

面向可穿戴设备的超声波手势识别方法

杨晓东^{1,2,3} 陈益强^{1,2} 于汉超^{1,2,3} 刘军发^{1,2} 李展歌^{1,2,4}

(中国科学院计算技术研究所 北京 100190)¹ (移动计算与新型终端北京市重点实验室 北京 100190)²
(中国科学院大学 北京 100049)³ (天津大学软件学院 天津 300072)⁴

摘要 针对传统的基于触摸屏和计算机视觉的手势交互方法无法应对智能可穿戴设备逐渐趋向小型化和低功耗等问题,依据声波的 Doppler 效应和运动物体能够改变声波频率的特性,提出了一种基于超声波的低功耗鲁棒手势识别方法。该方法利用 Goertzel 算法进行超声波频率特征提取和分析,得到手势的移动方向并形成时间序列,进而利用隐马尔科夫模型实现用户手势识别。在微软移动终端 Surface 上的实验表明,此方法在安静和嘈杂的环境下都能达到较高的手势识别准确率、更高的鲁棒性、更低的计算成本和功耗,能够满足可穿戴设备发展对于手势识别的需求。

关键词 可穿戴设备,手势识别,超声波,Goertzel 算法,隐马尔科夫模型,Doppler 效应

中图分类号 TP391.4 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.10.005

Ultrasonic Waves Based Gesture Recognition Method for Wearable Equipment

YANG Xiao-dong^{1,2,3} CHEN Yi-qiang^{1,2} YU Han-chao^{1,2,3} LIU Jun-fa^{1,2} LI Zhan-ge^{1,2,4}

(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)¹

(Beijing Key Laboratory of Mobile Computing and Pervasive Device, Beijing 100190, China)²

(University of Chinese Academy of Sciences, Beijing 100049, China)³

(School of Computer Software, Tianjin University, Tianjin 300072, China)⁴

Abstract Wearable equipment has several limitations such as the smaller shape, limited power and CPU performance, which the traditional methods of human-computer interaction based on the touch screen and the computer vision cannot deal with. Based on the Doppler effect of sound waves, we proposed a low-power robust method. The method depends on Goertzel algorithm to extract the features of soundwave's frequency-shifted, so that the moving direction of the user's hand can be got, and furthermore depends on the HMMs to classify the hand gestures. Using the proposed method, we conducted a series of comparative experiments on Surface Pro, one of the Microsoft mobile terminals. The experiment results show that no matter in the quiet environment or in the noisy one, the proposed method both has quite high precision rate, lower computational complexity which can lead to lower power consuming and better robustness. So the proposed method can meet the needs of wearable equipment development for gesture recognition.

Keywords Wearable equipment, Gesture recognition, Ultrasonic waves, Goertzel algorithm, Hidden Markov model, Doppler effect

1 引言

随着 Android、IOS 和 Windows 8 等移动操作系统的兴起,搭载移动操作系统的智能可穿戴设备发展势头迅猛,Google Glass、苹果 iWatch 智能手表以及 Jawbone UP 智能手环等设备的推出更是将可穿戴式设备的发展带入了新的时代。然而可穿戴设备在人机交互(Human-Computer Interaction)方面还存在诸多限制,例如设备屏幕偏小(如智能腕表)、不配备可触摸的交互屏幕(如 Google Glass)等,传统的按键式和触摸式手势操控方法已经不能胜任,可穿戴设备平

台迫切需要更为自然灵活的非接触式手势操控技术。

综合考虑到可穿戴设备小型化、性能和电量有限等特点,本文提出了一种基于超声波的低功耗鲁棒手势识别方法,利用设备自带的麦克风和扬声器,发射超声波并进行接收,对用户手势变化引起的声波信号频率变化的特征进行提取和分析,进而识别用户手势。人耳听觉可感知的频率范围在 20Hz 到 20kHz 之间^[1],一般情况下成年人听不到频率超过 18kHz 的声音(忽略超声波对动物和儿童的影响),因此,为了不影响人们的日常生活和设备的正常使用,我们令设备自带的扬声器发出 20kHz 的超声波信号(扬声器一般最高可以发出

到稿日期:2014-07-21 返修日期:2015-01-09 本文受国家国际科技合作专项项目(2014DFG12750),国家自然科学基金(61070110, 61173066)资助。

杨晓东(1991—),男,硕士生,主要研究领域为人机交互、普适计算,E-mail: yangxiaodong@ict.ac.cn;陈益强(1973—),男,博士,研究员,博士生导师,CCF 高级会员,主要研究领域为人机交互、普适计算;于汉超(1985—),男,博士生,CCF 学生会会员,主要研究领域为人机交互、普适计算;刘军发(1974—),男,博士,副研究员,CCF 高级会员,主要研究领域为智能感知交互、无线定位、机器学习应用等;李展歌(1991—),男,硕士生,主要研究领域为人机交互、网络安全。

18kHz~22kHz的声波),同时令麦克风进行声波信号采集。由于声波具有的 Doppler 效应,用户手在麦克风感应范围内运动时会引起声波的频率变化,使用 Goertzel 算法对采集的声波信号频率变化进行特征提取和分析,进而得到用户手势的移动方向,最后将得到的手势运动方向作为观察值输入隐马尔可夫模型(Hidden Markov Model, HMM)中进行手势识别。与传统的基于计算机视觉的非接触式手势交互方式相比,本文方法不易受环境光照和噪声的影响,降低了计算的复杂度并有效减少了设备耗电量,更适用于可穿戴设备。

2 相关工作

近年来出现了许多新颖的非接触式人机交互技术,其中大多数都基于计算机视觉,例如,利用微软的 Kinect 通过感知黑白光谱的方式还原 3D 模型来感知用户行为^[2];通过 RGB-Depth 摄像机捕获彩色图和深度图来进行相关的手势识别工作^[3]。但是基于计算机视觉的手势交互方法一般对光线比较敏感,算法复杂度较高、需要较多的系统资源,并且可穿戴设备的处理性能和电量都十分有限,大部分穿戴式设备不配备专业的摄像头或深度摄像头,故基于计算机视觉的非接触式手势交互技术大都不适用于可穿戴设备。

与基于计算机视觉的非接触式手势交互技术相比,基于超声波的手势识别技术则更适用于可穿戴设备。利用超声波进行人机交互由来已久,早在 1997 年,Joseph Paradiso 等人^[4]制作了利用雷达声波感知用户行为的沉浸式地毯,结合其他传感器在地毯交互空间内感知用户的各种动作并实时地转换音乐;美国西北大学的 Stephen P. Tarzia 等人^[5]通过分析用户身体反射的超声波的相位和角度变化,判断用户是否在设备前来进行人机交互。

近年来,在利用超声波进行手势识别方面,也有很多研究人员进行了相关的研究工作,例如,Georgia Tech^[6]的 Kaus-tubh KalgaoKar 模型使用分布在左、中、右 3 个位置的接收器来接收超声波信号,综合多个位置的信号变化特征创造出 3D 环境来判断用户手势,该方法使用频率为 40kHz 的超声波,采样频率高达 96kHz,普通的扬声器已经不能满足此要求,同时使用多个声波接收器也不适用于体积较小、形状多样的可穿戴设备;Microsoft 研究院的 Sidhand Gupta 等人^[7]针对 PC 平台设计了 SoundWave 系统,利用快速傅里叶变换(Fast Fourier Transformation, FFT)算法处理声波信号进而判断用户手势,取得了良好的效果,但用于处理声波信号的 FFT 算法要求一次处理的信号数量为 2 的整数次方,而在移动设备中不同的硬件平台和系统版本录音时所需要的最小缓冲区大小不同,此方法在可移植性和灵活性方面存在缺陷。

3 基于超声波的低功耗鲁棒手势识别方法

利用超声波进行手势识别的关键问题在于如何从声波信号中提取特征以及如何根据特征进行手势识别,本文采用对于处理样本点数更加灵活的 Goertzel 算法取代 FFT 算法进行时域-频域转换,然后根据频谱中 20kHz 邻域内谱峰的宽度和边界的变化进行特征提取,进一步分析得出用户手势移动的方向特性,并将其作为观测值输入 HMM 模型中,以实现用户手势的实时识别。图 1 为基于超声波的手势识别流程图。

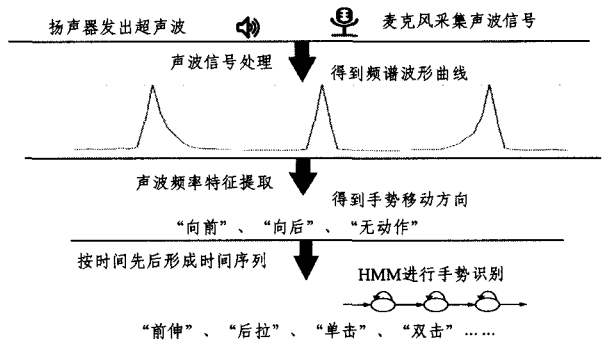


图 1 基于超声波的手势识别流程

3.1 声波特征提取与分析

声波的 Doppler 效应^[8]体现在其频率的变化上,频率的变化与运动物体的速度规律关系如式(1)所示。

$$f' = f \cdot \left(\frac{v \pm v_0}{v} \right) \quad (1)$$

其中, f' 为观察到的频率; f 为发射的原始频率(扬声器发出的声波为 20kHz); v 为声波在空气中传播的速度; v_0 为物体与接收端的相对移动速度。

简言之,当用户手朝向麦克风运动时,其反射的声波频率增大;远离麦克风运动时,其反射的声波频率减小。由经验测试知,人手在设备前移动的最大速度为 4m/s,声波在空气中的传播速度为 340m/s,根据式(1)可以得出声波频率的变化范围 $r \leq 300\text{Hz}$,因此,提取声波频率变化特征只需计算以频率 $f = 20\text{kHz}$ 为中心、 $r = 300\text{Hz}$ 为半径的邻域内的频率。

本文使用 Goertzel 算法^[9]进行信号预处理,将采集到的时域声波信号转变为频域信号。Goertzel 算法识别的频率分辨率是 f_s/N ,其中, f_s 是采样频率, N 是一次处理的采样点数。Goertzel 算法可以直接求取频率为 $k \cdot (f_s/N)$, $k = 1, 2, \dots, N-1$ 的频谱值,频谱值越大,表示该频率声波的能量越大。由于扬声器发出的高频声波信号集中在 20kHz,其周围高频区域的能量变化可以认为是由用户手移动所引起的。本文针对提取的频率变化特征,提出了两种有效的判断方法:判断谱峰边界的变化和是否生成第二谱峰。

用户手在设备前的运动会反射扬声器发出的 20kHz 的超声波,故 20kHz 处的频谱值是整个高频区域的峰值,本文通过计算周围频段的频谱值与峰值的比例大小关系来确定频域信号中谱峰的边界。

利用 Goertzel 算法计算峰值及其两边各频段的频谱值,如式(2)~式(4),其中, $Energy$ 表示频率为 20kHz 处的频谱值,函数 $Goertzel(f)$ 表示利用 Goertzel 算法进行频谱值的计算, $before$ 和 $after$ 表示两侧各频段的频谱值。

$$Energy = Goertzel(f) \quad (2)$$

$$before[k] = Goertzel\left(f - k \cdot \frac{f_s}{N}\right) \quad (3)$$

$$after[k] = Goertzel\left(f + k \cdot \frac{f_s}{N}\right) \quad (4)$$

其中, $k = 1, 2, \dots, \left\lfloor \frac{r \cdot N}{f_s} \right\rfloor$, $f = 20\text{kHz}$, $r = 300\text{Hz}$ 。

由于通过式(2)、式(3)计算得到的频谱值是离散的,因此,用频段的数量来度量谱峰的宽度,而不是具体的频率值范围。从频谱峰值频段向两边扫描,找到谱峰的前边界 $preEdge$ 和后边界 $tailEdge$ 。算法 1 为谱峰边界查找算法,其中 a 表示设定的边界值与峰值的比例系数(一般为 10%)。

算法 1 谱峰边界查找算法

输入: Energy, before[n], after[n], a

输出: preEdge, tailEdge

preEdge ← 1

tailEdge ← 1

WHILE before[preEdge-1] > a * Energy

DO preEdge ← preEdge + 1

WHILE after[tailEdge-1] > a * Energy

DO tailEdge ← tailEdge + 1

由于信号的频谱形状呈局部不规则的锯齿状,针对手势快速运动时的频率变化特性表现为产生第二谱峰的问题,传统的谱峰搜索算法复杂度高^[10],会影响识别的实时性,因此本文采用设置门限值的方法,从谱峰边界处向两侧继续扫描频谱值数组,若存在高于门限值的频谱值,则说明存在第二谱峰,其判断的准则如式(5)、式(6)所示,其中, b 为设定的第二谱峰值与峰值的门限比例系数。

假设事件 L :“低频方向存在第二谱峰”,事件 H :“高频方向存在第二谱峰”。

$$(\exists e \in \text{before}[\text{preEdge} \dots n], e > b \cdot \text{Energy}) \Rightarrow L \quad (5)$$

$$(\exists e \in \text{after}[\text{tailEdge} \dots n], e > b \cdot \text{Energy}) \Rightarrow H \quad (6)$$

从声波信号中提取出谱峰边界和第二谱峰的两项特征之后,进一步根据式(7)分析出手势的移动方向,其中, $Direction$ 表示手势移动方向,1 表示向前,2 表示向后,0 表示无动作, $BasicEdge$ 表示无动作时谱峰的边界(谱峰宽度一般约为 100Hz)。

$$Direction = \begin{cases} 1, & \text{if } L \wedge (\text{preEdge} > \text{BasicEdge}) \\ 2, & \text{if } H \wedge (\text{tailEdge} > \text{BasicEdge}) \\ 0, & \text{else} \end{cases} \quad (7)$$

3.2 基于隐马尔可夫模型的手势识别方法

按一定的时间间隔识别出手势移动的方向并形成时间序列后,进一步基于 HMM 识别出用户手势。基于 HMM 进行手势识别^[11],大多通过计算某测试样本与参考模型集的似然度,判断最大值以确定测试样本的类别^[12],如式(8)所示,其中 λ 表示 HMM 模型, O 表示观察值序列。

$$\hat{\lambda} = \arg \max P(O|\lambda) \quad (8)$$

使用 HMM 模型进行手势识别需要确定手势开始和结束边界,本文设定当识别的移动方向不为“无动作”时,手势开始,当连续识别到 3 个“无动作”时,手势结束。记录之间的手势移动方向的时间序列和序列长度,若长度小于 6,则视为噪声抛弃,不予识别^[13]。

在获取手势的特征观察序列 O 后,可以利用提前训练好的 HMM 模型进行手势识别,本文所实现的利用 HMM 模型进行手势识别的流程如图 2 所示。

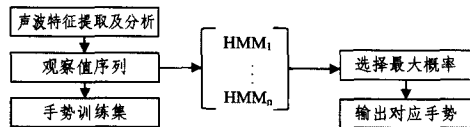


图 2 基于 HMM 的手势识别方法流程

为了计算简便,本文选择离散的 HMM 模型,观察值是提取声波特征后并分析得到的手势移动方向,接下来为每一个手势分别建立一个 HMM 模型。本文选择的 HMM 模型类型是 Bakis 模型(左-右模型),因为这种模型参数较少,拓扑结构相对简单,更重要的是它特别适合为时变信号建模^[14]。

对每个手势建立的 HMM 模型进行训练,首先初始化每

个手势模型的参数 $\lambda_i = (A, B, \pi)$,其中 A 表示状态转移矩阵, B 表示混淆矩阵, π 表示初始状态概率。每个手势执行过程中将通过声波特征提取和分析得到的观察值序列作为训练样本,利用 Baum-Welch 算法进行重新计算,调整参数直至模型收敛,得到最优的模型参数,确定手势模型。

在将 HMM 模型进行训练确定后,把通过声波特征提取和分析得到的观察值序列 O 作为输入,应用 Viterbi 算法计算观察序列 O 在每个模型下的 $P_i(O|\lambda_i)$,选择最大似然的手势模型输出。

4 实验与分析

4.1 实验描述

为了验证提出的基于超声波的低功耗鲁棒手势识别方法适用于可穿戴设备,本文选择了搭载 Windows 8 操作系统的微软 Surface Pro 移动终端作为测试平台,设计了多组对比实验,以验证本文提出的方法在保证识别准确率的情况下能有效地降低计算的时间和空间复杂度,进一步减少电量消耗。

本文使用频率为 20kHz 的声波信号,根据奈奎斯特采样定律^[15],采用 $f_s = 44.1 \text{ kHz}$ 的采样频率。采样点数 N 的选取决定了频率分辨率和录音时间窗口大小,经过多人实验测试得到手移动的最大速度约为 3.9m/s,手势移动的距离约为 20cm,因此手势实施的最小时间间隔约为 50ms,同时为了与 FFT 算法提取特征进行计算复杂度对比实验,本文选取采样点数 $N = 2048$,可得频率分辨率 $f_s/N = 21.5 \text{ Hz}$,采样 N 个点的时间窗口大小为 $N/f_s \approx 46.4 \text{ ms}$ 。经过多次实验观测得到,在频率变化特征提取过程中,边界搜索算法的边界比例系数约为 10%,第二谱峰搜索算法的门限比例系数约为 30%。

在 HMM 模型中,本文定义观察状态为 $\{f, b, \epsilon\}$,其中,“ f ”表示“向前”,“ b ”表示“向后”,“ ϵ ”表示“无动作”。隐藏状态为 $\{F, B, N, SF, SB\}$,其中,“ F ”表示“前进”,“ B ”表示“后退”,“ N ”表示“无动作”,“ SF ”表示“第二次前进”,“ SB ”表示“第二次后退”。对每个手势都建立一个 HMM 模型,其状态转移图如表 1 所列。

表 1 手势的 HMM 状态转移图

手势 HMM 状态转移图	识别手势描述
	前伸
	后拉
	单击
	双击

根据手势移动特性的时间序列进行手势识别的方法还有基于规则分析的手势识别方法。其原理是利用相同的手势移动方向出现的次数来度量其手势持续的时间并根据手势移动的方向和时间定义一系列手势识别规则产生式,以实现手势识别。

本文为将基于 HMM 的手势识别方法与基于规则分析的手势识别方法进行对比实验,定义了基于移动方向时间序列的手势识别规则集,如表 2 所列。为了方便描述,以下用“ f ”表示“向前”移动方向,“ b ”表示“向后”的移动方向,“ ϵ ”表示“无动作”。

表2 手势识别规则集

手势动作序列	识别手势描述
$f^p * \epsilon^m \epsilon^*$	前伸
$b^q * \epsilon^m \epsilon^*$	后拉
$f^p * b^q b^* * \epsilon^m \epsilon^*$	单击
$f^p f^* b^q b^* * f^p f^* b^q b^* * \epsilon^m \epsilon^*$	双击

其中 t, p, q, m 表示触发手势识别时相同的移动方向连续出现次数的阈值,通过阈值的设定来表示不同手势被识别所需对应同一移动方向保持时间的差异。具体识别过程描述如下:以识别单击手势的时间动作序列为例,当已识别的动作序列中连续出现至少 p 次的向前移动方向 f ,继而连续出现至少 q 次的向后移动方向 b ,以及连续出现至少 m 次无动作 ϵ 时,可识别出用户做了单击动作的手势,同理可完成其他用户手势的识别。

为了全面地比较本文提出的基于超声波的低功耗鲁棒手势识别方法与其他方法的性能,本文设计了以下几组对比实验,如表3所列。

表3 对比实验列表

实验序号	实验系统描述
1	Goertzel 算法提取特征+HMM 识别
2	Goertzel 算法提取特征+规则识别
3	FFT 算法提取特征+HMM 识别
4	FFT 算法提取特征+规则识别

注:以下实验结果描述使用序号标注。

4.2 算法的运行效率分析

从理论上进行复杂度分析,FFT 算法计算 N 个点需要有 $N \log_2 N$ 次复数乘法和 $0.5N \log_2 N$ 次复数加法。而 Goertzel 算法每次只计算一个频点的傅里叶分量,每计算一个频段的频谱值只需要 $N+1$ 次实数乘法和 $2N+1$ 次加法运算,尤其是需要计算的点数不超过 $2 \log_2 N$ 时,Goertzel 算法将比 FFT 算法效率更高^[14],而实验中仅需要计算 30 个频段的频谱值即可。并且,FFT 算法在计算过程中使用蝶形运算,每次计算需要递归 $\log_2 N$ 次,相比之下,Goertzel 算法无需递归,其空间复杂度更低。因此,使用 Goertzel 算法进行特征提取具有更低的系统消耗和更高的运行效率。

使用 HMM 模型进行手势识别,在模型训练完成后,解码过程的时间复杂度为 $O(N^2 T)$ ^[17],其中, N 为隐藏状态的数目(本文中 $N=5$), T 为观察序列的长度,因此与基于规则的识别方法一样,计算复杂度都为 $O(T)$,实际进行的实验也证明了理论上的分析。

在 Windows 8 平台上,以系统运行时的 CPU 占用率和内存占用率为指标来度量系统的时间复杂度和空间复杂度。关闭其他可能引起误差的后台程序,分别运行相应的系统各 10 分钟,在相同的时间点上实施相同的手势,同时每 20s 记录一次 CPU 使用率和内存占用量,最后取平均值,对比实验的结果如表 4 所列。

表4 CPU使用率和内存占用量的对比实验结果

实验序号	CPU 使用率	内存占用量(MB)
1	4.0%	19.2
2	4.4%	17.8
3	7.4%	24.9
4	7.9%	23

从实验 1 和实验 2、实验 3 和实验 4 的对比中可以看出,与基于规则的手势识别方法相比,使用 HMM 进行手势识别

几乎没有增加对系统的资源消耗;然而,实验 1 和实验 3、实验 2 和实验 4 的对比结果表明,使用 Goertzel 算法进行特征提取相比于使用 FFT 算法,识别系统的性能有显著提升,CPU 使用率降低了约 45%,内存占用量降低了约 23%。另外通过实验观察到,当有手势实施时,CPU 使用率平均增加约 20%,这是因为识别系统只有识别出有手势实施时才进行手势识别,否则只进行移动方向的判断,此机制可以进一步减少在无手势时的资源消耗。

在可穿戴设备中,电量和 CPU 处理能力一直是影响其发展的重要瓶颈之一,而在现有硬件的基础上进行改进的一个重要途径就是优化算法,降低其计算复杂度。因此,基于超声波的低功耗鲁棒手势识别方法能够很好地适用于可穿戴设备平台。

4.3 手势识别的准确性和鲁棒性分析

基于超声波进行手势识别,无论是使用 FFT 算法还是 Goertzel 算法进行声波频率特征提取,都只使用频率为 20kHz 左右的声波信号,对环境其他频率段可见噪声不敏感,具有良好的抗噪性和鲁棒性。

实验中邀请 5 位志愿者(3 男 2 女),每位志愿者交叉完成“前伸”、“后拉”、“单击”、“双击”4 种手势各 50 次,相邻手势之间的时间间隔不小于 1s。这样,共收集了 1000 个手势样本(每个手势各 250 个),利用 10-折交叉验证的方法,对 HMM 模型进行训练并测试,其结果如表 5 所列。

表5 HMM模型训练及测试结果

手势类别	训练误差	测试误差
前伸	0.00%	0.00%
后拉	0.00%	0.00%
单击	0.89%	2.00%
双击	0.22%	1.00%
全部手势	0.25%	0.75%

从表 5 中可以看出,经过训练的 HMM 模型在手势识别方面表现十分良好,具有很高的识别准确率,其中“单击”的识别准确率相对较低,经过实验观察,其错误识别为“双击”主要是因为手势移动方向识别的过程中存在一定误差,尤其是将“向前”错误识别为“向后”或将“向后”错误识别为“向前”两种情况,会造成将“单击”错误识别为“双击”。但在嘈杂环境下,移动方向的识别准确率约为 95%,所以此问题在一定程度上可以忽略。

为了对比基于规则模型和基于 HMM 模型的手势识别方法的准确率,本文交叉完成“前伸”、“后拉”、“单击”和“双击”4 种手势各 20 次,记录手势完成次序并采集手势移动方向的时间序列,再输入规则模型和 HMM 模型中,统计各自输出的手势,结果如图 3 所示。

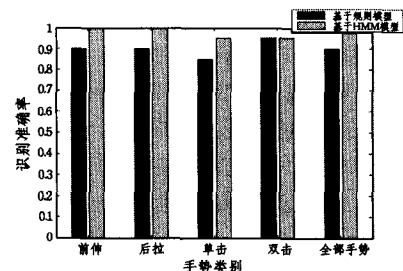


图3 手势识别模型准确率对比

从实验结果可以看出,无论是基于规则模型还是 HMM

模型的手势识别方法都具有很高的识别准确率,但基于HMM模型的手势识别方法总体准确率优于基于规则模型的方法。通过对手势移动方向的观察和分析,可以发现基于规则模型的方法要求相同移动方向连续出现一定次数才进行响应,由于手势完成过程中移动的速度是变化的,手势移动的方向序列有时是间断的,例如一次“前伸”手势的方向时间序列为“fεffεfεffεffεε”,手势“前伸”识别的阈值为4,则基于规则的手势识别模型无法正确识别;更有甚者,在“前伸”的时间序列中会存在错误识别的“向后”的方向,基于规则的手势识别模型无法对此类情况进行识别,而基于HMM模型的手势识别模型之前经过大量的样本训练,就可以弥补在特征提取和分析阶段存在的误差,有效提高了手势识别的准确率。相比之下,基于HMM模型的手势识别方法具有更好的鲁棒性。

为了验证本文提出的方法具有良好的抗噪性,我们分别在安静环境下(30±5dB)和喧闹的环境(70±5dB)下对表3中所列的4组系统进行准确率实验,同样分别交叉完成4种手势各20次,记录手势完成的次序和系统输出的手势类别,实验结果如图4所示。

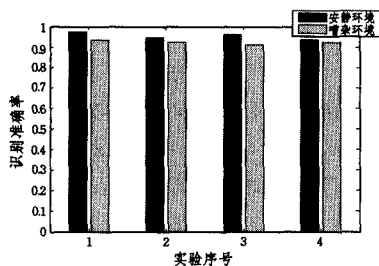


图4 低功耗鲁棒手势识别方法的抗噪性测试结果

从图4的结果可以看出,基于超声波手势识别的4种识别系统都具有较好的识别准确率,通过实验观察发现,在嘈杂环境下,手势移动方向的识别有一定错误率,大约为5%,影响了在嘈杂环境下的手势识别模型识别准确率。其原因可能是嘈杂环境中存在超声波信号的干扰。

除此之外,在实验过程中还加入了播放音乐和录制音频作为对照组,实验结果显示播放音乐和录制音频与超声波手势识别之间相互没有影响,因此本文提出的方法在实现手势识别的同时不会影响计算机中其他应用的正常使用。同时,实验发现,设备的音量和扬声器的位置对手势实施的位置有一定限制,进而对识别的准确率存在影响。当使用外接音响设备时,手势识别的范围有明显的扩大。

结束语 利用声波的Doppler效应,本文提出了一种基于超声波的低功耗鲁棒手势识别方法,使用Goertzel算法进行声波频率特征提取,仅需要计算20kHz周围少数频段的频谱值便可取得较好的识别准确度,达到增加计算效率和减小计算量的效果,从而降低了系统资源消耗并减少了设备耗电量,使用HMM模型进行手势识别能够进一步提高识别的准确性和鲁棒性。基于超声波的低功耗鲁棒手势识别方法具有对光线不敏感、抗噪能力强、计算复杂性低等优点。通过多组对比实验可以看出,本文提出的手势识别方法能够在保证较高识别准确率的同时有效降低时间和空间复杂度,运行效率高,适于可穿戴设备平台。

然而,本文提出的方法也存在待改进之处,例如使用设备自带的扬声器发出超声波限制了手势移动的范围,并且本文定义的手势集比较简单,有待进一步扩展。

- [1] Fausti S A, Erickson D A, Frey R H, et al. The effects of noise upon human hearing sensitivity from 8000 to 20000Hz[J]. The journal of the Acoustical society of America, 1981, 69(5): 1343-1349
- [2] Oikonomidis I, Kyriazis N, Argyros A A. Efficient model-based 3D tracking of hand articulations using Kinect[C]// BMVC. 2011: 1-11
- [3] Yao Yuan, Zhang Lin-jian, Qiao Wen-bao. Hand part labeling and gesture recognition from RGB-D data [J]. Journal of Computer-Aided Design & Computer Graphics, 2013, 25(12): 1818-1825
- [4] Kalgaonkar K, Raj B. One-handed gesture recognition using ultrasonic Doppler sonar[C]// IEEE International Conference on Acoustics, Speech and Signal Processing, 2009 (ICASSP 2009). IEEE, 2009: 1889-1892
- [5] Paradiso J, Ablner C, Hsiao K, et al. The magic carpet: physical sensing for immersive environments[C]// CHI '97 Extended Abstracts on Human Factors in Computing Systems. ACM, 1997: 277-278
- [6] Tarzia S P, Dick R P, Dinda P A, et al. Sonar-based measurement of user presence and attention[C]// Proceedings of the 11th International Conference on Ubiquitous Computing. ACM, 2009: 89-92
- [7] Gupta S, Morris D, Patel S, et al. Soundwave: using the doppler effect to sense gestures[C]// Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems. ACM, 2012: 1911-1914
- [8] Gill T P. The Doppler Effect: an introduction to the theory of the effect[M]. Logos Press, 1965
- [9] Wu Guo-feng, Ji Zhong-mei, Zhang Jing. Spectrum sensing method based on Goertzel algorithm[J]. Journal of Information Engineering University, 2010, 11(3): 272-275
- [10] Zeng Hao, Zhang Ying-hui, Feng Wen-jiang. New peak searching method in subspace spectrum[J]. Journal of Computer Applications, 2009, 29(9): 2546-2547
- [11] Bilal S, Akmeliawati R, Shafie A A, et al. Hidden Markov model for human to computer interaction: a study on human hand gesture recognition[J]. Artificial Intelligence Review, 2013, 40(4): 495-516
- [12] Freeman W T, Weissman C. Television control by hand gestures [C]// Proc. of Intl. Workshop on Automatic Face and Gesture Recognition. 1995: 179-183
- [13] Ramamoorthy A, Vaswani N, et al. Recognition of dynamic hand gestures[J]. Pattern Recognition, 36(9): 2069-2081
- [14] Rabiner L. A tutorial on hidden Markov models and selected applications in speech recognition[J]. Proceedings of the IEEE, 1989, 77(2): 257-286
- [15] Landau H J. Sampling, data transmission, and the Nyquist rate [J]. Proceedings of the IEEE, 1967, 55(10): 1701-1706
- [16] Wang Wei-bing, Zhu Qiu-ping, Xu Xin-yi, et al. An improved realization structure for Goertzel algorithm[J]. Journal of Wuhan University (Natural Science Edition), 2007, 53(3): 375-378
- [17] Johnson M T. Capacity and complexity of HMM duration modeling techniques[J]. IEEE Signal Processing Letters, 2005, 12(5): 407-410