



(12) 发明专利

(10) 授权公告号 CN 103310789 B

(45) 授权公告日 2016. 04. 06

(21) 申请号 201310239724. 7

(22) 申请日 2013. 06. 17

(66) 本国优先权数据

201310166660. 2 2013. 05. 08 CN

(73) 专利权人 北京大学深圳研究生院

地址 518055 广东省深圳市南山区西丽深圳
大学城北京大学校区

(72) 发明人 刘宏 王一 李晓飞

(74) 专利代理机构 北京君尚知识产权代理事务
所(普通合伙) 11200

代理人 余长江

(51) Int. Cl.

G10L 15/06(2013. 01)

G10L 15/02(2006. 01)

(56) 对比文件

CN 1819019 A, 2006. 08. 16,

CN 102426837 A, 2012. 04. 25,

US 6876966 B1, 2005. 04. 05,

US 2004/0138882 A1, 2004. 07. 15,

CN 1397929 A, 2003. 02. 19,

审查员 张飞弦

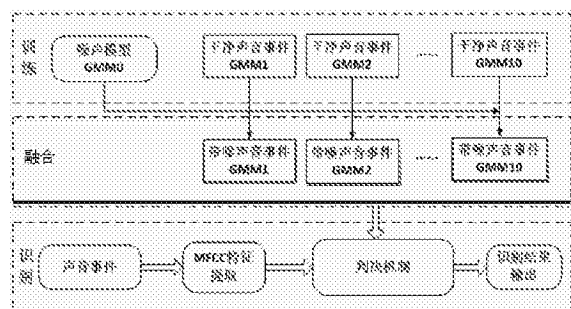
权利要求书2页 说明书6页 附图2页

(54) 发明名称

一种基于改进的并行模型组合的声音事件识别方法

(57) 摘要

本发明涉及一种基于改进的并行模型组合的声音事件识别方法,其步骤包括:1) 录制声音事件的数据,根据干净声音事件训练得到GMM高斯混合模型,建立干净声音事件模板;2) 在室内真实有噪声的环境下获取当前环境中的噪声数据,根据噪声数据训练得到GMM,建立噪声模板;3) 对噪声模板和干净声音事件模板采用改进的并行模型融合的方法,得到带噪声声音事件模板;4) 采样得到带噪声声音事件样本信号,根据带噪声声音事件模板中的参数对样本信号进行声音识别。本发明通过建立能够更好描述背景噪声特征分布的GMM来作为PMC方法中一个输入,建立5种声音事件的干净GMM作为PMC的另一个输入。同时本发明保证了识别系统对噪声的鲁棒性。



1. 一种基于改进的并行模型组合的声音事件识别方法,其步骤包括:

- 1) 根据干净声音事件训练得到 GMM 高斯混合模型,建立干净声音事件模板;
- 2) 根据噪声数据训练得到 GMM 高斯混合模型,建立噪声模板;

3) 对所述噪声模板和所述干净声音事件模板采用并行模型融合的方法,得到带噪声声音事件模板;包括以下分步骤:

3-1) 采用离散余弦逆变换将任一模型参数由倒谱域映射到线性谱域,得到对数谱域模型的均值 $\mu^{\log} = C^{-1} \mu$ 和方差 $\Sigma^{\log} = C^{-1} \Sigma (C^{-1})^T$, 其中, C 为离散余弦变换矩阵, μ , Σ 分别为模型的倒谱域的均值和方差;

3-2) 将对数谱域模型中的对数谱域均值和方差通过指数函数变换到线性谱域, $\mu_i^{\text{lin}} = \exp(\mu_i^{\log} + \frac{\Sigma_{ii}^{\log}}{2})$ 为线性谱域的均值向量的第 i 个元素, $\Sigma_{ij}^{\text{lin}} = \mu_i^{\text{lin}} \mu_j^{\text{lin}} [\exp(\Sigma_{ij}^{\log}) - 1]$ 为线性谱域的协方差矩阵的第 i 行第 j 列元素;其中, μ_i^{\log} 为对数谱域的均值向量的第 i 个元素, Σ_{ij}^{\log} 为对数谱域的协方差矩阵的第 i 行第 j 列元素;

3-3) 采用改进的并行模型组合方法,将干净声音事件模型参数与噪声模型参数在线性谱域进行融合, $\mu_{yk}^{\text{lin}} = g \mu_{xk}^{\text{lin}} + (1-g) \sum_{k=1}^K w_{nk} \mu_{nk}^{\text{lin}}$ 为融合后的带噪声声音事件模型在线性谱

域的均值, $\Sigma_{yk}^{\text{lin}} = g^2 \Sigma_{xk}^{\text{lin}} + (1-g)^2 \sum_{k=1}^K w_{nk} \Sigma_{nk}^{\text{lin}}$ 为融合后的带噪声声音事件模型在线性谱域的方差,其中 μ_{xk}^{lin} 为干净声音事件模型经所述步骤 3-1)、3-2) 变换后的线性谱域的均值, Σ_{xk}^{lin} 为干净声音事件模型经所述步骤 3-1)、3-2) 变换后的线性谱域的方差, μ_{nk}^{lin} 为噪声模型经所述步骤 3-1)、3-2) 变换后的线性谱域的均值, Σ_{nk}^{lin} 为噪声模型经所述步骤 3-1)、3-2) 变换后的线性谱域的方差;

3-4) 将融合后的带噪声声音事件模型的线性谱域模型的均值和方差经过上述步骤 3-2) 的逆变换得到对数谱域参数,再经过上述步骤 3-1) 逆变换得到倒谱域的参数,得到带噪声声音事件模型的均值向量和方差;

4) 采样得到带噪声声音事件样本信号,根据所述带噪声声音事件模板中的参数对样本信号进行声音识别。

2. 如权利要求 1 所述的基于改进的并行模型组合的声音事件识别方法,其特征在于,建立干净声音事件的模板的方法如下:

1) 在无噪安静室内的环境下录制声音事件的数据,对录制的声音事件进行预滤波、模数变换 后再进行分帧、加窗处理;

2) 提取 MFCC 梅尔倒谱系数特征,训练出声音事件的 GMM 高斯混合模板。

3. 如权利要求 1 所述的基于改进的并行模型组合的声音事件识别方法,其特征在于,所述高斯混合模型采用 EM 算法训练并更新高斯模型的参数,训练得到的干净声音事件的 GMM 参数为 $\lambda_x = \{w_{xk}, \mu_{xk}, \Sigma_{xk}\}$, $k = 1, 2, \dots, M$, 其中, w_{xk} 表示干净声音事件模型的混合权重, μ_{xk} 表示干净声音事件模型的均值, Σ_{xk} 表示干净声音事件模型的方差, M 表示混合高斯的阶数。

4. 如权利要求 1 所述的基于改进的并行模型组合的声音事件识别方法,其特征在于,在室内真实有噪声的环境下获取当前环境中的噪声数据,建立所述噪声模板方法为:提取 MFCC 特征,建立噪声的 GMM 模板,得到噪声模板 GMM 参数为 $\lambda_n = \{w_{nk}, \mu_{nk}, \Sigma_{nk}\}$, $k = 1, 2, \dots, M$, 其中, w_{nk} 表示噪声模型的混合权重, μ_{nk} 表示噪声模型的均值, Σ_{nk} 表示噪声模型的方差, M 表示混合高斯的阶数。

5. 如权利要求 1 所述的基于改进的并行模型组合的声音事件识别方法,其特征在于,带噪声音事件模型的参数 $\lambda_y = \{w_{yk}, \mu_{yk}, \Sigma_{yk}\}$, $k = 1, 2, \dots, M$, 其中 w_{yk} , μ_{yk} , Σ_{yk} 分别表示噪声模板的混合权重, 均值和方差, 其中混合权重没有线性谱域, 带噪声音事件模型的混合权重 w_{yk} 即为干净声音事件模板的权重 w_{xk} , M 表示混合高斯的阶数。

6. 如权利要求 1-5 任意一项所述的基于改进的并行模型组合的声音事件识别方法,其特征在于,所述噪声数据采用 NoiseX-92 的 babble 噪声和 / 或室内环境下空调噪声。

7. 如权利要求 1 所述的基于改进的并行模型组合的声音事件识别方法,其特征在于,根据所述带噪声音事件模型中的参数对样本信号进行声音识别的方法如下:

1) 对所述样本信号进行预滤波、模数变换,再进行分帧、加窗处理后提取多维 MFCC 特征得到样本信号特征序列;

2) 将样本信号的特征向量序列与所述带噪声音事件模型进行匹配,计算匹配似然度,最大似然度的匹配模板为识别结果。

一种基于改进的并行模型组合的声音事件识别方法

技术领域

[0001] 本发明属于智能监控中音频信号处理领域,涉及室内环境中声音事件识别方法,具体涉及一种基于改进的并行模型组合的声音事件识别方法。

背景技术

[0002] 相对于人工智能领域中成熟的语音识别方法来讲,利用计算机来进行声音事件的识别是近几年在比较新研究方向。声音事件识别针对的是自然环境中发声的具有一定含义的或是可以反映人们行为的声音事件,进行自动的判断和归类。在家庭智能监控系统中,声音事件的识别可以帮助人们远程的监控家庭室内环境中发生的情况,并及时告知用户产生了什么样的事件,有利于用户及时进行处理。但是,真实的环境中是存在着复杂的噪声的,要想实现在真实环境下的有效的监控,对噪声的处理是必须的和紧迫的。

[0003] 首先,声音事件的识别属于一个模式识别的问题,类似于自动语音识别。基本的方法是信号处理和模式识别。现有的声音事件识别方法包括以下几个步骤:

[0004] (1) 声音事件信号的录入,预滤波、模数变换。先把录入的模拟声音信号进行预滤波,高通滤波抑制 50HZ 电源噪声信号;低通滤波滤除声音信号中频率分量超过采样频率一半的部分,防止混叠干扰。对模拟声音信号进行采样和量化得到数字信号。

[0005] (2) 分帧、加窗。声音信号和语音信号一样,都具有整体非平稳,局部平稳的短时平稳性,类似语音信号,可以认为声音信号在 10 ~ 30ms 内是平稳的,可以把声音信号按照 30ms 的长度进行分帧。分帧时利用窗函数进行信号的提取,窗函数它的选择(形状和长度)对短时分析参数的特性影响很大,常用的窗函数包括矩形窗、汉宁窗和汉明窗等。一般选用汉明窗,可以很好地反应语音信号的特性变化。

[0006] (3) 特征提取。不同的声音事件的特征是不相同的,要想区分不同的声音信号,就要对声音信号的特征进行数学描述。常用的声音事件识别的特征有时域特征:短时能量,短时过零率。频域特征:子带能量,小波时频特性。倒谱域特征:线性预测倒谱系数(LPCC),梅尔频率倒谱系数(MFCC)等。

[0007] (4) 识别。声音事件的识别方法也是采用类似于语音识别的算法。常用的声音事件识别的方法有基于支持向量机(SVM)的分类,基于混合高斯模型(GMM)聚类方法,隐马尔可夫模型(HMM)方法,贝叶斯分类算法。

[0008] 其次,对噪声的处理。上述所讲述的识别方法在实际环境中应用时,识别系统的性能会随着训练数据和测试数据的失配急剧恶化,而导致所述失配的原因就是环境噪声的影响。由噪声引起的训练和测试的不匹配可以从信号空间、特征空间和模型空间三个空间来分析。常用的方法有类似于语音增强的声音增强法、鲁棒特征提取、特征补偿、模型补偿如并行模型组合(PMC)等方法对噪声进行处理,提高系统的鲁棒性。

[0009] 现有的方法大部分还是沿用语音识别的一套,对噪声的处理也不外乎以上几种方法,以上方法中基于 PMC 的方法能够充分描述环境噪声而被广泛采用,他们可以充分挖掘环境中的信息,提高系统识别的鲁棒性,但现有的 PMC 方法中对于噪声特征是用单高斯模

型 (SGM) 来描述的,对于噪声比较复杂的情况,SGM 并不能很好表征噪声的特性。所以,噪声复杂的情况下识别率不够理想。

发明内容

[0010] 为了解决上述技术问题,本发明的目的在于提供一种通过改进的模型参数融合的方法得到符合噪声环境的带噪声事件模型,对于实际噪声环境下的待识别声音事件进行识别。

[0011] 为了实现上述的目的,本发明技术方案为:一种基于改进的并行模型组合的声音事件识别方法,其步骤包括:

[0012] 1) 根据干净声音事件训练得到 GMM 高斯混合模型,建立干净声音事件模板;

[0013] 2) 根据噪声数据训练得到 GMM 高斯混合模型,建立噪声模板;

[0014] 3) 对所述噪声模板和所述干净声音事件模板采用并行模型融合的方法,得到带噪声事件模板;

[0015] 4) 采样得到带噪声事件样本信号,根据所述带噪声事件模板中的参数对样本信号进行声音识别。

[0016] 更进一步,建立干净声音事件的模板的方法如下:

[0017] 1) 在无噪安静室内的环境下录制声音事件的数据,对录制的声音事件进行预滤波、模数变换后再进行分帧、加窗处理;

[0018] 2) 提取 MFCC 梅尔倒谱系数特征,训练出声音事件的 GMM 高斯混合模板。

[0019] 更进一步,所述高斯混合模型采用 EM 算法训练并更新高斯模型的参数,训练得到的干净声音事件的 GMM 参数为 $\lambda_x = \{w_{xk}, \mu_{xk}, \Sigma_{xk}\}$, $k = 1, 2, \dots, M$, 其中, w_{xk} 表示干净声音事件模型的混合权重, μ_{xk} 表示干净声音事件模型的均值, Σ_{xk} 表示干净声音事件模型的方差, M 表示混合高斯的阶数。

[0020] 更进一步,在室内真实有噪声的环境下获取当前环境中的噪声数据,建立所述噪声模板方法为:提取 MFCC 特征,建立噪声的 GMM 模板,得到噪声模板 GMM 参数为 $\lambda_n = \{w_{nk}, \mu_{nk}, \Sigma_{nk}\}$, $k = 1, 2, \dots, M$, 其中, w_{nk} 表示噪声模型的混合权重, μ_{nk} 表示噪声模型的均值, Σ_{nk} 表示噪声模型的方差, M 表示混合高斯的阶数。

[0021] 更进一步,对所述噪声模板和所述干净声音事件模板采用改进的并行模型融合的方法如下:

[0022] (1) 采用离散余弦逆变换将任一模型参数由倒谱域映射到线性谱域,得到对数谱域模型的均值 $\mu^{\log} = C^{-1} \mu$, 和方差 $\Sigma^{\log} = C^{-1} \Sigma (C^{-1})^T$, 其中, C 为离散余弦变换矩阵, μ , Σ 分别为模型的倒谱域的均值和方差;

[0023] (2) 将对数谱域模型中的对数谱域均值和方差通过指数函数变换到线性谱域,

$\mu_i^{\text{lin}} = \exp(\mu_i^{\log} + \frac{\Sigma_{ii}^{\log}}{2})$ 为线性谱域的均值向量的第 i 个元素, $\Sigma_{ij}^{\text{lin}} = \mu_i^{\text{lin}} \mu_j^{\text{lin}} [\exp(\Sigma_{ij}^{\log}) - 1]$ 为线性谱域的协方差矩阵的第 i 行第 j 列元素。其中, μ_i^{\log} 为对数谱域的均值向量的第 i 个元素, Σ_{ij}^{\log} 为对数谱域的协方差矩阵的第 i 行第 j 列元素;

[0024] (3) 采用改进的并行模型组合方法,将干净声音事件模型参数与噪声模型参数在

线性谱域进行融合, $\mu_{yk}^{lin} = g\mu_{sk}^{lin} + (1-g)\sum_{k=1}^K w_{nk} \mu_{nk}^{lin}$ 为融合后的带噪声事件模型在线性谱域的均值, $\Sigma_{yk}^{lin} = g^2\Sigma_{sk}^{lin} + (1-g)^2\sum_{k=1}^K w_{nk} \Sigma_{nk}^{lin}$ 为融合后的带噪声事件模型在线性谱域的方差。其中 μ_{sk}^{lin} 为干净声音事件模型经所述步骤 (1) (2) 变换后的线性谱域的均值, Σ_{sk}^{lin} 为干净声音事件模型经所述步骤 (1) (2) 变换后的线性谱域的方差, μ_{nk}^{lin} 为噪声模型经所述步骤 (1) (2) 变换后的线性谱域的均值, Σ_{nk}^{lin} 为噪声模型经所述步骤 (1) (2) 变换后的线性谱域的方差;

[0025] (4) 将融合后的带噪声事件模型的线性谱域模型的均值和方差经过上述步骤 (2) 的逆变换得到对数谱域参数, 再经过上述步骤 (1) 逆变换得到倒谱域的特征参数, 得到带噪声事件模型的均值向量和方差。

[0026] 更进一步, 带噪声事件模型的参数为 $\lambda_y = \{w_{yk}, \mu_{yk}, \Sigma_{yk}\}$, $k = 1, 2, \dots, M$, 其中 w_{yk} , μ_{yk} , Σ_{yk} 分别表示噪声模板的混合权重, 均值和方差。其中混合权重没有线性谱域, 对数谱域以及倒谱域的差别。因此带噪声事件模型的混合权重 w_{yk} 即为干净声音事件模板的权重 w_{sk} , M 表示混合高斯的阶数。

[0027] 更进一步, 根据所述带噪声事件模型中的参数对样本信号进行声音识别的方法如下:

[0028] 1) 对所述样本信号进行预滤波、模数变换, 再进行分帧、加窗处理后提取多维 MFCC 特征得到样本信号特征序列;

[0029] 2) 将样本信号的特征向量序列与所述带噪声事件模型进行匹配, 计算匹配似然度, 最大似然度的匹配模板为识别结果。

[0030] 更进一步, 所述噪声数据采用 NoiseX-92 的 babble 噪声和 / 或室内环境下空调噪声。

[0031] 本发明的技术效果:

[0032] 本发明能够在复杂噪声背景下, 通过建立能够更好描述背景噪声特征分布的背景 GMM, 来作为 PMC 方法中一个输入, 建立 5 种声音事件的干净 GMM 作为 PMC 的另一个输入。经过改进的模型参数融合的方法得到符合噪声环境的带噪声事件模型, 对于实际噪声环境下的待识别声音事件, 有着良好的识别效果。本发明保证了识别系统对噪声的鲁棒性。

附图说明

[0033] 图 1 是本发明基于改进的并行模型组合的声音事件识别方法整体识别流程示意图。

[0034] 图 2 是本发明基于改进的并行模型组合的声音事件识别方法一实施例中融合方法示意图。

[0035] 图 3 是本发明基于改进的并行模型组合的声音事件识别方法一实施例中 5 种声音事件识别效果示意图。

[0036] 具体实施方法

[0037] 下面将结合本发明实施例中的附图, 对本发明实施例中的技术方案进行清楚、完

整地描述,可以理解的是,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0038] 本发明针对室内环境中经常发生的需要引起人们注意的 5 种声音事件进行识别。另外,充分考虑复杂噪声的情况(室内环境下录制的空调噪声、公共噪声数据库 NoiseX-92 的 babble 噪声),利用 GMM 混合高斯模型(GMM,《语音信号处理》第 2 版,赵力编著,机械工业出版社,228-230 页)来描述背景噪声信号,GMM 用多个高斯分布加权来描述背景的特征分布,可以更好的充分的描述背景噪声的信息。在模型层对利用所述的背景噪声模型参数,对干净声音事件模型参数进行补偿,得到带噪声事件的模型,以防止由于噪声带来的训练数据和测试数据的失配。

[0039] 本发明为一种基于改进的并行模型组合的声音事件识别方法,具体内容为:

[0040] 第一,建立干净声音事件的模板。

[0041] (1) 在安静环境下分别录制 5 种声音事件的数据,按照如前所述的语音信号处理步骤对加窗分帧等预处理。

[0042] (2) 再按如前所述,提取鲁棒的 MFCC 特征,分别训练出 5 种声音事件的高斯混合模型。高斯混合模型的训练采用 EM 算法更新高斯模型的参数。假设训练得到的干净声音事件的其中一种的 GMM 参数如下:

$$[0043] \quad \lambda_x = \{w_{xk}, \mu_{xk}, \Sigma_{xk}\}, k = 1, 2, \dots, M \quad (1)$$

[0044] 第二,获取当前环境中的噪声数据,提取 MFCC 特征,建立噪声的 GMM 模板。得到噪声模板参数如下:

$$[0045] \quad \lambda_n = \{w_{nk}, \mu_{nk}, \Sigma_{nk}\}, k = 1, 2, \dots, M \quad (2)$$

[0046] 第三,进行模型融合,由于本发明中用来训练 GMM 的数据都为 MFCC 特征,属于倒谱域特征,而背景噪声和声音事件模型参数只在线性谱域是可加的,所以要对以上两种模型同样做以下处理:(用 $\lambda = \{w_k, \mu_k, \Sigma_k\}, k = 1, 2, \dots, M$ 来统一表述干净声音的 GMM 和背景噪声的 GMM)

[0047] 1) 将模型参数由倒谱域映射到线性谱域,具体可以采用离散余弦变换的逆变换进行处理。这里我们不提取 MFCC 的差分系数,计算方法公式 (3) 和 (4):

$$[0048] \quad \mu^{\log} = C^{-1} \mu \quad (3)$$

$$[0049] \quad \Sigma^{\log} = C^{-1} \Sigma (C^{-1})^T \quad (4)$$

[0050] 其中 $\mu^{\log}, \Sigma^{\log}$ 为对数谱域模型的均值和方差, μ, Σ 为模型的倒谱域的均值和方差, C 为离散余弦变换矩阵。

[0051] 2) 将对数谱域的正态分布随机变量通过指数函数变换到线性谱域,计算方法如公式 (5) 和 (6):

$$[0052] \quad \mu_i^{\text{lin}} = \exp\left(\mu_i^{\log} + \frac{\sum_{jj}^{\log}}{2}\right) \quad (5)$$

$$[0053] \quad \Sigma_{ij}^{\text{lin}} = \mu_i^{\text{lin}} \mu_j^{\text{lin}} [\exp(\Sigma_{ij}^{\log}) - 1] \quad (6)$$

[0054] 其中, $\mu_i^{\text{lin}}, \Sigma_{ij}^{\text{lin}}$ 分别为线性谱域的均值向量的第 i 个元素和协方差矩阵的第 i 行

第 j 列元素； μ_i^{\log} , Σ_{ij}^{\log} 分别为对数谱域的均值向量的第 i 个元素和协方差矩阵的第 i 行第 j 列元素。

[0055] 3) 假设由上式计算后的干净声音事件模型的线性谱域的均值向量和方差分别为 μ_{sk}^{lin} , Σ_{sk}^{lin} ；噪声模型的线性谱域均值向量和方差分别为 μ_{nk}^{lin} , Σ_{nk}^{lin} ；采用公式 (7) 和 (8) 对两个模型进行融合：

$$[0056] \quad \mu_{yk}^{lin} = g\mu_{sk}^{lin} + (1-g)\sum_{k=1}^K w_{nk}\mu_{nk}^{lin} \quad (7)$$

$$[0057] \quad \Sigma_{yk}^{lin} = g^2\Sigma_{sk}^{lin} + (1-g)^2\sum_{k=1}^K w_{nk}\Sigma_{nk}^{lin} \quad (8)$$

[0058] 其中 μ_{yk}^{lin} , Σ_{yk}^{lin} 表示融合后的带噪声事件模型的均值向量和方差， g 表示增益因子。

[0059] 4) 将融合后的线性谱域模型参数经过公式 (5) (6) 的逆变换得到模型的对数谱域参数，再经过公式 (3) (4) 的逆变换得到模型的倒谱域的参数。将 5 种声音事件模型均做上述处理即可最后得到的即为融合后的 5 种带噪声事件模型参数。

[0060] 第四，对于在室内噪声环境声音中提取的带噪声事件信号样本，识别的目的是要确定当前样本属于 5 种声音事件的哪一个，即计算样本对 5 种模型的后验概率，其中最大的一个后验概率对应的模型即为样本的类别。根据贝叶斯公式，由于 5 种声音事件可能发生的概率相同，对于确定的观察矢量而言，上述最大后验概率的计算等同于计算改样本在 5 种声音事件模型下的概率，使得该概率最大的模型即为样本所属的类别。

[0061] 如图 1 所示是本发明基于改进的并行模型组合的声音事件识别方法整体识别流程图示意图，包括训练部分和识别部分。

[0062] 本发明考虑室内环境下经常发生并需要引起人们注意的 5 种声音事件，分别为关门声、敲门声、拍手声、说话声、尖叫声。5 种声音事件模板及噪声训练模板训练过程如下：

[0063] 1、在安静环境下录制 5 种声音事件数据库并进行标定。每种声音事件类型 100 个，由 5 男 5 女分别发声或产生动作得到。噪声采用 NoiseX-92 的 babble 噪声和室内环境下空调噪声。

[0064] 2、预滤波，高通滤波抑制 50Hz 电源噪声信号；低通滤波滤除声音信号中频率分量超过采样频率一半的部分；模数变换，采样频率为 11025Hz，采样精度为 16bits；

[0065] 3、对于每个完整的语音段，分帧、加窗。帧长为 256 采样点，帧移为 128 采样点。窗函数选取汉明窗；

[0066] 4、特征提取。提取 13 维 MFCC 特征；

[0067] 5、每种声音事件分别利用 60 个特征向量序列，噪声采用 10 个特征向量序列，基于期望最大化 (EM) 算法训练 5 种声音的 GMM 模板 λ_{sk} , $k = 1, 2, \dots, 5$ ，以及噪声的模板 λ_n ，模板采用 8 个高斯分量的高斯混合模型。

[0068] 本发明的模型融合过程如图 2 所示是本发明基于改进的并行模型组合的声音事件识别方法一实施例中融合方法示意图。

[0069] 具体步骤如下：

[0070] 1、采用所述公式 (3) (4) (5) (6) 将背景噪声模型和十个干净声音事件模型参数谱域变换至线性谱域。

[0071] 2、采用所述公式 (7) (8) 分别将十种干净声音事件的的线性谱域参数与噪声的线性谱域参数进行融合,这里 $g = 0.5$ 。

[0072] 3、将融合后的带噪声事件模型的线性谱域参数分别经 (5) (6) 公式的逆变换和 (3) (4) 的逆变换,得到带噪声事件的 5 个 GMM 模型 $\lambda_{y_k}, k = 1, 2, \dots, 5$ 。

[0073] 本发明的识别过程如下:

[0074] 1、在上述两种噪声条件下录制 5 种带噪声事件信号共 110 个。进行预滤波;模数变换,采样频率为 11025Hz,采样精度为 16bits。

[0075] 2、分帧、加窗。帧长为 256 采样点,帧移为 128 采样点。窗函数选取汉明窗。提取 13 维 MFCC 特征。

[0076] 3、模板匹配。当前声音信号的特征向量序列与 5 种带噪声事件模板进行匹配。特征矢量序列为 $X_k, k = 1, \dots, N$, 5 个模板为 $\lambda_{y_k}, k = 1, 2, \dots, 5$ 。计算匹配似然度,选择获得最大似然度的模板为识别结果。如图 3 所示是本发明基于改进的并行模型组合的声音事件识别方法一实施例中 5 种声音事件识别效果示意图。

[0077] 上述实例只是本发明的举例,尽管为说明目的公开了本发明的实例,但是本领域的技术人员可以理解:在不脱离本发明及所附的权利要求的精神和范围内,各种替换、变化和修改都是可能的。因此,本发明不应局限于该实例的内容。

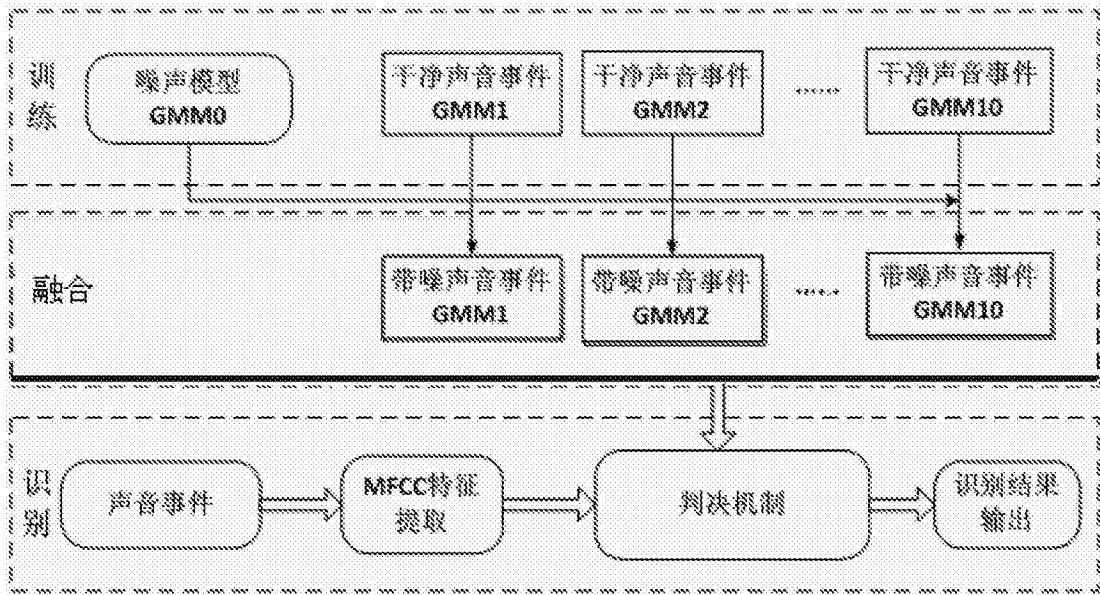


图 1

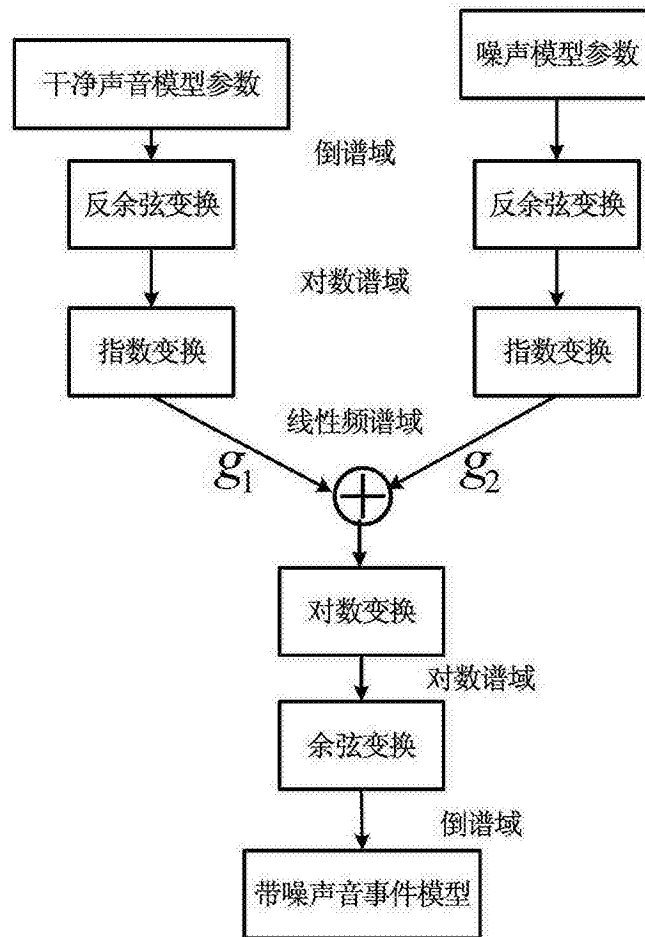


图 2

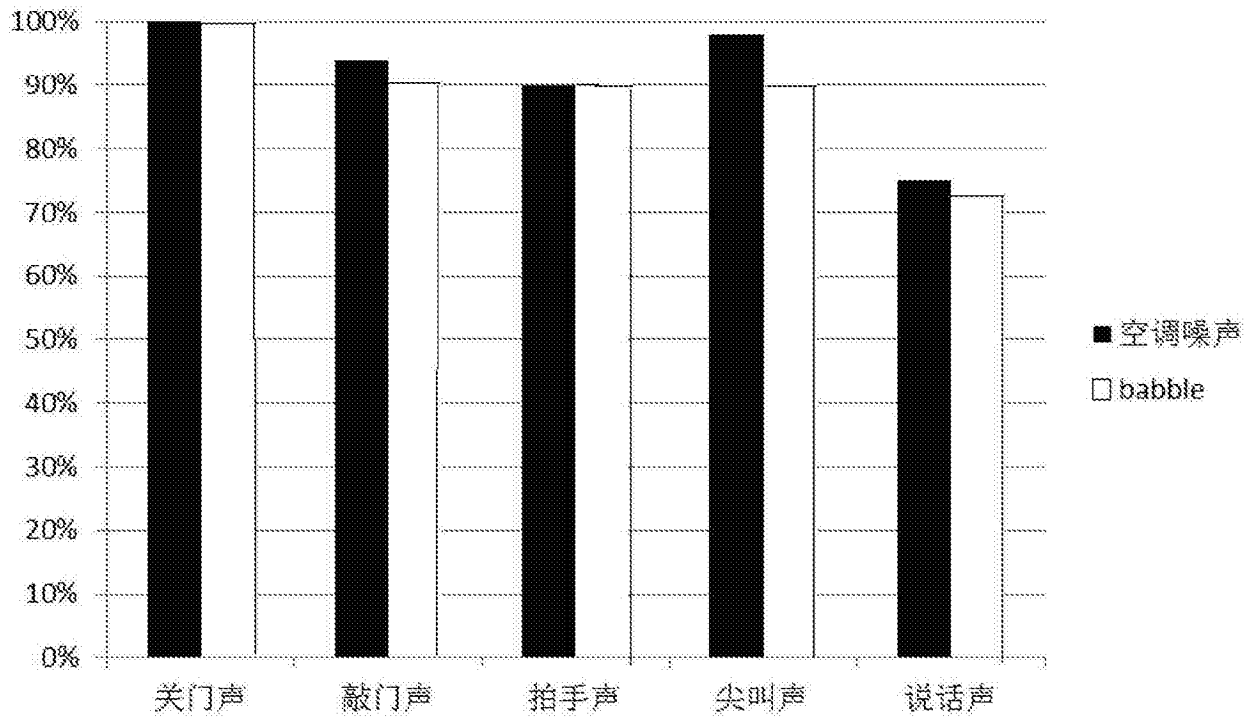


图 3