

基于声音空间梯度的高稳健性击键识别方法

刘影¹, 韩康康², 钱志鸿³

(1. 辽宁工程技术大学电子与信息工程学院, 辽宁 葫芦岛 125105; 2. 北京邮电大学电子工程学院, 北京 100876;
3. 吉林大学通信工程学院, 吉林 长春 130022)

摘 要: 针对声音信号受实际环境噪声影响引起耳蜗倒谱系数 (CFCC) 波动是导致击键内容识别率低的主要原因, 研究相邻键 CFCC 之间的空间特征, 建立基于点的 CFCC 空间梯度结构; 在此基础上, 在训练集和测试上研究 CFCC 空间梯度对击键内容识别的影响及确切邻域的点选取; 最后, 建立基于声音的高稳健性击键内容识别方法。在不同环境下进行的实验表明, 所提 CFCC 空间梯度声音特征效果较好, 识别准确率为 96.15%。

关键词: 击键检测; 声音识别; CFCC 空间梯度; 神经网络

中图分类号: TP391

文献标识码: A

doi:10.11959/j.issn.1000-436x.2020095

High-robustness keystroke recognition method based on acoustic spatial gradient

LIU Ying¹, HAN Kangkang², QIAN Zhihong³

1. School of Electronic and Information Engineering, Liaoning Technical University, Huludao 125105, China
2. School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
3. College of Communication Engineering, Jilin University, Changchun 130022, China

Abstract: For the fluctuation of CFCC caused by environmental noise is the main reason for the low accuracy of keystroke detection, the spatial characteristics of adjacent between CFCC were studied, and the spatial gradient structure of CFCC based on points was established. On this basis, the effect of CFCC spatial gradient on keystroke content recognition and the selection of precise neighborhood points were studied on training and testing. Finally, a high-robustness keystroke recognition algorithm based on acoustic signals was constructed. Extensive experiments in different environments demonstrate that the proposed CFCC spatial gradient sound feature achieves great performance and the recognition accuracy is 96.15%.

Key words: keystroke detection, acoustic recognition, CFCC spatial gradient, neural network

1 引言

电子设备和机械设备信号泄露一直是安全领域广泛关注与重点研究的问题^[1]。早在 1910 年, 德

国科学家就在法国的电话线上进行了实验^[2]。1943 年, 贝尔电话公司的一名工程师在使用 131-B2 型贝尔电话时, 用示波器发现了电磁信号, 这是美国陆军和海军用来传输战时通信对抗德国和日本

收稿日期: 2019-12-17; 修回日期: 2020-03-24

基金项目: 国家重点研发计划基金资助项目 (No.2018YFB1403300); 辽宁省教育厅一般项目资金资助项目 (No.LJ2017QL013); 辽宁省博士启动基金资助项目 (No.20170520098); 辽宁省自然科学基金资助项目 (No.2019-ZD-0038); 辽宁省创新人才支持计划基金资助项目 (No.LR2016045)

Foundation Items: The National Key Research and Development Program of China (No.2018YFB1403300), The Project of Educational Commission of Liaoning Province (No.LJ2017QL013), Liaoning Province Doctor Startup Fund (No.20170520098), The Natural Science Foundation of Liaoning Province (No.2019-ZD-0038), The Project of Innovative Talents Support of Liaoning Province (No.LR2016045)

密码分析的绝密加密电传终端^[3]。研究人员通过提取和破译设备产生的电磁干扰来窃听视频显示单元，该方法在很远的距离或者有障碍物遮挡下都能实现。这种探测之所以能成功，主要是因为使用了精密仪器，这也增加了部署难度^[4]。

智能手机终端除了具备通话功能外，也具备了PDA的大部分功能，还为软件运行和内容服务提供了平台，此外，各种传感器的显著增长使这些设备能够与周围的事物积极互动，并以非预期的方式使用。事实上，恶意软件可能会进入智能手机的摄像头并拍摄照片或视频，或者激活麦克风录制环境声音。近年来，键盘成为计算机系统中人机交互的常用工具，通过键盘输入的内容可以从智能手机麦克风获取的声音和加速度传感器的振动中恢复。当目标用户敲击键盘（以下简称击键）时，就可以记录下这些信号信息，根据收集的数据来恢复击键。如果这些信息被攻击者获取，将会给用户带来极大的危害。因此，针对键盘击键内容的研究对反监听有着重要意义。

迄今为止，研究人员提出了一系列的方案来识别用户击键内容，根据用户击键产生不同的音频信号，进而根据获取的音频信号可识别用户输入的文本信息^[5]。文献[6]使用声音信号来识别用户击键内容，以快速傅里叶变换(FFT, fast Fourier transform)来提取击键信号频谱图上的峰值作为声音信号的特征，通过训练神经网络对击键信号进行识别，准确率为80%。文献[7]通过分析人们在VoIP通话过程中经常从事的辅助活动(包括打字)，再通过传送的音频信息来重建输入——在远程键盘上键入的信息，准确率达到91.7%。由于在实际环境中噪声的影响很大，导致信噪比很小，因此准确率不是很高。文献[8]提出使用基于语义和音频信号感知用户键盘输入的内容，准确率达到95%。文献[4]提出使用同一手机的不同麦克风接收同一击键声音信号的到达时间，通过到达时间差(TDA, time difference of arrival)信息来精确找出按键的位置。利用麦克风阵列，基于到达时间差的传统声源定位算法在最好的情况下能实现厘米级的定位^[9]，但该类方法需要测量设备具有多个麦克风。

综上所述，基于声音的击键内容识别方法大多是直接提取单个按键的原始音频特征，但该音频特征往往存在时间不稳定性以及空间模糊性，导致提

取到的音频信号存在较大误差。因此本文提出一种基于声音空间梯度的击键识别系统(ASGK, acoustic-signal spatial gradient keystroke tracking system)，通过多信号空间特征代替已有的单信号空间特征。这种声音信号特征在空间上具有区分性，在时间上更加稳定，即来自邻域的声音信号特征的空间关系比来自单个键的声音信号特征更具有稳健性。例如，如果有语音信号干扰，那么键盘上其他键感知到的声音信号都将会受到影响，从而使相对声音信号变化小于绝对声音信号值。因此，本文提出的声音空间梯度刻画比普通的声音信号特征更加稳定和可识别，因为它可以很好地容忍环境变化和多样性引起的声音信号特征变化。

2 声音信号特征提取

语音特征参数提取是影响语音识别系统的关键技术之一。目前，语音特征参数提取大多是基于傅里叶变换或小波变换后提取语音信息，但是傅里叶变换仅包含语音的频率特性，而小波变换包含时域和频域信息^[10]。耳蜗倒谱系数(CFCC, cochlear filter cepstral coefficient)特征参数依据人耳听觉特性，具有时域和频域分析的优势，与梅尔频率倒谱系数(MFCC, Mel frequency cepstral coefficient)特征参数和感知线性预测(PLP, perceptual linear prediction)特征参数相比，在抗噪方面和识别效果方面具有良好的性能。因此，本文采用CFCC特征参数作为音频信号特征。

假设将 ω 个需要提取声音信号的键描述为 $k \in (1, \omega)$ ，每个键的击键声音用 P 个采样 $(f_1^k, f_2^k, \dots, f_p^k)$ 组成，这些击键声音反映了不同键的不同发音和语音韵律等特征。根据CFCC特征参数提取方法^[11]，将录入的击键声音信号 $(f_1^k, f_2^k, \dots, f_p^k)$ 代入CFCC听觉变换式，具体定义为

$$T(a, b) = \int_{-\infty}^{+\infty} f(t) \psi_{a,b}(t) dt \quad (1)$$

其中， $f(t)$ 为时域语音信号； $\psi_{a,b}(t)$ 为耳蜗滤波函数，具体计算式为

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \left(\frac{t-b}{a} \right)^\alpha \exp \left[-2\pi f_L \beta \left(\frac{t-b}{a} \right) \right] \cdot \cos \left[2\pi f_L \left(\frac{t-b}{a} + \theta \right) \right] u(t-b) \quad (2)$$

其中， $\alpha > 0$ 和 $\beta > 0$ 决定 $\psi_{a,b}(t)$ 的频域形状和宽度，

一般取 $\alpha=3$ 、 $\beta=0.2$ 时降噪效果良好； $u(t)$ 为单位阶跃函数； θ 为初始相位，控制冲激响应的角度，它的取值一般满足 $\int_{-\infty}^{+\infty} \psi_{a,b}(t) dt = 0$ ； a 为尺度变量， b 为随时间可变的实数，一般 a 可由滤波器组的中心频率 f_c 和最低中心频率 f_L 确定，即 $a = \frac{f_L}{f_c}$ 。

接下来，对 $T(a, b)$ 进行变换，得到基于人耳耳蜗的语音为

$$h(a, b) = T(a, b)^2, \forall T(a, b) \quad (3)$$

$$S(i, j) = \frac{1}{d_i} \sum_{d_i}^{\alpha+d_i-1} h(i, b) \quad (4)$$

$$y(i, j) = \sqrt[3]{s(i, j)} \quad (5)$$

其中， $\alpha = 1, 2F, \dots, \forall i, j$ ， F 是帧移，实验中一般取 $\frac{d}{2}$ ， $d = \max\{3.5\tau_i, 20\text{ms}\}$ 是可变窗长度， $\tau_i = \frac{1}{f_c}$ 是第 i 个滤波器中心频带中心频率的时间长度， j 是窗的个数。

最后将非线性响应变换输出 $y(i, j)$ 通过离散余弦变换 (DCT, discrete cosine transform) 式进行变换，去除通道之间的相关性，最终得到 CFCC 特征参数为

$$\text{CFCC}_{(n,j)} = \sqrt{\frac{2}{M} \sum_{m=1}^{M-1} y(m, j) \cos\left(\frac{\pi n \left(m - \frac{1}{2}\right)}{N}\right)} \quad (6)$$

其中， M 为耳蜗滤波器的个数， n 为特征变换后每帧特征的维数， $0 < n < N$ ， $0 \leq m \leq M$ 。

3 ASGK 系统的设计与实现

本文使用手机中的单麦克风感知用户键盘输入的内容，识别工作和娱乐中的计算机活动，ASGK 方案架构如图 1 所示。具体步骤如下。

步骤 1 采集多组击键时每个字母的声音信号，并给每组声音信号加上类别标签，形成样本集。

步骤 2 采用 CFCC 的方法提取每个样本的特征参数，与样本的类别标签形成新的样本集，并将其按一定比例划分为训练集和测试集。

步骤 3 通过 BP 神经网络对训练集进行训练，得到训练好的 BP 神经网络模型。

步骤 4 将测试集输入训练好的 BP 神经网络模型中进行测试，得到测试集中每个样本的类别。

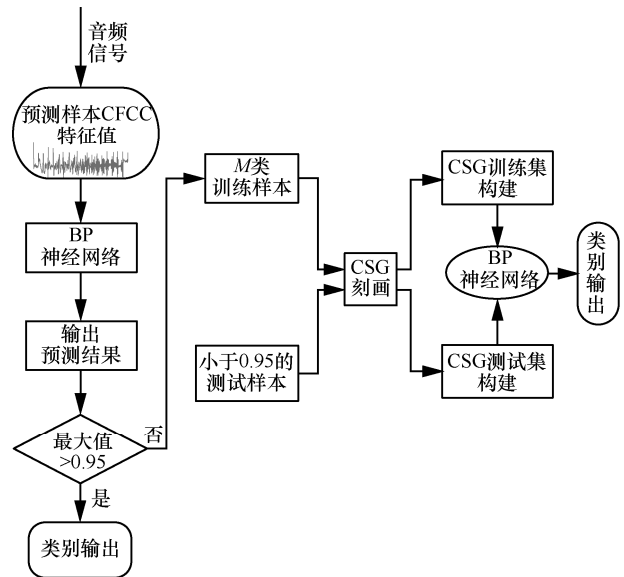


图 1 ASGK 方案架构

步骤 5 设定一个测试准确率阈值 μ (μ 取经验值 0.95)，计算同一类别样本的测试准确率并与阈值 μ 进行比较，若大于阈值 μ 则输出类别结果，否则执行步骤 6。

步骤 6 通过计算测试准确率低于阈值 μ 的 γ 键与其他键之间的曼哈顿距离及 CFCC 差值，构建 γ 键的 CFCC 空间梯度 (CSG, CFCC spatial gradient) 矩阵，详见 3.1 节。

步骤 7 基于 γ 键的 CSG 矩阵，构建新的训练集和测试集，详见 3.2 节。

步骤 8 根据步骤 7 得到新的训练集进行神经网络训练，将步骤 7 得到的测试样本集输入新训练好的神经网络中，从结果中选择分类概率最高的类别作为 γ 键的输出类别，详见 3.3 节。

考虑与现有方法的兼容，该方法的输入是基于 CFCC 的声音识别系统，其核心贡献是引入了 CSG 训练集和 CSG 测试集。

3.1 CSG 的定义

为了解决 CFCC 时间不稳定性以及空间模糊性，ASGK 基于传统的 CFCC 音频特征充分挖掘音频特征值相似度中的空间关系限制，并提出了 CSG 这种新的结构。通常提取到某个键的 CFCC 特征值会发生变化，但是它与相邻位置键的 CFCC 特征值之差会保持相对稳定。具体来说，G 键与 H 键相邻，在 t_1 时刻提取到 2 个键的 CFCC 特征值分别为 CFCC_G^1 与 CFCC_H^1 ，在 t_2 时刻提取到 2 个键的 CFCC 特征值分别为 CFCC_G^2 与 CFCC_H^2 ，尽管在不同时刻

收集到同一键的音频特征之间 $|CFCC_G^1 - CFCC_G^2|$ 会发生不小的变化，但是相邻键在不同时刻收集到的 CFCC 之差会保持相对稳定，即 $|CFCC_G^1 - CFCC_H^1| \approx |CFCC_G^2 - CFCC_H^2|$ ，这是因为即使 CFCC 音频特征会发生改变，但这个改变带来的影响在 2 个位置上会比较小，这样就可以找到在变化过程中的一个稳定量。

CSG 是利用键盘上多个键之间的空间特征，而不是单一使用 CFCC 特征值来表示的。如图 2 所示，对于一个特定的键，CSG 矩阵刻画了该键的 CFCC 与其相邻几个键 CFCC 的相似趋势。

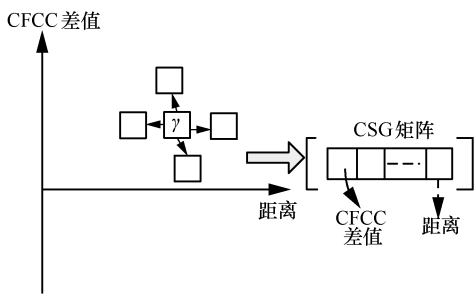


图 2 CSG 刻画

对于任意 γ 键的 CSG 矩阵定义为

$$g = \left\{ \left\{ d(\gamma, \mu), \phi(CFCC_\gamma, CFCC_\mu) \right\} \right\} \quad (7)$$

其中， $CFCC_\gamma$ 表示 γ 键提取的声音信号特征，有 P 个采样，即 $\{CFCC_\gamma^{(x)}, 1 \leq x \leq P\}$ ；同理 $CFCC_\mu$ 表示 μ 键提取的声音信号特征，有 P 个采样，即 $\{CFCC_\mu^{(x)}, 1 \leq x \leq P\}$ ； $\phi(CFCC_\gamma, CFCC_\mu)$ 表示 γ 键和 μ 键之间的 CFCC 差值，即 $\phi(CFCC_\gamma, CFCC_\mu) = CFCC_\gamma - CFCC_\mu$ ； $d(\gamma, \mu)$ 表示 γ 键和 μ 键的曼哈顿距离，从 γ 键周围空间中选择 i 个键，应该出现在 γ 键的周围子空间。在特定 γ 键上的相似度为 0，然后增大，对应到 γ 键的有序物理距离通过实验测试得到。

3.1.1 CSG 的优势

通常，对于 2 个相邻键采集到的声音信号较相似，而距离较远的击键信号特征差异较大，CSG 的构建关键在于采用较远键的空间特征。当面对信号变化时，各个信号的 CFCC 特征值可能发生变化，而 2 个键之间存在的某种空间特征（差值）能保持相对稳定。对于一个易受到环境等影响的特定键，CFCC 特征值可能存在着一一定的不确定性，当通过 CSG 进行刻画后，CFCC 信号相似性的空间趋势预计比单个键的 CFCC 本身更加稳定，如图 3 所示。G 键和 H 键的 CFCC 图谱极其相似，当 G 键和 H 键选择较近键作差时，得到的

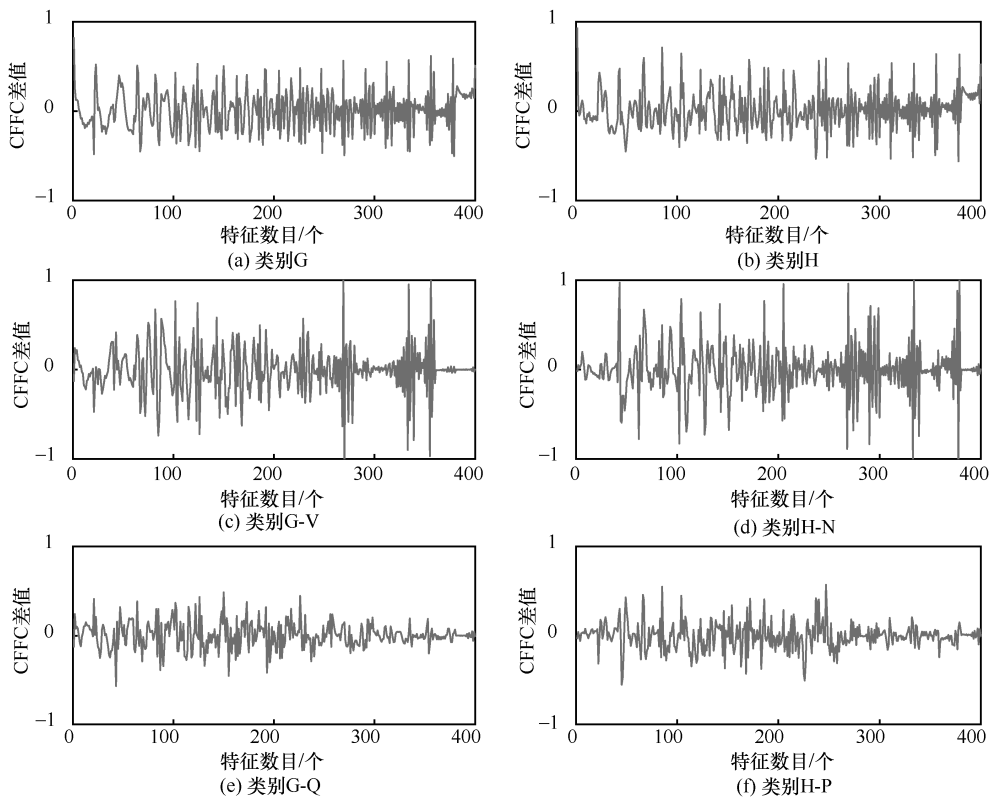


图 3 几种类别的 CFCC 差值

信号图谱也比较相似。如果 G 键选 Q 键、H 键选 P 键作为邻域，相比于 G 键选 V 键、H 键选 N 键作为邻域得到的信号图谱更具有区分性。因此，利用这种 CFCC 空间梯度作为更有利的特征，用于声音信号的映射。

3.1.2 CSG 的实现

每个键的 CSG 矩阵都构建在原始 CFCC 上，通常每个键的周围都有多个相邻的 CFCC。对于任意 G 键，首先形成一个包含当前键位置的、距离不超过 i 个键的所有相邻位置的邻域集合，然后计算 G 键与相邻集合中所有键的 CFCC 差值。CSG 构建方法为最大化全局梯度的构建方式，即选择曼哈顿距离较远的键作为邻域。根据键盘的形状建立如图 4 左下侧的坐标系，将每个字母用坐标系上的点表示，以此计算曼哈顿距离。如字母 Z 和字母 F 的坐标依次表示为(1,1)和(4,2)。曼哈顿距离计算式为 $distance = |1 - 4| + |1 - 2| = 4$ 。

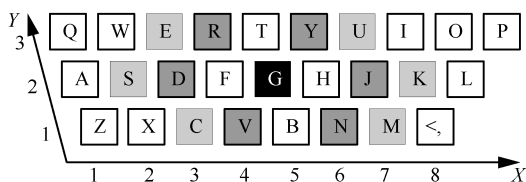


图 4 最大化全局梯度选取

当 $distance=2$ 时，以 G 键为例，取曼哈顿距离为 2 最小的键依次为字母 R、字母 D、字母 V、字母 N、字母 J 和字母 Y，曼哈顿距离为 3 最小的键依次为字母 E、字母 S、字母 C、字母 U、字母 K 和字母 M。曼哈顿距离由小到大，在字母 G 键周围选取一个距离 G 键最远的键作为差值样本。

3.2 基于 CSG 构建训练集和测试集

训练集的 CSG 构建，建立于第一次所有测试样本预测完成基础之上。例如，对于原始测试集 N 个样本中的一个样本，神经网络输出小于 0.95，测试样本命名为 X^* 。假设取神经网络输出 M (M 取经验值 3) 个较大值的类别作为“相似类别”，分别命名为类别 D、类别 B 和类别 F。其后根据图 4 中的差值选取规则，选取类别 D、类别 B 和类别 F 的邻域样本，每类有 100 个原始样本。将类别 D、类别 B 和类别 F 训练集中原始样本的 CFCC 值与得到的 M 个相似类别的邻域集合的训练集样本 CFCC 值分别作差，得到新的训练样本集。 $M=3$ 时测试集和训练集构建如图 5 所示。

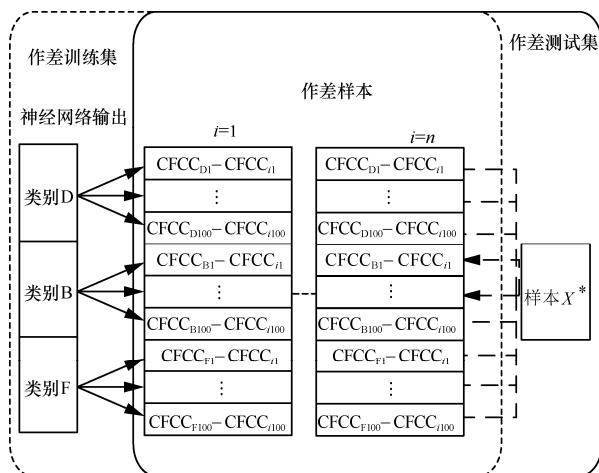


图 5 $M=3$ 时测试集和训练集构建

将 X^* 测试集中原始样本 CFCC 值与得到的 M 个相似类别的邻域集合的测试集原始样本 CFCC 值分别作差，得到新的测试集。

3.3 基于 BP 神经网络 CSG 击键声音特征识别

基于声音信号的击键内容识别，实质是对击键声音信号进行预测分类，运用分类方法完成从击键音频信号的 CFCC 特征数据到击键类别的映射。神经网络方法具有很强的非线性拟合能力，实现具有非线性关系的输入向量到输出向量之间的映射^[12]。因此本文分类算法采用神经网络算法。

基于 BP 神经网络的击键声音特征信号分类算法建模包括 BP 神经网络构建、BP 神经网络训练和 BP 神经网络分类。1) BP 神经网络的构建根据击键声音信号输入输出数据特点来确定 BP 神经网络的结构。在神经网络的设计中，采用三层神经网络的结构，分别是输入层、隐含层和输出层。输入层神经元的个数与输入 CFCC 特征值个数相同，为 400 个；输出层个数与字母类别相同，为 26 个；对于隐含层的个数，采用试凑法确定。2) 对于 BP 神经网络的训练，用训练数据训练 BP 神经网络。共 26×110 组击键声音信号，每类 100 个样本作为训练数据训练网络，每类剩余的 10 组数据作为测试数据训练网络分类能力。输入向量为 400×1 的列向量，输出向量为 26×1 的列向量，取输出向量最大值的行数作为样本输出类别。3) 测试集的实际分类与预测分类。测试样本共 26 类，每类 10 个样本，共 260 个样本。

4 仿真结果分析

4.1 仿真环境介绍

1) 实验平台搭建及数据采集

本文实验中使用一部 OPPO R7s 单麦克风智能

手机和一副星字-black Window 键盘，手机可放置在点画线上的任意位置，如图 6 所示。实验中使用智能手机的用户对单个键盘字母的击键声音进行采集，采样频率为手机系统默认的 48 000 Hz。采集样本为 26 个字母，每个字母采集 110 个样本，其中 26×100 个样本作为训练集，26×10 个样本作为测试集。

2) 系统结构介绍

基于声音信号的击键识别系统主要包括信号采集、信号数据处理、分类算法 3 个部分。第一部分为信号采集，为手机录下用户敲击电脑键盘上字母的音频。第二部分为数据处理，包括通过巴特沃斯和小波变换对击键音频信号进行降噪处理，其后提取 CFCC 特征参数 CSG。第三部分为分类算法，主要包括神经网络算法。

3) 评价指标

对于击键识别的性能分析，采用准确率和相关系数作为 2 个评价指标，计算式分别为

$$accuracy = \frac{\text{测试样本正确总数}}{\text{测试样本总数}} \quad (8)$$

$$R^2 = \frac{\left(L \sum_{i=1}^L \hat{y}_i y_i - \sum_{i=1}^L \hat{y}_i \sum_{i=1}^L y_i \right)^2}{\left(L \sum_{i=1}^L \hat{y}_i^2 - \left(\sum_{i=1}^L \hat{y}_i \right)^2 \right) \left(L \sum_{i=1}^L y_i^2 - \left(\sum_{i=1}^L y_i \right)^2 \right)} \quad (9)$$

其中， $\hat{y}_i (i=1,2,\dots,L)$ 为第 i 个样本的预测值， $y_i (i=1,2,\dots,L)$ 为第 i 个样本的真实值， L 为样本数目。相关系数在 [0,1] 范围内越接近 1，则表明模型的性能越好；反之，则表明模型性能越差。

4.2 参数影响

1) M 参数选择

M 取值与判定样本数量之间的关系如图 7 所示。当 M 为 1~7 时，曲线增速较快，此时构建 CSG 训练样本增加速度较快，可以满足训练样本数量。只要 M 大小适中，训练样本集的样本数量介于 100 左右，对于网络训练迭代次数影响不大。当 M 为 8~26 时，曲线逐渐趋于平缓，即重新判定样本的数量增加较少，但是给构建 CSG 样本增加了计算量，使神经网络增加了误判的概率。由此可知， M 的取值范围应在 1~7，考虑到 CSG 的构建与曼哈顿距离有关，因此将 2 个参数同时考虑，接下来通过仿真给出最佳的 M 值和曼哈顿距离。

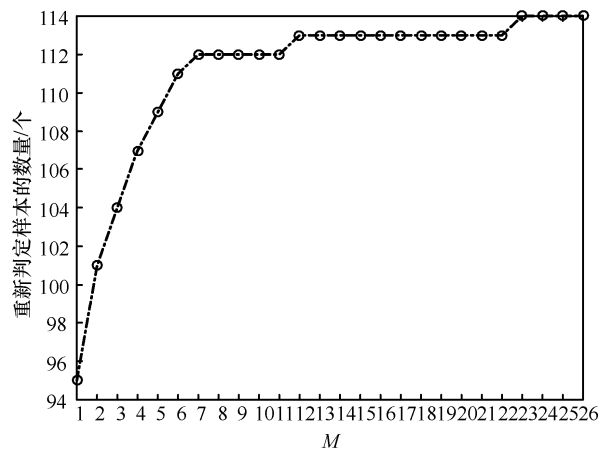


图 7 M 取值与判定样本数量之间的关系

2) 邻域数的影响

对于 CSG 的实现关键在于曼哈顿距离的选取，在不同的 M 取值下，随着曼哈顿距离的增加，准确

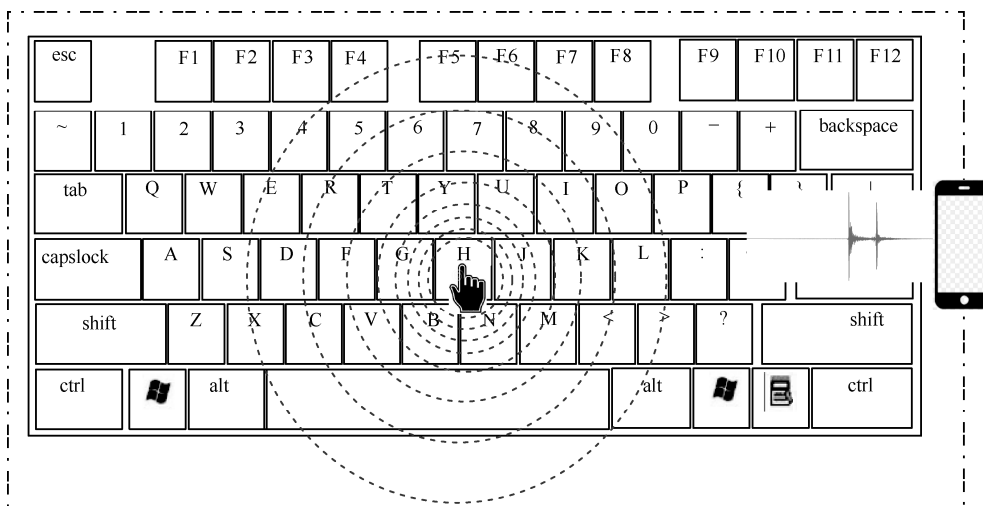


图 6 实验平台

率的变化趋势如图 8 所示。

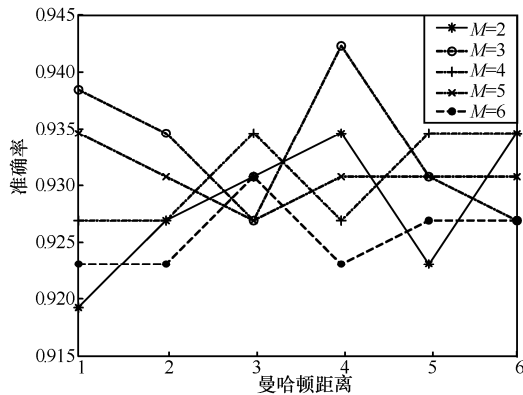


图 8 准确率的变化趋势

由图 8 可知，在同一个曼哈顿距离下，准确率不随 M 取值的增大而增大。因为在同一曼哈顿距离范围内，差值选取过多，空间特征具有更多相似性。在同一 M 取值下，曼哈顿距离的增大并未导致准确率的上升。所以在进行系统性能准确率提升方面， M 取值与曼哈顿距离均需要考虑。当 $M=3$ 、曼哈顿距离=4 时，准确率最高。

4.3 方法对比

为了验证本文方法的性能，采用 CFCC+SVM、CFCC+BP 和 CSG+BP 这 3 种方法进行对比。本文方法的主要目标是展示 CSG 相对于 CFCC 特征的优势，由表 1 可以看出，CSG+BP 方法的识别准确率达到 96.15%，在 3 种方法中表现最优。

方法	输入数据形式	准确率	相关系数
SVM	CFCC	43.46%	0.25
BP	CFCC	90%	0.85
BP	CSG	96.15%	0.9196

本文方法仅基于 CFCC 特征就可以获得显著的性能提升，而不需要额外的信息。如果通过附加信息（如多麦克风），则可以轻松地在 CSG+BP 方法中实现进一步的性能改进，并且 CSG+BP 方法与文献[5-8,13]相比具有较好的性能。

4.4 不同实验环境下系统性能分析

击键声音信号采集环境中的噪声会对击键信号的识别造成巨大干扰，因此在不同实验环境下对击键识别系统性能评估必不可少。图 9 为嘈杂环境下信噪比较低的声音样本频谱图，图 10 为安静环境下信噪比较高的声音样本频谱图。在这 2 种环境

下，分别对传统的 BP 神经网络以及基于 CFCC 差值的 CSP 网络进行评估，评估结果如表 2 所示。由表 2 可知，CSG+BP 比 CFCC+BP 具有一定的环境干扰抵抗能力。

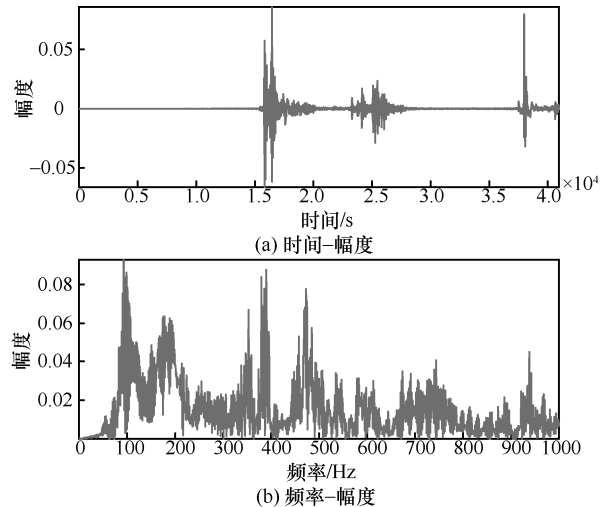


图 9 嘈杂环境下信噪比较低的声音样本频谱图

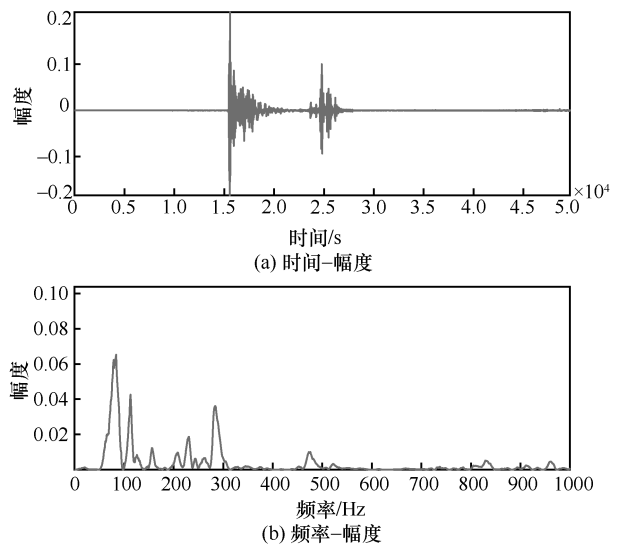


图 10 安静环境下信噪比较高的声音样本频谱图

表 2 不同环境下的准确率

方法	安静环境	嘈杂环境
CFCC+BP	91%	83%
CSG+BP	96%	90%

5 结束语

本文提出了一种基于 CFCC 空间梯度的击键内容识别方法的设计与实现，该系统在不影响普遍性的前提下，缓解了 CFCC 的空间模糊性和时间不稳

定性, 获得了稳定可靠的性能。其主要思想是利用 CFCC 值的空间感知能力, 将 CSG 表示为增强的声音信号特征, 并设计了用于表示、构造所提出的 CSG 矩阵的技术, 结合 BP 神经网络集成为一个完全实用的系统。在不同环境下进行算法验证, 实验中分析了 M 值和曼哈顿距离的选取, 并与多种系统进行对比分析。实验结果表明, 本文方法准确率达到 96.15%, 相关系数为 0.919 6, 均优于 CFCC+BP 和 SVM+CFCC 等当前主流的击键识别算法。

参考文献:

- [1] 张弛. 神经网络算法在电磁泄露及脑电信息认知中的应用[D]. 北京: 北京邮电大学, 2018.
ZHANG C. Application of neural network algorithm in electromagnetic[C]. Beijing: Beijing University of Posts and Telecommunications, 2018.
- [2] BAUER A O. Some aspects of military line communications as deployed by the german armed forces prior to 1945[C]//In Proceedings of 5th Annual Colloquium, The History of Military Communications. [S.n.: s.l.], 1999:1-21.
- [3] 张南. 计算机电磁信息泄露与视频信息认知研究[D]. 北京: 北京邮电大学, 2019.
ZHANG N. Research on computer electromagnetic information leakage and video[D]. Beijing: Beijing University of Posts and Telecommunications, 2019.
- [4] ZHU T, MA Q, ZHANG S F, et al. Context-free attacks using keyboard acoustic emanations[C]//ACM SIGSAC Conference on Computer and Communications Security. New York: ACM Press, 2014: 453-464.
- [5] ABHISHEK S, SAXENA N. Keyboard emanations in remote voice calls: password leakage and noise(less) masking defense[C]//ACM Conference on Data and Application Security and Privacy. New York: ACM Press, 2018: 103-110.
- [6] ASONOV D, AGRAWAL R. Keyboard acoustic emanations[C]//IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2004: 3-11.
- [7] COMPAGNO A, CONTI M, LAIN D, et al. Don't skype & type! Acoustic eavesdropping in voice-over-IP[C]//ACM on Asia Conference on Computer and Communications Security. New York: ACM Press, 2017: 703-715.
- [8] DU H, YU Z W, XIAO D. Sensing keyboard input for computer activity recognition with a smartphone[C]//ACM International Joint Conference on Pervasive and Ubiquitous Computing. New York: ACM Press, 2017: 25-28.
- [9] ZHENG X S, YANG S L, JIN N G, et al. Diva: distributed voronoi-based acoustic source localization with wireless sensor networks[C]//IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2016: 1-9.
- [10] 米建伟, 方晓莉, 仇原鹰. 非平稳背景噪声下声音信号增强技术[J]. 仪器仪表学报, 2017, 38(1): 17-38.
MI J W, FANG X L, QIU Y Y. Enhancement technology for the audio signal with nonstationary background noise[J]. Chinese Journal of Scientific Instrument, 2017, 38(1): 17-38.
- [11] 史燕燕, 白静. 融合 CFCC 和 Teager 能量算子倒谱参数的语音识别[J]. 计算机科学, 2019, 46(5): 286-289.
SHI Y Y, BAI J. Speech recognition combining CFCC and Teager energy operators cepstral coefficients[J]. Computer Science, 2019, 46(5): 286-289.
- [12] ZHANG J, ZUO L, GAO J W. Digital instruments recognition based on PCA-BP neural network[C]//IEEE Information Technology, Networking, Electronic, and Automation Control Conference. Piscataway: IEEE Press, 2017: 928-932.
- [13] 沈伟国, 王巍. 基于稳健线性判别分析的击键特征识别方法[J]. 通信学报, 2017, 38(Z2): 26-29.
SHEN W G, WANG W. Keystroke features recognition based on stable linear discriminant analysis[J]. Journal on Communications, 2017, 38(Z2): 26-29.

[作者简介]



刘影(1983-), 女, 吉林大安人, 博士, 辽宁工程技术大学副教授, 主要研究方向为无线定位、智能无线感知。



韩康康(1997-), 男, 河南商丘人, 北京邮电大学硕士生, 主要研究方向为认知无线电、智能信息处理。



钱志鸿(1957-), 男, 吉林长春人, 博士, 吉林大学教授、博士生导师, 主要研究方向为基于物联网、D2D、Wi-Fi、RFID 等的无线网络与通信技术。