

# 基于隐马尔科夫模型的人体动作压缩红外分类

关秋菊<sup>1</sup> 罗晓牧<sup>2</sup> 郭雪梅<sup>3</sup> 王国利<sup>3</sup>

**摘要** 人体动作产生的辐射能量变化 (Infrared radiation changes, IRC) 信号是动作识别的重要线索, 本文提出了一种基于隐马尔科夫模型的人体动作压缩红外分类新方法. 针对人体动作的自遮挡问题, 建立基于正交视角的压缩红外测量系统, 获取人体动作在主投影面和辅助投影面的 IRC 压缩信号; 然后, 采用隐马尔科夫模型 (Hidden Markov model, HMM) 双层特征建模算法进行压缩域动作分类. 实验结果表明双层特征建模的平均正确分类率高于主层特征建模, 平均正确分类率可达 95.71%. 该方法为环境辅助生活系统提供了人体动作识别的新途径.

**关键词** 环境辅助生活, 隐马尔科夫模型, 压缩感知, 热释电红外传感器, 动作分类

**引用格式** 关秋菊, 罗晓牧, 郭雪梅, 王国利. 基于隐马尔科夫模型的人体动作压缩红外分类. 自动化学报, 2017, 43(3): 398–406

**DOI** 10.16383/j.aas.2017.c160130

## Compressive Infrared Classification of Human Motion Using HMM

GUAN Qiu-Ju<sup>1</sup> LUO Xiao-Mu<sup>2</sup> GUO Xue-Mei<sup>3</sup> WANG Guo-Li<sup>3</sup>

**Abstract** Infrared radiation changes (IRC) induced by human motion can provide important clue for motion classification. This paper presents a hidden Markov model (HMM)-based compressive infrared classification method to recognize human motions. In order to solve the problem of self-occlusion, an orthogonal-view based compressive infrared sensing system is implemented for projecting the IRC to two orthogonal planes in the infrared radiation field. Then, a double-layer feature model using HMM classifier is trained to carry out motion recognition with the compressive measurements. Experimental results show that the mean correct classification rate with double-layer feature is 95.71%, which is better than that with main-layer feature. This method provides a new approach to classification of human motions for ambient assisted system.

**Key words** Ambient assisted living (AAL), hidden Markov model (HMM), compressive sensing, pyroelectric infrared sensors, motion classification

**Citation** Guan Qiu-Ju, Luo Xiao-Mu, Guo Xue-Mei, Wang Guo-Li. Compressive infrared classification of human motion using HMM. *Acta Automatica Sinica*, 2017, 43(3): 398–406

环境辅助生活 (Ambient assisted living, AAL) 系统是辅助居家老人延长独立生活的智能系统<sup>[1–2]</sup>.

收稿日期 2016-02-03 录用日期 2016-09-30  
Manuscript received February 3, 2016; accepted September 30, 2016

国家自然科学基金 (61375080, 61301294, 61601523), 广东省自然科学基金 (2015A030311049, 2016A030310238), 广东省教育厅青年创新人才项目 (2015KQNCX068) 资助

Supported by National Natural Science Foundation of China (61375080, 61301294, 61601523), Natural Science Foundation of Guangdong Province (2015A030311049, 2016A030310238), and Young Creative Talents Project of Guangdong Province Education Department (2015KQNCX068)

本文责任编辑 王启宁

Recommended by Associate Editor WANG Qi-Ning

1. 仲恺农业工程学院机电工程学院 广州 510225 2. 广州中医药大学医学信息工程学院 广州 510006 3. 中山大学数据科学与计算机学院 广州 510006

1. College of Mechanical and Electrical Engineering, Zhongkai University of Agriculture Engineering, Guangzhou 510225 2. School of Medical Information Engineering, Guangzhou University of Chinese Medicine, Guangzhou 510006 3. School of Data and Computer Science, Sun Yat-Sen University, Guangzhou 510006

人机物 (环境) 深度融合是“以人为中心” AAL 系统的显著特征, 而人体行为理解是实现人机物深度融合的重要纽带, 只有当系统能正确理解人体行为, 才能更好地给予老人生活协助. 行为理解是利用 AAL 技术应对日益严重的人口老龄化社会问题中需要解决的基础课题. 因其巨大的潜在应用前景和固有的科学挑战, 开展行为理解基础研究有着重要的学术价值和实际意义. 其中, 人体目标的动作识别是实现行为理解的基础和关键技术要素.

目前, 人体动作识别研究主要有两个方向: 基于视觉感知的动作识别和基于非视觉感知的动作识别. 基于视觉感知的动作识别是使用最为广泛的方法, 它通过视觉传感器采集人体动作视频或数字化图像序列, 再利用计算机视觉技术处理和分析图像数据, 进而实现动作识别<sup>[3–6]</sup>. 该方法通过从图像数据中提取出与运动线索相关的有用信息, 形成基于光流、轨迹、外观、形状或轮廓等线索或其组合策

略的多种动作特征提取及识别方法。值得关注的是, Johansson 等的经典实验证实, 生物视觉通过观察放置在肢体关节处的点光源的运动状态, 能自然地推断出步态的模式<sup>[7]</sup>。这表明, 视觉传感以同构成像方式获得的高维图像数据中含有相当规模的信息与行为感知无关。传统奈奎斯特规则支配的视觉传感模式耗费网络资源产生的大数据, 与面向动作识别任务的信息处理模式对特征数据简洁化的约束, 两者之间存在严重不对称的数据鸿沟。

基于非视觉传感的动作识别可进一步分为基于人体穿戴式传感和基于环境部署式传感两类。基于人体穿戴式传感的动作识别系统, 利用部署在人体上的动作传感器, 获取人体动作产生的位移、关节弯曲角度或加速度等运动信号, 并进行特征提取与分析, 实现动作的分析与理解。常用的动作传感器有陀螺仪、加速度计和角速率计等<sup>[8-9]</sup>。其中, 加速度传感器是最常用的穿戴式动作传感器, 特别适用于走、跑、爬楼梯等周期重复性动作的识别。但是, 穿戴式传感器测量时可能在人体表面滑动, 这不仅引入测量数据误差, 有时还会限制人体动作, 在一定程度上影响了用户体验。基于环境部署式传感的动作识别, 通常是在人体所处物理空间中(如, 墙壁、走廊或室内天花板)部署大量低成本、低功耗的微型传感器, 通过监测人体动作引起的环境状态变化, 识别人体动作。该方法基于“人-物交互”现象: 人体动作时, 必然会和周边环境中的物体产生交互, 而这种交互会引起环境中某些物理量状态的变化。常用的动作传感器有接触开关、压力传感器、热释电红外(Pyroelectric infrared, PIR)传感器等。其中, PIR传感器能以非接触的方式探测环境中人体运动引起的辐射场能量变化。2006年, Burchett等采用3个PIR传感单元组合的顶视感知方式, 测量人体运动产生的热辐射变化信号, 并基于主成分回归方法, 识别出常速走、慢走和挥臂三类典型人体动作<sup>[10]</sup>。基于环境部署式传感的动作识别方式, 便于进行大规模网络部署, 且使人体运动更自然, 尤其适用于AAL系统。

压缩感知(Compressive sensing, CS)是一种压缩与采样合并进行的信息获取新理论<sup>[11-12]</sup>, 它以远低于奈奎斯特频率进行采样, 采用非自适应线性投影保持信号的原始结构。人体作为红外辐射源, 运动产生的辐射能量变化(Infrared radiation changes, IRC)信息是动作识别的重要线索。PIR传感器是获取人体动作IRC的被动式传感手段, 与几何参考结构相结合, 形成压缩感知支配的非同构被动光学成像机制, 即红外动作的参考结构层析成像技术<sup>[13-14]</sup>。由此形成的压缩红外感知机制, 可以

直接面向特征信息进行压缩采样与处理, 避免了传统奈奎斯特规则支配的视觉模式下信息获取与处理之间的数据鸿沟。

基于CS的压缩分类是一种新的信息处理方法, 它通过分析处理低维压缩域测量值间接实现高维原始数据的分类<sup>[15-16]</sup>。这种新的信息获取与处理机制能有效避免无效或冗余信息的产生和处理过程, 有助于降低传感系统的复杂度和硬件成本, 易形成轻量级的信息处理算法, 便于构建资源受限的无线传感器网络。

目前采用压缩红外感知技术手段识别人体动作已有一些初步研究。2012年, Luo等通过在天花板上安装IRC压缩感知模块(由7个PIR传感器和几何参考结构组成), 将目标空间划分为17个采样胞元, 利用PIR传感数据流的特征分析, 识别人体摔倒动作<sup>[17]</sup>。2014年, Sun等在室内布置4个IRC测量节点, 每个节点由4个PIR传感器和几何参考结构组合而成, 用于获取辐射场空间14个采样胞元内的IRC信息, 基于传感输出信号, 实现人体运动场景识别<sup>[18]</sup>。上述研究中, 人体红外测量系统对感知空间的观测为粗粒度形态, 适用于人体移动层的运动识别。2014年, Guan等采用单个PIR传感阵列和几何参考结构组合的形式, 将辐射场空间划分为64个采样胞元, 基于压缩测量数据进行10类平面动作的识别<sup>[19]</sup>。该研究提高了感知空间的观测粒度, 易于实现肢体动作层识别; 但由于单一视角的限制, 仅适用于平面动作识别。

本文针对人体动作IRC信号的特点, 结合压缩分类理论, 提出一种新的红外辐射场人体动作分类方法。在动作信号获取方面, 利用PIR传感阵列和掩膜阵列组合来实现红外动作压缩感知, 该机制不仅提高了人体红外辐射场空间IRC信息的观测粒度, 还兼具数据降维效果; 采用正交视角融合的压缩红外感知方式获取IRC信息, 解决三维动作自遮挡问题。在动作信号分析方面, 针对红外动作IRC信号的时序特点, 在低维压缩测量域采用隐马尔科夫模型进行动作模式分类。

## 1 理论基础

### 1.1 压缩分类

压缩分类是压缩域信号处理的基本问题之一, 它的基本思想在于, 利用测量矩阵的限制等距性质(Restricted isometry property, RIP), 将高维空间的信号分类问题, 映射到低维测量空间来求解<sup>[16]</sup>。下面简要介绍压缩分类建模过程。

假设高维目标空间 $\mathbf{R}^N$ 中有 $L$ 类源信号

$\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_L\}$ , 目标信号  $\mathbf{x}$  的类别未知, 但已知它属于这  $L$  类信号中的其中一类. 建模目的是确定目标信号  $\mathbf{x}$  的类别, 即从  $L$  种假设情形中推断信号  $\mathbf{x}$  的真实类别:

$$\tilde{\mathcal{H}}_i: \mathbf{x} = \mathbf{s}_i + \mathbf{n}, \quad i = 1, 2, \dots, L \quad (1)$$

其中, 噪声  $\mathbf{n} \sim N(0, \sigma^2 \mathbf{I}_N)$ ,  $\sigma^2$  为噪声方差,  $\mathbf{I}_N$  为  $N$  维单位向量.

在高维目标空间内, 构造统计量:  $t_i = \|\mathbf{x} - \mathbf{s}_i\|_2$ , 即目标信号和源信号的距离. 根据最近邻分类准则:

$$i^* = \arg \min_{i=1, \dots, L} \|\mathbf{x} - \mathbf{s}_i\|_2 \quad (2)$$

$\mathbf{x}$  即为上式取最小值时对应的第  $i$  个信号. 但是, 在高维目标空间中直接进行信号分类带来的巨大计算量, 不容忽视.

为了降低计算量, 可以考虑在低维的压缩测量域构造统计量进行信号分类.

采用具有 RIP 的投影矩阵  $\Phi \in \mathbf{R}^{M \times N}$ ,  $M \ll N$ , 把源信号从目标空间  $\mathbf{R}^N$  投影至低维空间  $\mathbf{R}^M$ , 这样,  $L$  类信号的压缩投影值分别为  $\Phi \mathbf{s}_1, \Phi \mathbf{s}_2, \dots, \Phi \mathbf{s}_L$ , 目标信号  $\mathbf{x}$  的压缩投影值  $\mathbf{y} = \Phi \mathbf{x}$ . 同样的, 压缩信号  $\mathbf{y}$  的类别也有  $L$  种可能情形:

$$\tilde{\mathcal{H}}_i: \mathbf{y} = \Phi \mathbf{x} = \Phi(\mathbf{s}_i + \mathbf{n}), \quad i = 1, 2, \dots, L \quad (3)$$

假设以上情形等概率发生, 可以构造统计量<sup>[20]</sup>

$$t_i := (\mathbf{y} - \Phi \mathbf{s}_i)^\top (\Phi \Phi^\top)^{-1} (\mathbf{y} - \Phi \mathbf{s}_i) \quad (4)$$

分析信号  $\mathbf{y}$  的类别. 特别的, 当  $\Phi$  为正交矩阵 (或其行向量互相正交) 时,  $\Phi \Phi^\top = \mathbf{E}$ , 则式 (4) 可化简为  $t_i = \|\mathbf{y} - \Phi \mathbf{s}_i\|_2^2$ . 在低维压缩域中, 依最近邻原则:

$$i^* = \arg \min_{i=1, \dots, L} \|\mathbf{y} - \Phi \mathbf{s}_i\|_2^2 \quad (5)$$

压缩测量信号  $\mathbf{y}$  得以识别.

实际上, 利用投影矩阵  $\Phi$  的 RIP, 可使信号之间的距离在投影前后近似保持不变, 即  $\|\mathbf{x} - \mathbf{s}_i\|_2 \approx \|\mathbf{y} - \Phi \mathbf{s}_i\|_2$ , 从而实现将高维空间的分类问题投影到低维空间求解的目的.

## 1.2 基于 GMHMM 的压缩红外分类

利用红外压缩测量, 人体动作在高维红外辐射场空间产生的 IRC 可被压缩投影至低维测量空间. 人体动作产生的 IRC 是一个动态连续过程. 相应的, 在低维测量空间, PIR 传感数据是连续的电压时序信号. 隐马尔科夫模型 (Hidden Markov model,

HMM) 具有很强的时序推演建模能力, 是分析时间序列的有力工具, 已广泛用于语音<sup>[21]</sup>、视频时序信号处理<sup>[22]</sup>. 本文选用连续型的高斯混合隐马尔科夫模型 (Gaussian mixture HMM, GMHMM) 对红外动作在低维压缩测量域进行模式分类.

首先定义 GMHMM 中的模型参数. 已知人体动作是由一系列静止姿态在时域上动态变化而形成的. 这里, 将具有代表性的静止姿态定义为 GMHMM 的隐状态, 一个动作所包含的代表性静止姿态的数目, 即 GMHMM 隐状态数目. 连续动作, 从一个静止姿态变化转移到另一个静止姿态时, 会导致时空域辐射能量变化, 并以 PIR 传感器阵列输出时序信号的形式呈现, 即 GMHMM 的观测值序列. 连续动作各个隐状态之间的转移模式会在观测值序列中体现出来. 因此, 利用观测值序列作为样本进行模型训练, 可以得到各个动作的 GMHMM 模型, 它们可以很好地解释特定动作类别中的样本. 一个 GMHMM 可简记为  $\lambda = (H, M_G, \mathbf{A}, \mathbf{B}, \mathbf{\Pi})$ , 各参数意义如下:

1)  $H$ : 模型的隐状态数目;

2)  $M_G$ : 高斯概率密度函数的数目, 这里利用高斯概率密度函数来拟合隐状态被观测值观察到的概率分布, 同时为了提高拟合的准确性, 可以采用多个高斯概率密度函数混合表示输出观测值的概率分布;

3)  $\mathbf{A} = \{a_{ij}\}$ : 隐状态转移概率分布矩阵,  $\mathbf{A} \in \mathbf{R}^{H \times H}$ , 其中,

$$a_{ij} = P(q_t = S_j | q_{t-1} = S_i), \quad 1 \leq i, j \leq H$$

表示从隐状态  $S_i$  到  $S_j$  的转移概率, 其中  $q_t$  为  $t$  时刻的隐状态, 隐状态转移概率分布满足:  $a_{ij} \geq 0$  且  $\sum_{j=1}^H a_{ij} = 1$ ;

4)  $\mathbf{B} = \{b_i(\mathbf{V})\}$ : 观测值序列的概率密度分布,  $\mathbf{B} \in \mathbf{R}^{H \times M_G}$ , 当隐状态为  $S_i$  时, 模型输出观测值  $\mathbf{V}$  的概率为

$$b_i(\mathbf{V}) = \sum_{m=1}^{M_G} c_{im} G(\boldsymbol{\mu}_{im}, \boldsymbol{\Sigma}_{im}, \mathbf{V}) \quad (6)$$

其中,  $G$  为高斯概率密度函数,  $c_{im}$  为在隐状态  $S_i$  下, 第  $m$  个高斯函数的权重因子, 满足  $c_{im} \geq 0$  和  $\sum_{m=1}^{M_G} c_{im} = 1$ ,  $\boldsymbol{\mu}_{im}$  和  $\boldsymbol{\Sigma}_{im}$  分别为该高斯函数的均值向量和协方差矩阵;

5)  $\mathbf{\Pi} = \{\pi_i\}$ : 初始状态矢量,  $\mathbf{\Pi} \in \mathbf{R}^{H \times 1}$ , 其中,

$$\pi_i = P\{q_1 = S_i\}$$

满足约束  $\pi_i \geq 0$ ,  $\sum_{i=1}^H \pi_i = 1$ .

其次, 基于 GMHMM 建模的红外动作分类基本步骤如下: 1) 动作模型训练: 假设共有  $L$  类红外动作, 确定模型参数  $H, M_G$  后, 利用部分观测值序列作为训练样本, 对 GMHMM 进行训练, 得到各个动作的模型  $\lambda_i, i = 1, 2, \dots, L$ ; 2) 测试动作识别: 根据输入的未知动作压缩投影观测值  $\mathbf{V}$ , 计算每个 GMHMM  $\lambda$  输出观测值  $\mathbf{V}$  的似然概率;  $\mathbf{V}$  被归类为产生最大似然概率的模型对应的动作类别, 即

$$i^* = \arg \max_{i=1, 2, \dots, L} P(\mathbf{V}|\lambda_i)$$

## 2 压缩红外动作信号获取及 GMHMM 动作建模

### 2.1 基于正交视角的压缩红外动作测量

红外动作信号测量系统基于 IRC 压缩测量模块进行, 该模块主要由 PIR 传感器阵列、掩膜阵列和无线收发模块构成<sup>[19]</sup>. 基于正交视角的压缩红外动作测量系统共采用两个这样的模块. 主投影面内的压缩测量模块作为正视测量节点, 用于 IRC 主层特征的压缩投影, PIR 传感器阵列和掩膜阵列联合实现  $64 \times 9$  维压缩投影矩阵, 其中的掩膜阵列将感知区域剖分为 64 个感知胞元, 采用 9 个 PIR 传感器对这些胞元内的 IRC 进行压缩测量. 辅助投影面内的压缩测量模块作为顶视测量节点, 用于 IRC 辅助层特征的压缩投影, PIR 传感器阵列和掩膜阵列联合实现  $16 \times 5$  维压缩投影矩阵. 利用上述面向双层特征获取的 3D 动作压缩红外测量系统, 辐射场空间  $64 \times 16$  维动作特征经压缩投影, 转换为低维空间 14 维特征向量序列.

图 1 所示为 3D 动作压缩红外测量实验场景: 正视测量节点位于主投影面 ( $xy$  平面), 与人体的水平间距为 150 cm; 顶视测量节点位于辅助投影面 ( $yz$  平面), 挂顶方式部署在人体正上方的天花板, 距离地面 300 cm; 汇聚节点放置在桌面上方, 将接收到的传感数据实时传送给上位机. 实验中, 测试对象在限定空间范围内动作, 且人体正面朝向正视测量节点.

人体动作识别研究中, 由于缺乏统一动作数据集, 很多研究者通常在各自设计的数据集上验证动作识别方法, 这些数据集包含不同种类和数目的动作. 经对比分析现有动作数据集, 我们从中选择参考了文献 [23] 的数据集, 该动作集在受试者数目、动作类型数目和样本量等方面均具有合理的规模. 参照该数据集, 我们邀请 5 名受试者 (4 名男性, 1 名女性, 平均身高 170 cm, 平均体重 55 kg) 参与实验, 建立红外动作数据集. 其中 3 名受试者, 每人每个动



图 1 红外动作识别实验场景

Fig. 1 The experiment scenario for infrared motion classification

作做 10 遍, 共有 300 个动作样本, 用于建立动作模型的训练集. 另外 2 名受试者, 每人每个动作做 20 遍, 共有 400 个动作样本, 用于建立动作识别的测试集. 本文建立的动作集包含 10 类上肢动作: 挥右臂 (m1)、挥左臂 (m2)、挥双臂 (m3)、举右臂 (m4)、举左臂 (m5)、举双臂 (m6)、左臂下右臂上 (m7)、左臂上右臂下 (m8)、左臂前右臂后 (m9) 和左臂后右臂前 (m10). 其中, 前 8 类 (m1 ~ m8) 选自文献 [23] 中的动作类型; 为了进一步验证本文提出的动作识别方法对不同维度方向动作的辨别力, 我们在数据集中添加了两类动作: 左臂前右臂后 (m9)、左臂后右臂前 (m10). 总之, 这些动作具有不同维度的代表性, 既包含平行于  $xy$  平面的动作 (m1、m2 和 m3), 平行于  $xz$  平面的动作 (m7、m8、m9 和 m10), 也包含  $xyz$  三维空间内的动作 (m4、m5 和 m6). 图 2 是一名受试者重复做 5 遍动作 m3 “挥双臂” 时的 IRC 压缩信号时域波形, 横轴表示采样时间, 纵轴表示 PIR 传感器的输出电压值.

### 2.2 基于双层特征的 GMHMM 动作建模

特征选取是信号分类的前提, 应首先予以分析检验. 在 GMHMM 红外动作特征建模中, 为了避免模型过拟合或欠拟合问题, 需要采取措施提高模型泛化能力: 训练集选取中, 3 名受试对象具有不同的体型特征 (身高: 164 ~ 176 cm; 体重: 50 ~ 73 kg), 其动作具有多样性的样本分布特点, 更符合实际的

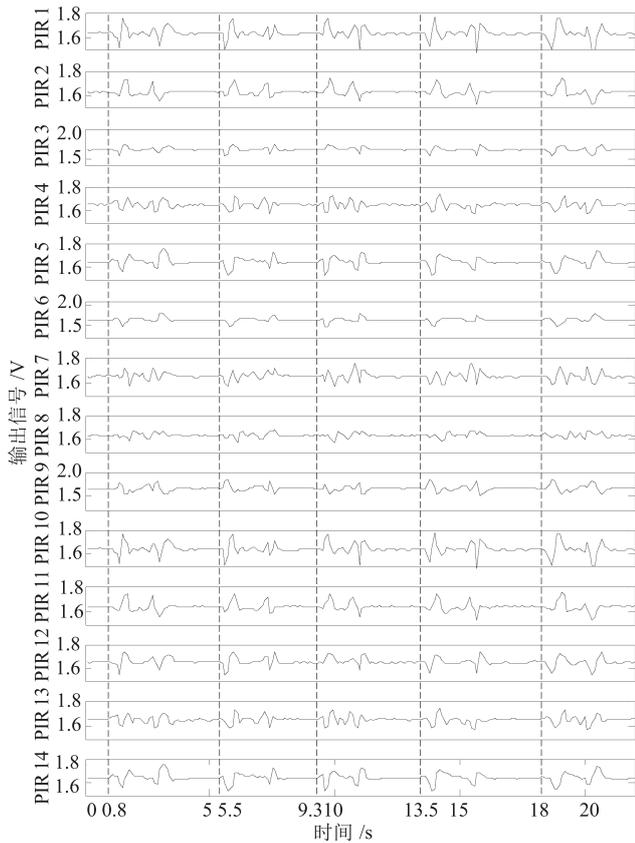


图2 人体动作压缩红外测量信号 (“m3挥双臂”动作)  
 Fig. 2 The PIR sensors output recorded for the motion of “waving two hands” (m3)

数据分布情况; 初始参数也会影响训练效果, 通过设置不同的参数组合 (隐状态数目、高斯模型数目等), 比较模型输出似然概率, 选择能产生似然概率最高

的模型. 本文中, 正视测量节点采用 9 个传感器, 顶视测量节点采用 5 个传感器, 即, 数据集中主层特征的压缩特征向量为 9 维, 双层特征的压缩特征向量为 14 维.

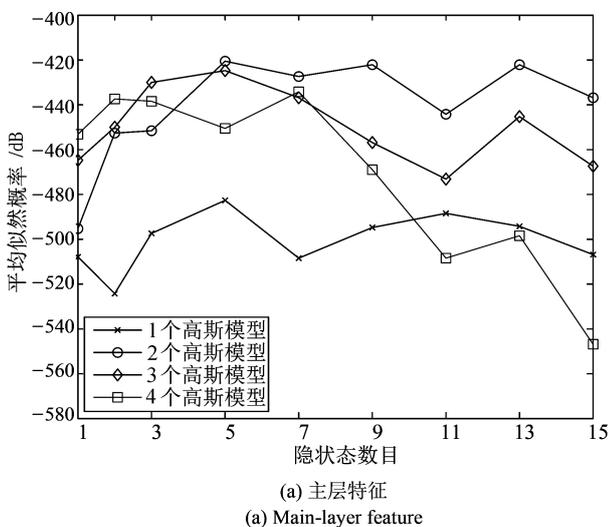
1) 基于不同特征层的动作模型训练

利用动作集中 300 个训练样本, 分别训练出两组 HMM 模型: 基于主层特征的主层 GMHMM 和基于双层特征的双层 GMHMM. 模型训练阶段的主要任务是模型  $\lambda = (H, M_G, A, B, \Pi)$  参数的确定.

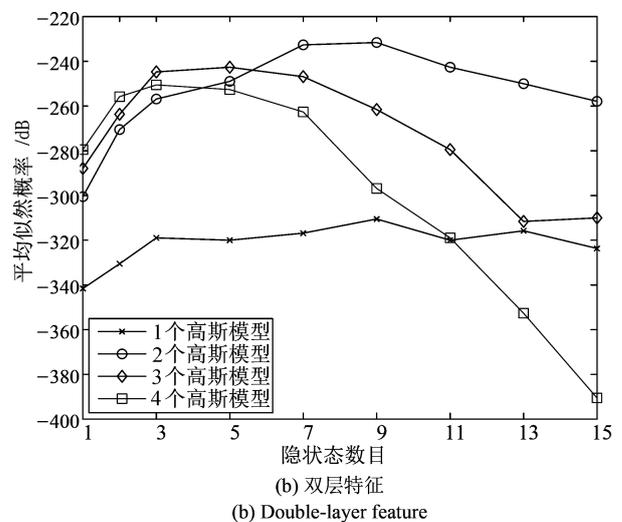
下面先观察参数隐状态数目  $H$ 、高斯模型数目  $M_G$  对模型输出似然概率的影响. 对于使用不同的参数组合, 动作 m1 的主层 GMHMM 建模平均似然概率输出, 如图 3(a) 所示; 双层 GMHMM 建模平均似然概率输出, 如图 3(b) 所示. 可以发现,  $H$  和  $M_G$  不同参数值组合下, 模型输出的平均似然概率不同. 即, 基于 GMHMM 建模的动作分类中, 隐状态数目  $H$  和高斯模型数目  $M_G$  参数变化, 对模型的识别性能影响较大.

为了保证动作分类的准确度, 基于每个动作模型的输出似然概率性能曲线, 选择可以使模型输出似然概率最大的隐状态数目  $H$  和高斯模型数目  $M_G$ . 各动作模型的最优配置参数如表 1 所示.

值得注意的是, 训练样本直接影响模型最优参数的取值. 为了正确选择最优参数, 我们在建立样本集时主要从三方面考虑: 样本量、样本分布和样本质量. 在样本量较少情况下, 最优参数受影响显著, 这就要求样本量应足够多, 确保建模的准确性和稳定性; 从样本分布角度看, 需要采集一个能够覆盖动作模式所有种类的样本集; 从样本质量角度看, 动作执行相对规范的样本质量较高, 更有利于准确建模.



(a) 主层特征  
 (a) Main-layer feature



(b) 双层特征  
 (b) Double-layer feature

图3 GMHMM 建模中隐状态数目、高斯模型数目对平均对数似然值的影响 (动作 m1)  
 Fig. 3 The influences of the HMM parameters on the average log likelihood (motion m1)

表1 GMHMMs 参数配置  
Table 1 The specification of GMHMMs

特征层模型	m1		m2		m3		m4		m5		m6		m7		m8		m9		m10	
	$M_G$	$H$																		
主层 GMHMM	2	5	2	7	2	9	4	3	2	7	2	6	2	5	2	3	2	5	2	6
双层 GMHMM	2	9	2	4	2	5	3	3	2	8	2	5	2	4	2	4	2	3	2	5

2) 特征层的分类性能对比

完成上述 GMHMM 训练后, 对动作的测试样本进行识别. 观察同一个测试样本 (m9) 在不同 GMHMM 下的识别结果: 图 4 为在 10 类动作主层 GMHMM 下的输出似然概率, 图 5 为在 10 类动作的双层 GMHMM 下的输出似然概率. 对比发现, 对同一个测试动作, 双层 GMHMM 输出似然概率普遍高于主层 GMHMM, 而且双层 GMHMM 对测试动作的区分度也更显著.

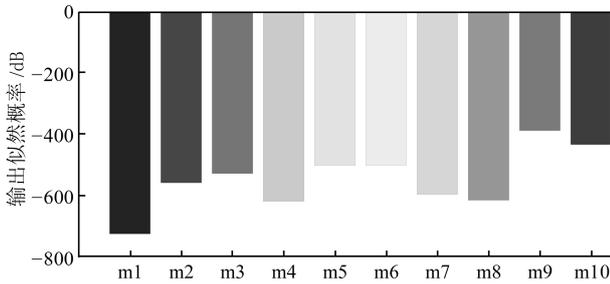


图 4 动作 m9 的一个测试样本在主层 GMHMM 下的对数似然概率

Fig. 4 Log likelihoods of a test sequence of motion m9 using main-layer GMHMM

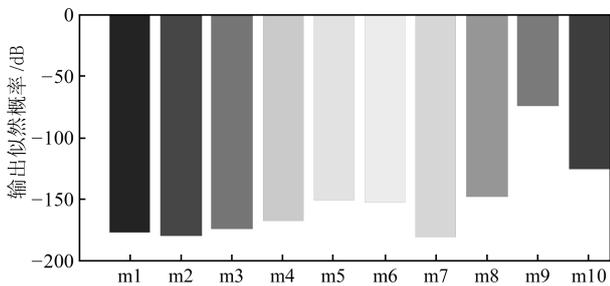


图 5 动作 m9 的一个测试样本在双层 GMHMM 下的对数似然概率

Fig. 5 Log likelihoods of a test sequence of motion m9 using double-layer GMHMM

3 实验及结果分析

测量维数和测量视角是压缩红外动作测量系统配置的两个关键参数, 它们不仅影响测量系统的复杂度, 还会直接影响动作分类的性能. 实验中, 采用正交视角红外动作测量平台获取实验数据, 基于

GMHMM 进行动作特征建模, 通过实验结果对比来分析测量维数和测量视角对 3D 动作分类性能的影响.

3.1 测量维数对动作识别性能的影响实验

测量维数 (PIR 传感器数目) 是体现传感效率的一个关键参数. 这里, 我们通过实验来分析测量维数对正确分类率 (Correct classification rate, CCR) 和分类时间的影响. 利用动作数据集中的 400 个测试样本和 GMHMM 分类方法, 得到如下的分析结果.

实验中, 针对任一测量维数  $M$ , 传感器组合的选择及相应 CCR 计算分为两种情况. 单一测量视角 (正视或顶视) 时, 从传感阵列中遍历选择  $M$  个传感器组合 (正视有  $C_9^M$  种形式, 顶视有  $C_5^M$  种形式), 对所有组合形式下的正确分类率求均值; 正交视角时, 正视传感阵列的 9 个传感器全部选用, 同时从顶视传感阵列中遍历选择  $M - 9$  个传感器组合 (共有  $C_5^{(M-9)}$  种形式), 二者进行数据融合后计算正确分类率, 并对所有组合形式下的正确分类率求均值.

图 6 为测量维数  $M$  增加时 CCR 的变化曲线.

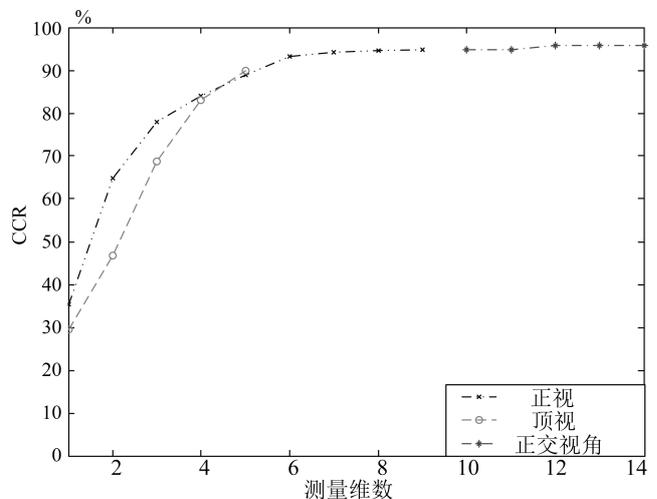


图 6 测量维数对正确分类率的影响

Fig. 6 The influence of the measurement dimension on the CCR

从整体上看,  $1 \leq M \leq 14$  范围内的 CCR 呈指数增长并趋于稳定. 进一步观察图中相应曲线, 可发现: 仅用正视测量节点时, 当  $1 \leq M \leq 6$  时, CCR 随  $M$  增加而迅速升高, 当  $6 \leq M \leq 9$  时, CCR 进一步缓慢提高; 仅用顶视测量节点时, 当  $1 \leq M \leq 5$  时, CCR 随  $M$  增加而迅速升高; 当正视、顶视测量节点协同工作时, 即  $10 \leq M \leq 14$ , 可以得到较满意且稳定的正确分类率.

为了考察动作分类的实时性问题, 需要统计模型训练时间和分类识别时间. 模型训练阶段, 当 PIR 传感器输出的 14 维压缩测量数据全部用来训练时, 建模所需时间为 7.583 s. 从图 7 中可以看出, 动作分类时间随测量维数的增加而略有延长, 但基本上可以满足动作分类的实时性需求.

测量维数的实际选择中, 须在保证 CCR 需求的前提下, 兼顾传感效率 (测量维数  $M$ )、识别时间 ( $t$ ) 以及测量系统的可靠性 (如, 个别 PIR 传感器出现异常的情况).

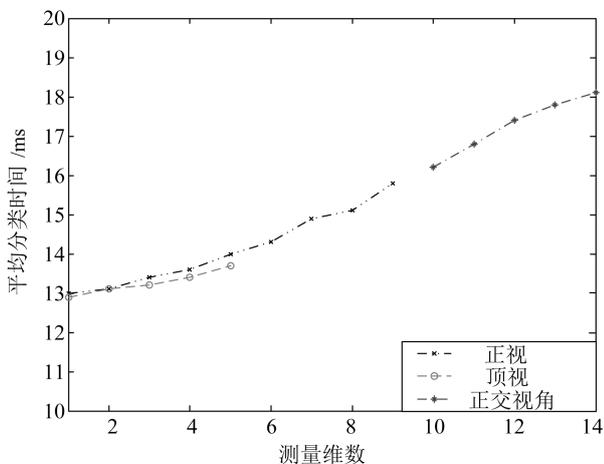


图 7 测量维数对分类时间的影响

Fig. 7 The influence of the measurement dimension on the running time

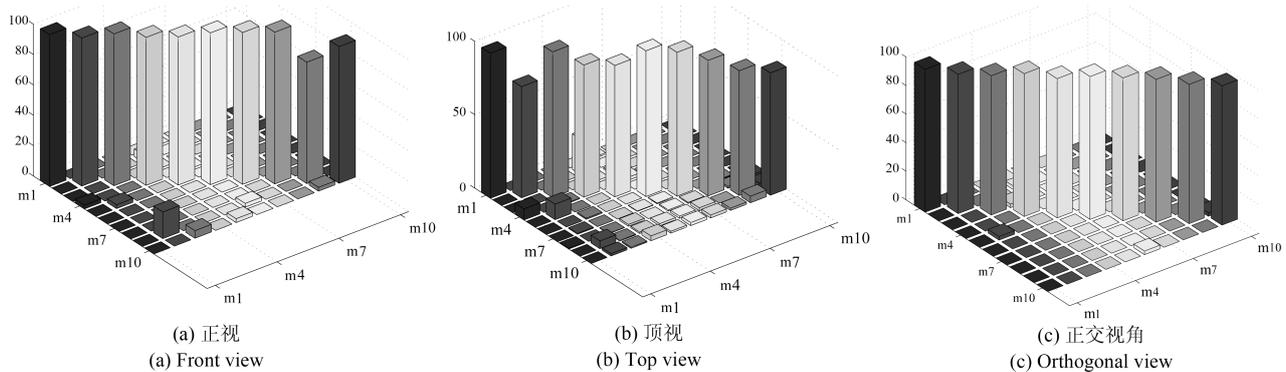


图 8 测量视角对正确识别率的影响

Fig. 8 The influences of the sensing view on the CCR

### 3.2 测量视角对 CCR 影响实验

首先, 考察正视、顶视和正交视角三种测量视角下的整体动作识别效果. 如表 2 所示, 正视视角下, 基于主层特征的 GMHMM 动作建模, CCR 可达 94.81%; 顶视视角下, 基于辅助投影面特征的 GMHMM 动作建模, CCR 偏低 (87.23%); 当采用正交视角获取双层特征时, CCR 最高达 95.71%. 可以发现, 从整体识别效果来看, 顶视测量视角对提高 CCR 的贡献并不大, 这是因为: 面向辅助投影面特征获取的顶视测量节点只具有粗粒度的观测效果, 而本实验动作集中的动作特征投影大多位于主投影面.

表 2 测量视角对 CCR 的影响

Table 2 Sensing view vs. CCR

测量视角	正视	顶视	正交视角
CCR (%)	94.81	87.23	95.71

其次, 进一步具体分析测量视角对单个动作的分类性能影响. 从图 8 所示的三维混淆矩阵中可发现, 正视视角下动作  $m1 \sim m8$  的 CCR 较高, 而  $m9$  和  $m10$  的 CCR 偏低. 这是因为前 8 个动作基本平行于主投影面, 即正视测量节点感知区域的敏感视场内; 而动作  $m9$  和  $m10$  在  $x$  和  $z$  两个方向上, IRC 特征无法全部投影至主投影面. 正视测量节点单独测量时, 只能获取  $m9$ 、 $m10$  的  $x$  方向特征数据. 再观察图 8 (b), 顶视测量节点单独测量时,  $m9$  和  $m10$  的 CCR 同样偏低; 这是因为辅助投影面的顶视传感阵列只能获取  $m9$ 、 $m10$  的  $z$  方向动作特征. 从图 8 (c) 中可以发现, 采用正交视角测量后, 动作  $m9$  和  $m10$  的 CCR 得到显著提高. 同时, 也可以观察到在动作  $m7$  和  $m8$  的识别中, 正交视角相对单独正视视角的改善效果.

可以看出, 对平面动作, 在动作投影面内布置单个测量节点, 即可达到满意的 CCR; 对 3D 动作, 基于正交视角测量方式获取的双层特征进行 GMHMM 动作建模分类, 更有助于提高 CCR.

## 4 结论

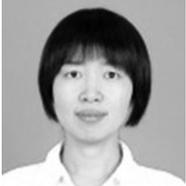
针对红外辐射场空间三维动作的自遮挡现象, 提出基于双层特征 GMHMM 建模的人体动作压缩红外分类方法. 该方法利用 PIR 传感阵列和参考结构组合构建的两个测量节点, 联合实现压缩红外正交投影测量系统, 该系统下, 三维动作  $64 \times 16$  维特征向量, 投影为测量空间 14 维特征向量. 基于测量系统获取的压缩红外测量值, 利用双层特征 GMHMM 动作建模来实现压缩域动作识别. 实验结果表明, 正交投影测量方式下, 采用双层特征建模得到的 CCR 达到 95.71%, 本文建立的动作识别系统可直接应用于固定位置的人机交互场景. 需要指出的是, 本系统的传感节点部署在有限空间的固定位置, 要求受试者在相应视场内接受动作识别. 这里, 视场受限的问题可以通过网络化分布式部署传感节点加以解决, 通过扩展传感视场满足实际的应用需求.

## References

- 1 Wichert R, Eberhardt B. Ambient assisted living. *Advanced Technologies & Societal Change*. Berlin, Germany: Springer-Verlag Berlin Heidelberg, 2012. 1145–1148
- 2 Chen L M, Hoey J, Nugent C D, Cook D J, Yu Z W. Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2012, **42**(6): 790–808
- 3 Aggarwal J K, Ryoo M S. Human activity analysis: a review. *ACM Computing Surveys*, 2011, **43**(3): 16
- 4 Hu W M, Tan T N, Wang L, Maybank S. A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2004, **34**(3): 334–352
- 5 Moeslund T B, Hilton A, Krüger V. A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 2006, **104**(2–3): 90–126
- 6 Turaga P, Chellappa R, Subrahmanian V S, Udrea O. Machine recognition of human activities: a survey. *IEEE Transactions on Circuits and Systems for Video Technology*, 2008, **18**(11): 1473–1488
- 7 Johansson G. Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 1973, **14**(2): 201–211
- 8 Shi G Y, Chan C S, Li W J, Leung K S, Zou Y X, Jin Y F. Mobile human airbag system for fall protection using MEMS sensors and embedded SVM classifier. *IEEE Sensors Journal*, 2009, **9**(5): 495–503
- 9 Yang A Y, Jafari R, Sastry S S, Bajcsy R. Distributed recognition of human actions using wearable motion sensor networks. *Journal of Ambient Intelligence and Smart Environments*, 2009, **1**(2): 103–115
- 10 Burchett J, Shankar M, Hamza A B, Guenther B D, Pitsianis N, Brady D J. Lightweight biometric detection system for human classification using pyroelectric infrared detectors. *Applied Optics*, 2006, **45**(13): 3031–3037
- 11 Candès E J, Wakin M B, Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 2008, **25**(2): 21–30
- 12 Duarte D F, Eldar Y C. Structured compressed sensing: from theory to applications. *IEEE Transactions on Signal Processing*, 2011, **59**(9): 4053–4085
- 13 Brady D J, Pitsianis N P, Sun X B. Reference structure tomography. *Journal of the Optical Society of America A*, 2004, **21**(7): 1140–1147
- 14 Peng M, Xiao Y. A survey of reference structure for sensor systems. *IEEE Communications Surveys & Tutorials*, 2012, **14**(3): 897–910
- 15 Wimalajeewa T, Chen H, Varshney P K. Performance limits of compressive sensing-based signal classification. *IEEE Transactions on Signal Processing*, 2012, **60**(6): 2758–2770
- 16 Davenport M A, Boufounos P T, Wakin M B, Baraniuk R G. Signal processing with compressive measurements. *IEEE Journal of Selected Topics in Signal Processing*, 2010, **4**(2): 445–460
- 17 Luo X M, Liu T, Liu J, Guo X M, Wang G L. Design and implementation of a distributed fall detection system based on wireless sensor networks. *EURASIP Journal on Wireless Communications and Networking*, 2012, **2012**(1): 118
- 18 Sun Q Q, Hu F, Hao Q. Mobile target scenario recognition via low-cost pyroelectric sensing system: Toward a context-enhanced accurate identification. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2014, **44**(3): 375–384
- 19 Guan Q J, Li C Y, Guo X M, Wang G L. Compressive classification of human motion using pyroelectric infrared sensors. *Pattern Recognition Letters*, 2014, **49**: 231–237
- 20 Kay S M. *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998.
- 21 Gales M J F. Maximum likelihood linear transformations for HMM-based speech recognition. *Computer Speech & Language*, 1998, **12**(2): 75–98

22 Babu R V, Anantharaman B, Ramakrishnan K R, Srinivasan S H. Compressed domain action classification using HMM. *Pattern Recognition Letters*, 2002, **23**(10): 1203–1213

23 Li H, Greenspan M. Model-based segmentation and recognition of dynamic gestures in continuous video streams. *Pattern Recognition*, 2011, **44**(8): 1614–1628

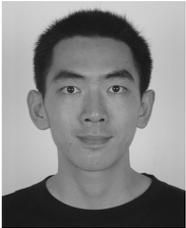


**关秋菊** 仲恺农业工程学院机电工程学院讲师. 主要研究方向为基于周边智能环境的人体运动行为识别.

E-mail: qiuju95@126.com

(**GUAN Qiu-Ju** Lecturer at the College of Mechanical and Electrical Engineering, Zhongkai University of

Agriculture and Engineering. Her main research interest is physical activity recognition based on ambient intelligence environment.)



**罗晓牧** 广州中医药大学医学信息工程学院讲师. 2012 年获得中山大学通信与信息系统专业博士学位. 主要研究方向为无线传感器网络, 人体动作识别和机器学习.

E-mail: woodwood2000@163.com

(**LUO Xiao-Mu** Lecturer at the School of Medical Information Engineering, Guangzhou

University of Chinese Medicine. He received his Ph.D. degree from Sun Yat-Sen University in 2012. His research interest covers wireless sensor networks, human activity recognition and machine learning.)



**郭雪梅** 中山大学数据科学与计算机学院副教授. 主要研究方向为嵌入式系统与移动计算, 信息获取与信息处理.

E-mail: guoxuem@mail.sysu.edu.cn

(**GUO Xue-Mei** Associate professor at the School of Data and Computer Science, Sun Yat-Sen University. Her

research interest covers embedded system, mobile computing, information acquisition and processing.)



**王国利** 中山大学数据科学与计算机学院教授. 主要研究方向是周边智能及周边智能感知辅助实现的人机协作. 本文通信作者.

E-mail: isswgl@mail.sysu.edu.cn

(**WANG Guo-Li** Professor at the School of Data and Computer Science,

Sun Yat-Sen University. His research interest covers ambient intelligence and ambient assisted human-robot collaboration. Corresponding author of this paper.)