

(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
2022 年 6 月 9 日 (09.06.2022)



(10) 国际公布号
WO 2022/116888 A1

(51) 国际专利分类号:
H04N 21/239 (2011.01)

(21) 国际申请号: PCT/CN2021/133035

(22) 国际申请日: 2021 年 11 月 25 日 (25.11.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202011390109.2 2020 年 12 月 2 日 (02.12.2020) CN

(71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(72) 发明人: 郭卉 (GUO, Hui); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(74) 代理人: 北京德琦知识产权代理有限公司 (DEQI INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区知春路 1 号学院国际大厦 7 层, Beijing 100083 (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, IT, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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,

(54) Title: METHOD AND DEVICE FOR VIDEO DATA PROCESSING, EQUIPMENT, AND MEDIUM

(54) 发明名称: 一种视频数据处理方法、装置、设备以及介质

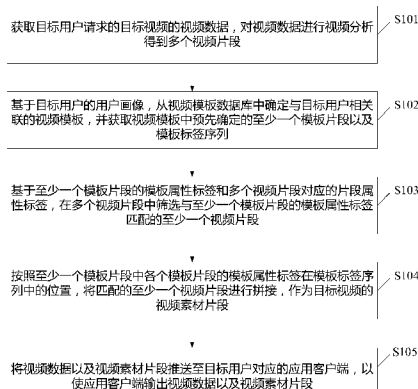


图 3

- S101 Acquire video data of target video requested by target user, and perform video analysis with respect to video data to produce multiple video clips
- S102 Determine, on the basis of user profile of target user, video template associated with target user in video template database, and acquire at least one predetermined template clip and template label sequence in video template
- S103 Screen, on the basis of template attribute label of at least one template clip and of clip attribute labels corresponding to multiple video clips, in multiple video clips for at least one video clip matching template attribute label of at least one template clip
- S104 According to position of each template attribute label of each template clip in at least one template clip in template label sequence, splice at least one matching video clip to serve as stock video clip of target video
- S105 Push video data and stock video clip to application client corresponding to target user, thus allowing application client to output video data and stock video clip

(57) Abstract: Provided in the embodiments of the present application are a method and device for video data processing, equipment, and a medium. The method relates to the field of artificial intelligence. The method comprises: acquiring video data of a target video requested by a target user, and performing a video analysis with respect to the video data to produce multiple video clips; determining, on the basis of a user profile of the target user, a video template associated with the target user in a video template database, and acquiring at least one predetermined template clip and a template label sequence in the video template; screening, on the basis of a template attribute label of the at least one template clip and clip attribute labels corresponding to the multiple video clips, in the multiple video clips for at least one video clip matching the template attribute label of the at least one template clip; according to the position of the template attribute label of each template clip in the at least one template clip in the template label sequence, splicing the at least one matching video clip to serve as a stock video clip of the target video; and pushing the video data and the stock video clip to an application client corresponding to the target user, thus allowing the application client to output the video data and the stock video clip.

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))。

(57) 摘要：本申请实施例提供了一种视频数据处理方法、装置、设备以及介质，该方法涉及人工智能领域，方法包括：获取目标用户请求的目标视频的视频数据，对视频数据进行视频分析得到多个视频片段；基于目标用户的用户画像，从视频模板数据库中确定与目标用户相关联的视频模板，并获取视频模板中预先确定的至少一个模板片段以及模板标签序列；基于至少一个模板片段的模板属性标签和多个视频片段对应的片段属性标签，在多个视频片段中筛选与至少一个模板片段的模板属性标签匹配的至少一个视频片段；按照至少一个模板片段中各个模板片段的模板属性标签在模板标签序列中的位置，将匹配的至少一个视频片段进行拼接，作为目标视频的视频素材片段；将视频数据以及视频素材片段推送至目标用户对应的应用客户端，以使应用客户端输出视频数据以及视频素材片段。

一种视频数据处理方法、装置、设备以及介质

本申请要求 2020 年 12 月 02 日提交的申请号为 202011390109.2、发明名称为“一种视频数据处理方法、装置、设备以及介质”的中国专利申请的优先权。

技术领域

本申请涉及计算机技术领域，尤其涉及一种视频数据处理方法、装置、设备以及介质。

背景技术

随着多媒体技术的发展，视频已成为人们日常生活中获取信息与享受娱乐的主要载体。因为各类视频播放平台的普及，衍生出了各式各样的短视频（即精彩视频集锦）。可以理解的是，这里的短视频是指在各类视频播放平台上播放的、适合在移动状态和短时休闲状态下观看的视频内容。

但是目前，在生成短视频的过程中，往往需要人工剪辑素材、人工合成视频、以及人工配乐、音频视频合成等。

发明内容

本申请实施例提供一种视频数据处理方法、装置、设备以及介质，通过对视频数据进行视频分析（例如，视频分镜和属性分析等），可以快速得到携带片段属性标签的一个或者多个视频片段，使用基于目标用户的用户画像确定的视频模板，进行目标视频的视频片段的属性标签匹配，生成目标视频的视频素材片段，可以随着视频模板的增加和更新，重复利用这些视频片段的分镜和属性信息，减少对目标视频中视频帧的识别和处理，提高短视频的生成效率，节省针对不同用户不断生成和分发大量短视频的计算成本，节省服务器的计算资源。

本申请实施例一方面提供了一种视频数据处理方法，包括：

获取目标用户请求的目标视频的视频数据，对视频数据进行视频分析得到多个视频片段，其中，视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析，多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段；

基于目标用户的用户画像，从视频模板数据库中确定与目标用户相关联的视频模板，并获取视频模板中预先确定的至少一个模板片段以及模板标签序列，所述模板标签序列由至少一个模板片段的模板属性标签构成；

基于至少一个模板片段的模板属性标签和多个视频片段对应的片段属性标签，在多个视频片段中筛选与至少一个模板片段的模板属性标签匹配的至少一个视频片段；

按照所述至少一个模板片段中各个模板片段的模板属性标签在模板标签序列中的位置，将匹配的至少一个视频片段进行拼接，作为目标视频的视频素材片段；

将视频数据以及视频素材片段推送至目标用户对应的应用客户端，以使应用客户端输出视频数据以及视频素材片段。

本申请实施例一方面提供了一种视频数据处理装置，包括：

片段生成模块，用于获取目标用户请求的目标视频的视频数据，对视频数据进行视频分析得到多个视频片段，其中，视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析，多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段；

模板获取模块，用于基于目标用户的用户画像，从视频模板数据库中确定与目标用户相关联的视频模

板，并获取视频模板中预先确定的至少一个模板片段以及模板标签序列，模板标签序列由至少一个模板片段的模板属性标签构成；

素材确定模块，用于基于至少一个模板片段的模板属性标签和多个视频片段对应的片段属性标签，在多个视频片段中筛选与至少一个模板片段的模板属性标签匹配的至少一个视频片段，按照所述至少一个模板片段中各个模板片段的模板属性标签在模板标签序列中的位置，将匹配的至少一个视频片段进行拼接，作为目标视频的视频素材片段；

数据发送模块，用于将视频数据以及视频素材片段推送至目标用户对应的应用客户端，以使应用客户端输出视频数据以及视频素材片段。

本申请实施例一方面提供了一种视频数据处理方法，包括：

响应目标用户针对应用客户端中的目标视频执行的播放操作，从服务器上获取目标视频的视频数据，以及与目标视频相关联的视频素材片段；视频素材片段是由服务器对视频数据进行视频分析得到多个视频片段，其中，视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析，多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段（也就是说，每个视频片段即为对应于一个片段属性标签的一个分镜片段）；基于目标用户的用户画像，从视频模板数据库中确定与目标用户相关联的视频模板，并获取视频模板中预先确定的至少一个模板片段以及模板标签序列，模板标签序列由至少一个模板片段的模板属性标签构成；基于至少一个模板片段的模板属性标签和多个视频片段对应的片段属性标签，在多个视频片段中筛选与至少一个模板片段的模板属性标签匹配的至少一个视频片段；按照至少一个模板片段中各个模板片段的模板属性标签在模板标签序列中的位置，将匹配的至少一个视频片段进行拼接得到的；

在应用客户端的应用显示界面中输出视频数据以及视频素材片段。

本申请实施例一方面提供了一种视频数据处理装置，包括：

数据获取模块，用户响应目标用户针对应用客户端中的目标视频执行的播放操作，从服务器上获取目标视频的视频数据，以及与目标视频相关联的视频素材片段；视频素材片段是由服务器对视频数据进行视频分析得到多个视频片段，其中，视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析，多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段；基于目标用户的用户画像，从视频模板数据库中确定与目标用户相关联的视频模板，并获取视频模板中预先确定的至少一个模板片段以及模板标签序列，模板标签序列由至少一个模板片段的模板属性标签构成；基于至少一个模板片段的模板属性标签和多个视频片段对应的片段属性标签，在多个视频片段中筛选与至少一个模板片段的模板属性标签匹配的至少一个视频片段；按照至少一个模板片段中各个模板片段的模板属性标签在模板标签序列中的位置，将匹配的至少一个视频片段进行拼接得到的；

数据输出模块，用于在应用客户端的应用显示界面中输出视频数据以及视频素材片段。

本申请实施例一方面提供了一种计算机设备，包括存储器和处理器，存储器存储有计算机程序，计算机程序被处理器执行时，使得处理器执行本申请实施例提供的方法。

本申请实施例一方面提供了一种计算机可读存储介质，计算机可读存储介质存储有计算机程序，计算机程序包括程序指令，程序指令当被处理器执行时，执行如本申请实施例提供的方法。

本申请实施例一方面提供了一种计算机程序产品或计算机程序，该计算机程序产品或计算机程序包括计算机指令，该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令，处理器执行该计算机指令，使得该计算机设备执行本申请实施例提供的方法。

附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要

使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图 1 是本申请实施例提供的一种网络架构的结构示意图；

图 2 是本申请实施例提供的一种进行数据交互的场景示意图；

图 3 是本申请实施例提供的一种视频数据处理方法的流程示意图；

图 4 是本申请实施例提供的一种查询视频数据的场景示意图；

图 5 是本申请实施例提供的一种进行分镜处理的场景示意图；

图 6 是本申请实施例提供的一种提取片段属性标签的流程示意图；

图 7 是本申请实施例提供的一种获取视频模板的场景示意图；

图 8A 是本申请实施例提供的一种对模板视频进行视频分析的场景示意图；

图 8B 是本申请实施例提供的一种对目标视频进行视频分析的场景示意图；

图 9 是本申请实施例提供的一种视频数据处理方法的流程示意图；

图 10 是本申请实施例提供的一种生成视频素材片段的流程示意图；

图 11 是本申请实施例提供的一种前后端交互的流程示意图；

图 12A 是本申请实施例提供的一种输出视频素材片段的场景示意图；

图 12B 是本申请实施例提供的一种更新视频素材片段的场景示意图；

图 13 是本申请实施例提供的一种视频数据处理装置的结构示意图；

图 14 是本申请实施例提供的一种视频数据处理装置的结构示意图；

图 15 是本申请实施例提供的一种计算机设备的结构示意图；

图 16 是本申请实施例提供的一种视频数据处理系统。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，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他实施例，都属于本申请保护的范围。

应当理解，人工智能(Artificial Intelligence, 简称 AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能，感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说，人工智能是计算机科学的一个综合技术，它企图了解智能的实质，并生产出一种新的能以人类智能相似的方式做出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法，使机器具有感知、推理与决策的功能。

人工智能技术是一门综合学科，涉及领域广泛，既有硬件层面的技术也有软件层面的技术。人工智能基础技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术。人工智能软件技术主要包括计算机视觉技术、语音处理技术、自然语言处理技术以及机器学习/深度学习等方向。

其中，计算机视觉(Computer Vision, 简称 CV)是一门研究如何使机器“看”的科学，更进一步的说，就是指用摄影机和电脑代替人眼对目标进行识别、跟踪和测量等机器视觉，并进一步做图形处理，使电脑处理成为更适合人眼观察或传送给仪器检测的图像。作为一个科学学科，计算机视觉研究相关的理论和技術，试图建立能够从图像或者多维数据中获取信息的人工智能系统。计算机视觉技术通常包括图像处理、图像识别、图像语义理解、图像检索、OCR、视频处理、视频语义理解、视频内容/行为识别、三维物体重建、3D 技术、虚拟现实、增强现实、同步定位与地图构建等技术，还包括常见的人脸识别、指纹识别等

生物特征识别技术。

具体的，请参见图 1，图 1 是本申请实施例提供的一种网络架构的结构示意图。如图 1 所示，该网络架构可以包括业务服务器 2000 和用户终端集群。其中，用户终端集群具体可以包括一个或者多个用户终端，这里将不对用户终端集群中的用户终端的数量进行限制。如图 1 所示，多个用户终端具体可以包括用户终端 3000a、用户终端 3000b、用户终端 3000c、...、用户终端 3000n。其中，用户终端 3000a、用户终端 3000b、用户终端 3000c、...、用户终端 3000n 可以分别与业务服务器 2000 通过有线或无线通信方式进行直接或间接地网络连接，以便于每个用户终端可以通过该网络连接与业务服务器 2000 之间进行数据交互。

其中，如图 1 所示的业务服务器 2000 可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群或者分布式系统，还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN、以及大数据和人工智能平台等基础云计算服务的云服务器。

应当理解，如图 1 所示的用户终端集群中的每个用户终端均可以集成安装有应用客户端，当该应用客户端运行于各用户终端中时，可以分别与上述图 1 所示的业务服务器 2000 之间进行数据交互。其中，该应用客户端可以理解作为一种能够加载并显示视频数据的应用，例如，这里的应用客户端具体可以包括：车载客户端、智能家居客户端、娱乐客户端（例如，游戏客户端）、多媒体客户端（例如，视频客户端）、社交客户端以及资讯类客户端（例如，新闻客户端）等。其中，为便于理解，本申请实施例可以在图 1 所示的多个用户终端中选择一个用户终端作为目标用户终端，该目标用户终端可以包括：智能手机、平板电脑、笔记本电脑、智能电视等具有视频数据加载功能的智能终端。例如，本申请实施例可以将图 1 所示的用户终端 3000a 作为目标用户终端。

为便于理解，比如，当用户 Y（即目标用户）在上述目标用户终端中需要播放某个视频（比如，该用户 Y 自己感兴趣的视频）时，该目标用户终端可以响应用户 Y 针对该视频的触发操作，向图 1 所示的业务服务器 2000 发送视频播放请求。这样，该业务服务器 2000 可以基于该视频播放请求在视频业务数据库中查找到这个视频的视频数据，进而可以将这个视频数据和与这个视频数据相关联的一个或者多个视频素材片段（例如，这个视频的视频花絮等）返回给目标用户终端，以在该目标用户终端中对该用户 Y 所请求播放的这个视频的视频数据进行播放处理。在一实施方式中，与此同时，该目标用户终端还可以在播放这个视频数据时，一并显示接收到的这些视频素材片段。可以理解的是，这里的视频素材片段可以由该业务服务器 2000 按照选取的视频模板的模板片段以及模板片段对应的模板标签序列，对目标视频的视频片段进行筛选后所得到的；此外，可以理解的是，这里的视频片段可以由该业务服务器 2000 对视频数据进行视频分析后所得到的；应当理解，这里的视频模板可以由该业务服务器 2000 基于该用户 Y（即目标用户）的用户画像所确定的。

其中，可以理解的是，本申请实施例可以将上述用户 Y（即目标用户）在应用客户端（例如，视频客户端 K）中所选择的贴合自己兴趣的视频（比如，电视剧或者短视频等）统称为目标视频。

其中，本申请实施例中的视频素材片段可以由上述业务服务器 2000 根据上述视频模板的模板片段和模板标签序列所智能生成的。比如，业务服务器 2000 可以通过本申请实施例所涉及的视频数据处理方法，智能生成上述用户 Y 在上述目标用户终端中所选择的目标视频（例如，电视剧 S1）的一个或多个视频素材片段。其中，视频素材片段的生成过程是指业务服务器 2000 可以将电视剧 S1 的视频片段的标签信息（即片段属性标签）与模板视频（例如，视频 M）的模板片段的标签信息（即模板属性标签）进行标签匹配以及内容相似度匹配的过程，进而可以根据标签匹配以及内容相似度匹配的结果，从电视剧 S1 的视频片段中筛选与视频 M 的模板片段中的每个模板片段具有相似视频播放效果的视频片段，从而可以根据筛选出的这些视频片段所构成的拼接视频数据以及模板片段的模板音频数据，智能生成与视频 M 相似的视频

素材片段。

应当理解，上述网络框架适用于人工智能领域（即 AI 领域），该 AI 领域所对应的业务场景可以为视频分类场景、视频推荐场景等，这里将不对具体的业务场景进行一一列举。

其中，这里的视频分类场景主要是指计算机设备（例如，上述业务服务器 2000）在通过上述视频数据处理方法，可以将同一视频下的视频片段存储于第一业务数据库。比如，计算机设备在基于某个视频模板生成视频素材片段（例如，基于视频模板 B1 生成的视频素材片段 A1 和基于视频模板 B2 生成的视频素材片段 A2）之后，还可以将视频素材片段 A1 和视频素材片段 A2 添加至相应的短视频推荐数据库，这里的短视频推荐数据库至少可以包含第一业务数据库和第二业务数据库。其中，这里的第一业务数据库可以用于存储与同一视频相关联的一个或者多个视频素材片段。比如，若视频素材片段 A1 和视频素材片段 A2 均属于同一视频（例如，视频 W）的视频片段，则可以将视频素材片段 A1 和视频素材片段 A2 添加至这个视频 W 所对应的第一业务数据库。在一实施方式中，若视频素材片段 A1 和视频素材片段 A2 分别属于不同视频的视频片段，比如，若视频素材片段 A1 对应的目标视频为用户 Y1 所请求的视频 W1，则可以将视频素材片段 A1 添加至这个视频 W1 所对应的第一业务数据库；若视频素材片段 A2 对应的目标视频为用户 Y2 所请求的视频 W2，则可以将视频素材片段 A2 添加至这个视频 W2 所对应的第一业务数据库。

其中，这里的第二业务数据库可以用于存储与同一视频模板相关联的一个或者多个视频素材片段。这意味着本申请实施例可以在不同视频的视频素材片段中，将使用同一视频模板的视频片段添加至第二业务数据库。例如，计算机设备在基于某个视频模板生成视频素材片段（例如，基于视频模板 B 生成视频素材片段 A）之后，还可以将视频素材片段 A 添加至视频模板 B 所对应的第二业务数据库。为便于理解，比如，若该视频模板 B 为表情集锦类，则可以将视频素材片段 A 添加至这个表情集锦类所对应的第二业务数据库。又比如，若该视频模板 B 为故事情节集锦类，则可以将该视频素材片段 A 添加至这个故事情节集锦类所对应的第二业务数据库。再比如，若该视频模板 B 为人物混剪集锦类，则可以将该视频素材片段 A 添加至这个人物混剪集锦类所对应的第二业务数据库。

此外，应当理解，在上述视频推荐场景下，计算机设备（例如，上述业务服务器 2000）在通过上述视频数据处理方法智能生成目标用户所请求的目标视频的视频素材片段（例如，基于视频模板 B1 生成的视频素材片段 A1 和基于视频模板 B2 生成的视频素材片段 A2）之后，还可以将上述同一视频 W 的这些视频片段（即上述第一业务数据库中的视频片段，例如，视频素材片段 A1 和视频素材片段 A2）添加至短视频推荐列表（例如，短视频推荐列表 1），以将该短视频推荐列表 1 中的这些视频片段智能推送给上述目标用户。这样，当目标用户在上述应用客户端中观看完上述视频 W（即目标视频）之后，还可以智能在应用客户端中为上述目标用户遍历播放上述短视频推荐列表 1 中的这些视频片段。比如，当目标用户在上述目标用户终端中观看完上述视频 W 时，计算机设备（例如，上述业务服务器 2000）还可以将短视频推荐列表 1 中的多个视频素材片段中的视频素材片段 A1 输出至应用客户端，以在该应用客户端中实现对该视频素材片段 A1 的智能播放。可以理解的是，该视频素材片段 A1 与上述视频模板 B1 所映射的模板片段具有相似的视频播放效果。

在一实施方式中，计算机设备（例如，上述业务服务器 2000）在通过上述视频数据处理方法智能生成视频素材片段（例如，基于视频模板 B1 生成的视频素材片段 A1 和视频素材片段 A3）之后，还可以将使用同一视频模板 B1 的这些视频片段（即上述第二业务数据库中的视频片段，例如，视频素材片段 A1 和视频素材片段 A3）添加至另一短视频推荐列表（例如，短视频推荐列表 2），以将该短视频推荐列表 2 中的这些视频片段智能推送给上述目标用户。这样，当目标用户在上述应用客户端中观看短视频推荐列表 2 中的视频片段（例如，视频素材片段 A1 之后，还可以智能在应用客户端中为上述目标用户遍历播放上述短视频推荐列表 2 中的这些视频片段。比如，当目标用户在上述目标用户终端中观看完上述视频素材片段

A1 时, 计算机设备(例如, 上述业务服务器 2000)还可以将短视频推荐列表 2 中的其他视频素材片段(例如, 视频素材片段 A3) 输出至应用客户端, 以在该应用客户端中实现对该视频素材片段 A3 的智能播放。可以理解的是, 由于该视频素材片段 A3 与上述视频模板 B1 使用的是同一视频模板, 所以, 当在应用客户端中播放该视频素材片段 A3 与上述视频模板 B1 时, 将为目标用户呈现出与上述模板片段 B1 所映射的模板片段相似的视频播放效果。

为便于理解, 进一步的, 请参见图 2, 图 2 是本申请实施例提供的一种进行数据交互的场景示意图。其中, 如图 2 所示的服务器可以为上述图 1 所对应实施例中的业务服务器 2000, 如图 2 所示的用户终端 X 可以为上述图 1 所对应实施例的用户终端集群中的任意一个用户终端。为便于理解, 本申请实施例以上述图 1 所示的用户终端 3000a 作为该用户终端 X 为例, 以阐述在该业务场景为视频推荐场景下, 为目标用户推荐视频素材片段的具体过程。

如图 2 所示, 视频推荐界面 200a 中可以包含多个推荐视频数据, 这里的多个推荐视频数据具体可以包括图 2 所示的视频数据 20a、视频数据 20b、视频数据 20c 和视频数据 20d。可以理解的是, 本申请实施例可以将展示在视频推荐界面 200a 中的视频数据 20a、视频数据 20b、视频数据 20c 和视频数据 20d 统称为推荐视频数据。

其中, 如图 2 所示, 当目标用户需要播放某个推荐视频数据(例如, 视频数据 20b)时, 可以将该目标用户从该视频推荐界面 200a 中所选取的视频数据 20b 统称为应用显示界面中的目标视频。此时, 用户终端可以响应针对应用显示界面中的该目标视频的播放操作, 向图 2 所示的服务器发送视频播放请求。此时, 服务器可以响应该视频播放请求, 以在应用客户端输出该目标视频对应的视频播放界面, 例如, 可以在应用客户端中输出视频数据 20b 对应的视频播放界面, 该视频数据 20b 对应的视频播放界面可以为图 2 所示的视频播放界面 200b。其中, 应用显示界面中可以包含用于播放目标视频的视频播放界面 200b, 还可以包括用于展示视频素材片段的短视频推荐列表, 该短视频推荐列表中可以包含与该目标视频相关联的视频素材片段。

其中, 可以理解的是, 服务器在接收目标用户通过用户终端发送的视频播放请求时, 可以从该视频播放请求中获取目标视频的视频标识, 并根据视频标识在视频业务数据库中查询目标视频的视频数据。在查询到目标视频的视频数据后, 服务器可以对该视频数据的视频序列进行上述视频分析, 以得到该视频数据的视频片段, 这里的视频片段具体可以包括图 2 所示的视频片段 100a、视频片段 100b、...、视频片段 100k, 这里的每个视频片段都可以对应一个片段属性标签。

进一步的, 服务器可以基于该目标用户的用户画像, 获取贴合该目标用户观影兴趣的视频模板, 进而可以获取该视频模板所映射的模板片段以及该模板片段所对应的模板标签序列, 以便于能够根据该模板标签序列, 从上述视频片段中筛选与每个模板片段相匹配的视频片段(即满足片段匹配条件的视频片段), 进而可以基于这些筛选出的满足片段匹配条件的视频片段, 得到视频素材片段。由此可见, 本申请实施例可以尽可能地从中获取与模板片段有相同标签序列特征的视频片段, 进而可以按照相同的标签序列(即上述模板标签序列)填充得到上述视频素材片段(比如, 可以得到上述目标视频的一个或者多个短视频), 以使用户终端可以将该视频素材片段和上述视频数据输出至应用客户端。其中, 可以理解的是, 一个视频模板, 可以对应一个或者多个视频素材片段, 比如, 这里将不对从目标视频的视频片段中所筛选出的具有相同标签序列特征的视频素材片段的数量进行限定。

为便于理解, 本申请实施例以一个视频模板对应一个视频素材片段为例。那么, 当服务器确定出贴合该目标用户的观影兴趣的视频模板的数量有多个(例如, N 个)时, 本申请实施例还可以用于将智能生成 N 个视频素材片段的 N 个视频模板统称为视频模板。应当理解, 通过 N 个视频模板智能生成其他视频素材片段的具体实现方式, 可以一并参见对智能生成上述生成视频素材片段的具体过程的描述, 这里将不再

继续进行赘述。

其中，在目标用户终端中推荐视频素材片段的具体实现方式，可以参见下述图 3-图 12B 所对应的实施例。

进一步的，请参见图 3，图 3 是本申请实施例提供的一种视频数据处理方法的流程示意图。如图 3 所示，该方法可以由应用客户端执行，也可以由服务器执行，还可以由应用客户端和服务器共同执行。该应用客户端可以为上述图 2 所对应实施例中的用户终端 X 中运行的应用客户端，该服务器可以为上述图 2 所对应实施例中的服务器。为便于理解，本实施例以该方法由服务器执行为例进行说明，以阐述在服务器中基于视频模板生成目标视频对应的视频素材片段的具体过程。其中，该方法至少可以包括以下步骤 S101-步骤 S105：

步骤 S101，获取目标用户请求的目标视频的视频数据，对视频数据进行视频分析得到多个视频片段，其中，所述视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析，所述多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段。

具体的，服务器可以获取目标用户请求的目标视频的视频数据以及与视频数据相关联的网络识别模型。进一步的，服务器可以通过视频切分组件将视频数据对应的视频序列进行分镜处理，得到与视频序列相关联的多个分镜片段。进一步的，服务器可以将多个分镜片段输入至网络识别模型，由网络识别模型基于多个预设的片段属性标签，对多个分镜片段进行属性分析，得到多个分镜片段对应的片段属性标签。进一步的，服务器可以将具备片段属性标签的多个分镜片段确定为视频数据的多个视频片段。其中，一个视频片段可以对应一个片段属性标签。

应当理解，在获取目标用户请求的目标视频的视频数据之前，服务器可以接收应用客户端发送的视频播放请求。其中，视频播放请求是由应用客户端响应目标用户针对目标视频执行的播放操作所生成的。进一步的，服务器可以从视频播放请求中提取目标视频的视频标识，基于视频标识在视频业务数据库中查找目标视频对应的业务视频数据，将查找到的业务视频数据作为应用客户端中的目标视频的视频数据。

为便于理解，请参见图 4，图 4 是本申请实施例提供的一种查询视频数据的场景示意图。如图 4 所示，这里的应用显示界面 400a 可以为上述图 2 所对应实施例中的应用显示界面 200a。在目标用户针对应用客户端的应用显示界面 400a 中的视频数据 40b 执行触发操作（即播放操作）时，应用客户端可以将该视频数据 40b 作为目标视频，并向服务器发送携带该视频数据 40b 的视频标识的视频播放请求，进而服务器可以接收应用客户端发送的视频播放请求，获取该视频播放请求所携带的关于视频数据 40b 的视频标识，并基于该视频标识在应用客户端对应的视频业务数据库中查找该视频标识对应的业务视频数据，并将查找到的业务视频数据作为视频数据 40b 对应的视频数据。

其中，可以理解的是，这里的目标视频可以为综艺节目、电影、电视剧等长视频，还可以为从长视频中截取的短视频等，本申请对此不做限制。

应当理解，服务器通过视频切分组件将视频数据对应的视频序列进行分镜处理，得到与视频序列相关联的多个分镜片段的具体过程可以描述为：该服务器可以在获取用于对视频数据的视频序列进行分镜处理的视频切分组件时，通过视频切分组件在视频序列中确定用于作为聚类质心的第一视频帧，并创建该第一视频帧所属的分镜簇的分镜簇信息（可以理解的是，这里的分镜簇信息可以为配置的相应分镜簇的标识）。进一步的，服务器可以在视频序列中将除第一视频帧之外的视频帧确定为第二视频帧，并可以基于轮询机制依次获取每个第二视频帧，以确定每个第二视频帧与第一视频帧的图像相似度。进一步的，若第一视频帧与某个第二视频帧的图像相似度大于或者等于聚类阈值，则服务器可以将图像相似度大于或者等于聚类阈值的该第二视频帧划分到第一视频帧所属的分镜簇。进一步的，若第一视频帧与某个第二视频帧的图像相似度小于聚类阈值，则服务器可以用图像相似度小于聚类阈值的该第二视频帧更新第一视频帧（例如，

将该第二视频帧作为更新后的第一视频帧),并创建更新后的第一视频帧所属的另一分镜簇的分镜簇信息,进而可以将更新后的第一视频帧依次与未匹配的第二视频帧进行图像相似度匹配,直到视频序列中的视频帧均完成图像相似度匹配时,可以得到视频序列中的视频帧所属的分镜簇的分镜簇信息(即可以划分得到视频序列中的每个视频帧所属的分镜簇)。进一步的,服务器可以基于视频序列中的视频帧所属的分镜簇的分镜簇信息,确定与视频序列相关联的分镜片段,即,将视频序列中的视频帧组成多个分镜片段。

可以理解的是,图像相似度匹配指的是对两幅图像之间内容的相似程度进行计算,可以得到用于判断图像内容的相似程度的图像相似度。若图像相似度越大,则表明这两幅图像越相似,若图像相似度越小,则表明这两幅图像越不相似。其中,两幅图像之间内容的相似程度可以使用不同的方法来衡量。比如,在使用余弦相似度的情况下,可以把图片表示成一个向量,通过计算向量之间的余弦距离来表征两张图片的相似度;直方图可以描述一幅图像中颜色的全局分布,直方图相似度是另一种图像相似度计算方法;结构相似度是一种全参考的图像质量评价指标,分别从亮度、对比度、结构三个方面度量图像相似性。应当理解,本申请对图像相似度匹配时具体使用的方法不做限制。

为便于理解,请参见图5,图5是本申请实施例提供的一种进行分镜处理的场景示意图。如图5所示的视频序列可以包括多个视频帧,具体可以包括图2所示的n个视频帧,n可以为大于1的正整数,这n个视频帧具体可以包括:视频帧10a、视频帧10b、视频帧10c、视频帧10d、...、视频帧10n。应当理解,通过聚类算法可以计算该视频序列中的视频帧之间的图像相似度,从而可以基于计算得到的视频帧之间的图像相似度,将该视频序列中的视频帧划分到不同的聚类簇(即分镜簇)。比如,通过聚类算法可以得到图5所示的k个聚类簇(即k个分镜簇),且这k个聚类簇具体可以包含图5所示的聚类簇20a、聚类簇20b、...、聚类簇20k。其中,可以理解的是,图5所示的k个聚类簇中每个聚类簇中均至少可以包括一个视频帧。

具体的,本申请可以在图5所示的视频序列中,将该视频序列中的首个视频帧(即视频帧10a)称之为能够用于作为聚类质心(即聚类质心1)的第一视频帧,并可以将该视频序列中除视频帧10a之外的视频帧确定为第二视频帧,并可以基于轮询机制依次获取第二视频帧(即视频帧10b、视频帧10c、...、视频帧10n),以依次计算第一视频帧与第二视频帧的图像相似度。其中,本申请可以创建该聚类质心1所属的分镜簇(即分镜簇1),进而可以将视频帧10b与视频帧10a进行图像相似度匹配。在视频帧10b与视频帧10a的图像相似度(例如,相似度1)大于或等于聚类阈值时,将相似度1所对应的视频帧10b划分到视频帧10a所属的分镜簇(即分镜簇1)中。同理,本申请可以将视频帧10c划分到视频帧10a所属的分镜簇(即分镜簇1)中。

进一步的,由于视频帧10d为视频帧10c的下一视频帧,因此,本申请可以将视频帧10d与视频帧10a进行图像相似度匹配。在视频帧10d与视频帧10a的图像相似度(例如,相似度2)小于聚类阈值时,根据视频帧10d更新第一视频帧,以将相似度2所对应的视频帧10d作为更新后的第一视频帧以及新的聚类质心(即聚类质心2),并可以创建该聚类质心2所属的分镜簇(即分镜簇2),进而,可以基于上述轮询机制依次获取未匹配的第二视频帧(即视频帧10e、...、视频帧10n),以依次计算更新后的第一视频帧与未匹配的第二视频帧的图像相似度。其中,本申请可以将视频帧10e划分到视频帧10d所属的分镜簇(即分镜簇2)中。

其中,可以理解的是,本申请在获取聚类质心1和聚类质心2之后,可以使用同样的方法获取聚类质心3、聚类质心4、...、聚类质心k。同理,本申请在获取分镜簇1(即聚类簇20a)和分镜簇2(即聚类簇20b)之后,可以使用同样的方法获取分镜簇3(即聚类簇20c)、分镜簇4(即聚类簇20d)、...、分镜簇k(即聚类簇20k)。此时,视频序列中的视频帧10a、视频帧10b、...、视频帧10n已经全部完成图像相似度匹配。

由此可见，通过对图 5 所示的视频序列中的视频帧进行聚类处理（即分镜处理），可以得到与该视频序列相关联的多个聚类簇（即分镜簇），从而可以将每个聚类簇中的视频帧构成一个分镜片段，进而可以得到图 5 所示的 k 个分镜片段。比如，可以将聚类簇 20a 中的视频帧 10a、视频帧 10b 和视频帧 10c 构成分镜簇 1 对应的分镜片段（即分镜片段 1），可以将聚类簇 20b 中的视频帧 10d 和视频帧 10e 构成分镜簇 2 对应的分镜片段（即分镜片段 2），...，可以将聚类簇 20k 中的视频帧 10(n-2)、视频帧 10(n-1) 和视频帧 10n 构成分镜簇 k 对应的分镜片段（即分镜片段 k）。

应当理解，将目标视频对应的视频序列划分为多个分镜片段的视频切分组件可以为 pyscenedetect 开源代码库，该 pyscenedetect 开源代码库是一个自动将视频数据分割为单个片段的工具，其中，第一视频帧（聚类质心）的选择可以不限于上述方式。可以理解的是，将目标视频对应的视频序列划分为多个分镜片段的方法还可以为鼓点识别的方式，例如，获取目标视频的音频数据，识别该音频数据中的鼓点，根据鼓点在音频数据中的位置，确定鼓点在目标视频的视频数据中的位置，以对视频数据的视频序列进行划分。其中，将视频序列划分为多个分镜片段的方法还可以为其他方式，本申请对视频分镜具体使用的分镜方法不做限制。

可以理解的是，上述网络识别模型至少包括：具有第一属性标签提取功能的第一网络模型、具有第二属性标签提取功能的第二网络模型和具有第三属性标签提取功能的第三网络模型。应当理解，服务器可以将多个分镜片段输入第一网络模型，通过第一网络模型对多个分镜片段中的每个分镜片段进行远近景分析，得到多个分镜片段的远近景标签，将多个分镜片段的远近景标签作为第一网络模型输出的第一属性标签，将具有第一属性标签的分镜片段作为第一类分镜片段。进一步的，服务器可以将第一类分镜片段输入第二网络模型，由第二网络模型对第一类分镜片段中的每个分镜片段进行人脸检测，得到人脸检测结果。进一步的，若人脸检测结果指示第一类分镜片段中存在目标角色的人脸，则服务器可以在第一类分镜片段中将存在目标角色的人脸所对应的分镜片段作为第二类分镜片段，通过第二网络模型确定第二类分镜片段中的目标角色所属的角色标签，将目标角色所属的角色标签确定为第二类分镜片段的第二属性标签。其中，目标角色为目标视频中的一个或者多个角色。进一步的，服务器可以在第一类分镜片段中将除第二类分镜片段之外的分镜片段，确定为第三类分镜片段，将第三类分镜片段输入第三网络模型，由第三网络模型对第一类分镜片段中的每个分镜片段进行场景检测，得到第三类分镜片段的第三属性标签。进一步的，服务器可以根据第一类分镜片段的第一属性标签、第二类分镜片段的第二属性标签、以及第三类分镜片段的第三属性标签，确定多个分镜片段中的每个分镜片段对应的片段属性标签。

可以理解的是，第一网络模型可以为远近景识别模型，第二网络模型可以为人脸识别模型，第三网络模型可以为场景识别模型。基于此，上述第一网络模型、第二网络模型和第三网络模型还可以为表情识别模型、动作识别模型等，本申请对网络识别模型的具体类型不做限制。同理，网络识别模型可以采用任意深度学习模型或机器学习模型，本申请对网络识别模型使用的具体模型不做限制。应当理解，通过预先训练的网络识别模型，可以提高目标视频的基础分析能力，进而可以快速得到目标视频对应的视频片段。

可以理解的是，可以通过人脸检测模型对分镜片段进行人脸检测，得到人脸检测结果，进而可以通过人脸识别模型确定人脸检测结果中的人脸所对应的角色标签。这里的人脸检测模型和人脸识别模型可以统称为第二网络模型。其中，人脸检测与人脸识别可以统称为图像检测。图像检测表示机器学习技术可以对标注样本数据（例如，图像中多个标注框与标签对的对应关系）进行学习后获得数学模型，在学习训练的过程中可以获得该数学模型的参数，识别预测时加载该数学模型的参数，并计算输入样本存在的实物标签的预测框以及该预测框属于指定范围内某个实物标签的概率，进而可以将具有最大概率的实物标签作为该预测框对应的标签。

其中，可以理解的是，将分镜片段直接输入远近景识别模型，可以获取该分镜片段对应的远近景标签

(即第一属性标签), 将分镜片段直接输入场景识别模型, 可以获取该分镜片段对应的场景标签(即第三属性标签)。在将分镜片段输入第三网络模型之前, 需要提前对人脸进行目标角色的检索, 即可以提前将分镜片段的目标角色输入第二网络模型, 通过该第二网络模型提取该目标角色的特征向量。因此, 在确定分镜片段的角色标签时, 可以在将分镜片段输入第二网络模型时, 提取该分镜片段中的全部视频帧, 并对这些视频帧进行人脸检测, 进而可以将检测到的人脸的特征向量与上述目标角色的特征向量进行比较。若特征向量比较得到的相似度结果大于阈值, 则认为该人脸是目标角色, 将目标角色的角色标签作为该检测到的人脸所在的分镜片段的角色标签(即第二属性标签)。

应当理解, 本申请借助于上述网络识别模型(即第一网络模型、第二网络模型和第三网络模型), 可以在无需知道分镜片段的标签信息(即不需要预先知道任何图像噪声信息)的情况下, 直接将分镜片段输入第一网络模型(即上述远近景识别模型), 以获取该分镜片段对应的第一属性标签(即根据模型自动进行样本噪声程度预测, 以给出每个分镜片段的新标签), 进而可以将具有第一属性标签的分镜片段统称为第一类分镜片段。可以理解的是, 本申请实施例在得到第一类分镜片段之后, 还可以进一步将新标签自动反馈到后续模型中进行学习, 以通过动态噪声预测及处理, 来避免上述网络识别模型陷入局部最优, 保证模型学习是向识别效果更佳的方向进行。另外, 当使用第一网络模型对目标视频的分镜片段进行远近景识别之后, 得到的第一类分镜片段, 在有新的视频模板时, 也可以重复使用, 不需要重复进行视频帧的识别和处理, 进而可以节省计算资源。

比如, 本申请实施例还可以进一步将第一类分镜片段输入至第二网络模型, 以使第二网络模型可以对所述第一类分镜片段中的每个分镜片段进行人脸检测和人脸识别, 进而可以在这些第一类分镜片段中挑选出所有的包含目标角色的人脸的第一类分镜片段。可以理解的是, 本申请实施例可以将挑选出的这些包含目标角色的人脸的第一类分镜片段统称为第二类分镜片段。此外, 可以理解的是, 该第二网络模型还可以用于输出每个第二分镜片段中的目标角色所属的角色标签。基于此, 本申请实施例可以将这里的角色标签所属的角色标签统称为第二类分镜片段的第二属性标签。应当理解, 这里的角色标签可以为目标视频中的一个或者多个角色, 这里将不对目标角色的数量进行限定。进一步的, 本申请实施例还可以在第二分镜片段中将除第二类分镜片段之外的其他分镜片段统称为第三类分镜片段, 进而可以将第三类分镜片段输入第三网络模型(即上述场景识别模型), 以获取该第三类分镜片段对应的第三属性标签。由此可见, 本申请实施例通过上述训练好的网络识别模型可以实时校正分镜片段所属的标签信息, 进而可以根据上述第一属性标签、第二属性标签和第三属性标签, 准确得到每个分镜片段的片段属性标签。

为便于理解, 请参见图 6, 图 6 是本申请实施例提供的一种提取片段属性标签的流程示意图。如图 6 所示的视频数据可以为上述目标视频的视频数据, 则得到分镜片段的标签信息的具体过程可以描述为: 对视频数据的视频序列进行视频分镜可以得到 k 个分镜片段, 进而可以将每个分镜片段输入网络识别模型, 以得到每个分镜片段在网络识别模型下的标签信息。这里的网络识别模型可以为图 6 所示的场景识别模型(即第三网络模型)、远近景识别模型(即第一网络模型)、人脸检测模型和人脸识别模型(即第二网络模型)。

其中, 可以理解的是, 如图 6 所示, 在将 k 个分镜片段输入远近景识别模型之后, 得到的每个分镜片段对应的远近景标签(即第一属性标签)可以为: {分镜 1: x_1 , 分镜 2: x_2 , ..., 分镜 k : x_k }, 这里的 x_1 表示分镜 1 对应的远近景标签为 x_1 , 这里的 x_2 表示分镜 2 对应的远近景标签为 x_2 , ..., 这里的 x_k 表示分镜 k 对应的远近景标签为 x_k 。其中, 上述远近景标签可以包括但不限于: 远景、人物近景、人物特写、物体全景、物体特写等。其中, 这里的分镜 1、分镜 2、...、分镜 k 可以为上述图 5 所对应实施例中的分镜片段 1、分镜片段 2、...、分镜片段 k 。

其中, 可以理解的是, 如图 6 所示, 在将 k 个分镜片段输入远近景识别模型之后, 可以将 k 个分镜片

段输入人脸检测模型和人脸识别模型。在将k个分镜片段输入人脸检测模型和人脸识别模型之后，得到的每个分镜片段对应的角色标签（即第二属性标签）可以为：{分镜1: y1, 分镜2: y2, 分镜4: y4, ..., 分镜k-1: yk-1}，这里的y1表示分镜1对应的角色标签为y1，这里的y2表示分镜2对应的角色标签为y2，这里的y4表示分镜4对应的角色标签为y4，..., 这里的yk-1表示分镜k-1对应的角色标签为yk-1。其中，上述角色标签可以包括但不限于：单人、双人等；上述角色标签还可以包括但不限于：男一、男二、女一、女二、小女孩A、小男孩B等。其中，分镜3、分镜5、...、分镜k不包括角色标签。

其中，可以理解的是，如图6所示，在将k个分镜片段输入人脸检测模型和人脸识别模型之后，可以将未获得检测或识别结果（即不包括角色标签）的分镜片段输入场景识别模型。在将未获得检测或识别结果的分镜片段输入场景识别模型之后，得到的每个分镜对应的场景标签（即第三属性标签）可以为：{分镜3: z3, 分镜5: z5, ..., 分镜k: zk}，这里的z3表示分镜3对应的场景标签为z3，这里的z5表示分镜5对应的场景标签为z5，..., 这里的zk表示分镜k对应的场景标签为zk。其中，上述场景标签可以包括但不限于：自然场景、室内场景、人物建筑、竹林、河边、游乐园等。

应当理解，对于k个分镜片段中的某一个分镜片段而言，可以使用该分镜片段的远近景标签和角色标签、或远近景标签和场景标签，来共同描述该分镜片段的片段属性标签。比如，对于k个分镜片段中的分镜1而言，可以使用该分镜1的远近景标签和角色标签，共同来描述该分镜1的片段属性标签（即片段属性标签1），例如分镜1对应的远近景标签为远景（即x1为远景），分镜1对应的角色标签为男一（即y1为男一），则该分镜1对应的片段属性标签1可以为：{远景、男一}。

步骤S102，基于目标用户的用户画像，从视频模板数据库中确定与目标用户相关联的视频模板，并获取视频模板中预先确定的至少一个模板片段以及模板标签序列，所述模板标签序列由所述至少一个模板片段的模板属性标签构成。

具体的，服务器可以获取目标用户的行为日志表，从行为日志表中提取与目标用户相关联的行为数据信息。进一步的，服务器可以对行为数据信息进行用户画像分析，得到用于表征目标用户的用户画像，基于目标用户的用户画像，从视频模板数据库中确定与目标用户相关联的视频模板。其中，视频模板可以携带模板片段的模板属性标签所构成的模板标签序列，模板片段是对模板视频进行视频分析后所得到的，模板视频是由行为数据信息所确定的。所述视频模板数据库可以是在服务器上，或者在独立于服务器的其他设备上。进一步的，服务器可以获取视频模板中包括的至少一个模板片段以及至少一个模板片段的模板属性标签构成的模板标签序列。其中，可以理解的是，本申请实施例可以将服务器在目标时长内所获取到的应用客户端中的不同用户的行为日志统称为行为日志表。

其中，可以理解的是，行为数据信息用于记录目标用户每次访问应用客户端时，产生的行为交互数据（访问，浏览，搜索，点击等）。这里的行为交互数据具体可以包括目标用户访问视频的类型、浏览视频的时间、浏览视频的次数、搜索视频的记录、点击视频的次数，以及目标用户收藏的视频、推荐的视频、点赞的视频、购买的视频、投币的视频等。

为便于理解，请参见图7，图7是本申请实施例提供的一种获取视频模板的场景示意图。如图7所示的日志管理系统70中具体可以包括多个数据库，多个数据库具体可以包含图7所示的数据库70a、数据库70b、...、数据库70n。这意味着该日志管理系统70可以用于存储应用客户端中的不同用户的行为日志。比如，数据库70a可以用于存储用户Y1（未在图上示出）的行为日志，数据库70b可以用于存储用户Y2（未在图上示出）的行为日志，..., 数据库70n可以用于存储用户Yn（未在图上示出）的行为日志。

其中，如图7所示，在目标用户为上述用户Y1（即目标用户）时，服务器可以在数据库10a中获取目标用户在目标时长内的行为日志表，可以进一步在行为日志表中获取行为数据信息。应当理解，服务器在获取到目标用户的行为数据信息之后，可以对目标时长内的行为数据信息进行用户画像分析，以得到用

于表征目标用户的用户画像。

其中，这里的用户画像可以包括目标用户对于某一个视频类型的喜爱程度，服务器进而可以选择这一视频类型的视频模板作为与目标用户相关联的视频模板。同理，这里的用户画像可以包括目标用户对于某一个视频的喜爱程度，服务器进而可以选择这一个视频对应的视频模板作为与目标用户相关联的视频模板。可以理解的是，这里的视频模板对应的模板数据可以为与目标视频的视频数据具有相同视频类型的数据。比如，在目标视频为动漫时，可以在动漫类的视频模板中选择与目标视频相关联的视频模板。又比如，在目标视频为真人剧时，可以在真人剧类的视频模板中选择与目标视频相关联的视频模板。这样，可以为目标视频选择最佳的视频模板，提高视频素材片段的显示效果。

其中，可以理解的是，图 7 所示的日志管理系统 70 可以在单个行为记录周期内（比如，以天为该行为记录周期的单位），为访问该应用客户端的目标用户建立一个行为日志表。例如，该日志管理系统 70 可以在检测到该目标用户当天首次访问该应用客户端时，为该目标用户建立一个行为日志表。此时，这个行为日志表中记录了当前访问该应用客户端的访问时间戳（例如，T1 时刻）。这意味着这个行为日志表中并不存在当前 T1 时刻之前的任何其他行为交互数据。进一步的，该日志管理系统 70 可以在当前这个行为记录周期达到记录周期阈值时，将为这个目标用户建立的这个行为日志表（例如，行为日志表 1）添加到对应的数据库（例如，图 7 所示的数据库 10a）进行存储。同理，日志管理系统 70 可以在目标用户的访问时间戳为其他时刻（例如，T2 时刻）时，将该 T2 时刻所对应行为日志表（例如，行为日志表 2）添加到对应的数据库（例如，图 7 所示的数据库 10a）进行存储。

应当理解，当目标用户在记录周期内访问客户端，且与该应用客户端之间产生交互行为时，则日志管理系统 70 可以在这个记录周期的行为日志表中，记录该目标用户与应用客户端之间的交互行为。可以理解的是，这里的目标时长具体可以包括：一个或者多个记录周期。所以，服务器在上述目标时长（即截止到本次访问应用客户端前的多个记录周期）内所获取到的目标用户的行为日志表，具体可以包括上述行为日志表 1、上述行为日志表 1 和行为日志表 2。

步骤 S103，基于至少一个模板片段的模板属性标签和多个视频片段对应的片段属性标签，在多个视频片段中筛选与至少一个模板片段的模板属性标签匹配的至少一个视频片段。

步骤 S104，按照至少一个模板片段中各个模板片段的模板属性标签在模板标签序列中的位置，将匹配的至少一个视频片段进行拼接，作为目标视频的视频素材片段。

根据本申请实施例，服务器可以基于至少一个模板片段以及所述模板标签序列，在多个视频片段中筛选满足片段匹配条件的视频片段，将满足片段匹配条件的视频片段作为目标视频的视频素材片段。

具体的，服务器可以将 N 个模板片段作为目标模板片段，在模板标签序列中将目标模板片段的队列位置（目标模板片段在 N 个模板片段形成的队列中的位置或顺序）确定为目标队列位置，将目标队列位置对应的模板属性标签确定为目标模板属性标签。其中，模板片段的数量可以为 N 个，这里的 N 可以为大于 1 的正整数。因此，模板标签序列可以包含 N 个序列位置，一个序列位置对应一个模板属性标签，且一个模板属性标签对应一个模板片段。进一步的，服务器可以在多个视频片段对应的片段属性标签中，筛选与目标模板属性标签相匹配的片段属性标签，将筛选出的片段属性标签所对应的一个或多个视频片段确定为候选视频片段。进一步的，服务器可以将候选视频片段中的每个候选视频片段与目标模板片段进行相似分析，得到每个候选视频片段与目标模板片段的相似阈值。在相似阈值中确定最大相似阈值，将最大相似阈值所对应的候选视频片段确定为目标模板片段相匹配的目标候选视频片段。进一步的，服务器可以基于目标模板片段在模板标签序列中的目标队列位置，确定目标候选视频片段对应的片段属性标签所构成的目标标签序列，将与所述目标标签序列相关联的所有目标候选视频片段进行拼接处理，作为所述视频素材片段，即，根据与目标标签序列相关联的所有目标候选视频片段，确定满足片段匹配条件的视频素材片段。其中，视

频素材片段的片段属性标签所构成的目标标签序列与模板标签序列相同。

可以理解的是，相似分析可以表示候选视频片段与目标模板片段之间的场景相似度。将候选视频片段输入第三网络模型，可以得到候选视频片段对应的候选特征向量。将目标模板片段输入第三网络模型，可以得到目标模板片段对应的目标特征向量。通过计算候选特征向量与目标特征向量之间的向量距离，可以得到候选视频片段与目标模板片段之间的相似度（即上述相似阈值）。考虑到该第三网络模型为场景识别模型，则这里的相似度可以表示场景相似度。其中，相似分析还可以表示候选视频片段与目标模板片段之间远近景相似度，相似分析还可以表示候选视频片段与目标模板片段之间人物相似度。

比如，可以将目标模板片段输入第三网络模型，得到目标模板片段的目标特征向量。假设存在2个候选视频片段，这2个候选视频片段具体可以包括：候选视频片段1和候选视频片段2，其中，将这2个候选视频片段输入第三网络模型，得到候选视频片段1的候选特征向量1以及候选视频片段2的候选特征向量2。在计算上述目标特征向量分别与2个候选特征向量之间的向量距离之后，若目标特征向量与候选特征向量2之间的距离最小，则表示目标模板片段与候选视频片段2之间的相似阈值为最大相似阈值，可以将候选特征向量2所对应的候选视频片段2作为与目标模板片段相匹配的目标候选视频片段。其中，相似分析还可以表示候选视频片段与目标模板片段之间的时长关系，本申请对相似分析的计算方法不做具体限制。

为便于理解，请参见图8A和图8B，图8A是本申请实施例提供的一种对模板视频进行视频分析的场景示意图，图8B是本申请实施例提供的一种对目标视频进行视频分析的场景示意图。对模板视频进行视频分析后可以得到图8A所示的N个模板片段，这里的N可以为大于1的正整数。例如，N等于4，则4个模板片段可以包括：模板片段80a、模板片段80b、模板片段80c和模板片段80d。其中，模板片段80a对应的模板属性标签为{远景}，模板片段80b对应的模板属性标签为{人物特写}，模板片段80c对应的模板属性标签为{人物近景}以及模板片段80d对应的模板属性标签为{物体近景}。对目标视频进行视频分析后可以得到图8B所示的M个视频片段，这里的M可以为大于1的正整数。例如，M等于8，则8个视频片段可以包括：视频片段800a、视频片段800b、视频片段800c、视频片段800d、视频片段800e、视频片段800f、视频片段800g和视频片段800h。其中，视频片段800a对应的片段属性标签为{远景}，视频片段800b对应的片段属性标签为{人物近景}，视频片段800c对应的片段属性标签为{远景}，视频片段800d对应的片段属性标签为{人物近景}，视频片段800e对应的片段属性标签为{人物近景}，视频片段800f对应的片段属性标签为{远景}，视频片段800g对应的片段属性标签为{物体近景}，视频片段800h对应的片段属性标签为{人物特写}。

其中，可以理解的是，若从图8A中的4个模板片段中获取模板片段80a作为目标模板片段（例如，目标模板片段1），则该目标模板片段1的队列位置可以为位置1（即目标队列位置为位置1），该目标模板片段的模板属性标签可以为{远景}（即目标模板属性标签为{远景}）。在图8B的8个视频片段中筛选出与该目标模板属性标签相匹配的片段属性标签为{远景}，{远景}对应的视频片段为视频片段800a、视频片段800c和视频片段800f，则目标模板片段1对应的候选视频片段为视频片段800a、视频片段800c和视频片段800f。进一步的，在计算这3个候选视频片段与目标模板片段1之间的相似阈值之后，若视频片段800a与该目标模板片段1之间的相似阈值为最大相似阈值，则将视频片段800a确定为与该目标模板片段1相匹配的目标候选视频片段（例如，目标候选视频片段1）。

同理，可以理解的是，若从图8A中的4个模板片段中获取模板片段80b作为目标模板片段（例如，目标模板片段2），则该目标模板片段2的队列位置可以为位置2（即目标队列位置为位置2），该目标模板片段的模板属性标签可以为{人物特写}（即目标模板属性标签为{人物特写}）。在图8B的8个视频片段中筛选出与该目标模板属性标签相匹配的片段属性标签为{人物特写}，{人物特写}对应的视频片段为视

频片段 800h, 则将视频片段 800h 确定为与该目标模板片段 2 相匹配的目标候选视频片段 (例如, 目标候选视频片段 2)。

同理, 可以理解的是, 若从图 8A 中的 4 个模板片段中获取模板片段 80c 作为目标模板片段 (例如, 目标模板片段 3), 则该目标模板片段 3 的队列位置可以为位置 3 (即目标队列位置为位置 3), 该目标模板片段的模板属性标签可以为{人物近景} (即目标模板属性标签为{人物近景})。在图 8B 的 8 个视频片段中筛选出与该目标模板属性标签相匹配的片段属性标签为{人物近景}, {人物近景}对应的视频片段为视频片段 800d 和视频片段 800e, 则目标模板片段 3 对应的候选视频片段为视频片段 800d 和视频片段 800e。进一步的, 在计算这 2 个候选视频片段与目标模板片段 3 之间的相似阈值之后, 若视频片段 800e 与该目标模板片段 3 之间的相似阈值为最大相似阈值, 则将视频片段 800e 确定为与该目标模板片段 3 相匹配的目标候选视频片段 (例如, 目标候选视频片段 3)。

同理, 可以理解的是, 若从图 8A 的 4 个模板片段中获取模板片段 80d 作为目标模板片段 (例如, 目标模板片段 4), 则该目标模板片段 4 的队列位置可以为位置 4 (即目标队列位置为位置 4), 该目标模板片段的模板属性标签可以为{物体近景} (即目标模板属性标签为{物体近景})。在图 8B 的 8 个视频片段中筛选出与该目标模板属性标签相匹配的片段属性标签为{物体近景}, {物体近景}对应的视频片段为视频片段 800g, 则将视频片段 800g 确定为与该目标模板片段 4 相匹配的目标候选视频片段 (例如, 目标候选视频片段 4)。

因此, 位置 1 对应的目标候选视频片段 1 为视频片段 800a, 位置 2 对应的目标候选视频片段 2 为视频片段 800h, 位置 3 对应的目标候选视频片段 3 为视频片段 800e, 位置 4 对应的目标候选视频片段 4 为视频片段 800g, 则可以基于位置 1、位置 2、位置 3 和位置 4, 由视频片段 800a、视频片段 800h、视频片段 800e 和视频片段 800g 确定视频素材片段。其中, 模板标签序列为模板片段对应的模板属性标签所构成的序列, 这里的模板标签序列可以表示为{远景、人物特写、人物近景、物体近景}; 目标标签序列为与模板片段匹配的视频片段对应的片段属性标签所构成的序列, 这里的目标标签序列可以表示为{远景、人物特写、人物近景、物体近景}。

其中, 可以理解的是, 目标模板片段 1 可以与目标候选视频片段 1 具有相似的视频播放效果, 目标模板片段 2 可以与目标候选视频片段 2 具有相似的视频播放效果, 目标模板片段 3 可以与目标候选视频片段 3 具有相似的视频播放效果, 目标模板片段 4 可以与目标候选视频片段 4 具有相似的视频播放效果, 因此, 视频素材片段可以与上述模板片段具有相同的视频播放效果。

应当理解, 服务器可以将与目标标签序列相关联的所有目标候选视频片段进行视频拼接处理, 得到与 N 个模板片段相关联的拼接视频数据。进一步的, 服务器可以获取与 N 个模板片段相关联的模板音频数据, 通过音视频合成组件将模板音频数据和拼接视频数据进行音视频合并处理, 得到满足片段匹配条件的视频素材片段。

其中, 将每个目标候选视频片段进行视频拼接处理, 以及将模板音频数据和拼接视频数据进行音视频合并处理的工具可以为同一个工具, 这个工具可以为上述音视频合成组件。这里的音视频合成组件可以为 ffmpeg 工具, 也可以为其他第三方具有视频解封装能力的软件工具。这里将不再对视频解封装组件进行一一举例。

步骤 S105, 将视频数据以及视频素材片段推送至目标用户对应的应用客户端, 以使应用客户端输出视频数据以及视频素材片段。

其中, 可以理解的是, 应用客户端在接收到视频数据以及视频素材片段后, 可以在应用显示界面中播放该视频数据以及视频素材片段。在一实施方式中, 当应用客户端播放所述视频数据的同时, 该应用客户端还可以用于展示每个视频素材片段的缩略图。这里将不对应用客户端输出视频素材片段的具体实现形式

进行限定。

在本申请实施例中，服务器在获取到目标用户请求的某个视频的视频数据时，可以对该视频数据进行视频分析，以得到该视频数据的一个或者多个视频片段。其中，可以理解的是，本申请实施例所涉及的视频分析主要包括：视频分镜和属性分析。其中，视频分镜主要是指可以将该视频数据划分成一个或者多个分镜片段，这样，服务器可以进一步对每个分镜片段的片段内容进行属性分析，以得到每个分镜片段的片段属性标签，从而将具备片段属性标签的分镜片段统称为前述视频片段，应当理解，一个视频片段可以对应一个片段属性标签。进一步的，服务器可以在获取到目标用户的用户画像时，快速根据该用户画像确定出与该目标用户相关联的视频模板，进而可以在获取到这个视频模板所映射的模板片段（比如，热门短视频）以及模板片段对应的模板标签序列时，智能在视频片段中筛选满足片段匹配条件的视频片段，从而可以将筛选出的满足片段匹配条件的视频片段作为目标视频的视频素材片段。其中，可以理解的是，这里的视频素材片段的片段属性标签所构成的目标标签序列可以与模板标签序列相同，以确保该视频素材片段与上述模板片段具有相同的视频播放效果。然后，服务器可以将上述视频数据以及视频素材片段智能推送至目标用户对应的应用客户端，以使应用客户端可以输出视频数据和视频素材片段。由此可见，本申请实施例通过视频分析（例如，视频分镜和属性分析等），可以快速得到携带片段属性标签的一个或者多个视频片段。这样，对于这些视频片段而言，可以在根据用户画像智能确定出一个或者多个视频模板时，根据这些视频模板的模板标签序列来分别对这些视频片段进行筛选，以快速得到与该视频模板具有相似视频播放效果的视频片段，进而可以快速合成得到视频素材片段（比如，可以快速得到能够推送给目标用户的短视频），并且可以随着视频模板的增加和更新，重复利用这些视频片段的分镜和属性信息，减少对目标视频中视频帧的识别和处理，提高短视频的生成效率，节省针对不同用户不断生成和分发大量短视频的计算成本，节省服务器的计算资源。

进一步的，请参见图 9，图 9 是本申请实施例提供的一种视频数据处理方法的流程示意图。如图 9 所示，该方法可以由应用客户端和服务器共同执行，该应用客户端可以为上述图 2 所对应实施例中的用户终端 X 中运行的应用客户端，该服务器可以为上述图 2 所对应实施例中的服务器。其中，该方法可以包括以下步骤：

步骤 S201，应用客户端可以响应目标用户针对目标视频执行的播放操作，生成用于请求播放目标视频的视频播放请求，将视频播放请求发送给服务器；

其中，视频播放请求中可以携带目标视频的视频标识，这里的视频标识用于指示服务器获取目标用户所请求播放的目标视频的视频数据。其中，播放操作可以包括点击、长按、滑动等接触性操作，也可以包括语音、手势等非接触性操作，本申请在此不做限定。

步骤 S202，服务器获取目标用户请求的目标视频的视频数据，对视频数据进行视频分析得到多个视频片段，其中，所述视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析，所述多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段；

步骤 S203，服务器基于目标用户的用户画像，从视频模板数据库中确定与目标用户相关联的视频模板，并获取视频模板中预先确定的至少一个模板片段以及模板标签序列，所述模板标签序列由所述至少一个模板片段的模板属性标签构成；

步骤 S204，服务器基于至少一个模板片段的模板属性标签和多个视频片段对应的片段属性标签，在多个视频片段中筛选与至少一个模板片段的模板属性标签匹配的至少一个视频片段；

步骤 S205，服务器按照至少一个模板片段中各个模板片段的模板属性标签在模板标签序列中的位置，将匹配的至少一个视频片段进行拼接，作为目标视频的视频素材片段；

步骤 S206，服务器将视频数据以及视频素材片段推送至目标用户对应的应用客户端；

为便于理解,请参见图 10,图 10 是本申请实施例提供的一种生成视频素材片段的流程示意图。如图 10 所示,服务器可以在获取到精彩短视频(即模板视频)时,对该精彩短视频进行视频分析,以得到该精彩短视频的一个或者多个视频片段,进而可以将该精彩短视频的一个或者多个视频片段作为模板片段。其中,可以理解的是,本申请实施例所涉及的视频分析主要包括:视频分镜和属性分析。其中,视频分镜主要是指可以将该精彩短视频的视频数据划分成一个或者多个分镜片段。这样,服务器可以进一步对每个分镜片段的片段内容进行属性分析(即分镜信息抽取),以得到每个分镜片段的模板属性标签(即图 10 所示的场景标签、人物标签(即角色标签)和远近镜标签),从而将具备模板属性标签的分镜片段统称为前述模板片段,从而可以基于模板属性标签确定热门集锦序列(即镜头序列记录)。应当理解,一个模板片段可以对应一个模板属性标签。其中,图 10 所示的集锦序列库中的热门集锦序列 1 可以为模板片段 1 对应的模板属性标签,热门集锦序列 2 可以为模板片段 2 对应的模板属性标签,热门集锦序列 3 可以为模板片段 3 对应的模板属性标签。

应当理解,本申请实施例可以将模板视频(即上述精彩短视频)的模板片段、模板片段的模板标签序列和模板音频数据(即音乐)统称为视频模板。

如图 10 所示,服务器可以在获取到电视剧(即目标视频)时,对该电视剧进行视频分镜和属性分析,以得到该电视剧的一个或者多个视频片段。应当理解,一个视频片段可以对应一个片段属性标签。这样,服务器可以从集锦序列库中获取一个或多个热门集锦序列(即序列采样),进而可以按照选取的热门集锦序列确定模板片段以及模板片段对应的模板标签序列,对目标视频的视频片段进行筛选和排序,以得到筛选后的视频片段(即基于素材匹配的分片段镜序列排列),进而可以根据筛选出的这些视频片段所构成的拼接视频数据以及模板片段的模板音频数据,智能生成与模板片段相似的视频素材片段。

其中,通过抽取各短视频平台中的精彩短视频,并获取这些精彩短视频对应的视频模板,可以实现视频模板的连续多日积累。将电视剧根据视频模板生成相应样式的一个或多个视频素材片段,可以丰富最终生成的视频素材片段的样式。其中,一个电视剧可以根据多个视频模板生成多种样式的视频素材片段,可以供视频推荐场景中千人千面的推荐选择,且对于每个视频模板,通过深度学习和图像分析算法可以对精彩短视频和电视剧进行视频分析与视频匹配,可以达到自动化分析的目标。此外,对于新的电视剧,只需要有限的迁移能力即可完成该电视剧的解析,使得新的电视剧的视频素材片段生成难度将低,生成视频素材片段的方法的可迁移性大。

应当理解,服务器对电视剧进行视频分镜和属性分析的具体过程,可以参见上述步骤 S102 的描述,这里将不再继续进行赘述。应当理解,服务器对精彩短视频进行视频分镜和属性分析的具体过程,可以参见服务器对电视剧进行视频分镜和属性分析的描述,这里将不再继续进行赘述。

步骤 S207,应用客户端在应用显示界面中输出视频数据以及视频素材片段。

具体的,应用客户端可以接收服务器基于视频播放请求返回的目标视频的视频数据,以及与目标视频相关联的视频素材片段,并可以在应用客户端的应用显示界面中确定用于播放视频数据的视频播放界面,进而可以在视频播放界面中播放视频数据。进一步的,应用客户端可以响应针对应用显示界面的触发操作,在应用客户端的应用显示界面中播放相应的视频素材片段。其中,该触发操作可以包括点击、长按、滑动等接触性操作,也可以包括语音、手势等非接触性操作,本申请在此不做限定。在一实施方式中,可以理解的是,应用客户端在获取到视频素材片段之后,还可以在应用显示界面中展示每个视频素材片段的缩略图,或者在应用显示界面中动态播放每个视频素材片段的动画,这里将不对这些视频素材片段的具体展示形式进行限定。

为便于理解,请参见图 11,图 11 是本申请实施例提供的一种前后端交互的流程示意图。可以理解的是,上述应用客户端可以运行在图 11 所示的前端 B。目标用户针对前端 B 的应用客户端中的目标视频(比

如，目标用户感兴趣的视频）执行的播放操作，即为前端 B 输入目标视频。进而服务器（即后端）可以基于视频模板，生成与该目标视频相关联的一个或多个视频素材片段（即后端生成）。进而服务器可以将这个目标视频的视频数据和与这个目标视频相关联的一个或者多个视频素材片段（例如，这个视频的视频花絮等）返回给前端 B，即在前端 B 的应用显示界面中显示服务器返回的视频数据和视频素材片段。应当理解，这里的视频模板可以是由该服务器基于该目标用户的用户画像所确定的。

可以理解的是，如图 11 所示，前端 A 可以为视频剪辑员对应的另一用户终端。在对前端 A 输入的精彩短视频进行视频分析后，该视频剪辑员可以在视频分析得到的视频片段中，选择一个或多个视频片段作为模板片段，进而可以基于这些模板片段确定视频模板（即挖掘精彩视频模板）。其中，前端 A 可以接收精彩短视频的输入，然后将该精彩短视频对应的视频模板（即精彩视频模板）上传给服务器保存（即后端保存）。

应当理解，上述前端 B 与前端 A 还可以为同一个用户终端，即前端 B（或前端 A）可以是精彩短视频的输入方，也可以是目标视频的输入方。

为便于理解，请参见图 12A，图 12A 是本申请实施例提供的一种输出视频素材片段的场景示意图。如图 12A 所示，这里的应用显示界面 120a 可以为上述图 2 所对应实施例中的应用显示界面。应用显示界面 120a 中可以包含用于播放目标视频的视频播放界面 1，还可以包括用于展示或播放视频素材片段的短视频推荐列表（例如，短视频推荐列表 1）。该短视频推荐列表 1 中至少可以包含与该目标视频相关联的视频素材片段。这里的视频素材片段可以为上述第一业务数据库中与该目标视频相关联的视频素材片段。在目标用户针对应用显示界面 120a 执行触发操作（例如，图 12A 所示的滑动操作）后，应用客户端可以在应用显示界面 120b 的集锦推荐部分展示或播放上述短视频推荐列表 1 中的视频素材片段。其中，在一实施方式中，当应用客户端在视频播放界面 1 中播放目标视频时，该应用客户端还可以遍历播放（或同步播放）短视频推荐列表 1 中的视频素材片段。如图 12A 所示，该视频推荐列表 1 中具体可以包括与该目标视频相关联的 N 个视频素材片段。这里的 N 个视频素材片段具体可以为图 12A 所示的 3 个视频素材片段。比如，这 3 个视频素材片段可以具体包括：视频素材片段 A1、视频素材片段 A2 和视频素材片段 A3。

在一实施方式中，在目标用户针对应用显示界面 120a 中的业务推荐控件执行触发操作（例如，点击操作）后，应用客户端可以在应用显示界面 120b 的集锦推荐部分展示或播放上述短视频推荐列表 1 中的视频素材片段，例如，应用显示界面 120b 中的视频素材片段 A1、视频素材片段 A2 和视频素材片段 A3 等。

为便于理解，请参见图 12B，图 12B 是本申请实施例提供的一种更新视频素材片段的场景示意图。如图 12B 所示，在目标用户针对上述图 12A 的视频素材片段 A1 执行触发操作（例如，点击操作）时，服务器可以将这个视频素材片段 A1 的视频数据（例如，视频数据 J）和与这个视频数据 J 相关联的一个或者多个视频素材片段（例如，视频素材片段 C1、视频素材片段 C2 和视频素材片段 C3）返回给应用客户端，以在应用客户端中播放这个视频数据 J。在一实施方式中，应用客户端还可以在播放视频素材片段 A1 的视频数据 J 时，一并显示接收到的这些视频素材片段，得到应用显示界面 120c。

这里的应用显示界面 120c 中可以包含用于播放视频数据 J 的视频播放界面 2，还可以包括用于展示视频素材片段的短视频推荐列表（例如，短视频推荐列表 2）。该短视频推荐列表 2 中至少可以包含与该视频数据 J 关联的视频素材片段。在目标用户针对应用显示界面 120c 中的业务推荐控件执行触发操作（例如，图 12B 所示的点击操作）后，应用客户端可以在应用显示界面 120d 的集锦推荐部分展示或播放上述短视频推荐列表 2 中的视频素材片段。这里的视频素材片段可以为上述第二业务数据库中与该视频素材片段 A1 具有同一视频模板的视频素材片段。如图 12B 所示，该短视频推荐列表 2 中具体可以包括与该视频数据 J 相关联的 M 个视频素材片段。这里的 M 个视频素材片段具体可以为图 12B 所示的 3 个视频素材片段。比

如，这3个视频素材片段可以具体包括：视频素材片段C1、视频素材片段C2和视频素材片段C3。

在一实施方式中，在目标用户针对应用显示界面120c执行触发操作（例如，滑动操作）后，应用客户端可以在应用显示界面120d的集锦推荐部分展示或播放上述短视频推荐列表2中的视频素材片段，例如，应用显示界面120d中的视频素材片段C1、视频素材片段C2和视频素材片段C3等。

应当理解，当目标用户在上述应用客户端中观看完上述视频素材片段A1之后，还可以智能在应用客户端中为上述目标用户遍历播放上述短视频推荐列表2中的这些视频素材片段。比如，当目标用户在上述应用客户端中观看完上述视频素材片段A1时，服务器还可以将短视频推荐列表2中的多个视频素材片段中的视频素材片段C1输出至应用客户端，以在该应用客户端中实现对该视频素材片段C1的智能播放。在一实施方式中，该应用客户端还可以在将应用客户端的视频播放界面1中所播放的视频数据更新为视频素材片段A1时，记录目标视频的当前播放进度（例如，时刻T），以在播放完视频素材片段A1后，从目标视频的時刻T开始继续对目标视频进行播放。

其中，可以理解的是，应用客户端可以根据目标视频的当前播放进度，实时动态调整视频素材片段在短视频推荐列表中的位置，以为目标用户推荐不同排序的视频素材片段。比如，若在当前播放进度之前，包括组成视频素材片段的全部视频片段，即组成视频素材片段的全部视频片段在当前时刻已经观看完成，则可以将该视频素材片段排列在短视频推荐列表的前面，即实现剧情回放。在一实施方式中，应用客户端还可以根据当前视频素材片段在其它用户终端中的应用客户端上的播放次数，来将视频推荐列表中的视频素材片段进行排序。若某个视频素材片段的播放总次数比较高，则表示这个视频素材片段的质量比较高，则可以为目标用户优先推荐该视频素材片段，即将该视频素材片段排列在短视频推荐列表的前面。

由此可见，本申请实施例通过对视频数据进行视频分析（例如，视频分镜和属性分析等），可以快速得到携带片段属性标签的一个或者多个视频片段。这样，对于这些视频片段而言，可以在根据用户画像准确地确定出一个或者多个视频模板时，智能根据这些视频模板的模板标签序列来分别对这些视频片段进行筛选，以快速得到与该视频模板具有相似视频播放效果的视频片段，进而可以快速合成得到视频素材片段（比如，可以快速得到能够展示给目标用户的短视频），并且可以随着视频模板的增加和更新，重复利用这些视频片段的分镜和属性信息，减少对目标视频中视频帧的识别和处理，提高短视频的生成效率，节省针对不同用户不断生成和分发大量短视频的计算成本，节省服务器的计算资源。

进一步的，请参见图13，图13是本申请实施例提供的一种视频数据处理装置的结构示意图。视频数据处理装置1可以包括：片段生成模块30、模板获取模块40、素材确定模块50、数据发送模块60；进一步的，视频数据处理装置1还可以包括：请求接收模块10、数据查找模块20。

片段生成模块30，用于获取目标用户请求的目标视频的视频数据，对视频数据进行视频分析得到多个视频片段，其中，所述视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析，所述多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段。

其中，片段生成模块30包括：模型获取单元301、分镜获取单元302、标签确定单元303、片段确定单元304。

模型获取单元301，用于获取目标用户请求的目标视频的视频数据以及与视频数据相关联的网络识别模型；

分镜获取单元302，用于通过视频切分组件将视频数据对应的视频序列进行分镜处理，得到与视频序列相关联的多个分镜片段；

其中，分镜获取单元302包括：组件获取子单元3021、图像匹配子单元3022、分镜创建子单元3023、匹配完成子单元3024、分镜确定子单元3025；

组件获取子单元3021，用于通过视频切分组件在视频序列中确定用于作为聚类质心的第一视频帧，创

建第一视频帧所属的分镜簇的分镜簇信息；

图像匹配子单元 3022，用于在视频序列中将除第一视频帧之外的视频帧确定为第二视频帧，基于轮询机制依次获取第二视频帧中的每个第二视频帧，确定每个第二视频帧与第一视频帧的图像相似度；

分镜创建子单元 3023，用于若第一视频帧与一第二视频帧的图像相似度大于或者等于聚类阈值，则将图像相似度大于或者等于聚类阈值的第二视频帧划分到第一视频帧所属的分镜簇；

匹配完成子单元 3024，用于若第一视频帧与一第二视频帧的图像相似度小于聚类阈值，则用图像相似度小于聚类阈值的第二视频帧更新第一视频帧，创建更新后的第一视频帧所属的分镜簇的分镜簇信息，将更新后的第一视频帧依次与未匹配的第二视频帧进行图像相似度匹配，直到视频序列中的视频帧均完成图像相似度匹配时，得到视频序列中的视频帧所属的分镜簇的分镜簇信息；

分镜确定子单元 3025，用于基于视频序列中的视频帧所属的分镜簇的分镜簇信息，将视频序列中的视频帧组成多个分镜片段。

其中，组件获取子单元 3021、图像匹配子单元 3022、分镜创建子单元 3023、匹配完成子单元 3024 以及分镜确定子单元 3025 的具体实现方式，可以参见上述图 3 所对应实施例中对步骤 S101 的描述，这里将不再进行赘述。

标签确定单元 303，用于将多个分镜片段输入至网络识别模型，由网络识别模型基于所述多个预设的片段属性标签，对多个分镜片段进行属性分析，得到多个分镜片段对应的片段属性标签。

其中，网络识别模型至少包括：具有第一属性标签提取功能的第一网络模型、具有第二属性标签提取功能的第二网络模型和具有第三属性标签提取功能的第三网络模型。

标签确定单元 303 包括：第一分析子单元 3031、人脸检测子单元 3032、第二分析子单元 3033、第三分析子单元 3034、标签分析子单元 3035；

第一分析子单元 3031，用于将多个分镜片段输入第一网络模型，通过第一网络模型对多个分镜片段中的每个分镜片段进行远近景分析，得到多个分镜片段的远近景标签，将多个分镜片段的远近景标签作为第一网络模型输出的第一属性标签，将具有第一属性标签的分镜片段作为第一类分镜片段；

人脸检测子单元 3032，用于将第一类分镜片段输入第二网络模型，由第二网络模型对第一类分镜片段中的每个分镜片段进行人脸检测，得到人脸检测结果；

第二分析子单元 3033，用于若人脸检测结果指示第一类分镜片段中存在目标角色的人脸，则在第一类分镜片段中将存在目标角色的人脸所对应的分镜片段作为第二类分镜片段，通过第二网络模型确定第二类分镜片段中的目标角色所属的角色标签，将目标角色所属的角色标签确定为第二类分镜片段的第二属性标签；目标角色为目标视频中的一个或者多个角色；

第三分析子单元 3034，用于在第一类分镜片段中将除第二类分镜片段之外的分镜片段，确定为第三类分镜片段，将第三类分镜片段输入第三网络模型，由第三网络模型对第一类分镜片段中的每个分镜片段进行场景检测，得到第三类分镜片段的第三属性标签；

标签分析子单元 3035，用于根据第一类分镜片段的第一属性标签、第二类分镜片段的第二属性标签、以及第三类分镜片段的第三属性标签，确定多个分镜片段中的每个分镜片段对应的片段属性标签。

其中，第一分析子单元 3031、人脸检测子单元 3032、第二分析子单元 3033、第三分析子单元 3034 以及标签分析子单元 3035 的具体实现方式，可以参见上述图 3 所对应实施例中对步骤 S101 的描述，这里将不再进行赘述。

片段确定单元 304，用于将具备片段属性标签的分镜片段确定为视频数据的视频片段。

其中，模型获取单元 301、分镜获取单元 302、标签确定单元 303 以及片段确定单元 304 的具体实现方式，可以参见上述图 3 所对应实施例中对步骤 S101 的描述，这里将不再进行赘述。

模板获取模块 40, 用于基于目标用户的用户画像, 从视频模板数据库中确定与目标用户相关联的视频模板, 并获取视频模板中预先确定的至少一个模板片段以及模板标签序列, 所述模板标签序列由所述至少一个模板片段的模板属性标签构成。

其中, 模板获取模块 40 包括: 行为提取单元 401、行为分析单元 402、模板分析单元 403;

行为提取单元 401, 用于获取目标用户的行为日志表, 从行为日志表中提取与目标用户相关联的行为数据信息;

行为分析单元 402, 用于对行为数据信息进行用户画像分析, 得到用于表征目标用户的用户画像, 基于目标用户的用户画像, 从视频模板数据库中确定与目标用户相关联的视频模板;

模板分析单元 403, 用于获取视频模板中预先确定的所述至少一个模板片段以及所述模板标签序列。

其中, 行为提取单元 401, 行为分析单元 402 以及模板分析单元 403 的具体实现方式, 可以参见上述图 3 所对应实施例中对步骤 S102 的描述, 这里将不再进行赘述。

素材确定模块 50, 用于基于至少一个模板片段的模板属性标签和多个视频片段对应的片段属性标签, 在多个视频片段中筛选与至少一个模板片段的模板属性标签匹配的一个视频片段, 按照至少一个模板片段中各个模板片段的位置, 将匹配的至少一个视频片段进行拼接, 作为目标视频的视频素材片段。

其中, 模板片段的数量为 N 个, N 为大于 1 的正整数; 模板标签序列包含 N 个序列位置, 一个序列位置对应一个模板属性标签, 且一个模板属性标签对应一个模板片段。

素材确定模块 50 包括: 标签确定单元 501、标签筛选单元 502、片段匹配单元 503、素材生成单元 504;

标签确定单元 501, 用于将 N 个模板片段作为目标模板片段, 在模板标签序列中将目标模板片段的队列位置确定为目标队列位置, 将目标队列位置对应的模板属性标签确定为目标模板属性标签;

标签筛选单元 502, 用于在多个视频片段对应的片段属性标签中, 筛选与目标模板属性标签相匹配的片段属性标签, 将筛选出的片段属性标签所对应的一个或多个视频片段确定为候选视频片段;

片段匹配单元 503, 用于将候选视频片段中的每个候选视频片段与目标模板片段进行相似分析, 得到每个候选视频片段与目标模板的相似阈值, 在相似阈值中确定最大相似阈值, 将最大相似阈值所对应的候选视频片段确定为目标模板片段相匹配的目标候选视频片段;

素材生成单元 504, 用于基于目标模板片段在模板标签序列中的目标队列位置, 确定目标候选视频片段对应的片段属性标签所构成的目标标签序列, 将与目标标签序列相关联的所有目标候选视频片段进行拼接处理, 得到视频素材片段。

其中, 素材生成单元 504 包括: 视频拼接子单元 5041、素材合成子单元 5042;

视频拼接子单元 5041, 用于将与目标标签序列相关联的所有目标候选视频片段进行视频拼接处理, 得到与 N 个模板片段相关联的拼接视频数据;

素材合成子单元 5042, 用于获取与 N 个模板片段相关联的模板音频数据, 通过音视频合成组件将模板音频数据和拼接视频数据进行音视频合并处理, 得到视频素材片段。

其中, 视频拼接子单元 5041 以及素材合成子单元 5042 的具体实现方式, 可以参见上述图 3 所对应实施例中对步骤 S103、S104 的描述, 这里将不再进行赘述。

其中, 标签确定单元 501、标签筛选单元 502、片段匹配单元 503 以及素材生成单元 504 的具体实现方式, 可以参见上述图 3 所对应实施例中对步骤 S103、S104 的描述, 这里将不再进行赘述。

数据发送模块 60, 用于将视频数据以及视频素材片段推送至目标用户对应的应用客户端, 以使应用客户端输出视频数据以及视频素材片段。

在一实施方式中, 请求接收模块 10, 用于接收应用客户端发送的视频播放请求; 视频播放请求是由应用客户端响应目标用户针对目标视频执行的播放操作所生成的;

数据查找模块 20, 用于从视频播放请求中提取目标视频的视频标识, 基于视频标识在视频业务数据库中查找目标视频对应的业务视频数据, 将查找到的业务视频数据作为应用客户端中的目标视频的视频数据。

其中, 片段生成模块 30、模板获取模块 40、素材确定模块 50 以及数据发送模块 60 的具体实现方式, 可以参见上述图 3 所对应实施例中对步骤 S101-步骤 S105 的描述, 这里将不再进行赘述。在一实施方式中, 请求接收模块 10 以及数据查找模块 20 的具体实现方式, 可以参见上述图 9 所对应实施例中对步骤 S201 和步骤 S207 的描述, 这里将不再进行赘述。另外, 对采用相同方法的有益效果描述, 也不再进行赘述。

进一步的, 请参见图 14, 图 14 是本申请实施例提供的一种视频数据处理装置的结构示意图。视频数据处理装置 2 可以包括: 数据获取模块 70、数据输出模块 80;

数据获取模块 70, 用于响应目标用户针对应用客户端中的目标视频执行的播放操作, 从服务器上获取目标视频的视频数据, 以及与目标视频相关联的视频素材片段; 视频素材片段是由服务器对视频数据进行视频分析得到多个视频片段, 其中, 所述视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析, 所述多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段; 基于目标用户的用户画像, 从视频模板数据库中确定与目标用户相关联的视频模板, 并获取视频模板中预先确定的至少一个模板片段以及模板标签序列, 所述模板标签序列由所述至少一个模板片段的模板属性标签构成; 基于至少一个模板片段的模板属性标签和多个视频片段对应的片段属性标签, 在多个视频片段中筛选与至少一个模板片段的模板属性标签匹配的至少一个视频片段; 按照至少一个模板片段中各个模板片段的模板属性标签在模板标签序列中的位置, 将匹配的至少一个视频片段进行拼接得到的。

其中, 数据获取模块 70 包括: 请求发送单元 701、数据接收单元 702;

请求发送单元 701, 用于响应目标用户针对应用客户端中的目标视频执行的播放操作, 生成用于请求播放目标视频的视频播放请求, 将视频播放请求发送给服务器; 视频播放请求中携带目标视频的视频标识; 视频标识用于指示服务器获取目标用户所请求播放的目标视频的视频数据;

数据接收单元 702, 用于接收服务器基于视频播放请求返回的视频数据, 以及与目标视频相关联的视频素材片段; 视频素材片段是由服务器在根据目标用户的用户画像确定出视频模板时, 根据视频模板对视频数据进行视频分析以及视频匹配后所得到的, 用户画像是由目标用户在应用客户端中的用户行为信息所确定的。

其中, 请求发送单元 701 以及数据接收单元 702 的具体实现方式, 可以参见上述图 9 所对应实施例中对步骤 S201 的描述, 这里将不再进行赘述。

数据输出模块 80, 用于在应用客户端的应用显示界面中输出视频数据以及视频素材片段。

其中, 数据输出模块包括: 视频播放单元 801、素材输出单元 802;

视频播放单元 801, 用于在应用客户端的应用显示界面中确定用于播放视频数据的视频播放界面, 在视频播放界面中播放视频数据;

素材输出单元 802, 用于响应针对应用显示界面的触发操作, 在应用显示界面中播放视频素材片段。

其中, 视频播放单元 801 以及素材输出单元 802 的具体实现方式, 可以参见上述图 9 所对应实施例中对步骤 S207 的描述, 这里将不再进行赘述。

其中, 数据获取模块 70 以及数据输出模块 80 的具体实现方式, 可以参见上述图 9 所对应实施例中对步骤 S201 和步骤 S207 的描述, 这里将不再进行赘述。另外, 对采用相同方法的有益效果描述, 也不再进行赘述。

请参见图 15, 图 15 是本申请实施例提供的一种计算机设备的结构示意图。如图 15 所示, 该计算机设备 2000 可以包括: 处理器 2001、网络接口 2004 和存储器 2005, 此外, 上述计算机设备 2000 还可以包括: 用户接口 2003 和至少一个通信总线 2002。其中, 通信总线 2002 用于实现这些组件之间的连接通信。用户

接口 2003 还可以包括标准的有线接口、无线接口。在一实施方式中，网络接口 2004 可以包括标准的有线接口、无线接口（如 WI-FI 接口）。存储器 2005 可以是高速 RAM 存储器，也可以是非不稳定的存储器（non-volatile memory），例如至少一个磁盘存储器。在一实施方式中，存储器 2005 还可以是至少一个位于远离前述处理器 2001 的存储装置。如图 15 所示，作为一种计算机可读存储介质的存储器 2005 中可以包括操作系统、网络通信模块、用户接口模块以及设备控制应用程序。

在如图 15 所示的计算机设备 2000 中，网络接口 2004 可提供网络通讯功能；而用户接口 2003 主要用于为用户提供输入的接口；而处理器 2001 可以用于调用存储器 2005 中存储的设备控制应用程序。

应当理解，本申请实施例中所描述的计算机设备 2000 可以为服务器或用户终端，这里将不对其进行限定。可以理解的是，该计算机设备 2000 可以用于执行前文图 3 或图 9 所对应实施例中对视频数据处理方法的描述，在此不再赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。

此外，这里需要指出的是：本申请实施例还提供了一种计算机可读存储介质，且计算机可读存储介质中存储有前文提及的视频数据处理装置 1 或视频数据处理装置 2 所执行的计算机程序，且计算机程序包括程序指令，当处理器执行程序指令时，能够执行前文图 3 或图 9 所对应实施例中对视频数据处理方法的描述，因此，这里将不再进行赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。对于本申请所涉及的计算机可读存储介质实施例中未披露的技术细节，请参照本申请方法实施例的描述。

进一步的，请参见图 16，图 16 是本申请实施例还提供一种视频数据处理系统。该视频数据处理系统 3 中可以包含服务器 3a 和用户终端 3b，所述服务器 3a 可以为前述图 13 所对应实施例中的视频数据处理装置 1；所述用户终端 3b 可以为前述图 14 所对应实施例中的视频数据处理装置 2。可以理解的是，对采用相同方法的有益效果描述，也不再进行赘述。

此外，需要说明的是：本申请实施例还提供了一种计算机程序产品或计算机程序，该计算机程序产品或者计算机程序可以包括计算机指令，该计算机指令可以存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令，处理器可以执行该计算机指令，使得该计算机设备执行前文图 3 或图 9 所对应实施例中对视频数据处理方法的描述，因此，这里将不再进行赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。对于本申请所涉及的计算机程序产品或者计算机程序实施例中未披露的技术细节，请参照本申请方法实施例的描述。

可以理解的是，在本申请的具体实施方式中，涉及到用户相关联的行为数据、用户画像等相关的数据，当本申请以上实施例运用到具体产品或技术中时，需要获得用户许可或者同意，且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，计算机程序可存储于一计算机可读存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，存储介质可为磁碟、光盘、只读存储存储器（Read-Only Memory，ROM）或随机存储存储器（Random Access Memory，RAM）等。

以上所揭露的仅为本申请较佳实施例而已，当然不能以此来限定本申请之权利范围，因此依本申请权利要求所作的等同变化，仍属本申请所涵盖的范围。

权利要求书

1、一种视频数据处理方法，由计算机设备执行，所述方法包括：

获取目标用户请求的目标视频的视频数据，对所述视频数据进行视频分析得到多个视频片段，其中，所述视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析，所述多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段；

基于所述目标用户的用户画像，从视频模板数据库中确定与所述目标用户相关联的视频模板，并获取所述视频模板中预先确定的至少一个模板片段以及模板标签序列，所述模板标签序列由所述至少一个模板片段的模板属性标签构成；

基于所述至少一个模板片段的模板属性标签和所述多个视频片段对应的片段属性标签，在所述多个视频片段中筛选与所述至少一个模板片段的模板属性标签匹配的至少一个视频片段；

按照所述至少一个模板片段中各个模板片段的模板属性标签在所述模板标签序列中的位置，将匹配的至少一个视频片段进行拼接，作为所述目标视频的视频素材片段；

将所述视频数据以及所述视频素材片段推送至所述目标用户对应的应用客户端，以使所述应用客户端输出所述视频数据以及所述视频素材片段。

2、根据权利要求 1 所述的方法，其中，在所述获取目标用户请求的目标视频的视频数据之前，所述方法还包括：

响应于接收到应用客户端发送的针对目标视频的视频播放请求，从所述视频播放请求中提取所述目标视频的视频标识；

基于所述视频标识，在视频业务数据库中查找所述目标视频对应的业务视频数据，将查找到的业务视频数据作为所述应用客户端中的目标视频的视频数据。

3、根据权利要求 1 所述的方法，其中，所述对所述视频数据进行视频分析得到多个视频片段，包括：

通过视频切分组件将所述视频数据对应的视频序列进行分镜处理，得到与所述视频序列相关联的多个分镜片段；

将所述多个分镜片段输入至网络识别模型，由所述网络识别模型基于所述多个预设的片段属性标签，对所述多个分镜片段进行属性分析，得到所述多个分镜片段对应的片段属性标签；

将具备片段属性标签的所述多个分镜片段确定为所述视频数据的所述多个视频片段。

4、根据权利要求 3 所述的方法，其中，所述通过视频切分组件将所述视频数据对应的视频序列进行分镜处理，得到与所述视频序列相关联的多个分镜片段，包括：

通过所述视频切分组件在所述视频序列中确定用于作为聚类质心的第一视频帧，创建所述第一视频帧所属的分镜簇的分镜簇信息；

在所述视频序列中将除所述第一视频帧之外的视频帧确定为第二视频帧，基于轮询机制依次获取所述第二视频帧中的每个第二视频帧，确定每个第二视频帧与所述第一视频帧的图像相似度；

若所述第一视频帧与一第二视频帧的图像相似度大于或者等于聚类阈值，则将所述图像相似度大于或者等于所述聚类阈值的第二视频帧划分到所述第一视频帧所属的分镜簇；

若所述第一视频帧与一第二视频帧的图像相似度小于所述聚类阈值，则用所述图像相似度小于所述聚类阈值的第二视频帧更新所述第一视频帧，创建更新后的第一视频帧所属的分镜簇的分镜簇信息，将所述更新后的第一视频帧依次与未匹配的第二视频帧进行图像相似度匹配，直到所述视频序列中的视频帧均完成所述图像相似度匹配时，得到所述视频序列中的视频帧所属的分镜簇的分镜簇信息；

基于所述视频序列中的视频帧所属的分镜簇的分镜簇信息，将所述视频序列中的视频帧组成所述多个分镜片段。

5、根据权利要求3所述的方法，其中，所述网络识别模型至少包括：具有第一属性标签提取功能的第一网络模型、具有第二属性标签提取功能的第二网络模型和具有第三属性标签提取功能的第三网络模型；

所述将所述多个分镜片段输入至网络识别模型，由所述网络识别模型基于所述多个预设的片段属性标签，对所述多个分镜片段进行属性分析，得到所述多个分镜片段对应的片段属性标签，包括：

将所述多个分镜片段输入所述第一网络模型，通过所述第一网络模型对所述多个分镜片段中的每个分镜片段进行远近景分析，得到所述多个分镜片段的远近景标签，将所述多个分镜片段的远近景标签作为所述第一网络模型输出的第一属性标签，将具有所述第一属性标签的分镜片段作为第一类分镜片段；

将所述第一类分镜片段输入所述第二网络模型，由所述第二网络模型对所述第一类分镜片段中的每个分镜片段进行人脸检测，得到人脸检测结果；

若所述人脸检测结果指示所述第一类分镜片段中存在目标角色的人脸，则在所述第一类分镜片段中将存在所述目标角色的人脸所对应的分镜片段作为第二类分镜片段，通过所述第二网络模型确定所述第二类分镜片段中的目标角色所属的角色标签，将所述目标角色所属的角色标签确定为所述第二类分镜片段的第二属性标签；所述目标角色为所述目标视频中的一个或者多个角色；

在所述第一类分镜片段中将除所述第二类分镜片段之外的分镜片段，确定为第三类分镜片段，将所述第三类分镜片段输入所述第三网络模型，由所述第三网络模型对所述第一类分镜片段中的每个分镜片段进行场景检测，得到所述第三类分镜片段的第三属性标签；

根据所述第一类分镜片段的第一属性标签、所述第二类分镜片段的第二属性标签、以及所述第三类分镜片段的第三属性标签，确定所述多个分镜片段中的每个分镜片段对应的片段属性标签。

6、根据权利要求1所述的方法，其中，所述基于目标用户的用户画像，从视频模板数据库中确定与所述目标用户相关联的视频模板，并获取所述视频模板中预先确定的至少一个模板片段以及模板标签序列，包括：

获取所述目标用户的行为日志表，从所述行为日志表中提取与所述目标用户相关联的行为数据信息；

对所述行为数据信息进行用户画像分析，得到用于表征所述目标用户的用户画像；

基于所述目标用户的用户画像，从所述视频模板数据库中确定与所述目标用户相关联的视频模板，并获取所述视频模板中预先确定的所述至少一个模板片段以及所述模板标签序列。

7、根据权利要求1所述的方法，其中，所述至少一个模板片段的数量为N个，所述N为大于1的正整数；所述模板标签序列包含N个序列位置，一个序列位置对应一个模板属性标签，且一个模板属性标签对应一个模板片段；

所述基于所述至少一个模板片段的模板属性标签和所述多个视频片段对应的片段属性标签，在所述多个视频片段中筛选与所述至少一个模板片段的模板属性标签匹配的至少一个视频片段，包括：

将N个所述模板片段作为目标模板片段，在所述模板标签序列中将所述目标模板片段的队列位置确定为目标队列位置，将所述目标队列位置对应的模板属性标签确定为目标模板属性标签；

在所述多个视频片段对应的片段属性标签中，筛选与所述目标模板属性标签相匹配的片段属性标签，将筛选出的片段属性标签所对应的一个或多个视频片段确定为候选视频片段；

将所述候选视频片段中的每个候选视频片段与所述目标模板片段进行相似分析，得到所述每个候选视频片段与所述目标模板片段的相似阈值，在所述相似阈值中确定最大相似阈值，将所述最大相似阈值所对应的候选视频片段确定为所述目标模板片段相匹配的目标候选视频片段；

所述按照所述至少一个模板片段中各个模板片段的模板属性标签在所述模板标签序列中的位置，将匹配的至少一个视频片段进行拼接，作为所述目标视频的视频素材片段，包括：

基于所述目标模板片段在所述模板标签序列中的目标队列位置，确定所述目标候选视频片段对应的片

段属性标签所构成的目标标签序列，将与所述目标标签序列相关联的所有目标候选视频片段进行拼接处理，得到所述视频素材片段。

8、根据权利要求 7 所述的方法，其中，所述将与所述目标标签序列相关联的所有目标候选视频片段进行拼接处理，得到所述视频素材片段，包括：

将与所述目标标签序列相关联的所有目标候选视频片段进行视频拼接处理，得到与所述 N 个模板片段相关联的拼接视频数据；

获取与所述 N 个模板片段相关联的模板音频数据，通过音视频合成组件将所述模板音频数据和所述拼接视频数据进行音视频合并处理，得到所述视频素材片段。

9、一种视频数据处理方法，包括：

响应目标用户针对应用客户端中的目标视频执行的播放操作，从服务器上获取所述目标视频的视频数据，以及与所述目标视频相关联的视频素材片段；所述视频素材片段是由所述服务器对所述视频数据进行视频分析得到多个视频片段，其中，所述视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析，所述多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段；基于目标用户的用户画像，从视频模板数据库中确定与目标用户相关联的视频模板，并获取视频模板中预先确定的至少一个模板片段以及模板标签序列，所述模板标签序列由所述至少一个模板片段的模板属性标签构成；基于至少一个模板片段的模板属性标签和多个视频片段对应的片段属性标签，在多个视频片段中筛选与至少一个模板片段的模板属性标签匹配的至少一个视频片段；按照所述至少一个模板片段中各个模板片段的模板属性标签在所述模板标签序列中的位置，将匹配的至少一个视频片段进行拼接得到的；

在所述应用客户端的应用显示界面中输出所述视频数据以及所述视频素材片段。

10、根据权利要求 9 所述的方法，其中，所述响应目标用户针对应用客户端中的目标视频执行的播放操作，从服务器上获取所述目标视频的视频数据，以及与所述目标视频相关联的视频素材片段，包括：

响应目标用户针对应用客户端中的目标视频执行的播放操作，生成用于请求播放所述目标视频的视频播放请求，将所述视频播放请求发送给服务器；所述视频播放请求中携带所述目标视频的视频标识；所述视频标识用于指示所述服务器获取所述目标用户所请求播放的目标视频的视频数据；

接收所述服务器基于所述视频播放请求返回的所述视频数据，以及与所述目标视频相关联的视频素材片段；所述视频素材片段是由所述服务器在根据所述目标用户的用户画像确定出视频模板时，根据所述视频模板对所述视频数据进行视频分析以及视频匹配后所得到的，所述用户画像由所述目标用户在所述应用客户端中的用户行为信息所确定的。

11、根据权利要求 9 所述的方法，其中，所述在所述应用客户端的应用显示界面中输出所述视频数据以及所述视频素材片段，包括：

在所述应用客户端的应用显示界面中确定用于播放所述视频数据的视频播放界面，在所述视频播放界面中播放所述视频数据；

响应针对所述应用显示界面的触发操作，在所述应用显示界面中播放所述视频素材片段。

12、一种视频数据处理装置，包括：

片段生成模块，用于获取目标用户请求的目标视频的视频数据，对所述视频数据进行视频分析得到多个视频片段，其中，所述视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析，所述多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段；

模板获取模块，用于基于所述目标用户的用户画像，从视频模板数据库中确定与所述目标用户相关联的视频模板，并获取所述视频模板中预先确定的至少一个模板片段以及模板标签序列，所述模板标签序列由所述至少一个模板片段的模板属性标签构成；

素材确定模块,用于基于所述至少一个模板片段的模板属性标签和所述多个视频片段对应的片段属性标签,在所述多个视频片段中筛选与所述至少一个模板片段的模板属性标签匹配的至少一个视频片段,按照所述至少一个模板片段中各个模板片段的模板属性标签在所述模板标签序列中的位置,将匹配的至少一个视频片段进行拼接,作为所述目标视频的视频素材片段;

数据发送模块,用于将所述视频数据以及所述视频素材片段推送至所述目标用户对应的应用客户端,以使所述应用客户端输出所述视频数据以及所述视频素材片段。

13、一种视频数据处理装置,包括:

数据获取模块,用于响应目标用户针对应用客户端中的目标视频执行的播放操作,从服务器上获取所述目标视频的视频数据,以及与所述目标视频相关联的视频素材片段;所述视频素材片段是由所述服务器对所述视频数据进行视频分析得到多个视频片段,所述视频分析包括分镜处理和基于多个预设的片段属性标签的属性分析,所述多个视频片段中的每一个视频片段对应一个片段属性标签和一个分镜片段;基于目标用户的用户画像,从视频模板数据库中确定与目标用户相关联的视频模板,并获取视频模板中预先确定的至少一个模板片段以及模板标签序列,所述模板标签序列由所述至少一个模板片段的模板属性标签构成;基于至少一个模板片段的模板属性标签和多个视频片段对应的片段属性标签,在多个视频片段中筛选与至少一个模板片段的模板属性标签匹配的至少一个视频片段;按照所述至少一个模板片段中各个模板片段的模板属性标签在所述模板标签序列中的位置,将匹配的至少一个视频片段进行拼接得到的;

数据输出模块,用于在所述应用客户端的应用显示界面中输出所述视频数据以及所述视频素材片段。

14、一种计算机设备,包括:处理器、存储器、网络接口;

所述处理器与存储器、网络接口相连,其中,网络接口用于提供数据通信功能,所述存储器用于存储计算机程序,所述处理器用于调用所述计算机程序,以执行权利要求 1-11 任一项所述的方法。

15、一种计算机可读存储介质,所述计算机可读存储介质存储有计算机程序,该计算机程序适于由处理器加载并执行权利要求 1-11 任一项所述的方法。

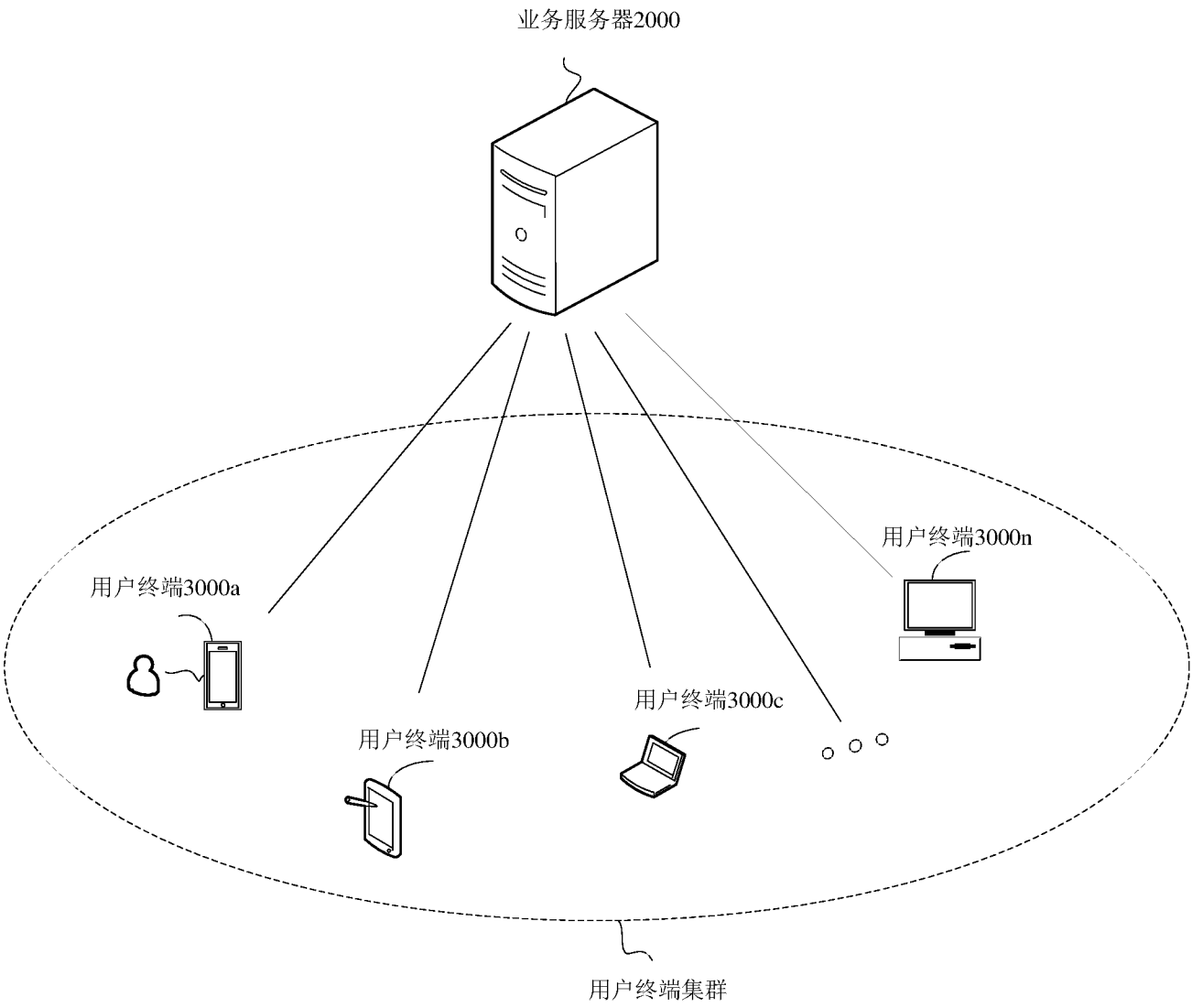


图 1

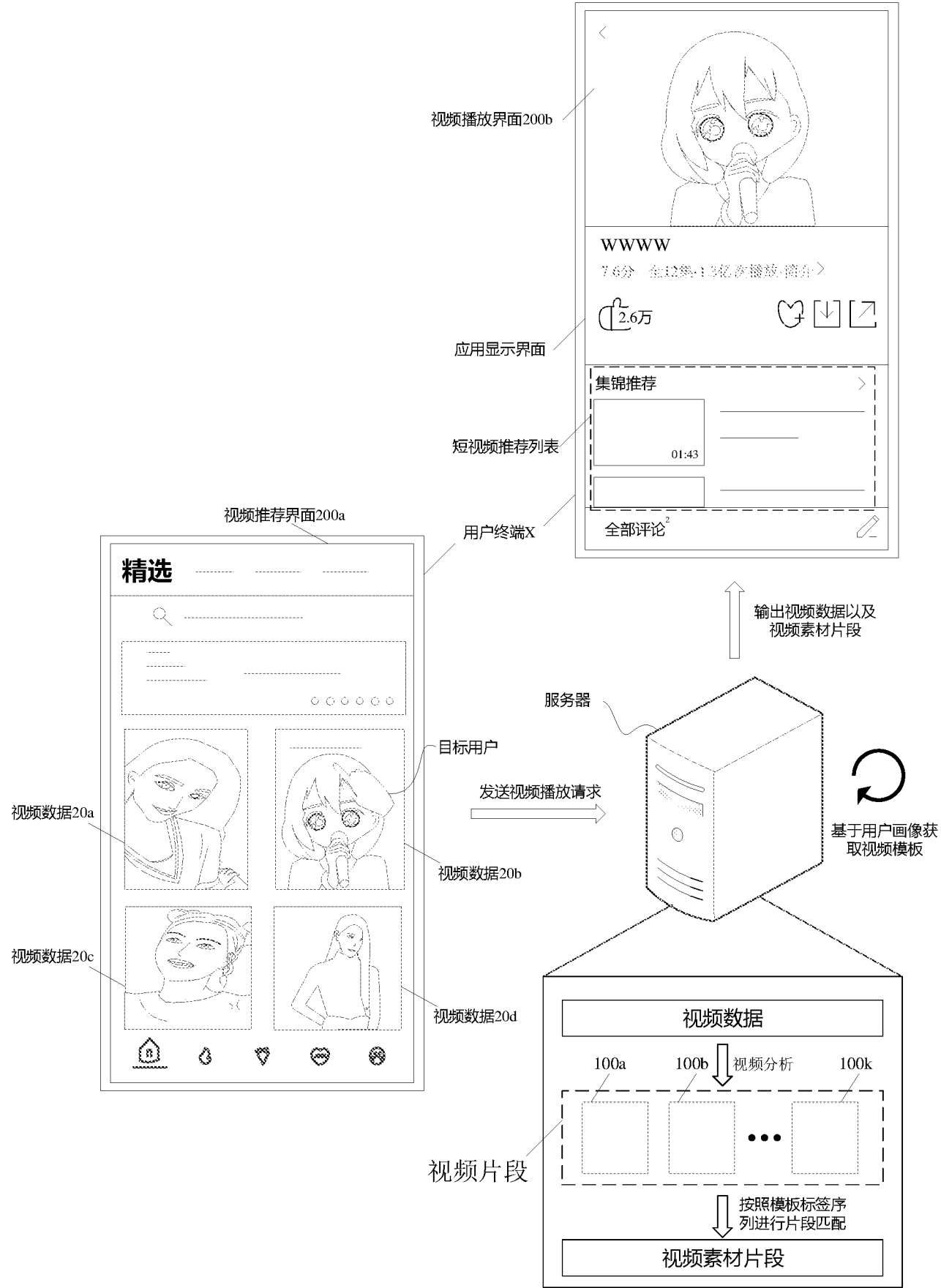


图 2

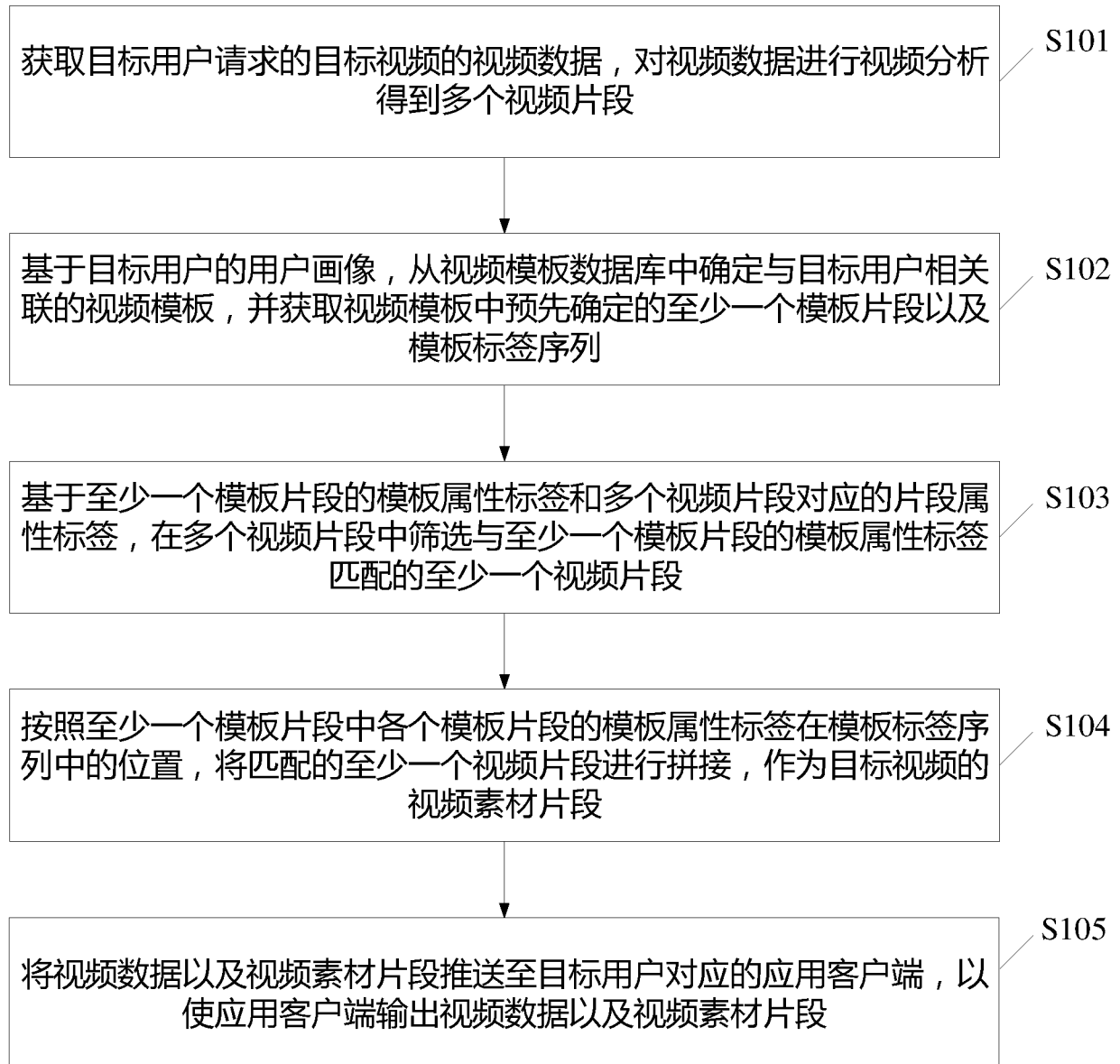


图 3

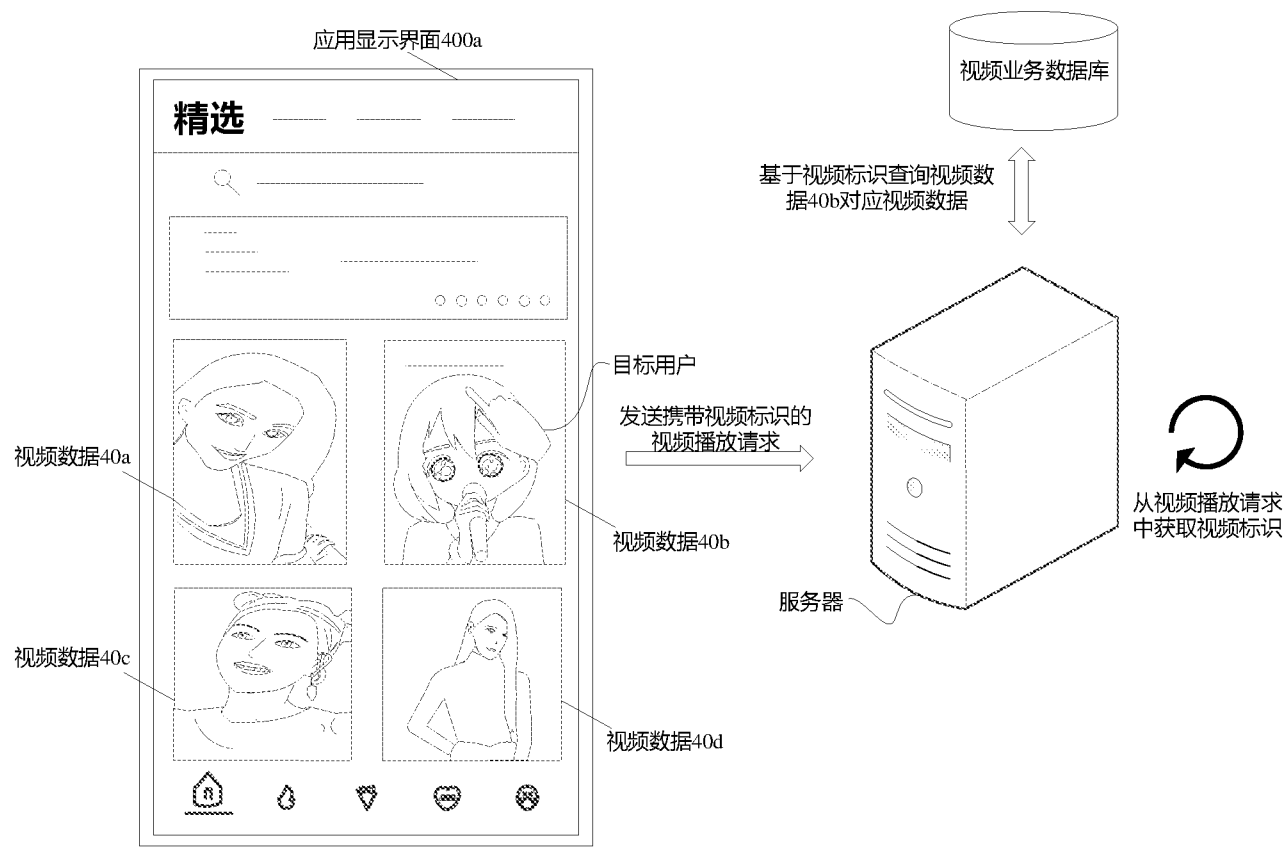


图 4

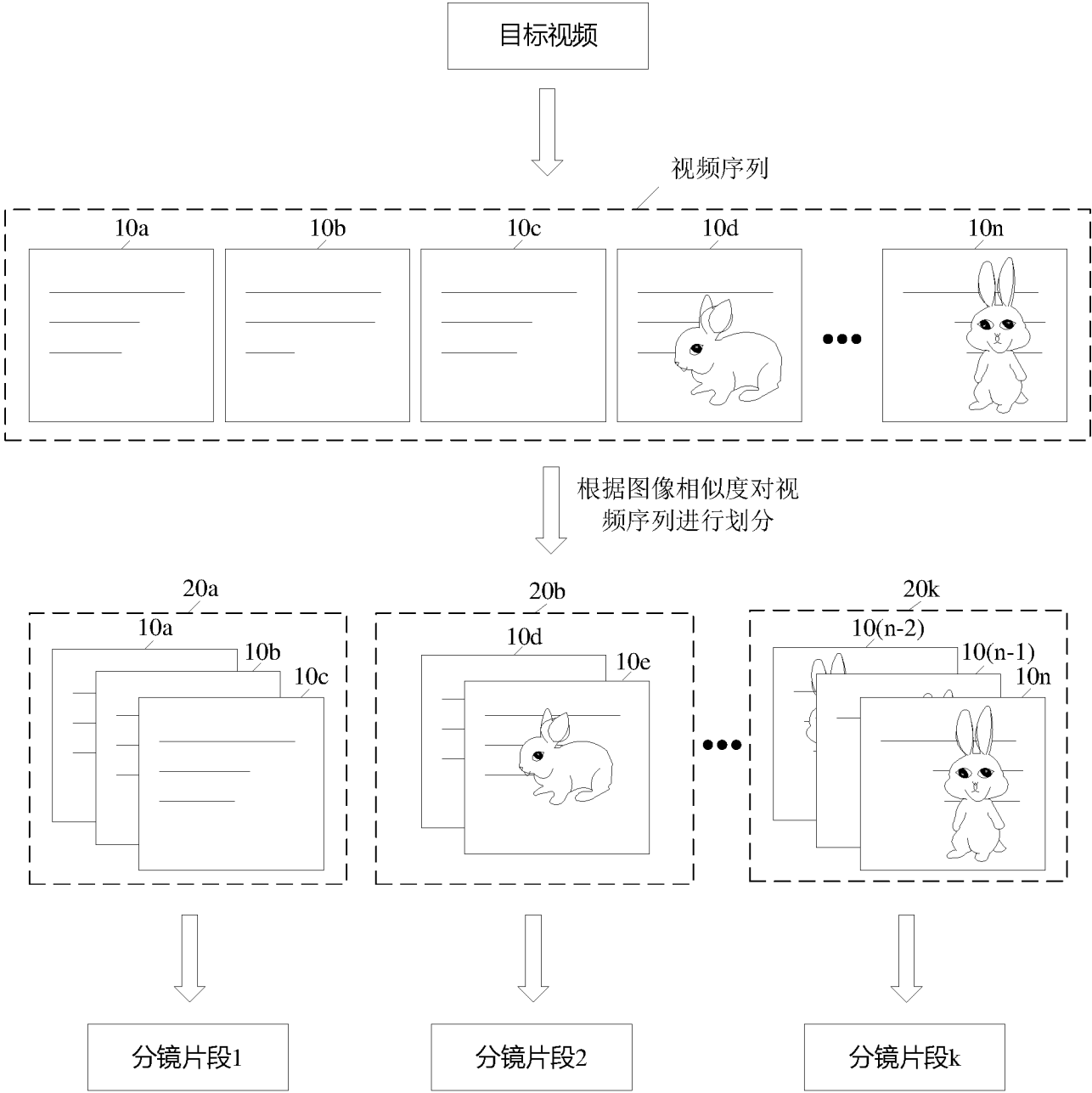


图 5

6/16

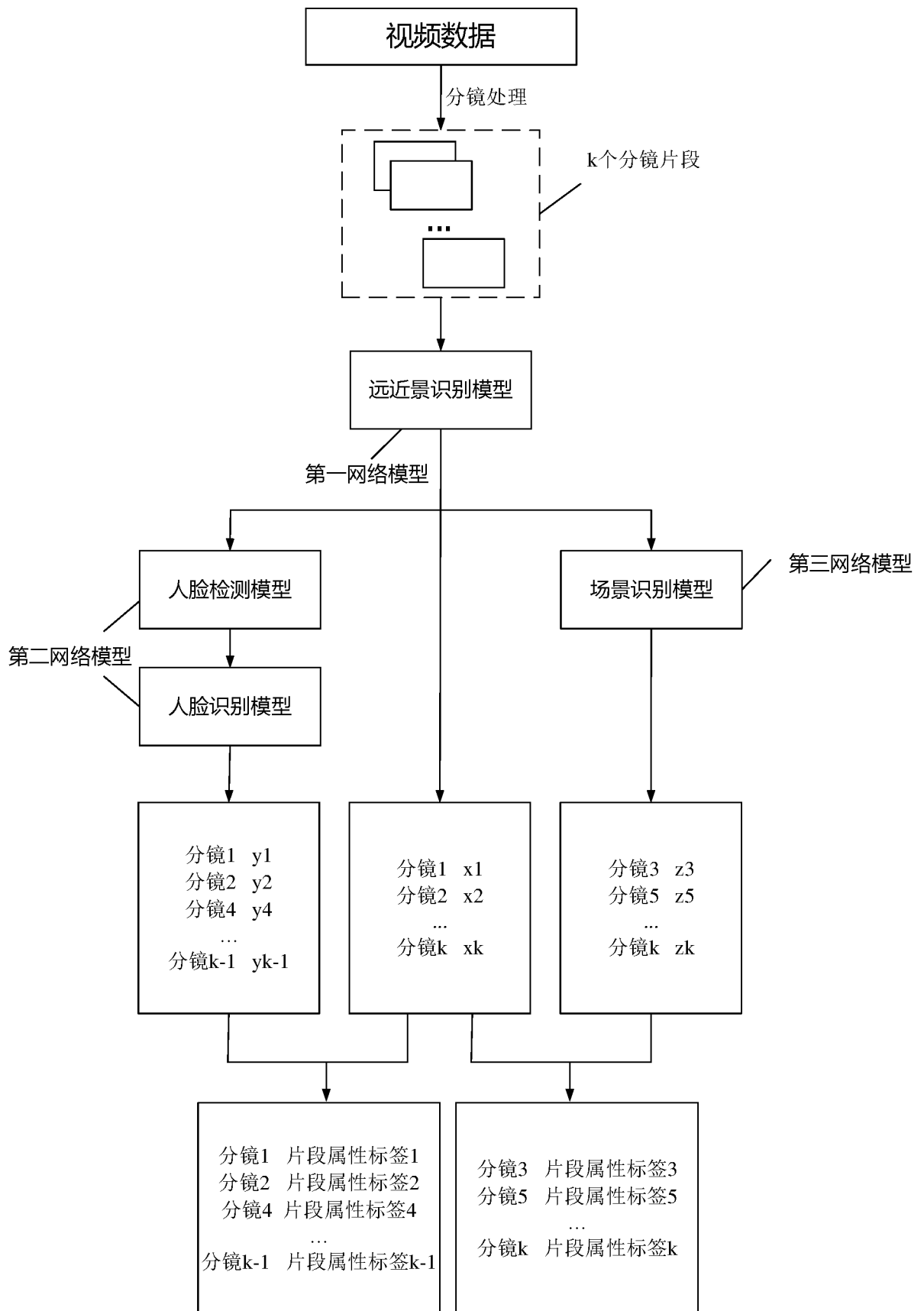


图 6

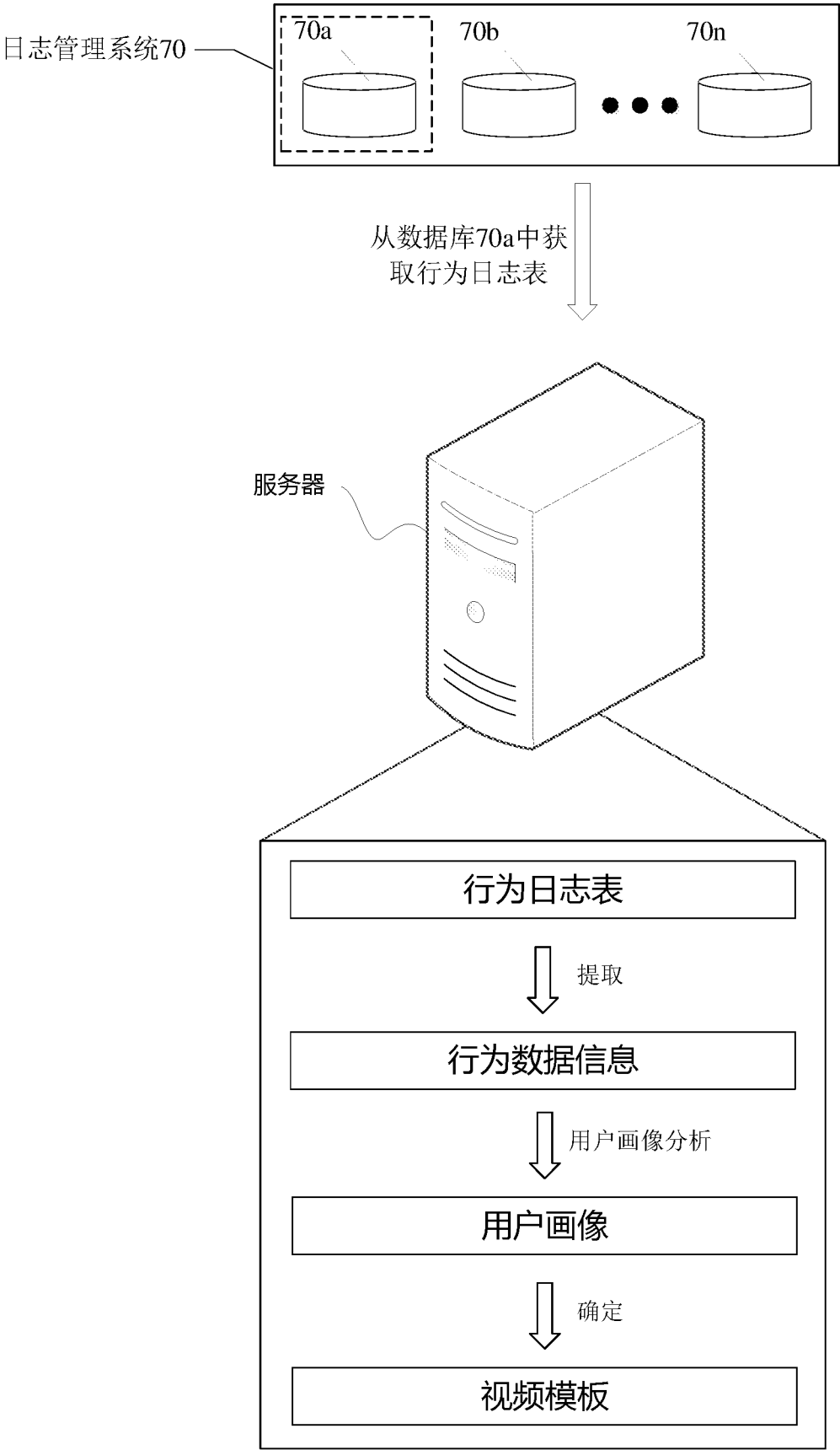


图 7

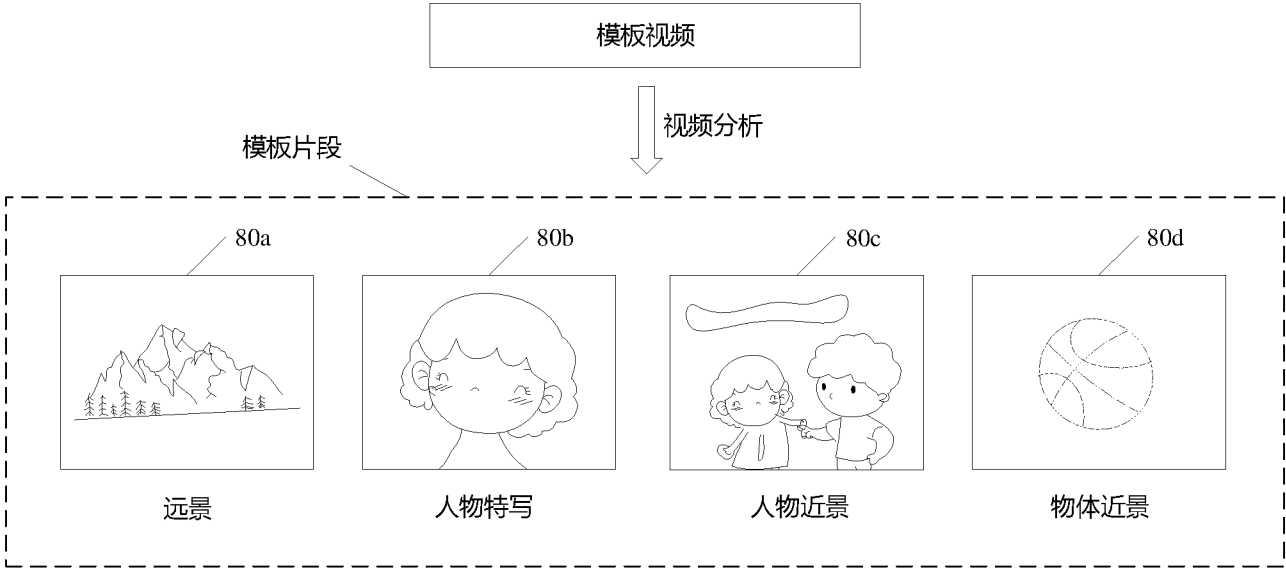


图 8A

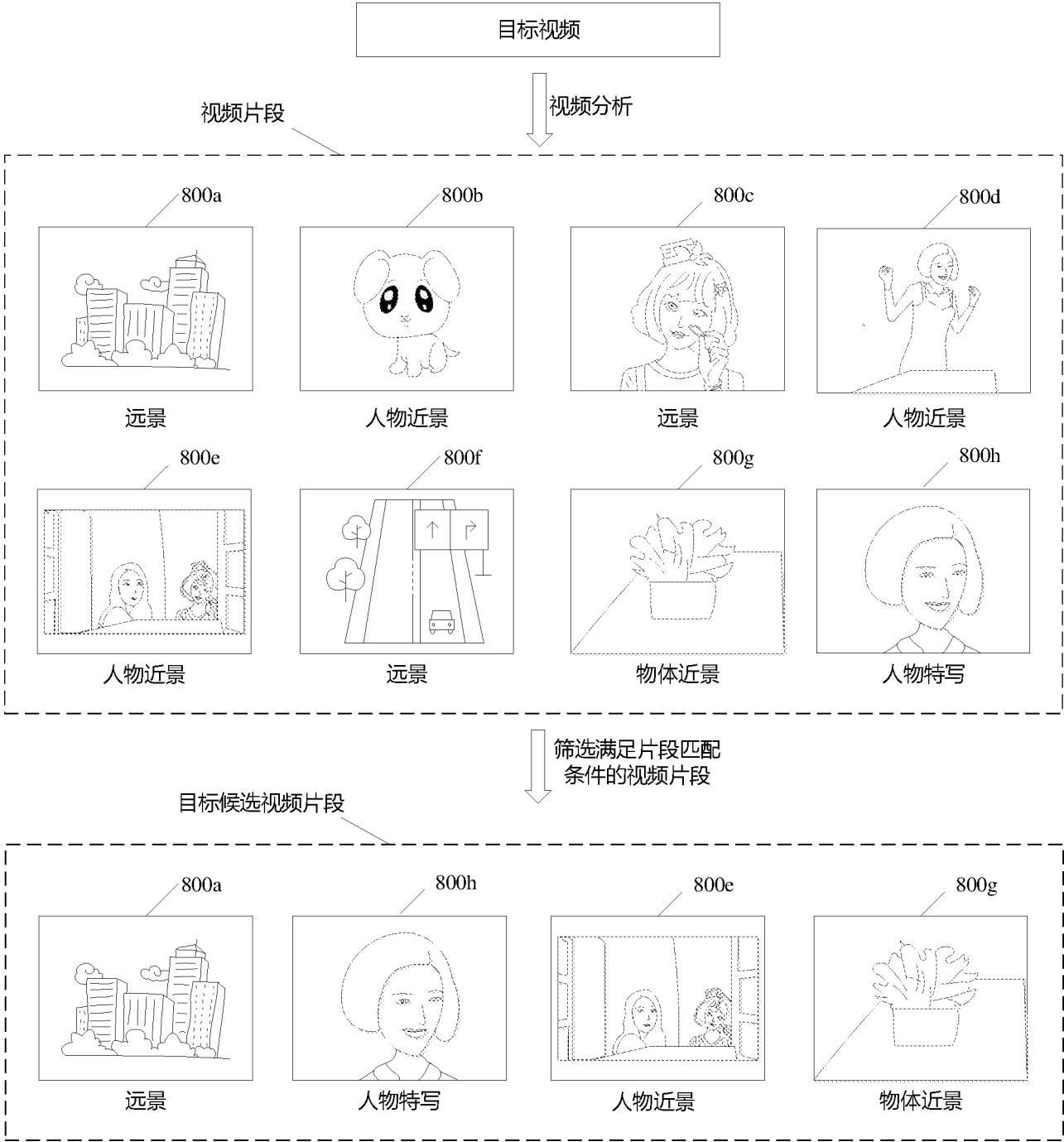


图 8B

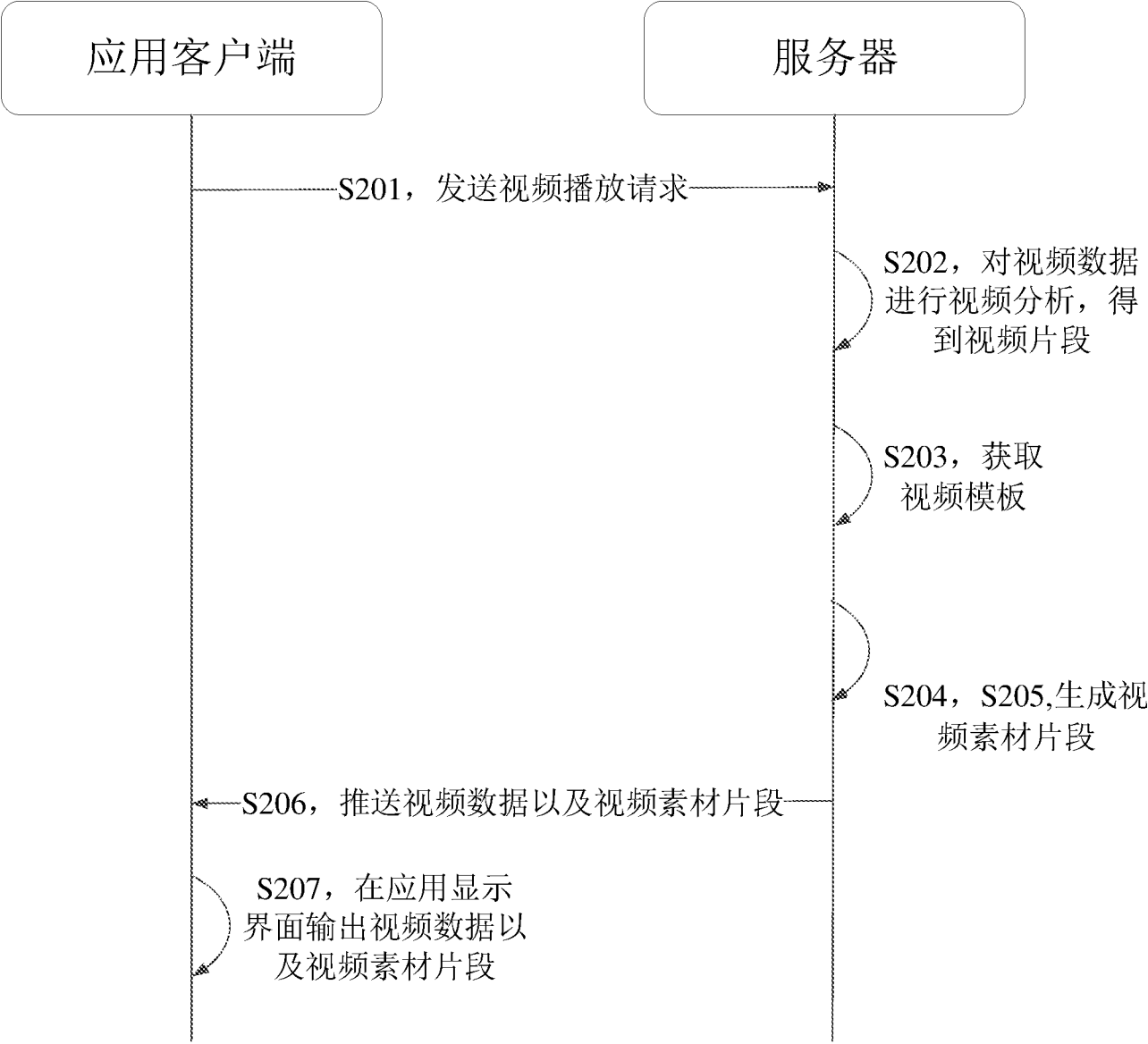


图 9

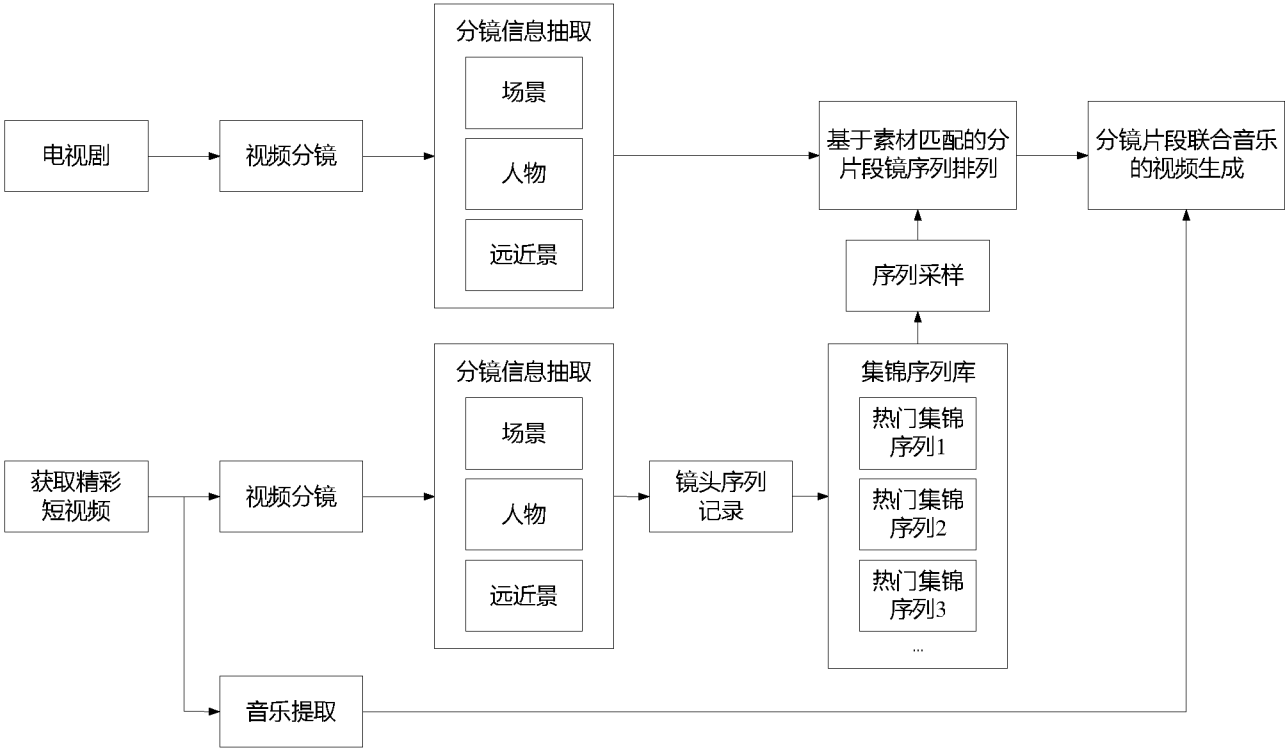


图 10

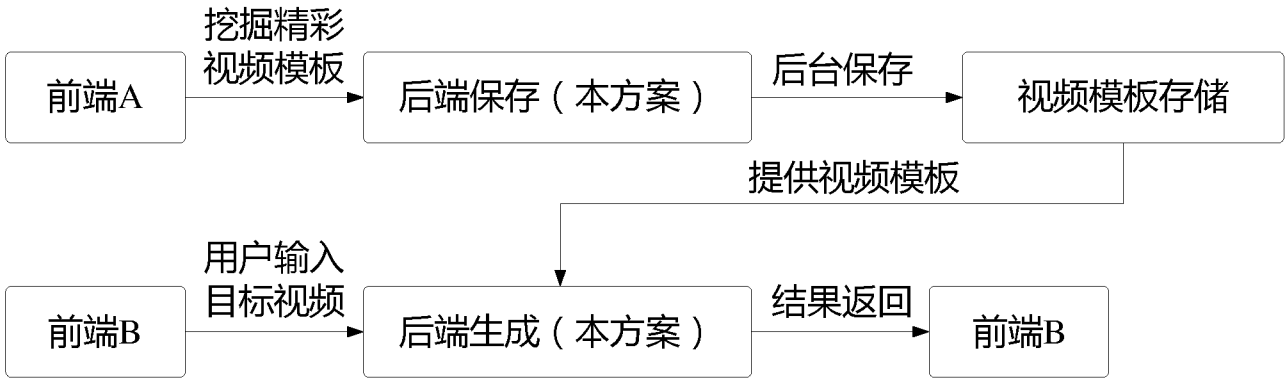


图 11

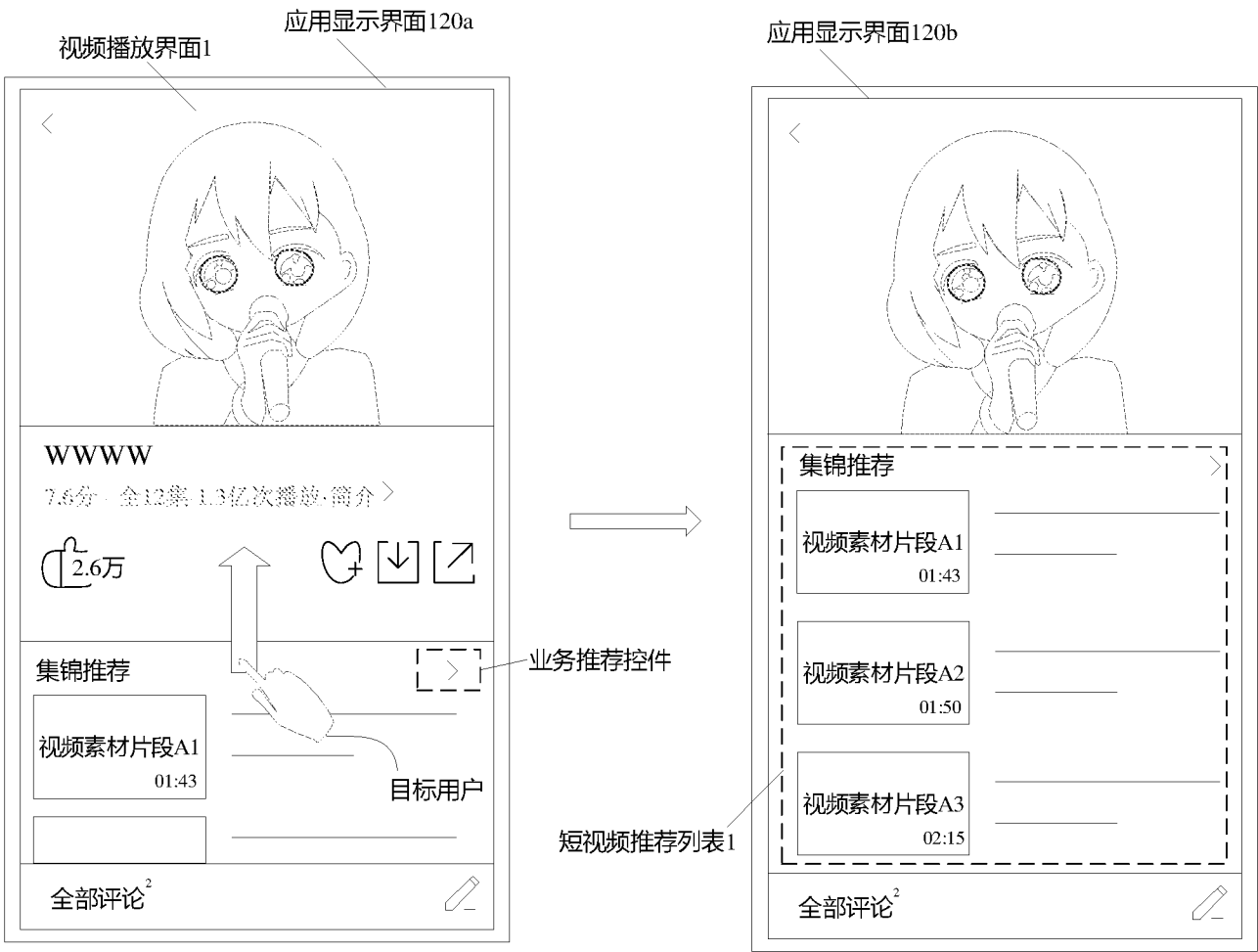


图 12A

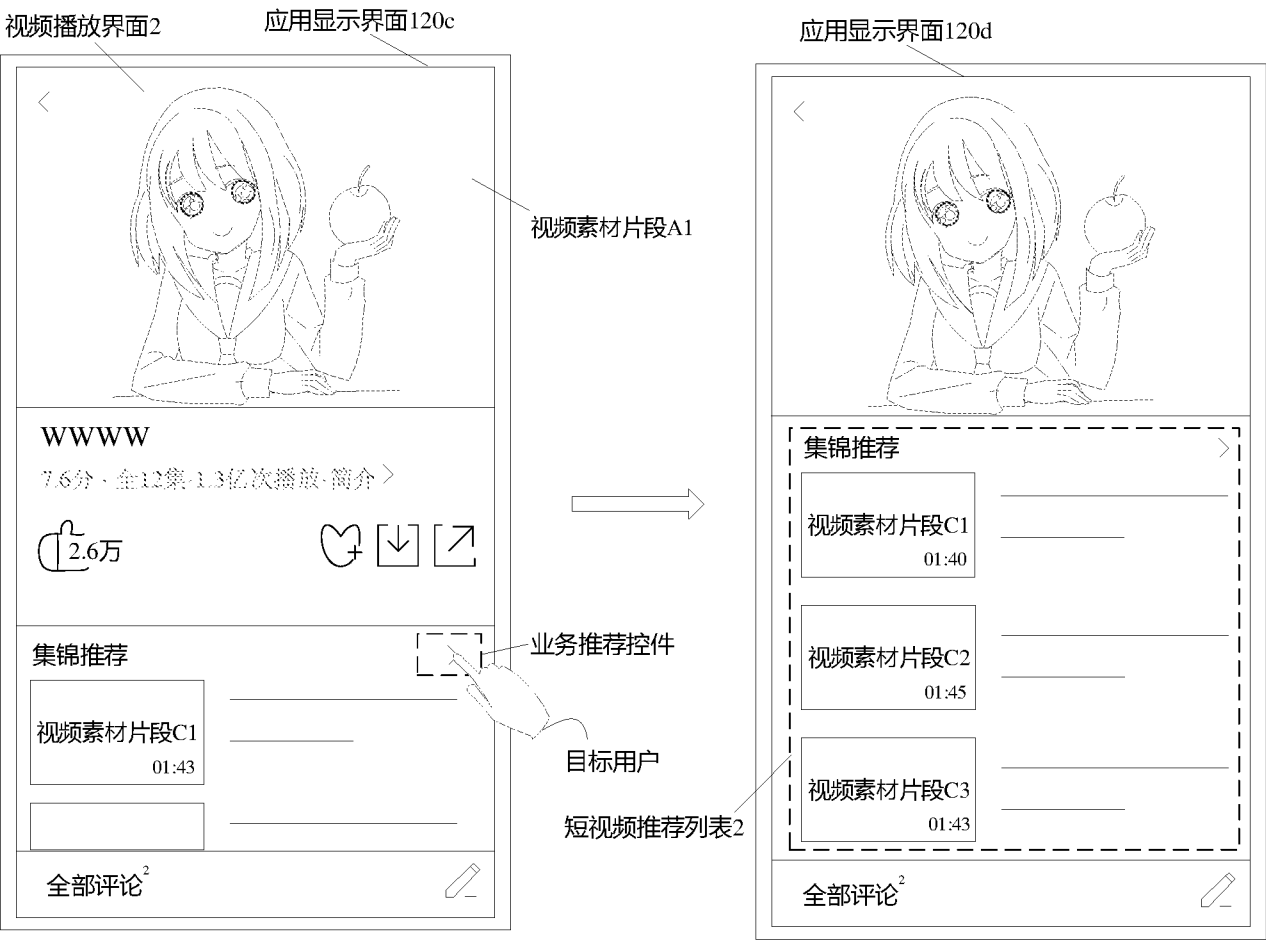


图 12B



图 13

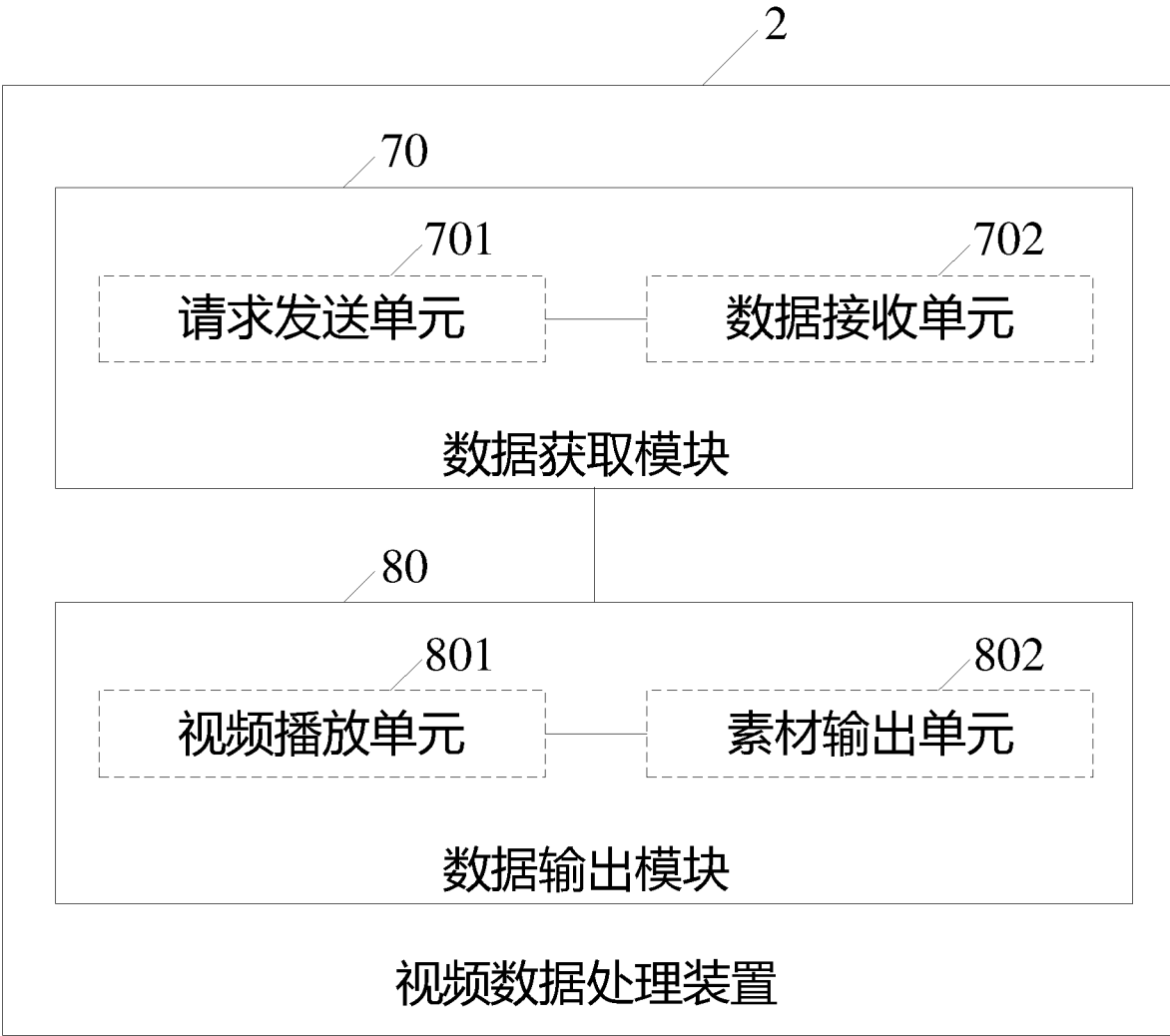


图 14

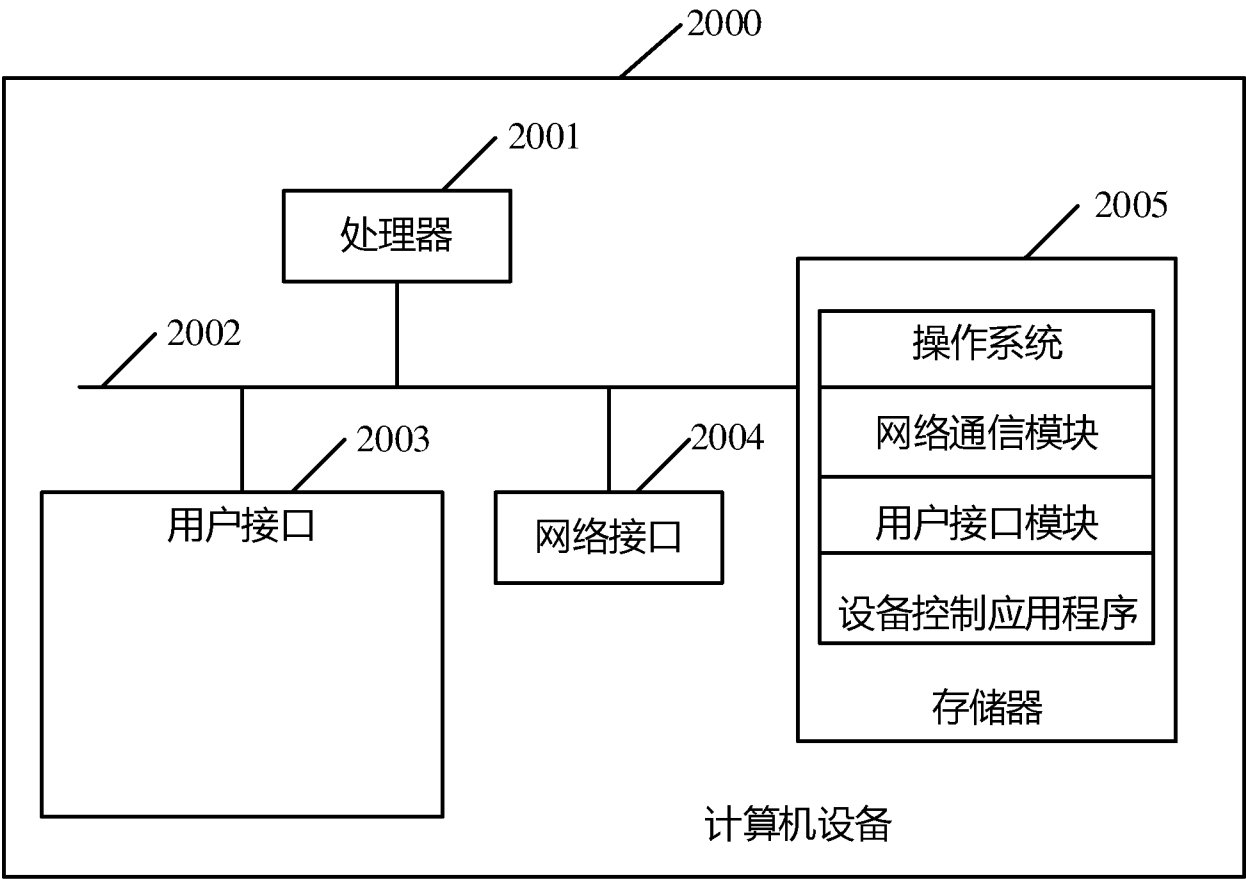


图 15

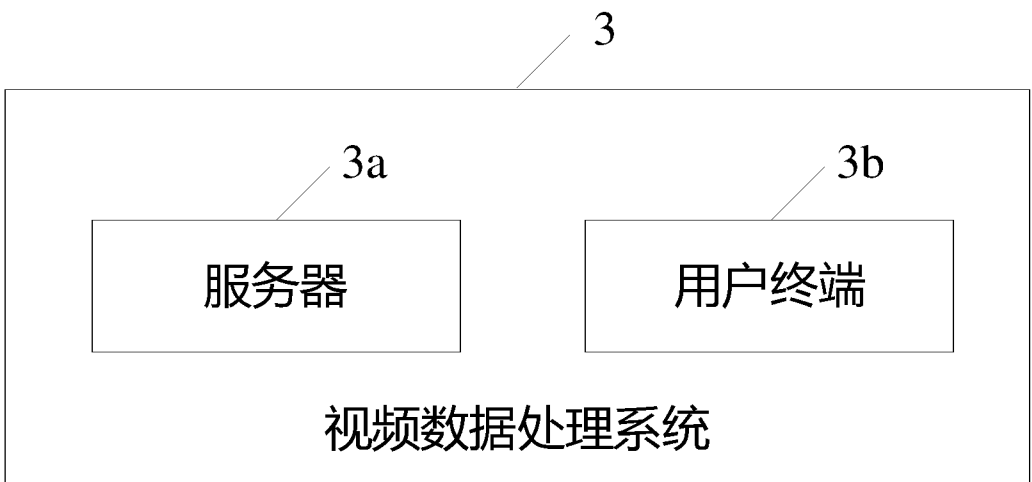


图 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/133035

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/239(2011.01)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)

H04N

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; DWPI; SIPOABS; VEN; EXTXT; EXTXTC; USTXT; WOTXT; EPTXT; CNKI: 视频, 属性, 分镜, 分段, 片段, 标签, 用户画像, 用户, 模板, 序列, 匹配, 素材, 顺序, 序号, 拼接, 推荐, 行为, 喜爱, video, value, nature, fragment, segment +, section, identifi+, tag, label, behavior, user, module, sequence, order, match+, recommend+, love, interest+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 112565825 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 26 March 2021 (2021-03-26) claims 1-15	1-15
Y	CN 111866585 A (BEIJING MEISHE NETWORK TECHNOLOGY CO., LTD.) 30 October 2020 (2020-10-30) description, paragraphs [0057]-[0168], and figures 1-5	1-4, 6-15
Y	CN 111105819 A (BEIJING DAJIA INTERCONNECTION INFORMATION TECHNOLOGY CO., LTD.) 05 May 2020 (2020-05-05) description, paragraphs [0081]-[0219], and figures 1-11	1-4, 6-15
Y	CN 103593464 A (HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY) 19 February 2014 (2014-02-19) description, paragraphs [0045]-[0107], and figures 1-3	3, 4
Y	CN 110139159 A (SHANGHAI MOXIANG NETWORK TECHNOLOGY CO., LTD.) 16 August 2019 (2019-08-16) description, paragraphs [0041]-[0118], and figures 1-5	6, 10



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

09 February 2022

Date of mailing of the international search report

18 February 2022

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

International application No.

PCT/CN2021/133035**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2020322684 A1 (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 08 October 2020 (2020-10-08) entire document	1-15
A	WO 2018107914 A1 (ZTE CORPORATION) 21 June 2018 (2018-06-21) entire document	1-15

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/133035

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	112565825	A	26 March 2021	None			
CN	111866585	A	30 October 2020	None			
CN	111105819	A	05 May 2020	None			
CN	103593464	A	19 February 2014	None			
CN	110139159	A	16 August 2019	None			
US	2020322684	A1	08 October 2020	JP	2020536455	A	10 December 2020
				WO	2019109643	A1	13 June 2019
				CN	108024145	A	11 May 2018
WO	2018107914	A1	21 June 2018	CN	106686404	A	17 May 2017

国际检索报告

国际申请号

PCT/CN2021/133035

A. 主题的分类

H04N 21/239(2011.01) 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)

H04N

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;DWPI;SIPOABS;VEN;EXTXT;EXTXTC;USTXT;WOTXT;EPTXT;CNKI:视频, 属性, 分镜, 分段, 片段, 标签, 用户画像, 用户, 模板, 序列, 匹配, 素材, 顺序, 序号, 拼接, 推荐, 行为, 喜爱, video, value, nature, fragment, segment+, section, identif+, tag, label, behavior, user, module, sequence, order, match+, recommend+, love, interest+

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
PX	CN 112565825 A (腾讯科技深圳有限公司) 2021年3月26日 (2021 - 03 - 26) 权利要求1-15	1-15
Y	CN 111866585 A (北京美摄网络科技有限公司) 2020年10月30日 (2020 - 10 - 30) 说明书第[0057]-[0168]段, 附图1-5	1-4, 6-15
Y	CN 111105819 A (北京达佳互联信息技术有限公司) 2020年5月5日 (2020 - 05 - 05) 说明书第[0081]-[0219]段, 附图1-11	1-4, 6-15
Y	CN 103593464 A (华中科技大学) 2014年2月19日 (2014 - 02 - 19) 说明书第[0045]-[0107]段, 附图1-3	3, 4
Y	CN 110139159 A (上海摩象网络科技有限公司) 2019年8月16日 (2019 - 08 - 16) 说明书第[0041]-[0118]段, 附图1-5	6, 10
A	US 2020322684 A1 (BEIJING BAIDU NETCOM SCI & TEC) 2020年10月8日 (2020 - 10 - 08) 全文	1-15
A	WO 2018107914 A1 (ZTE CORP) 2018年6月21日 (2018 - 06 - 21) 全文	1-15

☐ 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

2022年2月9日

国际检索报告邮寄日期

2022年2月18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

孙佳琛

电话号码 62411520

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/133035

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	112565825	A	2021年3月26日	无			
CN	111866585	A	2020年10月30日	无			
CN	111105819	A	2020年5月5日	无			
CN	103593464	A	2014年2月19日	无			
CN	110139159	A	2019年8月16日	无			
US	2020322684	A1	2020年10月8日	JP	2020536455	A	2020年12月10日
				WO	2019109643	A1	2019年6月13日
				CN	108024145	A	2018年5月11日
WO	2018107914	A1	2018年6月21日	CN	106686404	A	2017年5月17日