(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 4 月 30 日 (30.04.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/082573 A1

(51) 国际专利分类号:
G10H 1/32 (2006.01)

(21) 国际申请号: PCT/CN2018/123549

(22) 国际申请日: 2018 年 12 月 25 日 (25.12.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811257165.1 2018年10月26日 (26.10.2018) CN

(71) 申请人: 平安科技（深圳）有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(72) 发明人: 刘鼻智 (LIU, Aozhi); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。 王义文 (WANG, Yiwen); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。 王健宗 (WANG, Jianzong); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。 肖京 (XIAO, Jing); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(74) 代理人: 北京汇思诚业知识产权代理有限公司 (UNI-INTEL PATENT AND TRADEMARK

WO 2020/082573 A1

(54) **Title:** LONG-SHORT-TERM NEURAL NETWORK-BASED MULTI-PART MUSIC GENERATION METHOD AND DEVICE

(54) 发明名称: 基于长短时神经网络的多声部音乐生成方法及装置

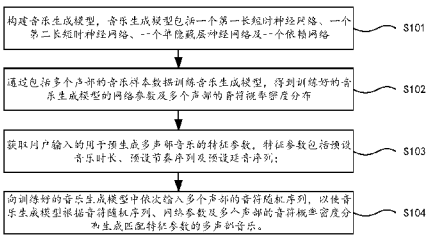

图 1

S101 Construct a music generation model, the music generation model comprising a first long-short-term neural network, a second long-short-term neural network, a single hidden layer neural network and a dependent network

S102 Train the music generation model by means of music sample data including a plurality of parts, to obtain network parameters of the trained music generation model and a music note probability density distribution of the plurality of parts

S103 Acquire characteristic parameters inputted by a user for pre-generating multi-part music, the characteristic parameters comprising a pre-set music duration, a pre-set rhythm sequence and a pre-set damper sequence

S104 Sequentially input, into the trained music generation model, a music note random sequence of the plurality of parts, so that the music generation model generates, according to the music note random sequence, the network parameters, and the music note probability density distribution of the plurality of parts, the multi-part music matching the characteristic parameters

(57) **Abstract:** A long-short-term neural network-based multi-part music generation method and device. Said method comprises: constructing a music generation model, the music generation model comprising a first long-short-term neural network, a second long-short-term neural network, a single hidden layer neural network and a dependent network (S101); training the music generation model by means of music sample data including a plurality of parts, to obtain network parameters of the trained music generation model and a music note probability density distribution of the plurality of parts (S102); acquiring characteristic parameters inputted by a user

[见续页]

LAW FIRM)；中国北京市海淀区高梁桥斜街59号中坤大厦603室, Beijing 100044 (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, 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))。

for pre-generating multi-part music, the characteristic parameters comprising a pre-set music duration, a pre-set rhythm sequence and a pre-set damper sequence (S103); and sequentially inputting, into the trained music generation model, a music note random sequence of the plurality of parts, so that the music generation model generates, according to the music note random sequence, the network parameters, and the music note probability density distribution of the plurality of parts, the multi-part music matching the characteristic parameters (S104).

(57) 摘要：基于长短时神经网络的多声部音乐生成方法及装置，该方法包括：构建音乐生成模型，音乐生成模型包括一个第一长短时神经网络、一个第二长短时神经网络、一个单隐藏层神经网络及一个依赖网络（S101）；通过包括多个声部的音乐样本数据训练音乐生成模型，得到训练好的音乐生成模型的网络参数及多个声部的音符概率密度分布（S102）；获取用户输入的用于预生成多声部音乐的特征参数，特征参数包括预设音乐时长、预设节奏序列及预设延音序列（S103）；向训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使音乐生成模型根据音符随机序列、网络参数及多个声部的音符概率密度分布生成匹配特征参数的多声部音乐（S104）。

基于长短时神经网络的多声部音乐生成方法及装置

本申请要求于2018年10月26日提交中国专利局、申请号为201811257165.1、申请名称为“基于长短时神经网络的多声部音乐生成方法及装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及人工智能技术领域，尤其涉及一种基于长短时神经网络的多声部音乐生成方法及装置。

背景技术

音乐通常由多个音轨组成，并具有各自的时间动态，音乐会随着时间的推移而相互依存地展开。自然语言生成和单音音乐生成的成功不容易普及到多音轨的音乐。现有的音乐生成方法通常是单旋律音乐，因为多个旋律之间的复杂的相互影响，很难生成多旋律的音乐。

因此，如何生成多个声部之间协调的音乐成为目前亟待解决的问题。

申请内容

有鉴于此，本申请实施例提供了一种基于长短时神经网络的多声部音乐生成方法及装置，用以解决现有技术中难以生成多个声部之间协调的音乐的问题。

为了实现上述目的，根据本申请的一个方面，提供了一种基于长短时神经网络的多声部音乐生成方法，所述方法包括：构建音乐生成模型，所述音乐生成模型包括一个第一长短时神经网络、一个第二长短时神经网络、一个单隐藏层神经网络及一个依赖网络；通过包括多个声部的

音乐样本数据训练所述音乐生成模型，得到训练好的所述音乐生成模型的网络参数及所述多个声部的音符概率密度分布；获取用户输入的用于预生成多声部音乐的特征参数，所述特征参数包括预设音乐时长、预设节奏序列及预设延音序列；向所述训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使所述音乐生成模型根据所述音符随机序列、所述网络参数及所述多个声部的音符概率密度分布生成匹配所述特征参数的多声部音乐。

为了实现上述目的，根据本申请的一个方面，提供了一种基于长短时神经网络的多声部音乐生成装置，所述装置包括：构建单元，用于构建音乐生成模型，所述音乐生成模型包括一个第一长短时神经网络、一个第二长短时神经网络、一个单隐藏层神经网络及一个依赖网络；第一获取单元，用于通过包括多个声部的音乐样本数据训练所述音乐生成模型，得到训练好的所述音乐生成模型的网络参数及所述多个声部的音符概率密度分布；第二获取单元，用于获取用户输入的用于预生成多声部音乐的特征参数，所述特征参数包括预设音乐时长、预设节奏序列及预设延音序列；生成单元，用于向所述训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使所述音乐生成模型根据所述音符随机序列、所述网络参数及所述多个声部的音符概率密度分布生成匹配所述特征参数的多声部音乐。

为了实现上述目的，根据本申请的一个方面，提供了一种计算机非易失性存储介质，所述存储介质包括存储的程序，在所述程序运行时控制所述存储介质所在设备执行上述的基于长短时神经网络的多声部音乐生成方法的步骤。

为了实现上述目的，根据本申请的一个方面，提供了一种计算机设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现上述的基于长短时神经网络的多声部音乐生成方法的步骤。

在本方案中，通过构建包括长短时神经网络的音乐生成模型，利用长短时神经网络直接对音符序列进行处理，同时能够利用序列时间前

后之间的相关性，得到音符概率密度分布；从而调整多个声部的音符序列，生成多声部之间协调的音乐，从而解决现有技术中难以生成多个声部之间协调的音乐的问题。

附图说明

为了更清楚地说明本申请实施例的技术方案，下面将对实施例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申请的一些实施例，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。

图 1 是根据本申请实施例的一种基于长短时神经网络的多声部音乐生成方法的流程图；

图 2 是根据本申请实施例的一种音乐生成模型的示意图；

图 3 是根据本申请实施例的一种基于长短时神经网络的多声部音乐生成装置的示意图；

图 4 是根据本申请实施例的一种计算机设备的示意图。

具体实施方式

为了更好的理解本申请的技术方案，下面结合附图对本申请实施例进行详细描述。

应当明确，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其它实施例，都属于本申请保护的范围。

在本申请实施例中使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本申请。在本申请实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。

图 1 是根据本申请实施例的一种基于长短时神经网络的多声部音乐

生成方法的流程图，如图 1 所示，该方法包括：

步骤 S101，构建音乐生成模型，音乐生成模型包括一个第一长短时神经网络、一个第二长短时神经网络、一个单隐藏层神经网络及一个依赖网络；

步骤 S102，通过包括多个声部的音乐样本数据训练音乐生成模型，得到训练好的音乐生成模型的网络参数及多个声部的音符概率密度分布；

步骤 S103，获取用户输入的用于预生成多声部音乐的特征参数，特征参数包括预设音乐时长、预设节奏序列及预设延音序列；

步骤 S104，向训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使音乐生成模型根据音符随机序列、网络参数及多个声部的音符概率密度分布生成匹配特征参数的多声部音乐。

在本方案中，通过构建包括长短时神经网络的音乐生成模型，利用长短时神经网络直接对音符序列进行处理，同时能够利用序列时间前后之间的相关性，得到音符概率密度分布；从而调整多个声部的音符序列，生成多声部之间协调的音乐，从而解决现有技术中难以生成多个声部之间协调的音乐的问题。

可选地，在通过包括多个声部的音乐样本数据训练音乐生成模型之前，方法还包括：获取多个音乐训练样本，其中，音乐训练样本包括多个声部信息；提取每个声部的音符序列、音乐训练样本的节奏序列及延音序列；其中，每个声部的音符序列表示为：$V_i = \{V_i^t\}$，$t \in T$，T 为音乐训练样本的时长，是十六分音符的整数倍；i 为声部；V_i^t 为当前时刻 t 的音符；将多个声部的音符序列、音乐训练样本的节奏序列及延音序列作为音乐样本数据。

可以理解地，每首曲子包括多个声部的音符序列、这个曲子的节奏序列及延音序列。将每首曲子随时间序列化处理，有利于长短时神经网络学习音符之间随时间尺度的依赖关系。

例如，收集 389 首众赞歌的 midi 数据，其中，每首曲子包含四个声部：女高音、女低音、男高音和男低音。相对音高较低的女低音、男高音和男低音给音高最高的女高音作伴奏。将其中 80%的 midi 数据用于音乐

训练样本，其中 20%的 midi 数据用于音乐评估样本。

可选地，获取多个音乐训练样本之后，方法还包括：剔除在一个声部内，有两个及以上的音符同时出现的曲子。

可选地，音符序列中用音符代码来表示音符，例如“C4，E5，C5”，音符序列中用“-”来表示音符的持续。延音序列 M 中，用“0”表示该处没有延音记号，用“1”则表示该处有延音记号。节奏序列 S 中，用“1，2，3，4”中的任意一个值来表示音符在这一拍中的位置。

可选地，通过包括多个声部的音乐样本数据训练音乐生成模型，得到训练好的音乐生成模型的网络参数及多个声部的音符概率密度分布，包括：向音乐生成模型中输入音乐样本数据；获取音乐生成模型输出的每个声部的音符概率密度函数：$p_i(V_i^t|V_{i,t},\mathcal{M},\theta_i)$，其中，$V_i^t$为当前时刻 t 的音符，$V_{i,t}$为音符序列中除去当前音符剩下的所有音符；$\mathcal{M}$为节奏序列及延音序列；$\theta_i$为依赖网络的参数；训练音乐生成模型使以下公式的值最大化：$\max_{\theta_i}\sum_t \log p_i(V_i^t|V_{i,t},\mathcal{M},\theta_i)$；获取当公式的值最大时音乐生成模型的网络参数及多个声部的音符概率密度分布。

长短时神经网络为循环神经网络，能够利用其内部的记忆来处理任意时序的输入序列。初始地，可以直接初始化音乐生成模型的各个网络参数，例如，随机生成并采集大数量的音乐样本数据，以对音乐生成模型进行训练。其后，可以通过随机梯度下降算法，使得其中的长短时神经网络的网络参数随之更新，如：层与层之间的连接权值和神经元偏置等，以达到音乐生成模型的音乐生成效果不断逼近最优的效果。

在训练期间，系统对长短时神经网络的参数值赋予约束条件，从而使其继续满足对神经网路的参数的要求。从而通过多次迭代，调节长短时神经网络的参数的值来对目标函数进行优化。

图 2 是根据本申请实施例的一种音乐生成模型的示意图，如图 2 所示，训练过程中，向音乐生成模型中输入音乐样本数据之后，音乐生成模型的第一长短时神经网络接收每个声部的音符序列中当前时刻音符前的预设时长的第一音符序列，并根据第一音符序列输出第一参数至依赖网络；第二长短时神经网络接收每个声部的音符序列中当前时刻音符后的预设

时长的第二音符序列，并根据第二音符序列输出第二参数至依赖网络；单隐藏层神经网络接收每个声部的音符序列中当前时刻音符并传递至依赖网络；依赖网络根据第一参数、第二参数及当前时刻音符输出每个声部的音符概率密度函数。

可选地，第一长短时神经网络接收每个声部的音符序列中当前时刻音符前的 16 个时间节点的第一音符序列，第二长短时神经网络接收每个声部的音符序列中当前时刻音符后的 16 个时间节点的第二音符序列。

具体地，每个声部的音符序列先通过嵌入层进行向量转换后输出至第一长短时神经网络或第二长短时神经网络；第一长短时神经网络输出的第一参数、第二长短时神经网络输出的第二参数及单隐藏层神经网络输出的当前时刻音符通过融合层进行融合后输入依赖网络中。

可选地，向训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使音乐生成模型根据音符随机序列、网络参数及多个声部的音符概率密度分布生成匹配特征参数的多声部音乐，包括：向训练好的音乐生成模型中依次输入第一声部、第二声部、第三声部、第四声部的音符随机序列；音乐生成模型基于第 i 声部的音符随机序列、网络参数、特征参数及第 i 声部的音符概率密度分布生成第 i 声部的多个音符，i 依次取一、二、三、四；根据第 i 声部的多个音符生成第 i 声部的音符新序列；将第一声部的音符新序列、第二声部的音符新序列、第三声部的音符新序列、第四声部的音符新序列组合形成多声部音乐。

可选地，用户输入的预设音乐时长与预设节奏序列及预设延音序列的序列时长相同，例如为 40 个十六音符的时长。

本申请实施例提供了一种基于长短时神经网络的多声部音乐生成装置，该装置用于执行上述基于长短时神经网络的多声部音乐生成方法，如图 3 所示，该装置包括：构建单元 10、第一获取单元 20、第二获取单元 30、生成单元 40。

构建单元 10，用于构建音乐生成模型，音乐生成模型包括一个第一长短时神经网络、一个第二长短时神经网络、一个单隐藏层神经网络及一

个依赖网络；

第一获取单元 20，用于通过包括多个声部的音乐样本数据训练音乐生成模型，得到训练好的音乐生成模型的网络参数及多个声部的音符概率密度分布；

第二获取单元 30，用于获取用户输入的用于预生成多声部音乐的特征参数，特征参数包括预设音乐时长、预设节奏序列及预设延音序列；

生成单元 40，用于向训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使音乐生成模型根据音符随机序列、网络参数及多个声部的音符概率密度分布生成匹配特征参数的多声部音乐。

在本方案中，通过构建包括长短时神经网络的音乐生成模型，利用长短时神经网络直接对音符序列进行处理，同时能够利用序列时间前后之间的相关性，得到音符概率密度分布；从而调整多个声部的音符序列，生成多声部之间协调的音乐，从而解决现有技术中难以生成多个声部之间协调的音乐的问题。

可选地，装置还包括：第三获取单元、提取单元、处理单元。

第三获取单元，用于获取多个音乐训练样本，其中，音乐训练样本包括多个声部信息；提取单元，用于提取每个声部的音符序列、音乐训练样本的节奏序列及延音序列；其中，每个声部的音符序列表示为：$V_i = \{V_i^t\}_t$，$t \in [T]$，T 为音乐训练样本的时长，是十六分音符的整数倍；i 为声部；V_i^t 为当前时刻 t 的音符；处理单元，用于将多个声部的音符序列、音乐训练样本的节奏序列及延音序列作为音乐样本数据。

可以理解地，每首曲子包括多个声部的音符序列、这个曲子的节奏序列及延音序列。将每首曲子随时间序列处理，有利于长短时神经网络学习音符之间随时间尺度的依赖关系。

例如，收集 389 首众赞歌的 midi 数据，其中，每首曲子包含四个声部：女高音、女低音、男高音和男低音。相对音高较低的女低音、男高音和男低音给音高最高的女高音作伴奏。将其中 80%的 midi 数据用于音乐训练样本，其中 20%的 midi 数据用于音乐评估样本。

可选地，音符序列中用音符代码来表示音符，例如“C4，E5，C5”，

音符序列中用“-”来表示音符的持续。延音序列M中，用“0”表示该处没有延音记号，用“1”则表示该处有延音记号。节奏序列S中，用“1，2，3，4”中的任意一个值来表示音符在这一拍中的位置。

可选地，第一获取单元20，包括输入子单元、第一获取子单元、训练子单元、第二获取子单元。

输入子单元，用于向音乐生成模型中输入音乐样本数据；第一获取子单元，用于获取音乐生成模型输出的每个声部的音符概率密度函数：$p_i(V_t^i|V_{i,t},\mathcal{M},\theta_i)$，其中，$V_t^i$为当前时刻t的音符，$V_{i,t}$为音符序列中除去当前音符剩下的所有音符；$\mathcal{M}$为节奏序列及延音序列；$\theta_i$为依赖网络的参数；训练子单元，用于训练音乐生成模型使以下公式的值最大化：$\max_{\theta_i}\sum_t \log p_i(V_t^i|V_{i,t},\mathcal{M},\theta_i)$；第二获取子单元，用于获取当公式的值最大时音乐生成模型的网络参数及多个声部的音符概率密度分布。

长短时神经网络为循环神经网络，能够利用其内部的记忆来处理任意时序的输入序列。初始地，可以直接初始化音乐生成模型的各个网络参数，例如，随机生成并采集大数量的音乐样本数据，以对音乐生成模型进行训练。其后，可以通过随机梯度下降算法，使得其中的长短时神经网络的网络参数随之更新，如：层与层之间的连接权值和神经元偏置等，以达到音乐生成模型的音乐生成效果不断逼近最优的效果。

在训练期间，系统对长短时神经网络的参数值赋予约束条件，从而使其继续满足对神经网路的参数的要求。从而通过多次迭代，调节长短时神经网络的参数的值来对目标函数进行优化。

可选地，音乐生成模型如图2所示，训练过程中，向音乐生成模型中输入音乐样本数据之后，音乐生成模型的第一长短时神经网络接收每个声部的音符序列中当前时刻音符前的预设时长的第一音符序列，并根据第一音符序列输出第一参数至依赖网络；第二长短时神经网络接收每个声部的音符序列中当前时刻音符后的预设时长的第二音符序列，并根据第二音符序列输出第二参数至依赖网络；单隐藏层神经网络接收每个声部的音符序列中当前时刻音符并传递至依赖网络；依赖网络根据第一参数、第二参数及当前时刻音符输出每个声部的音符概率密度函数。

可选地，第一长短时神经网络接收每个声部的音符序列中当前时刻音符前的16个时间节点的第一音符序列，第二长短时神经网络接收每个声部的音符序列中当前时刻音符后的16个时间节点的第二音符序列。

具体地，每个声部的音符序列先通过嵌入层进行向量转换后输出至第一长短时神经网络或第二长短时神经网络；第一长短时神经网络输出的第一参数、第二长短时神经网络输出的第二参数及单隐藏层神经网络输出的当前时刻音符通过融合层进行融合后输入依赖网络中。

具体地，生成新音乐过程中，生成单元40包括输入子单元，用于向训练好的音乐生成模型中依次输入第一声部、第二声部、第三声部、第四声部的音符随机序列；音乐生成模型基于第i声部的音符随机序列、网络参数、特征参数及第i声部的音符概率密度分布生成第i声部的多个音符，i依次取一、二、三、四；根据第i声部的多个音符生成第i声部的音符新序列；将第一声部的音符新序列、第二声部的音符新序列、第三声部的音符新序列、第四声部的音符新序列组合形成多声部音乐。

本申请实施例提供了一种计算机非易失性存储介质，存储介质包括存储的程序，其中，在程序运行时控制存储介质所在设备执行以下步骤：

构建音乐生成模型，音乐生成模型包括一个第一长短时神经网络、一个第二长短时神经网络、一个单隐藏层神经网络及一个依赖网络；通过包括多个声部的音乐样本数据训练音乐生成模型，得到训练好的音乐生成模型的网络参数及多个声部的音符概率密度分布；获取用户输入的用于预生成多声部音乐的特征参数，特征参数包括预设音乐时长、预设节奏序列及预设延音序列；向训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使音乐生成模型根据音符随机序列、网络参数及多个声部的音符概率密度分布生成匹配特征参数的多声部音乐。

可选地，在程序运行时控制存储介质所在设备还执行以下步骤：获取多个音乐训练样本，其中，音乐训练样本包括多个声部信息；提取每个声部的音符序列、音乐训练样本的节奏序列及延音序列；其中，每个声部的音符序列表示为：$V_i = \{V_i^t\}_t$，$t \in [T]$，T为音乐训练样本的时长，是十

六分音符的整数倍；i 为声部；V_t^i为当前时刻 t 的音符；将多个声部的音符序列、音乐训练样本的节奏序列及延音序列作为音乐样本数据。

可选地，在程序运行时控制存储介质所在设备还执行以下步骤：向音乐生成模型中输入音乐样本数据；获取音乐生成模型输出的每个声部的音符概率密度函数：$p_i(V_t^i|V_{i,t},M,\theta_i)$，其中，$V_t^i$为当前时刻 t 的音符，$V_{i,t}$为音符序列中除去当前音符剩下的所有音符；$M$为节奏序列及延音序列；$\theta_i$为依赖网络的参数；训练音乐生成模型使以下公式的值最大化：$\max_{\theta_i}\sum_t \log p_i(V_t^i|V_{i,t},M,\theta_i)$；获取当公式的值最大时音乐生成模型的网络参数及多个声部的音符概率密度分布。

可选地，在程序运行时控制存储介质所在设备还执行以下步骤：音乐生成模型的第一长短时神经网络接收每个声部的音符序列中当前时刻音符前的预设时长的第一音符序列，并根据第一音符序列输出第一参数至依赖网络；第二长短时神经网络接收每个声部的音符序列中当前时刻音符后的预设时长的第二音符序列，并根据第二音符序列输出第二参数至依赖网络；单隐藏层神经网络接收每个声部的音符序列中当前时刻音符并传递至依赖网络；依赖网络根据第一参数、第二参数及当前时刻音符输出每个声部的音符概率密度函数。

可选地，在程序运行时控制存储介质所在设备还执行以下步骤：向训练好的音乐生成模型中依次输入第一声部、第二声部、第三声部、第四声部的音符随机序列；音乐生成模型基于第 i 声部的音符随机序列、网络参数、特征参数及第 i 声部的音符概率密度分布生成第 i 声部的多个音符，i 依次取一、二、三、四；根据第 i 声部的多个音符生成第 i 声部的音符新序列；将第一声部的音符新序列、第二声部的音符新序列、第三声部的音符新序列、第四声部的音符新序列组合形成多声部音乐。

如图 4 所示，本申请实施例提供了一种计算机设备 100，包括存储器 102、处理器 101 以及存储在所述存储器 102 中并可在所述处理器 101 上运行的计算机程序 103，处理器执行计算机程序时实现以下步骤：

构建音乐生成模型，音乐生成模型包括一个第一长短时神经网络、一

个第二长短时神经网络、一个单隐藏层神经网络及一个依赖网络；通过包括多个声部的音乐样本数据训练音乐生成模型，得到训练好的音乐生成模型的网络参数及多个声部的音符概率密度分布；获取用户输入的用于预生成多声部音乐的特征参数，特征参数包括预设音乐时长、预设节奏序列及预设延音序列；向训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使音乐生成模型根据音符随机序列、网络参数及多个声部的音符概率密度分布生成匹配特征参数的多声部音乐。

可选地，处理器执行计算机程序时还实现以下步骤：获取多个音乐训练样本，其中，音乐训练样本包括多个声部信息；提取每个声部的音符序列、音乐训练样本的节奏序列及延音序列；其中，每个声部的音符序列表示为：$V_i = (v_i^t)_t$，$t \in [T]$，T为音乐训练样本的时长，是十六分音符的整数倍；i为声部；v_i^t为当前时刻t的音符；将多个声部的音符序列、音乐训练样本的节奏序列及延音序列作为音乐样本数据。

可选地，处理器执行计算机程序时还实现以下步骤：向音乐生成模型中输入音乐样本数据；获取音乐生成模型输出的每个声部的音符概率密度函数：$p_i(v_i^t | V_{i,<t}, M, \theta_i)$，其中，$v_i^t$为当前时刻t的音符，$V_{i,<t}$为音符序列中除去当前音符剩下的所有音符；$M$为节奏序列及延音序列；$\theta_i$为依赖网络的参数；训练音乐生成模型使以下公式的值最大化：$\max_{\theta_i} \sum_t \log p_i(v_i^t | V_{i,<t}, M, \theta_i)$；获取当公式的值最大时音乐生成模型的网络参数及多个声部的音符概率密度分布。

可选地，处理器执行计算机程序时还实现以下步骤：音乐生成模型的第一长短时神经网络接收每个声部的音符序列中当前时刻音符前的预设时长的第一音符序列，并根据第一音符序列输出第一参数至依赖网络；第二长短时神经网络接收每个声部的音符序列中当前时刻音符后的预设时长的第二音符序列，并根据第二音符序列输出第二参数至依赖网络；单隐藏层神经网络接收每个声部的音符序列中当前时刻音符并传递至依赖网络；依赖网络根据第一参数、第二参数及当前时刻音符输出每个声部的音符概率密度函数。

可选地，处理器执行计算机程序时还实现以下步骤：向训练好的音乐

生成模型中依次输入第一声部、第二声部、第三声部、第四声部的音符随机序列；音乐生成模型基于第 i 声部的音符随机序列、网络参数、特征参数及第 i 声部的音符概率密度分布生成第 i 声部的多个音符，i 依次取一、二、三、四；根据第 i 声部的多个音符生成第 i 声部的音符新序列；将第一声部的音符新序列、第二声部的音符新序列、第三声部的音符新序列、第四声部的音符新序列组合形成多声部音乐。

需要说明的是，本申请实施例中所涉及的终端可以包括但不限于个人计算机（Personal Computer，PC）、个人数字助理（Personal Digital Assistant，PDA）、无线手持设备、平板电脑（Tablet Computer）、手机、MP3 播放器、MP4 播放器等。

可以理解的是，所述应用可以是安装在终端上的应用程序（nativeApp），或者还可以是终端上的浏览器的一个网页程序（webApp），本申请实施例对此不进行限定。

在本申请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，也可以采用硬件加软件功能单元的形式实现。

上述以软件功能单元的形式实现的集成的单元，可以存储在一个计算机可读取存储介质中。上述软件功能单元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装置（可以是个人计算机，服务器，或者网络装置等）或处理器（Processor）执行本申请各个实施例所述方法的部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（Read-Only Memory，ROM）、随机存取存储器（Random Access Memory，RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述仅为本申请的较佳实施例而已，并不用以限制本申请，凡在本申请的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本申请保护的范围之内。

权 利 要 求 书

1、一种基于长短时神经网络的多声部音乐生成方法，其特征在于，所述方法包括：

构建音乐生成模型，所述音乐生成模型包括一个第一长短时神经网络、一个第二长短时神经网络、一个单隐藏层神经网络及一个依赖网络；

通过包括多个声部的音乐样本数据训练所述音乐生成模型，得到训练好的所述音乐生成模型的网络参数及所述多个声部的音符概率密度分布；

获取用户输入的用于预生成多声部音乐的特征参数，所述特征参数包括预设音乐时长、预设节奏序列及预设延音序列；

向所述训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使所述音乐生成模型根据所述音符随机序列、所述网络参数及所述多个声部的音符概率密度分布生成匹配所述特征参数的多声部音乐。

2、根据权利要求 1 所述的方法，其特征在于，在所述通过包括多个声部的音乐样本数据训练所述音乐生成模型之前，所述方法还包括：

获取多个音乐训练样本，其中，所述音乐训练样本包括多个声部信息；

提取每个声部的音符序列、所述音乐训练样本的节奏序列及延音序列；其中，所述每个声部的音符序列表示为：$V_i = (V_i^t)_t$，$t \in [T]$，T 为所述音乐训练样本的时长，是十六分音符的整数倍；i 为声部；V_i^t 为当前时刻 t 的音符；

将所述多个声部的音符序列、所述音乐训练样本的节奏序列及延音序列作为所述音乐样本数据。

3、根据权利要求 2 所述的方法，其特征在于，所述通过包括多个声部的音乐样本数据训练所述音乐生成模型，得到训练好的所述音乐生成模型的网络参数及所述多个声部的音符概率密度分布，包括：

向所述音乐生成模型中输入所述音乐样本数据；

获取所述音乐生成模型输出的每个声部的音符概率密度函数：$p_i(V_i^t | V_{i,-t}, M, \theta_i)$，其中，$V_i^t$ 为当前时刻 t 的音符，$V_{i,-t}$ 为音符序列中除去当前音符剩下的所有音符；M 为所述节奏序列及延音序列；θ_i 为所述依赖网络的参数；

训练所述音乐生成模型使以下公式的值最大化：

$$\max_{\theta_i} \sum_{t} \log p_i(V_t^i | V_{t-1}, M, \theta_i)$$；

获取当所述公式的值最大时所述音乐生成模型的网络参数及所述多个声部的音符概率密度分布。

4、根据权利要求 3 所述的方法，其特征在于：

所述向所述音乐生成模型中输入所述音乐样本数据之后，所述音乐生成模型的所述第一长短时神经网络接收每个声部的音符序列中当前时刻音符前的预设时长的第一音符序列，并根据所述第一音符序列输出第一参数至所述依赖网络；

所述第二长短时神经网络接收每个声部的音符序列中所述当前时刻音符后的预设时长的第二音符序列，并根据所述第二音符序列输出第二参数至所述依赖网络；

所述单隐藏层神经网络接收每个声部的音符序列中所述当前时刻音符并传递至所述依赖网络；

所述依赖网络根据所述第一参数、所述第二参数及所述当前时刻音符输出所述每个声部的音符概率密度函数。

5、根据权利要求 1 所述的方法，其特征在于，所述向所述训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使所述音乐生成模型根据所述音符随机序列、所述网络参数及所述多个声部的音符概率密度分布生成匹配所述特征参数的多声部音乐，包括：

向所述训练好的音乐生成模型中依次输入第一声部、第二声部、第三声部、第四声部的音符随机序列；

所述音乐生成模型基于第 i 声部的音符随机序列、所述网络参数、所述特征参数及所述第 i 声部的音符概率密度分布生成所述第 i 声部的多个音符，i 依次取一、二、三、四；

根据所述第 i 声部的多个音符生成所述第 i 声部的音符新序列；

将所述第一声部的音符新序列、所述第二声部的音符新序列、所述第三声部的音符新序列、所述第四声部的音符新序列组合形成所述多声部音

乐。

6、一种基于长短时神经网络的多声部音乐生成装置，其特征在于，所述装置包括：

构建单元，用于构建音乐生成模型，所述音乐生成模型包括一个第一长短时神经网络、一个第二长短时神经网络、一个单隐藏层神经网络及一个依赖网络；

第一获取单元，用于通过包括多个声部的音乐样本数据训练所述音乐生成模型，得到训练好的所述音乐生成模型的网络参数及所述多个声部的音符概率密度分布；

第二获取单元，用于获取用户输入的用于预生成多声部音乐的特征参数，所述特征参数包括预设音乐时长、预设节奏序列及预设延音序列；

生成单元，用于向所述训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使所述音乐生成模型根据所述音符随机序列、所述网络参数及所述多个声部的音符概率密度分布生成匹配所述特征参数的多声部音乐。

7、根据权利要求 6 所述的装置，其特征在于，所述装置还包括：

第三获取单元，用于获取多个音乐训练样本，其中，所述音乐训练样本包括多个声部信息；

提取单元，用于提取每个声部的音符序列、所述音乐训练样本的节奏序列及延音序列；其中，所述每个声部的音符序列表示为：$V_i = \{V_i^t\}_t$，$t \in [T]$，T 为所述音乐训练样本的时长，是十六分音符的整数倍；i 为声部；V_i^t 为当前时刻 t 的音符；

处理单元，用于将所述多个声部的音符序列、所述音乐训练样本的节奏序列及延音序列作为所述音乐样本数据。

8、根据权利要求 7 所述的装置，其特征在于，所述第一获取单元，包括：

输入子单元，用于向所述音乐生成模型中输入所述音乐样本数据；

第一获取子单元，用于获取所述音乐生成模型输出的每个声部的音符概率密度函数：$p_i(V_i^t | V_{i,<t}, M, \theta_i)$，其中，$V_i^t$ 为当前时刻 t 的音符，$V_{i,<t}$

为音符序列中除去当前音符剩下的所有音符；$\mathcal{M}$为所述节奏序列及延音序列；θ_i为所述依赖网络的参数；

训练子单元，用于训练所述音乐生成模型使以下公式的值最大化：

$$\max_{\theta_i} \sum_t \log p_i(V_i^t | V_{i,t}, \mathcal{M}, \theta_i)$$

；

第二获取子单元，用于获取当所述公式的值最大时所述音乐生成模型的网络参数及所述多个声部的音符概率密度分布。

9、根据权利要求 8 所述的装置，其特征在于：所述音乐生成模型的所述第一长短时神经网络接收每个声部的音符序列中当前时刻音符前的预设时长的第一音符序列，并根据所述第一音符序列输出第一参数至所述依赖网络；

所述第二长短时神经网络接收每个声部的音符序列中所述当前时刻音符后的预设时长的第二音符序列，并根据所述第二音符序列输出第二参数至所述依赖网络；

所述单隐藏层神经网络接收每个声部的音符序列中所述当前时刻音符并传递至所述依赖网络；

所述依赖网络根据所述第一参数、所述第二参数及所述当前时刻音符输出所述每个声部的音符概率密度函数。

10、根据权利要求 6 所述的装置，其特征在于，所述生成单元包括：

输入子单元，用于向所述训练好的音乐生成模型中依次输入第一声部、第二声部、第三声部、第四声部的音符随机序列；

所述音乐生成模型基于第 i 声部的音符随机序列、所述网络参数、所述特征参数及所述第 i 声部的音符概率密度分布生成所述第 i 声部的多个音符，i 依次取一、二、三、四；

根据所述第 i 声部的多个音符生成所述第 i 声部的音符新序列；

将所述第一声部的音符新序列、所述第二声部的音符新序列、所述第三声部的音符新序列、所述第四声部的音符新序列组合形成所述多声部音乐。

11、一种计算机设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序，其特征在于，所述处理器执行所述计算机程序时实现以下步骤：

构建音乐生成模型，所述音乐生成模型包括一个第一长短时神经网络、一个第二长短时神经网络、一个单隐藏层神经网络及一个依赖网络；

通过包括多个声部的音乐样本数据训练所述音乐生成模型，得到训练好的所述音乐生成模型的网络参数及所述多个声部的音符概率密度分布；

获取用户输入的用于预生成多声部音乐的特征参数，所述特征参数包括预设音乐时长、预设节奏序列及预设延音序列；

向所述训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使所述音乐生成模型根据所述音符随机序列、所述网络参数及所述多个声部的音符概率密度分布生成匹配所述特征参数的多声部音乐。

12、根据权利要求 11 所述的计算机设备，其特征在于，所述处理器执行所述计算机程序时还实现以下步骤：

获取多个音乐训练样本，其中，所述音乐训练样本包括多个声部信息；

提取每个声部的音符序列、所述音乐训练样本的节奏序列及延音序列；其中，所述每个声部的音符序列表示为：$V_i = (v_i^t)_t$，$t \in [T]$，T 为所述音乐训练样本的时长，是十六分音符的整数倍；i 为声部；v_i^t 为当前时刻 t 的音符；

将所述多个声部的音符序列、所述音乐训练样本的节奏序列及延音序列作为所述音乐样本数据。

13、根据权利要求 12 所述的计算机设备，其特征在于，所述处理器执行所述计算机程序时还实现以下步骤：

向所述音乐生成模型中输入所述音乐样本数据；

获取所述音乐生成模型输出的每个声部的音符概率密度函数：$p_i(v_i^t | V_{i,<t}, M, \theta_i)$，其中，$v_i^t$ 为当前时刻 t 的音符，$V_{i,<t}$ 为音符序列中除去当前音符剩下的所有音符；M 为所述节奏序列及延音序列；θ_i 为所述依赖网络的参数；

训练所述音乐生成模型使以下公式的值最大化：

$$\max_{\theta_i} \sum_t \log p_i(V_i^t | V_{i,t}, \mathcal{M}, \theta_i)$$;

获取当所述公式的值最大时所述音乐生成模型的网络参数及所述多个声部的音符概率密度分布。

14、根据权利要求 13 所述的计算机设备，其特征在于，所述处理器执行所述计算机程序时还实现以下步骤：

所述音乐生成模型的所述第一长短时神经网络接收每个声部的音符序列中当前时刻音符前的预设时长的第一音符序列，并根据所述第一音符序列输出第一参数至所述依赖网络；

所述第二长短时神经网络接收每个声部的音符序列中所述当前时刻音符后的预设时长的第二音符序列，并根据所述第二音符序列输出第二参数至所述依赖网络；

所述单隐藏层神经网络接收每个声部的音符序列中所述当前时刻音符并传递至所述依赖网络；

所述依赖网络根据所述第一参数、所述第二参数及所述当前时刻音符输出所述每个声部的音符概率密度函数。

15、根据权利要求 11 所述的计算机设备，其特征在于，所述处理器执行所述计算机程序时还实现以下步骤：

向所述训练好的音乐生成模型中依次输入第一声部、第二声部、第三声部、第四声部的音符随机序列；

所述音乐生成模型基于第 i 声部的音符随机序列、所述网络参数、所述特征参数及所述第 i 声部的音符概率密度分布生成所述第 i 声部的多个音符，i 依次取一、二、三、四；

根据所述第 i 声部的多个音符生成所述第 i 声部的音符新序列；

将所述第一声部的音符新序列、所述第二声部的音符新序列、所述第三声部的音符新序列、所述第四声部的音符新序列组合形成所述多声部音乐。

16、一种计算机非易失性可读存储介质，所述存储介质包括存储的程序，其特征在于，在所述程序运行时控制所述存储介质所在设备执行以下

步骤：

构建音乐生成模型，所述音乐生成模型包括一个第一长短时神经网络、一个第二长短时神经网络、一个单隐藏层神经网络及一个依赖网络；

通过包括多个声部的音乐样本数据训练所述音乐生成模型，得到训练好的所述音乐生成模型的网络参数及所述多个声部的音符概率密度分布；

获取用户输入的用于预生成多声部音乐的特征参数，所述特征参数包括预设音乐时长、预设节奏序列及预设延音序列；

向所述训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使所述音乐生成模型根据所述音符随机序列、所述网络参数及所述多个声部的音符概率密度分布生成匹配所述特征参数的多声部音乐。

17、根据权利要求 16 所述的计算机非易失性可读存储介质，其特征在于，在所述程序运行时控制所述存储介质所在设备执行以下步骤：

获取多个音乐训练样本，其中，所述音乐训练样本包括多个声部信息；

提取每个声部的音符序列、所述音乐训练样本的节奏序列及延音序列；其中，所述每个声部的音符序列表示为：$V_i = (V_i^t)_t$，$t \in [T]$，T 为所述音乐训练样本的时长，是十六分音符的整数倍；i 为声部；V_i^t 为当前时刻 t 的音符；

将所述多个声部的音符序列、所述音乐训练样本的节奏序列及延音序列作为所述音乐样本数据。

18、根据权利要求 17 所述的计算机非易失性可读存储介质，其特征在于，在所述程序运行时控制所述存储介质所在设备执行以下步骤：

向所述音乐生成模型中输入所述音乐样本数据；

获取所述音乐生成模型输出的每个声部的音符概率密度函数：$p_i(V_i^t | V_{i,t}, \mathcal{M}, \theta_i)$，其中，$V_i^t$ 为当前时刻 t 的音符，$V_{i,t}$ 为音符序列中除去当前音符剩下的所有音符；$\mathcal{M}$ 为所述节奏序列及延音序列；θ_i 为所述依赖网络的参数；

训练所述音乐生成模型使以下公式的值最大化：

$$\max_{\theta_i} \sum_t \log p_i(V_i^t | V_{i,t}, \mathcal{M}, \theta_i)$$

；

获取当所述公式的值最大时所述音乐生成模型的网络参数及所述多个声部的音符概率密度分布。

19、根据权利要求 18 所述的计算机非易失性可读存储介质，其特征在于，在所述程序运行时控制所述存储介质所在设备执行以下步骤：

所述音乐生成模型的所述第一长短时神经网络接收每个声部的音符序列中当前时刻音符前的预设时长的第一音符序列，并根据所述第一音符序列输出第一参数至所述依赖网络；

所述第二长短时神经网络接收每个声部的音符序列中所述当前时刻音符后的预设时长的第二音符序列，并根据所述第二音符序列输出第二参数至所述依赖网络；

所述单隐藏层神经网络接收每个声部的音符序列中所述当前时刻音符并传递至所述依赖网络；

所述依赖网络根据所述第一参数、所述第二参数及所述当前时刻音符输出所述每个声部的音符概率密度函数。

20、根据权利要求 16 所述的计算机非易失性可读存储介质，其特征在于，在所述程序运行时控制所述存储介质所在设备执行以下步骤：

向所述训练好的音乐生成模型中依次输入第一声部、第二声部、第三声部、第四声部的音符随机序列；

所述音乐生成模型基于第 i 声部的音符随机序列、所述网络参数、所述特征参数及所述第 i 声部的音符概率密度分布生成所述第 i 声部的多个音符，i 依次取一、二、三、四；

根据所述第 i 声部的多个音符生成所述第 i 声部的音符新序列；

将所述第一声部的音符新序列、所述第二声部的音符新序列、所述第三声部的音符新序列、所述第四声部的音符新序列组合形成所述多声部音乐。

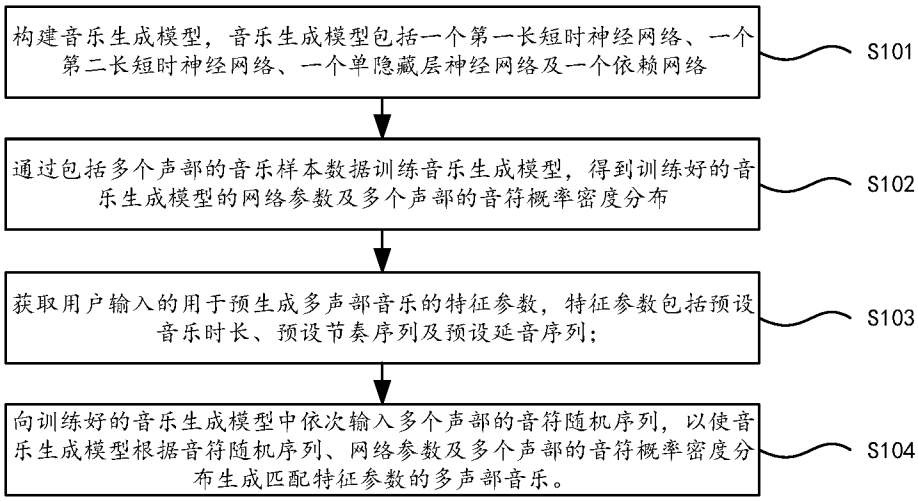
构建音乐生成模型，音乐生成模型包括一个第一长短时神经网络、一个第二长短时神经网络、一个单隐藏层神经网络及一个依赖网络
S101
通过包括多个声部的音乐样本数据训练音乐生成模型，得到训练好的音乐生成模型的网络参数及多个声部的音符概率密度分布
S102
获取用户输入的用于预生成多声部音乐的特征参数，特征参数包括预设音乐时长、预设节奏序列及预设延音序列；
S103
向训练好的音乐生成模型中依次输入多个声部的音符随机序列，以使音乐生成模型根据音符随机序列、网络参数及多个声部的音符概率密度分布生成匹配特征参数的多声部音乐。
S104

图 1

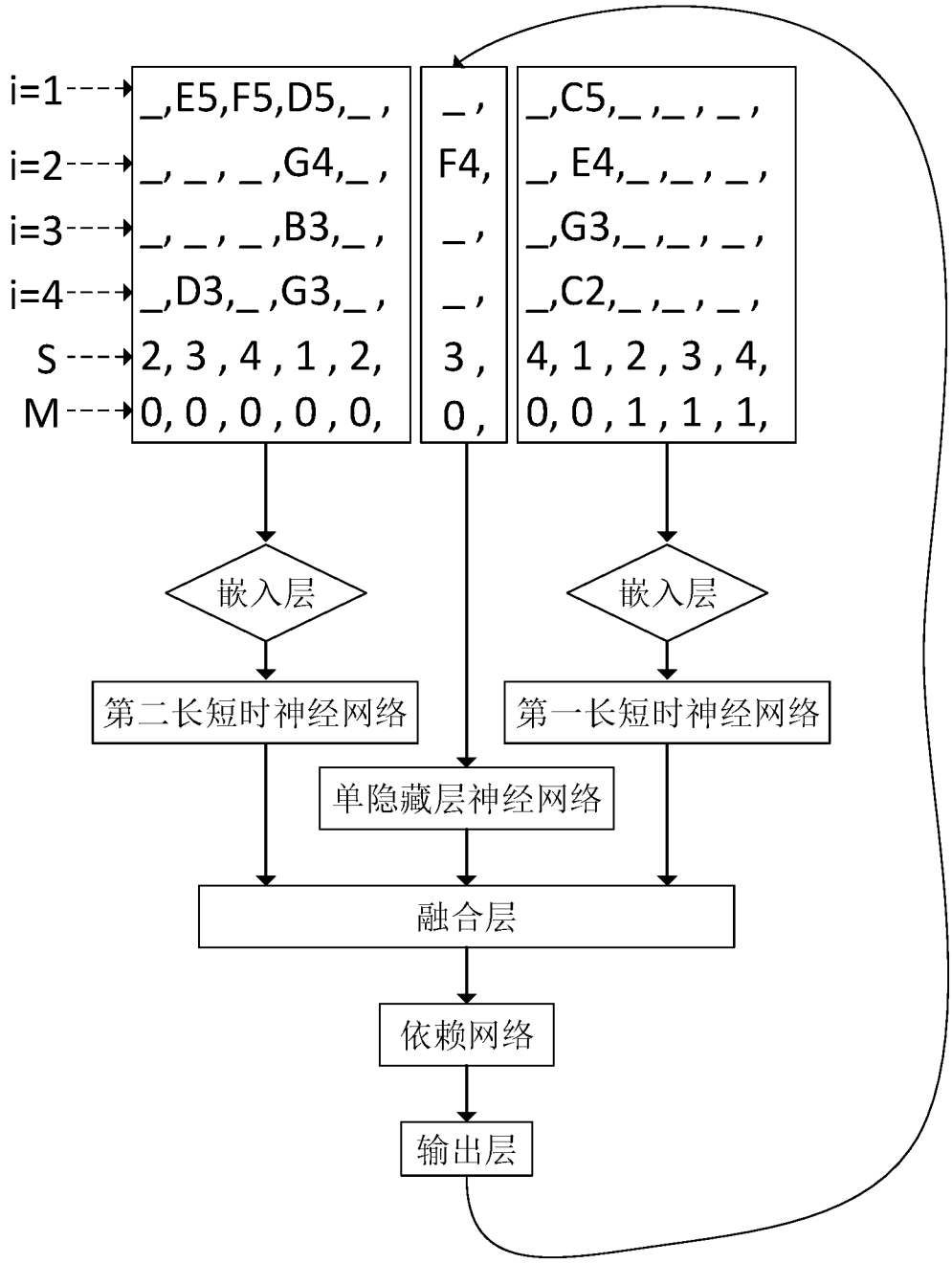
i=1
i=2
i=3
i=4
S
M
_,E5,F5,D5,_ ,
_, _ , _ ,G4,_ ,
_, _ , _ ,B3,_ ,
_,D3,_ ,G3,_ ,
2, 3 , 4 , 1 , 2,
0, 0 , 0 , 0 , 0,
_ ,
F4,
_ ,
_ ,
3 ,
0 ,
_,C5,_ ,_ , _ ,
_, E4,_ ,_ , _ ,
_,G3,_ ,_ , _ ,
_,C2,_ ,_ , _ ,
4, 1 , 2 , 3 , 4,
0, 0 , 1 , 1 , 1,
嵌入层
嵌入层
第二长短时神经网络
第一长短时神经网络
单隐藏层神经网络
融合层
依赖网络
输出层

图 2

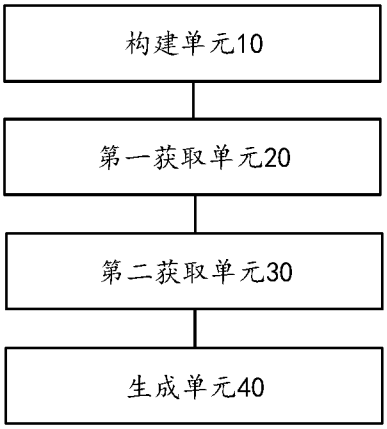
构建单元10
第一获取单元20
第二获取单元30
生成单元40

图 3

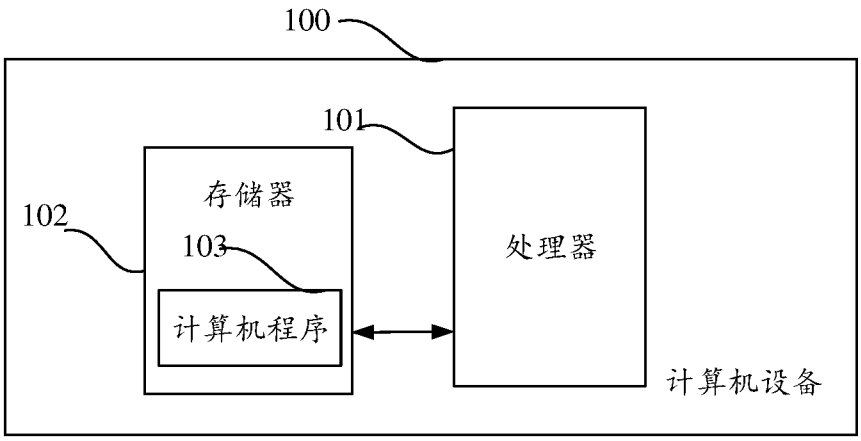
100
101
存储器
102
103
计算机程序
处理器
计算机设备

图 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/123549

A. CLASSIFICATION OF SUBJECT MATTER

G10H 1/32(2006.01)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)

G10H

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)

CNKI, CNPAT, WPI, EPODOC, IEEE: 音乐, 生成, 长短时神经网络, 多声部, 模型, 隐藏, 训练, music, generation, LSTM, polyphonic, model, hidden, train

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	王程等 (WANG, Cheng et al.). "面向自动音乐生成的深度递归神经网络方法 (Recurrent Neural Network Method for Automatic Generation of Music)" *小型微型计算机系统 (Journal of Chinese Computer Systems),* Vol. 38, No. 10, 31 October 2017 (2017-10-31), pp. 2412-2416	1-20
Y	BRUNNER, G. et al. "JamBot: Music Theory Aware Chord Based Generation of Polyphonic Music with LSTMs," *2017 International Conference on Tools with Artificial Intelligence,* 08 November 2017 (2017-11-08), ISSN: 2375-0197, pp. 519-526	1-20
A	CN 107644630 A (TSINGHUA UNIVERSITY) 30 January 2018 (2018-01-30) entire document	1-20
A	CN 107123415 A (WU, ZHENGUO ET AL.) 01 September 2017 (2017-09-01) entire document	1-20

☐ 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	Date of mailing of the international search report
16 July 2019	**08 August 2019**
Name and mailing address of the ISA/CN **National Intellectual Property Administration, PRC (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/CN2018/123549

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 107644630 A	30 January 2018	None	
CN 107123415 A	01 September 2017	None	

国际检索报告

国际申请号
PCT/CN2018/123549

A. 主题的分类

G10H 1/32(2006.01)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)

G10H

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））

CNKI, CNPAT, WPI, EPODOC, IEEE: 音乐, 生成, 长短时神经网络, 多声部, 模型, 隐藏, 训练, music, generation, LSTM, polyphonic, model, hidden, train

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
Y	王程 等,. "面向自动音乐生成的深度递归神经网络方法," 《小型微型计算机系统》, 第38卷，第10期，2017年 10月 31日 (2017-10-31), 第2412-2416页	1-20
Y	BRUNNER, Gino 等,. "JamBot: Music Theory Aware Chord Based Generation of Polyphonic Music with LSTMs," 《2017 International Conference on Tools with Artificial Intelligence》, 2017年 11月 8日 (2017-11-08), ISSN: 2375-0197, 第519-526页	1-20
A	CN 107644630 A (清华大学) 2018年 1月 30日 (2018-01-30) 全文	1-20
A	CN 107123415 A (吴振国 等) 2017年 9月 1日 (2017-09-01) 全文	1-20

☐ 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	国际检索报告邮寄日期
2019年 7月 16日	2019年 8月 8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 田越 电话号码 86-(10)-53961347

表 PCT/ISA/210 (第2页) (2015年1月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/123549

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 107644630 A	2018年 1月 30日	无	
CN 107123415 A	2017年 9月 1日	无	