



(12) 发明专利

(10) 授权公告号 CN 109146002 B

(45) 授权公告日 2021.06.01

(21) 申请号 201811153290.8

(22) 申请日 2018.09.30

(65) 同一申请的已公布的文献号
申请公布号 CN 109146002 A

(43) 申请公布日 2019.01.04

(73) 专利权人 佛山科学技术学院
地址 528000 广东省佛山市南海区狮山镇
仙溪水库西路佛山科学技术学院

(72) 发明人 周伟力

(74) 专利代理机构 广州嘉权专利商标事务所有
限公司 44205

代理人 王国标

(51) Int. Cl.
G06K 9/62 (2006.01)

(56) 对比文件

CN 108269573 A, 2018.07.10
US 2011262033 A1, 2011.10.27
CN 102324232 A, 2012.01.18
CN 102509547 A, 2012.06.20
CN 102800316 A, 2012.11.28
CN 103794219 A, 2014.05.14

林子明. 基于GMM和VQ的说话人识别系统的研究.《中国优秀硕士学位论文全文数据库 信息科技辑》.2012,

审查员 王菲

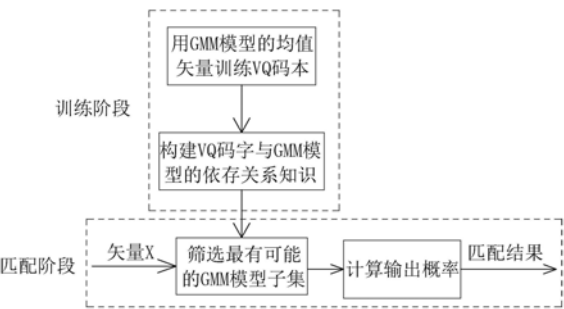
权利要求书1页 说明书6页 附图3页

(54) 发明名称

一种GMM识别器的快速识别方法

(57) 摘要

本发明公开了一种GMM识别器的快速识别方法,包括以GMM模型中的均值矢量为VQ码本的训练数据,构建VQ码本;构建VQ码字与GMM模型的依存关系知识;利用VQ码字与GMM模型的依存关系知识筛选出待识别的GMM模型,筛选出的GMM模型组成GMM模型子集;计算待识别的矢量X与GMM模型子集中各个GMM模型的输出概率,获得识别结果。本发明通过GMM模型中的Gaussian函数的均值矢量构建信号特征空间的描述,相对于现有技术,通过构建基于信号特征空间的VQ码本,依照VQ码字与Gaussian函数的归属关系筛选获得部分GMM模型有更大的概率属于测试样本的GMM模型,提高入选GMM模型为测试样本模型的准确度;同时本技术方案只遍历最有可能的部分GMM模型的模型分量,降低似然度的计算量,减少识别时间。



1. 一种GMM识别器的快速识别方法,其特征在于,包括以下步骤:

步骤1,以待识别的GMM模型中的Gaussian函数中的均值矢量为VQ码本的训练数据,采用无监督聚类方法构建VQ码本;

步骤2,构建VQ码本中的VQ码字与GMM模型的依存关系知识;

步骤3,待识别的矢量X首先依据邻近原则选择N个VQ码字,利用VQ码字与GMM模型的依存关系知识筛选出待识别的GMM模型,筛选出的GMM模型组成GMM模型子集;

步骤4,计算待识别的矢量X与GMM模型子集中各个GMM模型的输出概率,获得待识别矢量X的GMM识别器的识别结果。

2. 根据权利要求1所述的一种GMM识别器的快速识别方法,其特征在于:步骤1中,各个所述GMM模型的Gaussian函数中的均值矢量共同构成均值矢量集 μ_n^m ,其中n为GMM模型编号,m为GMM模型的Gaussian分量索引;按照Gaussian函数空间邻近聚类原则,采用K均值聚类算法对均值矢量集 μ_n^m 构建K个VQ码字的VQ码本 $W = \{w_i, i=1, 2, \dots, K\}$,其中 w_i 表示第i个VQ码字。

3. 根据权利要求2所述的一种GMM识别器的快速识别方法,其特征在于:步骤2中VQ码本中的VQ码字 w_i 与GMM模型的依存关系知识由该VQ码字 w_i 包含的Gaussian模型分量 $d_l^{(i,j)}$ 、Gaussian分量权重和 C_j^i 以及GMM模型编号 G_j^i 构成。

4. 根据权利要求3所述的一种GMM识别器的快速识别方法,其特征在于:所述Gaussian模型分量 $d_l^{(i,j)}$ 按与VQ码字距离从小到大存储;Gaussian分量权重和 C_j^i 按从大到小存储;GMM模型编号 G_j^i 按Gaussian分量权重和从大到小存储。

5. 根据权利要求4所述的一种GMM识别器的快速识别方法,其特征在于:步骤3中,采用欧氏距离计算待识别的矢量X与所有VQ码字的距离序列,确定矢量X在VQ码本构建的信号特征空间的位置;距离序列从小到大存储,选取距离序列中最小的前N个VQ码字,N个VQ码字组成矢量X的近邻码字子集。

6. 根据权利要求5所述的一种GMM识别器的快速识别方法,其特征在于:步骤3中,利用近邻码字子集中的VQ码字与GMM模型的依存关系知识,计算待识别矢量X的近邻码字子集中GMM模型的Gaussian分量权重和,选取Gaussian分量权重和最大的前M个GMM模型,所选的M个GMM模型组成GMM模型子集。

7. 根据权利要求6所述的一种GMM识别器的快速识别方法,其特征在于:步骤4中,对待识别矢量X与GMM模型子集中各个GMM模型的Gaussian模型分量进行输出概率计算,输出概率最大的GMM模型作为待识别矢量X的GMM识别器的识别结果。

8. 根据权利要求6所述的一种GMM识别器的快速识别方法,其特征在于:步骤1中VQ码字数K、步骤3中近邻码字子集的VQ码字数N、以及GMM模型子集中GMM模型个数M均可自由配置。

一种GMM识别器的快速识别方法

技术领域

[0001] 本发明涉及智能识别技术领域,更具体地说涉及一种GMM(高斯混合模型)识别器快速识别方法。

背景技术

[0002] GMM(高斯混合模型)采用多个高斯概率密度函数(正态分布曲线)精确地量化事物,即将一个事物分解为若干的基于高斯概率密度函数形成的数学模型。目前基于GMM模型的识别分类器广泛应用于智能模式识别,并且已在说话人识别(Speaker Recognition)、音频事件识别(Acoustic Event Recognition)、语音识别(Speech Recognition)、语种识别(Language Recognition)、手写字识别(Handwriting Recognition)等应用领域成为了主流的识别分类方法。

[0003] 随着技术发展和应用需求的不断增加,智能模式识别系统需要建立的识别分类对象规模也在不断增加。如一个功能强大的说话人识别系统,手写字识别系统等需要建立上千个以上的说话人和手写体GMM模型,因此降低测试样本分类识别所需的计算时间已成为实时模式识别系统最需要解决的问题之一。识别所需的计算时间取决于测试样本与GMM模型的似然度计算过程。对于实时的分类识别,系统必须快速对未知的测试样本进行GMM的模型识别。

[0004] 传统的GMM识别器识别方法(例如公开号为CN103914811A的发明专利基于GMM模型进行SAR图像检索分类,公开号为CN102324232A的发明专利基于GMM模型进行声纹识别)在训练阶段对已知对象(如已知说话人,已知音频事件等)进行GMM模型训练,测试阶段通过提取测试样本的特征参数,计算特征参数与GMM模型集合的似然度值,最后最大的似然值对应的GMM模型为测试样本的识别结果。传统的GMM识别器在训练对象较少的情况下(如需要训练的说话人数量较少)能够取得较好的性能,但是在训练对象规模较大的情况下,实时性能就不能得到保证,究其原因在于测试样本与GMM模型的似然度计算实质是测试样本与GMM模型的多个高斯概率密度函数的似然度计算,当训练对象较多的情况下,需要识别的GMM模型越多,因此遍历识别的时间越长。

[0005] 针对上述问题,目前GMM识别器主流快速识别方法主要基于只识别最有可能的GMM模型的思路,这类方法(如Vijendra et al.2009;Sarkar et al.2010)通过一些快速方法(例如最大似然度线性回归,Maximum Likelihood Linear Regression,MLLR)找出与测试样本最可能的GMM模型集合,在计算似然度的过程中只遍历这些最有可能的GMM模型,以达到节省时间的效果。这类方法在识别测试样本时仍需要遍历GMM模型集合的所有模型分量,因此在训练对象规模较大的情况下,实时性就不如理想。GMM模型采用多个Gaussian(高斯)分量描述信号的特征空间,然而由于每个测试样本存在差异,每个Gaussian分量对于测试样本的贡献程度是不一样的,即存在某些Gaussian分量与测试样本的似然度值较大,某些Gaussian分量与测试样本的似然度值可以忽略不计。目前GMM识别器主流快速识别方法忽略了GMM模型该重要特性,因此算法效率和实时性能仍有较大的改善空间。

发明内容

[0006] 本发明要解决的技术问题是：如何提高GMM识别器对测试样本进行识别的实时性。

[0007] 本发明解决其技术问题的解决方案是：

[0008] 一种GMM识别器的快速识别方法，包括以下步骤：

[0009] 步骤1，以待识别的GMM模型中的Gaussian (高斯) 函数中的均值矢量为VQ码本的训练数据，采用无监督聚类方法构建VQ码本；

[0010] 步骤2，构建VQ码本中的VQ码字与GMM模型的依存关系知识；

[0011] 步骤3，待识别的矢量X首先依据邻近原则选择N个VQ码字，利用VQ码字与GMM模型的依存关系知识筛选出待识别的GMM模型，筛选出的GMM模型组成GMM模型子集；

[0012] 步骤4，计算待识别的矢量X与GMM模型子集中各个GMM模型的输出概率，获得待识别矢量X的GMM识别器的识别结果。

[0013] 作为上述技术方案的进一步改进，步骤1中，各个所述GMM模型的Gaussian函数中的均值矢量共同构成均值矢量集 μ_n^m ，其中n为GMM模型编号，m为GMM模型的Gaussian分量索引；按照Gaussian函数空间邻近聚类原则，采用K均值聚类算法对均值矢量集 μ_n^m 构建K个VQ码字的VQ码本 $W = \{w_i, i=1, 2, \dots, K\}$ ，其中 w_i 表示第i个VQ码字。

[0014] 作为上述技术方案的进一步改进，步骤2中VQ码本中的VQ码字 w_i 与GMM模型的依存关系知识由该VQ码字 w_i 包含的Gaussian模型分量 $d_l^{(i,j)}$ 、Gaussian分量权重和 C_j^i 以及GMM模型编号 G_j^i 构成。其中 $d_l^{(i,j)}$ 表示VQ码字 w_i 的第j个GMM模型的第l个Gaussian模型分量， C_j^i 表示VQ码字 w_i 的第j个GMM模型的Gaussian分量权重和， G_j^i 表示码字 w_i 的第j个GMM模型编号。

[0015] 作为上述技术方案的进一步改进，所述Gaussian模型分量 $d_l^{(i,j)}$ 按与VQ码字距离从小到大存储；Gaussian分量权重和 C_j^i 按照权重和从大到小存储；GMM模型编号 G_j^i 按Gaussian分量权重和从大到小存储。

[0016] 作为上述技术方案的进一步改进，步骤3中，采用欧氏距离计算待识别的矢量X与所有VQ码字的距离序列，确定矢量X在VQ码本构建的信号特征空间的位置；距离序列从小到大存储，选取距离序列中最小的前N个VQ码字，N个VQ码字组成矢量X的近邻码字子集。

[0017] 作为上述技术方案的进一步改进，步骤3中，利用近邻码字子集中的VQ码字与GMM模型的依存关系知识，计算待识别矢量X的近邻码字子集中GMM模型的Gaussian分量权重和，选取Gaussian分量权重和最大的前M个GMM模型，所选的M个GMM模型组成GMM模型子集。

[0018] 作为上述技术方案的进一步改进，步骤4中，对待识别矢量X与GMM模型子集中各个GMM模型的Gaussian模型分量进行输出概率计算，输出概率最大的GMM模型作为待识别矢量X的GMM识别器的识别结果。

[0019] 作为上述技术方案的进一步改进，步骤1中VQ码字个数K、步骤3中近邻码字子集的VQ码字个数N、以及GMM模型子集中GMM模型个数M均可自由配置。

[0020] 本发明的有益效果是：本发明通过GMM模型中的Gaussian函数的均值矢量构建信号特征空间的描述，相对于现有技术，通过构建基于信号特征空间的VQ码本，依照VQ码字与

Gaussian函数的归属关系筛选获得部分GMM模型有更大的概率属于测试样本的GMM模型,提高入选GMM模型为测试样本模型的准确度;同时本技术方案只遍历最有可能的部分GMM模型的模型分量,降低似然度的计算量,减少识别时间。

附图说明

[0021] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单说明。显然,所描述的附图只是本发明的一部分实施例,而不是全部实施例,本领域的技术人员在不付出创造性劳动的前提下,还可以根据这些附图获得其他设计方案和附图。

[0022] 图1是本发明的识别方法流程图;

[0023] 图2是本发明构建信号特征空间VQ码本的流程图;

[0024] 图3是VQ码字与GMM模型的依存关系知识的构建图;

[0025] 图4是本发明对测试样本(矢量X)进行识别的流程图。

具体实施方式

[0026] 以下将结合实施例和附图对本发明的构思、具体结构及产生的技术效果进行清楚、完整的描述,以充分地理解本发明的目的、特征和效果。显然,所描述的实施例只是本发明的一部分实施例,而不是全部实施例,基于本发明的实施例,本领域的技术人员在不付出创造性劳动的前提下所获得的其他实施例,均属于本发明保护的范围。

[0027] 参照图1~图4,本发明创造公开了一种GMM识别器的快速识别方法,包括以下步骤:

[0028] 步骤1,以待识别的GMM模型中的Gaussian(高斯)函数中的均值矢量为VQ码本的训练数据,采用无监督聚类方法构建VQ码本;

[0029] 步骤2,构建VQ码本中的VQ码字与GMM模型的依存关系知识;

[0030] 步骤3,待识别的矢量X首先依据邻近原则选择N个VQ码字,利用VQ码字与GMM模型的依存关系知识筛选出待识别的GMM模型,筛选出的GMM模型组成GMM模型子集;

[0031] 步骤4,计算待识别的矢量X与GMM模型子集中各个GMM模型的输出概率,获得待识别矢量X的GMM识别器的识别结果。

[0032] 具体地,本发明通过GMM模型中的Gaussian函数的均值矢量构建信号特征空间的描述,相对于现有技术,通过构建基于信号特征空间的VQ码本,依照VQ码字与Gaussian函数的归属关系筛选获得部分GMM模型有更大的概率属于测试样本的GMM模型,提高入选GMM模型为测试样本模型的准确度;同时本技术方案只遍历最有可能的部分GMM模型的模型分量,降低似然度的计算量,减少识别时间。

[0033] 步骤1中通过GMM模型的Gaussian函数中的均值矢量描述信号特征空间,即VQ码本的训练数据由全部GMM模型的均值矢量集 μ_n^m 构成,其中n为GMM模型编号,m为GMM模型的Gaussian分量索引;按照Gaussian函数空间邻近聚类原则,采用K均值聚类算法对均值矢量集 μ_n^m 构建K个VQ码字的VQ码本 $W = \{w_i, i=1, 2, \dots, K\}$,其中 w_i 表示第i个VQ码字,VQ码字的数量K不作为固定值进行设置,可按照计算复杂度和识别准确度的权衡需求进行配置。

[0034] 步骤2中VQ码本中的VQ码字 w_i 与GMM模型的依存关系知识由该VQ码字 w_i 包含的Gaussian模型分量 $d_l^{(i,j)}$ 、Gaussian分量权重和 C_j^i 以及GMM模型编号 G_j^i 构成。其中 $d_l^{(i,j)}$ 表示VQ码字 w_i 的第j个GMM模型的第l个Gaussian模型分量, C_j^i 表示VQ码字 w_i 的第j个GMM模型的Gaussian分量权重和, G_j^i 表示码字 w_i 的第j个GMM模型编号。所述Gaussian模型分量 $d_l^{(i,j)}$ 按与VQ码字距离从小到大存储; Gaussian分量权重和 C_j^i 按照权重和从大到小存储; GMM模型编号 G_j^i 按Gaussian分量权重和从大到小存储。

[0035] 步骤3中采用欧氏距离计算待识别的矢量X与所有VQ码字的距离序列, 确定矢量X在VQ码本构建的信号特征空间的位置; 距离序列从小到大存储, 选取距离序列中最小的前N个VQ码字, N个VQ码字组成矢量X的近邻码字子集; 利用近邻码字子集中的VQ码字与GMM模型的依存关系知识, 计算待识别矢量X的近邻码字子集中GMM模型的Gaussian分量权重和, 选取Gaussian分量权重和最大的前M个GMM模型, 所选的M个GMM模型组成GMM模型子集。近邻码字子集的VQ码字数N以及GMM模型子集中GMM模型个数M同样可按照计算复杂度和识别准确度的权衡需求进行配置。

[0036] 步骤4中对待识别矢量X与GMM模型子集中各个GMM模型的Gaussian模型分量进行输出概率计算, 输出概率最大的GMM模型作为待识别矢量X的GMM识别器的识别结果。

[0037] 本发明创造中, 步骤1中VQ码字数K、步骤3中近邻码字子集的VQ码字数N、以及GMM模型子集中GMM模型个数M均可依据计算复杂度和识别准确度的权衡需求自由配置, 从而可在不同需求的应用场景能够获得最优化的性能。

[0038] 本发明创造所述的GMM识别器快速识别方法具体过程如下:

[0039] 训练阶段:

[0040] 构建信号特征空间矢量量化VQ码本, 图2给出了构件信号特征空间VQ码本的过程, 首先分别对训练对象(如已知说话人, 已知音频事件)进行GMM建模, 获取每个GMM模型的均值矢量; 接着构建GMM模型的均值矢量集, 所述均值矢量集作为信号特征空间VQ码本的训练数据, 按照Gaussian函数空间邻近聚类原则, 采用K均值聚类算法对均值矢量集 μ_n^m 构建K个VQ码字的VQ码本 $W = \{w_i, i=1, 2, \dots, K\}$;

[0041] 构建VQ码本中的VQ码字与GMM模型的依存关系知识, 如图3所示, 首先统计VQ码字 w_i 包含的Gaussian模型分量以及Gaussian模型分量所属的GMM模型, 然后计算GMM模型的Gaussian分量权重和, Gaussian模型分量按与VQ码字距离从小到大存储, 即第一分量与VQ码字最靠近, Gaussian分量权重和按照从大到小顺序存储, GMM模型编号按Gaussian分量权重和从大到小存储, 即第一编号的GMM模型的Gaussian分量权重和最大;

[0042] 识别阶段:

[0043] 筛选待识别矢量X最有可能的GMM模型集, 图4中分出了相应的流程图。首先基于步骤1中所构建的VQ码本对待识别矢量X进行空间定位, 采用欧氏距离计算待识别矢量X与所有VQ码字的距离序列, 然后按照近邻原则选择距离最小的前N个VQ码字作为近邻码字子集; 利用近邻码字子集中的VQ码字与GMM模型的依存关系知识, 计算待识别矢量X的近邻码字子集中GMM模型分量权重和, 选取分量权重和最大的前M个GMM模型。

[0044] 计算待识别矢量X与GMM模型子集中各个GMM模型的Gaussian模型分量的输出概率,输出概率最大的GMM模型作为待识别矢量X的GMM识别器的识别结果。

[0045] 本发明创造的原理如下所示:应用GMM识别器中的Gaussian函数的均值矢量构建VQ码本,并利用VQ码本生成过程中VQ码字与Gaussian函数的归属关系构建VQ码字与GMM模型的依存关系;待识别矢量利用该依存关系筛选出最有可能的部分GMM模型及其Gaussian模型分量进行识别,从而避免遍历识别,达到提高识别速度的目的。

[0046] 以下结合一个与音频相关的具体例子进行说明。

[0047] 训练阶段:

[0048] 构建音频事件GMM模型的均值矢量码本。具体实现如下,假定共有18类待识别的音频事件类别,首先分别对每类音频事件进行GMM建模,GMM模型参数 $\lambda = \{q_n, \mu_n, \Sigma_n\}$, $n=1, 2, \dots, 18$,其中 q 、 μ 、 Σ 分别为GMM模型的权重、均值和方差。接着构建音频事件GMM模型的均值矢量集 μ_n^m 构成,其中 n 为GMM模型编号, m 为GMM模型的Gaussian分量索引(即混合数),每个音频事件GMM模型中混合数可取不同的值。采用Kmeans聚类算法对均值矢量集构建具有256个VQ码字的VQ码本 $W = \{w_i, i=1, 2, \dots, K\}$ 。

[0049] 构建256个VQ码字与音频事件GMM模型的依存关系知识。首先统计256个VQ码字中每个VQ码字包含的Gaussian模型分量以及Gaussian模型分量所属的GMM模型,然后计算GMM模型的Gaussian分量权重和,Gaussian模型分量按与VQ码字距离从小到大存储,即第一分量与VQ码字最靠近,Gaussian分量权重和按照从大到小顺序存储,GMM模型编号按Gaussian分量权重和从大到小存储,即第一编号的GMM模型的Gaussian分量权重和最大。

[0050] 识别阶段:

[0051] 筛选待识别矢量X最有可能的GMM模型集。首先提取音频事件测试特征,即矢量X,

采用欧氏距离计算矢量X与所有VQ码字的欧氏距离序列 dis , $dis_i = \sqrt{\sum_{t=1}^D (X^t - w_i^t)^2}$,其

中 D 表示矢量X的维度数, i 表示VQ码字索引,然后选择距离最小的前15个VQ码字,利用步骤2构建的VQ码字与音频事件GMM模型的依存关系知识,计算前15个VQ码字中每个GMM模型的Gaussian分量权重和,选择Gaussian分量权重和最大的前10个GMM模型,得到矢量X的最有可能的GMM模型子集 $\{G_1, G_2, \dots, G_{10}\}$ 。

[0052] 计算矢量X与GMM模型子集中各个GMM模型的Gaussian模型分量的输出概率:

$$[0053] \quad \log(p(X | \lambda_s)) = \log\left(\sum_{j=1}^r q_j p_j(X)\right)$$

[0054] 其中 λ_s 表示第 s 个入选的GMM模型参数, r 表示GMM模型入选的Gaussian模型分量数, q_j 表示入选的Gaussian分量权重和, $p_j(X)$ 表示入选的Gaussian模型分量的输出概率:

$$[0055] \quad p_j(X) = \frac{1}{2\pi^{D/2} |\Sigma_j|^{1/2}} \exp\left\{-\frac{1}{2}(X - \mu_j)'(\Sigma_j)^{-1}(X - \mu_j)\right\}$$

[0056] 最后输出概率最大的GMM模型作为待识别矢量X的GMM识别器的识别结果:

$$s = \arg \max_{1 \leq s \leq 10} (\log(p(X | \lambda_s)))。$$

[0057] 以上对本发明的较佳实施方式进行了具体说明,但本发明创造并不限于所述实施例,熟悉本领域的技术人员在不违背本发明精神的前提下还可作出种种的等同变型或替换,这些等同的变型或替换均包含在本申请权利要求所限定的范围内。

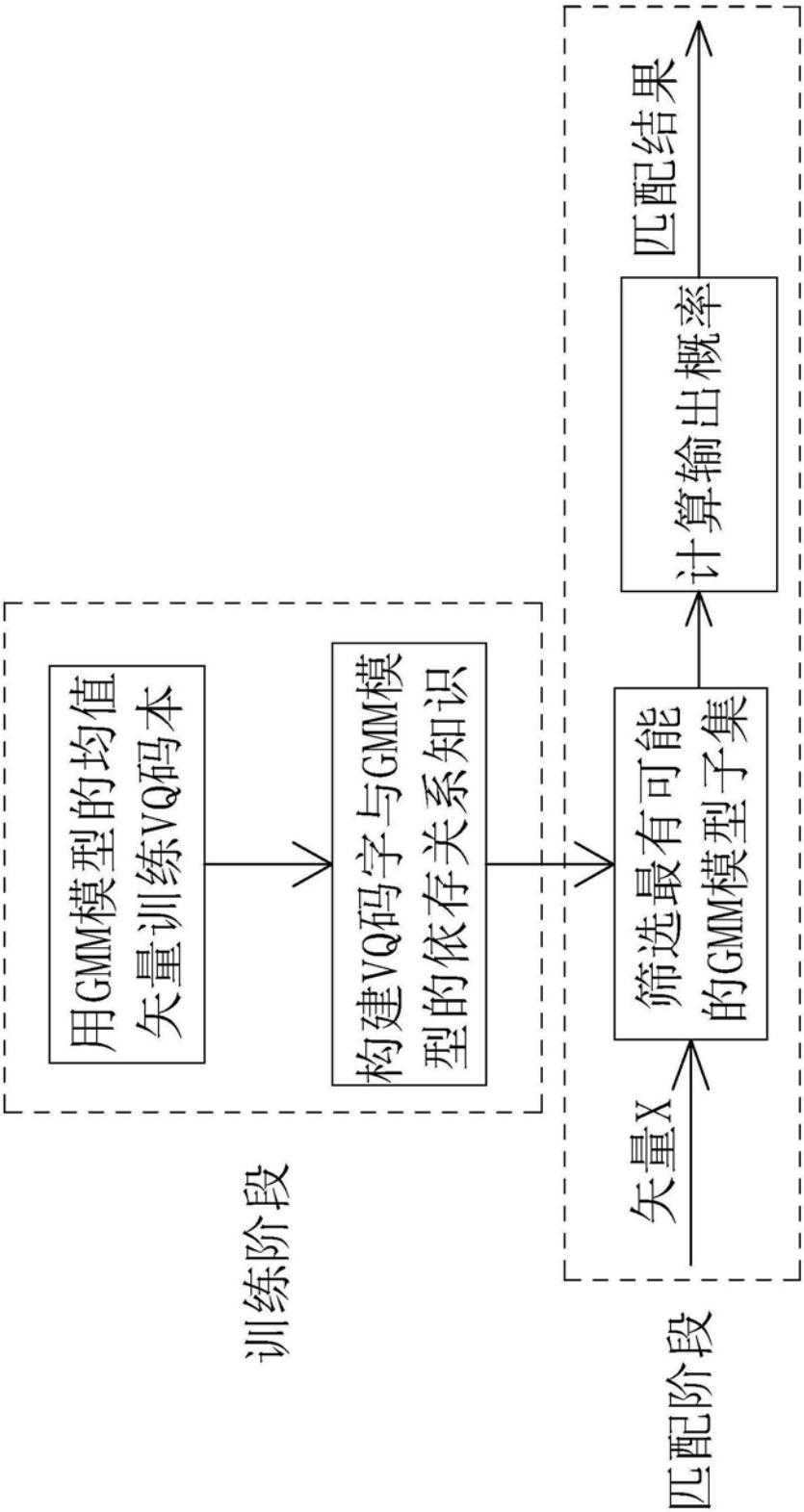


图1

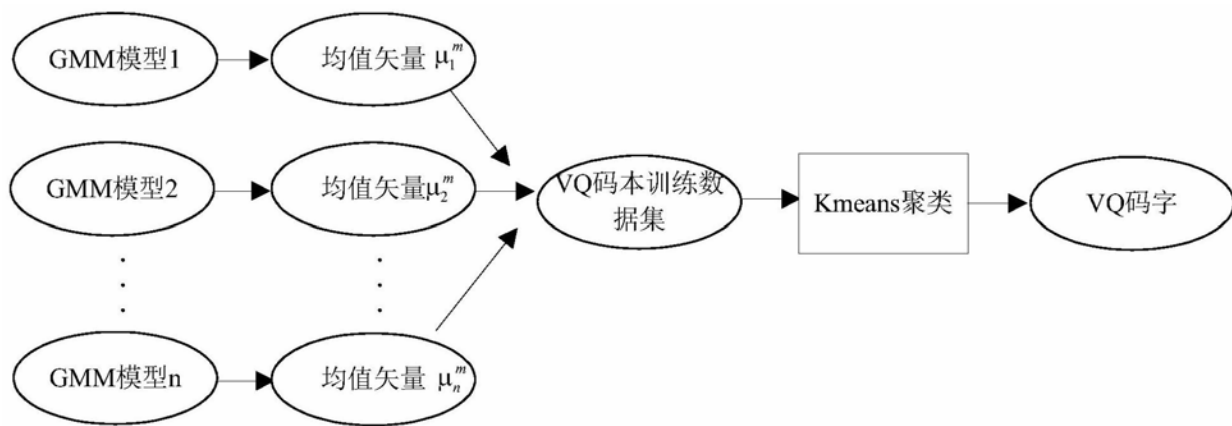


图2

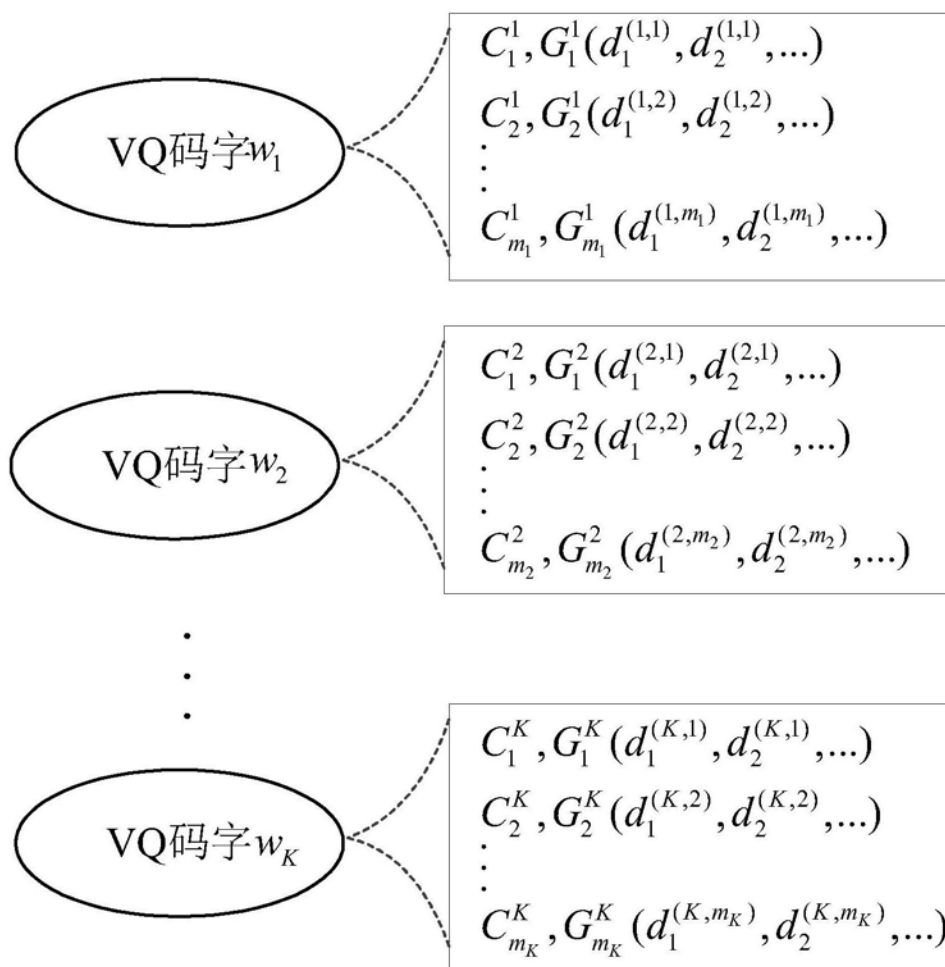


图3

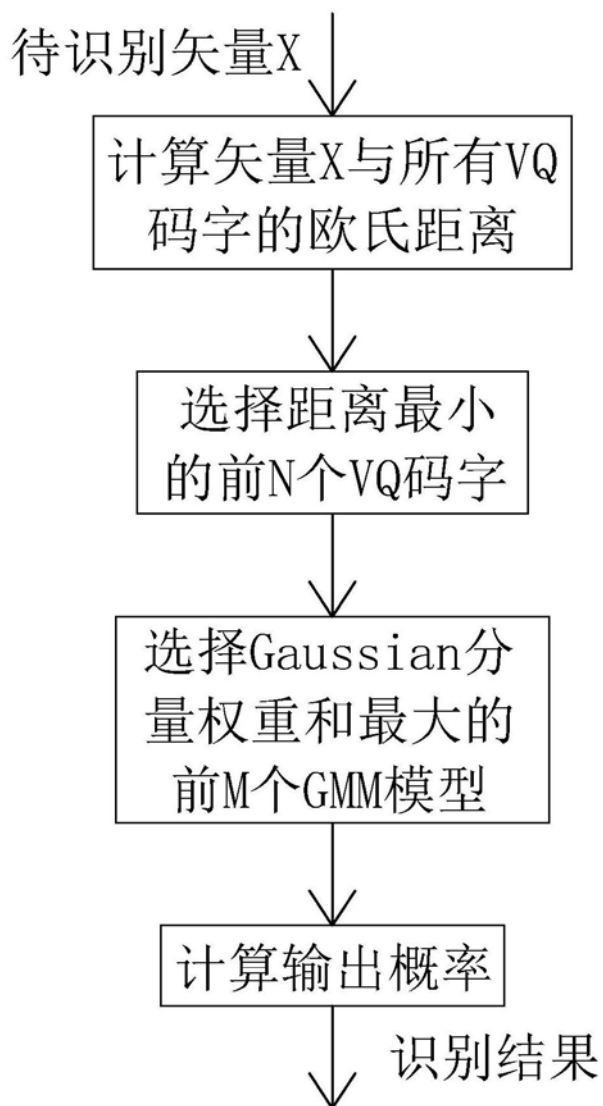


图4