

(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 5 月 7 日 (07.05.2020)



(10) 国际公布号
WO 2020/087979 A1

(51) 国际专利分类号:
G06F 16/73 (2019.01)

(21) 国际申请号: PCT/CN2019/095735

(22) 国际申请日: 2019 年 7 月 12 日 (12.07.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811273701.7 2018年10月30日 (30.10.2018) 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)。

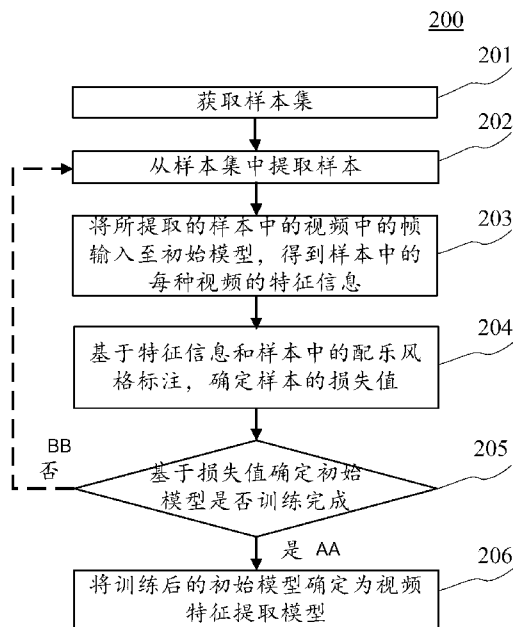
(72) 发明人: 袁泽寰 (YUAN, Zehuan); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。 卮春光 (QIE, Chunguang); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。 王长虎 (WANG, Changhu); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。

(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,

(54) Title: METHOD AND APPARATUS FOR GENERATING MODEL

(54) 发明名称: 生成模型的方法和装置



201 Acquire a sample set
202 Extract a sample from the sample set
203 Input frames in videos in the extracted sample into an initial model to obtain feature information of each video in the sample
204 On the basis of the feature information and soundtrack style label in the sample, determine a loss value of the sample
205 On the basis of the loss value, determine whether training of the initial model is complete
206 Determine that the trained initial model is a video feature extraction model
AA Yes
BB No

图 2

(57) Abstract: Disclosed in the embodiments of the present invention are a method and an apparatus for generating a model. One embodiment of the present method comprises: acquiring a sample set; extracting a sample from the sample set and executing the following training process: inputting frames in videos in the extracted sample into an initial model to obtain feature information of each video in the sample; on the basis of the feature information and soundtrack style labels in the sample, determining a loss value of the sample; on the basis of the loss value, determining whether training of the initial model is complete; and, in response to determining

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))。

that training of the initial model is complete, determining that the trained initial model is a video feature extraction model.

(57) 摘要: 本申请实施例公开了生成模型的方法和装置。该方法的一示例实施方式包括: 获取样本集; 从该样本集中提取样本, 执行如下训练过程: 将所提取的样本中的视频中的帧输入至初始模型, 得到样本中的每种视频的特征信息; 基于特征信息和样本中的配乐风格标注, 确定样本的损失值; 基于该损失值确定初始模型是否训练完成; 响应于确定初始模型训练完成, 将训练后的初始模型确定为视频特征提取模型。

生成模型的方法和装置

本申请要求在2018年10月30日提交中国专利局、申请号为201811273701.7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

本申请实施例涉及计算机技术领域，例如生成模型的方法和装置。

背景技术

随着计算机技术的发展，短视频类应用应运而生。用户可以利用短视频类应用上传、发布视频。利用短视频应用上传视频时，通常需要用户选择一个音乐作为视频配乐。

相关技术中的方式，通常是用户手动从配乐列表中选取配乐。

发明内容

本申请实施例提出了生成模型的方法和装置。

第一方面，本申请实施例提供了一种生成模型的方法，该方法包括：获取样本集，其中，样本集中的样本包括第一视频、第二视频和第三视频，所述第一视频和所述第二视频的配乐相同且带有相同的配乐风格标注，所述第一视频与所述第三视频的配乐不同且带有不同的配乐风格标注；从样本集中提取样本，执行如下训练过程：将所提取的样本中的视频中的帧输入至初始模型，得到样本中的每种视频的特征信息；基于所述特征信息和样本中的配乐风格标注，确定样本的损失值；基于所述损失值确定初始模型是否训练完成；响应于确定初始模型训练完成，将训练后的初始模型确定为视频特征提取模型。

第二方面，本申请实施例提供了一种生成模型的装置，该装置包括：获取单元，被配置成获取样本集，其中，样本集中的样本包括第一视频、第二视频和第三视频，所述第一视频和所述第二视频的配乐相同且带有相同的配乐风格标注，所述第一视频与所述第三视频的配乐不同且带有不同的配乐风格标注；训练单元，被配置成从样本集中提取样本，执行如下训练过程：将所提取的样本中的视频中的帧输入至初始模型，得到样本中的每种视频的特征信息；基于所述特征信息和样本中的配乐风格标注，确定样本的损失值；基于所述损失值确定初始模型是否训练完成；响应于确定初始模型训练完成，将训练后的初始模型确定为视频特征提取模型。

第三方面，本申请实施例提供了一种推送信息的方法，包括：响应于接收

到目标视频，将目标视频中的帧输入采用如上述第一方面中任一实施例所描述的方法生成的视频特征提取模型，得到目标视频的目标特征信息；将目标特征信息与预置的视频库中的视频的特征信息进行相似度计算，按照相似度从大到小的顺序，从视频库中选取预设数量的视频作为候选视频；获取候选视频的配乐信息，将所述配乐信息进行推送。

第四方面，本申请实施例提供了一种推送信息的装置，包括：接收单元，被配置成响应于接收到目标视频，将目标视频中的帧输入采用如上述第一方面中任一实施例所描述的方法生成的视频特征提取模型，得到目标视频的目标特征信息；选取单元，被配置成将目标特征信息与预置的视频库中的视频的特征信息进行相似度计算，按照相似度从大到小的顺序，从视频库中选取预设数量的视频作为候选视频；推送单元，被配置成获取候选视频的配乐信息，将所述配乐信息进行推送。

第五方面，本申请实施例提供了一种电子设备，包括：至少一个处理器；存储装置，其上存储有至少一个程序，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上述第一方面和第三方面中任一实施例的方法。

第六方面，本申请实施例提供了一种计算机可读介质，其上存储有计算机程序，所述程序被处理器执行时实现如上述第一方面和第三方面中任一实施例的方法。

附图说明

图 1 是本申请的一个实施例可以应用于其中的示例性系统架构图；
图 2 是根据本申请的生成模型的方法的一个实施例的流程图；
图 3 是根据本申请的生成模型的方法的一个应用场景的示意图；
图 4 是根据本申请的生成模型的方法的又一个实施例的流程图；
图 5 是根据本申请的生成模型的装置的一个实施例的结构示意图；
图 6 是根据本申请的推送信息的方法的一个实施例的流程图；
图 7 是根据本申请的推送信息的装置的一个实施例的结构示意图；
图 8 是适于用来实现本申请实施例的电子设备的计算机系统的结构示意图。

具体实施方式

下面结合附图和实施例对本申请作详细说明。可以理解的是，此处所描述的实施例仅仅用于解释相关发明，而非对该申请的限定。另外还需要说明的是，为了便于描述，附图中仅示出了与有关申请相关的部分。

需要说明的是，在不冲突的情况下，本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本申请。

图 1 示出了可以应用本申请的生成模型的方法或生成模型的装置的示例性系统架构 100。

如图 1 所示，系统架构 100 可以包括终端设备 101、102、103，网络 104 和服务器 105。网络 104 用以在终端设备 101、102、103 和服务器 105 之间提供通信链路的介质。网络 104 可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

用户可以使用终端设备 101、102、103 通过网络 104 与服务器 105 交互，以接收或发送消息等。终端设备 101、102、103 上可以安装有各种通讯客户端应用，例如视频录制类应用、视频播放类应用、语音交互类应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

终端设备 101、102、103 可以是硬件，也可以是软件。在终端设备 101、102、103 为硬件的情况下，可以是具有显示屏的各种电子设备，包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等等。在终端设备 101、102、103 为软件的情况下，可以安装在上述所列举的电子设备中。其可以实现成多个软件或软件模块（例如用来提供分布式服务），也可以实现成单个软件或软件模块。在此不做限定。

在终端设备 101、102、103 为硬件的情况下，其上还可以安装有图像采集设备。图像采集设备可以是各种能实现采集图像功能的设备，如摄像头、传感器等等。用户可以利用终端设备 101、102、103 上的图像采集设备，来采集视频。

服务器 105 可以是提供各种服务的服务器，例如用于对终端设备 101、102、103 上传的视频进行存储、管理或者分析的视频处理服务器。视频处理服务器可以获取样本集。样本集中可以包含大量的样本。其中，上述样本集中的样本可以包括第一视频、第二视频和第三视频。第一视频和第二视频的配乐相同且带有相同的配乐风格标注。第一视频与第三视频的配乐不同且带有不同的配乐风格标注。此外，视频处理服务器可以利用样本集中的样本，对初始模型进行训练，并可以将训练结果（如生成的视频特征提取模型）进行存储。这样，在用户利用终端设备 101、102、103 上传视频后，服务器 105 可以确定用户所上传的视频的特征信息，进而，可以对该视频进行配乐信息的选取和推送等操作。

需要说明的是，服务器 105 可以是硬件，也可以是软件。在服务器为硬件的情况下，可以实现成多个服务器组成的分布式服务器集群，也可以实现成单个服务器。在服务器为软件的情况下，可以实现成多个软件或软件模块

(例如用来提供分布式服务),也可以实现成单个软件或软件模块。在此不做限定。

需要说明的是,本申请实施例所提供的生成模型的方法一般由服务器 105 执行,相应地,生成模型的装置一般设置于服务器 105 中。

应该理解,图 1 中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器的。

参考图 2,示出了根据本申请的生成模型的方法的一个实施例的流程 200。该生成模型的方法,包括步骤 201 至步骤 206。

在步骤 201 中,获取样本集。

在本实施例中,生成模型的方法的执行主体(例如图 1 所示的服务器 105)可以通过多种方式来获取样本集。例如,执行主体可以通过有线连接方式或无线连接方式,从用于存储样本的另一服务器(例如数据库服务器)中获取存储于其中的样本集。再例如,用户可以通过终端设备(例如图 1 所示的终端设备 101、102、103)来收集样本。这样,上述执行主体可以接收终端所收集的样本,并将这些样本存储在本地,从而生成样本集。需要指出的是,上述无线连接方式可以包括但不限于 3G/4G 连接、WiFi 连接、蓝牙连接、WiMAX 连接、Zigbee 连接、UWB (ultra wideband) 连接、以及其他现在已知或将来开发的无线连接方式。

此处,样本集中可以包括大量的样本。其中,样本可以包括第一视频、第二视频和第三视频。需要说明的是,第一视频和第二视频的配乐相同且带有相同的配乐风格标注。第一视频与第三视频的配乐不同且带有不同的配乐风格标注。配乐风格标注可以是用于指示和区分配乐的风格的信息。配乐风格可以预先划分为多种,例如悲伤、欢快、舒缓等。

在本实施例的一些实现方式中,上述样本集中的样本可以通过如下步骤生成:首先,可以从预置的视频库中随机提取视频作为第一视频,其中,上述视频库中的视频带有配乐标注和配乐风格标注。实践中,配乐标注可以设置为指示和区分配乐。例如,配乐标注可以是配乐的名称。接着,可以从上述视频库中随机提取与上述第一视频具有相同的配乐标注且具有相同的配乐风格标注的视频作为第二视频。接着,可以从上述视频库中随机选取与上述第一视频带有不同的配乐标注且具有不同的配乐风格标注的视频作为第三视频。最后,可以将上述第一视频、上述第二视频、上述第三视频汇总为样本。

需要说明的是,上述样本中的样本还可以利用其他方式生成,例如人工选取等方式,此处不再赘述。

在步骤 202 中,从样本集中提取样本。

在本实施例中,执行主体可以从步骤 201 中获取的样本集中提取样本,

以及执行步骤 203 至步骤 206 的训练过程。其中，样本的提取方式和提取数量在本申请中并不限制。例如，可以是随机提取至少一个样本，也可以是从中提取样本的视频的清晰度较好（即样本中的视频的帧的像素较高）的样本。

在步骤 203 中，将所提取的样本中的视频中的帧输入至初始模型，得到样本中的每种视频的特征信息。

在本实施例中，上述执行主体可以将步骤 202 中提取的样本中的视频中的帧输入至初始模型。由于所提取的样本中包括第一视频、第二视频、第三视频，因而，可以分别得到第一视频的特征信息、第二视频的特征信息、第三视频的特征信息。初始模型可以通过对视频中的帧进行分析等，输出视频的特征信息。实践中，特征信息可以以向量或者矩阵的形式来表示。

需要说明的是，所输入的视频中的帧，可以是随机抽取的一帧或多帧；也可以是按照指定时间间隔（例如 1s 或者 2s 等）从视频中抽取的多帧。此处不作限定。

在本实施例中，初始模型可以是基于机器学习技术而创建的各种具有图像特征提取功能的模型。初始模型可以对视频中的帧进行特征提取，而后对所提取的多帧的特征进行融合、分析等处理，最终输出视频的特征信息。

作为示例，初始模型可以是使用各种相关技术中的结构（例如 DenseBox、VGGNet、ResNet、SegNet 等）的卷积神经网络。实践中，卷积神经网络（Convolutional Neural Network，CNN）是一种前馈神经网络，它的人工神经元可以响应一部分覆盖范围内的周围单元，对于图像处理有出色表现，因而，可以利用卷积神经网络进行视频中的帧的特征的提取。

在本示例中，所建立的神经网络可以包含卷积层、池化层、特征融合层、全连接层等。其中，卷积层可以设置为提取图像特征。池化层可以设置为对输入的信息进行降采样（downsample）。特征融合层可以设置为将所得到的多帧对应的图像特征（例如，可以是特征矩阵的形式，或者特征向量的形式）进行融合。例如，可以将不同帧对应的特征矩阵中的相同位置的特征值取平均，以进行特征融合，生成一个融合后的特征矩阵。全连接层可以设置为对融合后的特征进行处理。

需要说明的是，上述初始模型也可以是具有图像特征提取功能的其他模型，并不限于上述示例，对于模型结构此处不作限定。

在步骤 204 中，基于特征信息和样本中的配乐风格标注，确定样本的损失值。

在本实施例中，训练初始模型的目标是使之从配乐相同且带有相同的配乐风格标注的视频的帧中提取的特征信息的差异尽可能小，同时，从配乐不同且带有不同配乐风格标注的视频的帧中提取的特征信息的差异尽可能大。

由此，可以将从配乐相同且带有相同的配乐风格标注的视频的帧中提取的特征信息的差异称为第一差异。可以将从配乐不同且带有不同配乐风格标注的视频的帧中提取的特征信息的差异称为第二差异。上述执行主体可以首先根据特征信息确定出第一差异和第二差异的值。实践中，由于所提取的样本中的第一视频与第二视频配乐相同且带有相同的配乐风格标注，因此，第一差异可以是所提取的样本中的第一视频的特征信息与第二视频的特征信息的差异。由于所提取的样本中的第一视频与第三视频配乐不同且带有不同的配乐风格标注，第二视频与第三视频的配乐不同且带有不同的配乐风格标注。因此，第二差异可以是所提取的样本中的第一视频的特征信息与第三视频的特征信息的差异，或者，可以是第二视频的特征信息与第三视频的特征信息的差异。此处，特征信息的差异可以利用欧氏距离、余弦相似度等方式来确定。特征信息越相似，差异越小。特征信息越不相似，差异越大。

在得到第一差异与第二差异之后，上述执行主体可以将第一差异、第二差异输入至预先建立的损失函数，确定样本的损失值。损失函数是一个非负实值函数。一般情况下，损失函数的值（损失值）越小，模型的鲁棒性就越好。损失函数可以根据实际需求来设置。作为示例，上述损失函数可以是用于表征第二差异与第一差异的差异程度的函数，例如三元组损失函数（triplet loss）。

在本实施例的一些实现方式中，上述执行主体可以通过如下步骤确定所提取的样本的损失值：

第一步，确定带有相同的配乐风格标注的视频的特征信息的欧氏距离，将该距离作为第一欧氏距离。并确定不同的配乐风格标注的视频的特征信息的欧氏距离，将该距离作为第二欧氏距离。

第二步，确定上述第二欧氏距离与上述第一欧氏距离的差值。

第三步，将上述差值与第一预设数值进行比较，确定样本的损失值，其中，上述第一预设数值为正数（例如 0.2）。此处，第一预设数值可以是技术人员基于大量数据统计和计算而预先指定的数值。

在一实施例中，响应于确定上述第二欧氏距离与上述第一欧氏距离的差值大于上述第一预设数值，上述执行主体可以将第二预设数值（例如 0）确定为样本的损失值。可以理解的是，在上述第二欧氏距离与上述第一欧氏距离的差值大于上述第一预设数值的情况下，可以认为第二欧氏距离与第一欧氏距离的数值关系满足预期。此时，损失值可以设置为较小的数，使该样本的损失值对梯度下降的影响较小或不参与梯度下降。由于此时第二欧氏距离与上述第一欧氏距离的差值大于上述第一预设数值，因而该差值与第一预设数值的差（此处可称为目标值）大于 0，因此，可以将上述第二预设数值设

置为小于上述目标值的数，例如 0。

在一实施例中，响应于确定上述第二欧氏距离与上述第一欧氏距离的差值小于或等于上述第一预设数值，上述执行主体可以将上述第一预设数值与上述差值的差确定为样本的损失值。可以理解的是，在上述第二欧氏距离与上述第一欧氏距离的差值小于或等于第一预设数值的情况下，可以认为第二欧氏距离与第一欧氏距离的数值关系不满足预期。此时，由于第一预设数值与该差值的差大于零，因此，可以将上述第一预设数值与上述差值的差确定为样本的损失值，参与梯度下降。

在步骤 205 中，基于损失值确定初始模型是否训练完成。

在本实施例中，上述执行主体可以基于步骤 204 所确定的损失值，确定初始模型是否训练完成。作为示例，上述执行主体可以确定损失值是否已收敛。在确定损失值收敛的情况下，则可以确定此时的初始模型已训练完成。作为又一示例，上述执行主体可以首先将损失值与目标值进行比较。响应于确定损失值小于或等于目标值，可以统计最近执行的预设数量（例如 100）次训练过程所确定的损失值中，小于或等于上述目标值的损失值的数量所占比例。在该比例大于预设比例（例如 95%）时，可以确定初始模型训练完成。需要说明的是，目标值一般可以设置为表示预测值与真实值之间的不一致程度的理想情况。也就是说，在损失值小于或等于目标值的情况下，可以认为预测值接近或近似真实值。目标值可以根据实际需求来设置。

需要说明的是，响应于确定初始模型已训练完成，则可以继续执行步骤 206。响应于确定初始模型未训练完成，可以基于所确定的损失值，更新初始模型中的参数，从上述样本集中重新提取样本，使用更新参数后的初始模型作为初始模型，继续执行上述训练过程。此处，可以利用反向传播算法求得损失值相对于模型参数的梯度，而后利用梯度下降算法基于梯度更新模型参数。需要说明的是，上述反向传播算法、梯度下降算法以及机器学习方法是目前广泛研究和应用的公知技术，在此不再赘述。需要指出的是，这里的提取方式在本申请中也不限制。例如在样本集中有大量样本的情况下，执行主体可以从中提取未被提取过的样本。

在步骤 206 中，响应于确定初始模型训练完成，将训练后的初始模型确定为视频特征提取模型。

在本实施例中，响应于确定初始模型训练完成，上述执行主体可以将训练后的初始模型确定为视频特征提取模型。

在本实施例的一些实现方式中，在训练得到视频特征提取模型之后，响应于接收到目标视频，将上述目标视频中的帧输入至该视频特征提取模型，得到上述目标视频的目标特征信息；而后，可以将上述目标特征信息与预置

的视频库中的视频的特征信息进行相似度计算,按照相似度从大到小的顺序,从上述视频库中选取预设数量的视频作为候选视频;最后,可以获取上述候选视频的配乐信息,将所选取的配乐信息进行推送。

参见图 3,图 3 是根据本实施例的生成模型的方法的应用场景的一个示意图。在图 3 的应用场景中,图 3 的应用场景中,用户所使用的终端设备 301 上可以安装有模型训练类应用。在用户打开该应用,并上传样本集或样本集的存储路径后,对该应用提供后台支持的服务器 302 可以运行生成模型的方法,包括:

首先,可以获取样本集。其中,上述样本集中的样本包括第一视频、第二视频和第三视频,第一视频和第二视频的配乐相同且带有相同的配乐风格标注,第一视频与第三视频的配乐不同且带有不同的配乐风格标注。之后,可以从上述样本集中提取样本,执行如下训练过程:对所提取的样本中的视频进行抽帧。将所抽取的第一视频中的帧 303、第二视频中的帧 304、第三视频中的帧 305 输入至初始模型 306,得到样本中的每种视频的特征信息。基于特征信息和样本中的配乐风格标注 307,确定样本的损失值 308。基于上述损失值确定初始模型是否训练完成。响应于确定初始模型训练完成,将训练后的初始模型确定为视频特征提取模型 309。

通过获取样本集,可以从中提取样本以进行初始模型的训练。其中,上述样本集中的样本可以包括第一视频、第二视频和第三视频。第一视频和第二视频的配乐相同且带有相同的配乐风格标注。第一视频与第三视频的配乐不同且带有不同的配乐风格标注。这样,将提取的样本中的视频中的帧输入至初始模型,便可以得到样本中的每种视频的特征信息。之后,可以基于特征信息和样本中的配乐风格标注,确定样本的损失值。最后,可以基于所确定的损失值确定初始模型是否训练完成。如果初始模型训练完成,就可以将训练后的初始模型确定为视频特征提取模型。从而能够得到一种可以用于提取视频特征的模型,且该模型所提取的视频特征有助于视频配乐的自动选取。

参考图 4,其示出了生成模型的方法的又一个实施例的流程 400。该生成模型的方法的流程 400,包括步骤 401 至步骤 408。

在步骤 401 中,获取样本集。

在本实施例中,生成模型的方法的执行主体(例如图 1 所示的服务器 105)可以获取样本集。样本集中可以包括大量的样本。其中,样本可以包括第一视频、第二视频和第三视频。需要说明的是,第一视频和第二视频的配乐相同且带有相同的配乐风格标注。第一视频与第三视频的配乐不同且带有不同的配乐风格标注。配乐风格标注可以是用于指示和区分配乐的风格的信息。配乐风格可以预先划分为多种,例如悲伤、欢快、舒缓等。

在本实施例中，上述样本集中的样本可以通过如下步骤生成：首先，可以从预置的视频库中随机提取视频作为第一视频，其中，上述视频库中的视频带有配乐标注和配乐风格标注，配乐风格标注设置为指示配乐的風格。接着，可以从上述视频库中随机提取与上述第一视频具有相同的配乐标注且具有相同的配乐风格标注的视频作为第二视频。接着，可以从上述视频库中随机选取与上述第一视频带有不同的配乐标注且具有不同的配乐风格标注的视频作为第三视频。最后，可以将上述第一视频、上述第二视频、上述第三视频汇总为样本。

在步骤 402 中，从样本集中提取样本。

在本实施例中，执行主体可以从步骤 401 中获取的样本集中提取样本，以及执行步骤 403 至步骤 408 的训练过程。其中，样本的提取方式和提取数量在本申请中并不限制。例如，可以是随机提取至少一个样本，也可以是从中提取样本的视频的清晰度较好（即样本中的视频的帧的像素较高）的样本。

在步骤 403 中，将所提取的样本中的视频中的帧输入至初始模型，得到样本中的每种视频的特征信息。

在本实施例中，上述执行主体可以将步骤 402 中提取的样本中的视频中的帧输入至初始模型，分别得到第一视频的特征信息、第二视频的特征信息、第三视频的特征信息。实践中，特征信息可以以向量或者矩阵的形式来表示。

在本实施例中，初始模型可以是基于机器学习技术而创建的卷积神经网络。初始模型可以对视频中的帧进行特征提取，而后对所提取的多帧的特征进行融合、分析等处理，最终输出视频的特征信息。此处，所建立的神经网络可以包含卷积层、池化层、特征融合层、全连接层等。

在步骤 404 中，确定第一欧氏距离和第二欧氏距离。

在本实施例中，上述执行主体可以确定第一欧氏距离和第二欧氏距离。其中，上述第一欧氏距离为带有相同的配乐风格标注的视频的特征信息的欧氏距离，上述第二欧氏距离为带有不同的配乐风格标注的视频的特征信息的欧氏距离。

此处，由于所提取的样本中的第一视频与第二视频的配乐相同且带有相同的配乐风格标注，因此，第一差异可以是所提取的样本中的第一视频的特征信息与第二视频的特征信息的差异。由于所提取的样本中的第一视频与第三视频的配乐不同且带有不同的配乐风格标注，第二视频与第三视频带有不同的配乐风格标注。因此，第二差异可以是所提取的样本中的第一视频的特征信息与第三视频的特征信息的差异，或者，第二视频的特征信息与第三视频的特征信息的差异。

在步骤 405 中，确定第二欧氏距离与第一欧氏距离的差值。

在本实施例中，上述执行主体可以确定上述第二欧氏距离与上述第一欧氏距离的差值。

在步骤 406 中，将上述差值与第一预设数值进行比较，确定样本的损失值。

在本实施例中，上述执行主体可以将上述差值与第一预设数值进行比较，确定样本的损失值。其中，上述第一预设数值为正数（例如 0.2）。此处，第一预设数值可以是技术人员基于大量数据统计和计算而预先指定的数值。

在一实施例中，响应于确定上述第二欧氏距离与上述第一欧氏距离的差值大于上述第一预设数值，上述执行主体可以将第二预设数值（例如 0）确定为样本的损失值。可以理解的是，在上述第二欧氏距离与上述第一欧氏距离的差值大于上述第一预设数值的情况下，可以认为第二欧氏距离与第一欧氏距离的数值关系满足预期。此时，损失值可以设置为较小的数值，使该样本的损失值对梯度下降的影响较小或不参与梯度下降。由于此时第二欧氏距离与上述第一欧氏距离的差值大于上述第一预设数值，因而该差值与第一预设数值的差（此处可称为目标值）大于 0，因此，可以将上述第二预设数值设置为小于上述目标值的数，例如 0。

响应于确定上述第二欧氏距离与上述第一欧氏距离的差值小于或等于上述第一预设数值，上述执行主体可以将上述第一预设数值与上述差值的差确定为样本的损失值。可以理解的是，在上述第二欧氏距离与上述第一欧氏距离的差值小于或等于第一预设数值的情况下，可以认为第二欧氏距离与第一欧氏距离的数值关系不满足预期。此时，由于第一预设数值与该差值的差大于零，因此，可以将上述第一预设数值与上述差值的差确定为样本的损失值，参与梯度下降。

在步骤 407 中，基于损失值确定初始模型是否训练完成。

在本实施例中，上述执行主体可以基于步骤 406 所确定的损失值，确定初始模型是否训练完成。作为示例，上述执行主体可以确定损失值是否已收敛。在确定损失值收敛的情况下，可以确定此时的初始模型已训练完成。

需要说明的是，响应于确定初始模型已训练完成，则可以继续执行步骤 408。响应于确定初始模型未训练完成，可以基于所确定的损失值，更新初始模型中的参数，从上述样本集中重新提取样本，使用更新参数后的初始模型作为初始模型，继续执行上述训练过程。此处，可以利用反向传播算法求得损失值相对于模型参数的梯度，而后利用梯度下降算法基于梯度更新模型参数。需要说明的是，上述反向传播算法、梯度下降算法以及机器学习方法是目前广泛研究和应用的公知技术，在此不再赘述。需要指出的是，这里的提取方式在本申请中也不限制。例如在样本集中有大量样本的情况下，执行主

体可以从中提取未被提取过的样本。

在步骤 408 中，响应于确定初始模型训练完成，将训练后的初始模型确定为视频特征提取模型。

在本实施例中，响应于确定初始模型训练完成，上述执行主体可以将训练后的初始模型确定为视频特征提取模型。

从图 4 中可以看出，与图 2 对应的实施例相比，本实施例中的生成模型的方法的流程 400 涉及了确定所提取的样本的损失值的一种方式。由此，本实施例描述的方案可以使模型从配乐相同且带有相同的配乐风格标注的视频的帧中提取的特征信息的差异尽可能小，同时，从配乐不同且带有不同配乐风格标注的视频的帧中提取的特征信息的差异尽可能大。由此，能够得到一种可以提取视频特征的模型，且该模型所提取的视频特征有助于视频配乐的自动选取。

参考图 5，作为对上述各图所示方法的实现，本申请提供了一种生成模型的装置的一个实施例，该装置实施例与图 2 所示的方法实施例相对应，该装置可以应用于各种电子设备中。

如图 5 所示，本实施例所述的生成模型的装置 500 包括：获取单元 501，被配置成获取样本集，其中，上述样本集中的样本包括第一视频、第二视频和第三视频，第一视频和第二视频的配乐相同且带有相同的配乐风格标注，第一视频与第三视频的配乐不同且带有不同的配乐风格标注；训练单元 502，被配置成从上述样本集中提取样本，执行如下训练过程：将所提取的样本中的视频中的帧输入至初始模型，得到样本中的每种视频的特征信息；基于特征信息和样本中的配乐风格标注，确定样本的损失值；基于上述损失值确定初始模型是否训练完成；响应于确定初始模型训练完成，将训练后的初始模型确定为视频特征提取模型。

在本实施例的一些实现方式中，上述训练单元 502 可以被配置成确定第一欧氏距离和第二欧氏距离，其中，上述第一欧氏距离为带有相同的配乐风格标注的视频的特征信息的欧氏距离，上述第二欧氏距离为带有不同的配乐风格标注的视频的特征信息的欧氏距离；确定上述第二欧氏距离与上述第一欧氏距离的差值；将上述差值与第一预设数值进行比较，确定样本的损失值，其中，上述第一预设数值为正数。

在本实施例的一些实现方式中，上述训练单元 502 可以被配置成：响应于确定上述差值大于第一预设数值，将第二预设数值确定为样本的损失值，其中，上述第二预设数值小于上述差值与上述第一预设数值的差。

在本实施例的一些实现方式中，上述训练单元 502 可以被配置成：响应于确定上述差值小于或等于第一预设数值，将上述第一预设数值与上述差值

的差确定为样本的损失值。

在本实施例的一些实现方式中,上述样本集中的样本通过如下步骤生成:从预置的视频库中随机提取视频作为第一视频,其中,上述视频库中的视频带有配乐标注和配乐风格标注,配乐风格标注设置为指示配乐的風格;从上述视频库中随机提取与上述第一视频具有相同的配乐标注且具有相同的配乐风格标注的视频作为第二视频;从上述视频库中随机选取与上述第一视频带有不同的配乐标注且具有不同的配乐风格标注的视频作为第三视频;将上述第一视频、上述第二视频、上述第三视频汇总为样本。

在本实施例的一些实现方式中,该装置还可以包括更新单元(图中未示出)。其中,上述更新单元可以被配置成响应于确定初始模型未训练完成,基于上述损失值,更新初始模型中的参数,从上述样本集中重新提取样本,使用更新参数后的初始模型作为初始模型,继续执行上述训练过程。

本申请的上述实施例提供的装置,通过获取单元 501 获取样本集,可以从中提取样本以进行初始模型的训练。其中,上述样本集中的样本可以包括第一视频、第二视频和第三视频。第一视频和第二视频的配乐相同且带有相同的配乐风格标注。第一视频与第三视频的配乐不同且带有不同的配乐风格标注。这样,训练单元 502 将提取的样本中的视频中的帧输入至初始模型,便可以得到样本中的每种视频的特征信息。之后,可以基于特征信息和样本中的配乐风格标注,确定样本的损失值。最后,可以基于所确定的损失值确定初始模型是否训练完成。如果初始模型训练完成,就可以将训练后的初始模型确定为视频特征提取模型。从而能够得到一种可以提取视频特征的模型,且该模型所提取的视频特征有助于视频配乐的自动选取。

请参见图 6,其示出了本申请提供的推送信息的方法的一个实施例的流程 600。该推送信息的方法可以包括步骤 601 至步骤 603。

在步骤 601 中,响应于接收到目标视频,将目标视频中的帧输入视频特征提取模型,得到目标视频的目标特征信息。

在本实施例中,推送信息的执行主体(例如图 1 所示的服务器 105,或者存储有视频特征提取模型的其他服务器)响应于接收到目标视频,可以将上述目标视频中的帧输入视频特征提取模型,得到上述目标视频的目标特征信息。此处,目标视频可以是终端设备上传的尚未配乐的視頻。

在本实施例中,视频特征提取模型可以是采用如上述图 2 实施例所描述的方法而生成的。生成过程可以参见图 2 实施例的相关描述,此处不再赘述。

在步骤 602 中,将目标特征信息与预置的视频库中的视频的特征信息进行相似度计算,按照相似度从大到小的顺序,从视频库中选取预设数量的视频作为候选视频。

在本实施例中，上述执行主体可以上述目标特征信息与预置的视频库中的视频的特征信息进行相似度计算，按照相似度从大到小的顺序，从上述视频库中选取预设数量（例如 5 个）的视频作为候选视频。

在步骤 603 中，获取候选视频的配乐信息，将所选取的配乐信息进行推送。

在本实施例中，上述执行主体可以获取上述候选视频的配乐信息。此处，配乐信息可以包括但不限于以下至少一项：配乐的音频文件、配乐的名称、配乐的风格名称。最后，可以将所选取的配乐信息进行推送，例如推送至上述终端设备，以供用户选择。

需要说明的是，本实施例推送信息的方法可以用于测试上述各实施例所生成的视频特征提取模型。进而根据测试结果可以不断地优化视频特征提取模型。该方法也可以是上述各实施例所生成的视频特征提取模型的实际应用方法。采用上述各实施例所生成的视频特征提取模型，来提取目标视频的目标特征信息，基于所提取的目标视频的特征信息，选取配乐信息，可以对没有配乐的 video 进行配乐推荐，实现了富于针对性的信息推送。

参见图 7，作为对上述图 6 所示方法的实现，本申请提供了一种推送信息的装置的一个实施例。该装置实施例与图 6 所示的方法实施例相对应，该装置可以应用于各种电子设备中。

如图 7 所示，本实施例所述的推送信息的装置 700 包括：接收单元 701，被配置成响应于接收到目标视频，将上述目标视频中的帧输入采用如上述图 2 实施例所描述的方法生成的视频特征提取模型，得到上述目标视频的目标特征信息；选取单元 702，被配置成将上述目标特征信息与预置的视频库中的视频的特征信息进行相似度计算，按照相似度从大到小的顺序，从上述视频库中选取预设数量的视频作为候选视频；推送单元 703，被配置成获取上述候选视频的配乐信息，将所选取的配乐信息进行推送。

可以理解的是，该装置 700 中记载的诸单元与参考图 6 描述的方法中的各个步骤相对应。由此，上文针对方法描述的操作、特征同样适用于装置 700 及其中包含的单元，在此不再赘述。

下面参考图 8，其示出了适于用来实现本申请实施例的电子设备的计算机系统 800 的结构示意图。图 8 示出的电子设备仅仅是一个示例，不对本申请实施例的功能和使用范围带来任何限制。

如图 8 所示，计算机系统 800 包括中央处理单元（Central Processing Unit, CPU）801，其可以根据存储在只读存储器（Read-Only Memory, ROM）802 中的程序或者从存储部分 808 加载到随机访问存储器（Random Access

Memory, RAM) 803 中的程序而执行各种适当的动作和处理。在 RAM 803 中, 还存储有系统 800 操作所需的各种程序和数据。CPU 801、ROM 802 以及 RAM 803 通过总线 804 彼此相连。输入/输出 (Input/Output, I/O) 接口 805 也连接至总线 804。

以下部件连接至 I/O 接口 805: 包括键盘、鼠标等的输入部分 806; 包括诸如阴极射线管 (Cathode Ray Tube, CRT)、液晶显示器 (Liquid Crystal Display, LCD) 等以及扬声器等的输出部分 807; 包括硬盘等的存储部分 808; 以及包括诸如局域网 (Local Area Network, LAN) 卡、调制解调器等网络接口卡的通信部分 809。通信部分 809 经由诸如因特网的网络执行通信处理。驱动器 810 也根据需要连接至 I/O 接口 805。可拆卸介质 811, 诸如磁盘、光盘、磁光盘、半导体存储器等等, 根据需要安装在驱动器 810 上, 以便于从其上读出的计算机程序根据需要被安装入存储部分 808。

特别地, 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, 其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部分 809 从网络上被下载和安装, 和/或从可拆卸介质 811 被安装。在该计算机程序被中央处理单元 (CPU) 801 执行时, 执行本申请的方法中限定的上述功能。需要说明的是, 本申请所述的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更具体的例子可以包括但不限于: 具有至少一个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器 (RAM)、只读存储器 (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。在本申请中, 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。而在本申请中, 计算机可读的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, 其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。计算机可读的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：无线、电线、光缆、射频（Radio Frequency，RF）等等，或者上述的任意合适的组合。

附图中的流程图和框图，图示了按照本申请各种实施例的系统、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，该模块、程序段、或代码的一部分包含至少一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。也要注意的，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。

描述于本申请实施例中所涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现。所描述的单元也可以设置在处理器中，例如，可以描述为：一种处理器包括获取单元和训练单元。其中，这些单元的名称在某种情况下并不构成对该单元本身的限定，例如，获取单元还可以被描述为“获取样本集的单元”。

作为另一方面，本申请还提供了一种计算机可读介质，该计算机可读介质可以是上述实施例中描述的装置中所包含的；也可以是单独存在，而未装配入该装置中。上述计算机可读介质承载有至少一个程序，在上述至少一个程序被该装置执行时，使得该装置：获取样本集；从该样本集中提取样本，执行如下训练过程：将所提取的样本中的视频中的帧输入至初始模型，得到样本中的每种视频的特征信息；基于特征信息和样本中的配乐风格标注，确定样本的损失值；基于该损失值确定初始模型是否训练完成；响应于确定初始模型训练完成，将训练后的初始模型确定为视频特征提取模型。

1、一种生成模型的方法，包括：

获取样本集，其中，所述样本集中的样本包括第一视频、第二视频和第三视频，所述第一视频和所述第二视频的配乐相同且带有相同的配乐风格标注，所述第一视频与所述第三视频的配乐不同且带有不同的配乐风格标注；

从所述样本集中提取样本，执行如下训练过程：将所提取的样本中的视频中的帧输入至初始模型，得到样本中的每种视频的特征信息；基于所述特征信息和样本中的配乐风格标注，确定样本的损失值；基于所述损失值确定初始模型是否训练完成；响应于确定初始模型训练完成，将训练后的初始模型确定为视频特征提取模型。

2、根据权利要求1所述的生成模型的方法，其中，所述基于所述特征信息和样本中的配乐风格标注，确定样本的损失值，包括：

确定第一欧氏距离和第二欧氏距离，其中，所述第一欧氏距离为带有相同的配乐风格标注的视频的特征信息的欧氏距离，所述第二欧氏距离为带有不同的配乐风格标注的视频的特征信息的欧氏距离；

确定所述第二欧氏距离与所述第一欧氏距离的差值；

将所述差值与第一预设数值进行比较，确定样本的损失值，其中，所述第一预设数值为正数。

3、根据权利要求2所述的生成模型的方法，其中，所述将所述差值与第一预设数值进行比较，确定样本的损失值，包括：

响应于确定所述差值大于所述第一预设数值，将第二预设数值确定为样本的损失值，其中，所述第二预设数值小于所述差值与所述第一预设数值的差。

4、根据权利要求2所述的生成模型的方法，其中，所述将所述差值与第一预设数值进行比较，确定样本的损失值，包括：

响应于确定所述差值小于或等于所述第一预设数值，将所述第一预设数值与所述差值的差确定为样本的损失值。

5、根据权利要求1所述的生成模型的方法，其中，所述样本集中的样本通过如下步骤生成：

从预置的视频库中随机提取视频作为第一视频，其中，所述视频库中的视频带有配乐标注和配乐风格标注；

从所述视频库中随机提取与所述第一视频具有相同的配乐标注且具有相同的配乐风格标注的视频作为第二视频；

从所述视频库中随机选取与所述第一视频带有不同的配乐标注且具有不同的配乐风格标注的视频作为第三视频；

将所述第一视频、所述第二视频、所述第三视频汇总为样本。

6、根据权利要求1所述的生成模型的方法，还包括：

响应于确定初始模型未训练完成，基于所述损失值，更新初始模型中的参数，从所述样本集中重新提取样本，使用更新参数后的初始模型作为初始模型，继续执行所述训练过程。

7、一种生成模型的装置，包括：

获取单元，被配置成获取样本集，其中，所述样本集中的样本包括第一视频、第二视频和第三视频，所述第一视频和所述第二视频的配乐相同且带有相同的配乐风格标注，所述第一视频与所述第三视频的配乐不同且带有不同的配乐风格标注；

训练单元，被配置成从所述样本集中提取样本，执行如下训练过程：将所提取的样本中的视频中的帧输入至初始模型，得到样本中的每种视频的特征信息；基于所述特征信息和样本中的配乐风格标注，确定样本的损失值；基于所述损失值确定初始模型是否训练完成；响应于确定初始模型训练完成，将训练后的初始模型确定为视频特征提取模型。

8、根据权利要求7所述的生成模型的装置，其中，所述训练单元被配置成：

确定第一欧氏距离和第二欧氏距离，其中，所述第一欧氏距离为带有相同的配乐风格标注的视频的特征信息的欧氏距离，所述第二欧氏距离为带有不同的配乐风格标注的视频的特征信息的欧氏距离；

确定所述第二欧氏距离与所述第一欧氏距离的差值；

将所述差值与第一预设数值进行比较，确定样本的损失值，其中，所述第一预设数值为正数。

9、根据权利要求8所述的生成模型的装置，其中，所述训练单元被配置成：

响应于确定所述差值大于第一预设数值，将第二预设数值确定为样本的损失值，其中，所述第二预设数值小于所述差值与所述第一预设数值的差。

10、根据权利要求8所述的生成模型的装置，其中，所述训练单元被配置成：

响应于确定所述差值小于或等于第一预设数值，将所述第一预设数值与所述差值的差确定为样本的损失值。

11、根据权利要求7所述的生成模型的装置，其中，所述样本集中的样本通过如下步骤生成：

从预置的视频库中随机提取视频作为第一视频，其中，所述视频库中的视频带有配乐标注和配乐风格标注；

从所述视频库中随机提取与所述第一视频具有相同的配乐标注且具有相同的配乐风格标注的视频作为第二视频；

从所述视频库中随机选取与所述第一视频带有不同的配乐标注且具有不同的配乐风格标注的视频作为第三视频；

将所述第一视频、所述第二视频、所述第三视频汇总为样本。

12、根据权利要求7所述的生成模型的装置，还包括：

更新单元，被配置成响应于确定初始模型未训练完成，基于所述损失值，更新初始模型中的参数，从所述样本集中重新提取样本，使用更新参数后的初始模型作为初始模型，继续执行所述训练过程。

13、一种推送信息的方法，包括：

响应于接收到目标视频，将所述目标视频中的帧输入采用如权利要求1-6任一所述的方法生成的视频特征提取模型中，得到所述目标视频的目标特征信息；

将所述目标特征信息与预置的视频库中的视频的特征信息进行相似度计算，按照相似度从大到小的顺序，从所述视频库中选取预设数量的视频作为候选视频；

获取所述候选视频的配乐信息，将所述的配乐信息进行推送。

14、一种推送信息的装置，包括：

接收单元，被配置成响应于接收到目标视频，将所述目标视频中的帧输入采用如权利要求1-6任一所述的方法生成的视频特征提取模型，得到所述目标视频的目标特征信息；

选取单元，被配置成将所述目标特征信息与预置的视频库中的视频的特征信息进行相似度计算，按照相似度从大到小的顺序，从所述视频库中选取预设数量的视频作为候选视频；

推送单元，被配置成获取所述候选视频的配乐信息，将所述配乐信息进行推送。

15、一种电子设备，包括：

至少一个处理器；

存储装置，其上存储有至少一个程序，

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1-6、13中任一项所述的方法。

16、一种计算机可读介质，其上存储有计算机程序，其中，所述程序被处理器执行时实现如权利要求1-6、13中任一项所述的方法。

100

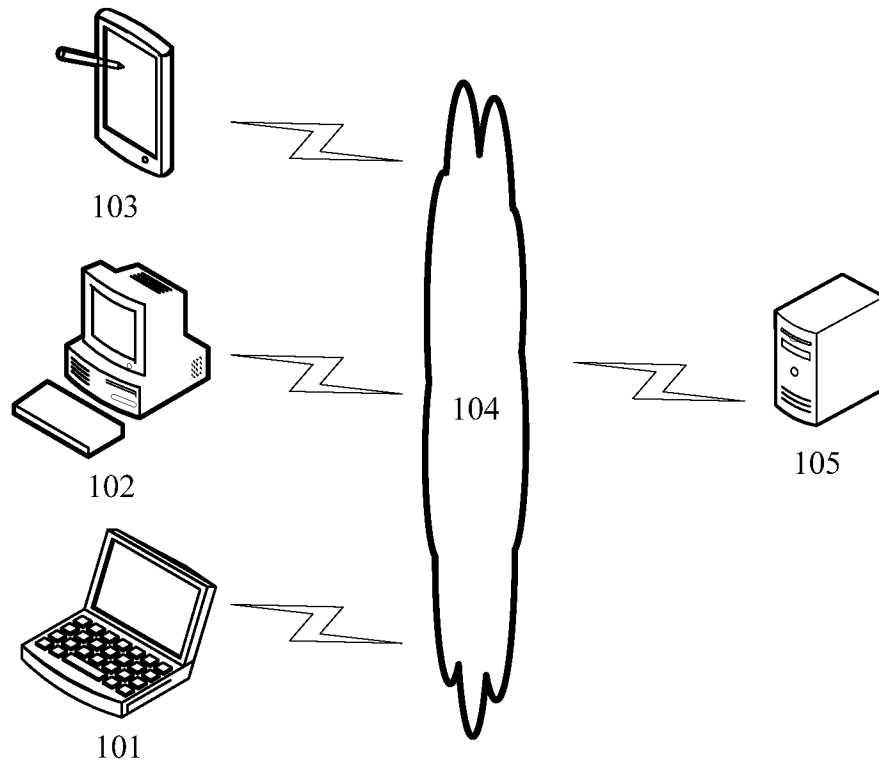


图 1

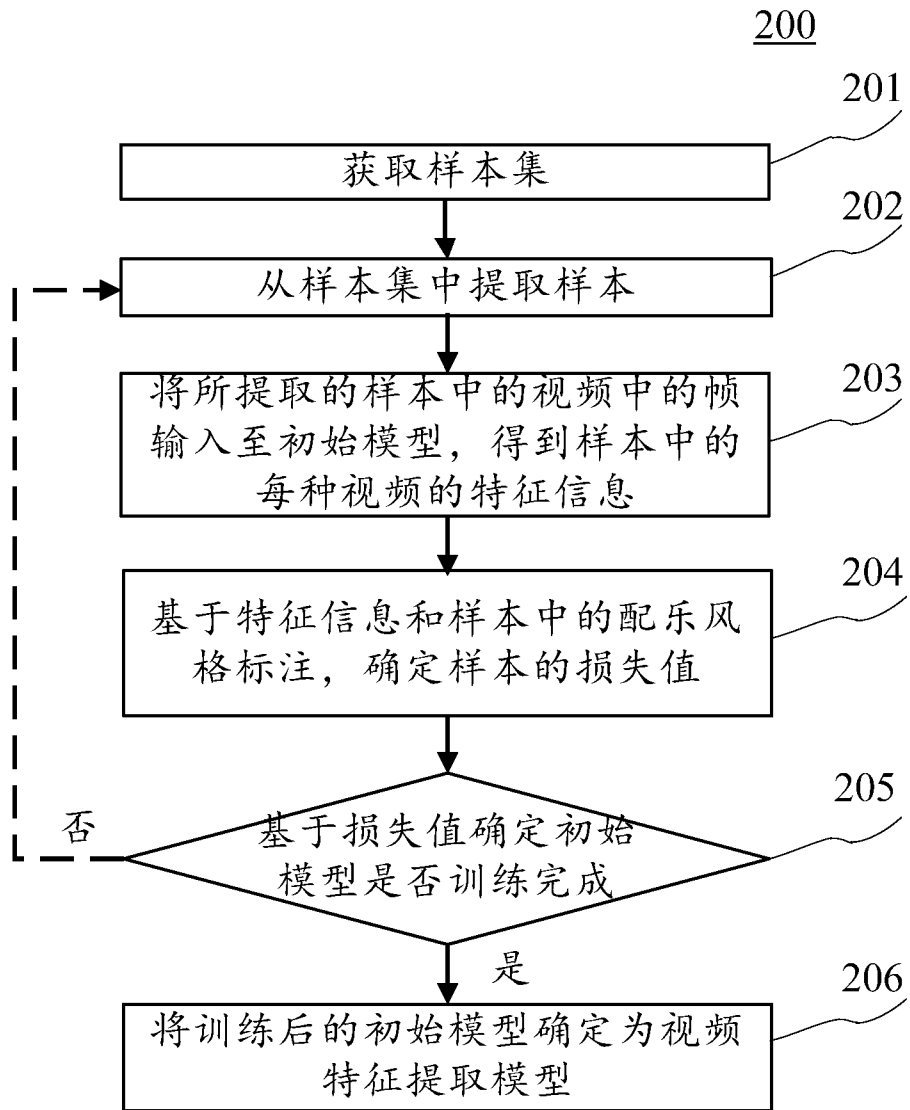


图 2

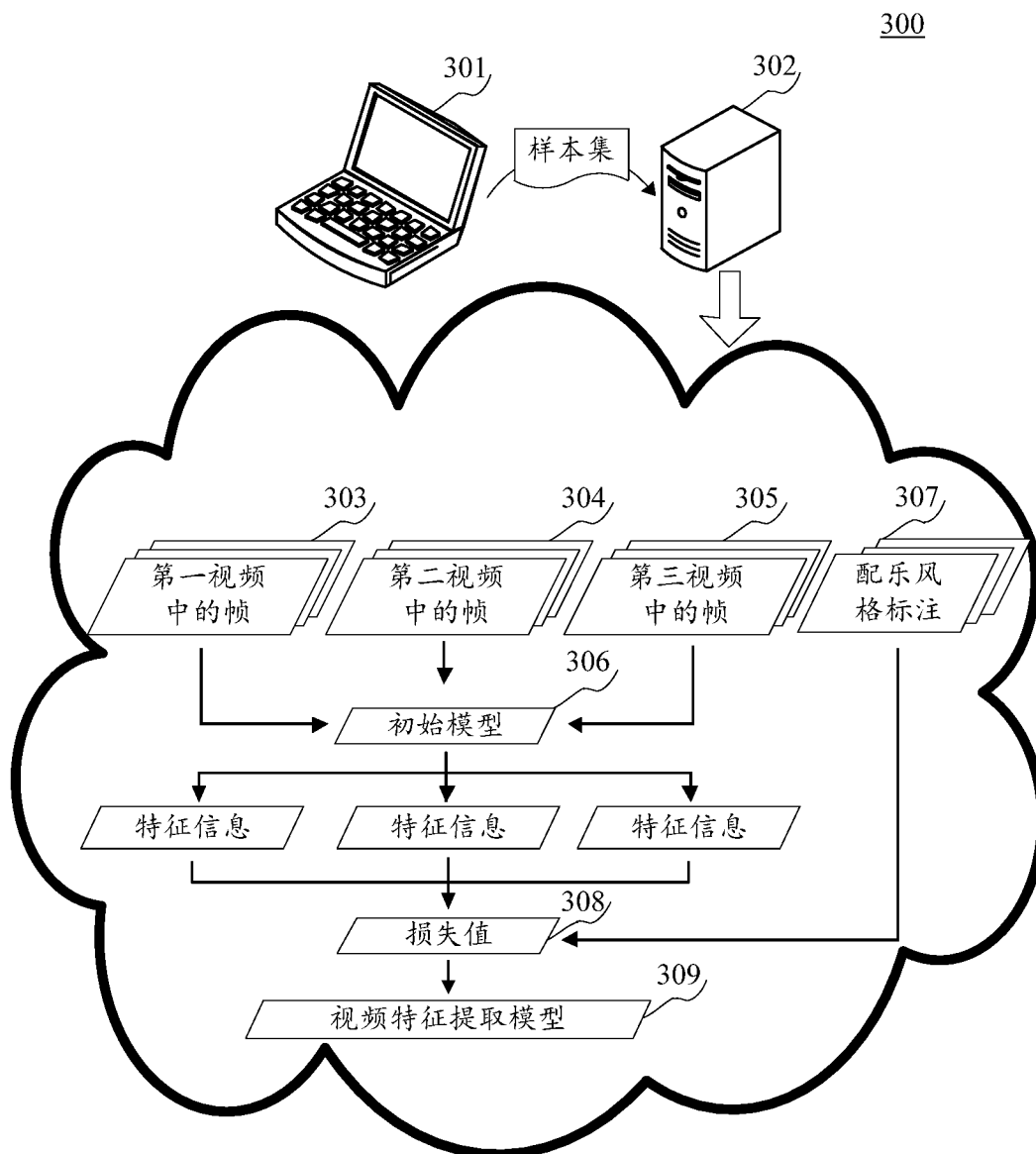


图 3

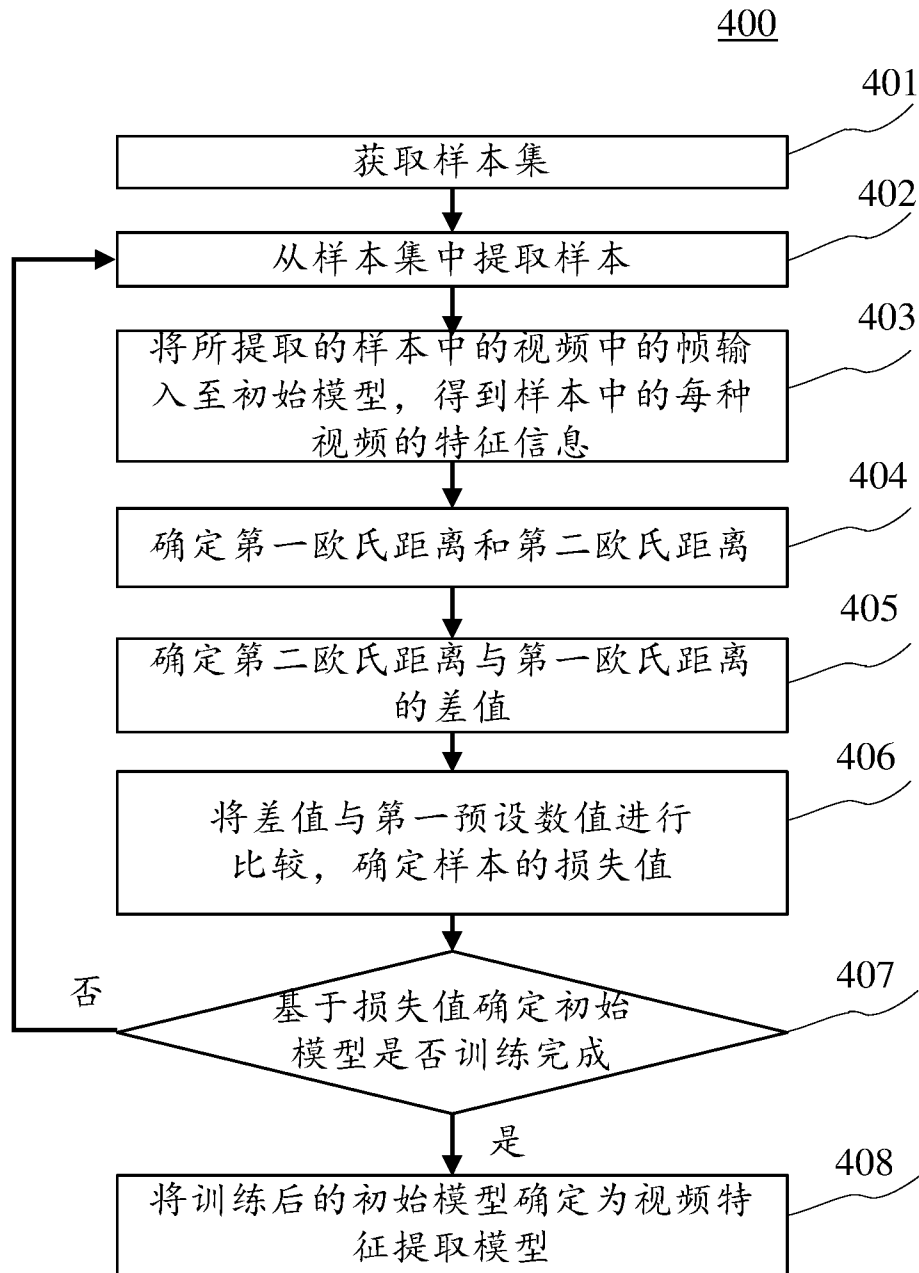


图 4

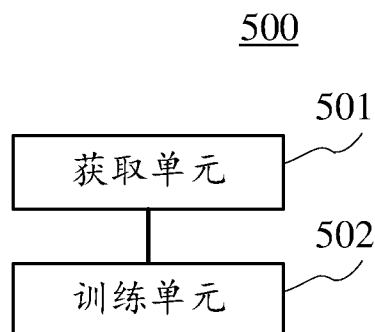


图 5

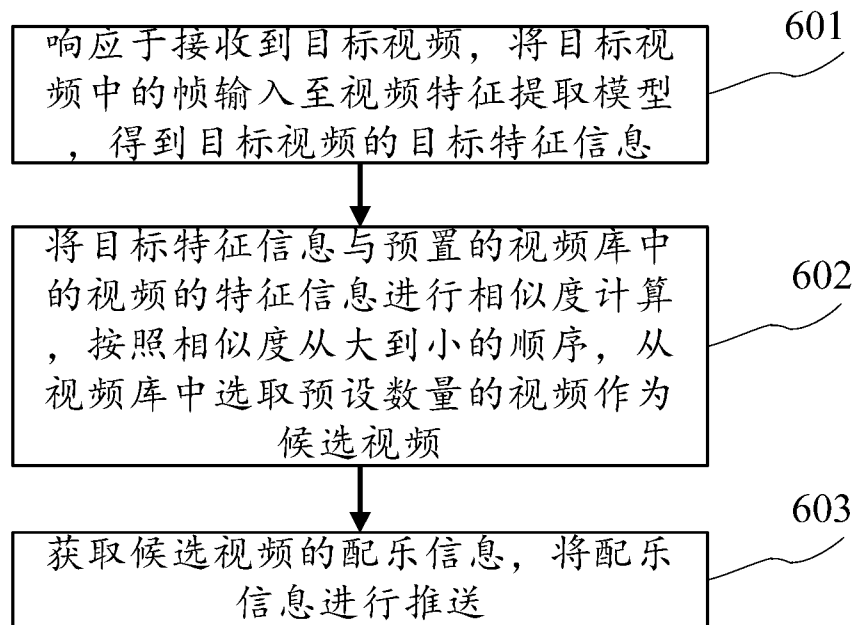
600

图 6

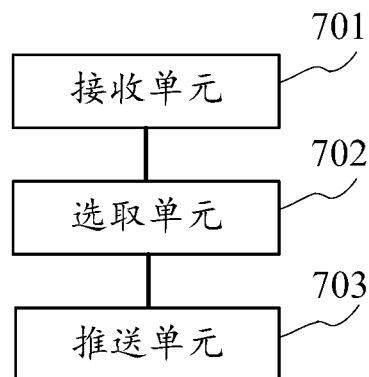
700

图 7

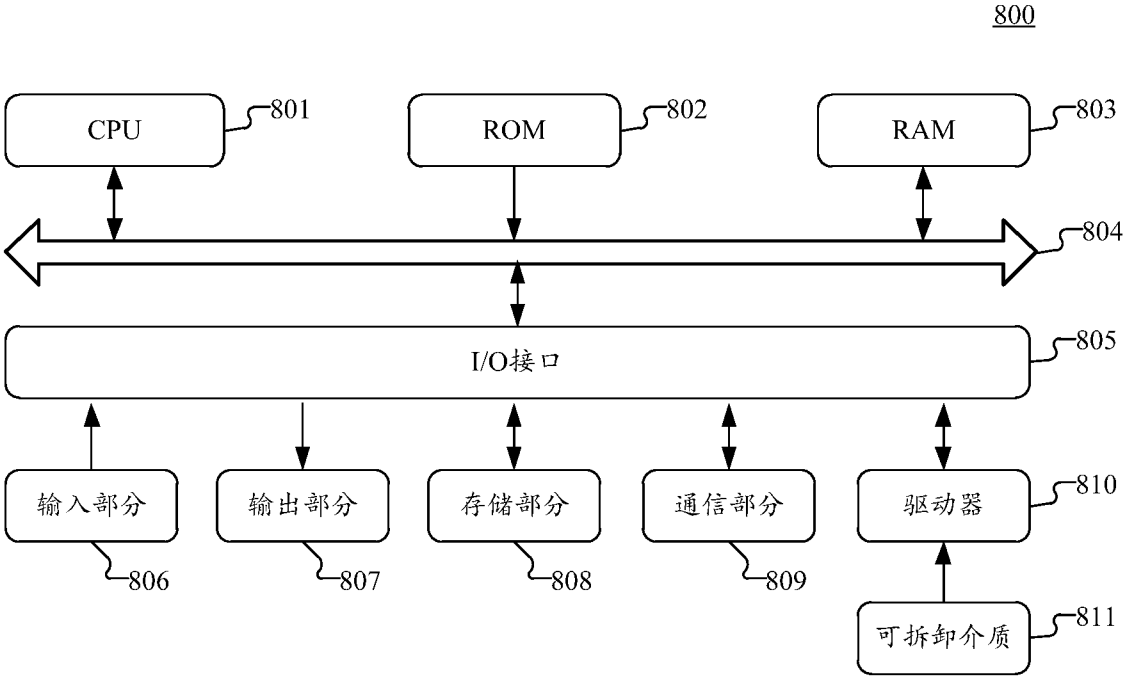


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/095735

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/73(2019.01)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)

G06F16;G06F17/30

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)

CNABS; CNTXT; VEN; DWPI; CNKI; IEEE: 学习, 神经网络, 训练, 配乐, 背景, 音乐, 视频, 音频, learn, neural network, train, soundtrack, background, music, video, audio

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	郝子涵等 (QIE, Zihan et al.). "视频背景音乐选配的人工神经网络模型 (An Artificial Neural Network Model for Video Background Music Selection)" <i>电脑知识与技术 (Computer Knowledge and Technology)</i> , Vol. 13, No. (21), 25 July 2017 (2017-07-25), ISSN: 1009-3044, pages 173-175	1, 5-7, 11-12, 15-16
A	CN 107888843 A (SHENZHEN THUNDER NETWORK TECHNOLOGY CO., LTD.) 06 April 2018 (2018-04-06) entire document	1-16
A	CN 107273458 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 20 October 2017 (2017-10-20) entire document	1-16
A	EP 3276617 A1 (BOOKTRACK HOLDINGS LTD.) 31 January 2018 (2018-01-31) entire document	1-16
A	US 2017125059 A1 (FACEBOOK INC.) 04 May 2017 (2017-05-04) 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

10 October 2019

Date of mailing of the international search report

17 October 2019

Name and mailing address of the ISA/CN

China National Intellectual Property Administration
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/095735

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109492128 A (BEIJING ZIJIE TIAODONG NETWORK TECHNOLOGY CO., LTD.) 19 March 2019 (2019-03-19) claims 1-16	1-16

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/095735

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	107888843	A	06 April 2018	None			
CN	107273458	A	20 October 2017	None			
EP	3276617	A1	31 January 2018	US	2018032611	A1	01 February 2018
US	2017125059	A1	04 May 2017	US	10381044	B2	13 August 2019
CN	109492128	A	19 March 2019	None			

国际检索报告

国际申请号

PCT/CN2019/095735

A. 主题的分类 G06F 16/73 (2019.01) 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) G06F16;G06F17/30 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)) CNABS;CNTXT;VEN;DWPI;CNKI;IEEE:学习, 神经网络, 训练, 配乐, 背景, 音乐, 视频, 音频, learn, neural network, train, soundtrack, background, music, video, audio																							
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>X</td> <td>鄧子涵等. “视频背景音乐选配的人工神经网络模型” 电脑知识与技术, 第13卷, 第21期, 2017年 7月 25日 (2017 - 07 - 25), ISSN: 1009-3044, 第173-175页</td> <td>1, 5-7, 11-12, 15-16</td> </tr> <tr> <td>A</td> <td>CN 107888843 A (深圳市迅雷网络技术有限公司) 2018年 4月 6日 (2018 - 04 - 06) 全文</td> <td>1-16</td> </tr> <tr> <td>A</td> <td>CN 107273458 A (百度在线网络技术北京有限公司) 2017年 10月 20日 (2017 - 10 - 20) 全文</td> <td>1-16</td> </tr> <tr> <td>A</td> <td>EP 3276617 A1 (BOOKTRACK HOLDINGS LTD) 2018年 1月 31日 (2018 - 01 - 31) 全文</td> <td>1-16</td> </tr> <tr> <td>A</td> <td>US 2017125059 A1 (FACEBOOK INC) 2017年 5月 4日 (2017 - 05 - 04) 全文</td> <td>1-16</td> </tr> <tr> <td>PX</td> <td>CN 109492128 A (北京字节跳动网络技术有限公司) 2019年 3月 19日 (2019 - 03 - 19) 权利要求1-16</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	X	鄧子涵等. “视频背景音乐选配的人工神经网络模型” 电脑知识与技术, 第13卷, 第21期, 2017年 7月 25日 (2017 - 07 - 25), ISSN: 1009-3044, 第173-175页	1, 5-7, 11-12, 15-16	A	CN 107888843 A (深圳市迅雷网络技术有限公司) 2018年 4月 6日 (2018 - 04 - 06) 全文	1-16	A	CN 107273458 A (百度在线网络技术北京有限公司) 2017年 10月 20日 (2017 - 10 - 20) 全文	1-16	A	EP 3276617 A1 (BOOKTRACK HOLDINGS LTD) 2018年 1月 31日 (2018 - 01 - 31) 全文	1-16	A	US 2017125059 A1 (FACEBOOK INC) 2017年 5月 4日 (2017 - 05 - 04) 全文	1-16	PX	CN 109492128 A (北京字节跳动网络技术有限公司) 2019年 3月 19日 (2019 - 03 - 19) 权利要求1-16	1-16
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
X	鄧子涵等. “视频背景音乐选配的人工神经网络模型” 电脑知识与技术, 第13卷, 第21期, 2017年 7月 25日 (2017 - 07 - 25), ISSN: 1009-3044, 第173-175页	1, 5-7, 11-12, 15-16																					
A	CN 107888843 A (深圳市迅雷网络技术有限公司) 2018年 4月 6日 (2018 - 04 - 06) 全文	1-16																					
A	CN 107273458 A (百度在线网络技术北京有限公司) 2017年 10月 20日 (2017 - 10 - 20) 全文	1-16																					
A	EP 3276617 A1 (BOOKTRACK HOLDINGS LTD) 2018年 1月 31日 (2018 - 01 - 31) 全文	1-16																					
A	US 2017125059 A1 (FACEBOOK INC) 2017年 5月 4日 (2017 - 05 - 04) 全文	1-16																					
PX	CN 109492128 A (北京字节跳动网络技术有限公司) 2019年 3月 19日 (2019 - 03 - 19) 权利要求1-16	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 10月 10日		国际检索报告邮寄日期 2019年 10月 17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		授权官员 许菲菲 电话号码 62411752																					

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/095735

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	107888843	A	2018年 4月 6日	无	
CN	107273458	A	2017年 10月 20日	无	
EP	3276617	A1	2018年 1月 31日	US 2018032611 A1	2018年 2月 1日
US	2017125059	A1	2017年 5月 4日	US 10381044 B2	2019年 8月 13日
CN	109492128	A	2019年 3月 19日	无	