

AccelWord: Energy Efficient Hotword Detection Through Accelerometer

JAN 19TH, 2016

论文下载: <http://spirit.cs.ucdavis.edu/pubs/conf/li-mobisys15.pdf>

简介

随着智能设备和可穿戴设备越来越流行，声音控制的应用也越来越流行。声音控制设备侦听关键词，例如，“Okey Google”或者“Hi Galaxy”，来区分用户的语音命令或者其他命令。然而，该类应用需要利用麦克风持续地侦听语音信号，这将大大地增加智能设备的功耗。

为了解决该问题，我们设计了AccelWord的系统。它利用加速度传感器来侦听用户的语音。我们发现用户说话时的语音在加速度传感器引起的变化，足够用来检测用户说的关键词。我们在Android设备上实现了Accelword应用。与基于麦克风的关键词检测应用程序，例如Google Now和Samsung S voice相比，该系统在静态环境中可以达到 98%的精确度，在动态环境汇总可得到 92%的精确度，而仅仅消耗一半的能量。

我们主要解决了以下两个挑战性的问题：（1）从加速度数据中提取关键词的特征；同时，特征的计算开销不能太高
（2）减少用户移动引起的干扰。

系统介绍

- 回顾之前应用使用的方法：基于麦克风进行关键词检测到方法。用实验数据表明，该方法非常耗能。

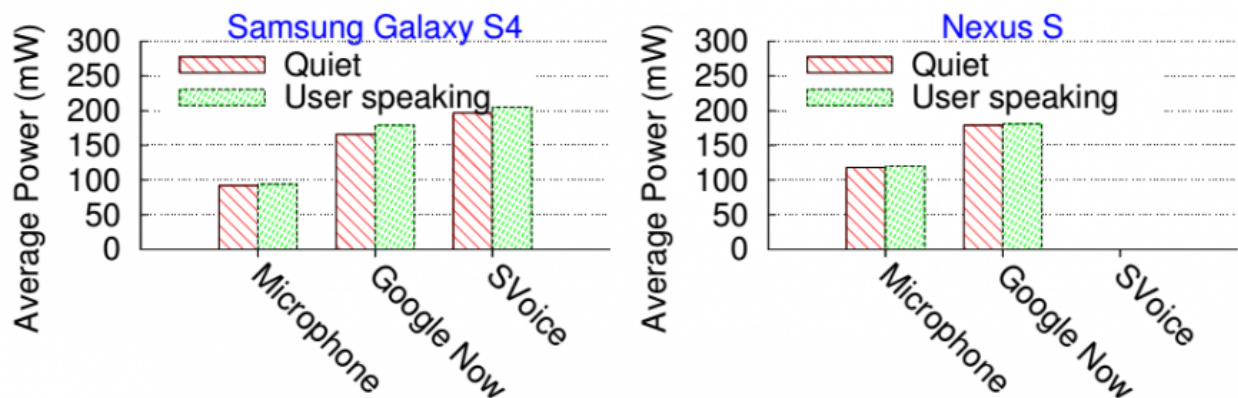


Figure 3: The Power Consumption of Current Hotwords Detection Apps.

- 设计目标：
 - 1 精度性：关键词检测方法
 - 2 稳健性：（1）用户移动的情况下（2）不同的语音频率下（3）噪声环境下
 - 3 节省能量：sensing方式和关键词检测方法

可行性分析

- 语音信号将会碰撞到加速仪，导致非常小的变化。（与陀螺仪原理类似）

实验测试

男性说话的语音频率为85HZ到155HZ, 女性说话的语音频率为165HZ到255HZ, Android 4.4OS要求加速仪的采用频率为200HZ以减少能量损耗。我们利用两部手机，分别演示说话

方和接收方。根据说话方手机的输出音量，我们在接收方产生不同的声音级别（SPL），收集加速度数据。

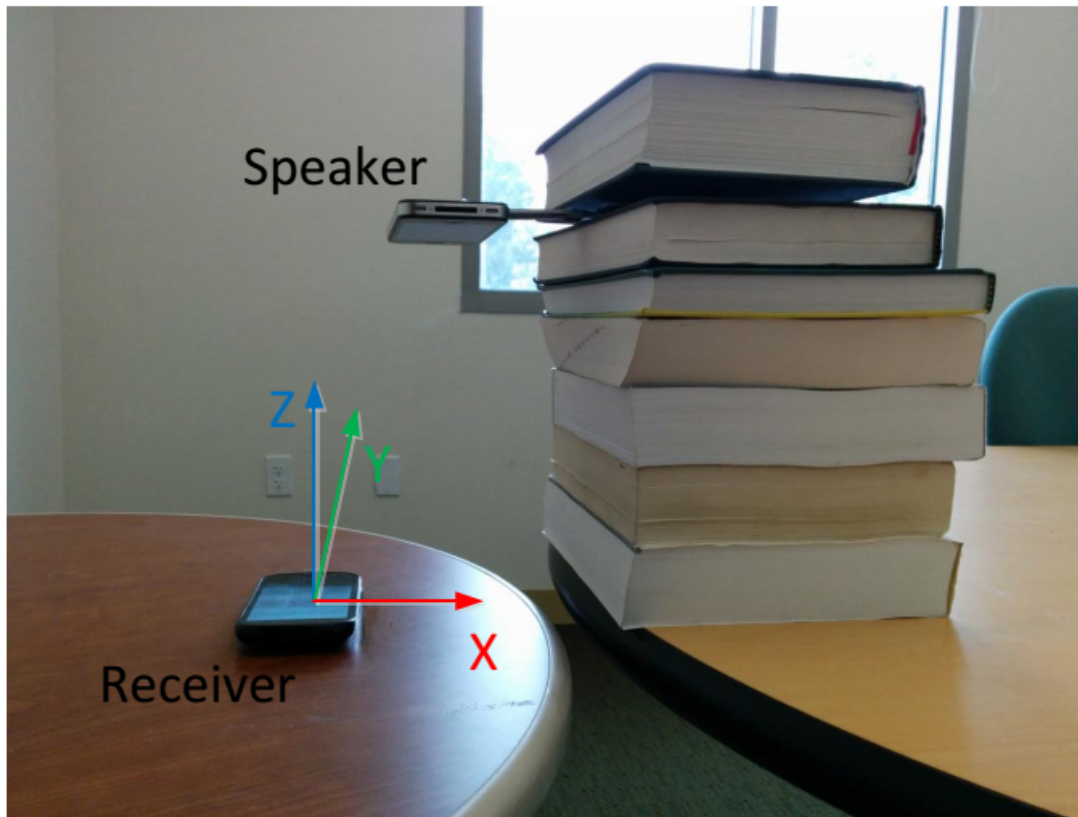


Figure 6: Experiment Setup

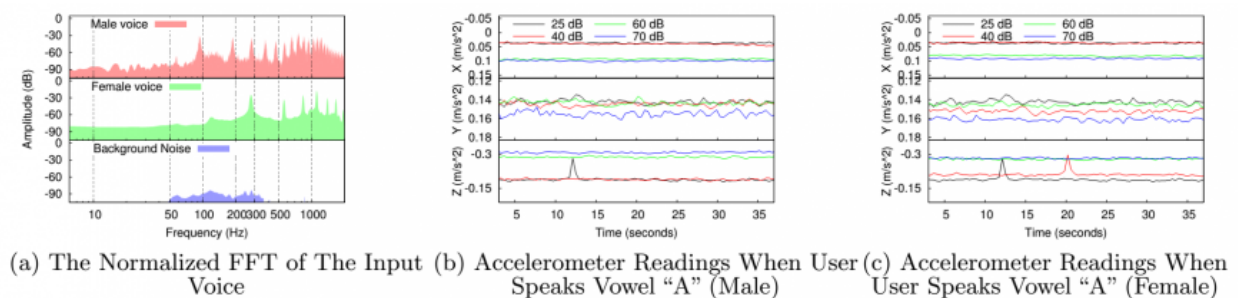
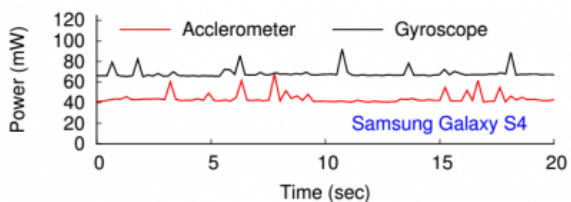
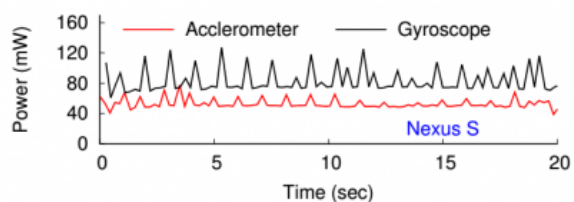


Figure 7: The Impact of Speaking Vowel "A" on Accelerometer

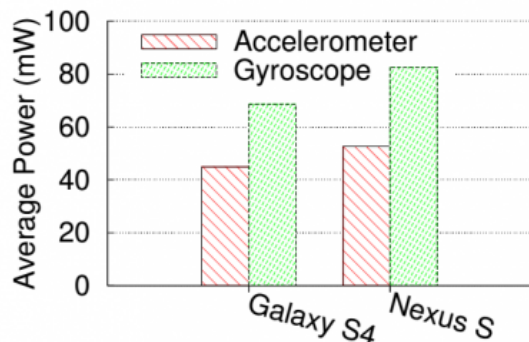
- 为什么选择加速仪，而不选择陀螺仪：
对加速仪，陀螺仪，以及麦克风的功耗进行统计，加速仪的功耗最低



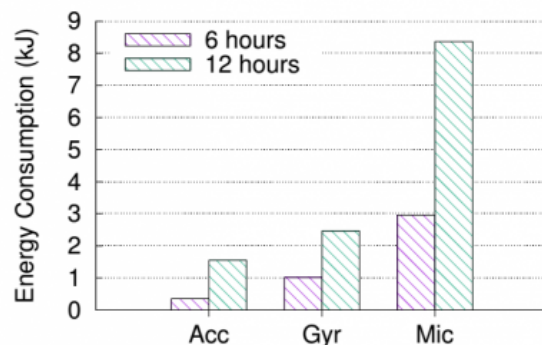
(a) Example: The Power Monitor Trace of Galaxy S4



(b) Example: The Power Monitor Trace of Nexus S



(c) The Average Energy Consumption of a 30 minutes Trace



(d) The Long Term Total Energy Consumption of Three Sensors

Figure 8: The Energy Consumption of Accelerometer, Gyroscope and Microphone

AccelWord关键词检测

- 提取关键词签名

基于信息增益的特征选取：

$$E(G) = -p_H \cdot \log_2 p_H - p_N \cdot \log_2 p_N$$

$$I(F) = E(G) - \sum_{f \in V(F)} \frac{|G_f|}{|G|} E(G_f)$$

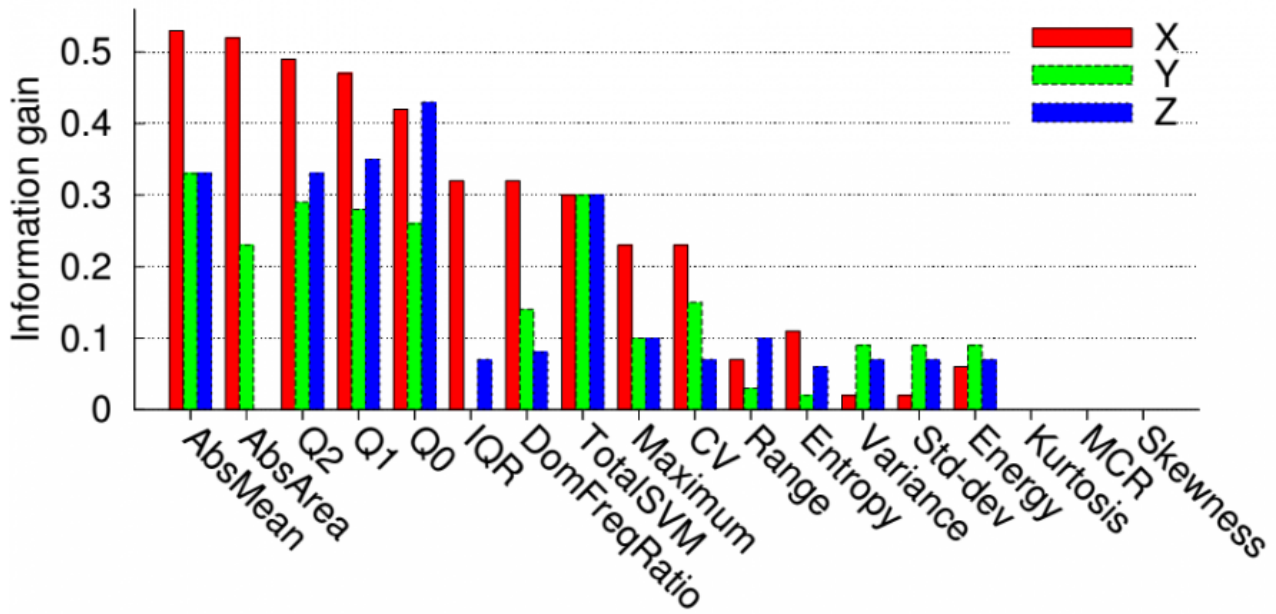


Figure 9: Information gain of candidate features

- 去除移动干扰

大多数人类的行为产生低频影响。因此，在计算特征之前采用高通滤波器，即可从加速度的信号中滤去移动的影响。

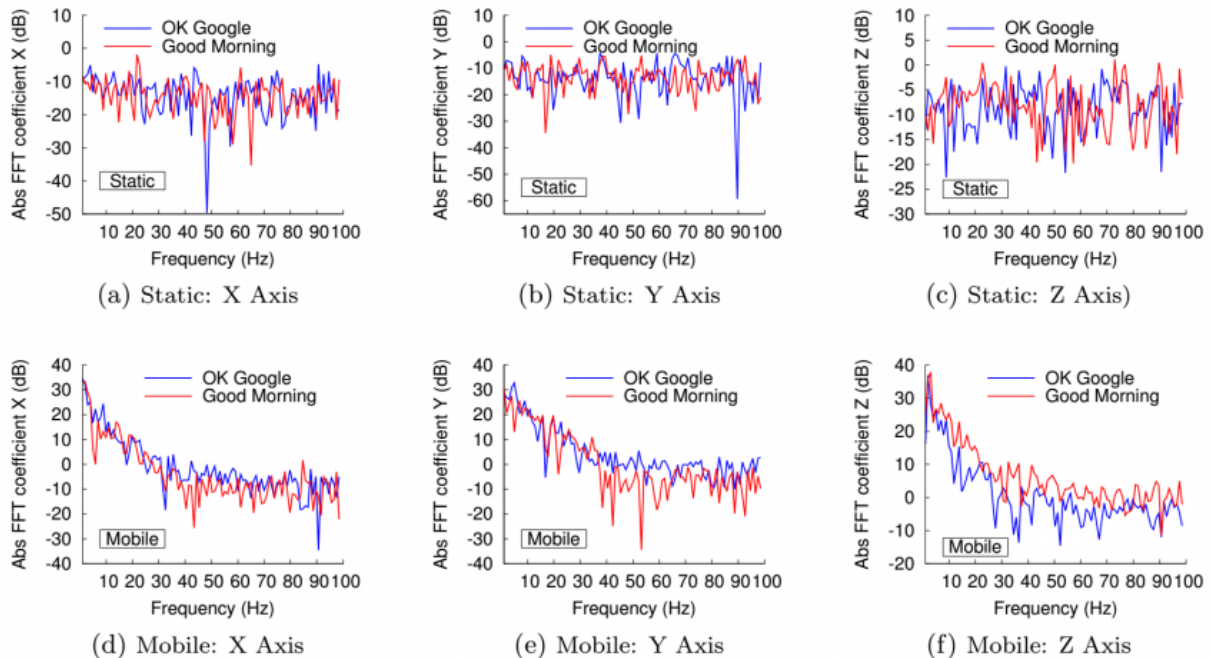
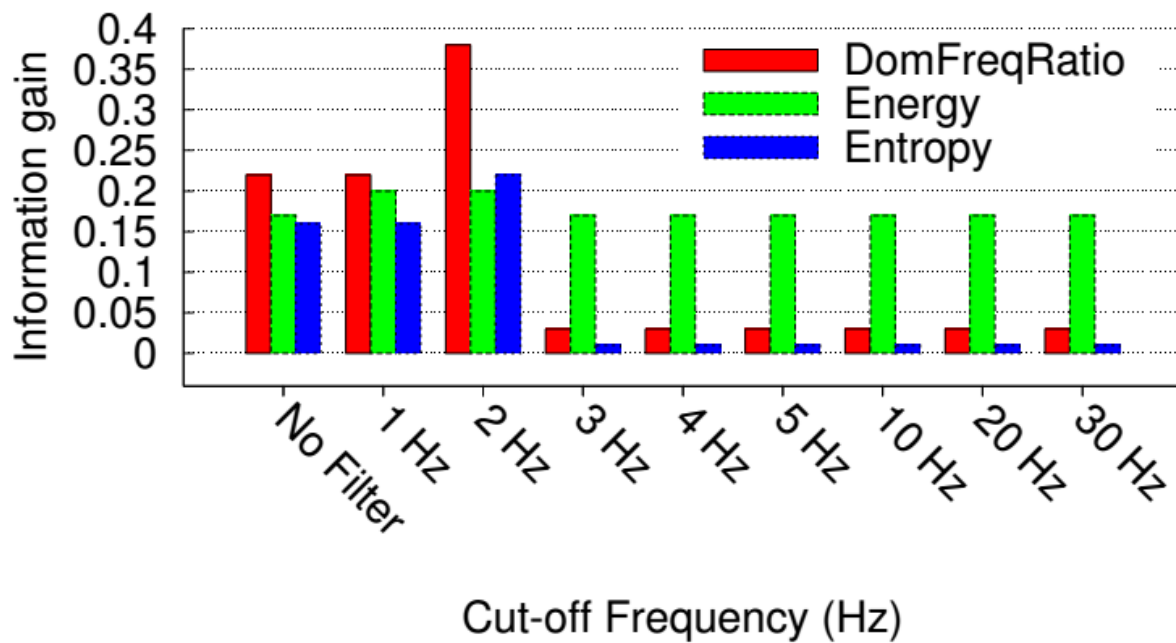


Figure 12: The FFT of the Static and Mobile Scenarios

通过实验观察，我们选择截止频率为2HZ的高通滤波器。



- 训练分类器

收集训练例子，计算特征，建立分类器

- 决定树分类器

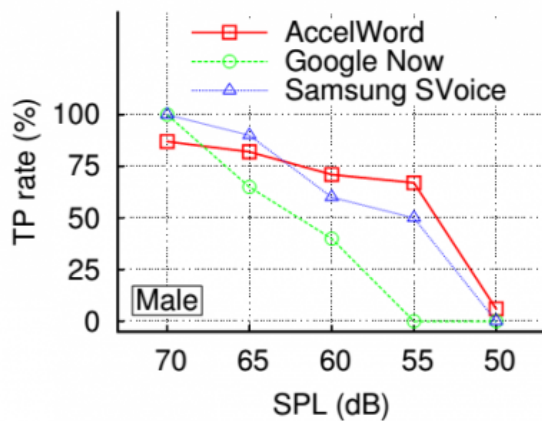
采用简单的基于滑窗的方式，对t秒的加速度信号进行实时分类。

实验

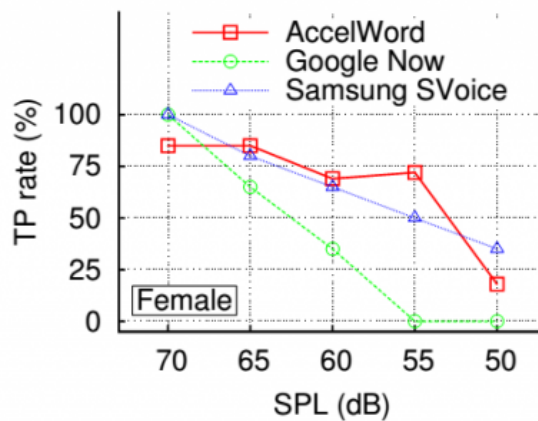
- 我们将AccleWord作为单独的app运行在Android上。为了有效计算特征，我们利用JAVA数据库定义的数据结构。
- 采用10个志愿者（5位男性，5位女性），关键词为“Okay Google”或者“Hi Galaxy”，说关键词10次，20次其他语句。
- 评估精确度：准确度（True Positive），误判率(False Positive) 以及能量

(1) 精确度

不同SPL



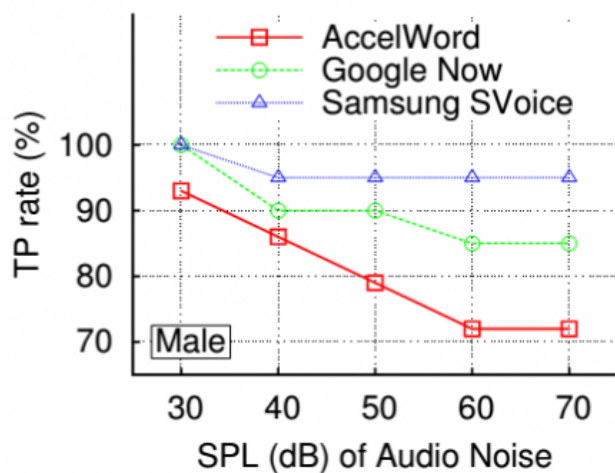
(a) Male



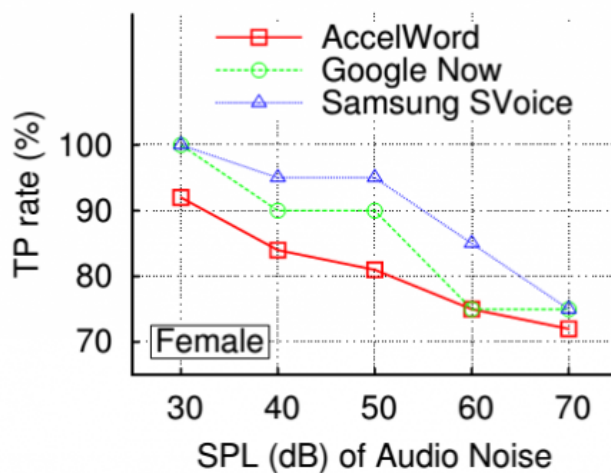
(b) Female

Figure 18: Accuracy Comparison with Google Now and Samsung S Voice

不同噪声背景



(a) Male



(b) Female

Figure 19: True Positive Rate vs. Audio Noise SPL

移动场景

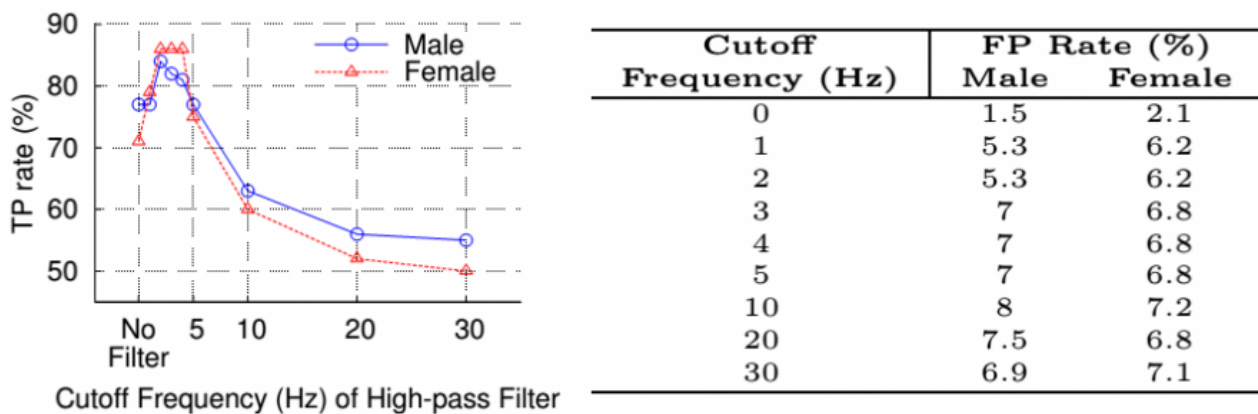


Figure 20: Change in TP and FP Rates for Different Cutoff Frequency

(2) 能量

	Energy Saving (%)	
	Galaxy S4	Nexus S
Google Now	46.19%	53.85%
Samsung S Voice	57.14%	N/A

Table 5: The percentage of energy saved

结语

这篇文章的优点在于可读性强，实验非常完善。可以看出，基于加速度进行关键词检测的方案在功耗上远远低于利用麦克风进行关键词检测的方法。然而，在噪声环境下，该方案的精确度会受到影响。利用麦克风进行关键词检测可以利用去除噪声的方法来去除环境噪声，然而在基于加速度进行关键词检测的方案中，该类型方法并不能直接使用。

在文章中，作者提出利用加速度可以精确区分出男女以及粗略区分使用者，因此基于加速度进行关键词检测的方法可以结合认证技术，发展出专门用于用户的语音控制应用。