion and HMM-CART Model [J]. *Journal of Physics: Conference Series*, 2017, 910(1): 012037.
- [95] TUNTAKURN A, THONGVIGITMANEE S S, SA-ING V, et al. Natural interactive 3D medical image viewer based on finger and arm gestures [C] // *Biomedical Engineering International Conference*. IEEE, 2013: 1-5.
- [96] ABDELNABY M, ELAZEM M A, ALY H A, et al. Augmented Reality Maintenance Training with Intel Depth Camera [C] // *International Conference on Machine Vision and Information Technology*. IEEE, 2017: 116-122.
- [97] SONG L, WEN F. The Human-Computer Interaction Design and Research of Virtual Experiment [C] // *Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics*. IEEE, 2014: 50-53.
- [98] MOLCHANOV P, GUPTA S, KIM K, et al. Hand gesture recognition with 3D convolutional neural networks [C] // *Computer Vision and Pattern Recognition Workshops*. IEEE, 2015: 1-7.