

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 6 月 18 日 (18.06.2020)



(10) 国际公布号
WO 2020/119150 A1

- (51) 国际专利分类号:
G06F 16/632 (2019.01) **G10L 25/51** (2013.01)
G06F 16/63 (2019.01)
- (21) 国际申请号: PCT/CN2019/099640
- (22) 国际申请日: 2019 年 8 月 7 日 (07.08.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201811519398.4 2018 年 12 月 12 日 (12.12.2018) CN
- (71) 申请人: 北京字节跳动网络技术有限公司 (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市

石景山区实兴大街 30 号院 3 号楼 2 层 B-0035 房间, Beijing 100041 (CN)。

(72) 发明人: 范旭 (FAN, Xu); 中国北京市海淀区知春路 63 号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。

(74) 代理人: 北京品源专利代理有限公司 (BEYOND ATTORNEYS AT LAW); 中国北京市海淀区莲花池东路 39 号西金大厦 6 层, Beijing 100036 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,

(54) **Title:** RHYTHM POINT RECOGNITION METHOD AND APPARATUS, ELECTRONIC DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 节奏点识别方法、装置、电子设备及存储介质

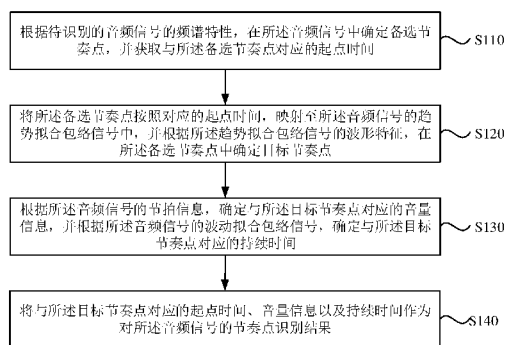


图 1a

- S110 According to spectral characteristics of an audio signal to be recognized, determine an alternative rhythm point in said audio signal, and obtain a starting time point corresponding to the alternative rhythm point
- S120 Map the alternative rhythm point into a trend fitting envelope signal of said audio signal according to the corresponding starting time point, and determine a target rhythm point in the alternative rhythm point according to waveform characteristics of the trend fitting envelope signal
- S130 Determine volume information corresponding to the target rhythm point according to beat information of said audio signal, fit an envelope signal according to the fluctuation of said audio signal, and determine the duration corresponding to the target rhythm point
- S140 Use the starting time point, the volume information, and the duration corresponding to the target rhythm point as the result of recognizing the rhythm point of said audio signal

(57) **Abstract:** Disclosed are a rhythm point recognition method and apparatus, an electronic device, and a storage medium. The method comprises: according to spectral characteristics of an audio signal to be recognized, determining an alternative rhythm point in said audio signal, and obtaining a starting time point corresponding to the alternative rhythm point; mapping the alternative rhythm point into a trend fitting envelope signal of said audio signal according to the corresponding starting time point, and determining a target rhythm point in the alternative rhythm point according to waveform characteristics of the trend fitting envelope signal; determining volume information corresponding to the target rhythm point according to beat information of said audio signal, fitting an envelope signal according to the fluctuation of said audio signal, and determining the duration corresponding to the target rhythm point; and using the starting time point, the volume information, and the duration corresponding to the target rhythm point as the result of recognizing the rhythm point of said audio signal.

(57) **摘要:** 本文公开了一种节奏点识别方法、装置、电子设备及存储介质。该方法包括: 根据待识别的音频信号的频谱特性, 在所述音频信号中确定备选节奏点, 并获取与所述备选节奏点对应的起点时间; 将所述备选节奏点按照对应的起点时间, 映射至所述音频信号的趋势拟合包络信号中, 并根据所述趋势拟合包络信号的波形特征, 在所述备选节奏点中确定目标节奏点; 根据所述音频信号的节拍信息, 确定与所述目标节奏点对应的音量信息, 并根据所述音频信号的波动拟合包络信号, 确定与所述目标节奏点对应的持续时间; 将与目标节奏点对应的起点时间、音量信息和持续时间作为对所述音频信号的节奏点识别结果。

LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,
MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,
PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

节奏点识别方法、装置、电子设备及存储介质

本申请要求在2018年12月12日提交中国专利局、申请号为201811519398.4的中国专利申请的优先权，该申请的全部内容通过引用结合在本申请中。

技术领域

本公开实施例涉及数据处理技术领域，例如涉及一种节奏点识别方法、装置、电子设备及存储介质。

背景技术

随着通信技术和电子设备的发展，电子设备例如手机、平板电脑等已经成为了人们工作和生活中不可或缺的一部分，而且随着电子设备的日益普及，安装于电子设备中的交互应用成为一种沟通和娱乐的主要渠道。

相关技术中，音乐交互应用根据音乐的节奏点，将交互提示展示给用户，用户根据交互提示输入交互操作，从而，实现激活视频特效并显示视频特效的功能。但相关技术中的节奏点一般是通过人工标注确定的，导致节奏点识别时间成本高，音乐交互应用中的音乐更新周期长。

发明内容

本公开实施例提供一种节奏点识别方法、装置、电子设备及存储介质，可以自动准确识别节奏点，提高节奏点识别效率。

本公开实施例提供了一种节奏点识别方法，该方法包括：

根据待识别的音频信号的频谱特性，在所述音频信号中确定备选节奏点，并获取与所述备选节奏点对应的起点时间；

将所述备选节奏点按照对应的起点时间，映射至所述音频信号的趋势拟合包络信号中，并根据所述趋势拟合包络信号的波形特征，在所述备选节奏点中确定目标节奏点；

根据所述音频信号的节拍信息，确定与所述目标节奏点对应的音量信息，并根据所述音频信号的波动拟合包络信号，确定与所述目标节奏点对应的持续时间；

将与所述目标节奏点对应的起点时间、音量信息以及持续时间作为对所述音频信号的节奏点识别结果。

本公开实施例还提供了一种节奏点识别装置，该装置包括：

备选节奏点确定模块，设置为根据待识别的音频信号的频谱特性，在所述音频信号中确定备选节奏点，并获取与所述备选节奏点对应的起点时间；

目标节奏点确定模块，设置为将所述备选节奏点按照对应的起点时间，映射至所述音频信号的趋势拟合包络信号中，并根据所述趋势拟合包络信号的波形特征，在所述备选节奏点中确定目标节奏点；

音量信息和持续时间确定模块，设置为根据所述音频信号的节拍信息，确定与所述目标节奏点对应的音量信息，并根据所述音频信号的波动拟合包络信号，确定与所述目标节奏点对应的持续时间；

节奏点识别结果确定模块，设置为将与所述目标节奏点对应的起点时间、音量信息以及持续时间作为对所述音频信号的节奏点识别结果。

本公开实施例还提供了一种电子设备，该电子设备包括：

一个或多个处理器；

存储装置，设置为存储一个或多个程序；

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如本公开实施例所述的节奏点识别方法。

本公开实施例还提供了一种计算机可读存储介质，存储有计算机程序，该程序被处理器执行时实现如本公开实施例所述的节奏点识别方法。

附图说明

图 1a 是本公开实施例一提供的一种节奏点识别方法的流程图；

图 1b 是本公开实施例一提供的一种音频信号的示意图；

图 2a 是本公开实施例二提供的一种节奏点识别方法的流程图；

图 2b 是本公开实施例二提供的一种音频信号的示意图；

图 3 是本公开实施例三提供的一种节奏点识别方法的流程图；

图 4 是本公开实施例四提供的一种节奏点识别装置的结构示意图；

图 5 是本公开实施例五提供的一种电子设备的结构示意图。

具体实施方式

下面结合附图和实施例对本公开进行说明。此处所描述的具体实施例仅仅用于解释本公开，而非对本公开的限定。另外，为了便于描述，附图中仅示出

了与本公开相关的部分而非全部结构。

实施例一

图 1a 为本公开实施例一提供的一种节奏点识别方法的流程图，本实施例可适用于识别一段音频信号中的节奏点的情况，该方法可以由节奏点识别装置来执行，该装置可以采用软件和/或硬件的方式实现，该装置可以配置于电子设备中，例如计算机等。如图 1a 所示，该方法包括如下步骤：

S110，根据待识别的音频信号的频谱特性，在所述音频信号中确定备选节奏点，并获取与所述备选节奏点对应的起点时间。

待识别的音频信号是指原始音频信号经过预处理生成的音频信号。一实施例中，原始音频信号是指连续的时域信号，但由于计算机只能处理离散信号，因此，需要对原始音频信号进行采样和量化，得到便于分析的离散数字信号。可以通过对原始音频信号按照设定频率进行采样，得到离散的时域信号，一实施例中，设定频率为 44.1kHz。也就是说，音频信号实际为经过采样的离散信号点形成的信号。

在本公开实施例中，频谱特性主要是指音频信号的频率、频域幅值和频域相位等参数变化的信息。

一实施例中，时域幅值和频域中计算得到的幅值不同，时域信号中某一个信号点处的幅值是该信号点对应的时间点映射在频域信号中不同频率正弦成分的信号的叠加，其中，每个信号点对应的时域幅值实际是包括多个频率信号对应的幅值信息和相位信息的叠加，并非多个频率信号对应的幅值信息的简单相加。

音频信号是一种声波信号，节奏点可以用于表示声波信号的节奏特征。通常，节奏点用于表征音符，示例性的，将在音频信号中的距离音符起始的时间点最近的信号点作为节奏点。音符的节奏特征是持续一段时间，且具有设定音量值的音符，相应的，节奏点的分析结果包括节奏点的起点时间、持续时间和音量值。

本实施例中，节奏点的起点时间可以是指该节奏点在音频信号中的开始时刻对应的时间点；持续时间可以是节奏点持续的时间长度，另外，起点时间也是该节奏点的持续时间的起点时间；音量信息可以是指节奏点的音强，用于表征该节奏点对应的声音强弱，一个音符的音强在持续时间内不是一个固定值，例如音强是不断衰减。此时，可以将音频信号中在持续时间内的信号点的时域幅值均值作为音强。

备选节奏点可以是指从音频信号中粗筛出来的节奏点。

根据待识别的音频信号的频谱特性，在所述音频信号中确定至少一个备选节奏点。例如可以是对音频信号依次进行差分处理、傅里叶变换和差分处理，并基于短时能量法确定备选节奏点以及对应的起点时间。

此外，本公开实施例还可以通过其他方法确定备选节奏点，本公开对此不作限制。

S120，将所述备选节奏点按照对应的起点时间，映射至所述音频信号的趋势拟合包络信号中，并根据所述趋势拟合包络信号的波形特征，在所述备选节奏点中确定目标节奏点。

趋势拟合包络信号可以是指对音频信号在时域中幅值特性进行拟合的信号，用于表征音频信号的时域幅值变化趋势。一实施例中，趋势拟合包络信号可以通过希尔伯特（Hilbert）变换得到。趋势拟合包络信号的波形特征可以是指音频信号的时域幅值变化趋势特征。趋势拟合包络信号的波形特征可以包括趋势拟合包络信号的波峰和波谷，对应音频信号中时域幅值波峰和时域幅值波谷。根据趋势拟合包络信号的波形特征对备选节奏点进行筛选，由于节奏点是用于表征音符，可以认为每个波峰实际就是一个音符，可以根据趋势拟合包络信号的波峰和波谷确定筛选目标节奏点，例如，在每个波峰与该波峰之前的相邻波谷之间的备选节奏点中，选择在时间上，距离该波峰最近的备选节奏点作为目标节奏点，从而可以根据每个波峰确定一个目标节奏点。

一实施例中，所述将所述备选节奏点按照对应的起点时间，映射至所述音频信号的趋势拟合包络信号中，根据所述趋势拟合包络信号的波形特征，在所述备选节奏点中确定目标节奏点，可以包括：根据所述音频信号的趋势拟合包络信号的波形特征，识别所述趋势拟合包络信号中的波峰点；将所述备选节奏点按照对应的起点时间，映射至所述趋势拟合包络信号中，并将在时间上，与所述波峰点最近的备选节奏点作为目标节奏点。

根据趋势拟合包络信号的波形特征，在趋势拟合包络信号中识别波峰点。一实施例中，若一个信号点之前和之后的信号点的时域幅值均小于该信号点，该信号点即为波峰点。将每个备选节奏点按照对应的起点时间，映射至所述趋势拟合包络信号中，可以确定每个备选节奏点对应的起点时间与波峰点对应的的时间之间的时间关系。一般来说，每个波峰可认为是一个音符，由此，根据每个波峰点从至少一个备选节奏点中筛选一个匹配的备选节奏点作为目标节奏点。一实施例中，是选择在时间上与波峰点最近的备选节奏点作为该波峰点匹配的目标节奏点。

通过根据趋势拟合包络信号的波形特征对备选节奏点进行筛选，从中确定目标节奏点，以对节奏点进一步筛选，实现提高节奏点识别的准确性。

S130, 根据所述音频信号的节拍信息, 确定与所述目标节奏点对应的音量信息, 并根据所述音频信号的波动拟合包络信号, 确定与所述目标节奏点对应的持续时间。

在音频信号中, 节拍用于表示强音和弱音规律性变化的周期, 节拍信息可以是指音乐中强音和弱音有规律地和周期性地循环反复的特征信息。一实施例中, 节拍信息包括每分钟节拍数目 (beat per minute, bpm)。其中, 音符一般以一拍为单位, 根据音频信号中的每分钟节拍数目可以确定一拍对应的时长, 映射到音频信号中, 可以确定一个节奏点 (音符) 对应的持续时间, 以及对应包括的多个信号点, 根据该多个信号点的时域幅值, 可以确定该节奏点的音量信息。

根据节拍信息可以确定节奏点对应的音符在音频信号中持续的信号区间, 本实施例中, 该信号区间是离散信号点形成的数组区间。从而根据信号区间中多个信号点的时域幅值, 确定与节奏点对应的音量信息。示例性的, 将信号区间中多个信号点的时域幅值的均值作为与节奏点对应的音量值。

波动拟合包络信号可以是指对音频信号在时域中的幅值特性进行拟合的信号, 而且, 波动拟合包络信号的波形特征也是指音频信号的时域幅值变化趋势特征。本实施例中, 波动拟合包络信号更加波动, 趋势拟合包络信号更加平滑, 趋势拟合包络信号可以是在波动拟合包络信号的基础上经过平滑操作处理完成。

一实施例中, 如图 1b 所示, 待识别的音频信号 101 的趋势拟合包络信号 102 比波动拟合包络信号 103 更加平缓。

根据节拍信息可以确定节奏点对应的音符在音频信号中持续的信号区间, 将节拍信息映射到波动拟合包络信号中得到与节拍信息对应的信号区间, 根据波动拟合包络信号的波形特征, 在上述确定的与节拍信息对应的信号区间中确定该音符的结束信号点, 进而将结束信号点对应的的时间作为该节奏点的终点时间, 并根据该节奏点的起点时间, 可以确定与该节奏点对应的持续时间。一实施例中, 该音符的结束信号点可以是与节拍信息对应的信号区间中任意一个波谷点。

一实施例中, 所述根据所述音频信号的节拍信息, 确定与所述目标节奏点对应的音量信息, 可以包括: 根据所述目标节奏点对应的起点时间以及所述音频信号的节拍信息, 确定所述目标节奏点匹配的音量区间; 根据所述音量区间中多个信号点的信号时域特征参数, 计算与所述目标节奏点对应的音量信息。

一实施例中, 根据音频信号的节拍信息, 确定与目标节奏点对应的音量信

息，可以是，将目标节奏点对应的起点时间作为起始端点，同时将音频信号的节拍信息中的一拍对应的时长作为区间长度，确定目标节奏点的音量区间。并根据音频信号在音量区间中多个信号点的时域幅值确定该目标节奏点的音量信息。示例性的，可以将音频信号在音量区间中多个信号点的时域幅值的均值作为该目标节奏点的音量值。此外，还可以是计算音频信号在音量区间中多个信号点的时域幅值的平方，将该多个信号点的时域幅值的平方中的最大值作为该目标节奏点的音量值，对此，本公开实施例不做限制。

本实施例中，bpm 可以通过复域谱差函数、光谱差函数和节拍强调函数等中至少一种进行计算，一实施例中，可以采用多种函数，并从 bpm 的计算结果中进行筛选，确定需要的 bpm。此外，还可以采用其他方法计算 bpm，对此，本公开实施例不做限制。

通过音频信号的节拍信息确定目标节奏点匹配的音量区间，可以准确确定目标节奏点的音量信息。

一实施例中，所述根据所述音频信号的波动拟合包络信号，确定与所述目标节奏点对应的持续时间，包括：将任意相邻两个目标节奏点按照对应的起点时间，映射至所述音频信号的波动拟合包络信号中，并根据所述波动拟合包络信号的波形特征，确定与所述相邻两个目标节奏点匹配的信号点对应的起点时间；将所述相邻两个目标节奏点中的首个目标节奏点对应的起点时间，以及与所述相邻两个目标节奏点匹配的信号点对应的起点时间之间的时长，作为所述相邻两个目标节奏点中的首个目标节奏点对应的持续时间。

首个目标节奏点是指，两个目标节奏点中起点时间在前的目标节奏点。任意一个节奏点的持续时间小于该节奏点对应的起点时间与相邻后一个节奏点对应的起点时间之间的时长。一实施例中，一个音符结束时，该音符的能量最小，此时反映在音频信号中即幅值最小，可以将相邻两个目标节奏点之间的波谷点作为两个目标节奏点匹配的信号点，并将波谷点对应的起点时间（实际是波谷点对应的时间点）与相邻两个目标节奏点中首个目标节奏点对应的起点时间之间的时长作为该首个目标节奏点对应的持续时间。而且音频信号的波动拟合包络信号比趋势拟合包络信号更符合音频信号的幅值变化情况，从而，可以根据音频信号的波动拟合包络信号确定相邻两个目标节奏点之间的波谷点。一实施例中，若一个信号点之前和之后的信号点的幅值均大于该信号点，该信号点即为波谷点。

通过波动拟合包络信号的波形特征确定目标节奏点的持续时间，可以准确找到相邻两个目标节奏点中首个目标节奏点对应的终点时间，从而准确确定首个目标节奏点对应的持续时间。

S140, 将与所述目标节奏点对应的起点时间、音量信息以及持续时间作为对所述音频信号的节奏点识别结果。

节奏是由多种不同时值的音符组合在一起,形成的多样化的形态,和音符的长短、强弱有着密切关系。为了表示节奏的特征,每个节奏点识别结果包括目标节奏点对应的起点时间、音量信息以及持续时间。

本公开实施例通过根据音频信号的频谱特性确定音频信号的至少一个备选节奏点,以及对应的起点时间,并根据音频信号的趋势拟合包络信号的波形特征从至少一个备选节奏点筛选目标节奏点,最后根据音频信号的波动拟合包络信号以及所述音频信号的节拍信息确定目标节奏点对应的音量信息和持续时间,确定目标节奏点的识别结果,解决了相关技术中人工标注节奏点的时间成本高效率低的问题,实现自动识别节奏点,而且对节奏点进行多次筛选,提高节奏点识别的准确率。

在上述实施例的基础上,在将与所述目标节奏点对应的起点时间、音量信息以及持续时间作为对所述音频信号的节奏点识别结果之后,还可以包括:在与所述目标节奏点对应的起点时间处,根据所述目标节奏点的音量信息和持续时间,添加与所述目标节奏点匹配的音乐特效。

在得到音频信号的节奏点识别结果之后,针对每个目标节奏点,在起点时间处,开始添加音乐特效,本实施例中,音乐特效的持续时间与该目标节奏点的持续时间相同,音乐特效的音量信息与该目标节奏点的音量信息匹配,例如,该目标节奏点的音量由 35 分贝逐渐衰减,添加的音乐特效的音量相应由 35 分贝逐渐衰减。此外,每个目标节奏点匹配的音乐特效可以相同也可以不同。

通过在识别音频信号中的节奏点之后,添加与目标节奏点匹配的音乐特效,为音频信号增加特殊效果,提高音频信号的丰富度。

实施例二

图 2a 为本公开实施例二提供的一种节奏点识别方法的流程图。本实施例以上述实施例中可选方案为基础进行说明。在本实施例中,根据待识别的音频信号的频谱特性,在所述音频信号中确定备选节奏点,并获取与所述备选节奏点对应的起点时间,可以包括:将所述音频信号中的多个信号点进行分组处理,确定多个分组,其中,每个分组中包括设定数量的相邻信号点,不同分组中包括的信号点相异或者部分重叠;根据每个分组中多个信号点的信号频域特征参数,计算与每个分组对应的分组频域特征参数;根据与每个分组对应的分组频域特征参数,以及预设的特征筛选条件,在所述多个分组中筛选出目标分

组，并根据目标分组中的多个信号点确定一个备选节奏点；在目标分组中的多个信号点对应的时间区间中，选择一个时间点作为与所述目标分组对应的备选节奏点的起点时间。

本实施例的方法可以包括：

S210，将音频信号中的多个信号点进行分组处理，确定多个分组。

本实施例中，每个分组中包括设定数量的相邻信号点，不同分组中包括的信号点相异或者部分重叠。

音频信号为离散信号，设定数量可以为 1024，分组处理可以是每隔 511 个信号点连续取 1024 个相邻信号点作为一组。一实施例中，将音频信号包括的离散信号按照时间顺序依次编号，第一个信号点的编号为 0，第二个信号点的编号为 1，以此类推，则第 1 分组对应的离散信号的编号为[0, 1023]，第 2 分组对应的离散信号的编号为[512, 512+1023]，第 3 分组对应的离散信号的编号为[1024, 1024+1023]，以此类推，本实施例中，每个数组中对应的数值为每个信号点对应的时域幅值。

本实施例中的音频信号、频谱特性、备选节奏点、起点时间、趋势拟合包络信号、节拍信息、音量信息、波动拟合包络信号和节奏点识别结果等均可以参考上述实施例中的描述。

S220，根据每个分组中多个信号点的信号频域特征参数，计算与所述每个分组对应的分组频域特征参数。

信号频域特征参数可以是指音频信号由时域信号转换为频域信号时得到的频域相位和频域幅值。分组频域特征参数可以是指每组对应的节奏点特征值，节奏点特征值用于识别节奏点。

一实施例中，傅里叶变换可以实现音频信号由时域信号到频域信号的转换，为了避免将音频信号中不同频率的信号混杂在一起难以分辨，提高音频信号的分辨率，对音频信号先进行加窗，即前述的分组处理和窗函数处理，实现将无限长的时间片段截断为多个短片段，再针对每个分组进行傅里叶变换。

一实施例中，在对音频信号中的多个信号点进行分组处理之后，在每个分组中，以编号居中的信号点为基准，将对称位置上的信号点的数据进行对调，并乘以预设的窗函数，再进行傅里叶变换。如前例，第 1 分组对应的离散信号的编号为[0,1023]，以编号为 512 的点作为基准，将编号为[0, 511]和编号为[512, 1023]的信号点对应的时域幅值进行对调，并乘以汉宁窗（hann）系数，得到傅里叶变换前的多个分组数据，在对每个分组数据进行傅里叶变换，得到每个分组中多个信号点对应的频域相位和频域幅值作为该分组中多个信号点的信号频

域特征参数。

根据每个分组中多个信号点的信号频域特征参数计算与每个分组对应的分组频域特征参数，可以是采用起始点（onset）检测方法进行计算。可以根据每个分组中多个信号点的信号频域特征参数，并基于如下公式计算每个分组中每个信号点的节奏点特征值：

$$\text{Onset}[i]=2 \times D[i] \times \sin((P[i]-2 \times P[i-1]+P[i-2]) \times 0.5)$$

$$\text{Onset}[i]=\text{Onset}[i] \times \text{Onset}[i]$$

本实施例中， i 表示第 i 个信号点， $\text{Onset}[i]$ 为第 i 个信号点的节奏点特征值， $D[i]$ 为第 i 个信号点的幅值， $P[i]$ 为第 i 个信号点的相位。若 $i-1$ 小于 0， $P[i-1]$ 为 0；若 $i-2$ 小于 0， $P[i-2]$ 为 0。每个分组对应的分组频域特征参数为该分组中多个信号点的节奏点特征值之和。

此外，可以对多个分组对应的分组频域特征参数进行归一化处理 and 窗口平滑处理，并根据处理后的结果修正每个分组频域特征参数。一实施例中，归一化处理是，每个分组对应的分组频域特征参数除以多个分组对应的分组频域特征参数中最大的分组频域特征参数；窗口平滑处理可以是无限脉冲响应（Infinite Impulse Response, IIR）平滑处理。一实施例，窗口平滑处理中的窗口为 5。

S230，根据与每个分组对应的分组频域特征参数，以及预设的特征筛选条件，在所述多个分组中筛选出目标分组，并根据目标分组中的多个信号点确定一个备选节奏点。

一实施例中，特征筛选条件可以包括至少一个筛选步骤，用于从多个分组中确定至少一个目标分组，同时每个目标分组确定一个备选节奏点，实现音频信号中节奏点初步识别。示例性的，特征筛选条件可以是超过设定阈值的分组频域特征参数对应的分组作为目标分组。此外，特征筛选条件还可以是其他条件，对此，本公开实施例不作限制。

一实施例中，根据与每个分组对应的分组频域特征参数，以及预设的特征筛选条件，在所述多个分组中筛选出目标分组，可以包括：将连续设定数量的分组作为一个分组集合，确定多个分组集合；针对每个分组集合，在确定所述每个分组集合满足频域特征阈值条件的情况下，将所述每个分组集合中的首个分组作为备选目标分组；从多个备选目标分组中剔除满足相邻剔除条件的备选目标分组，将剩下的备选目标分组作为目标分组。

一实施例中，频域特征阈值条件可以是限定分组集合中多个分组频域特征参数大小关系的条件。例如，一个分组集合包括 5 个分组，按照时间顺序依次编号，频域特征阈值条件如下：

$$\text{Onsets\_ma}[i] > \text{Onsets\_ma}[i+2] + 0.1 \times \left(\sum_{j=i}^{i+4} \text{Onsets\_ma}[j] \right) \div 5$$

本实施例中，该分组集合包括由 i 到 $i+4$ 共五个分组， $\text{Onsets\_ma}[i]$ 表示第 i 个分组的分组频域特征参数。当满足上述不等式时，该分组集合满足频域特征阈值条件，同时，将首个分组即 $\text{Onsets\_ma}[i]$ 作为备选目标分组。

此外还可以在确定分组集合的时候，对每个分组进行修正。一实施例中，在连续设定数量的分组作为一个分组集合之前，还可以包括：将低于设定阈值的分组频域特征参数修正为 0。通过对分组进行修正，根据修正后的分组确定分组集合，并根据频域特征阈值条件确定备选目标分组，减少对备选目标分组进行判断的数据量，从而提高筛选备选目标分组的效率。

相邻剔除条件可以是指限定备选目标分组之间的相邻关系的条件。一实施例中，如果两个节奏点的间隔时间极短，说明这两个节奏点在时间上是相邻的，通常，出现相邻的两个节奏点是由于噪音，而并非真正的节奏点，而且由于一个分组可以确定一个节奏点，因此，可以将相邻的分组从备选目标分组中剔除，实现对节奏点的识别。本实施例中，分组相邻是指两个及以上的分组的信号点对应的起点时间在时间上是相邻的，或者说两个及以上的分组的信号点对应的起点时间之间不存在其他分组的信号点对应的起点时间。

一实施例中，根据多个备选目标分组中首个信号点对应的起点时间，确定起点时间相邻的至少两个备选目标分组满足相邻剔除条件，并将所述至少两个备选目标分组剔除，将剩下的备选目标分组作为目标分组。

一实施例中，第 30 个备选目标分组中首个信号点对应的起点时间，与第 31 个备选目标分组中首个信号点对应的起点时间之间的区间中，不存在其他分组中首个信号点对应的起点时间，确定第 30 个备选目标分组和第 31 个备选目标分组满足相邻剔除条件。若第 32 个备选目标分组和第 31 个备选目标分组也满足相邻剔除条件，将第 30 个备选目标分组、第 31 个备选目标分组和第 32 个备选目标分组均剔除。本实施例中，其他分组并不是局限于备选目标分组，其他分组是指在前述进行分组时，形成的分组。

也就是说，经过相邻剔除条件筛选后的目标分组之间不存在相邻情况。

在一个具体的例子中，如图 2b 所示，根据趋势拟合包络信号 202 确定音频信号中的信号点 201 为目标节奏点。通过对多个分组分别进行阈值筛选和相邻筛选的两步筛选，最终确定目标分组，实现对节奏点的两步筛选，提高节奏点识别的准确性。

一实施例中，根据目标分组中的多个信号点确定一个备选节奏点可以是根

据预设规则从目标分组中的多个信号点中确认一个备选节奏点。一实施例中，将目标分组中的任意一个信号点作为备选节奏点。

S240，在目标分组中的多个信号点对应的时间区间中，选择一个时间点作为与目标分组对应的备选节奏点的起点时间。

时间区间可以是指目标分组中首个信号点对应的时间到目标分组中终点信号点对应的时间之间形成的区间。从该区间中选择一个时间点作为该目标分组对应的备选节奏点的起点时间，一实施例中，可以将首个信号点对应的时间点作为该备选节奏点对应的起点时间。

S250，将所述备选节奏点按照对应的起点时间，映射至所述音频信号的趋势拟合包络信号中，并根据所述趋势拟合包络信号的波形特征，在所述备选节奏点中确定目标节奏点。

S260，根据所述音频信号的节拍信息，确定与目标节奏点对应的音量信息，并根据所述音频信号的波动拟合包络信号，确定与目标节奏点对应的持续时间。S270，将与目标节奏点对应的起点时间、音量信息以及持续时间作为对所述音频信号的节奏点识别结果。

本公开实施例通过对音频信号进行分组处理，以及获取每个分组中多个信号点的信号频域特征参数，由此确定每个分组对应的分组频域特征参数，并根据分组频域特征参数对每个分组进行筛选，确定目标分组，对应每个目标分组确定一个备选节奏点，实现在确定备选节奏点之前对分组进行筛选，减少备选节奏点的数量，提高节奏点识别的效率和准确性。

实施例三

图 3 为本公开实施例三提供的一种节奏点识别方法的流程图。本实施例以上述实施例中可选方案为基础进行说明。

本实施例的方法可以包括：

S310，将所述音频信号中的多个信号点进行分组处理，确定多个分组。

本实施例中，每个分组中包括设定数量的相邻信号点，不同分组中包括的信号点相异或者部分重叠。

S320，根据每个分组中多个信号点的信号频域特征参数，计算与每个分组对应的分组频域特征参数。

S330，将连续设定数量的分组作为一个分组集合，确定多个分组集合。

S340，针对每个分组集合，在确定所述每个分组集合满足频域特征阈值条

件的情况下，将所述每个分组集合中的首个分组作为备选目标分组。

S350，从所述备选目标分组中剔除满足相邻剔除条件的备选目标分组，将剩下的备选目标分组作为目标分组。

S360，根据目标分组中的多个信号点确定一个备选节奏点。

S370，在目标分组中的多个信号点对应的时间区间中，选择一个时间点作为与所述每个目标分组对应的备选节奏点的起点时间。

S380，根据所述音频信号的趋势拟合包络信号的波形特征，识别所述趋势拟合包络信号中的波峰点。

S390，将所述备选节奏点按照对应的起点时间，映射至所述趋势拟合包络信号中，并将在时间上，与所述波峰点最近的备选节奏点作为目标节奏点。

S3100，根据所述音频信号的节拍信息，确定与所述目标节奏点对应的音量信息，并根据所述音频信号的波动拟合包络信号，确定与所述目标节奏点对应的持续时间。

S3110，将与所述目标节奏点对应的起点时间、音量信息以及持续时间作为对所述音频信号的节奏点识别结果。

S3120，在与所述目标节奏点对应的起点时间处，根据所述目标节奏点的音量信息和持续时间，添加与所述目标节奏点匹配的音乐特效。

实施例四

图 4 为本公开实施例提供的一种节奏点识别装置的结构示意图，本实施例可适用于识别一段音频信号中的节奏点的情况。该装置可以采用软件和/或硬件的方式实现，该装置可以配置于电子设备中。如图 4 所示，该装置可以包括：备选节奏点确定模块 410、目标节奏点确定模块 420、音量信息和持续时间确定模块 430 和节奏点识别结果确定模块 440。

备选节奏点确定模块 410，设置为根据待识别的音频信号的频谱特性，在所述音频信号中确定备选节奏点，并获取与所述备选节奏点对应的起点时间；目标节奏点确定模块 420，设置为将所述备选节奏点按照对应的起点时间，映射至所述音频信号的趋势拟合包络信号中，并根据所述趋势拟合包络信号的波形特征，在所述备选节奏点中确定目标节奏点；音量信息和持续时间确定模块 430，设置为根据所述音频信号的节拍信息，确定与所述目标节奏点对应的音量信息，并根据所述音频信号的波动拟合包络信号，确定与所述目标节奏点对应的持续时间；节奏点识别结果确定模块 440，设置为将与所述目标节奏点对应的起点时

间、音量信息以及持续时间作为对所述音频信号的节奏点识别结果。

本公开实施例通过根据音频信号的频谱特性确定音频信号的至少一个备选节奏点，以及对应的起点时间，并根据音频信号的趋势拟合包络信号的波形特征从至少一个备选节奏点筛选目标节奏点，最后根据音频信号的波动拟合包络信号以及所述音频信号的节拍信息确定目标节奏点对应的音量信息和持续时间，确定目标节奏点的识别结果，解决了相关技术中人工标注节奏点的时间成本效率低的问题，实现自动识别节奏点，而且对节奏点进行多次筛选，提高节奏点识别的准确率。

一实施例中，所述备选节奏点确定模块 410，包括：分组模块，设置为将所述音频信号中的多个信号点进行分组处理，确定多个分组，其中，每个分组中包括设定数量的相邻信号点，不同分组中包括的信号点相异或者部分重叠；频域特征参数计算模块，设置为根据每个分组中多个信号点的信号频域特征参数，计算与每个分组对应的分组频域特征参数；备选节奏点筛选模块，设置为根据与每个分组对应的分组频域特征参数，以及预设的特征筛选条件，在所述多个分组中筛选出目标分组，并根据目标分组中的多个信号点确定一个备选节奏点；起点时间确定模块，设置为在目标分组中的多个信号点对应的时间区间中，选择一个时间点作为与目标分组对应的备选节奏点的起点时间。

一实施例中，所述备选节奏点筛选模块，包括：分组集合确定模块，设置为将连续设定数量的分组作为一个分组集合，确定多个分组集合；备选目标分组确定模块，设置为针对每个分组集合，在确定所述每个分组集合中的分组对应的分组频域特征参数满足频域特征阈值条件的情况下，将所述每个分组集合中的首个分组作为备选目标分组；目标分组确定模块，设置为从多个备选目标分组中剔除满足相邻剔除条件的备选目标分组，将剩下的备选目标分组作为目标分组。

一实施例中，所述目标节奏点确定模块 420，包括：波峰点识别模块，设置为根据所述音频信号的趋势拟合包络信号的波形特征，识别所述趋势拟合包络信号中的波峰点；目标节奏点筛选模块，设置为将所述备选节奏点按照对应的起点时间，映射至所述趋势拟合包络信号中，并将在时间上，与所述波峰点最近的备选节奏点作为目标节奏点。

一实施例中，所述音量信息和持续时间确定模块 430，包括：音量区间确定模块，设置为根据所述目标节奏点对应的起点时间以及所述音频信号的节拍信息，确定所述目标节奏点匹配的音量区间；音量信息计算模块，设置为根据所述音量区间中多个信号点的信号时域特征参数，计算与目标节奏点对应的音量信息。

一实施例中，所述音量信息和持续时间确定模块 430，包括：终点时间确定模块，设置为将任意相邻两个目标节奏点按照对应的起点时间，映射至所述音频信号的波动拟合包络信号中，并根据所述波动拟合包络信号的波形特征，确定与所述相邻两个目标节奏点匹配的信号点的起点时间；持续时间计算模块，设置为将所述相邻两个目标节奏点中的首个目标节奏点对应的起点时间，以及与所述相邻两个目标节奏点匹配的信号点的起点时间之间的时长，作为所述相邻两个目标节奏点中的首个目标节奏点对应的持续时间。

一实施例中，所述节奏点识别装置，还包括：音乐特效添加模块，设置为在与所述目标节奏点对应的起点时间处，根据所述目标节奏点的音量信息和持续时间，添加与所述目标节奏点匹配的音乐特效。

本公开实施例提供的节奏点识别装置，与实施例一提供的节奏点识别方法属于同一发明构思，未在本公开实施例中详尽描述的技术细节可参见实施例一，并且本公开实施例与实施例一具有相同的有益效果。

实施例五

本公开实施例提供了一种电子设备，下面参考图 5，其示出了适于用来实现本公开实施例的电子设备（例如客户端或服务端）500 的结构示意图。本公开实施例中的电子设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、个人数字助理（Personal Digital Assistant, PDA）、平板电脑（Portable Android Device, PAD）、便携式多媒体播放器（Portable Media Player, PMP）、车载终端（例如车载导航终端）等等的移动终端以及诸如数字电视机（Television, TV）、台式计算机等等的固定终端。图 5 示出的电子设备仅仅是一个示例，不对本公开实施例的功能和使用范围带来任何限制。

如图 5 所示，电子设备 500 可以包括处理装置（例如中央处理器、图形处理器等）501，处理装置可以根据存储在只读存储器（Read-only Memory, ROM）502 中的程序或者从存储装置 508 加载到随机访问存储器（Random Access Memory, RAM）503 中的程序而执行至少一种适当的动作和处理。在 RAM 503 中，还存储有电子设备 500 操作所需的至少一种程序和数据。处理装置 501、ROM 502 以及 RAM 503 通过总线 504 彼此相连。输入/输出（Input/Output, I/O）接口 505 也连接至总线 504。

一实施例中，以下装置可以连接至 I/O 接口 505：包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置 506；包括例如液晶显示器（Liquid Crystal Display, LCD）、扬声器、振动器等等的输出装置 507；

包括例如磁带、硬盘等的存储装置 508；以及通信装置 509。通信装置 509 可以允许电子设备 500 与其他设备进行无线或有线通信以交换数据。虽然图 5 示出了具有多种装置的电子设备 500，并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。

根据本公开的实施例，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，该计算机程序产品包括承载在计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信装置 509 从网络上被下载和安装，或者从存储装置 508 被安装，或者从 ROM 502 被安装。在该计算机程序被处理装置 501 执行时，执行本公开实施例的方法中限定的上述功能。

实施例六

本公开实施例还提供了一种计算机可读存储介质，计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是一——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、RAM、ROM、可擦式可编程只读存储器（Erasable Programmable Read Only Memory，EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（Compact Disc Read-Only Memory，CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开中，计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，计算机可读信号介质中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：电线、光缆、射频（Radio Frequency，RF）等等，或者上述的任意合适的组合。

上述计算机可读介质可以是上述电子设备中所包含的；也可以是单独存在，而未装配入该电子设备中。

上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序

被该电子设备执行时，使得该电子设备：根据待识别的音频信号的频谱特性，在所述音频信号中确定备选节奏点，并获取与所述备选节奏点对应的起点时间；将所述备选节奏点按照对应的起点时间，映射至所述音频信号的趋势拟合包络信号中，并根据所述趋势拟合包络信号的波形特征，在所述备选节奏点中确定目标节奏点；根据所述音频信号的节拍信息，确定与所述目标节奏点对应的音量信息，并根据所述音频信号的波动拟合包络信号，确定与所述目标节奏点对应的持续时间；将与所述目标节奏点对应的起点时间、音量信息以及持续时间作为对所述音频信号的节奏点识别结果。

可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码，上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++，还包括常规的过程式程序设计语言—诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户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。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网(Local Area Network, LAN)或广域网(Wide Area Network, WAN)—连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

附图中的流程图和框图，图示了本公开至少一种实施例的系统、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。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，该模块、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。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。

描述于本公开实施例中所涉及到的模块可以通过软件的方式实现，也可以通过硬件的方式来实现。其中，模块的名称在某种情况下并不构成对该模块本身的限定，例如，备选节奏点确定模块还可以被描述为“根据待识别的音频信号的频谱特性，在所述音频信号中确定备选节奏点，并获取与所述备选节奏点对应的起点时间的模块”。

权利要求书

1、一种节奏点识别方法，包括：

根据待识别的音频信号的频谱特性，在所述音频信号中确定备选节奏点，并获取与所述备选节奏点对应的起点时间；

将所述备选节奏点按照对应的起点时间，映射至所述音频信号的趋势拟合包络信号中，并根据所述趋势拟合包络信号的波形特征，在所述备选节奏点中确定目标节奏点；

根据所述音频信号的节拍信息，确定与所述目标节奏点对应的音量信息，并根据所述音频信号的波动拟合包络信号，确定与所述目标节奏点对应的持续时间；

将与所述目标节奏点对应的起点时间、音量信息以及持续时间作为对所述音频信号的节奏点识别结果。

2、根据权利要求1所述的方法，其中，所述根据待识别的音频信号的频谱特性，在所述音频信号中确定备选节奏点，并获取与所述备选节奏点对应的起点时间，包括：

将所述音频信号中的多个信号点进行分组处理，确定多个分组，其中，每个分组中包括设定数量的相邻信号点，不同分组中包括的信号点相异或者部分重叠；

根据每个分组中多个信号点的信号频域特征参数，计算与所述每个分组对应的分组频域特征参数；

根据与每个分组对应的分组频域特征参数，以及预设的特征筛选条件，在所述多个分组中筛选出目标分组，并根据所述目标分组中的多个信号点确定一个备选节奏点；

在所述目标分组中的多个信号点对应的时间区间中，选择一个时间点作为与所述目标分组对应的备选节奏点的起点时间。

3、根据权利要求2所述的方法，其中，所述根据与每个分组对应的分组频域特征参数，以及预设的特征筛选条件，在所述多个分组中筛选出目标分组，包括：

将连续设定数量的分组作为一个分组集合，确定多个分组集合；

针对每个分组集合，在确定所述每个分组集合中的分组对应的分组频域特征参数满足频域特征阈值条件的情况下，将所述每个分组集合中的首个分组作为备选目标分组；

从多个备选目标分组中剔除满足相邻剔除条件的备选目标分组，将剩下的

备选目标分组作为目标分组。

4、根据权利要求 1-3 任一项所述的方法，其中，所述将所述备选节奏点按照对应的起点时间，映射至所述音频信号的趋势拟合包络信号中，并根据所述趋势拟合包络信号的波形特征，在所述备选节奏点中确定目标节奏点，包括：

根据所述音频信号的趋势拟合包络信号的波形特征，识别所述趋势拟合包络信号中的波峰点；

将所述备选节奏点按照对应的起点时间，映射至所述趋势拟合包络信号中，并将在时间上，与所述波峰点最近的备选节奏点作为目标节奏点。

5、根据权利要求 1-4 任一项所述的方法，其中，所述根据所述音频信号的节拍信息，确定与所述目标节奏点对应的音量信息，包括：

根据所述目标节奏点对应的起点时间以及所述音频信号的节拍信息，确定所述目标节奏点匹配的音量区间；

根据所述音量区间中多个信号点的信号时域特征参数，计算与所述目标节奏点对应的音量信息。

6、根据权利要求 1-5 任一项所述的方法，其中，所述根据所述音频信号的波动拟合包络信号，确定与所述目标节奏点对应的持续时间，包括：

将任意相邻两个目标节奏点按照对应的起点时间，映射至所述音频信号的波动拟合包络信号中，并根据所述波动拟合包络信号的波形特征，确定与所述相邻两个目标节奏点匹配的信号点的起点时间；

将所述相邻两个目标节奏点中的首个目标节奏点对应的起点时间，以及与所述相邻两个目标节奏点匹配的信号点的起点时间之间的时长，作为所述相邻两个目标节奏点中的首个目标节奏点对应的持续时间。

7、根据权利要求 1-6 任一项所述的方法，，在所述将与所述目标节奏点对应的起点时间、音量信息以及持续时间作为对所述音频信号的节奏点识别结果之后，还包括：

在与所述目标节奏点对应的起点时间处，根据所述目标节奏点的音量信息和持续时间，添加与所述目标节奏点匹配的音乐特效。

8、一种节奏点识别装置，包括：

备选节奏点确定模块，设置为根据待识别的音频信号的频谱特性，在所述音频信号中确定备选节奏点，并获取与所述备选节奏点对应的起点时间；

目标节奏点确定模块，设置为将所述备选节奏点按照对应的起点时间，映射至所述音频信号的趋势拟合包络信号中，并根据所述趋势拟合包络信号的波

形特征，在所述备选节奏点中确定目标节奏点；

音量信息和持续时间确定模块，设置为根据所述音频信号的节拍信息，确定与所述目标节奏点对应的音量信息，并根据所述音频信号的波动拟合包络信号，确定与所述目标节奏点对应的持续时间；

节奏点识别结果确定模块，设置为将与目标节奏点对应的起点时间、音量信息以及持续时间作为对所述音频信号的节奏点识别结果。

9、根据权利要求8所述的装置，其中，所述备选节奏点确定模块，包括：

分组模块，设置为将所述音频信号中的多个信号点进行分组处理，确定多个分组，其中，每个分组中包括设定数量的相邻信号点，不同分组中包括的信号点相异或者部分重叠；

频域特征参数计算模块，设置为根据每个分组中多个信号点的信号频域特征参数，计算与每个分组对应的分组频域特征参数；

备选节奏点筛选模块，设置为根据与每个分组对应的分组频域特征参数，以及预设的特征筛选条件，在所述多个分组中筛选出目标分组，并根据所述目标分组中的多个信号点确定一个备选节奏点；

起点时间确定模块，设置为在所述目标分组中的多个信号点对应的时间区间中，选择一个时间点作为与所述目标分组对应的备选节奏点的起点时间。

10、根据权利要求9所述的装置，其中，所述备选节奏点筛选模块，包括：

分组集合确定模块，设置为将连续设定数量的分组作为一个分组集合，确定多个分组集合；

备选目标分组确定模块，设置为针对每个分组集合，在确定所述每个分组集合中的分组对应的分组频域特征参数满足频域特征阈值条件的情况下，将所述每个分组集合中的首个分组作为备选目标分组；

目标分组确定模块，设置为从多个备选目标分组中剔除满足相邻剔除条件的备选目标分组，将剩下的备选目标分组作为目标分组。

11、根据权利要求8-10任一项所述的装置，其中，所述目标节奏点确定模块，包括：

波峰点识别模块，设置为根据所述音频信号的趋势拟合包络信号的波形特征，识别所述趋势拟合包络信号中的波峰点；

目标节奏点筛选模块，设置为将每个备选节奏点按照对应的起点时间，映射至所述趋势拟合包络信号中，并将在时间上，与所述波峰点最近的备选节奏点作为目标节奏点。

12、根据权利要求 8-11 任一项所述的装置，其中，所述音量信息和持续时间确定模块，包括：

音量区间确定模块，设置为根据所述目标节奏点对应的起点时间以及所述音频信号的节拍信息，确定所述目标节奏点匹配的音量区间；

音量信息计算模块，设置为根据所述音量区间中多个信号点的信号时域特征参数，计算与所述目标节奏点对应的音量信息。

13、根据权利要求 8-12 任一项所述的装置，其中，所述音量信息和持续时间确定模块，包括：

终点时间确定模块，设置为将任意相邻两个目标节奏点按照对应的起点时间，映射至波动拟合包络信号中，并根据所述波动拟合包络信号的波形特征，确定与所述相邻两个目标节奏点匹配的信号点对应的起点时间；

持续时间计算模块，设置为将所述相邻两个目标节奏点中的首个目标节奏点对应的起点时间，以及与所述相邻两个目标节奏点匹配的信号点对应的起点时间之间的时长，作为所述相邻两个目标节奏点中的首个目标节奏点对应的持续时间。

14、根据权利要求 8-13 任一项所述的装置，还包括：

音乐特效添加模块，设置为在与所述目标节奏点对应的起点时间处，根据所述目标节奏点的音量信息和持续时间，添加与所述目标节奏点匹配的音乐特效。

15、一种电子设备，包括：

至少一个处理器；

存储装置，设置为存储至少一个程序；

当所述至少一个程序被所述至少一个处理器执行，使得所述至少一个处理器实现如权利要求 1-7 任一所述的节奏点识别方法。

16、一种计算机可读存储介质，存储有计算机程序，所述计算机程序被处理器执行时实现如权利要求 1-7 任一所述的节奏点识别方法。

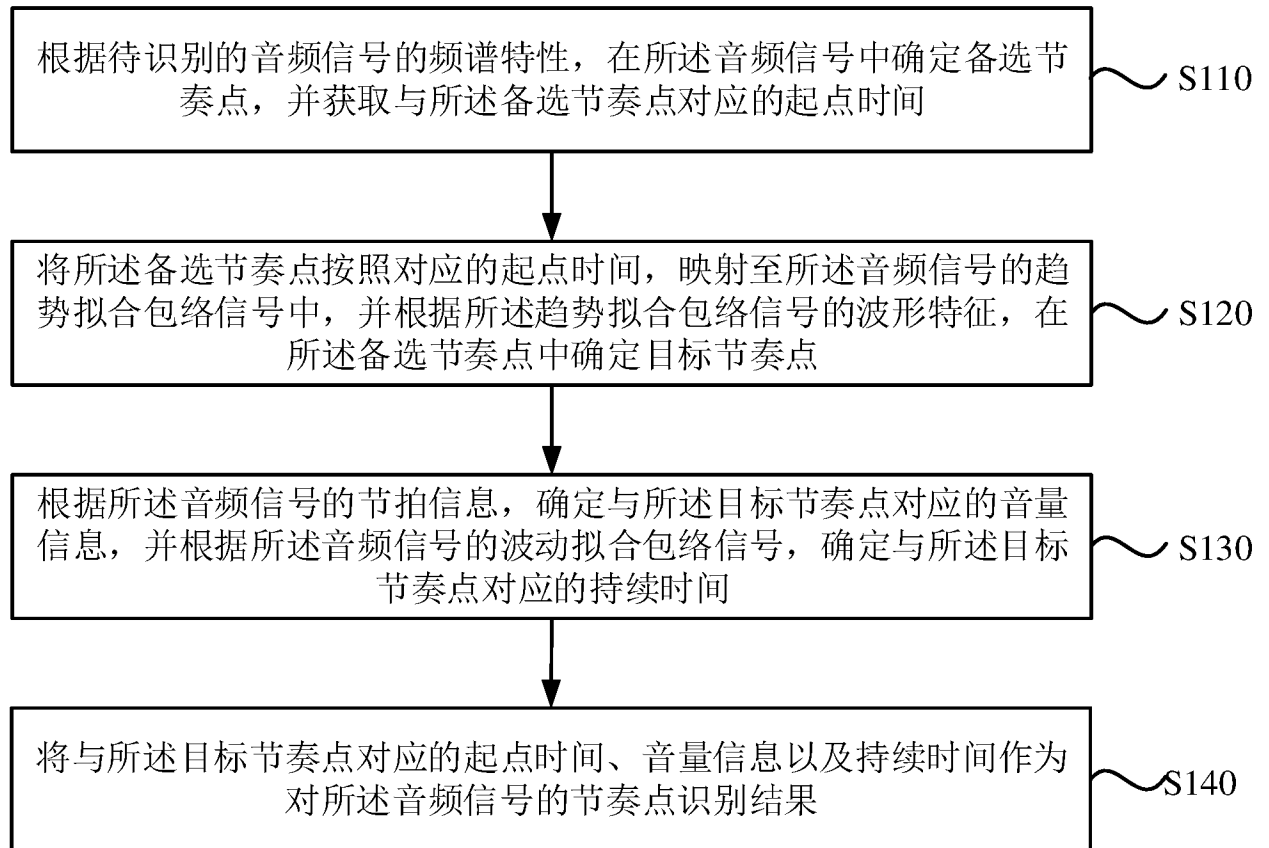


图 1a

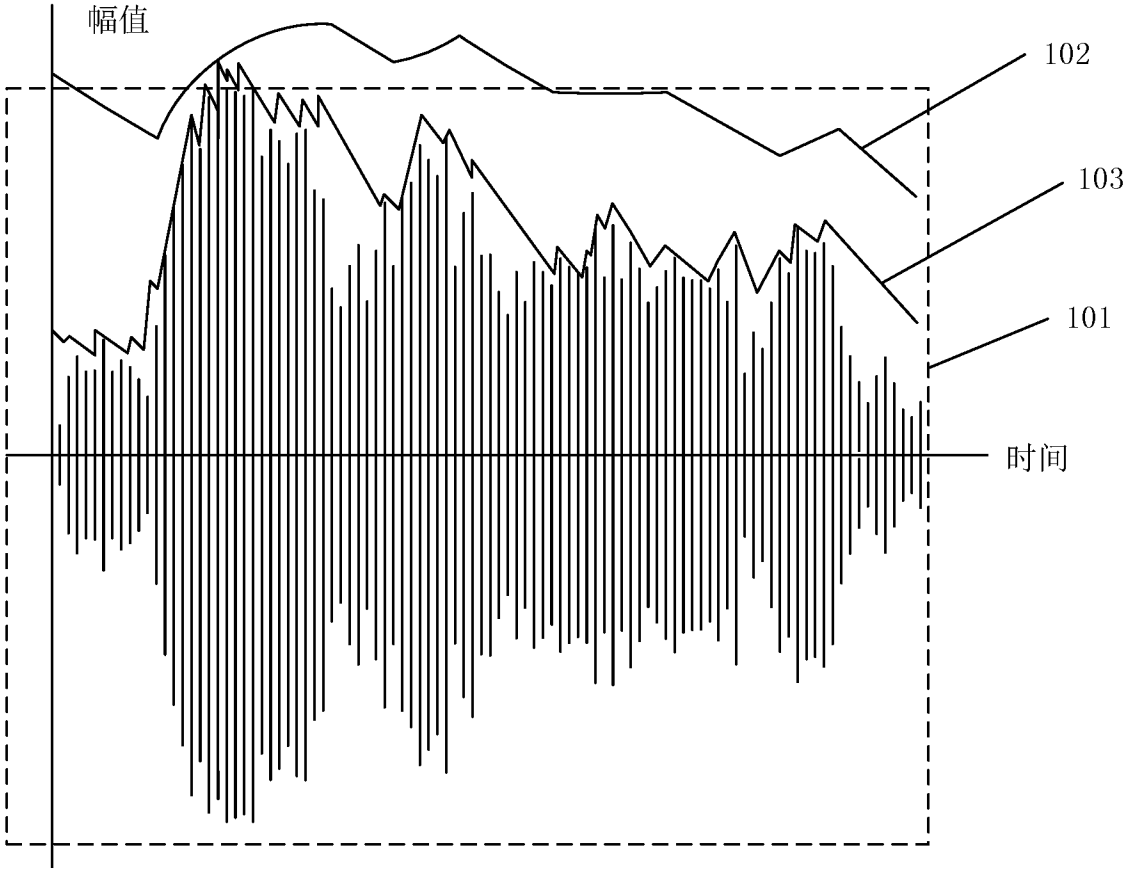


图 1b

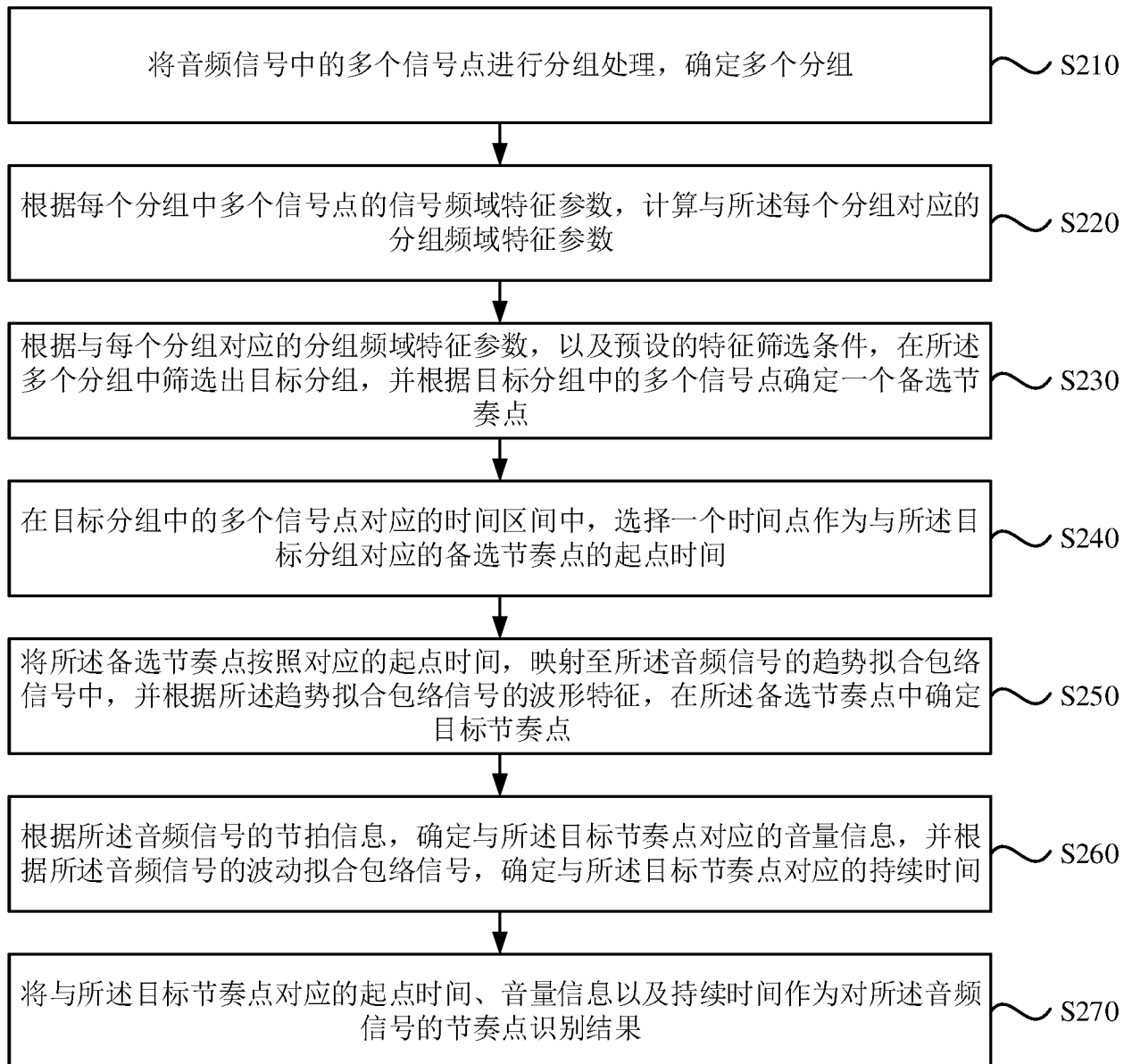


图 2a

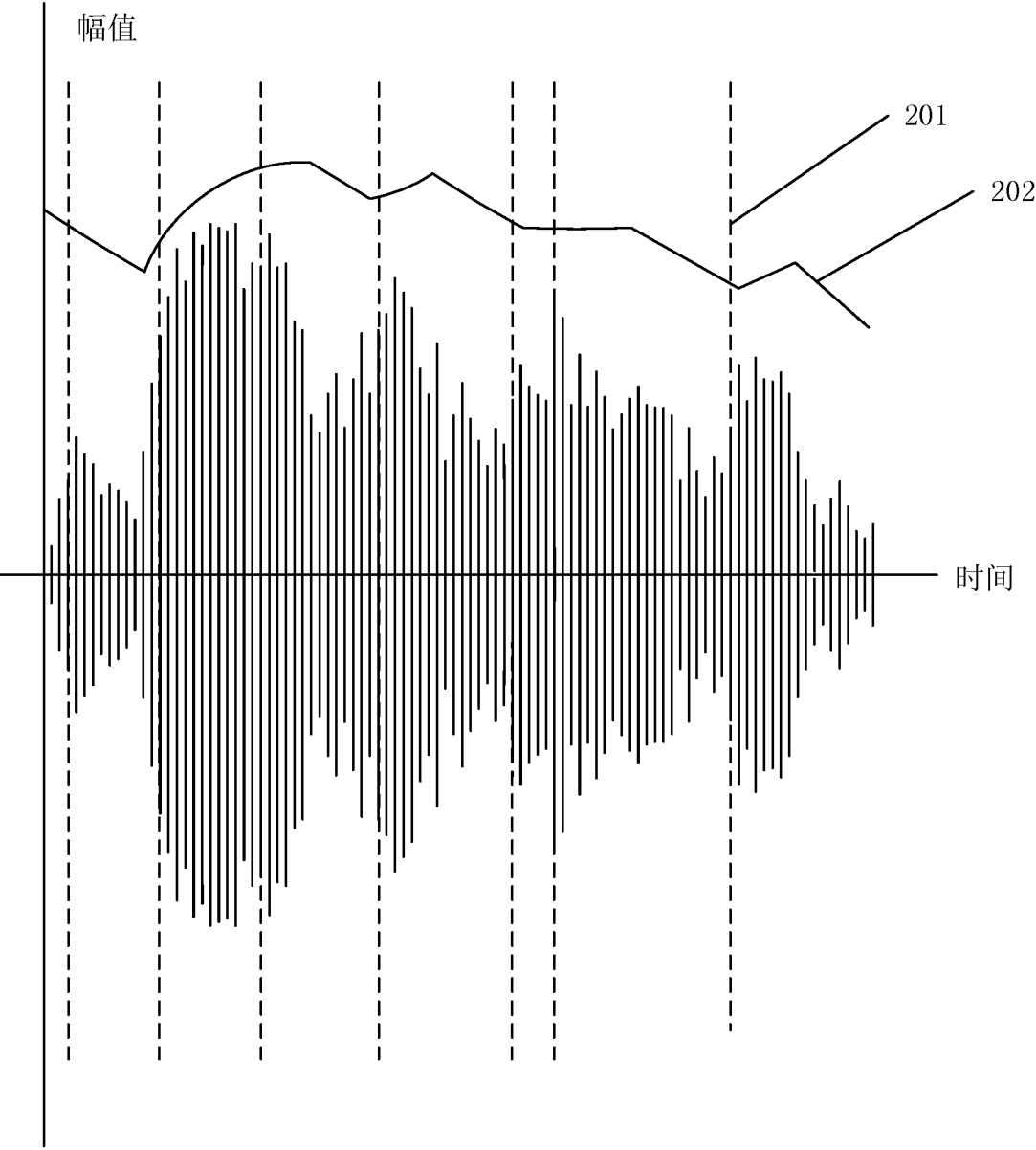


图 2b

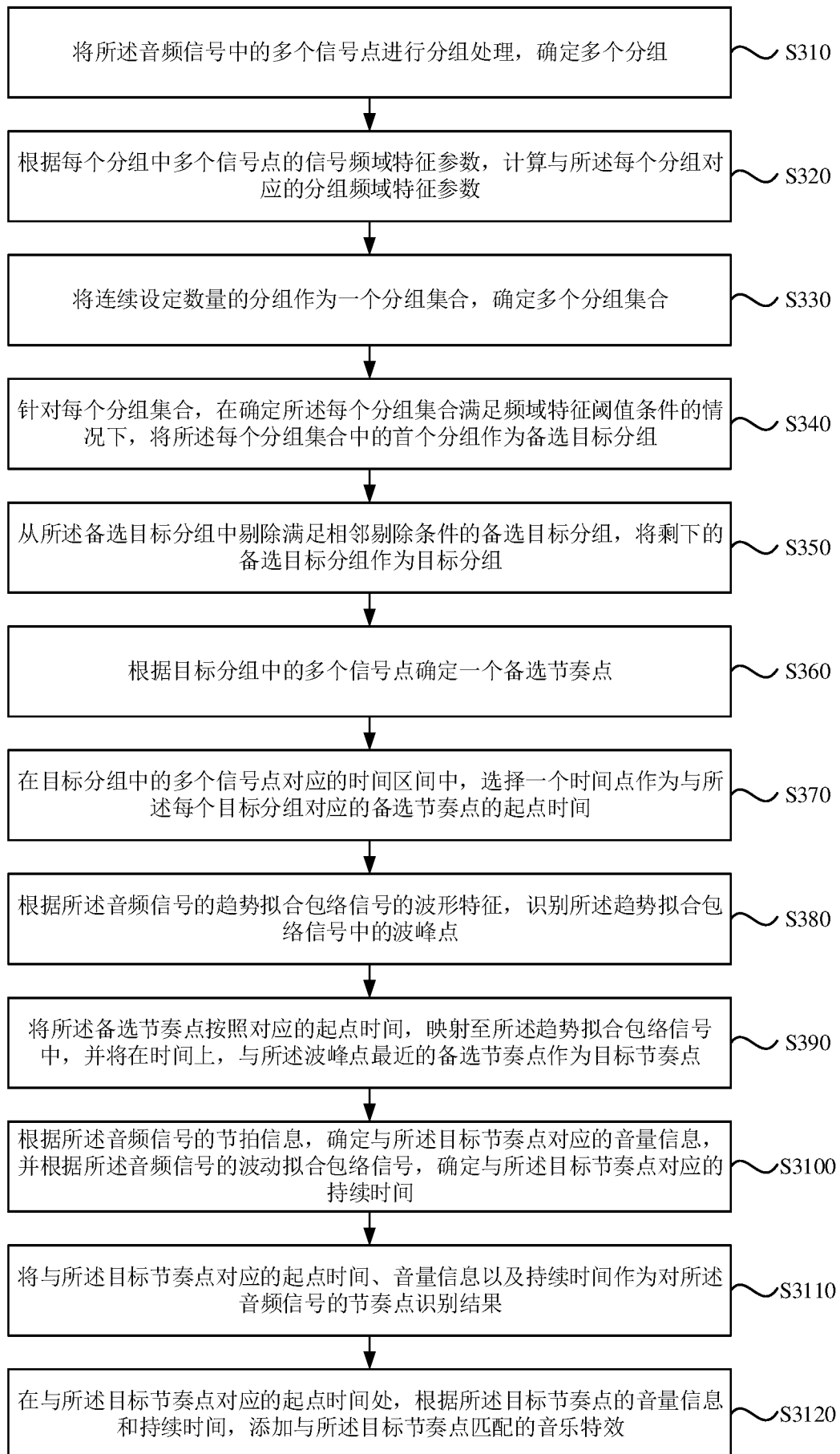


图 3

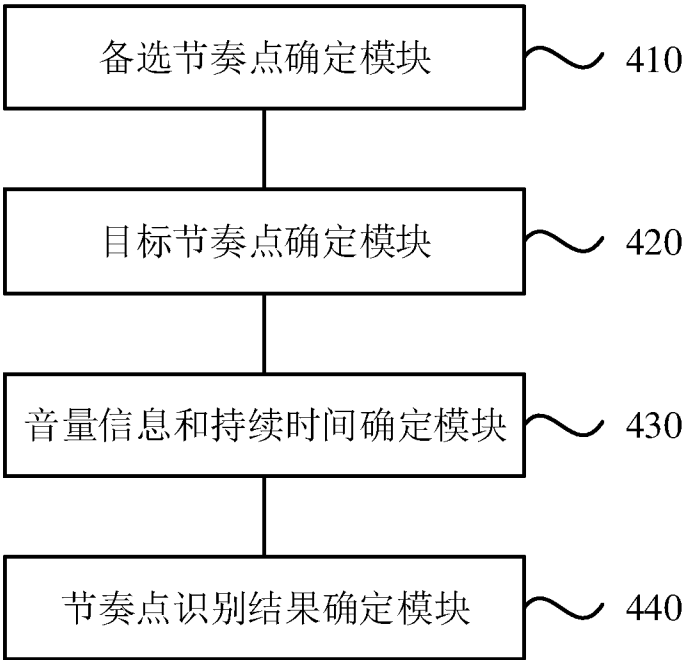


图 4

500

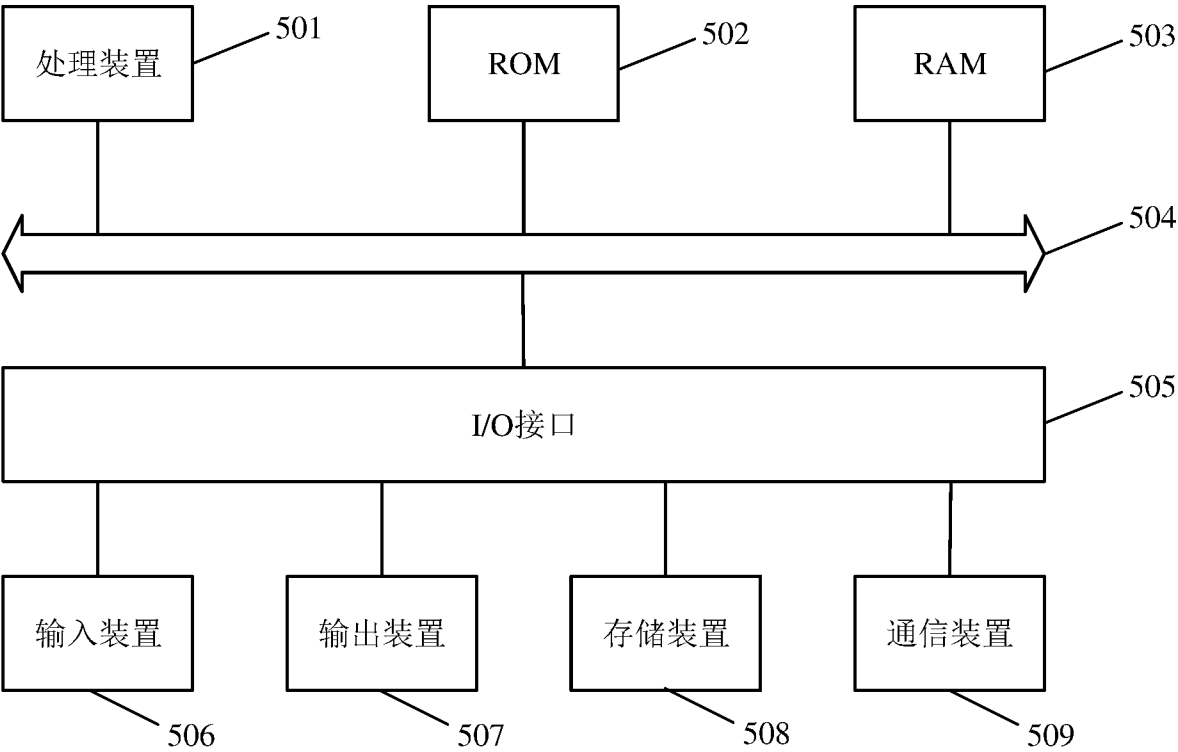


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/099640

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/632(2019.01)i; G06F 16/63(2019.01)i; G10L 25/51(2013.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F; G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT; CNKI; EPODOC; WPI; GOOGLE: 节奏, 节拍, 识别, 频谱, 备选, 多个, 起点, 时间, 包络, 目标, 音量, tempo, beat, identify, frequency, multiple, start point, time, envelope, curve, volume

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109670074 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 23 April 2019 (2019-04-23) claims 1-16	1-16
X	CN 108364660 A (TENCENT MUSIC ENTERTAINMENT TECHNOLOGY (SHENZHEN) CO., LTD.) 03 August 2018 (2018-08-03) description, the abstract, paragraphs [0005]-[0045], and figure 1	1-16
A	CN 108320730 A (GUANGZHOU BIGO INFORMATION TECHNOLOGY CO., LTD.) 24 July 2018 (2018-07-24) entire document	1-16
A	JP 3674950 B2 (YAMAHA CORPORATION) 27 July 2005 (2005-07-27) entire document	1-16



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

23 September 2019

Date of mailing of the international search report

28 October 2019

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/099640

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109670074	A	23 April 2019	None			
CN	108364660	A	03 August 2018	None			
CN	108320730	A	24 July 2018	WO	2019137115	A1	18 July 2019
JP	3674950	B2	27 July 2005	None			

国际检索报告

国际申请号

PCT/CN2019/099640

A. 主题的分类 G06F 16/632(2019.01)i; G06F 16/63(2019.01)i; G10L 25/51(2013.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类																	
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G06F; G10L 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNPAT;CNKI;EPDOC;WPI;GOOGLE: 节奏, 节拍, 识别, 频谱, 备选, 多个, 起点, 时间, 包络, 目标, 音量, tempo, beat, identify, frequency, multiple, start point, time, envelope, curve, volume																	
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 109670074 A (北京字节跳动网络技术有限公司) 2019年 4月 23日 (2019 - 04 - 23) 权利要求1-16</td> <td>1-16</td> </tr> <tr> <td>X</td> <td>CN 108364660 A (腾讯音乐娱乐科技深圳有限公司) 2018年 8月 3日 (2018 - 08 - 03) 说明书摘要, 第[0005]-[0045]段, 附图1</td> <td>1-16</td> </tr> <tr> <td>A</td> <td>CN 108320730 A (广州市百果园信息技术有限公司) 2018年 7月 24日 (2018 - 07 - 24) 全文</td> <td>1-16</td> </tr> <tr> <td>A</td> <td>JP 3674950 B2 (ヤマハ株式会社) 2005年 7月 27日 (2005 - 07 - 27) 全文</td> <td>1-16</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 109670074 A (北京字节跳动网络技术有限公司) 2019年 4月 23日 (2019 - 04 - 23) 权利要求1-16	1-16	X	CN 108364660 A (腾讯音乐娱乐科技深圳有限公司) 2018年 8月 3日 (2018 - 08 - 03) 说明书摘要, 第[0005]-[0045]段, 附图1	1-16	A	CN 108320730 A (广州市百果园信息技术有限公司) 2018年 7月 24日 (2018 - 07 - 24) 全文	1-16	A	JP 3674950 B2 (ヤマハ株式会社) 2005年 7月 27日 (2005 - 07 - 27) 全文	1-16
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
PX	CN 109670074 A (北京字节跳动网络技术有限公司) 2019年 4月 23日 (2019 - 04 - 23) 权利要求1-16	1-16															
X	CN 108364660 A (腾讯音乐娱乐科技深圳有限公司) 2018年 8月 3日 (2018 - 08 - 03) 说明书摘要, 第[0005]-[0045]段, 附图1	1-16															
A	CN 108320730 A (广州市百果园信息技术有限公司) 2018年 7月 24日 (2018 - 07 - 24) 全文	1-16															
A	JP 3674950 B2 (ヤマハ株式会社) 2005年 7月 27日 (2005 - 07 - 27) 全文	1-16															
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																	
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																	
国际检索实际完成的日期 2019年 9月 23日		国际检索报告邮寄日期 2019年 10月 28日															
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		受权官员 万沙沙 电话号码 86-(10)-53961576															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/099640

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	109670074	A	2019年 4月 23日	无	
CN	108364660	A	2018年 8月 3日	无	
CN	108320730	A	2018年 7月 24日	W0 2019137115	A1 2019年 7月 18日
JP	3674950	B2	2005年 7月 27日	无	