

融合关节旋转特征和指尖距离特征的手势识别

缪永伟¹⁾ 李佳颖¹⁾ 刘家宗¹⁾ 陈佳舟²⁾ 孙树森¹⁾

¹⁾(浙江理工大学信息学院 杭州 310018)

²⁾(浙江工业大学计算机科学与技术学院 杭州 310023)

摘要 作为人机交互的重要方式,手势交互和识别由于其具有的高自由度而成为计算机图形学、虚拟现实与人机交互等领域的研究热点.传统直接提取手势轮廓或手部关节点位置信息的手势识别方法,其提取的特征通常难以准确表示手势之间的区别.针对手势识别中不同手势具有的高自由度以及由于手势图像分辨率低、背景杂乱、手被遮挡、手指形状尺寸不同、个体差异性导致手势特征表示不准确等问题,本文提出了一种新的融合关节旋转特征和指尖距离特征的手势特征表示与手势识别方法.首先从手势深度图中利用手部模板并将手部看成链段结构提取手部20个关节点的3D位置信息;然后利用手部关节点位置信息提取四元数关节旋转特征和指尖距离特征,该表示构成了手势特征的内在表示;最后利用一对一支持向量机对手势进行有效识别分类.本文不仅提出了一种新的手势特征表示与提取方法,该表示融合了关节旋转信息和指尖距离特征;而且从理论上证明了该特征表示能唯一地表征手势关节点的位置信息;同时提出了基于一对一SVM多分类策略进行手势分类与识别.对ASTAR静态手势深度图数据集中8类中国数字手势和21类美国字母手势数据集分别进行了实验验证,其分类识别准确率分别为99.71%和85.24%.实验结果表明,本文提出的基于关节旋转特征和指尖距离特征的融合特征能很好地表示不同手势的几何特征,能准确地表征静态手势并进行手势识别.

关键词 手势识别;人机交互;关节点位置;四元数特征;支持向量机

中图法分类号 TP391 DOI号 10.11897/SP.J.1016.2020.00078

Hand Gesture Recognition Based on Joint Rotation Feature and Fingertip Distance Feature

MIAO Yong-Wei¹⁾ LI Jia-Ying¹⁾ LIU Jia-Zong¹⁾ CHEN Jia-Zhou²⁾ SUN Shu-Sen¹⁾

¹⁾(College of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018)

²⁾(College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023)

Abstract As a mainstream and important interactive mode of human-computer interaction (HCI), due to its high degree of freedom, hand gesture interaction and gesture recognition become a hot research topic in the literature of Computer Graphics, Virtual Reality and HCI, etc. The traditional hand gesture recognition method which directly extracts the gesture contour or the position information of hand joints is usually difficult to represent the differences between hand gestures accurately. In order to solve the issues of high degree of freedom of different gestures and the inaccurate feature representation in gesture recognition due to low resolution of gesture images, messy image background, occlusive hands, different shape and size of fingers and also individual differences, a new feature representation of hand gesture and the corresponding gesture recognition method is proposed in this paper, which combines the joint rotation feature

收稿日期:2019-02-03;在线出版日期:2019-10-16. 本课题得到国家自然科学基金(61972458)、浙江省基础公益研究计划项目(LY18F020035, GG19F020006)和浙江理工大学科研基金(17032001-Y)资助. 缪永伟, 博士, 教授, 博士生导师, 中国计算机学会(CCF)会员, 主要研究领域为计算机图形学、数字几何处理、计算机视觉、机器学习. E-mail: ywmiao@zstu.edu.cn. 李佳颖, 硕士研究生, 主要研究方向为计算机图形学、机器学习. 刘家宗, 硕士研究生, 主要研究方向为计算机图形学、计算机视觉. 陈佳舟, 博士, 副教授, 主要研究方向为计算机图形学、计算机视觉. 孙树森, 博士, 副教授, 主要研究方向为计算机图形学、虚拟现实.

and fingertip distance feature. First, the 3D position information of 20 hand joints is calculated from the depth map of the underlying hand gesture by using hand template, and also considering the hand as a segment structures which consists of 19 segments. Then, the quaternion joint rotation feature and the fingertip distance features are extracted by using the position information of the hand joints, which constitutes the intrinsic representation of hand gesture feature. Finally, the hand gesture can be effectively recognized and classified by using one-to-one support vector machine (SVM) classifier. This paper not only presents a new feature representation and the corresponding feature extraction approach of static hand gesture, but also proves theoretically that the proposed gesture features can uniquely represent the 3D position information of hand joints. At the same time, a multi-classification strategy based on one-to-one SVM is also adopted to classify and recognize different hand gestures. We evaluate the proposed algorithm on the ASTAR dataset with annotated hand-depth images of static gestures, which including 8 kinds of Chinese digital gestures and 21 classes of American alphabet gestures. We compared our gesture recognition method with the existing methods from the following two aspects. One is for different feature representations of hand gesture, such as HOG (histogram of oriented gradients feature), SURF (speeded-up robust feature), Feature 1 (quaternions joint rotation with rotation axis + fingertip distance feature), Feature 2 (quaternions without rotation axis), and Feature 3 (quaternions without rotation axis + fingertip distance feature). The other is for different classifier selections for gesture classification and recognition, such as *K*-Nearest Neighbor (*KNN*), Naive Bayes (*NB*), one-to-rest SVM, one-to-one SVM. Our experiments show the efficiency and advantage of our proposed static gesture recognition method by using the Feature 3 + one-to-one SVM in terms of the accuracy of gesture classification and recognition. The accuracy of our hand gesture recognition method is 99.71% for Chinese digital gestures and 85.24% for American alphabet gestures, respectively. Experimental results demonstrate that the proposed feature representation based on joint rotation feature and fingertip distance feature can reflect the geometric properties of different hand gestures, and also can accurately represent the static gestures and effectively recognize different hand gestures.

Keywords hand gesture recognition; human-computer interaction; position of hand joints; quaternion feature; support vector machine

1 引言

作为计算机图形学、虚拟现实与人机交互中重要的交互方式,手势交互提供了一种方便直观、简单便捷的交互体验.手势交互与识别对于虚拟现实交互^[1]、3D体感游戏^[2]、辅助医疗手术^[3]、老年人护理监测^[4]等应用具有重要意义;但是,由于不同手势具有较高自由度^[5],获取的手势图像数据通常具有分辨率低、背景杂乱、手被遮挡、手指形状尺寸不同以及个体差异性等特点,使得难以准确表示不同手势特征,从而给手势识别带来困难与挑战^[6].

传统的手势识别通常是基于相机照片和2D手势进行识别与分类,如朱继玉等人^[7]对分割得到的

人手根据其不同部分在几何尺寸上的变化,从低到高逐次分析图像金字塔中各种分辨率的图像,以获取手势的全局和局部结构特征并应用于2D手势识别. Panwar^[8]利用普通相机拍照获取的彩色照片,首先分割出手的形状并提取手的质心位置,然后对每个手指进行检测并实现手势识别,但是该方法仅仅能够识别每个手指是否伸出,对于手指弯曲程度的不同或相似手势难以正确区分识别. 手势识别中的关键问题是手势特征的表示与提取. Ren等人^[9]利用手势形状提出了一种有效的距离度量 FEMD (Finger-Earth Mover's Distance),该度量能够比较不同手势的形状差异. 杨学文等人^[10]提出一种基于手势主方向和类 Hausdorff 距离模板匹配的手势识别方法. 该方法对于手势的主方向有较高限制,要求

获取的手势主方向与训练库中的相近手势主方向保持一致,限制了方法的适用性.

随着深度相机和深度传感器的出现,基于深度图像的 3D 手势分类与识别得到了研究者的广泛关注^[11].传统直接提取手势轮廓或手部关节位置信息的手势识别方法,其提取的特征通常难以准确表示手势之间的区别.为了克服不同手势具有的高自由度、特征表示不准确等局限性,本文提出一种新的融合关节旋转特征和指尖距离特征的手势特征表示和相应的静态手势分类识别方法.

本文主要贡献在于:(1)提出了一种新的手势特征表示与提取方法,该表示融合了关节旋转信息和指尖距离特征;(2)理论上证明了该特征表示能唯一地表征手势关节的位置信息;(3)采用了基于一对一 SVM 多分类策略进行手势的分类与识别.本文方法基于手部关节的 3D 位置信息,将手势视为一种链段结构,提取关节旋转特征作为局部特征表示,再利用指尖距离特征作为全局特征表示,充分考虑了不同手势的几何特征,并且该特征在手指间无相互作用力时能够唯一表征一个静态手势.在 8 类中国数字手势和 21 类美国字母手势数据集上进行的实验表明,本文的手势特征表示能唯一地表征手势关节的位置信息,在手势识别中具有很好的鲁棒性,附录中给出了该特征表示内在唯一性的详细证明.

本文第 2 节介绍手势识别的 3 种类型方法;第 3 节概述本文手势识别的方法流程;第 4 节详细解释关节 3D 位置信息获取的具体过程;第 5 节阐述手势关节旋转特征和指尖距离特征的提取过程;第 6 节阐述融合关节内在特征的手势分类和识别;第 7 节给出实验结果和分析,并将本文方法和已有方法进行比较讨论;最后一节给出结论和研究展望.

2 相关工作

针对手势的识别方法通常可以分为判别方法、生成方法以及混合 3D 模型方法^[6,11].

基于判别方法的手势识别中,将学习一个从手势深度图像到手势分类类别的映射,其中的关键问题是映射函数的选取. Qian 等人^[12]提出一种手跟踪系统,使用 48 个球面集模拟 26 个自由度的手几何模型,同时加上手的运动学限制,并利用随机梯度下降的方法实现快速收敛,但是该方法容易陷入局部最优,并不适用于非刚体情形.为解决手势识别中

噪声数据和遮挡问题, Tang 等人^[13]基于合成数据和真实数据,提出一种半监督转导回归森林 STR 算法 (Semi-supervised Transductive Regression forests) 并用来实时估计手势.通过转导学习,STR 森林能够从少量标注的真实数据中识别各种手势,但是该方法对于背景的变化较敏感.

基于生成方法的手势识别亦称为基于模型的方法,该类方法首先需要生成大量的手势数据,然后基于提取的手势特征从数据集中选取与当前手势最匹配的深度图像手势.然而,基于模型的方法通常会因拍摄角度的不同而导致两个相近手势由于其深度图像差别较大而判为不同手势. Tompson 等人^[14]提出了一种从单张深度图中恢复连续手势的方法,该方法利用卷积神经网络为手部关节位置生成热图,并利用关节位置热图和反向动力学推断出手部关节的位置信息,但是该方法无法处理手部关节被隐藏遮挡的情形.为解决手势识别中的自遮挡、纹理或阴影信息缺失等问题, de La Gorce 等人^[15]提出一种从单目视频中估计 3D 手势的方法.该方法首先用 18 个封闭可定向的三角面片建立手部模型,视频跟踪中最小化相邻两帧手势图像之间的误差,并在手部纹理保持静止的前提下,确定每帧的手势参数和光照参数并更新其手部纹理.该方法在手指被遮挡或手指之间交错穿插时,最小化过程容易收敛到一个较差的局部极小值.

基于混合 3D 模型的方法则结合判别方法和生成方法. Ye 等人^[16]提出一种混合迭代改进的手势估计方法.先利用判别方法在深度神经网络的特征空间和输出层中应用分层策略生成手势;再利用生成方法在层次结构中各层次之间采用粒子群算法并加入运动学约束用于纠正手势估计结果. Xu 等人^[17]提出了一个从深度图像中估计手势的 Dhand 方法.该方法先利用 Shotton 等人^[18]基于部件的人体姿态识别方法与 Girshick 等人^[19]的 Hough 森林判别回归模型得到初始的手势估计,再利用由生成方法得到的 27 个自由度的手部模型进行验证,结果表明通过结合判别方法与生成方法可以较好地估计手势.

此外,为了克服手势识别结果易受实验环境光照的影响, Feng 等人^[20]先对手势深度图提取梯度方向直方图 HOG 特征,再利用 SVM 分类器训练特征向量.然而由于梯度提取的特点,该特征描述其对噪声较敏感且难以处理遮挡问题,导致对不同手势的分类效果并不理想. Bagdanov 等人^[21]提出 SURF

特征与 SVM 分类器相结合的手势识别方法. 该方法对传统 SIFT 特征进行了改进, 并采用 Hessian 矩阵获取图像局部极值, 但是由于该特征过于依赖局部邻域像素的梯度方向导致其主方向计算并不准确, 而其特征向量提取及匹配则严重依赖于主方向, 故该方法难以进行准确的特征匹配.

为了更好地进行手势特征提取和手势分类识别, 克服不同手势具有的高自由度、传统方法对手势特征表示的不准确等缺陷, 需要提出能准确表示手势之间本质区别的特征表示与特征提取方法. 本文提出了一种新的融合关节旋转特征和指尖距离特征的手势特征表示和相应的静态手势分类识别方法, 并从理论上证明了该特征表示能唯一地表征手势关节的位置信息.

3 本文手势识别方法

考虑到手势 RGB-D 图像中提取的特征存在视角特点的影响, 本文从手势的 3D 关节位置信息中提取手势特征. 输入静态手势 RGB-D 数据中的深度图, 根据手部模型的特点获取手部 20 个关节的位置信息, 根据手部的解剖特性, 将手部看作是由 19 段链段构成的链段结构. 本文利用手部关节旋转特征和指尖距离特征来唯一刻画静态手势, 该方法既融合了局部特征和全局特征, 又充分考虑了手势的解剖约束和几何特性, 手势表示不仅形象直观而且计算高效. 本文手势识别方法的具体流程如下:

(1) 手势关节 3D 位置信息的获取预处理.

输入手势 RGB-D 图像中的深度图, 通过手势模型得到关节的 3D 位置信息, 由于具有生物差异性的特点, 不同的人其手指关节长度不一样, 而且所处的局部坐标系也随相机位置改变而改变. 所以需要对手指关节长度进行长度归一化处理, 并将手腕关节设置为原点建立坐标系, 消除相机位置对手势特征提取的影响.

(2) 融合关节旋转特征和指尖距离特征的手势特征表示与提取.

在手指关节长度一定的条件下, 手势的不同源于手指关节之间的旋转信息的区别. 在特征提取中, 首先将手势关节之间的四元数作为旋转信息特征, 通过将手指关节段表示为向量并考虑两个向量之间的旋转, 提取四元数并利用四元数中的角度表示局部手指的特征. 为限制角度信息带来的累积误差, 使用 5 个指尖相对于手腕的距离特征和相邻指尖之间的

距离特征作为全局特征. 本文将关节旋转特征和指尖距离特征进行融合作为手势内在特征的一种表示.

(3) 基于特征表示的一对一 SVM 多分类手势识别.

利用融合关节旋转特征和指尖距离特征的手势内在特征表示, 可以采用一对一 SVM 多分类策略对数据集中的手势进行识别分类. 对于待识别的手势, 通过步骤(1)和(2)进行数据预处理和特征提取, 然后使用同一个 SVM 模型识别该手势属于数据集中哪一个类别, 并输出识别结果.

具体的手势识别算法如下:

算法 1. 基于内在特征表示的手势识别.

输入: 静态待识别手势的单张 RGB-D 深度图

输出: 待识别手势对应于数据集中的类别

1. 根据手势的单张深度图计算关节点相对位置信息 W_0, \dots, W_{19} ;
2. 对关节长度进行归一化, 得到归一化后关节点位置信息:

$$\bar{W}_i \leftarrow \text{normalize}(W_i), i=0, 1, 2, \dots, 19;$$

3. 获取关节旋转特征四元数:

$$Q_i \leftarrow \text{quaternion}(\bar{W}_i), i=0, 1, 2, \dots, 13;$$

4. 计算手指指尖距离特征:

$$S_0, S_1, S_2, S_3, D_0, D_1, D_2, D_3, D_4;$$

5. 对关节旋转特征和指尖距离特征的特征进行融合 $F = [Q_0, \dots, Q_{13}, S_0, \dots, S_3, D_0, \dots, D_4]$;

6. 对内在特征信息进行归一化:

$$\bar{F} \leftarrow \text{norm}(F);$$

7. 利用提供的手势数据集, 采用一对一 SVM 多分类策略识别输出其手势类别.

4 手势关节 3D 位置信息的获取预处理

根据静态手势 RGB-D 数据中深度图的输入, 本文采用由 20 个关节组成的铰链式手势模型获取手部各关节的位置信息, 同时为了使手势特征描述真正反映手势的特征, 需要对关节位置信息进行归一化预处理如下:

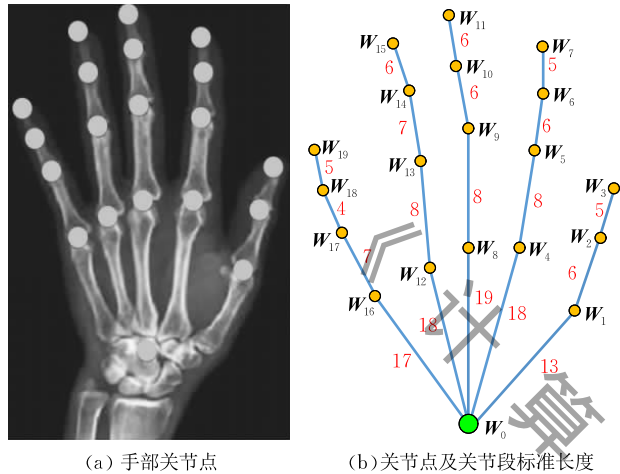
首先, 根据输入的静态手势 RGB-D 深度图, 将手势看作由骨架链构成的链段结构, 采用 Zhou 等人^[22]的类似手势模型得到手势 3D 关节位置信息.

其次, 由于得到的手部关节位置信息相对于摄像机坐标系, 其将依赖于相机的运动. 为了对不同手势的特征描述相对于不同个体的手指长短变化保持不变, 需要采用关节的相对位置表示, 并在数据预处理中对关节段长度进行归一化处理.

4.1 关节点相对位置表示

如图 1 所示,图 1(a)为人手 20 个关节点,图 1(b)为关节点位置及各关节段标准长度.图中用圆点标记出 20 个关节点的位置,利用 w_i 表示第 i 个关节点位置, w_0 表示根关节点位置.由于静态手势的具有平移不变性,本文通过减去定义的根本关键点位置坐标来获取关节的相对三维坐标.各关节相对位置表示为

$$W_i = w_i - w_0, \quad i = 0, 1, 2, \dots, 19.$$



(a) 手部关节点

(b) 关节点及关节段标准长度

图 1 手部关节点及各关节段标准长度

4.2 手部关节长度归一化

为了排除手的个体差异性,将手部关节段进行长度归一化处理:将关节段的每一段长度规定到各自的规范长度,不改变关节段的方向(即关节对之间的角度).在一些利用神经网络,或者深度学习的论文中,考虑将手部各关节长度归一化成单位长度,Zimmermann 等人^[23]将所有的手部关节对都归一化成单位长度.但是这样的处理会使得所有手指长度都一样,虽然方便计算,但是部分手势会失真,例如:当考虑握拳手势时,由于指尖到距离指尖最近的关节点的距离过长,会造成指尖穿过手掌的异常情况,不符合手的物理特性,而且在计算关节相对距离时将造成干扰,使得对于一些手势的区分度变弱. De Smedt 等人^[24]将手指长度归一化为数据集中手势关节长度的平均长度,该方法增加了计算量,本文将手指长度归一化成一个标准手指的长度,各关节段对应标准长度如图 1(b)中的数字所示.

手部关节长度归一化过程中,利用 19 个向量表示 20 个关节点构成的 19 个关节对如下:

$$V_i = \begin{cases} W_i - W_{i-1}, & 1 \leq i \leq 19, i \neq 4, 8, 12, 16 \\ W_i - W_0, & i = 4, 8, 12, 16 \end{cases},$$

则进行归一化为

$$\bar{V}_i = L_i \frac{V_i}{\|V_i\|},$$

$$\bar{W}_i = \begin{cases} W_0, & i = 0 \\ \bar{V}_i + W_{i-1}, & 1 \leq i \leq 19 \text{ 且 } i \neq 4, 8, 12, 16. \\ \bar{V}_i, & i = 4, 8, 12, 16 \end{cases}$$

本文的归一化方法基于标准手指长度进行,标准手指长度参考 ACT hand 关节段^[24]确立长度,其中 L_i 为对应第 i 节关节段标准长度,如图 1(b)所示.以中国数字 1 手势为例,进行数据预处理.如图 2 所示,图 2(a)为归一化前 3D 手势图,图 2(b)为归一化后 3D 手势图.该两张图中的关节段方向一致,仅关节长度发生了变化.

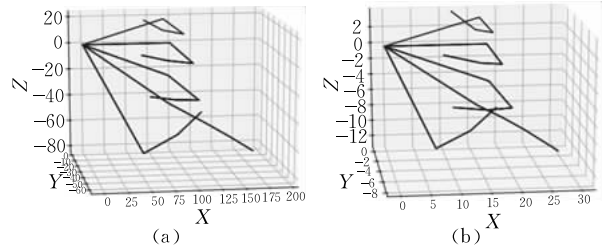


图 2 手势 3D 图示

5 融合关节旋转特征和指尖距离特征的手势特征表示与提取

根据手解剖结构的特点,人手形成一个以关节为节点,由骨骼和关节共同构成的层次结构.手不同类型运动实质上是各部分骨骼在不同关节处的旋转运动构成的复合体.为了完整地表示一个手势,本文提出利用关节的旋转和距离信息作为手势内在特征的表示方法.不同于利用彩色图或深度图直接提取手势特征(如手部关节点位置信息)的方法,本文的手势内在特征表示包含两部分:手部关节对之间的旋转特征和手指指尖距离特征.实际上,附录中给出的详细证明表明,该特征表示能够与基于关节点位置表示的静态手势唯一对应,从而可以唯一地刻画一个静态手势.

5.1 手部关节对之间的旋转特征

空间中常用的三种旋转表示方法分别是方向余弦旋转矩阵、欧拉旋转、四元数^[25].其中,方向余弦矩阵使用了一个 4×4 的矩阵来表示绕任意轴旋转的变换矩阵,数据具有较大冗余,难以实现对多限制条件的手势跟踪与识别.欧拉旋转是按照一定的坐标轴顺序,是一系列坐标轴旋转变换的组合.该方法需要按照一个固定的坐标轴顺序旋转,不同的旋转顺序将导致不同的识别结果.

本文采用四元数刻画手部关节对之间的旋转特征^[25]. 四元数包括旋转轴和旋转角度信息, 能够用 4 维数据表示沿任意轴的旋转, 具体可以表示为: $\mathbf{q} = q_0 + q_1 \cdot \mathbf{i} + q_2 \cdot \mathbf{j} + q_3 \cdot \mathbf{k}$, 其中 q_0 是实数, $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 可以理解成一种几何旋转正交坐标 $(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$. 为了表示关节对旋转特征和表征手势的唯一性, 本文采用四元数来表示手部关节对之间的旋转.

以 $\bar{\mathbf{V}}_0(x_0, y_0, z_0)\bar{\mathbf{V}}_1(x_1, y_1, z_1)$ 关节段之间的四元数为例, 计算如下:

$$\begin{cases} q_0 = \cos(\cos^{-1}(\bar{\mathbf{V}}_0 \times \bar{\mathbf{V}}_1)/2) \\ q_1 = \frac{k_1}{\sqrt{k_1^2 + k_2^2 + k_3^2}} \cdot \sin(\cos^{-1}(\bar{\mathbf{V}}_0 \times \bar{\mathbf{V}}_1)/2) \\ q_2 = \frac{k_2}{\sqrt{k_1^2 + k_2^2 + k_3^2}} \cdot \sin(\cos^{-1}(\bar{\mathbf{V}}_0 \times \bar{\mathbf{V}}_1)/2) \\ q_3 = \frac{k_3}{\sqrt{k_1^2 + k_2^2 + k_3^2}} \cdot \sin(\cos^{-1}(\bar{\mathbf{V}}_0 \times \bar{\mathbf{V}}_1)/2) \end{cases},$$

其中:

$$\begin{cases} k_1 = (y_1 - y_0)(z_2 - z_1) - (y_2 - y_1)(z_1 - z_0) \\ k_2 = (z_1 - z_0)(x_2 - x_1) - (z_2 - z_1)(x_1 - x_0) \\ k_3 = (x_1 - x_0)(y_2 - y_1) - (x_2 - x_1)(y_1 - y_0) \end{cases}$$

由于静态手势具有整体旋转平移不变性, 也为了简化特征向量, 本文中采用 q_0 作为旋转特征, 表示旋转的角度信息, 在四元数中其余的 3 个数表示旋转轴的信息, 旋转轴过于依赖关节点的位置信息, 相同的手势因关节点位置信息不同, 旋转轴的信息也会相差很远. 在方法比较中, 本文也将包含旋转轴信息的四元数作为特征进行了实验, 分析发现旋转轴信息不仅会造成信息重复, 而且将对手势分类带来干扰导致分类不准确. 因此, 本文采用四元数的角度信息作为特征, 将 19 个关节对所对应的 14 个旋转特征表示为 $Q_i (i=0, 1, 2, \dots, 13)$. 如图 3 所示, 其中弧线标记关节段之间的旋转角度.



图 3 手势旋转特征图示

5.2 手指指尖距离特征

一般来说, 仅考虑关节段之间旋转角度信息作为特征将出现旋转误差累积的现象. 不妨将每个手指抽象化为由 3 或 4 根链条组成的一段链条段, 链条与链条之间旋转角度的细微误差, 累积到链条顶端将造成较大的旋转累积误差, 因而需要考虑将 5 个指尖相对于手腕的距离作为一个特征来限制指关节旋转的累积误差. 此外, 利用关节对之间的旋转信息和 5 个指尖相对于手腕的距离信息, 可以确立 5 个手指在各自局部坐标系中坐标, 但是要将 5 个手指的局部坐标系统一到手部的全局坐标系中, 需要 5 个手指的相对距离信息, 因而需要进一步采用 5 个手指相邻指尖之间的距离作为另一个距离特征. 具体地说, 手指指尖距离特征表征如下:

(1) 手指相邻指尖之间的距离:

$$\begin{aligned} S_0 &= \|\bar{\mathbf{W}}_7 - \bar{\mathbf{W}}_3\|, S_1 = \|\bar{\mathbf{W}}_{11} - \bar{\mathbf{W}}_7\|, \\ S_2 &= \|\bar{\mathbf{W}}_{15} - \bar{\mathbf{W}}_{11}\|, S_3 = \|\bar{\mathbf{W}}_{19} - \bar{\mathbf{W}}_{15}\|. \end{aligned}$$

(2) 手指指尖相对于手腕的距离:

$$\begin{aligned} D_0 &= \|\bar{\mathbf{W}}_3\|, D_1 = \|\bar{\mathbf{W}}_7\|, D_2 = \|\bar{\mathbf{W}}_{11}\|, \\ D_3 &= \|\bar{\mathbf{W}}_{15}\|, D_4 = \|\bar{\mathbf{W}}_{19}\|. \end{aligned}$$

5.3 特征融合方法

由于上述提出的关节旋转特征和指尖距离特征将共同表征手势的内在几何特点, 类似于 Luvizon 等人^[20]将特征进行串联融合的思路, 本文将关节旋转特征和指尖距离特征进行串联, 融合为单个手势的 23 维特征向量:

$$\mathbf{F} = [Q_0, \dots, Q_{13}, S_0, \dots, S_3, D_0, \dots, D_4].$$

在含有 N 个样本的数据集中, 分别得到 N 个手势特征向量如下:

$$\mathbf{F}_i = [Q_{i,0}, \dots, Q_{i,13}, S_{i,0}, \dots, S_{i,3}, D_{i,0}, \dots, D_{i,4}].$$

记为

$$\mathbf{F}_i = [f_{i,0}, \dots, f_{i,13}, f_{i,14}, \dots, f_{i,17}, f_{i,18}, \dots, f_{i,22}],$$

$$i = 1, 2, \dots, N.$$

并对 N 个特征向量中各维特征分别归一化如下:

$$\bar{f}_{i,j} = \frac{f_{i,j} - f_j}{\sigma_j},$$

其中:

$$f_j = \frac{\sum_{i=1}^N f_{i,j}}{N}, \quad \sigma_j = \sqrt{\frac{\sum_{i=1}^N (f_{i,j} - f_j)^2}{N}}.$$

从而得到 N 个手势的归一化特征向量如下:

$$\bar{\mathbf{F}}_i = [\bar{f}_{i,0}, \dots, \bar{f}_{i,13}, \bar{f}_{i,14}, \dots, \bar{f}_{i,17}, \bar{f}_{i,18}, \dots, \bar{f}_{i,22}],$$

$$i = 1, 2, \dots, N.$$

定理. 本文融合关节旋转特征和指尖距离特征

的手势特征能唯一地表征手势关节的位置信息.

详细证明请见附录.

6 基于内在特征表示的一对一 SVM 多分类手势识别

作为一种有监督分类方法,支持向量机方法是建立在统计学习理论的 VC 维理论和结构风险最小原理基础上,其能够根据有限的样本信息在模型对特定训练样本的学习精度和学习能力之间寻求最佳的折衷,在解决小样本、非线性及高维度模式识别问题中具有一定的优势^[27].

对于多分类的 SVM,有两种方法选择.(1)一对多.依次将 K 个数据类别中的每一类作为正例,其余作为负例,构造 K 个二分类 SVM,将测试集在这 K 个二分类 SVM 做分类,相应的数据类别评分 +1,取评分最高的手势类别作为最终分类类别;(2)一对一.将 K 个数据类别,两两做 SVM 分类,将测试集在两者之间做分类,一共有 $K \times (K-1)/2$ 个 SVM,相应的数据类别评分 +1,取评分最高的手势类别作为最终分类类别.

De Smedt 等人^[24]采用的是一对多法进行动态手势识别,但是该方法由于负类样本数量远远大于正类样本的数量导致样本不对称情况,对于有一些异常情况的分类情况并不理想且这种不理想随着训练数据的增加而趋向严重,同时会出现分类重叠现象或者不可分类现象,手势分类与识别欠准确.为保证手势识别的准确性,本文采用一对一 SVM 多分类策略进行基于手势内在特征的分类与识别,既避免了数据样本不对称情况,也避免了不可分类的情况.

综上所述,基于本文的手势内在特征表示的 SVM 识别算法如下:

算法 2. 基于手势特征的 SVM 识别.

输入:静态待测试手势的特征向量

输出:该手势对应于数据集中的类别

1. 利用核函数 $H(\cdot)$ 将手势特征向量从原始空间映射到特征空间;

2. 从训练集中任取两类训练集的手势特征,利用 SMO (Sequential Minimal Optimization) 算法^[28] 确立两类手势之间的参数;

3. 重复步骤 2,直到所有类别均分类完成;

4. 将训练集中的样本放入模型进行训练,并保存模型;

5. 将测试集中手势作为待识别手势,并将其特征向量,利用核函数 $H(\cdot)$ 进行映射;

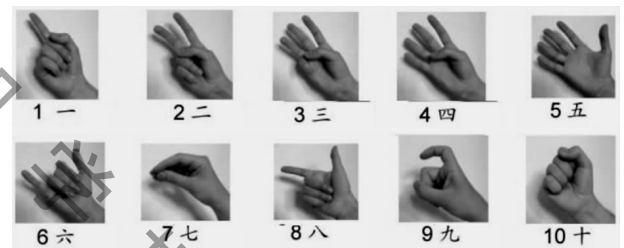
6. 将待测试的手势放入 SVM 模型中,并与训练集中任意 2 类手势之间进行分类,分类得到的相应手势类别评分加 1: $S_k = S_k + 1$,其中 k 为分类到的手势类别;

7. 输出评分最高的手势类别,作为待识别手势的类别.

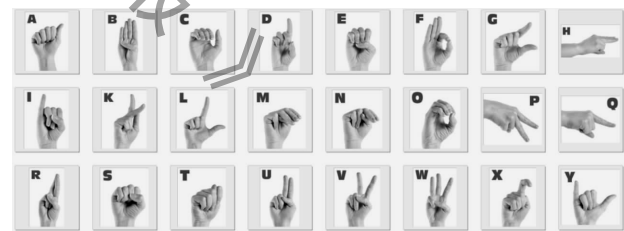
7 实验结果与分析

7.1 实验环境及数据

本文算法在 Ubuntu16.04 并配置 Python3 环境上得到了实现,利用 Opencv2 函数库和 Caffe 框架,借助 ASTAR 手势数据集^[17]对 RGB-D 深度图像获取的不同手势进行了分类和识别. ASTAR 静态手势数据集^[17]包含来自多个志愿者的带标注信息的 870 个静态手势深度图,其中手势深度图由 TOF 相机获取,手势标注信息则使用数据手套获得.在不失一般性的前提下,仅关注右手.这些照片来自 30 名志愿者,年龄各异(18~60 岁)、性别(15 男,15 女)、种族和手掌大小各不相同.在捕捉过程中,29 类手势,每个志愿者仅完成一次,可分为两组.这些图片来自中国数字计数系统(从 1 到 10,不包括 3 和 7),而其余来自美国手语字母(从 A 到 Z,不包括 J, R, T, W 和 Z),具体手势如图 4 所示.本文除了在该数据集上进行手势识别实验,还在此基础上添加噪声和扰动将该数据集扩充一倍进行了对比实验.



(a) 中国手语数字手势



(b) 美国手语字母手势

图 4 数据集手势图^[17]

7.2 手势识别分类实例

考虑到中国数字与美国手语字母的手势模型中有重复手势(如中国数字手势 2 与美国手语字母 V),所以在本文实验中将 2 个数据集分别进行交叉验证实验.对于中国手语数字手势数据集,如图 4(a),手势的区分度很明显.训练集和测试集的比例为 9:1,随机从数据集中抽取 10% 作为测试集,本文利用训练集来构造手势库,将测试集作为待识别手势数据.利用本文方法达到 99.71% 的分类识别准确率,中国手语数字手势分类的混淆矩阵如表 1 所示.

表 1 使用本文方法的中国手语数字分类混淆矩阵

类别	1	2	3	4	5	6	7	8	精确率
Num1(1)	30					1			0.9677
Num2(2)		33							1
Num4(3)			30						1
Num5(4)				21					1
Num6(5)					35				1
Num8(6)						21			1
Num9(7)						1	32		0.9697
Num10(8)								36	1
召回率	1	1	1	1	1	0.913	1	1	0.9971

对于美国手语字母数据集,如图 4(b),训练集和测试集的比例为 9:1,随机从数据集中抽取 10% 作为测试集.将旋转角度作为特征进行 SVM 分类时发现,对于美国手语字母 a 和 c,由于参与测试的人员做手势时,并不是严格按照手势图进行比对,故对于不同的测试人员具有较大差异性.利用本文方法达到 85.24% 的分类识别准确率,美国手语字母手势分类的混淆矩阵如表 2 所示.

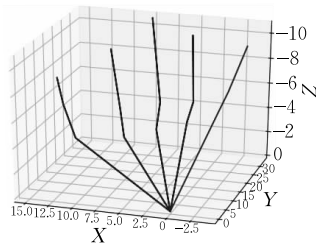
表 2 使用本文方法的美国手语字母分类混淆矩阵

类别	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	精确率	
A(1)	20		2											4			4					0.67	
B(2)		26														4						0.87	
C(3)	2		28											1			1					0.88	
D(4)				26																		1	
E(5)					30																	1	
F(6)						32																1	
G(7)							28															1	
H(8)								17										9				0.65	
I(9)									27													1	
K(10)										29						1						0.97	
L(11)											30											1	
M(12)												17		6			2					0.68	
N(13)													25				7					0.78	
O(14)														8								0.78	
P(15)								6							23			1	6			0.64	
Q(16)				5																		0.82	
S(17)		4																				0.77	
U(18)								4														0.84	
V(19)															9			4	24			0.65	
X(20)																					29	1	
Y(21)																						30	1
召回率	0.91	1	0.82	0.84	1	1	1	0.63	1	1	1	0.68	1	0.67	0.72	0.82	0.63	0.6	0.8	1	1	0.8524	

为不失一般性,利用微软 Kinect 2.0 相机,拍摄获取的一张手势深度图如图 5(a)所示,利用本文融合关节旋转特征和指尖距离特征的手势识别方法,能成功识别为中国手语数字 5,如图 5(b)所示.



(a) 手势深度图



(b) 归一化手势3维骨架

图 5 真实数据的手势识别

7.3 方法比较与讨论

基于相同的手势数据集,将本文方法与 Xu 等人^[17]提出的 Dhand 方法进行比较. Dhand 方法中将中国数字数据与美国字母数据合并进行分类,准确率为 53%. 本文在该合并数据集上进行分类,准确

率为 70.69%,在 Dhand 方法上提高了 17.69% 准确率.然而,由于在中国手语数字和美国手语字母中有些手势表示一模一样,比如 2 和 V、6 和 Y、10 和 S,因而本文还将中国数字和美国字母分开训练.在对比时,本文将 Dhand 方法的混淆矩阵中大类别分错的情况(如将中国数字分类到美国字母类别中的,或将美国字母分类到中国数字类别中),将其设为是分类正确的情况.需要说明的是,本文实验中的测试集不包含在训练集中,测试集与训练集相对独立.

在方法比较中,除了与 Dhand 方法^[17]对比外,将 HOG+SVM^[20]、SURF+SVM^[21]与本文的不同特征提取方法进行对比,以证实本文方法的有效性.其中 HOG+SVM^[20]方法在静态深度图中提取 HOG 特征并利用 SVM 进行分类,该方法直接针对深度图提取特征,对于轮廓相似的英文字母数据由于提取的深度图梯度特征较接近而难以正确区分,其识别率较低. SURF+SVM^[21]方法则利用手势深

度图提取的 SURF 描述作为特征并利用 SVM 分类,该方法针对差别较大的手势具有较好的识别率,而对于具有相似深度图的美国字母手势则难以正确区分.相较于已有方法,本文提出的手势特征表示由于具有更多可供区别的特征信息,能够更好地区别不同手势.表 3 中,根据是否包含手部关节旋转轴信息以及是否加入距离特征信息列出 3 种不同特征选取,特征集 1 使用“四元数特征(含旋转轴信息),加距离特征信息”;特征集 2 使用“四元数特征(不含旋转轴信息),不加距离特征信息”;特征集 3 则使用“四元数特征(不含旋转轴信息),加距离特征信息”.在中国数字手势识别中,利用含旋转轴信息的四元数和相对距离作为手势特征时,从表 4 可以看到旋

表 3 特征集及其描述

特征集	特征描述
特征集 1	四元数特征(含旋转轴信息), 加距离特征信息
特征集 2	四元数特征(不含旋转轴信息), 不加距离特征信息
特征集 3	四元数特征(不含旋转轴信息), 加距离特征信息

表 4 不同方法的手势识别率比较

方法	准确率/%	
	中国数字手势	美国字母手势
Dhand 方法 ^[17]	87.00	62.50
特征集 1+SVM	75.42	61.59
特征集 2+SVM	97.92	78.73
特征集 3+SVM (一对多)	97.08	72.70
特征集 3+SVM	99.71	85.24
特征集 3+KNN	—	78.41
特征集 3+NB	—	83.01
HOG+SVM ^[20]	87.92	35.87
SURF+SVM ^[21]	73.75	34.92
特征集 3+SVM(扩充数据集)	99.38	83.02

转轴信息将对数据分类造成较大干扰,虽然手部关节旋转轴有助于确定关节位置信息,但是由于手势的空间差异性,相同的手势关节对旋转轴信息将会有较大差别.

为体现 SVM 一对多分类与一对一分类的区别,添加 SVM 一对多分类作为对比,本文中除特别说明外,SVM 均指一对一分类.从表 4 可以看出,一对一分类可以获得更高的准确率,有效避免样本不对成和不可分类的情况.针对美国字母手势,Xu 等人^[17]提出的 Dhand 方法识别准确率为 62.5%;而采用不加距离特征的特征信息,其识别准确率仅 78.73%;添加距离特征作为特征信息,其识别准确率达到 85.24%,表明距离特征能提高手势识别准确率.同时,基于特征集 3,本文分别采用 K 邻近(KNN)、朴素贝叶斯(NB)另 2 种分类器进行美国字母手势的分类识别,其识别准确率分别为 78.41%、83.01%.实验表明,针对本文提出的特征集 3,使用 SVM 分类能获得更高的识别准确率.

在考虑特征提取时,不仅考虑其要与手势唯一对应,更要考虑手势特征对于手势区分度的影响.图 6 显示了中国数字手势的 14 维特征分布,可以看到每一类手势特征都用雷达图表示,手势具有各自的分布特点,其区分度较大,不同的数字手势对应伸出不同的手指,伸出手指的关节对旋转特征较小,所以可以利用 SVM 进行很好地区分.分类混淆矩阵如表 1 所示,其中每一列表示数据的预测类别,每一行表示数据的真实归属类别.表 1 中数字 m_{ij} 表示实际归属为第 i 类的实例被预测为第 j 类的数目,如表 1 中第一行第一列中 $m_{11}=30$ 表示有 30 个实际归属为第一类的实例被预测为第一类.从表 1 中可以看出,共 238 个数字手势测试实例集均能被正确识别与分

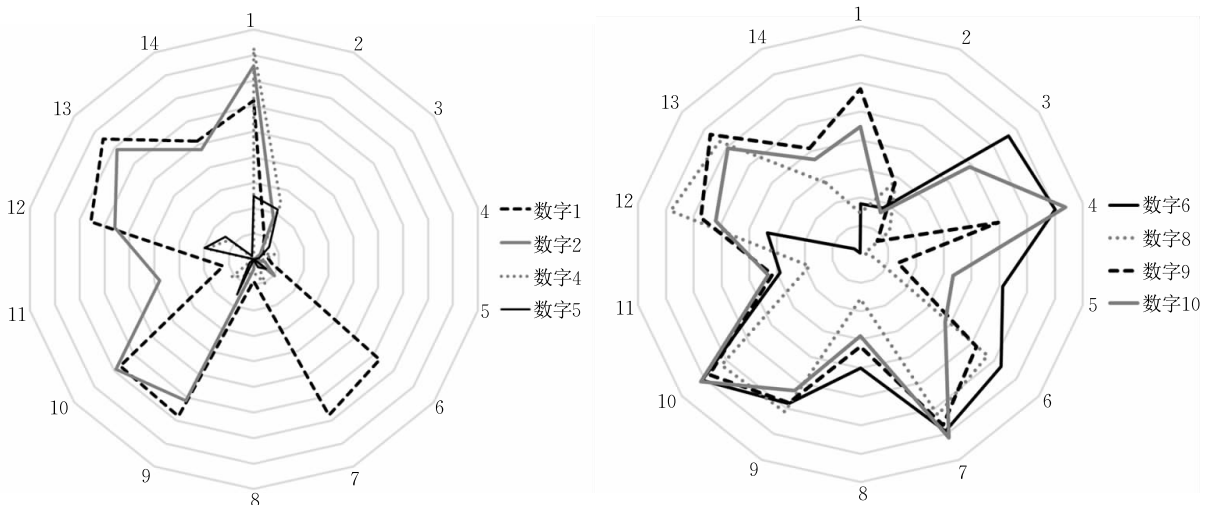


图 6 中国数字的 14 维关节对旋转特征

类. 另外, 表 1 给出了手势分类与识别的召回率和精确率作为衡量模型的参数, 召回率是指所有正确被检索到的样本数量占有应该被检索到的样本数量的比例, 体现了分类模型对正样本的识别能力; 精确率是指所有正确被检索到的样本数量占有实际被检索到的样本数量的比例, 体现了分类模型对负样本的识别能力; 混淆矩阵显示本文方法对于中国数字手势的分类识别准确率达 99.71%.

对于美国手语字母手势, 由于部分手势之间的区分度非常小, 如图 4(b) 中所示 A、C、S 这 3 个手语字母手势之间的相似度非常高. 若仅考虑旋转信息的特征, 每个关节之间的一些细微小差别, 对于整个手势来说却导致很大变化. 例如字母手势 A 和字母 C, 对于关节的旋转信息仅存在细微差别, 但是如果考虑到手指关节的相对位置信息, 则将发现字母手势 A 的手指相对于手腕的距离与字母手势 C 的手指相对于手腕的距离有较大差别, 因而, 需要考虑手指关节的相对距离信息, 以避免旋转信息特征的累计误差而导致的识别错误.

针对美国手语字母手势, 本文方法的分类混淆矩阵见表 2, 其中每一列表示数据的预测类别, 每一

行表示数据的真实归属类别, 数字 m_{ij} 表示实际归属为第 i 类的实例被预测为第 j 类的数目, 其混淆矩阵显示了本文方法对于美国手语字母手势的分类识别准确率达 85.24%. 在手语字母手势识别中, 有 93 组字母手势分类识别错误, 其中以较为典型的 P 和 V、H 和 P、N 和 S 为例进行说明. 观察这 3 组手势, 可以发现其中 P 和 V 手势的唯一区别在于手势的方向, 而本文中提取的特征强调手关节之间的旋转特征和距离特征, 没有考虑手势的方向问题, 如图 7 所示, H 手势和 P 手势的区别在于中指与食指的距离, 由于 Xu 等人^[17] 主要使用该数据集估计手部关节位置信息, 没有特别注重手势之间的区别, 而且收集手势的对象存在极大的年龄差异, 年较长的人对于手势重现的精确度把握较难, 导致这 2 个手势十分相似. 如图 8 所示, 字母 N 和 S 手势的最大区别在于大拇指在食指与中指的里侧还是外侧, 但由于收集手势对象的手指长度存在差异, 可能会造成手指长度归一化后, 手指长度变短导致指尖的位置发生变化, 从而使得手势之间区分度较小, 易发生手势分类错误.

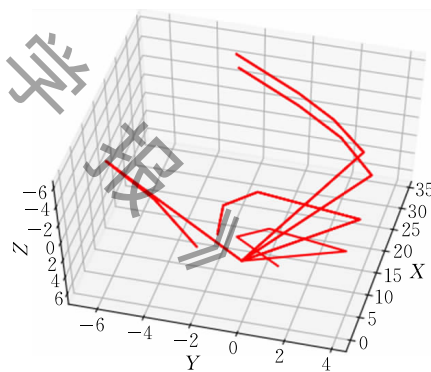
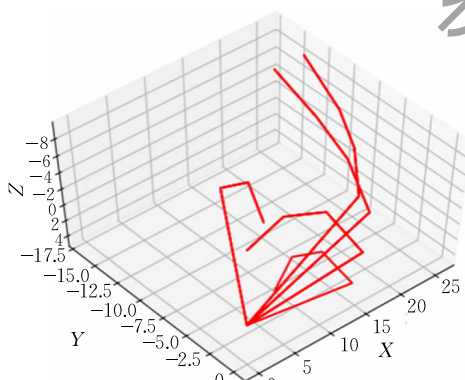


图 7 字母 H 和 P 的手势关节骨架图

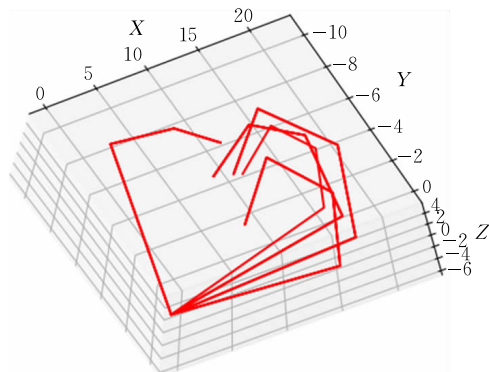
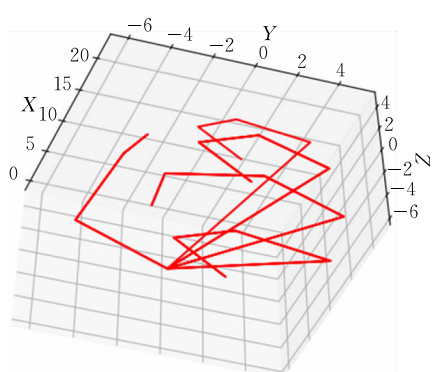


图 8 字母 N 和 S 的手势关节骨架图

8 结 语

本文从手势深度图中提取得到的手部 20 个关节节点的 3D 位置信息作为输入, 结合关节旋转特征和指尖距离特征, 提出了一种新的手势特征表示与手势识别的方法. 实验表明, 该特征表示能够很好地表征不同手势, 在静态手势的识别中利用该特征表示能够达到较高的分类识别准确率.

然而, 由于人机交互中的不同手势通常具有一定的运动属性, 未来将考虑结合手势运动方向特征的手势分类与识别, 并考虑基于动态特征表示的动态手势识别.

参 考 文 献

- [1] Liu Y, Yin Y, Zhang S. Hand gesture recognition based on HU moments in interaction of virtual reality // Proceedings of 4th IEEE International Conference on Intelligent Human-Machine Systems and Cybernetics. Nanchang, China, 2012: 145-148
- [2] Bannach D, Amft O D, Kunze K S, et al. Waving real hand gestures recorded by wearable motion sensors to a virtual car and driver in a mixed-reality parking game // Proceedings of the IEEE Symposium on Computational Intelligence and Games. Honolulu, Hawaii, USA, 2007: 32-39
- [3] Tao L, Zappella L, Hager G D, et al. Surgical gesture segmentation and recognition // Proceedings of the 16th International Conference on Medical Image Computing and Computer Assisted Intervention. Nagoya, Japan, 2013: 339-346
- [4] Moschetti A, Fiorini L, Esposito D, et al. Toward an unsupervised approach for daily gesture recognition in assisted living applications. *IEEE Sensors Journal*, 2017, 17(24): 8395-8403
- [5] Rautaray S S, Agrawal A. Vision based hand gesture recognition for human computer interaction: A survey. *Artificial Intelligence Review*, 2015, 43(1): 1-54
- [6] Cheng H, Yang L, Liu Z. Survey on 3D hand gesture recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016, 26(9): 1659-1673
- [7] Zhu Ji-Yu, Wang Xi-Ying, Wang Wei-Xin, et al. Hand gesture recognition based on structure analysis. *Chinese Journal of Computers*, 2006, 29(12): 2130-2137 (in Chinese) (朱继玉, 王西颖, 王威信等. 基于结构分析的手势识别. *计算机学报*, 2006, 29(12): 2130-2137)
- [8] Panwar M. Hand gesture recognition based on shape parameters // Proceedings of the IEEE International Conference on Computing, Communication and Applications. Dindigul, India, 2012: 1-6
- [9] Ren Z, Yuan J, Meng J, et al. Robust part-based hand gesture recognition using Kinect sensor. *IEEE Transactions on Multimedia*, 2013, 15(5): 1110-1120
- [10] Yang Xue-Wen, Feng Zhi-Quan, Huang Zhong-Zhu, et al. Gesture recognition based on combining main direction of gesture and Hausdorff-like distance. *Journal of Computer-Aided Design & Computer Graphics*, 2016, 28(1): 75-81 (in Chinese) (杨学文, 冯志全, 黄忠柱等. 结合手势主方向和类-Hausdorff 距离的手势识别. *计算机辅助设计与图形学学报*, 2016, 28(1): 75-81)
- [11] Zhang Z. Microsoft Kinect sensor and its effect. *IEEE Multimedia*, 2012, 19(2): 4-10
- [12] Qian C, Sun X, Wei Y, et al. Realtime and robust hand tracking from depth // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 1106-1113
- [13] Tang D, Yu T H, Kim T K. Real-time articulated hand pose estimation using semi-supervised transductive regression forests // Proceedings of the IEEE International Conference on Computer Vision. Sydney, Australia, 2013: 3224-3231
- [14] Tompson J, Stein M, Lecun Y, et al. Real-time continuous pose recovery of human hands using convolutional networks. *ACM Transactions on Graphics*, 2014, 33(5): Article No. 169
- [15] de La Gorce M, Fleet D J, Paragios N. Model-based 3D hand pose estimation from monocular video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(9): 1793-1805
- [16] Ye Q, Yuan S, Kim T K. Spatial attention deep net with partial BSO for hierarchical hybrid hand pose estimation // Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands, 2016: 346-361
- [17] Xu C, Nanjappa A, Zhang X, et al. Estimate hand poses efficiently from single depth images. *International Journal of Computer Vision*, 2016, 116(1): 21-45
- [18] Shotton J, Fitzgibbon A, Cook M, et al. Real-time human pose recognition in parts from single depth images // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Colorado, USA, 2011: 1297-1304
- [19] Girshick R, Shotton J, Kohli P, et al. Efficient regression of general-activity human poses from depth images // Proceedings of the IEEE International Conference on Computer Vision. Barcelona, Spain, 2011: 415-422
- [20] Feng K, Yuan F. Static hand gesture recognition based on HOG characters and support vector machines // Proceedings of the 2nd International Symposium on Instrumentation and Measurement, Sensor Network and Automation. Toronto, Canada, 2013: 936-938
- [21] Bagdanov A D, Del Bimbo A, Seidenari L, et al. Real-time hand status recognition from RGB-D imagery // Proceedings of the 21st International Conference on Pattern Recognition. Tsukuba, Japan, 2012: 2456-2459

- [22] Zhou X, Wan Q, Zhang W, et al. Model-based deep hand pose estimation//Proceedings of the 25th International Joint Conference on Artificial Intelligence. New York, USA, 2016: 2421-2427
- [23] Zimmermann C, Brox T. Learning to estimate 3D hand pose from single RGB images//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy, 2017: 4913-4921
- [24] De Smedt Q, Wannous H, Vandeborste J P. Skeleton-based dynamic hand gesture recognition//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Las Vegas, USA, 2016: 1206-1214
- [25] Spring K W. Euler parameters and the use of quaternion algebra in the manipulation of finite rotations: A review. Mechanism and Machine Theory, 1986, 21(5): 365-373
- [26] Luvizon D C, Tabia H, Picard D. Learning features combination for human action recognition from skeleton sequences. Pattern Recognition Letters, 2017, 99: 13-20
- [27] Kotsiantis S B, Zaharakis I, Pintelas P. Supervised machine learning: A review of classification techniques. Emerging Artificial Intelligence Applications in Computer Engineering, 2007, 160: 3-24
- [28] Cao L J, Keerthi S S, Ong C J, et al. Parallel sequential minimal optimization for the training of support vector machines. IEEE Transactions on Neural Networks, 2006, 17(4): 1039-1049
- [29] Kamel N. A novel approach to dynamic signature verification using sensor-based data glove. American Journal of Applied Sciences, 2009, 6(2): 233-240
- [30] Dipietro L, Sabatini A M, Dario P. A survey of glove-based systems and their applications. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Application and Reviews, 2008, 38(4): 461-482

附录. 本文手势特征表示的内在唯一性.

根据参考文献[29-30]中对于手指各主要关节数据的描述,可知手部关节点满足以下模型假设:

- (1) 每个手指的关节可单独视为在同一平面内;
- (2) 除大拇指外,其余四指掌端关节点在同一平面,且该平面与中指平面垂直;
- (3) 手指骨架之间没有发生相交;
- (4) 大拇指手腕关节点 \bar{W}_0 的屈曲自由度与掌端关节点 \bar{W}_1 的屈曲自由度 (inflection degree) 存在如下关系:

$$\text{flex}(\bar{W}_0) = 2 \cdot \left(\text{flex}(\bar{W}_1) - \frac{\pi}{6} \right).$$

引理 1. 手指平面中,根据单个手指关节段的长度和关节段与段之间角度,可计算确定手指关节点的坐标.

证明. 设关节长度为 l_0, l_1, l_2, l_3 , 关节段与段之间角度为 $\theta_0, \theta_1, \theta_2$, 则手指 4 个关节点 X_1, X_2, X_3, X_4 可确定如下:

不妨以手腕关节点为原点 $(0, 0)$, 以手腕和掌根关节的骨架作为 y 轴建立坐标系, 则掌端关节点 (Metacarpophalangeal Joint, MCJ) X_1 坐标 (x_1, y_1) 为 $(0, l_0)$.

已知掌端关节段与指中关节段间的夹角为 θ_0 , 则近端指节关节点 (Proximal Interphalangeal Joint, PIJ) X_2 坐标 (x_2, y_2) 为 $(l_1 \cdot \sin\theta_0, l_0 - l_1 \cdot \cos\theta_0)$;

远端指节关节点 (Distal Interphalangeal Joint, DIJ) X_3 坐标 (x_3, y_3) 为 $(x_2 + l_2 \cdot \sin\theta_1, y_2 - l_2 \cdot \cos\theta_1)$;

指尖关节点 (Fingertip Joint, FTJ) X_4 坐标 (x_4, y_4) 为

$$(x_3 + l_3 \cdot \sin\theta_2, y_3 - l_3 \cdot \cos\theta_2).$$

引理 2. 空间坐标系中,已知单个手指关节段的长度和关节段与段之间角度,以及该指指尖到某一固定点 (非原点) 的距离,且该指掌端关节保持在 YOZ 平面内,则可计算确定该手指关节点的坐标.

证明. 不妨设关节长度为 l_0, l_1, l_2, l_3 , 关节段与段之间的角度为 $\theta_0, \theta_1, \theta_2$, 指尖 X_4 到点 $X_0(x_0, y_0, z_0)$ 距离为 S_0 ,

且 X_1 在 YOZ 平面内,则手指 4 个关节点 X_1, X_2, X_3, X_4 在空间坐标系中坐标可确定如下.

由引理 1 可知,该手指在二维平面内的坐标是唯一确定的. 由引理 1 中公式,不妨设 X_1 坐标 $(0, m_1, n_1)$, X_2 坐标 $(x_1, m_1, y_1 - l_0 + n_1)$, X_3 坐标 $(x_2, m_1, y_2 - l_0 + n_1)$, X_4 坐标 $(x_3, m_1, y_3 - l_0 + n_1)$.

由于指尖 X_4 与 X_0 距离为 S_0 , 掌端关节段长度为 L_0 , 且 m_1, n_1 的正负与所在象限有关,不妨都设为正数,故 m_1, n_1 需满足:

$$\begin{cases} S_0^2 = (x_3 - x_0)^2 + (m_1 - y_0)^2 + (y_3 - l_0 + n_1 - z_0)^2 \\ l_0^2 = m_1^2 + n_1^2 \\ m_1 > 0, n_1 > 0 \end{cases}$$

解得:

$$\begin{cases} n_1 = \frac{-ab + \sqrt{a^2 b^2 - (a^2 + 1)(b^2 - l_0^2)}}{a^2 + 1} \\ m_1 = b + an_1 \end{cases}$$

其中:

$$\begin{cases} a = \frac{y_3 - l_0 - z_0}{-y_0} \\ b = \frac{S_0^2 - l_0^2 - (x_3 - x_1)^2 - y_0^2 - (y_3 - l_0 - z_0)^2}{-2y_0} \end{cases}$$

定理 1. 本文特征表示可以确定一个手势的 20 个关节点的位置信息.

证明. 已知手势特征如下:

$$F = [Q_0, \dots, Q_{13}, S_0, \dots, S_3, D_0, \dots, D_4],$$

以及手指各关节长度:

$$L = [L_0, L_1, L_2, \dots, L_{18}].$$

需要求解手部关节点坐标:

$$[\bar{W}_0, \bar{W}_1, \bar{W}_2, \dots, \bar{W}_{19}].$$

由模型假设(1)可知,每个手指都有各自的指平面,故先

证明手指在各自局部的二维指平面内形状唯一,再将五个手指平面放入同一三维空间坐标系中.

(1) 各指关节在指平面局部坐标系中的坐标确定

由引理 1 可知,根据关节段长度和关节段之间的角度,可以确定在手指平面局部坐标系内手指 4 个关节点坐标.

先以食指平面为例:手腕坐标为(0,0),掌端关节点为(0, L_3),近端指节关节点(x_{21}, y_{21})为

$$(2Q_2 \cdot L_4 \cdot \sin(\arccos Q_2), L_3 - L_4 \cdot (2Q_2^2 - 1)) \quad (1)'$$

远端指节关节点(x_{22}, y_{22})为

$$(x_{21} + 2Q_3 \cdot L_5 \cdot \sin(\arccos Q_3), y_{21} - L_5 \cdot (2Q_3^2 - 1)) \quad (2)'$$

指尖关节点(x_{23}, y_{23})为

$$(x_{22} + 2Q_4 \cdot L_6 \cdot \sin(\arccos Q_4), y_{22} - L_6 \cdot (2Q_4^2 - 1)) \quad (3)'$$

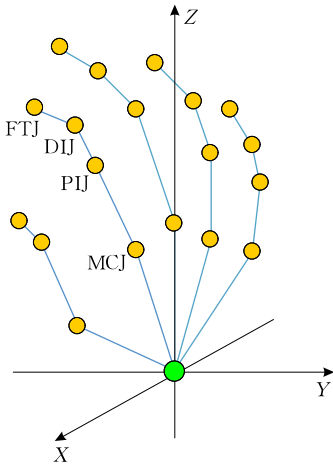
同理,类似式(1)'、(2)'和式(3)',可计算得其余四指关节在各自指平面局部坐标系中的坐标记为

$$\begin{cases} (0, L_0) & (x_{11}, y_{11}) & (x_{12}, y_{12}) \\ (0, L_7) & (x_{31}, y_{31}) & (x_{32}, y_{32}) & (x_{33}, y_{33}) \\ (0, L_{11}) & (x_{41}, y_{41}) & (x_{42}, y_{42}) & (x_{43}, y_{43}) \\ (0, L_{15}) & (x_{51}, y_{51}) & (x_{52}, y_{52}) & (x_{53}, y_{53}) \end{cases}$$

利用距离特征 D_0, \dots, D_4 可检验指尖关节点坐标的准确性.

(2) 各指关节在手部全局坐标系中的坐标确定

不妨以手腕关节点为原点,中指指平面所在平面为 XOZ 平面,且手腕关节与中指掌端关节点所在直线为 x 轴正半轴,坐标系建立如附图 1 所示.



附图 1 坐标系的建立

① 中指关节点坐标为:

$$\begin{aligned} \bar{W}_0 &= (0, 0, 0), \bar{W}_8 = (0, 0, L_7), \bar{W}_9 = (x_{31}, 0, y_{31}), \\ \bar{W}_{10} &= (x_{32}, 0, y_{32}), \bar{W}_{11} = (x_{33}, 0, y_{33}) \end{aligned} \quad (4)'$$

② 食指关节点坐标为:

由假设(2)可知,除大拇指外,另四指掌端关节点共面,且该面与中指指平面垂直,故可知另四指掌端关节点所在平面与 $y=0$ 平面垂直,即在 $x=0$ 平面内.食指指尖与中指指尖距离应为 S_1 ,食指掌端关节段长度为 L_3 ,且食指在中指的左边,位于 y 轴的负半轴, z 轴的正半轴.由引理 2 类似推导,关节点坐标确定如下:

不妨设:

$$\bar{W}_4 = (0, -m_1, n_1),$$

$$\begin{aligned} \bar{W}_5 &= (x_{21}, -m_1, y_{21} - L_3 + n_1), \\ \bar{W}_6 &= (x_{22}, -m_1, y_{22} - L_3 + n_1), \\ \bar{W}_7 &= (x_{23}, -m_1, y_{23} - L_3 + n_1) \end{aligned} \quad (5)'$$

则 m_1, n_1 需满足:

$$\begin{cases} S_1^2 = (x_{23} - x_{33})^2 + (-m_1)^2 + (y_{23} - L_3 + n_1 - y_{33})^2 \\ L_3^2 = m_1^2 + n_1^2 \\ m_1 > 0, n_1 > 0 \end{cases},$$

解得:

$$\begin{cases} n_1 = \frac{S_1^2 - L_3^2 - (x_{23} - x_{33})^2 - (y_{23} - y_{33} - L_3)^2}{2(y_{23} - y_{33} - L_3)} \\ m_1 = \sqrt{L_3^2 - n_1^2} \end{cases}.$$

③ 无名指关节点坐标为:

与食指坐标类似推导,有

$$\begin{aligned} \bar{W}_{12} &= (0, m_2, n_2), \\ \bar{W}_{13} &= (x_{41}, m_2, y_{41} - L_{11} + n_2), \\ \bar{W}_{14} &= (x_{42}, m_2, y_{42} - L_{11} + n_2), \\ \bar{W}_{15} &= (x_{43}, m_2, y_{43} - L_{11} + n_2) \end{aligned} \quad (6)'$$

其中:

$$\begin{cases} n_2 = \frac{S_2^2 - L_{11}^2 - (x_{43} - x_{33})^2 - (y_{43} - y_{33} - L_{11})^2}{2(y_{43} - y_{33} - L_{11})} \\ m_2 = \sqrt{L_{11}^2 - n_2^2} \end{cases}.$$

④ 小指关节点坐标为:

由引理 2 类似推导,小指指尖与无名指指尖间距离为 S_3 ,小指掌端关节段长度为 L_{15} ,且小指在无名指右边,不妨设:

$$\begin{aligned} \bar{W}_{16} &= (0, m_3, n_3), \\ \bar{W}_{17} &= (x_{51}, m_3, y_{51} - L_{15} + n_3), \\ \bar{W}_{18} &= (x_{52}, m_3, y_{52} - L_{15} + n_3), \\ \bar{W}_{19} &= (x_{53}, m_3, y_{53} - L_{15} + n_3) \end{aligned} \quad (7)'$$

则 m_3, n_3 需满足:

$$\begin{cases} S_3^2 = (m_3 - m_2)^2 + (x_{53} - x_{43})^2 + (y_{53} - L_{15} + n_3 - y_{43} + L_{11} - n_2)^2 \\ L_{15}^2 = m_3^2 + n_3^2 \\ m_3 > 0, n_3 > 0 \end{cases},$$

解得:

$$\begin{cases} n_3 = \frac{-a_3 b_3 + \sqrt{a_3^2 b_3^2 - (a_3^2 + 1)(b_3^2 - L_{15}^2)}}{a_3^2 + 1}, \\ m_3 = b_3 + a_3 n_3 \end{cases},$$

其中:

$$\begin{cases} a_3 = \frac{y_{53} - y_{43} - L_{15} + L_{11} - n_2}{m_2} \\ b_3 = \frac{S_3^2 - L_{15}^2 - m_2^2 - (x_{53} - x_{43})^2 - (y_{53} - y_{43} - L_{15} + L_{11} - n_2)^2}{-2m_2} \end{cases}.$$

⑤ 大拇指关节点坐标为:

由于大拇指掌根关节与其余 4 指不同,该关节存在 2 个自由度,所以在计算大拇指坐标时,分为 2 步考虑.第 1 步:先不考虑大拇指的屈曲度,即将大拇指指平面平移至大拇指掌端关节与其余四指掌端关节在同一平面内,得到大拇指的初始坐标 $\bar{W}'_1, \bar{W}'_2, \bar{W}'_3$;第 2 步:将第一步中的大拇指指平面旋转至满足大拇指掌跟关节自由度,得到大拇指的最终坐标 $\bar{W}_1, \bar{W}_2, \bar{W}_3$.

(1) 不考虑大拇指关节屈曲度的情形

将大拇指指平面平移至大拇指掌端关节与其余四指掌端关节在同一平面内, 得到大拇指关节点 $\bar{W}'_1, \bar{W}'_2, \bar{W}'_3$, 由引理 2 得:

$$\begin{aligned}\bar{W}'_1 &= (0, -m_4, n_4), \\ \bar{W}'_2 &= (x_{11}, -m_4, y_{11} - L_0 + n_4), \\ \bar{W}'_3 &= (x_{12}, -m_4, y_{12} - L_0 + n_4),\end{aligned}$$

其中:

$$\begin{cases} n_4 = \frac{-a_4 b_4 + \sqrt{a_4^2 b_4^2 - (a_4^2 + 1)(b_4^2 - L_0^2)}}{a_4^2 + 1} \\ m_4 = b_4 + a_4 n_4 \end{cases}$$

而:

$$\begin{cases} a_4 = \frac{y_{11} - y_{22} - L_0 + L_3 - n_1}{m_1} \\ b_4 = \frac{S_0'^2 - L_0^2 - m_1^2 - (x_{11} - x_{22})^2 - (y_{11} - y_{22} - L_0 + L_3 - n_1)^2}{-2m_1} \end{cases}$$

其中 S_0' 表示大拇指掌端关节与其余 4 指的掌端关节在同一平面内时, 大拇指指尖与食指指尖的距离待确定。

(2) 将大拇指指平面旋转至掌端关节自由度情形:

由于大拇指掌端关节点 \bar{W}_1 的屈曲自由度为

$$\text{flex}(\bar{W}_1) = \pi - 2 \cdot \arccos Q_0,$$

由假设 4 可知, 大拇指手腕关节点 \bar{W}_0 的屈曲自由度为

$$\text{flex}(\bar{W}_0) = 2 \left(\pi - 2 \cdot \arccos Q_0 - \frac{\pi}{6} \right) = \frac{5\pi}{3} - 4 \cdot \arccos Q_0.$$

由参考文献[29-30]可知, 大拇指的屈曲自由度是相对于大拇指完全伸展开而言, 所以大拇指发生屈曲时其旋转轴经过手腕关节点且垂直于掌面, 屈曲方向朝向掌心。已知旋转轴 N 和旋转角度 $\text{flex}(\bar{W}_0)$, 由罗德里格斯旋转公式 (Rodrigue Rotational Formula)^[25], 可知旋转矩阵为

$$R = [\cos \text{flex}(\bar{W}_0)] I + (1 - \cos \text{flex}(\bar{W}_0)) NN^T + [\sin \text{flex}(\bar{W}_0)] N^\wedge,$$

其中 I 表示单位矩阵, N^\wedge 表示 N 的反对称转换矩阵, 即若

$$N = (r_1, r_2, r_3), \text{ 则 } N^\wedge = \begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix}.$$

故 $\bar{W}_1^T = R\bar{W}'_1^T, \bar{W}_2^T = R\bar{W}'_2^T, \bar{W}_3^T = R\bar{W}'_3^T$, 可解得 $\bar{W}_1, \bar{W}_2, \bar{W}_3$ 关于 S_0' 的表达式, 并利用 $\|\bar{W}_3 - \bar{W}_7\| = S_0$, 可解得 S_0' , 并得到 $\bar{W}_1, \bar{W}_2, \bar{W}_3$ 的坐标。

定理 2. 2 个不同的特征表示将表征 2 个不同手势的 20 个关节点位置信息。

证明. 若特征为 F 对应的关节点位置信息为 $[\bar{W}_0, \bar{W}_1, \bar{W}_2, \dots, \bar{W}_{19}]$; 则特征为 $F + \Delta F$ 对应的关节点位置信息设为 $[\bar{W}_0^*, \bar{W}_1^*, \bar{W}_2^*, \dots, \bar{W}_{19}^*]$ 如下:

$$F + \Delta F = [Q_0 + \Delta Q_0, \dots, Q_{13} + \Delta Q_{13}, S_0 + \Delta S_0, \dots, S_3 + \Delta S_3, D_0 + \Delta D_0, \dots, D_4 + \Delta D_4],$$

则当 $[\Delta Q_0, \dots, \Delta Q_{13}, \Delta S_0, \dots, \Delta S_3, \Delta D_0, \dots, \Delta D_4]$ 至少有一个不为 0 时, 下面需证明 $[\bar{W}_0, \bar{W}_1, \bar{W}_2, \dots, \bar{W}_{19}]$ 与 $[\bar{W}_0^*, \bar{W}_1^*, \bar{W}_2^*, \dots, \bar{W}_{19}^*]$ 中至少有一个不同。

(1) 中指关节点坐标:

由式(4)'可知分别为

$$\begin{aligned}\bar{W}_0^* &= (0, 0, 0), \bar{W}_8^* = (0, 0, L_7), \bar{W}_9^* = (x_{31}^*, 0, y_{31}^*), \\ \bar{W}_{10}^* &= (x_{32}^*, 0, y_{32}^*), \bar{W}_{11}^* = (x_{33}^*, 0, y_{33}^*).\end{aligned}$$

与式(1)'推导类似可得:

$$\begin{aligned}x_{31}^* &= L_8 \cdot \cos \left(2 \cdot \arccos(Q_5 + \Delta Q_5) - \frac{\pi}{2} \right), \\ y_{31}^* &= L_7 + L_8 \cdot \sin \left(2 \cdot \arccos(Q_5 + \Delta Q_5) - \frac{\pi}{2} \right).\end{aligned}$$

而 $\bar{W}_0 = (0, 0, 0), \bar{W}_8 = (0, 0, L_7), \bar{W}_9 = (x_{31}, 0, y_{31}), \bar{W}_{10} = (x_{32}, 0, y_{32}), \bar{W}_{11} = (x_{33}, 0, y_{33})$ 。

其中:

$$x_{31} = L_8 \cdot \cos \left(2 \cdot \arccos Q_5 - \frac{\pi}{2} \right),$$

$$y_{31} = L_7 + L_8 \cdot \sin \left(2 \cdot \arccos Q_5 - \frac{\pi}{2} \right).$$

若 $x_{31}^* = x_{31}$, 则 $\Delta Q_5 = 2K\pi, K = \dots, -2, -1, 0, 1, 2, \dots$ 。

由于 $0 \leq Q_5 + \Delta Q_5 < \pi, 0 \leq Q_5 < \pi$, 故若 $\Delta Q_5 \neq 0$, 则 $x_{31}^* \neq x_{31}, y_{31}^*$ 情形同理可得。

与式(2)'推导类似可得:

$$x_{32}^* = x_{31}^* + L_5 \cdot \cos \left(2 \cdot \arccos(Q_6 + \Delta Q_6) - \frac{\pi}{2} \right),$$

$$x_{32} = x_{31} + L_5 \cdot \cos \left(2 \cdot \arccos Q_6 - \frac{\pi}{2} \right).$$

若 $\Delta Q_5 \neq 0$, 则有 $x_{31}^* \neq x_{31}$, 无需再讨论 x_{32}^* 。

若 $\Delta Q_5 = 0$, 且 $\Delta Q_6 \neq 0$, 则有 $x_{32}^* \neq x_{32}$ 。

同理可知, 若 $\Delta Q_5, \Delta Q_6, \Delta Q_7$ 中有一个不为 0, 则中指关节点坐标中至少有一个不同。

(2) 食指关节点坐标:

由式(5)'和式(6)'可知为

$$\begin{aligned}\bar{W}_4 &= (-m_1, 0, n_1), \bar{W}_5 = (-m_1, x_{21}, y_{21} - L_3 + n_1), \\ \bar{W}_6 &= (-m_1, x_{22}, y_{22} - L_3 + n_1), \\ \bar{W}_7 &= (-m_1, x_{23}, y_{23} - L_3 + n_1).\end{aligned}$$

$$\text{其中 } n_1 = \frac{S_1^2 - L_3^2 - (x_{23} - x_{33})^2 - (y_{23} - y_{33} - L_3)^2}{2(y_{23} - y_{33} - L_3)}.$$

若 $\Delta S_1 \neq 0$, 则 $n_1^* \neq n_1, x_{21}^*, x_{22}^*$ 与中指 x_{31}^*, x_{32}^* 同理, 可得若 $\Delta Q_2, \Delta Q_3, \Delta Q_4, \Delta S_1$ 中有一个不为 0, 则食指关节点坐标中至少有一个不同。

(3) 无名指关节点坐标: $\bar{W}_{12}^*, \bar{W}_{13}^*, \bar{W}_{14}^*, \bar{W}_{15}^*$ 与食指关节点坐标 $\bar{W}_4^*, \bar{W}_5^*, \bar{W}_6^*, \bar{W}_7^*$ 可类似证明。

(4) 小指关节点坐标:

由式(7)'可得为

$$\bar{W}_{16} = (m_3, 0, n_3), \bar{W}_{17} = (m_3, x_{51}, y_{51} - L_{15} + n_3),$$

$$\bar{W}_{18} = (m_3, x_{52}, y_{52} - L_{15} + n_3),$$

$$\bar{W}_{19} = (m_3, x_{53}, y_{53} - L_{15} + n_3).$$

$$\text{其中 } n_3 = \frac{-a_3 b_3 + \sqrt{a_3^2 b_3^2 - (a_3^2 + 1)(b_3^2 - L_{15}^2)}}{a_3^2 + 1},$$

而:

$$\begin{cases} a_3 = \frac{y_{53} - y_{43} - L_{15} + L_{11} - n_2}{m_2} \\ b_3 = \frac{S_3^2 - L_{15}^2 - m_2^2 - (x_{53} - x_{43})^2 - (y_{53} - y_{43} - L_{15} + L_{11} - n_2)^2}{-2m_2} \end{cases}$$

若 $m_2^* \neq m_2$ 或 $n_2^* \neq n_2$, 则 $\bar{W}_{12}^* \neq \bar{W}_{12}$, 即有一坐标不同. 故不妨设 $m_2^* = m_2$ 且 $n_2^* = n_2$, 若 $\Delta S_3 \neq 0$, 则 $b_3^* \neq b_3$, 且 $a_3^* = a_3$ 始终成立.

假设存在 $b_3^* \neq b_3$, 使 $n_3^* = n_3$, 但由于 $m_3 = b_3 + a_3 n_3$, 故 $m_3^* \neq m_3$, 故 $\bar{W}_{16}^* \neq \bar{W}_{16}$; 若不存在 $b_3^* \neq b_3$, 使得 $n_3^* = n_3$, 亦有 $\bar{W}_{16}^* \neq \bar{W}_{16}$.

对于 $\bar{W}_{17}^*, \bar{W}_{18}^*, \bar{W}_{19}^*$, 同理可得: 若 $\Delta Q_{11}, \Delta Q_{12}, \Delta Q_{13}, \Delta S_4$ 中有一个不为 0, 则小指关节坐标中至少有一个不同.

(5) 大拇指关节坐标:

由定理 1 可知, 大拇指坐标需要分 2 步求解, 第一步不考虑大拇指关节的屈曲, 即先将大拇指指平面平移至大拇指掌端关节与其余 4 指掌端关节在同一平面内, 得到大拇指关节点 $\bar{W}'_1, \bar{W}'_2, \bar{W}'_3$; 第二步根据大拇指的屈曲自由度对大拇指指平面进行旋转得到最终的大拇指关节坐标 $\bar{W}_1, \bar{W}_2, \bar{W}_3$.

已知若平面旋转前各关节点相对位置发生了改变, 由于旋转变换是刚性变换, 则旋转后得到的各点坐标将发生变化, 即若 $\bar{W}'_1, \bar{W}'_2, \bar{W}'_3$ 与 $\bar{W}'_1, \bar{W}'_2, \bar{W}'_3$ 不同, 则旋转后所得 3 点坐标亦不可能与 $\bar{W}_1, \bar{W}_2, \bar{W}_3$ 相同. 故仅需讨论第一步中大拇指关节点 $\bar{W}'_1, \bar{W}'_2, \bar{W}'_3$ 是否与 $\bar{W}'_1, \bar{W}'_2, \bar{W}'_3$ 相同.

在不考虑大拇指关节屈曲时, 关节点坐标 $\bar{W}'_1, \bar{W}'_2, \bar{W}'_3$ 与 $\bar{W}_{17}^*, \bar{W}_{18}^*, \bar{W}_{19}^*$ 同理计算, 可得结论如下: 若 $\Delta Q_0, \Delta Q_1, \Delta S_0$ 中

至少有一个不为 0, 则等式 $\bar{W}'_1 = \bar{W}'_1, \bar{W}'_2 = \bar{W}'_2, \bar{W}'_3 = \bar{W}'_3$ 中至少有一个不成立. 其中 $\Delta S'_0$ 表示大拇指掌端关节与其余四指的掌端关节在同一平面内时大拇指指尖与食指指尖的距离变化量.

由于屈曲自由度未发生变化, 故旋转矩阵亦相同, 可得:

$$\begin{cases} \|\bar{W}_3^* - \bar{W}_7\| = S_0 + \Delta S_0 \\ \|\bar{W}_3'^* - \bar{W}_7\| = S'_0 + \Delta S'_0 \\ \bar{W}_3^{*T} = R\bar{W}_3'^{*T} \end{cases} \quad \begin{cases} \|\bar{W}_3 - \bar{W}_7\| = S_0 \\ \|\bar{W}_3' - \bar{W}_7\| = S'_0 \\ \bar{W}_3^T = R\bar{W}_3'^T \end{cases}$$

若 $\bar{W}'_1, \bar{W}'_2, \bar{W}'_3$ 与 $\bar{W}_1, \bar{W}_2, \bar{W}_3$ 相同, 则 $\Delta Q_0, \Delta Q_1, \Delta S_0$ 一定全为 0. 因而, 若 $\Delta Q_0, \Delta Q_1, \Delta S_0$ 中有一个不为 0, 则大拇指关节坐标中至少有一个不同.

此外, 特征 $\Delta D_0, \dots, \Delta D_4$ 表示 5 个手指指尖到手腕关节点间距离的变化, 而 5 个手指关节坐标确定后, 关节点间距离唯一确定, 该距离特征不能单独产生变化量.

综上所述: $[\bar{W}_0^*, \bar{W}_1^*, \bar{W}_2^*, \dots, \bar{W}_{19}^*]$ 中关节坐标至少有一个与 $[\bar{W}_0, \bar{W}_1, \bar{W}_2, \dots, \bar{W}_{19}]$ 中不同, 即 2 个不同特征表征 2 个不同手势的 20 个关节点位置信息. 结合定理 1 和定理 2 可得, 本文的手势特征能唯一地表征手势关节点的位置信息.



MIAO Yong-Wei, Ph. D., professor, Ph. D. supervisor. His current research interests include computer graphics, digital geometry processing, computer vision, machine learning.

LI Jia-Ying, M. S. candidate. Her current research interests include computer graphics, machine learning.

LIU Jia-Zong, M. S. candidate. His current research interests include computer graphics, computer vision.

CHEN Jia-Zhou, Ph. D., associate professor. His current research interests include computer graphics, computer vision.

SUN Shu-Sen, Ph. D., associate professor. His current research interests include computer graphics, virtual reality.

Background

As a mainstream interactive mode and most powerful method of HCI (Human-Computer Interaction), due to its high degree of freedom, hand gesture and its recognition has become a hot research topic in the literature of Computer Graphics, Virtual Reality and HCI, etc. Unlike the traditional of two-dimensional hand gesture recognition from photographs or color images, the three-dimensional hand gesture recognition has attracted increasing research interests. The static 3D hand pose estimation and recognition methods can be classified into following types: (1) discriminative approaches, (2) generative approaches, and (3) hybrid of discriminative and generative approaches. To solve the problems of high degree of freedom and inaccurate feature representation of different hand gestures in static gesture classification and recognition, we

proposed a new intrinsic feature representation of 3D hand gestures in this paper, which combining joint rotation feature and fingertip distance feature. Extensive experimental results demonstrated that the proposed method can achieve high accuracy in static gesture classification and recognition. The research result can be applied for future applications on RGB-D depth camera based 3D human-computer interaction.

This research is supported by the National Natural Science Foundation of China under Grant No. 61972458, the Science and Technology Planning Project of Zhejiang Province under Grant Nos. LY18F020035, GG19F020006 and the Science Foundation of Zhejiang Sci-Tech University under Grant No. 17032001-Y.