

基于 GMM 和枪声的军事环境判别<sup>\*</sup>

刘力维 袁高高 潘志刚 董 俊

(63898 部队 济源 454650)

摘 要 枪声在军事环境下是一个令人高度重视的声音信号。针对其所含信息量的重要性, 提出了一种基于枪声检测的军事环境判断方法。首先利用基于中值滤波的算法检测出枪声的起止点; 采取反应人耳听觉机理的声音信号动态特征—Mel 倒谱系数描述声音信号; 然后利用基于高斯混合模型的枪声检测及识别系统完成对枪声的分析从而实现当前军事环境的判断。实验结果证明了该方法的准确性和有效性。

关键词 中值滤波; GMM; MFCC; 最大似然估计  
中图分类号 0211

Military Environment Discrimination Based on  
GMM and Gunshot Detection

Liu Liwei Yuan Gaogao Pan Zhigang Dong Jun

(No. 63898 Troops of PLA, Jiyuan 454650)

**Abstract** The sound of gunshots is very important for the military environment. According to the importance of gunshots, a novel method is proposed to discriminate the military environment, based on the detection of gunshots. First, the start and the end of gunshots is detected utilizing the median filtering. And then the MFCC parameter is choosed to describe the sound signal. The military environment is discriminated by the recognition of gunshots with the help of GMM. The experimental results conducted on the database of gunshot sounds show that the proposed method substantially performs very well not only in recognition rate but also in efficiency.

**Key words** median filter, GMM, MFCC, MLE

**Class Number** 0211

1 引言

声音信号是包含信息量最多的信号之一, 且传播媒介多种多样分布非常广泛, 获取也比较方便。同时声音也是人们交流最基本最有效的方式。自 20 世纪以来, 声音处理尤其是语音识别技术有了长足的进步和成功的应用; 如说话人识别技术在门禁系统、网上银行和刑事侦查中的应用<sup>[1]</sup>。枪声作为声音信号中非常特别的一种, 一直受到关注和重

视。在军事领域, 枪声检测和识别技术有着极其重要的意义。枪声识别技术能够有效及时地反应军事环境下的异常信息, 并且与雷达探测技术相比有着抗干扰、不易被发现、成本低等特点。同样, 在现今社会枪声的检测和识别技术对于环境监控, 安全防护也有着不可限量的作用<sup>[2]</sup>。

高斯混合模型(GMM)是比较成熟的语音识别技术, 可以很好地描述说话人在不同环境和生理条件下的声音特征, 已广泛地应用于说话人辨认系

<sup>\*</sup> 收稿日期: 2009 年 2 月 14 日, 修回日期: 2009 年 3 月 22 日  
作者简介: 刘力维, 男, 工程师, 研究方向: 电子装备试验技术应用。袁高高, 男, 工程师, 研究方向: 电子装备试验技术应用。潘志刚, 男, 工程师, 研究方向: 电子装备试验技术应用。董俊, 男, 工程师, 研究方向: 通信装备试验技术应用。

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

统<sup>[3]</sup>。它是用多个高斯分布的概率密度函数的组合来描述特征矢量在概率空间的分布状况。每一类声音信号对应着一个 GMM<sup>[4]</sup>。高斯混合模型(GMM)具有简单高效的特点,因此本文利用高斯混合模型检测环境中的声音信号是否包含了枪声,从而实现军事环境的判断。

## 2 枪声的检测和识别

### 2.1 枪声的检测

在安全状态的军事环境下,背景声音属于低频噪声。而枪声则是突然发生的与紧急事故相关的声音。首先,枪声相对于其周边的正常声音强度要大;其次,枪声的突发性决定了其在发生起始时刻点的能量必然高于其周围的背景声音能量。因此,通过检测声音信号能量的变化可以检测出枪声的发生。

表 1 异常声音库

名称	文件数	文件长度(ms)
枪声	200	65—2138

枪声的声音信号经过 16000Hz 的采样频率采样和 16 位精度的 A/D 转换量化后得到离散的数字信号。表 1 所示为本研究所包括的枪声数据库。枪声的主要检测方法是利用非线性的中值滤波来分析声音信号在时域的能量变化,当检测到能量变化超过阈值时则判定有可能发生枪声,并把可疑信号带入 GMM 模型的识别系统进行判断。枪声检测的主要步骤是:首先对枪声信号进行分帧,本研究中取帧长 128ms 包括 2048 个采样点,帧移位 96ms;并计算出每一帧的能量值。然后,对所求能量序列利用 10 阶的中值滤波其进行滤波得到新的能量序列,该能量序列是经过归一化的,并且序列中的枪声段被突显出来。最后,利用动态自适应阈值判断每一帧能量序列,确定是否有枪声发生。图 1 所示为枪声检测系统的流程框图。

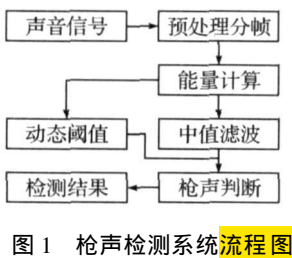


图 1 枪声检测系统流程图

最后,利用动态自适应阈值判断每一帧能量序列,确定是否有枪声发生。图 1 所示为枪声检测系统的流程框图。

### 2.2 枪声的识别

#### 2.2.1 枪声特征参数的选取

声音信号特征的 Mel 倒谱系数(MFCC)模拟人耳听觉感知特性:不同频率的声音会在耳内基础膜的不同位置振动。这样人耳就可以很容易分辨出各种声音,所以此参数作为异常声音的识别参数

能更好地反应各种异常声音信号的特性,从而大大提高鲁棒性和识别率<sup>[5]</sup>。因此,本研究中选用枪声的 MFCC 作为特征参数表征枪声的特性。

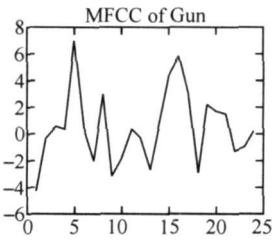


图 2 枪声帧信号的 MFCC 系数

MFCC 参数是基于 Mel 频率尺度的实对数短时能量的余弦变换,以模拟人的听觉感知。它的计算首先用 FFT 将时域信号转化为频域。然后对其对数能量谱用按 Mel 刻度分布的三角波滤波器组进行卷积,最后对每个滤波器的输出构成的向量进行离散余弦变换 DCT,取前  $N$  个(为 MFCC 参数的阶数)系数<sup>[6]</sup>。其分析属于利用数字滤波器组进行的频谱分析,数字滤波器组可看作是人类听觉系统初始传导阶段的粗糙模型,其设计基于以 Mel 频率尺度为线性标度的频率映射:

$$f_{mel} = 2595 \log_{10} \left( 1 + \frac{f}{700} \right) \quad (1)$$

其中:  $f_{mel}$  频率,  $f$  为线性频率。图 2 所示为一帧枪声信号的 MFCC 参数(24 阶)。

#### 2.2.2 枪声的识别系统

高斯混合模型用多个具有高斯分布的概率密度的加权和来表示,该概率密度函数的个数称之为高斯模型的混合数<sup>[7]</sup>。一个具有  $M$  个混合数的  $d$  维 GMM, 可以表示为:

$$P(x|\lambda) = \sum_{i=1}^M \omega_i p_i(x; \mu_i, \Sigma_i) \quad (2)$$

其中:  $x$  为  $d$  维观察矢量;  $\omega_i$  为混合权重,且满足  $\sum_{i=1}^M \omega_i = 1$ ;  $p_i(x; \mu_i, \Sigma_i)$  为  $d$  维高斯函数,表示 GMM 模型的第  $i$  个高斯分量,  $\mu_i$  为该高斯分量的均值矢量,  $\Sigma_i$  为协方差矩阵。整个 GMM 由各混合分量的均值矢量、协方差矩阵以及混合权重来描述,用  $\lambda$  来表示该模型,有:

$$\lambda = \{ \omega_i, \mu_i, \Sigma_i \} \quad (i=1, 2, \dots, M) \quad (3)$$

GMM 模型参数的训练一般采用最大似然估计(Maximum Likelihood Estimation, MLE)的方法。假设训练样本的观察矢量序列  $X = \{x_t, t=1, 2, \dots, T\}$  中各观察矢量  $x_t$  是独立不相关的,对于 GMM 模型  $\lambda$  的似然度可表示为:

$$L(\lambda|X) = P(X|\lambda) = \prod_{t=1}^T P(x_t|\lambda) \quad (4)$$

训练的的目的是找到一组模型参数  $\lambda$  使得  $L(\lambda|$

$X)$ 最大, 即:

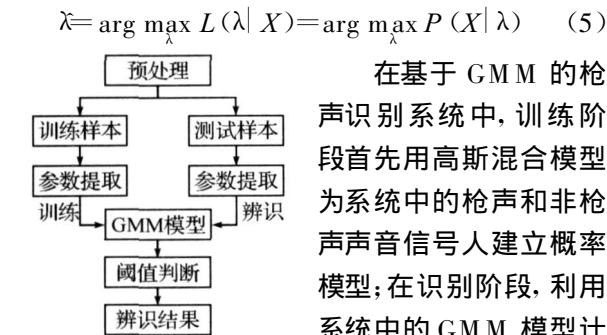


图3 枪声识别系统流程图

然准则做出判决。枪声识别系统的流程如图3所示。这种辨识方法只能用于与GMM模型中的相似类型的声音信号的分类, 但是日常生产生活中, 声音信号千变万化, 容易产生误报。因此, 采用了输出概率阈值来辨识异常声音, 当GMM输出最大概率大于阈值时确定为该类异常, 否则判别为非异常声音。

通过枪声的检测和识别两个阶段, 可以判别军事基地是否发生的枪声, 进而推断是否有紧急情况发生, 也即对军事环境做出了正确的判断。

### 3 实验

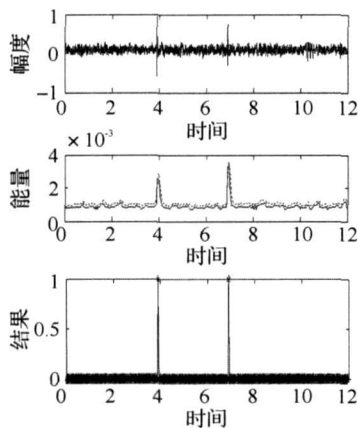


图4 枪声检测结果图

声环境下的枪声检测结果图。表2为枪声在不同信噪比的噪声环境下的检测结果。

表2 噪声条件下的枪声检测率

噪声类型	0db	10db	20db
白噪声	72.48%	85.36%	97.30%
部队背景声	76.84%	84.78%	97.56%

表3 噪声条件下枪声识别率

噪声类型	0db	10db	20db
白噪声	46.00%	82.00%	92.00%
部队背景声	44.00%	84.00%	90.00%

军事环境判断的一个重要过程就是枪声的识别, 该实验过程中, GMM的高斯混合的数目M经验地设置为64。在训练阶段, 将表1中200个枪声的150个随机抽取为训练样本, 训练样本为纯净的枪声不含任何噪声; 首先对训练样本进行预处理, 分帧; 然后提取每一帧信号的MFCC; 最后将MFCC参数序列带入系统的模型中进行迭代计算确定GMM的各项参数。在识别阶段, 枪声中剩余的50个作为测试样本; 首先对测试样本进行加噪, 加入不同信噪比的噪声; 然后再预处理、分帧和计算MFCC; 最后带入模型进行识别。表3为枪声在不同信噪比噪声环境下的识别效率。

### 4 结语

本文通过枪声的检测和识别, 为军事环境的判断提供了正确的依据。文章首先利用枪声的突发性和高声强特性检测军事环境下有无枪声的发生; 再运用GMM和MFCC来识别枪声为军事环境的判断提供准确的依据。实验结果表明了该方法的可行性和有效性。

#### 参考文献

[ 1 ] 杨澄宇, 赵文, 杨鉴. 基于高斯混合模型的说话人确认系统[ J ]. 计算机应用, 2001, 21(4): 7~8

[ 2 ] 潘燕, 梁文革等. 环境声音信号分析[ J ]. 中国医学物理学杂志, 2004, 21(3): 161~164

[ 3 ] Thomas F. Quatieri. Discrete-Time Speech Signal Processing: Principles and Practice[ M ]. 北京: 电子工业出版社, 2004

[ 4 ] 谢霞, 李宏, 郑俊. 基于GMM的说话人辨认系统及其改进[ J ]. 电脑与信息技术, 2006, 4: 48~51

[ 5 ] 刘辉, 杨俊安, 许学忠. 基于MFCC参数和HMM的低空目标声识别方法研究[ J ]. 弹箭与制导学报, 2007, 27(5): 217~222

[ 6 ] 于明, 袁玉倩等. 一种基于MFCC和LPCC的文本相关说话人识别方法[ J ]. 计算机应用, 2006, (4): 883~885

[ 7 ] Douglas A. Reynolds. Speaker identification and verification using Gaussian mixture speaker models [ J ]. Speech Communication, 1995, 17: 97~103