



(12) 发明专利

(10) 授权公告号 CN 102541965 B

(45) 授权公告日 2015. 05. 20

(21) 申请号 201010624747. 6

CN 101226526 A, 2008. 07. 23, 全文.

(22) 申请日 2010. 12. 30

CN 101499268 A, 2009. 08. 05, 全文.

(73) 专利权人 国际商业机器公司

审查员 张伯

地址 美国纽约阿芒克

(72) 发明人 包胜华 陈健 杨新颖 苏中

(74) 专利代理机构 北京市金杜律师事务所

11256

代理人 鄞迅 姜彦

(51) Int. Cl.

G06F 17/30(2006. 01)

G06F 17/22(2006. 01)

(56) 对比文件

US 2007131094 A1, 2007. 06. 14, 说明书段 50, 10-11, 52-61, 75-83, 第 8-10 页表格.

US 2003023421 A1, 2003. 01. 30, 全文.

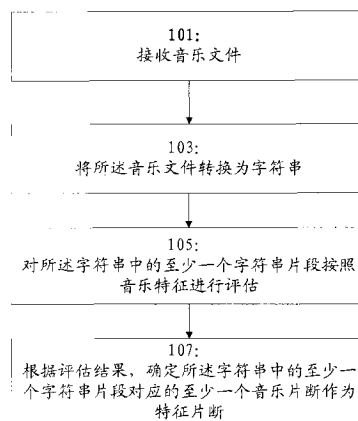
权利要求书2页 说明书10页 附图5页

(54) 发明名称

自动获得音乐文件中的特征片断的方法和系统

(57) 摘要

本发明提供了一种自动获得音乐文件中的特征片断的方法,包括:接收音乐文件;将所述音乐文件转换为字符串;对所述字符串中的至少一个字符串片段按照音乐特征进行评估;以及根据评估结果,确定所述字符串中的至少一个字符串片段对应的至少一个音乐片断作为特征片断。并且本发明还提供了一种自动获得音乐文件中的特征片断的系统。通过本发明中的自动获得音乐文件中的特征片断的技术,可以自动的获得音乐文件中最具代表性的片段。对这些有代表性的片段进行搜索可以在保证搜索效果的同时大大降低搜索的用时,同时减少待搜索数据的存储量,节省存储空间。



1. 一种自动获得音乐文件中的特征片断的方法,包括:
接收音乐文件;
将所述音乐文件转换为字符串;
对所述字符串中的至少一个字符串片段按照音乐特征进行评估;以及
根据评估结果,确定所述字符串中的至少一个字符串片段对应的至少一个音乐片断作为特征片断,从而自动的获得音乐文件中具代表性的片段。
2. 如权利要求 1 所述的方法,其中所述转换步骤进一步包括:
按照音调或者音节长度中的至少一个对所述字符串进行标准化表示。
3. 如权利要求 1 所述的方法,所述评估步骤进一步包括:
对所述字符串中超过预定长度的至少一个字符串片段进行评估。
4. 如权利要求 1 所述的方法,其中所述音乐特征包括音乐片段重复次数,并且所述评估步骤进一步包括:
处理所述字符串以获得所述字符串的差分表示;以及
计算所述差分表示中的至少一个字符串片段的重复次数。
5. 如权利要求 1-4 中任意一个所述的方法,其中所述音乐特征包括平均音高,并且所述评估步骤进一步包括:
计算所述字符串中的至少一个字符串片段所对应的音乐片段的平均音高。
6. 如权利要求 1-4 中任意一个所述的方法,其中所述音乐特征包括片段出现位置,并且所述评估步骤进一步包括:
判断所述字符串中的至少一个字符串片段在整个字符串中的出现位置。
7. 如权利要求 1-4 中任意一个所述的方法,其中所述音乐特征包括片段长度,并且所述评估步骤进一步包括:
计算所述字符串中的至少一个字符串片段的片段长度。
8. 如权利要求 1-4 中任意一个所述的方法,其中所述音乐特征包括半音比重,并且所述评估步骤进一步包括:
计算所述字符串中的至少一个字符串片段所对应的音乐片段的半音比重。
9. 如权利要求 1-4 中任意一个所述的方法,其中所述确定步骤进一步包括:
使用经训练数据训练过的分类器来确定特征片断,其中所述训练数据包括给定的音乐文件及其特征片段。
10. 如权利要求 1-4 中任意一个所述的方法,其中所述确定步骤进一步包括:
对所述字符串中的不同字符串片段的评估结果进行比较,进而根据比较结果确定特征片段。
11. 如权利要求 1-4 中任意一个所述的方法,其中所述确定步骤进一步包括:
将所述字符串中的至少一个字符串片段对应的多个音乐片断组合作为特征片断。
12. 一种自动获得音乐文件中的特征片断的系统,包括:
接收装置,用于接收音乐文件;
转换装置,用于将所述音乐文件转换为字符串;
评估装置,用于对所述字符串中的至少一个字符串片段按照音乐特征进行评估;以及
确定装置,用于根据评估结果,确定所述字符串中的至少一个字符串片段对应的至少

一个音乐片断作为特征片断,从而自动的获得音乐文件中具代表性的片段。

13. 如权利要求 12 所述的系统,其中所述转换装置进一步包括:

标准化装置,用于按照音调或者音节长度中的至少一个对所述字符串进行标准化表示。

14. 如权利要求 12 所述的系统,其中所述评估装置进一步用于对所述字符串中超过预定长度的至少一个字符串片段进行评估。

15. 如权利要求 12 所述的系统,其中所述音乐特征包括音乐片段重复次数,并且所述评估装置进一步包括:

处理装置,用于处理所述字符串以获得所述字符串的差分表示;以及

重复次数计算装置,用于计算所述差分表示中的至少一个字符串片段的重复次数。

16. 如权利要求 12-15 中任意一个所述的系统,其中所述音乐特征包括平均音高,并且所述评估装置进一步包括:

平均音高计算装置,用于计算所述字符串中的至少一个字符串片段所对应的音乐片段的平均音高。

17. 如权利要求 12-15 中任意一个所述的系统,其中所述音乐特征包括片段出现位置,并且所述评估装置进一步包括:

位置判断装置,用于判断所述字符串中的至少一个字符串片段在整个字符串中的出现位置。

18. 如权利要求 12-15 中任意一个所述的系统,其中所述音乐特征包括片段长度,并且所述评估装置进一步包括:

片段长度计算装置,用于计算所述字符串中的至少一个字符串片段的片段长度。

19. 如权利要求 12-15 中任意一个所述的系统,其中所述音乐特征包括半音比重,并且所述评估装置进一步包括:

比重计算装置,用于计算所述字符串中的至少一个字符串片段所对应的音乐片段的半音比重。

20. 如权利要求 12-15 中任意一个所述的系统,其中所述确定装置进一步使用经训练数据训练过的分类器来确定特征片断,其中所述训练数据包括音乐文件及指定的特征片段。

21. 如权利要求 12-15 中任意一个所述的系统,其中所述确定装置进一步对所述字符串中的不同字符串片段的评估结果进行比较,进而根据比较结果确定特征片段。

22. 如权利要求 12-15 中任意一个所述的系统,其中所述确定装置进一步将所述字符串中的至少一个字符串片段对应的多个音乐片断组合作为特征片断。

自动获得音乐文件中的特征片断的方法和系统

技术领域

[0001] 本发明总体上涉及对音乐文件进行处理的方法和系统,特别的本发明涉及自动获得音乐文件中的特征片断的方法和系统。

背景技术

[0002] 传统的歌曲搜索方式主要是以文字输入为主,比如用户通过输入歌名、作曲、词作者等信息从而获得返回的歌曲结果。近年来哼唱搜索 (Query by Humming, QbH),打破了传统的文字搜索方式,提供了一种全新的声音搜索方式。用户哼唱一段歌曲或者一段旋律作为系统的输入,哼唱搜索系统通过检索数据库得到返回的歌曲或乐曲结果。Musipedia 是一种以哼唱或者键盘敲击等为用户输入的 QbH 产品。Midomi 是一种 QbH 网络服务,其目前主要提供英语、日语、中文和西班牙文的哼唱检索服务。Tunebot 是一种以哼唱、曲调、旋律等为用户输入的 QbH 产品。MPEG-7 标准中也将 QbH 搜索列入其中。如何在海量的音乐文件内进行快速有效的搜索是哼唱搜索技术发展中一直追求的一个目标。

发明内容

[0003] 本发明的发明人发现如果能够减少哼唱搜索中音乐文件的搜索数据量则可以提高检索的速度。通常情况下,每个音乐文件都有自己最有代表性的一些片段(比如歌曲的副歌部分),而这些最有代表性的片段往往更容易被人们所记住,因此在哼唱搜索中,这些片段更容易被用户所哼唱作为搜索的检索式(query)。如何寻找音乐文件中的有代表性的片段(特征片段)是一个需要解决的问题。如果能够通过自动的方式确定出音乐文件中的特征片段,则可以大大节省人力成本,从而确保通过搜索特征片段进行哼唱搜索更具操作性。

[0004] 本发明提供了一种自动获得音乐文件中的特征片断的方法,包括:接收音乐文件;将所述音乐文件转换为字符串;对所述字符串中的至少一个字符串片段按照音乐特征进行评估;以及根据评估结果,确定所述字符串中的至少一个字符串片段对应的至少一个音乐片断作为特征片断。

[0005] 本发明还提供了一种自动获得音乐文件中的特征片断的系统,包括:接收装置,用于接收音乐文件;转换装置,用于将所述音乐文件转换为字符串;评估装置,用于对所述字符串中的至少一个字符串片段按照音乐特征进行评估;以及确定装置,用于根据评估结果,确定所述字符串中的至少一个字符串片段对应的至少一个音乐片断作为特征片断。

[0006] 通过本发明中的自动获得音乐文件中的特征片断的技术,可以自动的获得音乐文件中最具代表性的片段。对这些有代表性的片段进行搜索可以在保证搜索效果的同时大大降低搜索的用时,同时减少待搜索数据的存储量,节省存储空间。特别是对于一些移动终端而言,其存储空间有限,如果可以仅仅存储音乐文件的特征片段,则可以在本地实现快速的哼唱搜索。

[0007] 除了哼唱搜索以外,自动获得音乐文件中的特征片断的技术还有很多其它的应用

需求。比如,自动的实现手机彩铃(color ring)的提取。众所周知,手机彩铃的时长是有一定限制的,因此自动提取音乐文件中的特征片段作为手机彩铃可以节省大量的人力成本。

[0008] 再如,本发明中的技术还可以被用来实现音乐文件的预览,从而让浏览者可以很快确定某个音乐文件是否是自己想要的。特别的,按照很多国家或者地区的著作权法,未经权利人同意在网络上提供音乐文件的复制件是侵权行为,但是仅提供一小段音乐文件的特征片段作为摘要(snippet)供网络用户在进行网络搜歌过程中预览也许在一些国家和地区能够被认为是合理使用。

[0009] 上述仅仅简单列举了本发明几种可能的应用场景,但是本发明并不限于任何上述应用场景,而是可以适用于任何需要自动获得音乐文件特征片断的应用。

附图说明

[0010] 本说明中所参考的附图只用于示例本发明的典型实施例,不应该认为是对本发明范围的限制。

[0011] 图 1 示出了一种自动获得音乐文件中的特征片断的方法的流程图。

[0012] 图 2 示出了按照本发明的一个实施例对字符串中的至少一个字符串片段按照音乐特征进行评估的流程图。

[0013] 图 3 示出了按照本发明的一个实施例根据评估结果确定字符串中的至少一个字符串片段对应的至少一个音乐片断作为特征片断的流程图。

[0014] 图 4 示出了一种自动获得音乐文件中的特征片断的系统的框图。

[0015] 图 5 示出了按照本发明的一个实施例的评估装置的系统框图。

具体实施方式

[0016] 下列讨论中,提供大量具体的细节以帮助彻底了解本发明。然而,很显然对于本领域技术人员来说,即使没有这些具体细节,并不影响对本发明的理解。并且应该认识到,使用如下的任何具体术语仅仅是为了方便描述,因此,本发明不应当局限于只用在这样的术语所表示和 / 或暗示的任何特定应用中。

[0017] 图 1 示出了一种自动获得音乐文件中的特征片断的方法的流程图。在步骤 101 中接收音乐文件。所述音乐文件可以包括由歌手演唱的歌曲也可以包括由用乐器弹出的乐曲。所述音乐文件可以以各种文件格式(如 WAV、MP3、WMA 等等)存储。

[0018] 在步骤 103 中将所述音乐文件转换为字符串。现有技术中已经披露了一些将音乐文件转换为字符串的方法,比如软件产品 IntelliScore Ensemble(<http://3d2f.com/programs/14-967-intelliscore-ensemble-mp3-to-midi-converter-download.shtml>) 就可以将多种格式的音乐文件转换为 MIDI 音乐。进一步对 MIDI 音乐进行符号化,则可以将其转换成真正的字符串。有很多的软件产品能够用于将 MIDI 音乐转换为字符串,比如 MIDIFile2Text 1.1(<http://www.softpedia.com/downloadTag/MIDI+to+Text>)。所述字符串既可以是音乐文件的简谱(比如 $\underline{3} \underline{5} \underline{3} \underline{2} | \underline{3} \underline{2} \underline{3} | \dots$),也可以是对简谱的进一步处理结果,比如进一步对所述音乐文件的简谱中的音高或者音节长度中任意一个进行标准化表示。对音高进行标准化表示可以是对音调高低进行统一的数字化表示,如 $\dot{2} \rightarrow 9$ 。对音节长度进行标准化表示可以为对音节长度进行统一的数字化表示,如 $3 \cdot \rightarrow \underline{3} \underline{3} \underline{3}$ 。此外,还可以对

简谱进行其它的标准化表示以进一步使其数字化,如 $9\ 10\ 23\ |3\ 3\ 3\ 5 \rightarrow 9, 10, 2, 3, 3, 3, 3, 5$ 。上述列举了对音乐文件的简谱进行进一步处理的几个实例,这些进一步处理方案可以进一步提高确定特征片段的准确性。但是上述实例并不构成对本发明的限制,实际上根据实际应用的需求,以及考虑到对计算成本的要求,也可以不进行上述对简谱的进一步处理,比如可以简单的将“3·”表示为“3”,而不考虑音节长度的问题。

[0019] 在步骤 105 中对所述字符串中的至少一个字符串片段按照音乐特征进行评估。所述音乐特征可以是音乐片段重复次数、平均音高、片段出现位置、片段长度、半音比重或者其它的音乐特征中的一种或多种的组合。对所述字符串中的至少一个字符串片段进行评估的具体内容将在后文中展开更详细的说明。

[0020] 在步骤 107 中根据评估结果,确定所述字符串中的至少一个字符串片段对应的至少一个音乐片断作为特征片断。

[0021] 图 2 示出了按照本发明的一个实施例对字符串中的至少一个字符串片段按照音乐特征进行评估的流程图。步骤 201, 203 描述了按照本发明的一个实施例如果所述音乐特征为音乐片段的重复,所述评估步骤的进一步实现细节。步骤 211 描述了按照本发明的一个实施例如果所述音乐特征为平均音高,所述评估步骤的进一步实现细节。步骤 213 描述了按照本发明的一个实施例如果所述音乐特征为片段出现位置,所述评估步骤的进一步实现细节。步骤 215 描述了按照本发明的一个实施例如果所述音乐特征为片段长度,所述评估步骤的进一步实现细节。步骤 217 描述了按照本发明的一个实施例如果所述音乐特征为半音比重,所述评估步骤的进一步实现细节。

[0022] 按照本发明的一个实施例,如果所述音乐特征为音乐片段重复次数,在所述评估步骤中将确定重复的音乐片段,比如歌曲的副歌部分,并对字符串片段的重复次数进行计算,得到的计算结果作为该音乐特征(即音乐片段重复次数)的取值。比如《北京欢迎你》这首歌中,存在两个重复的字符串片段,分别为字符串片段 A 和 B,片段 A 为“35323233……2033”,片段 B 为“35856……2811”,全曲排列是 AABAABBBBB。因此片段 A 的重复次数为 4,片段 B 的重复次数为 6。

[0023] 按照本发明的另一个实施例,不仅能够计算音乐片段的简单重复,还能够计算音乐片段的变调重复(比如升调重复或者降调重复)。比如《雪绒花》(Edelweiss)这首歌中,从字符串片段“345666”到字符串片段“567888”就属于升调重复。类似的变调重复现象在很多歌曲中都是很比较常见的,对变调重复进行统计可以更全面准确的确定音乐文件中有代表性的片段。为了能够统计变调重复,在步骤 201 中处理所述字符串以获得所述字符串的差分表示。比如对《雪绒花》中的字符串片段“345666”进行处理,使后一个数字减去前一个数字,得到差分表示后的字符串片段为“11100”,同理对字符串片段“567888”按照同样的规则进行处理,也可以得到差分表示后的字符串片段“11100”。对字符串进行差分表示可以寻找出音乐文件中的变调重复的片段。当然本发明并不限于上述列举的差分计算规则,而是还可以使用任何差分计算规则,比如用前一个数字减去后一个数字,或者用相隔的两个数字进行相减等等。获得字符串片段的差分表示后,在步骤 203 中计算所述差分表示中的至少一个字符串片段的重复次数。比如在《雪绒花》中字符串片段“345666”出现了 1 次,“567888”出现了 2 次,利用本实施例中的差分表示方法,可以计算字符串片段“345666”的重复次数为 3 次。

[0024] 按照本发明的另一个实施例,所述音乐特征为平均音高 (average pitch),在步骤 211 中计算所述字符串中的至少一个字符串片段所对应的音乐片的平均音高,作为该音乐特征的取值。比如《北京欢迎你》这首歌中片段 A 为“35323233……2033”,其字符串长度为 52,片段 B 为“35856……2811”,其字符串长度为 43。片段 A 的平均音高为 $(3+5+3+2+3+2+3+3+3+\dots+2+0+3+3)/52 = 3.3$,片段 B 的平均音高为 $(3+5+8+5+6+\dots+2+8+1+1)/43 = 4.2$ 。在很多情况下,平均音高较高的音乐片段成为特征片段的可能性会比较大,比如很多副歌部分的平均音高都比较高。因此对平均音高进行计算可以更准确的确定特征片段。

[0025] 按照本发明的另一个实施例,所述音乐特征为片段出现位置 (segment location),在步骤 221 中判断所述字符串中的至少一个字符串片段在整个字符串中的出现位置。通常情况下,排在曲尾的比非排在曲尾的音乐片段更有可能成为特征片段。因此可以在步骤 221 中进一步根据字符串片段的出现位置对其进行打分,作为该音乐特征 (片段出现位置) 的取值。比如《北京欢迎你》这首歌中全曲排列是 AABAABBBBB,片段 B 在曲尾,打分为 1,片段 A 不在曲尾,打分为 0。当然本发明并不限于上述打分方式,而是可以应用其它的打分方式,比如将音乐文件,分为曲头 (如全曲长度的前 20%)、曲中 (如全曲长度的中间 60%)、曲尾 (如全曲长度的后 20%),并且根据音乐片段的出现位置按照一定条件进行打分。

[0026] 按照本发明的另一个实施例,所述音乐特征为片段长度 (segment length),在步骤 231 中计算所述字符串中的至少一个字符串片段的片段长度,作为该音乐特征 (片段长度) 的取值。比如《北京欢迎你》中片段 A 的字符串长度为 52,片段 B 的字符串长度为 43。

[0027] 按照本发明的另一个实施例,所述音乐特征为半音比重 (semitone percentage),在步骤 241 中计算所述字符串中的至少一个字符串片段所对应的音乐片的半音比重,作为该音乐特征 (半音比重) 的取值。在一些音乐文件中,音乐片段中所含半音的数量越多,音乐片段成为特征片段的可能性就越大。12345671 为一个八度的八个音,各音之音的关系并不相同,分为全音和半音两种,34 之间和 71 之间是半音,其它相邻两音间是全音。因此长度为 8 的字符串片段“34345677”中存在两个半音,其半音比重为 $2/8$ 。

[0028] 上述仅以五种音乐特征为例进行说明,本发明还可以适用于其它的音乐特征。并且本发明不仅适用于一种音乐特征,而是还可以将多种音乐特征进行结合,对所述字符串中的至少一个字符串片段进行综合评估。

[0029] 按照本发明的一个实施例,在步骤 105 中对所述字符串中的至少一个字符串片段进行评估可以进一步包括对所述字符串中超过预定长度的至少一个字符串片段进行评估。理论上,一个长度为 N 的字符串可以分割为长度为 n 的字符串片段 (n 大于等于 1 且小于 N)。实际上,将字符串分割为长度过小的字符串片段对分析音乐文件中的特征片段可能没有太多作用。比如在实际操作中,将长度为 N 的字符串分割为 N 个长度为 1 的字符串片段的可能性比较小。因此按照本实施,可以只对所述字符串中超过预定长度 (比如 5) 的字符串片段进行上述评估。

[0030] 图 3 示出了按照本发明的一个实施例根据评估结果确定字符串中的至少一个字符串片段对应的至少一个音乐片断作为特征片断的流程图。在步骤 301 中获得训练数据包括给定的音乐文件及其特征片段。经过对训练数据中音乐文件的特征片段的在音乐特征上

的取值进行分析可以确定成为特征片段的一些分类规则,从而将这些分类规则适用于新的音乐文件,以找出新的音乐文件中的特征片段。在一个例子中,训练数据中的音乐文件可以是一些歌曲,其特征片段可以这些歌曲的彩铃,所述彩铃可以是手工指定的歌曲中比较有代表性的音乐片段。

[0031] 在步骤 303 中将所述音乐文件及其特征片段转化为字符串,并进行文本对齐 (textual alignment),也就是将特征片段的字符串在音乐文件全曲对应的字符串中进行定位,比如定位“567”在“1234567”的尾部。

[0032] 可选的,在步骤 305 中还可以通过声音验证 (sound validation) 对上述对齐的文本进行校正。在 303 步骤中的文本对齐可能存在误差。导致误差的原因可能有很多,比如在转换成字符串的过程中出现错误等等。因此为了确保正确的将特征片段的字符串片段定位到整个歌曲的字符串中,还可以在步骤 305 中进行声音验证,比如确定字符串“567”所对应的特征片段的的声音确实出现在字符串“1234567”所对应的特征片段的尾部。比较两个声音片段是否一样,可以通过比较两个声音片段的音频、音调、强度等特征来进行。一些现有产品或技述已经能够提供类似的声音比较,比如产品 Similarity。

[0033] 在步骤 307 中计算特征片段在一种或多种音乐特征上的取值。所述音乐特征可以是音乐片段重复次数 (包括音乐片段的简单重复或表调重复)、平均音高、片段出现位置、片段长度、半音比重或者其它的音乐特征中的一种或多种的组合。上文中已经详细介绍了如何计算每种音乐特征上的取值,这部分内容在此不再赘述。

[0034] 在步骤 309 中将所述特征片段在一种或多种音乐特征上的取值输入分类器,形成分类规则。决策树是一种最简单的分类器。通过记录大量特征片段在一种或多种音乐特征上的取值来构建决策树,可以获得成为特征片段的一些规则,作为分类规则,比如下面所示的分类规则 1:分类规则 1 = (音乐片段的重复次数 > 5) & (平均音高 > 4) & (片段出现位置 > 0) & (片段长度 > 20) & (半音比重 >= 0)

[0035] 所述分类规则可以用来判断某个字符串片段所对应的音乐片段是否是特征片段。仍然以《北京欢迎你》这首歌为例进行说明,全曲字符串分为 A、B 两个片段,片段 A 为“35323233……2033”,片段 B 为“35856……2811”,其全曲排列为 AABAABBBBB。A、B 两个字符串片段在各个音乐特征上的取值如下表 1 所示:

[0036]

	A	B
音乐片段的重复次数	4	6
平均音高	3.3	4.2
片段出现位置	0	1
片段长度	43	52
半音比重	0	0

[0037] 表 1

[0038] 利用获得的分类规则 1, 可以判断出片段 B 所对应的音乐片段为歌曲《北京欢迎你》的特征片段, 而片段 A 所对应的音乐片段不是特征片段。

[0039] 决策树仅仅是分类器中的比较简单的实现方式。一些比较复杂的分类模型可能构建更加复杂的分类规则, 比如朴素贝叶斯模型 (Naive Bayes)、支持向量机模型 (SVM) 等。上述实施例仅仅示意出分类规则的一种实施例, 并不构成对本发明的限制。使用任何经训练数据训练过的分类器都可以用来确定特征片断。

[0040] 按照本发明的另一个实施例, 通过对所述字符串中的不同字符串片段的评估结果进行比较, 进而根据比较结果确定特征片段。仍然以《北京欢迎你》这首歌为例进行说明。通过比较片段 A 和片段 B 在一个或多个音乐特征上的取值的综合评估结果, 可以确定哪个片段更适合成为特征片段。比如按照如下表 2 对五个音乐特征的权重进行分配:

[0041]

音乐片段的重复次数	40%
平均音高	30%
片段出现位置	10%
片段长度	10%
半音比重	10%

[0042] 表 2

[0043] 并且对表 1 中的片段 A 和片段 B 在各个音乐特征上的取值按照归一化规则进行归一化处理, 如下表 3 所示:

[0044]

	A	B
音乐片段的重复次数	$0.67 = (4/6)$	1
平均音高	$0.79 = (3.3/4.2)$	1
片段出现位置	0	1
片段长度	1	0.8
半音比重	0	0

[0045] 表 3

[0046] 其中所述归一化规则为, 对音乐片段重复次数而言, 以字符串片段的最大重复次数 (在本例中为 6) 为 1, 对其它字符串片段的重复次数按比例缩小, 片段 B 的重复次数为 6, 其归一化后的值为 1, 片段 A 的重复次数为 4, 其归一化后的值为 0.67。

[0047] 对平均音高而言, 以字符串片段的最大平均音高为 1, 对其它字符串片段的平均音高按比例缩小, 片段 B 的平均音高为 42, 其归一化后的值为 1, 片段 A 的平均音高为 3.3, 其

归一化后的值为 0.79。

[0048] 对片段出现位置而言,出现在曲尾的取值为 1,非出现在曲尾的取值为 0,因此归一化后,片段 A 的取值仍为 0,片段 B 的取值仍为 1。

[0049] 对片段长度而言,可以定义归一化规则为长度在 30-50 之间的片段的归一化的值为 1,长度在 50 以上的片段的取值为 0.8,长度在 30 以下的片段的取值为 0.4,因此片段 A 的长度为 43,归一化后的值为 1,片段 B 的长度为 52,归一化后的值为 0.8。

[0050] 对半音比重而言,可以以字符串片段的半音比重作为该音乐特征的取值,在本例中片段 A 和片段 B 都没有出现半音,因此其半音比重都为 0,归一化后也为 0。

[0051] 上述归一化规则仅仅是示例性的,本发明还可以应用其它不同的归一化规则。

[0052] 经过上面的计算,片段 A 的综合评估结果为:

[0053] $0.67*40\% + 0.79*30\% + 0*10\% + 1*10\% + 0*10\% = 0.605$

[0054] 片段 B 的综合评估结果为:

[0055] $1*40\% + 1*30\% + 1*10\% + 0.8*10\% + 0*10\% = 0.88$

[0056] 可见片段 B 的综合评估结果大于片段 A 的综合评估结果,因此片段 B 所对应的音乐片段将被确定为特征片段。

[0057] 由于《北京欢迎你》这首歌全曲排列是 AABAABBBBB,因此字符串片段 B 对应于 6 段音乐片段,选择一段音乐片段作为特征片段,还是选择多段音乐片段组合成为特征片段,本发明对此不作限定,而是可以根据实际应用场景的需要实现为各种方式。

[0058] 上述实施例中将《北京欢迎你》这首歌的字符串分为两个字符串片段,片段 A 和片段 B。本发明还可以在其它的实施例中字符串片段构建更加复杂的分割方式,如将字符串片段 A 进一步划分为 A1、A2、A3 三个子段,将字符串片段 B 进一步划分为 B1、B2、B3 三个子段,并且分别计算 A1、A2、A3 及 B1、B2、B3 在各个音乐特征上的取值,再进一步综合计算片段 A 和片段 B 在各个音乐特征上的取值。

[0059] 图 4 示出了一种自动获得音乐文件中的特征片断的系统的框图。图 4 中的系统包括接收装置、转换装置、评估装置及确定装置,其中接收装置用于接收音乐文件,转换装置用于将所述音乐文件转换为字符串,评估装置用于对所述字符串中的至少一个字符串片段按照音乐特征进行评估;确定装置用于根据评估结果确定所述字符串中的至少一个字符串片段对应的至少一个音乐片断作为特征片断。

[0060] 图 5 示出了按照本发明的一个实施例的评估装置的系统框图。如果所述音乐特征包括音乐片段重复次数,则所述评估装置进一步包括处理装置及重复次数计算装置。所述处理装置用于处理所述字符串以获得所述字符串的差分表示,所述重复次数计算装置用于计算所述差分表示中的至少一个字符串片段的重复次数。获得差分表示是为了在对字符串片段进行评估时将变调重复也包括在评估音素之内。按照本发明的另一实施例,也可以不考虑变调重复而是仅仅考虑简单的片段重复,这样所述评估装置进一步包括确定至少一个字符串片段的重复次数的装置(图中未示出)。

[0061] 如果所述音乐特征包括平均音高,则所述评估装置进一步包括平均音高计算装置。所述平均音高计算装置用于计算所述字符串中的至少一个字符串片段所对应的音乐片段的平均音高。

[0062] 如果所述音乐特征包括片段出现位置,则所述评估装置进一步包括位置判断装

置。所述位置判断装置用于判断所述字符串中的至少一个字符串片段的在整个字符串中的出现位置。

[0063] 如果所述音乐特征包括片段长度,则所述评估装置进一步包括片段长度计算装置。所述片段长度计算装置用于计算所述字符串中的至少一个字符串片段的片段长度。

[0064] 如果所述音乐特征包括半音比重,则所述评估装置进一步包括比重计算装置。所述比重计算装置用于计算所述字符串中的至少一个字符串片段所对应的音乐片段的半音比重。

[0065] 按照本发明的一个实施例,图 4 中的转换装置还可以进一步包括标准化装置(图 4 中未示出)。所述标准化装置用于对所述音乐文件的音调或者音节长度中任意一个进行标准化表示。

[0066] 按照本发明的另一个实施例,图 4 中的评估装置还可以进一步用于对所述字符串中超过预定长度的至少一个字符串片段进行评估。

[0067] 按照本发明的另一个实施例,图 4 中的确定装置进一步使用经训练数据训练过的分类器来确定特征片断,其中所述训练数据包括给定的音乐文件及其特征片段。

[0068] 按照本发明的另一个实施例,图 4 中的确定装置进一步对所述字符串中的不同字符串片段的评估结果进行比较,进而根据比较结果确定特征片段。

[0069] 按照本发明的另一个实施例,图 4 中的确定装置进一步将所述字符串中的至少一个字符串片段对应的多个音乐片断组合作为特征片断。

[0070] 本发明中各个系统装置实现的其它具体细节(包括其功能、实例等)与方法流程中相应的步骤对应,重复内容在此不再赘述。

[0071] 所属技术领域的技术人员知道,本发明可以体现为系统、方法或计算机程序产品。因此,本发明可以具体实现为以下形式,即,可以是完全的硬件、完全的软件(包括固件、驻留软件、微代码等)、或者本文一般称为“电路”、“模块”或“系统”的软件部分与硬件部分的组合。此外,本发明还可以采取体现在任何有形的表达介质(medium of expression)中的计算机程序产品的形式,该介质中包含计算机可用的程序码。

[0072] 可以使用一个或多个计算机可用的或计算机可读的介质的任何组合。计算机可用的或计算机可读的介质例如可以是——但不限于——电的、磁的、光的、电磁的、红外线的、或半导体的系统、装置、器件或传播介质。计算机可读介质的更具体的例子(非穷举的列表)包括以下:有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM 或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、诸如支持因特网或内部网的传输介质、或者磁存储器件。注意计算机可用的或计算机可读的介质甚至可以是上面印有程序的纸张或者其它合适的介质,这是因为,例如可以通过电扫描这种纸张或其它介质,以电子方式获得程序,然后以适当的方式加以编译、解释或处理,并且必要的话在计算机存储器中存储。在本文件的语境中,计算机可用的或计算机可读的介质可以是任何含有、存储、传达、传播、或传输供指令执行系统、装置或器件使用的或与指令执行系统、装置或器件相联系的程序的介质。计算机可用的介质可包括在基带中或者作为载波一部分传播的、由其体现计算机可用的程序码的数据信号。计算机可用的程序码可以用任何适当的介质传输,包括——但不限于——无线、电线、光缆、RF 等等。

[0073] 用于执行本发明的操作的计算机程序码,可以以一种或多种程序设计语言的任何组合来编写,所述程序设计语言包括面向对象的程序设计语言—诸如 Java、Smalltalk、C++ 之类,还包括常规的过程式程序设计语言—诸如“C”程序设计语言或类似的设计语言。程序码可以完全地在用户的计算机上执行、部分地在用户的计算机上执行、作为一个独立的软件包执行、部分在用户的计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在后一种情形中,远程计算机可以通过任何种类的网络—包括局域网(LAN)或广域网(WAN)—连接到用户的计算机,或者,可以(例如利用因特网服务提供商来通过因特网)连接到外部计算机。

[0074] 以下参照按照本发明实施例的方法、装置(系统)和计算机程序产品的流程图和/或框图描述本发明。要明白的是,流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合,都可以由计算机程序指令实现。这些计算机程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器,从而生产出一种机器,使得通过计算机或其它可编程数据处理装置执行的这些指令,产生实现流程图和/或框图中的方框中规定的功能/操作的装置。

[0075] 也可以把这些计算机程序指令存储在能指令计算机或其它可编程数据处理装置以特定方式工作的计算机可读介质中,这样,存储在计算机可读介质中的指令产生一个包括实现流程图和/或框图中的方框中规定的功能/操作的指令装置(instruction means)的制品。

[0076] 也可以把计算机程序指令加载到计算机或其它可编程数据处理装置上,使得在计算机或其它可编程数据处理装置上执行一系列操作步骤,以产生计算机实现的过程,从而在计算机或其它可编程装置上执行的指令就提供实现流程图和/或框图中的方框中规定的功能/操作的过程。

[0077] 附图中的流程图和框图,图示了按照本发明各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,所述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0078] 本文中所述的术语,仅仅是为了描述特定的实施例,而不意图限定本发明。本文中所述的单数形式的“一”和“该”,旨在也包括复数形式,除非上下文中明确地另外指出。还要知道,“包含”一词在本说明书中使用,说明存在所指出的特征、整体、步骤、操作、单元和/或组件,但是并不排除存在或增加一个或多个其它特征、整体、步骤、操作、单元和/或组件,以及/或者它们的组合。

[0079] 权利要求中的对应结构、材料、操作以及所有功能性限定的装置(means)或步骤的等同替换,旨在包括任何用于与在权利要求中具体指出的其它单元相组合地执行该功能的结构、材料或操作。所给出的对本发明的描述其目的在于示意和描述,并非是穷尽性的,

也并非是要把本发明限定到所表述的形式。对于所属技术领域的普通技术人员来说,在不偏离本发明范围和精神的情况下,显然可以作出许多修改和变型。对实施例的选择和说明,是为了最好地解释本发明的原理和实际应用,使所属技术领域的普通技术人员能够明了,本发明可以有适合所要的特定用途的具有各种改变的各种实施方式。

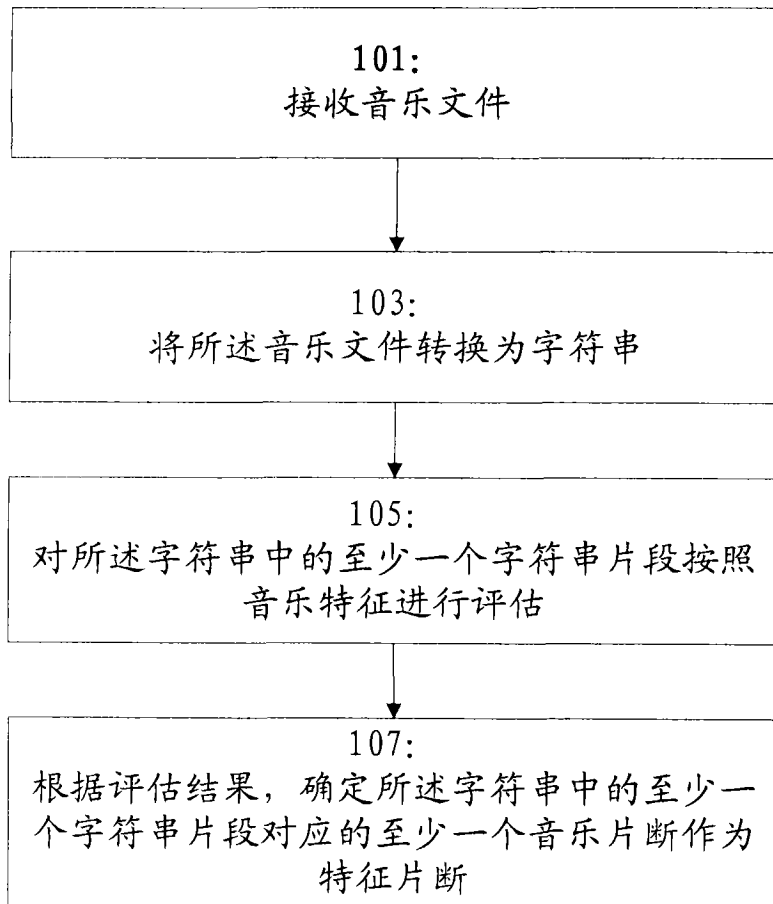


图 1

105

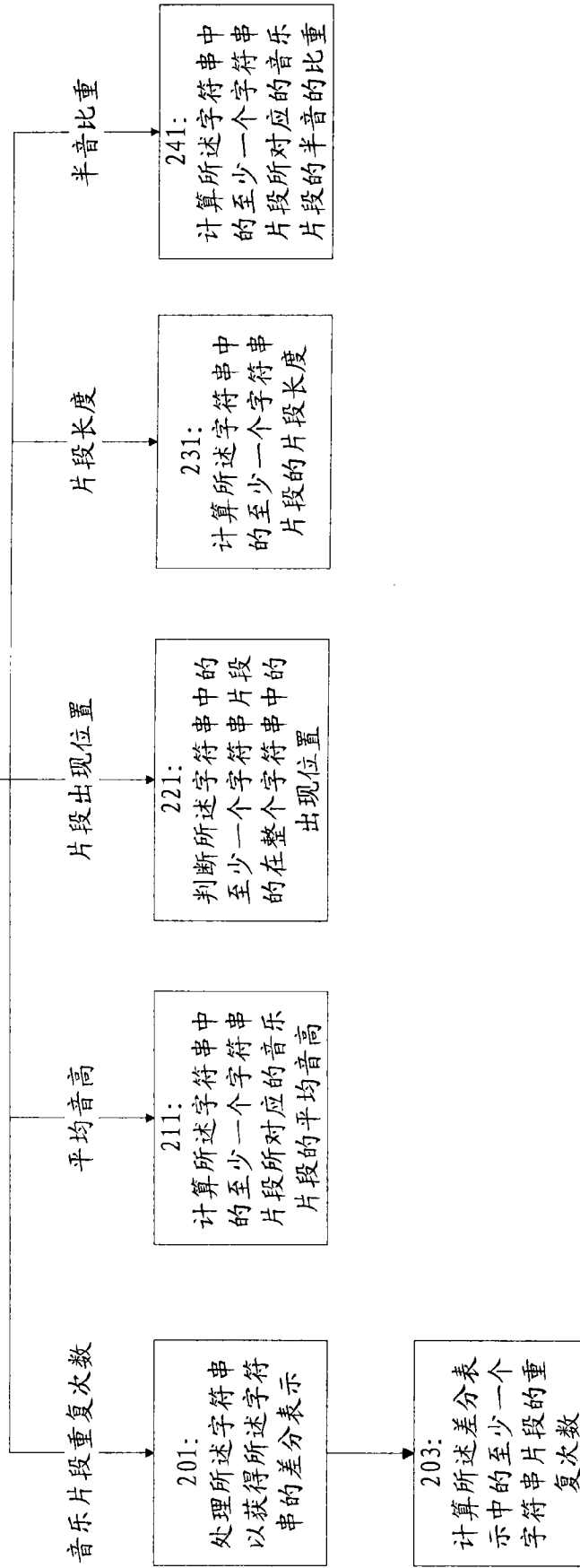


图 2

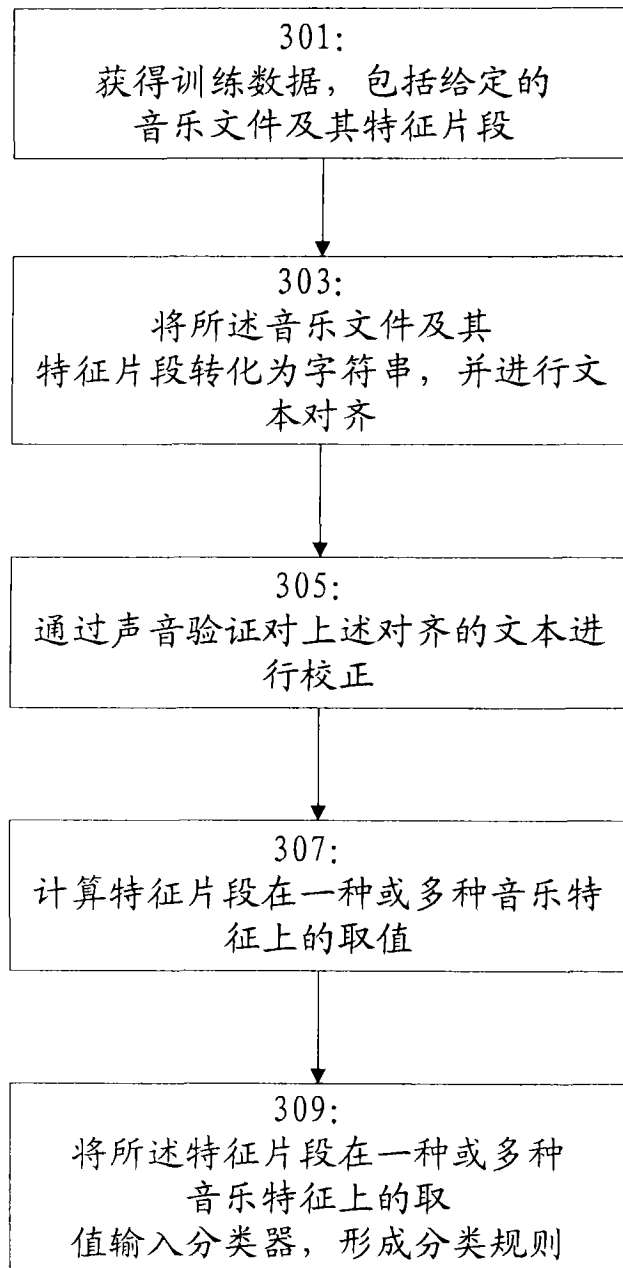


图 3

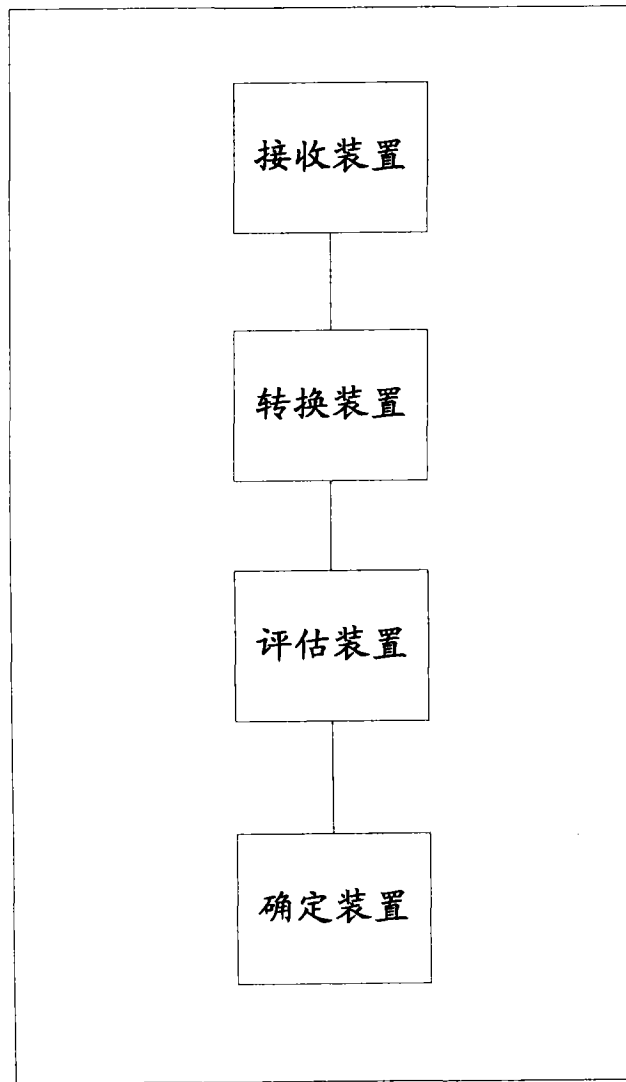


图 4

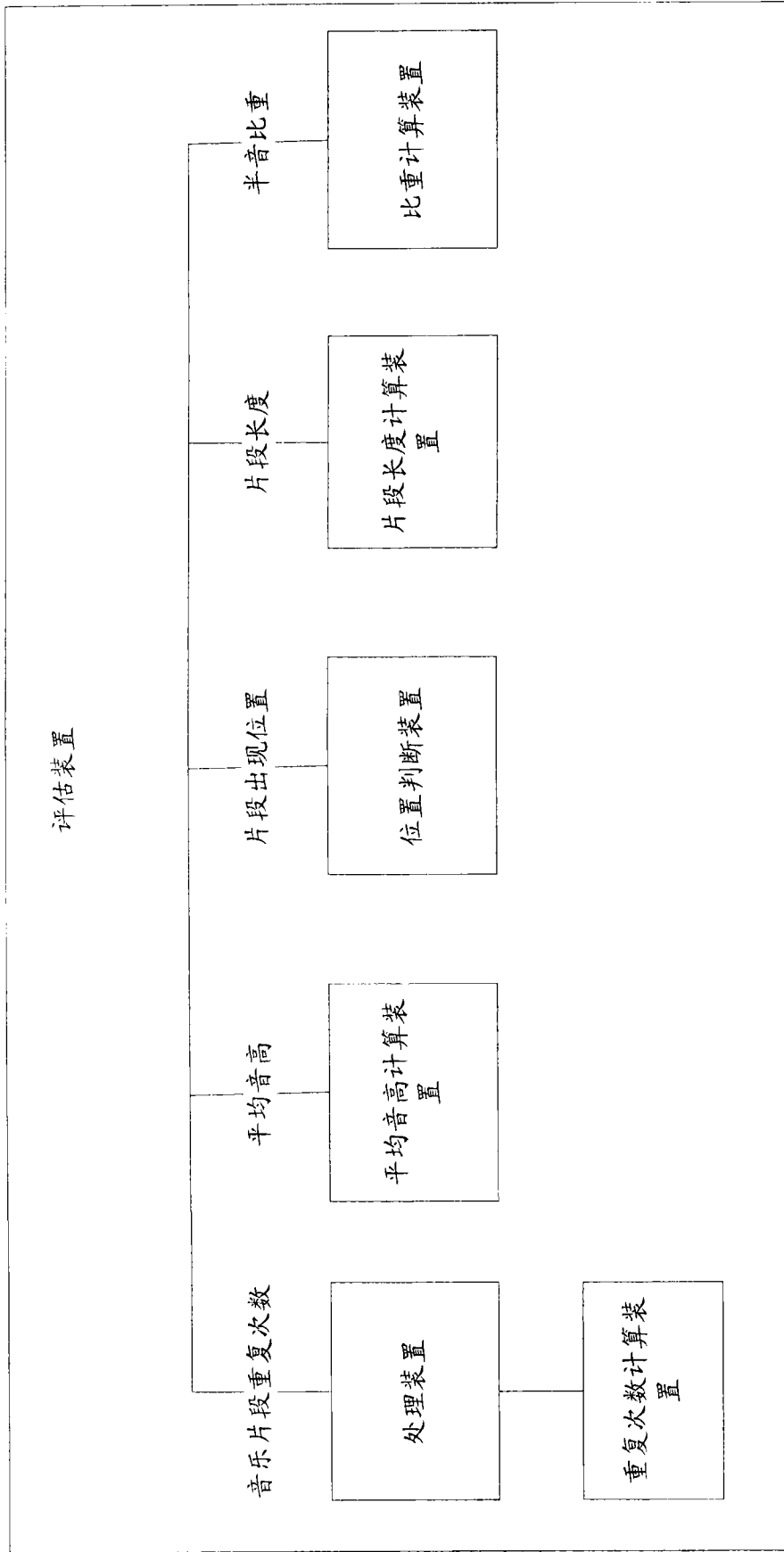


图 5