



(12)发明专利申请

(10)申请公布号 CN 108597535 A
(43)申请公布日 2018.09.28

(21)申请号 201810268151.3

(22)申请日 2018.03.29

(71)申请人 华南理工大学

地址 510640 广东省广州市天河区五山路
381号

(72)发明人 韦岗 张楚翘 马碧云 曹燕

(74)专利代理机构 广州粤高专利商标代理有限
公司 44102

代理人 何淑珍 江裕强

(51)Int.Cl.

G10L 25/30(2013.01)

G10L 25/51(2013.01)

G10H 1/00(2006.01)

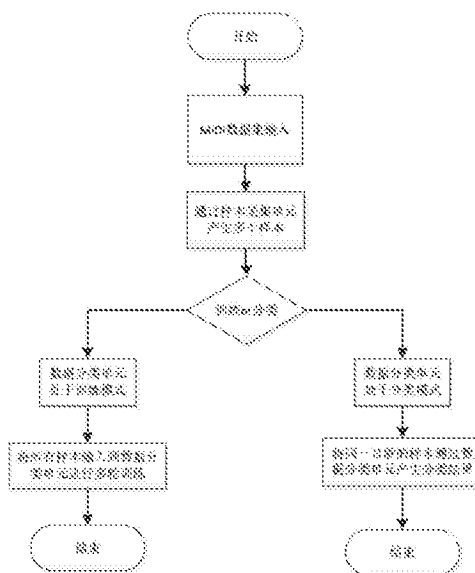
权利要求书1页 说明书4页 附图4页

(54)发明名称

一种融合伴奏的MIDI钢琴曲风格分类方法

(57)摘要

本发明公开一种融合伴奏的MIDI钢琴曲风格分类方法,采用样本采集单元以及数据分类单元实现,样本采集单元主要从一首钢琴曲中提取出含有主旋律和伴奏信息的音乐片段样本,其中每一个音乐片段样本由固定数量的音符向量组成,每个音符向量都涵盖了主旋律音符音高、该主旋律音符在该片段的起始时间、该主旋律音符的持续时间以及该主旋律音符可能存在的伴奏音符的音高信息,数据分类单元用于接收样本采集单元的样本,进行训练和分类。本发明实施的分类过程中,考虑到了伴奏对音乐风格的影响,并且通过样本采集单元增多了样本的数量,从而提高了分类的准确性。



1. 一种融合伴奏的MIDI钢琴曲风格分类方法,其特征不在于采用样本采集单元以及数据分类单元实现,样本采集单元主要从一首钢琴曲中提取出含有主旋律和伴奏信息的音乐片段样本,其中每一个音乐片段样本由固定数量的音符向量组成,每个音符向量都涵盖了主旋律音符音高、该主旋律音符在该片段的起始时间、该主旋律音符的持续时间以及该主旋律音符可能存在的伴奏音符的音高信息,数据分类单元用于接收样本采集单元的样本,进行训练和分类。

2. 根据权利要求1所述的一种融合伴奏的MIDI钢琴曲风格分类方法,其特征不在于所述的样本采集单元主要接收整个MIDI文件,产生若干个含有主旋律和伴奏信息的音乐片段样本,其中每一个音乐片段样本由固定数量的音符向量组成;每个音符向量都包含八个维度;第一个维度为主旋律音符的音高,对音高 p 进行归一化操作,作 $p=(p-21)/88$ 的归一化操作;第二个维度为该主旋律音符在该所在音乐片段的起始时间;第三个维度为该主旋律音符的持续时间;第四到第八维分别保存该主旋律音符的伴奏音符组里所有音符的音高信息,按从大到小排列,与第一个维度相同,也同样需要进行 $p=(p-21)/88$ 的归一化操作,如果该主旋律音符的伴奏音符不足5个,则多余的维度的值为0,如果超过5个,则舍弃音高较低的伴奏音符。

3. 根据权利要求1所述的一种融合伴奏的MIDI钢琴曲风格分类方法,其特征不在于所述的数据分类单元主要由深度神经网络构成,该深度神经网络用于接收样本采集单元产生的音乐片段样本,作为其输入特征,并输出风格标签的独热编码;数据分类单元主要有训练模式和分类模式两种工作模式,在训练模式下,它接收音乐片段集的输入,以音乐片段样本所在乐曲的风格标签作为样本的标签,以对深度神经网络进行训练;在分类模式下,它接收若干个音乐片段样本的输入,将这些样本依次输入至深度神经网络中得到多个风格标签,取出现次数最多的标签作为该乐曲的风格标签。

4. 根据权利要求1所述的一种融合伴奏的MIDI钢琴曲风格分类方法,其特征不在于所述训练包括如下步骤:

步骤1:获取训练用的钢琴曲MIDI文件训练数据集,其中每一个文件都已经打上风格标签;

步骤2:步骤1的文件训练数据集依次输入到样本采集单元,每个文件都会产生若干个含有主旋律和伴奏信息的音乐片段样本,给这些音符片段都打上与该文件相同的标签,得到样本数据集;

步骤3:将数据分类单元调至训练模式,取步骤2产生的样本数据集输入到数据分类单元进行多轮训练,以达到良好的分类效果。

5. 根据权利要求1所述的一种融合伴奏的MIDI钢琴曲风格分类方法,其特征不在于所述分类包括如下步骤:

步骤1:将一个未分类的MIDI文件通过样本采集单元产生若干个含有主旋律和伴奏信息的音乐片段样本;

步骤2:将数据分类单元调至分类模式,取步骤1产生的所有音乐片段样本输入至数据分类单元,产生对应的风格标签。

一种融合伴奏的MIDI钢琴曲风格分类方法

[0001]

技术领域

[0002] 本发明涉及数据处理技术领域,具体涉及一种融合伴奏的MIDI钢琴曲风格分类方法。

背景技术

[0003] 钢琴起源于18世纪初,是西洋古典音乐的一种键盘乐器,由88个琴键和金属弦音板构成。在经过莫扎特,勃拉姆斯,李斯特等优秀音乐家的推广下,逐渐由了乐器之王的美称。如今,它为音乐家的创作提供了条件,音乐家创作发展的钢琴文化又促进了钢琴结构功能的进一步成熟。

[0004] 钢琴曲是指用钢琴弹奏的乐曲,有时还配以其它的乐器同时演奏。随着互联网时代的到来,多媒体的应用领域越来越广,钢琴曲已经走进了我们的生活。小到音乐创作者的创作过程,大到影视作品的伴奏选取,都几乎离不开钢琴曲。如今,随着钢琴曲的进一步发展,对其质量、便携性、易操作性有了更高的要求。以往的钢琴曲制作都需要通过演奏者完完整整地演奏完,稍有失误都会导致创作的失败。随着乐器数字接口(MIDI, Musical Instrument Digital Interface)技术的诞生,钢琴曲的创作不再需要这么麻烦,除此之外,还大大减少了钢琴曲的存储空间。

[0005] MIDI是数字音乐国际的标准,它定义了计算机音乐程序,合成器,及其他电子设备交换信息和电子信号的方式,解决了不同电子乐器之间不兼容的问题。相比wav, mp3等格式的音乐,由于MIDI储存的不是波形,因此大大地减少文件的空间,一首可以播放5分钟左右的MIDI歌曲,其容量只有百余K字节。而同样这首歌的波形音乐文件如WAV,则高达50MB左右,即使是经过MP3技术进行高比例压缩处理,也有5MB大小。因此,MIDI便成为了钢琴曲的理想存储方式。

[0006] 音乐风格是听众区分乐曲的重要特征,也是各大音乐网站管理音乐的标签,以使用户检索音乐文件。其中也包括MIDI文件。但是,随着网上的MIDI文件资源越来越多,对海量的文件一一标注风格将耗费大量的人力和时间。而且当人工标注有误时,会影响用户检索的效率。但是,如果把这项工作交给计算机,不仅能省下人力,还能节省大量的时间。因此,机器学习成为了解决MIDI文件分类问题的新的思路。

[0007] 目前,有不少学者也在研究基于机器学习的MIDI的分类问题,对于特征提取方面,他们都先提取了整首MIDI文件中代表音乐主旋律音符组,抛弃掉乐曲里的伴奏音符,由于主旋律音符组里的音符非常多,因此他们其统计特征,比如音符密度,音符平均音高,音符音高方差等等以压缩输入数据特征的维度,并令这些特征作为输入分类器的参数。但是,这些特征都是根据主旋律提取出来的,并没有考虑到伴奏旋律,对于钢琴曲而言,伴奏中往往包含着和声,和声是音乐的血肉,决定歌曲的进行和构成,是决定音乐风格的又一项不可忽视的因素,因此,如果想达到更好的分类效果,乐曲的伴奏信息不应该丢弃。

[0008] 除此之外,这些统计特征不能完全表征原来的曲子的旋律,旋律一段包含音高,响度,时长信息的音符序列,音符出现的顺序对旋律的风格有着极大的影响。但是,上述研究的所使用的特征,无论是平均音高,还是音高方差,都没有反映出这一特点。

[0009] 在现实生活,一个懂音乐的业内人士往往仅凭一个音乐片段就能推断出整首乐曲的风格。这是因为,一首曲子往往有着许多重复的旋律片段。如果取整首歌作为样本,则会增加样本的冗余度。因此,本发明以一段音乐旋律当做样本,这样做不仅能减少原来样本的维度,还能增加了样本的个数。

[0010] 上述研究中基本都选择传统机器学习的分类器进行分类,如支持向量机(SVM),K邻近算法(KNN),神经网络等等,它们都有着严格的逻辑证明,而且训练时间短,在数据量少的情况下能发挥出不俗的效果。但是,本发明使用的特征将包含伴奏信息维度,在特征维度增多的情况下,更适合用深度学习进行处理这些样本,深度学习是机器学习的新领域,随着近几年计算机能力的不断提高和数据量的爆发性增长,它的优势越来越明显,由于它本身具有提取特征的能力,因此不需要我们再人工提取数据的统计特征,只需要输入的数据尽可能涵盖原始数据的信息即可。它大体上能分为卷积神经网络(CNN)和循环神经网络(RNN)。目前,深度学习已经在图像识别,语音识别等领域大放异彩。本发明也选择适用深度学习的方式实现MIDI钢琴曲风格分类。

发明内容

[0011] 本发明的目的在于克服现有技术存在的上述不足,提供一种融合伴奏的MIDI钢琴曲风格分类方法,以提高分类准确率。

[0012] 本发明的目的通过如下技术方案实现。

[0013] 本发明实施一种融合伴奏的MIDI钢琴曲风格分类方法,采用样本采集单元以及数据分类单元实现,样本采集单元主要从一首钢琴曲中提取出含有主旋律和伴奏信息的音乐片段样本,其中每一个音乐片段样本由固定数量的音符向量组成,每个音符向量都涵盖了主旋律音符音高、该主旋律音符在该片段的起始时间、该主旋律音符的持续时间以及该主旋律音符可能存在的伴奏音符的音高信息,数据分类单元用于接收样本采集单元的样本,进行训练和分类。

[0014] 这里先引入伴奏音符组的概念,假设一个音符 n_0 的起始时间为 s_0 ,结束时间为 t_0 ,另一个音符 n_1 的起始时间为 s_1 ,结束时间为 t_1 ,如果同时满足 $s_0 \geq s_1$ 和 $t_0 \leq t_1$,则称 n_1 为 n_0 的伴奏音符。一个音符的所有伴奏音符的集合成为伴奏音符组。

[0015] 进一步的,样本采集单元主要接收整个MIDI文件,产生若干个含有主旋律和伴奏信息的音乐片段样本,其中每一个音乐片段样本由固定数量的音符向量组成。每个音符向量都包含八个维度。第一个维度为主旋律音符的音高,为了提高学习速率,需要对音高进行归一化操作,由于钢琴的键盘由88个键组成,最低音的音符序号为21,最高音的音符序号为108,所以可以作 $p = (p - 21) / 88$ 的归一化操作。第二个维度为该主旋律音符在该所在音乐片段的起始时间。第三个维度为该主旋律音符的持续时间。第四到第八维分别保存该主旋律音符的伴奏音符组里所有音符的音高信息,按从大到小排列,与第一个维度相同,也同样需要进行 $p = (p - 21) / 88$ 的归一化操作,如果该主旋律音符的伴奏音符不足5个,则多余的维度的值为0,如果超过5个,则舍弃音高较低的伴奏音符。

[0016] 进一步的,数据分类单元主要由深度神经网络构成,该深度神经网络可以但不限于是一维卷积神经网络,它用于接收样本采集单元产生的音乐片段样本,作为其输入特征,其输出为音乐风格标签的独热编码,即使用与标签种类数量相同的寄存器来对风格标签进行编码。数据分类单元主要有训练模式和分类模式两种模式。在训练模式下,音乐片段样本集的输入,以音乐片段样本所在乐曲的风格标签作为样本的标签,以对深度神经网络进行训练。在分类模式下,它接收一首乐曲下的所有音乐片段样本的输入,将这些样本依次输入至深度神经网络中得到多个风格标签,取出现次数最多的标签作为该乐曲的风格标签。

[0017] 进一步的,在分类前,需要对数据分类单元内的深度神经网络进行训练,以达到良好的分类效果。训练步骤包括:

步骤1:获取训练用的钢琴曲MIDI文件训练数据集,其中每一个文件都已经打上风格标签;

步骤2:步骤1的文件合集依次输入到样本采集单元,每个文件都会产生若干个含有主旋律和伴奏信息的音乐片段样本,给这些音符片段都打上与该文件相同的标签,得到样本数据集;

步骤3:将数据分类单元调至训练模式,取步骤2产生的样本数据集输入到数据分类单元进行多轮训练,以达到良好的分类效果。

[0018] 进一步的,本发明提出的分类方法,分类步骤包括:

步骤1:将一个未分类的MIDI文件通过样本采集单元产生若干个含有主旋律和伴奏信息的音乐片段样本;

步骤2:将数据分类单元调至分类模式,取步骤1产生的所有音乐片段样本输入至数据分类单元,产生对应的风格标签。

[0019] 与现有技术相比,本发明具有如下优点和技术效果:

1. 本发明使用一段音乐片段作为样本而不是一整首乐曲,既减少了样本的维度,也增加了样本的数量;

2. 本发明选取的特征既包含了主旋律信息,也包含了可能存在的伴奏信息,囊括了乐曲片段的大部分信息,有利于提高分类准确率;

3. 本发明采用了深度学习的方式进行分类,在数据量大的情况下分类效果优于传统方法。

附图说明

[0020] 图1为所述分类方法的整个工作流程。

[0021] 图2为样本采集单元工作的主要步骤。

[0022] 图3为数据分类单元训练模式下的主要步骤。

[0023] 图4为数据分类单元分类模式下的主要步骤。

具体实施方式

[0024] 本实施实例结合附图对本发明的具体实施方式作进一步说明,但本发明的实施不限于此,需指出的是,以下若有未特别详细说明之处,均是本领域技术人员可根据现有技术实现的。

[0025] 图1为所述分类方法的整个工作流程。具体流程如下：

步骤1: MIDI数据集通过样本采集单元, 产生多个样本;

步骤2: 如果需要训练, 则跳到步骤3, 如果需要分类, 则跳到步骤5;

步骤3: 将数据分类模块调至训练模式;

步骤4: 将步骤1产生的所有样本输入到数据分类单元进行多轮训练, 直到数据分类模块的深度神经网络达到良好的分类效果, 整个流程结束;

步骤5: 将数据分类模块调至分类模式;

步骤6: 在步骤1产生的样本中, 将同一首歌产生的所有样本输入至数据分类模块, 得到这首歌的分类标签。

[0026] 图2为样本采集单元工作的主要步骤。步骤如下:

步骤1: 提取出MIDI乐曲中的主旋律音符组, 音符组由一系列的音符向量组成, 音符向量是一个八维的向量, 前三个维度分别代表主旋律音符的音高, 起始时间, 持续时间, 第四到第八个维度为伴奏维度, 用于保存可能存在的主旋律音符的伴奏音符的音高, 其初始值为21, 以便于之后的归一化操作。音符组里的音符向量按照起始时间从小到大排列;

步骤2: 遍历由步骤1得到的音符组里面的每个音符向量, 得到该音符向量在这首乐曲中的伴奏音符组, 将伴奏音符组里的音符里的音高从大到小依次填入该音符向量的第四到第八个维度, 得到了新的含有伴奏信息的音符组;

步骤3: 将步骤2所得到的音符组分成若干个音符片段, 每一个片段各有k个音符向量。并对其中每一个音符向量的音高维度以及伴奏维度作 $p = (p - 21) / 88$ 的归一化操作, 并让每个音符的起始时间都减去该片段第一个音符的起始时间, 以加快之后学习速率。这样, 就生成了若干个音符片段, 每一个音符片段都作为输入的特征。

[0027] 图3为数据分类单元训练模式下的主要步骤。具体步骤如下:

步骤1: 获取从样本采集单元中得到的样本集;

步骤2: 将步骤1的样本集中的每个样本都打上该样本所在乐曲的标签;

步骤3: 将样本输入到深度神经网络进行多轮训练, 直到该深度神经网络达到良好的分类效果。

[0028] 图4为数据分类单元分类模式下的主要步骤。具体步骤如下:

步骤1: 从样本采集单元中得到一首歌所有样本;

步骤2: 将步骤1得到的样本依次输入到深度神经网络, 得到多个标签;

步骤3: 在步骤2得到的多个标签中, 取出现次数最多的标签作为输出标签。

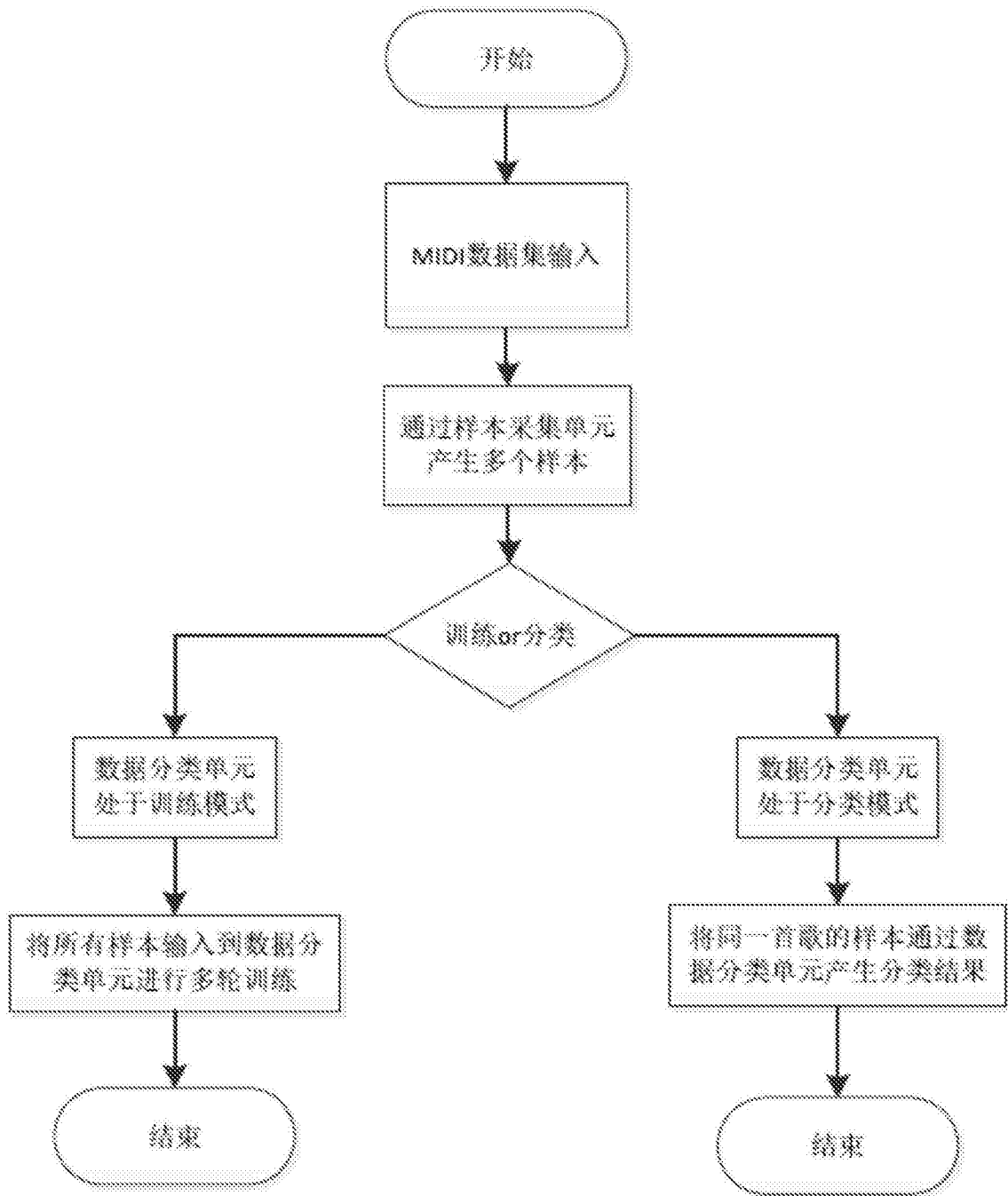


图1

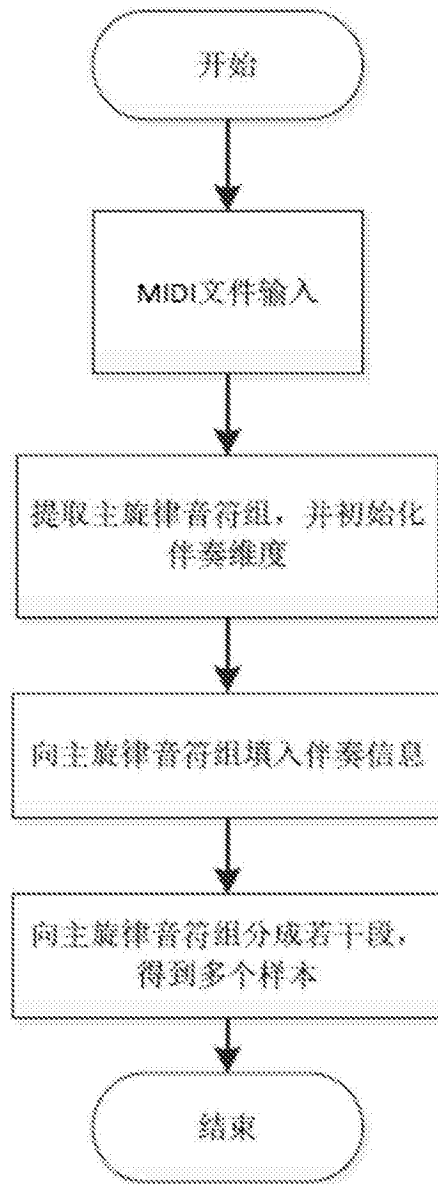


图2

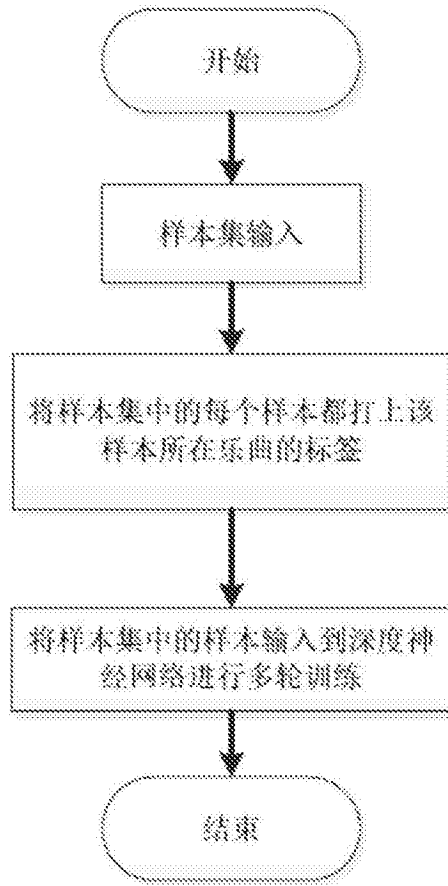


图3

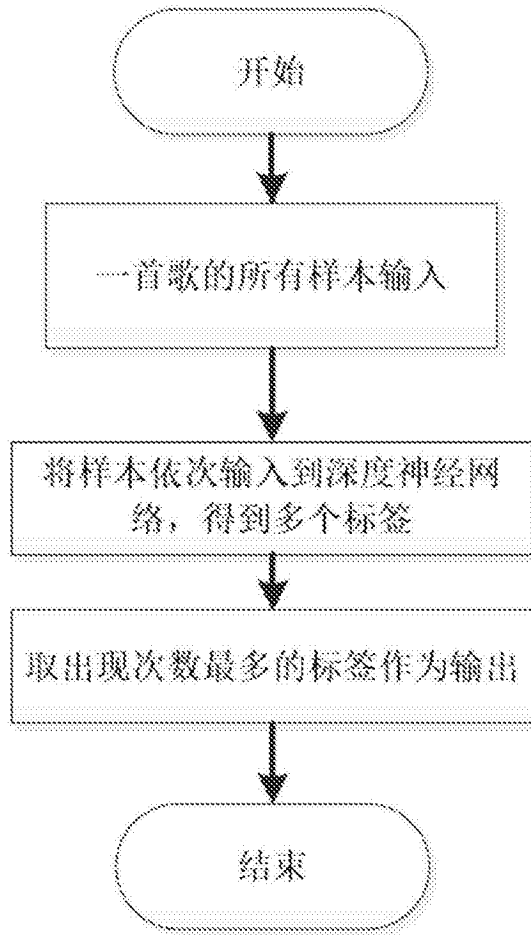


图4