



(12)发明专利

(10)授权公告号 CN 108665901 B

(45)授权公告日 2020.06.30

(21)申请号 201810421815.5

G10L 17/04(2013.01)

(22)申请日 2018.05.04

审查员 张天然

(65)同一申请的已公布的文献号

申请公布号 CN 108665901 A

(43)申请公布日 2018.10.16

(73)专利权人 广州国音科技有限公司

地址 510000 广东省广州市南沙区丰泽东
路106号(自编1号楼)X1301-A4423(集
群注册)(JM)

(72)发明人 陈昊亮 陈丹

(74)专利代理机构 深圳市深佳知识产权代理事

务所(普通合伙) 44285

代理人 王仲凯

(51)Int.Cl.

G10L 17/02(2013.01)

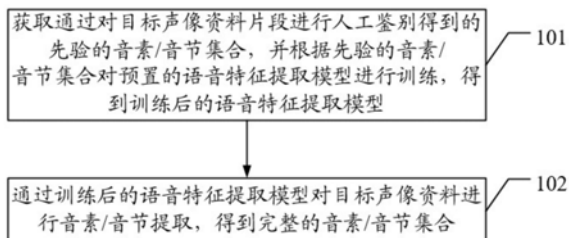
权利要求书1页 说明书6页 附图2页

(54)发明名称

一种音素/音节提取方法及装置

(57)摘要

本发明公开了一种音素/音节提取方法及装置,包括:S1:获取通过对目标声像资料片段进行人工鉴别得到的先验的音素/音节集合,根据先验的音素/音节集合对预置的语音特征提取模型进行训练,得到训练后的语音特征提取模型;S2:通过训练后的语音特征提取模型对目标声像资料进行音素/音节提取,得到完整的音素/音节集合。本发明的音素/音节提取方法将人工鉴别和机器提取相结合,与纯人工鉴别相比,减少了人力和耗时,提高了提取效率,与纯机器提取相比,以先验的音素/音节集合对语音特征提取模型进行训练,通过训练后的语音特征提取模型进行音素/音节提取,提高了准确性,解决了当前的音素/音节提取方法无法同时兼顾准确性和效率的技术问题。



1. 一种音素/音节提取方法,其特征在于,包括:

S1:获取通过对目标声像资料片段进行人工鉴别得到的先验的音素/音节集合,并根据先验的音素/音节集合对预置的语音特征提取模型进行训练,得到训练后的语音特征提取模型;

S2:通过训练后的语音特征提取模型对目标声像资料进行音素/音节提取,得到完整的音素/音节集合。

2. 根据权利要求1所述的一种音素/音节提取方法,其特征在于,步骤S2具体包括:

S21:对目标声像资料进行初步音素/音节提取得到未筛选的音素/音节集合;

S22:通过训练后的语音特征提取模型对未筛选的音素/音节集合进行模型匹配,提取未筛选的音素/音节集合中与训练后的语音特征提取模型的匹配度超过预置阈值的音素/音节,得到完整的音素/音节集合。

3. 根据权利要求1所述的一种音素/音节提取方法,其特征在于,步骤S1之前还包括:S0;

S0:通过人工鉴别对目标声像资料片段进行音素/音节提取,得到先验的音素/音节集合。

4. 根据权利要求1所述的一种音素/音节提取方法,其特征在于,步骤S2之后还包括:S3;

S3:根据完整的音素/音节集合进行语音同一性鉴定,得到语音同一性鉴定意见。

5. 一种音素/音节提取装置,其特征在于,包括:

模型训练单元,用于获取通过对目标声像资料片段进行人工鉴别得到的先验的音素/音节集合,并根据先验的音素/音节集合对预置的语音特征提取模型进行训练,得到训练后的语音特征提取模型;

特征提取单元,用于通过训练后的语音特征提取模型对目标声像资料进行音素/音节提取,得到完整的音素/音节集合。

6. 根据权利要求5所述的一种音素/音节提取装置,其特征在于,特征提取单元包括:

初始子单元,用于对目标声像资料进行初步音素/音节提取得到未筛选的音素/音节集合;

筛选子单元,用于通过训练后的语音特征提取模型对未筛选的音素/音节集合进行模型匹配,提取未筛选的音素/音节集合中与训练后的语音特征提取模型的匹配度超过预置阈值的音素/音节,得到完整的音素/音节集合。

7. 根据权利要求5所述的一种音素/音节提取装置,其特征在于,还包括:片段提取单元;

片段提取单元,用于通过人工鉴别对目标声像资料片段进行音素/音节提取,得到先验的音素/音节集合。

8. 根据权利要求5所述的一种音素/音节提取装置,其特征在于,还包括:语音鉴定单元;

语音鉴定单元,用于根据完整的音素/音节集合进行语音同一性鉴定,得到语音同一性鉴定意见。

一种音素/音节提取方法及装置

技术领域

[0001] 本发明涉及声纹领域,尤其涉及一种音素/音节提取方法及装置。

背景技术

[0002] 随着科技的发展,声纹 (Voiceprint) 识别在生活中的应用逐渐增多,其中,声纹是指用电声仪器显示的携带言语信息的声波频谱。

[0003] 现代科学研究表明,声纹不仅具有特定性,而且具有相对稳定性的特点。实验证明,成年以后人的声音可保持长期相对稳定不变,且无论讲话者是故意模仿他人声音和语气,还是耳语轻声讲话,即使模仿得惟妙惟肖,其声纹却始终不相同。基于声纹的这两个特征,侦查人员就可将获取的犯罪分子的声纹和嫌疑人的声纹,通过声纹鉴定技术进行检验对比,迅速认定罪犯,为侦查破案提供可靠的证据

[0004] 声纹鉴定又称语音同一性鉴定,或话者识别/鉴定,指的通过比较、分析和对声像资料记载的语音的同一性问题所进行的科学判断。在实际公安及司法工作中,鉴定人员常常需要对涉案语音(如勒索、恐吓电话录音,经济纠纷中当事双方的谈话录音等)进行检验,分析说话人的身份和判断涉案语音(检材语音)与特定对象语音(样本语音)是否来源于同一人,并作出科学的评判性书面意见—语音同一性鉴定意见,进而为案件的调查提供线索和方向,为法庭诉讼提供证据。

[0005] 声纹鉴定主要分成两类:即话者辨认 (Speaker Identification) 和话者确认 (Speaker Verification)。前者用以判断某段语音是若干人中的哪一个所说的,是“多选一”问题,而后者用以确认某段语音是否是指定的某个人所说的,是“一对一判别”问题。如缩小刑侦范围时可能需要辨认技术,而银行交易时则需要确认技术。

[0006] 不管是辨认还是确认,都需要先对说话人的声纹进行建模。建模需要提取声像资料中的目前对象的音素或者音节。当前主要采用人工鉴定的方法或者纯机器提取的方法,人工鉴定的方法准确性高,但是需要的人力多,且耗时长,效率低下,通过纯机器提取的方法,效率高,但是准确性低。

[0007] 因此,导致了当前的音节/音节提取方法无法同时兼顾准确性和效率的技术问题。

发明内容

[0008] 本发明提供了一种音素/音节提取方法及装置,解决了当前的音节/音节提取方法无法同时兼顾准确性和效率的技术问题。

[0009] 本发明提供了一种音素/音节提取方法,包括:

[0010] S1:获取通过对目标声像资料片段进行人工鉴别得到的先验的音素/音节集合,并根据先验的音素/音节集合对预置的语音特征提取模型进行训练,得到训练后的语音特征提取模型;

[0011] S2:通过训练后的语音特征提取模型对目标声像资料进行音素/音节提取,得到完整的音素/音节集合。

- [0012] 优选地,步骤S2具体包括:
- [0013] S21:对目标声像资料进行初步音素/音节提取得到未筛选的音素/音节集合;
- [0014] S22:通过训练后的语音特征提取模型对未筛选的音素/音节集合进行模型匹配,提取未筛选的音素/音节集合中与训练后的语音特征提取模型的匹配度超过预置阈值的音素/音节,得到完整的音素/音节集合。
- [0015] 优选地,步骤S1之前还包括:S0;
- [0016] S0:通过人工鉴别对目标声像资料片段进行音素/音节提取,得到先验的音素/音节集合。
- [0017] 优选地,步骤S2之后还包括:S3;
- [0018] S3:根据完整的音素/音节集合进行语音同一性鉴定,得到语音同一性鉴定意见。
- [0019] 本发明提供了一种音素/音节提取装置,包括:
- [0020] 模型训练单元,用于获取通过对目标声像资料片段进行人工鉴别得到的先验的音素/音节集合,并根据先验的音素/音节集合对预置的语音特征提取模型进行训练,得到训练后的语音特征提取模型;
- [0021] 特征提取单元,用于通过训练后的语音特征提取模型对目标声像资料进行音素/音节提取,得到完整的音素/音节集合。
- [0022] 优选地,特征提取单元包括:
- [0023] 初始子单元,用于对目标声像资料进行初步音素/音节提取得到未筛选的音素/音节集合;
- [0024] 筛选子单元,用于通过训练后的语音特征提取模型对未筛选的音素/音节集合进行模型匹配,提取未筛选的音素/音节集合中与训练后的语音特征提取模型的匹配度超过预置阈值的音素/音节,得到完整的音素/音节集合。
- [0025] 优选地,还包括:片段提取单元;
- [0026] 片段提取单元,用于通过人工鉴别对目标声像资料片段进行音素/音节提取,得到先验的音素/音节集合。
- [0027] 优选地,还包括:语音鉴定单元;
- [0028] 语音鉴定单元,用于根据完整的音素/音节集合进行语音同一性鉴定,得到语音同一性鉴定意见。
- [0029] 从以上技术方案可以看出,本发明具有以下优点:
- [0030] 本发明提供了一种音素/音节提取方法,包括:S1:获取通过对目标声像资料片段进行人工鉴别得到的先验的音素/音节集合,并根据先验的音素/音节集合对预置的语音特征提取模型进行训练,得到训练后的语音特征提取模型;S2:通过训练后的语音特征提取模型对目标声像资料进行音素/音节提取,得到完整的音素/音节集合。
- [0031] 本发明公开的音素/音节提取方法将人工鉴别和机器提取相结合,与纯人工鉴别相比,大大减少了人力和耗时,提高了音素/音节提取的效率,同时,与传统的纯机器提取方法相比,以人工鉴别得到的先验的音素/音节集合对语音特征提取模型进行训练,通过训练后的语音特征提取模型对目标声像资料进行音素/音节提取,极大地提高了音素/音节提取的准确性,解决了当前的音节/音节提取方法无法同时兼顾准确性和效率的技术问题。

附图说明

[0032] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其它的附图。

[0033] 图1为本发明实施例提供的一种音素/音节提取方法的一个实施例的流程示意图;

[0034] 图2为本发明实施例提供的一种音素/音节提取方法的另一个实施例的流程示意图;

[0035] 图3为本发明实施例提供的一种音素/音节提取装置的一个实施例的结果示意图。

具体实施方式

[0036] 本发明实施例提供了一种音素/音节提取方法及装置,解决了当前的音节/音节提取方法无法同时兼顾准确性和效率的技术问题。

[0037] 为使得本发明的发明目的、特征、优点能够更加的明显和易懂,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,显然,下面所描述的实施例仅仅是本发明一部分实施例,而非全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例,都属于本发明保护的范围。

[0038] 请参阅图1,本发明实施例提供了一种音素/音节提取方法的一个实施例,包括:

[0039] 步骤101:获取通过对目标声像资料片段进行人工鉴别得到的先验的音素/音节集合,并根据先验的音素/音节集合对预置的语音特征提取模型进行训练,得到训练后的语音特征提取模型;

[0040] 需要说明的是,音素是语音中的最小的单位,依据音节里的发音动作来分析,一个动作构成一个音素,音素分为元音、辅音两大类;

[0041] 音节是语音中最自然的结构单位。确切地说,音节是音位组合构成的最小的语音结构单位,它的构成分头腹尾三部分,因而音节之间具有明显可感知的界限;

[0042] 进行语音同一性鉴定时,可以根据需要选择提取音节进行鉴定或者提取音素进行鉴定;

[0043] 获取通过对目标声像资料片段进行人工鉴别得到的先验的音素/音节集合,并根据先验的音素/音节集合对预置的语音特征提取模型进行训练,则可以得到训练后的语音特征提取模型。

[0044] 步骤102:通过训练后的语音特征提取模型对目标声像资料进行音素/音节提取,得到完整的音素/音节集合。

[0045] 需要说明的是,以训练后的语音特征提取模型对目标声像资料进行音素/音节提取,则可以提取到目标声像资料中与目标对象有关的完整的音素/音节集合。

[0046] 本实施例公开的音素/音节提取方法将人工鉴别和机器提取相结合,与纯人工鉴别相比,大大减少了人力和耗时,提高了音素/音节提取的效率,同时,与传统的纯机器提取方法相比,以人工鉴别得到的先验的音素/音节集合对语音特征提取模型进行训练,通过训练后的语音特征提取模型对目标声像资料进行音素/音节提取,极大地提高了音素/音节提

取的准确性,解决了当前的音节/音节提取方法无法同时兼顾准确性和效率的技术问题。

[0047] 以上为本发明实施例提供的一种音素/音节提取方法的一个实施例,以下为本发明实施例提供的一种音素/音节提取方法的另一个实施例。

[0048] 请参阅图2,本发明实施例提供了一种音素/音节提取方法的另一个实施例,包括:

[0049] 步骤201:通过人工鉴别对目标声像资料片段进行音素/音节提取,得到先验的音素/音节集合;

[0050] 需要说明的是,如果纯人工鉴别进行音素/音节提取,则需要的时间太长,因此,只是通过人工鉴别对目标声像资料片段进行音素/音节提取,将得到的音素/音节集合作为先验信息。

[0051] 步骤202:获取通过对目标声像资料片段进行人工鉴别得到的先验的音素/音节集合,并根据先验的音素/音节集合对预置的语音特征提取模型进行训练,得到训练后的语音特征提取模型;

[0052] 需要说明的是,通过先验的音素/音节集合可对预置的语音特征提取模型进行训练,预置的语音特征提取模型为通用的语音特征提取模型,所以难以根据不同目标对象提取准确的音素/音节集合,通过输入先验信息进行训练,训练后的语音特征提取模型会使音素/音节的提取更具备针对性,从而提取准确的音素/音节集合。

[0053] 步骤203:对目标声像资料进行初步音素/音节提取得到未筛选的音素/音节集合;

[0054] 需要说明的是,得到训练后的语音特征提取模型之后,可以直接使用语音特征提取模型提取目标声像资料中的音素/音节集合,也可以先使用纯机器提取方式提取目标声像资料中的音素/音节集合,再通过语音特征提取模型进行筛选。

[0055] 步骤204:通过训练后的语音特征提取模型对未筛选的音素/音节集合进行模型匹配,提取未筛选的音素/音节集合中与训练后的语音特征提取模型的匹配度超过预置阈值的音素/音节,得到完整的音素/音节集合;

[0056] 需要说明的是,通过训练后的语音特征提取模型对未筛选的音素/音节集合进行模型匹配,根据匹配度对未筛选的音素/音节集合进行筛选,只有匹配度超过预置阈值才被提取进入完整的音素/音节集合。

[0057] 步骤205:根据完整的音素/音节集合进行语音同一性鉴定,得到语音同一性鉴定意见。

[0058] 需要说明的是,得到目标声像资料的完整的音素/音节集合之后,则可以进行语音同一性鉴定,判断目标声像资料与待鉴定声像资料是否是同一个人,得到语音同一性鉴定意见。

[0059] 传统的人工鉴别方式虽然准确性极高,可以视为100%正确,但是需要耗费大量的人力和时间,效率很低,传统的机器提取效率高,但是因为语音发音受地域、时间和个体影响较大,难以使用一个通用的语音特征提取模型进行针对性提取,所以实际应用中及其提取的准确性在60%左右;

[0060] 音素/音节自动提取或标注算法一方面希望算法模型适配尽可能多的场景,适配尽可能多的地域、时间和个体范围;另一方面,又要求算法对任意单一的场景有较高的准确性,客观上,这属于性能上对立的两极,比较困难同时满足以上两项要求;

[0061] 而本实施例公开的音素/音节提取方法将人工鉴别和机器提取相结合,与纯人工

鉴别相比,大大减少了人力和耗时,提高了音素/音节提取的效率,同时,与传统的纯机器提取方法相比,以人工鉴别得到的先验的音素/音节集合对语音特征提取模型进行训练,通过训练后的语音特征提取模型对目标声像资料进行音素/音节提取,极大地提高了音素/音节提取的准确性;

[0062] 本实施例的音素/音节提取方法相比传统的人工鉴别方式,大大提高了提取效率,同时因为采用部分人工鉴别的结果作为机器提取的先验信息,因此可以得到较高的准确性,解决了当前的音节/音节提取方法无法同时兼顾准确性和效率的技术问题。

[0063] 以上为本发明实施例提供的一种音素/音节提取方法的另一个实施例,以下为本发明实施例提供的一种音素/音节提取装置的一个实施例。

[0064] 请参阅图3,本发明实施例提供了一种音素/音节提取装置的一个实施例,包括:

[0065] 模型训练单元301,用于获取通过对目标声像资料片段进行人工鉴别得到的先验的音素/音节集合,并根据先验的音素/音节集合对预置的语音特征提取模型进行训练,得到训练后的语音特征提取模型;

[0066] 特征提取单元302,用于通过训练后的语音特征提取模型对目标声像资料进行音素/音节提取,得到完整的音素/音节集合。

[0067] 进一步地,特征提取单元302包括:

[0068] 初始子单元3021,用于对目标声像资料进行初步音素/音节提取得到未筛选的音素/音节集合;

[0069] 筛选子单元3022,用于通过训练后的语音特征提取模型对未筛选的音素/音节集合进行模型匹配,提取未筛选的音素/音节集合中与训练后的语音特征提取模型的匹配度超过预置阈值的音素/音节,得到完整的音素/音节集合。

[0070] 进一步地,还包括:片段提取单元300;

[0071] 片段提取单元300,用于通过人工鉴别对目标声像资料片段进行音素/音节提取,得到先验的音素/音节集合。

[0072] 进一步地,还包括:语音鉴定单元303;

[0073] 语音鉴定单元303,用于根据完整的音素/音节集合进行语音同一性鉴定,得到语音同一性鉴定意见。

[0074] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统,装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0075] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统,装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0076] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0077] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0078] 所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-OnlyMemory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0079] 以上所述,以上实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

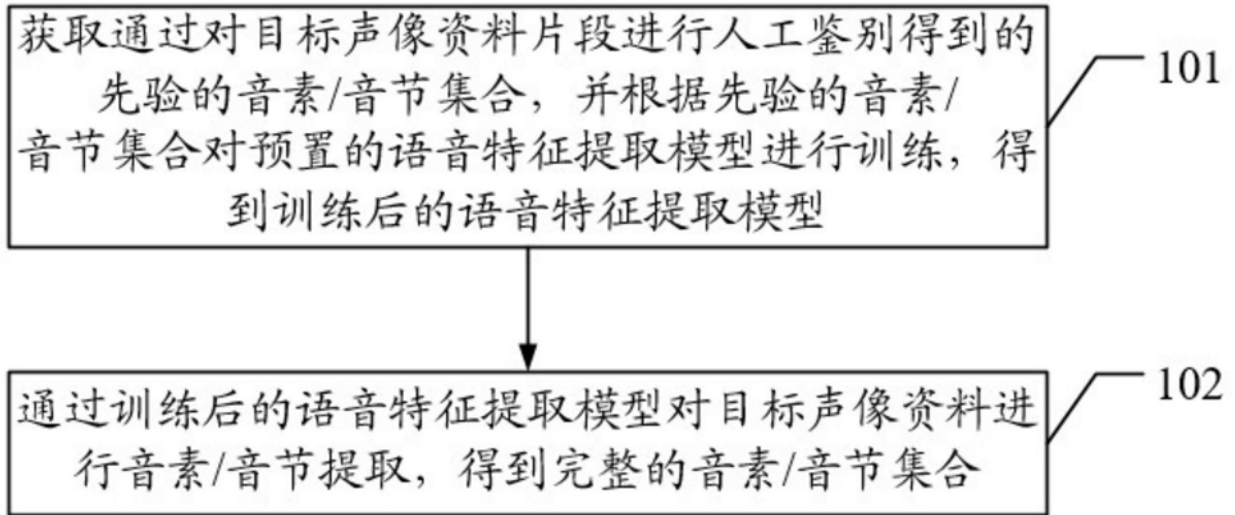


图1

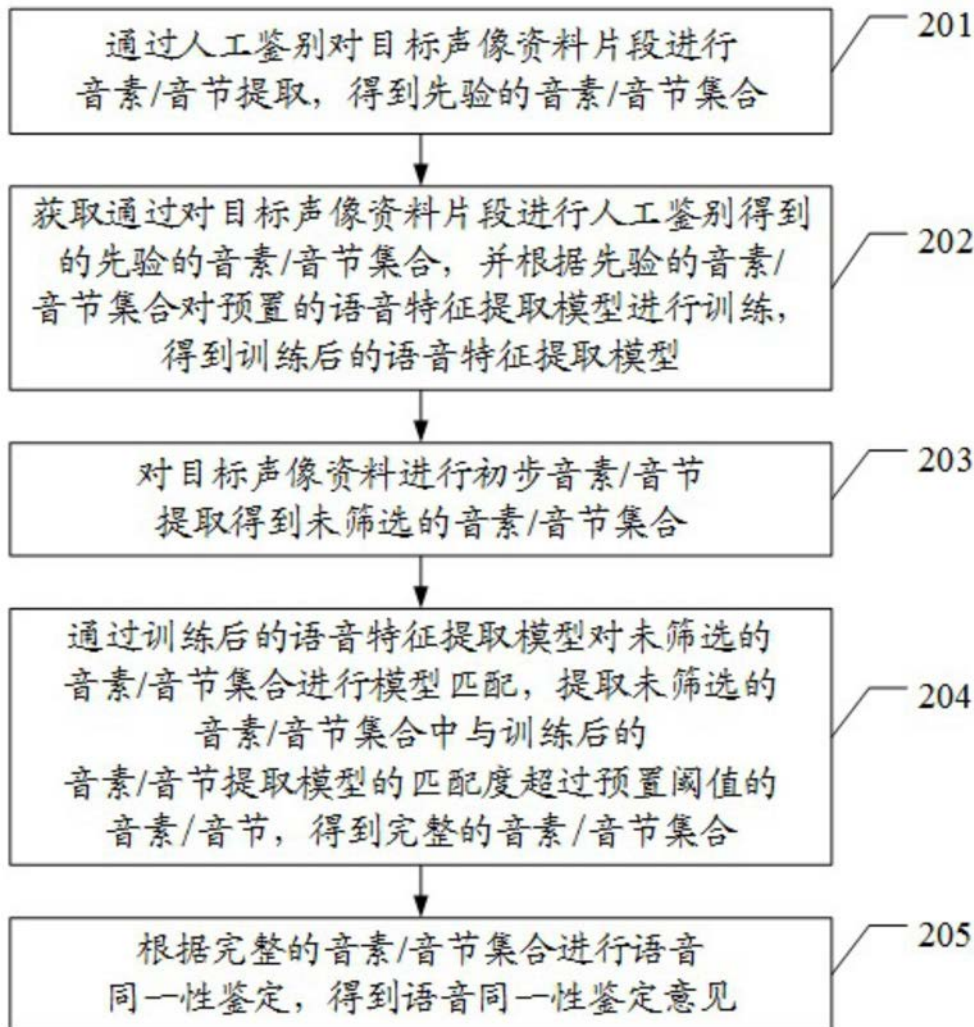


图2

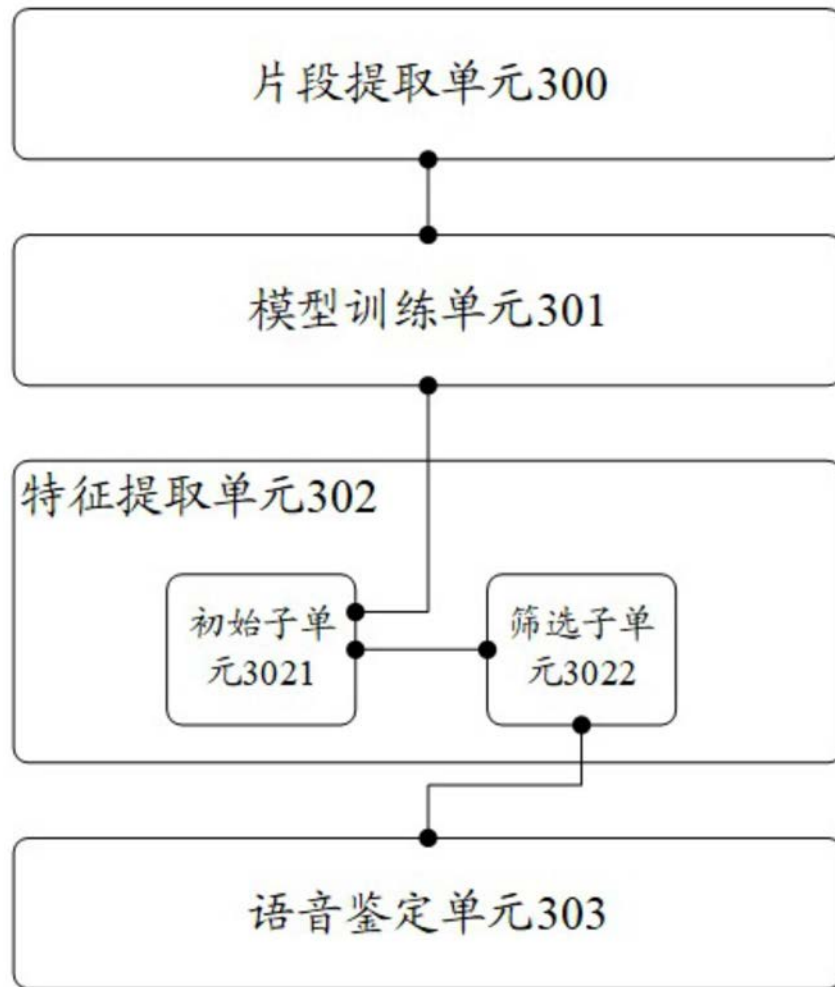


图3