

(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
2017 年 5 月 4 日 (04.05.2017)



(10) 国际公布号
WO 2017/071086 A1

- (51) 国际专利分类号:
H04N 7/18 (2006.01)
- (21) 国际申请号: PCT/CN2015/099593
- (22) 国际申请日: 2015 年 12 月 29 日 (29.12.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510713118.3 2015 年 10 月 28 日 (28.10.2015) CN
- (71) 申请人: 小米科技有限责任公司 (XIAOMI INC.)
[CN/CN]; 中国北京市海淀区清河中街 68 号华润五彩城购物中心二期 13 层, Beijing 100085 (CN)。
- (72) 发明人: 张涛 (ZHANG, Tao); 中国北京市海淀区清河中街 68 号华润五彩城购物中心二期 13 层由小米科技有限责任公司转交, Beijing 100085 (CN)。 陈志军 (CHEN, Zhijun); 中国北京市海淀区清河中街 68 号华润五彩城购物中心二期 13 层由小米科技有限责任公司转交, Beijing 100085 (CN)。 龙飞 (LONG, Fei); 中国北京市海淀区清河中街 68 号华

润五彩城购物中心二期 13 层由小米科技有限责任公司转交, Beijing 100085 (CN)。

- (74) 代理人: 北京派特恩知识产权代理有限公司 (CHINA PAT INTELLECTUAL PROPERTY OFFICE); 中国北京市海淀区海淀南路 21 号中关村知识产权大厦 B 座 2 层, Beijing 100080 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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,

[见续页]

(54) Title: METHOD AND DEVICE USED FOR VIDEO PLAYBACK

(54) 发明名称: 用于视频播放的方法及装置

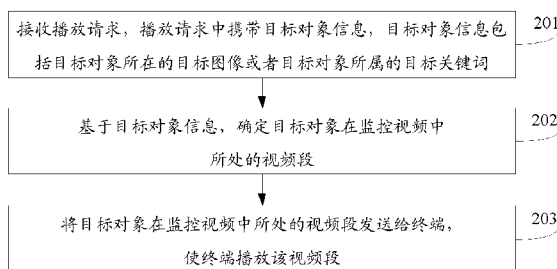


图 2

- 201 RECEIVING A PLAYBACK REQUEST, THE PLAYBACK REQUEST CARRYING TARGET OBJECT INFORMATION, AND THE TARGET OBJECT INFORMATION COMPRISING A TARGET IMAGE IN WHICH THE TARGET OBJECT IS LOCATED OR A TARGET KEYWORD TO WHICH THE TARGET OBJECT BELONGS;
- 202 ON THE BASIS OF THE TARGET OBJECT INFORMATION, DETERMINING THE VIDEO SEGMENT, OF THE MONITORING VIDEO, IN WHICH THE TARGET OBJECT APPEARS;
- 203 SENDING THE VIDEO SEGMENT, OF THE MONITORING VIDEO, IN WHICH THE TARGET OBJECT APPEARS, TO A TERMINAL, CAUSING THE TERMINAL TO PLAY THE VIDEO SEGMENT.

(57) Abstract: The embodiments of the present disclosure relate to a method and device used for video playback, belonging to the technical field of the Internet. The method comprises: receiving a playback request, said playback request carrying target object information, and said target object information comprising a target image in which the target object is located or a target keyword to which said target object belongs; on the basis of said target object information, determining the video segment, of the monitoring video, in which the target object appears; sending said video segment to a terminal, causing said terminal to play the video segment.

(57) 摘要: 本公开实施例是关于一种用于视频播放的方法及装置, 属于互联网技术领域。所述方法包括: 接收播放请求, 所述播放请求中携带目标对象信息, 所述目标对象信息包括目标对象所在的目标图像或者所述目标对象所属的目标关键词; 基于所述目标对象信息, 确定所述目标对象在监控视频中所处的视频段; 将所述视频段发送给终端, 使所述终端播放所述视频段。



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))。

用于视频播放的方法及装置

相关申请的交叉引用

本申请基于申请号为 201510713118.3、申请日为 2015 年 10 月 28 日的中国专利申请提出，并要求该中国专利申请的优先权，该中国专利申请的
5 全部内容在此引入本申请作为参考。

技术领域

本公开涉及互联网技术领域，尤其涉及一种用于视频播放的方法及装置。

背景技术

10 随着摄像头的普及，利用摄像头进行实时监控越来越流行。由于利用摄像头实时监控得到的监控视频具有时间上的连续性，所以，当用户想要查看监控视频中的特定对象的行为时，比如用户想要查看监控视频中婴儿的行为时，用户往往需要在监控视频播放的过程中，手动地对监控视频的播放进度进行调整，使监控视频前进或者后退到特定对象所在的视频画面，
15 从而查看特定对象的行为。而用户手动对监控视频的播放进度进行调整的操作比较繁琐，降低了视频播放效率，因此，亟需一种提高视频播放效率的用于视频播放的方法。

发明内容

为克服相关技术中存在的问题，本公开实施例提供一种用于视频播放
20 的方法及装置。

根据本公开实施例的第一方面，提供一种用于视频播放的方法，所述

方法包括:

接收播放请求, 所述播放请求中携带目标对象信息, 所述目标对象信息包括目标对象所在的目标图像或者所述目标对象所属的目标关键词;

5 基于所述目标对象信息, 确定所述目标对象在监控视频中所处的视频段;

将所述视频段发送给终端, 使所述终端播放所述视频段。

结合第一方面, 在上述第一方面的第一种可能的实现方式中, 所述基于所述目标对象信息, 确定所述目标对象在监控视频中所处的视频段, 包括:

10 当所述目标对象信息包括目标对象所在的目标图像时, 基于指定分类模型和所述目标图像, 确定所述目标对象所属的目标类别;

基于所述目标类别, 确定所述目标对象所属的目标关键词;

基于所述目标关键词, 确定所述目标对象在所述监控视频中所处的视频段。

15 结合第一方面, 在上述第一方面的第二种可能的实现方式中, 所述基于所述目标对象信息, 确定所述目标对象在监控视频中所处的视频段, 包括:

基于所述目标对象信息对应的所述目标关键词和存储的索引库, 获取所述目标对象在所述监控视频中所处的至少一帧视频图像;

20 将所述至少一帧视频图像组成所述目标对象在所述监控视频中所处的视频段。

结合第一方面的第二种可能的实现方式, 在上述第一方面的第三种可能的实现方式中, 所述基于所述目标对象信息对应的所述目标关键词和存储的索引库, 获取所述目标对象在所述监控视频中所处的至少一帧视频图像, 包括:

25

当所述索引库中存储关键词与监控时间点之间的对应关系时，基于所述目标对象信息对应的所述目标关键词，从所述关键词与监控时间点之间的对应关系中，获取至少一个监控时间点；

5 基于所述至少一个监控时间点，从所述监控视频中，获取至少一帧视频图像。

结合第一方面的第二种可能的实现方式，在上述第一方面的第四种可能的实现方式中，所述基于所述目标对象信息对应的所述目标关键词和存储的索引库，获取所述目标对象在所述监控视频中所处的至少一帧视频图像，包括：

10 当所述索引库中存储关键词与视频图像之间的对应关系时，基于所述目标对象信息对应的所述目标关键词，从所述关键词与视频图像之间的对应关系中，获取至少一帧视频图像。

结合第一方面的第二种可能的实现方式至第一方面的第四种可能的实现方式中任一可能的实现方式，在上述第一方面的第五种可能的实现方式
15 中，所述基于所述目标对象信息对应的所述目标关键词和存储的索引库，获取所述目标对象在所述监控视频中的至少一帧视频图像之前，还包括：

获取监控视频；

对于所述监控视频中每帧视频图像，基于指定分类模型，确定所述视频图像包括的对象所属的对象类别；

20 基于所述对象类别，确定所述视频图像包括的对象所属的关键词；

基于所述关键词和所述监控视频，生成所述索引库。

结合第一方面的第五种可能的实现方式，在上述第一方面的第六种可能的实现方式中，所述基于所述对象类别，确定所述视频图像包括的对象所属的关键词，包括：

25 当所述对象类别为人时，对所述视频图像包括的对象进行人脸识别，

得到人脸特征;

基于所述人脸特征,从存储的人脸特征与身份标识之间的对应关系中,获取对应的身份标识;

将所述身份标识确定为所述视频图像包括的对象所属的关键词。

- 5 结合第一方面的第五种可能的实现方式,在上述第一方面的第七种可能的实现方式中,所述基于所述关键词和所述监控视频,生成所述索引库,包括:

从所述监控视频中,确定所述视频图像所在的监控时间点;

- 10 将所述关键词和所述监控时间点存储在所述索引库包括的关键词与监控时间点之间的对应关系中。

结合第一方面的第五种可能的实现方式,在上述第一方面的第八种可能的实现方式中,所述基于所述关键词和所述监控视频,生成所述索引库,包括:

- 15 将所述关键词和所述视频图像存储在所述索引库包括的关键词与视频图像之间的对应关系中。

根据本公开实施例的第二方面,提供一种用于视频播放的装置,所述装置包括:

- 20 接收模块,配置为接收播放请求,所述播放请求中携带目标对象信息,所述目标对象信息包括目标对象所在的目标图像或者所述目标对象所属的目标关键词;

确定模块,配置为基于所述目标对象信息,确定所述目标对象在监控视频中所处的视频段;

发送模块,配置为将所述视频段发送给终端,使所述终端播放所述视频段。

- 25 结合第二方面,在上述第二方面的第一种可能的实现方式中,所述确

定模块包括:

第一确定单元,配置为当所述目标对象信息包括目标对象所在的目标图像时,基于指定分类模型和所述目标图像,确定所述目标对象所属的目标类别;

5 第二确定单元,配置为基于所述目标类别,确定所述目标对象所属的目标关键词;

第三确定单元,配置为基于所述目标关键词,确定所述目标对象在所述监控视频中所处的视频段。

10 结合第二方面,在上述第二方面的第二种可能的实现方式中,所述确定模块包括:

第一获取单元,配置为基于所述目标对象信息对应的所述目标关键词和存储的索引库,获取所述目标对象在所述监控视频中所处的至少一帧视频图像;

15 组成单元,配置为将所述至少一帧视频图像组成所述目标对象在所述监控视频中所处的视频段。

结合第二方面的第二种可能的实现方式,在上述第二方面的第三种可能的实现方式中,所述第一获取单元包括:

20 第一获取子单元,配置为当所述索引库中存储关键词与监控时间点之间的对应关系时,基于所述目标对象信息对应的所述目标关键词,从所述关键词与监控时间点之间的对应关系中,获取至少一个监控时间点;

第二获取子单元,配置为基于所述至少一个监控时间点,从所述监控视频中,获取至少一帧视频图像。

结合第二方面的第二种可能的实现方式,在上述第二方面的第四种可能的实现方式中,所述第一获取单元包括:

25 第三获取子单元,配置为当所述索引库中存储关键词与视频图像之间

的对应关系时，基于所述目标对象信息对应的所述目标关键词，从所述关键词与视频图像之间的对应关系中，获取至少一帧视频图像。

结合第二方面的第二种可能的实现方式至第二方面的第四种可能的实现方式中任一可能的实现方式，在上述第二方面的第五种可能的实现方式
5 中，所述确定模块还包括：

第二获取单元，配置为获取监控视频；

第四确定单元，配置为对于所述监控视频中每帧视频图像，基于指定分类模型，确定所述视频图像包括的对象所属的对象类别；

第五确定单元，配置为基于所述对象类别，确定所述视频图像包括的
10 对象所属的关键词；

生成单元，配置为基于所述关键词和所述监控视频，生成所述索引库。

结合第二方面的第五种可能的实现方式，在上述第二方面的第六种可能的实现方式中，所述第五确定单元包括：

识别子单元，配置为当所述对象类别为人时，对所述视频图像包括的
15 对象进行人脸识别，得到人脸特征；

第四获取子单元，配置为基于所述人脸特征，从存储的人脸特征与身份标识之间的对应关系中，获取对应的身份标识；

第一确定子单元，配置为将所述身份标识确定为所述视频图像包括的对象所属的关键词。

结合第二方面的第五种可能的实现方式，在上述第二方面的第七种可能的实现方式中，所述生成单元包括：

第二确定子单元，配置为从所述监控视频中，确定所述视频图像所在的监控时间点；

第一存储子单元，配置为将所述关键词和所述监控时间点存储在所述
25 索引库包括的关键词与监控时间点之间的对应关系中。

结合第二方面的第五种可能的实现方式，在上述第二方面的第八种可能的实现方式中，所述生成单元包括：

第二存储子单元，配置为将所述关键词和所述视频图像存储在所述索引库包括的关键词与视频图像之间的对应关系中。

5 根据本公开实施例的第三方面，提供一种用于视频播放的装置，所述装置包括：

处理器；

配置为存储处理器可执行指令的存储器；

其中，所述处理器被配置为：

10 接收播放请求，所述播放请求中携带目标对象信息，所述目标对象信息包括目标对象所在的目标图像或者所述目标对象所属的目标关键词；

基于所述目标对象信息，确定所述目标对象在监控视频中所处的视频段；

将所述视频段发送给终端，使所述终端播放所述视频段。

15 在本公开实施例中，服务器接收播放请求，该播放请求中携带目标对象信息，服务器基于目标对象信息确定目标对象在监控视频中所处的视频段，并将该视频段发送给终端，使终端可以直接播放该监控视频中存在目标对象的视频段，无需播放该监控视频中除该目标对象之外的其他对象所在的视频段，进而避免用户手动对该监控视频的播放进行调整来观看目标
20 对象所处的视频，操作简单，且提高了视频播放效率。

应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本公开。

附图说明

此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本
25 发明的实施例，并与说明书一起用于解释本发明的原理。

图 1 是根据一示例性实施例示出的一种用于视频播放的方法所涉及的实施环境的示意图；

图 2 是根据一示例性实施例示出的一种用于视频播放的方法的流程图；

图 3 是根据一示例性实施例示出的另一种用于视频播放的方法的流程图；

5 图；

图 4 是根据一示例性实施例示出的一种用于视频播放的装置的框图；

图 5 是根据一示例性实施例示出的一种确定模块的框图；

图 6 是根据一示例性实施例示出的另一种确定模块的框图；

图 7 是根据一示例性实施例示出的一种第一获取单元的框图；

10 图 8 是根据一示例性实施例示出的又一种确定模块的框图；

图 9 是根据一示例性实施例示出的一种第五确定单元的框图；

图 10 是根据一示例性实施例示出的一种生成单元的框图；

图 11 是根据一示例性实施例示出的另一种用于视频播放的装置的框图。

15 具体实施方式

这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本发明相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本
20 发明的一些方面相一致的装置和方法的例子。

图 1 是根据一示例性实施例示出的一种用于视频播放的方法所涉及的实施环境的示意图。如图 1 所示，该实施环境可以包括：服务器 101、智能摄像设备 102 和终端 103。服务器 101 可以是一台服务器，或者是由若干台服务器组成的服务器集群，或者是一个云计算服务中心，智能摄像设备 102
25 可以是智能摄像机，终端 103 可以是移动电话，计算机，平板设备等。服

服务器 101 和智能摄像设备 102 之间可以通过网络进行连接，服务器 101 与终端 103 之间也可以通过网络进行连接。服务器 101 配置为接收终端发送的播放请求，基于该播放请求获取对应的视频，并将该视频发送给终端。智能摄像设备 102 配置为采集监控区域内的监控视频，并将监控视频发送给服务器。终端 103 配置为接收服务器发送的视频，并播放该视频。

图 2 是根据一示例性实施例示出的一种用于视频播放的方法的流程图，如图 2 所示，该方法用于服务器中，包括以下步骤。

在步骤 201 中，接收播放请求，播放请求中携带目标对象信息，目标对象信息包括目标对象所在的目标图像或者目标对象所属的目标关键词。

在步骤 202 中，基于目标对象信息，确定目标对象在监控视频中所处的视频段。

在步骤 203 中，将目标对象在监控视频中所处的视频段发送给终端，使终端播放该视频段。

在本公开实施例中，服务器接收播放请求，该播放请求中携带目标对象信息，服务器基于目标对象信息确定目标对象在监控视频中所处的视频段，并将该视频段发送给终端，使终端可以直接播放该监控视频中存在目标对象的视频段，无需播放该监控视频中除该目标对象之外的其他对象所在的视频段，进而避免用户手动对该监控视频的播放进行调整来观看目标对象所处的视频，操作简单，且提高了视频播放效率。

在本公开的另一实施例中，基于目标对象信息，确定目标对象在监控视频中所处的视频段，包括：

当目标对象信息包括目标对象所在的目标图像时，基于指定分类模型和目标图像，确定目标对象所属的目标类别；

基于目标类别，确定目标对象所属的目标关键词；

基于目标关键词，确定目标对象在监控视频中所处的视频段。

其中，服务器基于指定分类模型和目标图像，确定目标对象所属的目标类别，并基于该目标类别，确定目标对象所属的目标关键词，可以便于服务器基于该目标关键词快速确定目标对象在监控视频中所处的视频段。

在本公开的另一实施例中，基于目标对象信息，确定目标对象在监控
5 视频中所处的视频段，包括：

基于目标对象信息对应的目标关键词和存储的索引库，获取目标对象在监控视频中所处的至少一帧视频图像；

将至少一帧视频图像组成目标对象在监控视频中所处的视频段。

由于终端发送播放请求是为了获取目标对象在监控视频中所处的视频
10 段，而视频段可以由至少一帧视频图像组成，因此，服务器基于目标关键词和存储的索引库，可以快速获取目标对象在监控视频中所处的至少一帧视频图像，提高了视频获取效率。

在本公开的另一实施例中，基于目标对象信息对应的目标关键词和存储的索引库，获取目标对象在监控视频中所处的至少一帧视频图像，包括：

15 当索引库中存储关键词与监控时间点之间的对应关系时，基于目标对象信息对应的目标关键词，从关键词与监控时间点之间的对应关系中，获取至少一个监控时间点；

基于至少一个监控时间点，从监控视频中，获取至少一帧视频图像。

由于监控视频中包括该监控视频包括的每帧视频图像对应的监控时间
20 点，因此，服务器可以确定目标关键词对应的至少一个监控时间点，并从监控视频中，获取该至少一个监控时间点对应的至少一帧视频图像，提高了视频图像的获取准确性。

在本公开的另一实施例中，基于目标对象信息对应的目标关键词和存储的索引库，获取目标对象在监控视频中所处的至少一帧视频图像，包括：

25 当索引库中存储关键词与视频图像之间的对应关系时，基于目标对象

信息对应的目标关键词，从关键词与视频图像之间的对应关系中，获取至少一帧视频图像。

其中，服务器基于目标关键词，直接获取目标对象对应的至少一帧视频图像，提高了视频图像获取效率。

- 5 在本公开的另一实施例中，基于目标对象信息对应的目标关键词和存储的索引库，获取目标对象在监控视频中的至少一帧视频图像之前，还包括：

获取监控视频；

- 10 对于监控视频中每帧视频图像，基于指定分类模型，确定视频图像包括的对象所属的对象类别；

基于对象类别，确定视频图像包括的对象所属的关键词；

基于关键词和监控视频，生成索引库。

- 15 其中，服务器基于关键词和监控视频，生成索引库，可以便于服务器在接收到播放请求时，基于该索引库，快速获取目标对象在监控视频中所处的至少一帧视频图像，提高了视频图像的获取效率。

在本公开的另一实施例中，基于对象类别，确定视频图像包括的对象所属的关键词，包括：

当对象类别为人时，对视频图像包括的对象进行人脸识别，得到人脸特征；

- 20 基于人脸特征，从存储的人脸特征与身份标识之间的对应关系中，获取对应的身份标识；

将身份标识确定为视频图像包括的对象所属的关键词。

- 25 其中，当对象类别为人时，服务器将对象的身份标识确定为该对象所述的目标关键词，可以便于终端获取具有特定身份的人在监控视频中所处的至少一帧视频图像，可以针对性地获取某一个人的视频段。

在本公开的另一实施例中，基于关键词和监控视频，生成索引库，包括：

从监控视频中，确定视频图像所在的监控时间点；

将关键词和监控时间点存储在索引库包括的关键词与监控时间点之间的对应关系中。

由于监控视频中包括每帧视频图像对应的监控时间点，因此，服务器将关键词和监控时间点存储在索引库包括的关键词与监控时间点之间的对应关系中，可以便于服务器基于关键词，获取对应的监控时间点，进而从监控视频中，获取该监控时间点对应的视频图像，提高了视频图像的获取准确性。

在本公开的另一实施例中，基于关键词和监控视频，生成索引库，包括：

将关键词和视频图像存储在索引库包括的关键词与视频图像之间的对应关系中。

其中，服务器将关键词和视频图像存储在索引库包括的关键词与视频图像之间的对应关系中，可以便于服务器基于关键词，直接获取对应的视频图像，提高了视频图像获取效率。

上述所有可选技术方案，均可按照任意结合形成本公开的可选实施例，本公开实施例对此不再一一赘述。

图 3 是根据一示例性实施例示出的一种用于视频播放的方法的流程图，如图 3 所示，该方法包括以下步骤。

在步骤 301 中，服务器接收播放请求，该播放请求中携带目标对象信息，目标对象信息包括目标对象所在的目标图像或者目标对象所属的目标关键词。

需要说明的是，该播放请求可以由终端直接发送，当然，该播放请求

也可以由终端先发送给其它设备，再由其它设备发送给服务器，本公开实施例对此不做具体限定。

另外，当终端发送播放请求时，该终端可以在接收到播放指令时，向服务器或者其它设备发送播放请求。

- 5 需要说明的是，播放指令配置为获取目标对象在监控视频中所处的视频段，且播放指令可以由用户触发，用户可以通过指定操作触发，该指定操作可以为点击操作、滑动操作、语音操作等等，本公开实施例对此不做具体限定。

10 另外，目标图像为包括目标对象的图像，且该目标图像可以为目标对象的照片，或者是终端在播放监控视频的过程中，当终端基于监控视频的视频图像接收到选择指令时，该选择指令中携带目标对象的图像，当然，目标图像还可以通过其它方式获取得到，本公开实施例对此不做具体限定。

15 再者，目标关键词与目标对象唯一对应，且目标关键词可以为目标对象所属的类别、目标对象的身份标识等等，本公开实施例对此不做具体限定。

在步骤 302 中，服务器基于目标对象信息，确定目标对象在监控视频中所处的视频段。

20 由于目标对象信息中包括目标对象所在的图像或者目标对象所属的目标关键词，因此，根据目标对象信息包括的不同内容，服务器基于目标对象信息，确定目标对象在监控视频中所处的视频段可以有如下两种方式：

第一种方式：基于目标对象信息对应的目标关键词和存储的索引库，获取目标对象在监控视频中所处的至少一帧视频图像，并将该至少一帧视频图像组成目标对象在该监控视频中所处的视频段。

25 由于终端发送播放请求是为了获取目标对象在监控视频中所处的视频段，而视频段可以由至少一帧视频图像组成，因此，服务器可以获取目标

对象在监控视频中所处的至少一帧视频图像，并将该至少一帧视频图像组成目标对象在监控视频中所处的视频段。

需要说明的是，当目标对象信息包括目标关键词时，目标对象信息对应的目标关键词可以为该目标对象信息中包括的目标关键词，当目标对象
5 信息中包括目标图像时，目标对象信息对应的目标关键词可以通过该目标图像获取得到。

其中，服务器基于目标关键词和存储的索引库，获取目标对象在监控视频中所处的至少一帧视频图像时，当索引库中存储关键词与监控时间点之间的对应关系时，服务器基于目标关键词，从该关键词与监控时间点之
10 间的对应关系中，获取至少一个监控时间点，并基于该至少一个监控时间点，从监控视频中，获取至少一帧视频图像。或者，当索引库中存储关键词与视频图像之间的对应关系时，服务器基于目标关键词，从该关键词与视频图像之间的对应关系中，获取至少一帧视频图像。

由于监控视频中包括该监控视频包括的每帧视频图像对应的监控时间
15 点，因此，当服务器基于目标关键词，该关键词与监控时间点之间的对应关系中，获取该目标关键词对应的至少一个监控时间点之后，服务器可以基于该至少一个监控时间点，从该监控视频中，获取该至少一个监控时间点对应的至少一帧视频图像。

其中，服务器基于该至少一个监控时间点，从监控视频中，获取至少
20 一帧视频图像的过程可以参考相关技术，本公开实施例在此不进行详细阐述。

例如，当索引库中存储关键词与监控时间点之间的对应关系时，假设目标关键词为杨乐乐，则服务器可以基于目标关键词杨乐乐，从如下表 1 所示的关键词与监控时间点之间的对应关系中，获取杨乐乐对应的至少一个
25 个 监 控 时 间 点 为 2015/02/03-21:08:31 、 2015/03/05-11:08:11 、

2015/08/03-09:05:31，之后，服务器可以从监控视频中，获取2015/02/03-21:08:31、2015/03/05-11:08:11、2015/08/03-09:05:31分别对应的视频图像。

表 1

关键词	监控时间点
杨乐乐	2015/02/03-21:08:31、2015/03/05-11:08:11、2015/08/03-09:05:31
郭杰	2015/05/13-21:10:02、2015/05/25-21:03:06、2015/08/11-19:18:26
... ..	

5

需要说明的是，在本公开实施例中，仅以上述表 1 所示的关键词与监控时间点之间的对应关系为例进行说明，上述表 1 并不对本公开实施例构成限定。

再例如，当索引库中存储关键词与视频图像之间的对应关系时，假设
10 目标关键词为杨乐乐，服务器可以基于目标关键词杨乐乐，从如下表 2 所示关键词与视频图像之间的对应关系中，获取杨乐乐对应的至少一帧视频图像为 1. JPEG、2. JPEG、3. JPEG。

表 2

关键词	视频图像
杨乐乐	1. JPEG、2. JPEG、3. JPEG
郭杰	1. JPEG、2. JPEG、3. JPEG
... ..	

15 需要说明的是，在本公开实施例中，仅以上述表 2 所示的关键词与视频图像之间的对应关系为例进行说明，上述表 2 并不对本公开实施例构成限定。

其中，服务器将该至少一帧视频图像组成目标对象在监控视频中所处的视频段的过程可以参考相关技术，本发明实施例在此不进行详细阐述。

进一步地，服务器基于目标关键词和存储的索引库，获取目标对象在监控视频中的至少一帧视频图像之前，还可以通过如下步骤（1）-（4）来生成索引库，包括：

（1）服务器获取监控视频。

5 需要说明的是，服务器可以从智能摄像设备中获取该监控视频，当然，该智能摄像设备也可以将该监控视频发送到其它设备中，以使服务器可以从其它设备中获取该监控视频，本公开实施例对此不做具体限定。

其中，智能摄像设备配置为采集监控区域内的监控视频，且智能摄像设备采集监控区域内的监控视频的过程可以参考相关技术，本公开实施例
10 在此不进行详细阐述。

另外，智能摄像设备可以通过有线网络或者无线网络和服务器或者其它设备进行通信，而当智能摄像设备通过无线网络和服务器或者其它设备进行通信时，智能摄像设备可以通过内置的无线保真（英文：Wireless-Fidelity，简称：WIFI）、蓝牙或者其它无线通信芯片来和服务器或
15 者其它设备进行通信，本公开实施例对此不做具体限定。

（2）对于该监控视频中每帧视频图像，服务器基于指定分类模型，确定该视频图像包括的对象所属的对象类别。

需要说明的是，指定分类模型用于确定图像所对应的对象类别，且指定分类模型可以预先建立。

20 由于指定分类模型可以预先建立，且为了提高对象类别确定的效率，指定分类模型一般可以对预设尺寸的图像进行处理，以确定该图像包括的对象所属的对象类别，因此，服务器基于指定分类模型，确定该视频图像包括的对象所属的对象类别时，服务器可以在监控视频的视频图像中，对该对象所在的区域进行裁剪，得到对象图像，并将对象图像的尺寸处理为
25 预设尺寸，之后，基于指定分类模型和处理后的对象图像，确定该对象所

属的对象类别。

其中，服务器在监控视频的视频图像中，对该对象所在的区域进行裁剪，得到对象图像时，可以从该对象所在的视频图像中，截取该对象的外接矩形，并将该外接矩形确定为该对象在监控视频中所处的图像区域，即对象图像。当然，服务器也可以以其它方式对该对象所在的区域进行裁剪，得到对象图像，本公开实施例对此不做具体限定。

另外，预设尺寸可以预先设置，如预设尺寸可以为 224*224 像素、300*300 像素等等，本公开实施例对此不做具体限定。

需要说明的是，服务器基于指定分类模型和处理后的对象图像，确定该对象所属的对象类别的过程可以参考相关技术，本公开实施例在此不进行详细阐述。

(3) 服务器基于该视频图像包括的对象所属的对象类别，确定该视频图像包括的对象所属的关键词。

由于该视频图像包括的对象所属的对象类别可能为人，也可能为宠物，还可能为其它物体，而当该视频图像包括的对象所属的对象类别为人或宠物时，终端需要获取具有特定身份的人或宠物在监控视频中所处的视频段，因此，服务器基于该视频图像包括的对象所属的对象类别，确定该视频图像包括的对象所属的关键词的操作可以为：当视频图像包括的对象所属的对象类别为人时，服务器对该对象进行人脸识别，得到人脸特征，之后，基于该人脸特征，从存储的人脸特征与身份标识之间的对应关系中，获取对应的身份标识，并将该身份标识确定为该视频图像包括的对象所属的关键词。当视频图像包括的对象所属的对象类别为宠物时，服务器基于该视频图像，获取宠物标识，并基于该宠物标识，从存储的宠物标识与身份标识的对应关系中，获取对应的身份标识，并将该身份标识确定为该视频图像包括的对象所属的关键词。当视频图像包括的对象所属的对象类别不为

人，也不为宠物时，服务器可以直接将该对象类别确定为该视频图像包括的对象所属的关键词。

需要说明的是，服务器对该对象进行人脸识别，得到人脸特征的过程可以参考相关技术，本公开实施例在此不进行详细阐述。

- 5 还需要说明的是，宠物标识用于唯一标识该宠物，且该宠物标识可以通过置于该宠物上的二维码、条形码或者其他可识别标识获取得到，本公开实施例对此不做具体限定。

例如，当视频图像包括的对象所属的对象类别为人时，服务器对该对象进行人脸识别，得到人脸特征为 A，之后，基于该人脸特征 A，从如下表 3 所示的人脸特征与身份标识之间的对应关系中，获取 A 对应的身份标识为杨乐乐，则服务器可以将杨乐乐确定为该视频图像包括的对象所属的关键词。

表 3

人脸特征	身份标识
A	杨乐乐
B	郭杰
... ..	

- 15 需要说明的是，在本公开实施例中，仅以上述表 3 所示的人脸特征与身份标识之间的对应关系为例进行说明，上述表 3 并不对本公开实施例构成限定。

再例如，当视频图像包括的对象所属的对象类别为宠物时，服务器基于该视频图像，扫描置于该宠物上的二维码、条形码或者其他可识别标识，获取宠物标识为 ID1，之后，基于该宠物标识 ID1，从如下表 4 所示的宠物标识与身份标识之间的对应关系中，获取 ID1 对应的身份标识为豆豆，则服务器可以将豆豆确定为该视频图像包括的对象所属的关键词。

表 4

宠物标识	身份标识
ID1	豆豆
ID2	欢欢
... ..	

需要说明的是，在本公开实施例中，仅以上述表 4 所示的宠物标识与身份标识之间的对应关系为例进行说明，上述表 4 并不对本公开实施例构成限定。

进一步地，服务器基于该人脸特征，从存储的人脸特征与身份标识之间的对应关系中，获取对应的身份标识之前，服务器可以接收终端发送的第一设置信息，该第一设置信息中携带该对象的身份标识以及该对象的人脸图像，服务器对该人脸图像进行特征提取，得到该对象的人脸特征，并将该人脸特征与该身份标识存储在人脸特征与身份标识之间的对应关系中。

例如，终端发送的第一设置信息中携带身份标识和人脸图像，且该身份标识为杨乐乐，服务器对该人脸图像进行特征提取，得到人脸特征为 A，则服务器可以将 A 和杨乐乐存储在如表 3 所示的人脸特征与身份标识之间的对应关系中。

进一步地，服务器基于该宠物标识，从存储的宠物标识与身份标识之间的对应关系中，获取对应的身份标识之前，服务器可以接收终端发送的第二设置信息，该第二设置信息中携带该对象的身份标识以及该对象的宠物标识，服务器该宠物标识与该身份标识存储在宠物标识与身份标识之间的对应关系中。

例如，终端发送的第二设置信息中携带身份标识和宠物标识，且该身份标识为豆豆，宠物标识为 ID1，则服务器可以将 ID1 和豆豆存储在如表 4

所示的宠物标识与身份标识之间的对应关系中。

(4) 服务器基于该关键词和该监控视频，生成索引库。

由于索引库中既可以存储关键词与监控时间点之间的对应关系，也可以存储关键词与视频图像之间的对应关系，因此，服务器基于该关键词和该监控视频，生成索引库的操作可以为：当索引库中存储关键词与监控时间点之间的对应关系时，服务器从该监控视频中，确定该视频图像所在的监控时间点，并将该关键词和该监控时间点存储在索引库包括的关键词与监控时间点之间的对应关系中。当索引库中存储关键词与视频图像之间的对应关系时，服务器将该关键词和该视频图像存储在索引库包括的关键词与视频图像之间的对应关系中。

由于监控视频中包括该监控视频中每帧视频图像对应的监控时间点，因此，当服务器确定了存在某个对象的视频图像时，可以基于该视频图像，从该视频图像所属的监控视频中获取该视频图像对应的监控时间点，之后，服务器可以将该对象所属的关键词和该监控时间点存储在索引库包括的关键词与监控时间点之间的对应关系中。

例如，该对象所属的关键词为杨乐乐，服务器从该监控视频中，确定存在杨乐乐的视频图像所在的监控时间点为 2015/08/03-09:05:31，则服务器可以将杨乐乐和 2015/08/03-09:05:31 存储在如表 1 所示的关键词与监控时间点之间的对应关系中。

再例如，该对象所属的关键词为杨乐乐，该监控视频中存在杨乐乐的视频图像为 3. JPEG，则服务器可以将杨乐乐和 3. JPEG 存储在如表 2 所示的关键词与视频图像之间的对应关系中。

第二种方式：当目标对象信息包括目标对象所在的目标图像时，服务器基于指定分类模型和目标图像，确定目标对象所属的目标类别，并基于目标类别，确定目标对象所属的目标关键词，进而基于目标关键词，确定

目标对象在监控视频中所处的视频段。

其中，服务器基于指定分类模型和目标图像，确定目标对象所属的目标类别时，服务器可以将目标图像的尺寸处理为预设尺寸，并基于指定分类模型和处理后的目标图像，确定该目标图像包括的目标对象所属的目标类别。

需要说明的是，服务器基于指定分类模型和处理后的目标图像，确定该目标对象所属目标类别的过程可以参考相关技术，本公开实施例在此不进行详细阐述。

需要说明的是，服务器基于目标类别，确定目标对象所属的目标关键词的过程与步骤 302 第一种方式中步骤（3）的确定过程类似，本公开实施例对此不再赘述。

还需要说明的是，服务器基于目标关键词，确定目标对象在监控视频中所处的视频段的过程与步骤 302 第一种方式的确定过程相同，本公开实施例对此不再赘述。

在步骤 303 中，服务器将目标对象在监控视频中所处的视频段发送给终端，使终端播放该视频段。

需要说明的是，终端播放该视频段时，不仅可以通过终端中设置的播放模块来播放，当然，终端还可以通过自身安装的播放应用程序来播放，本公开实施例对此不做具体限定。

在本公开实施例中，服务器获取监控视频，并基于指定分类模型确定该监控视频的视频图像中包括的对象所属的对象类别，进而基于该对象类别，确定该对象所属的关键词，并将该关键词和该关键词对应的监控时间点或者将该关键词和该关键词对应的视频图像存储在索引库中，之后，服务器在接收到播放请求时，基于该播放请求携带的目标对象信息，确定目标对象所属的目标关键词，并基于该目标关键词和存储的索引库，获取目

标对象在该监控视频中所处的至少一帧视频图像，并将该至少一帧视频图像组成目标对象在该监控视频中所处的视频段，进而将该视频段发送给终端，使终端可以直接播放该监控视频中存在目标对象的视频段，无需播放该监控视频中除该目标对象之外的其他对象所在的视频段，进而避免用户手动对该监控视频的播放进行调整来观看目标对象所处的视频，操作简单，且提高了视频播放效率。

图 4 是根据一示例性实施例示出的一种用于视频播放的装置的框图。

参照图 4，该装置包括接收模块 401，确定模块 402，发送模块 403。

接收模块 401，配置为接收播放请求，播放请求中携带信息，目标对象信息包括目标对象所在的目标图像或者目标对象所属的目标关键词；

确定模块 402，配置为基于目标对象信息，确定目标对象在监控视频中所处的视频段；

发送模块 403，配置为将视频段发送给终端，使终端播放视频段。

在本公开的另一实施例中，参照图 5，该确定模块 402 包括：

第一确定单元 4021，配置为当目标对象信息包括目标对象所在的目标图像时，基于指定分类模型和目标图像，确定目标对象所属的目标类别；

第二确定单元 4022，配置为基于目标类别，确定目标对象所属的目标关键词；

第三确定单元 4023，配置为基于目标关键词，确定目标对象在监控视频中所处的视频段。

在本公开的另一实施例中，参照图 6，该确定模块 402 包括第一获取单元 4024，组成单元 4025。

第一获取单元 4024，配置为基于目标对象信息对应的目标关键词和存储的索引库，获取目标对象在监控视频中所处的至少一帧视频图像；

组成单元 4025，配置为将至少一帧视频图像组成目标对象在监控视频

中所处的视频段。

在本公开的另一实施例中，参照图 7，该第一获取单元 4024 包括第一获取子单元 40241，第二获取子单元 40242。

5 第一获取子单元 40241，配置为当索引库中存储关键词与监控时间点之间的对应关系时，基于目标对象信息对应的目标关键词，从关键词与监控时间点之间的对应关系中，获取至少一个监控时间点；

第二获取子单元 40242，配置为基于至少一个监控时间点，从监控视频中，获取至少一帧视频图像。

在本公开的另一实施例中，该第一获取单元 4024 包括：

10 第三获取子单元，配置为当索引库中存储关键词与视频图像之间的对应关系时，基于目标对象信息对应的目标关键词，从关键词与视频图像之间的对应关系中，获取至少一帧视频图像。

在本公开的另一实施例中，参照图 8，该确定模块 402 还包括第二获取单元 4026，第四确定单元 4027，第五确定单元 4028，生成单元 4029。

15 第二获取单元 4026，配置为接收智能摄像设备发送的监控视频；

第四确定单元 4027，配置为对于监控视频中每帧视频图像，基于指定分类模型，确定视频图像包括的对象所属的对象类别；

第五确定单元 4028，配置为基于对象类别，确定视频图像包括的对象所属的关键词；

20 生成单元 4029，配置为基于关键词和监控视频，生成索引库。

在本公开的另一实施例中，参照图 9，该第五确定单元 4028 包括识别子单元 40281，第四获取子单元 40282，第一确定子单元 40283。

识别子单元 40281，配置为当对象类别为人时，对视频图像包括的对象进行人脸识别，得到人脸特征；

25 第四获取子单元 40282，配置为基于人脸特征，从存储的人脸特征与身

份标识之间的对应关系中，获取对应的身份标识；

第一确定子单元 40283，配置为将身份标识确定为视频图像包括的对象所属的关键词。

在本公开的另一实施例中，参照图 10，该生成单元 4029 包括第二确定子单元 40291，第一存储子单元 40292。

第二确定子单元 40291，配置为从监控视频中，确定视频图像所在的监控时间点；

第一存储子单元 40292，配置为将关键词和监控时间点存储在索引库包括的关键词与监控时间点之间的对应关系中。

在本公开的另一实施例中，该生成单元 4029 包括：

第二存储子单元，配置为将关键词和视频图像存储在索引库包括的关键词与视频图像之间的对应关系中。

在本公开实施例中，服务器接收播放请求，该播放请求中携带目标对象信息，服务器基于目标对象信息确定目标对象在监控视频中所处的视频段，并将该视频段发送给终端，使终端可以直接播放该监控视频中存在目标对象的视频段，无需播放该监控视频中除该目标对象之外的其他对象所在的视频段，进而避免用户手动对该监控视频的播放进行调整来观看目标对象所处的视频，操作简单，且提高了视频播放效率。

关于上述实施例中的装置，其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述，此处将不做详细阐述说明。

图 11 是根据一示例性实施例示出的一种用于视频播放的装置 1100 的框图。例如，装置 1100 可以被提供为一服务器。参照图 11，装置 1100 包括处理组件 1122，其进一步包括一个或多个处理器，以及由存储器 1132 所代表的存储器资源，配置为存储可由处理组件 1122 的执行的指令，例如应用程序。存储器 1132 中存储的应用程序可以包括一个或一个以上的每一个

对应于一组指令的模块。

装置 1100 还可以包括一个电源组件 1126 被配置为执行装置 1100 的电源管理，一个有线或无线网络接口 1150 被配置为将装置 1100 连接到网络，和一个输入输出（I/O）接口 1158。装置 1100 可以操作基于存储在存储器 5 1132 的操作系统，例如 Windows Server™，Mac OS X™，Unix™，Linux™，FreeBSD™ 或类似。

此外，处理组件 1122 被配置为执行指令，以执行下述视频播放方法，所述方法包括：

接收播放请求，播放请求中携带目标对象信息，目标对象信息包括目标对象所在的目标图像或者目标对象所属的目标关键词。 10

基于目标对象信息，确定目标对象在监控视频中所处的视频段。

将目标对象在监控视频中所处的视频段发送给终端，使终端播放该视频段。

在本公开的另一实施例中，基于目标对象信息，确定目标对象在监控 15 视频中所处的视频段，包括：

当目标对象信息包括目标对象所在的目标图像时，基于指定分类模型和目标图像，确定目标对象所属的目标类别；

基于目标类别，确定目标对象所属的目标关键词；

基于目标关键词，确定目标对象在监控视频中所处的视频段。

20 在本公开的另一实施例中，基于目标对象信息，确定目标对象在监控视频中所处的视频段，包括：

基于目标对象信息对应的目标关键词和存储的索引库，获取目标对象在监控视频中所处的至少一帧视频图像；

将至少一帧视频图像组成目标对象在监控视频中所处的视频段。

25 在本公开的另一实施例中，基于目标对象信息对应的目标关键词和存

储的索引库，获取目标对象在监控视频中所处的至少一帧视频图像，包括：

当索引库中存储关键词与监控时间点之间的对应关系时，基于目标对象信息对应的目标关键词，从关键词与监控时间点之间的对应关系中，获取至少一个监控时间点；

5 基于至少一个监控时间点，从监控视频中，获取至少一帧视频图像。

在本公开的另一实施例中，基于目标对象信息对应的目标关键词和存储的索引库，获取目标对象在监控视频中所处的至少一帧视频图像，包括：

当索引库中存储关键词与视频图像之间的对应关系时，基于目标对象信息对应的目标关键词，从关键词与视频图像之间的对应关系中，获取至少一帧视频图像。

在本公开的另一实施例中，基于目标对象信息对应的目标关键词和存储的索引库，获取目标对象在监控视频中的至少一帧视频图像之前，还包括：

获取监控视频；

15 对于监控视频中每帧视频图像，基于指定分类模型，确定视频图像包括的对象所属的对象类别；

基于对象类别，确定视频图像包括的对象所属的关键词；

基于关键词和监控视频，生成索引库。

在本公开的另一实施例中，基于对象类别，确定视频图像包括的对象所属的关键词，包括：

当对象类别为人时，对视频图像包括的对象进行人脸识别，得到人脸特征；

基于人脸特征，从存储的人脸特征与身份标识之间的对应关系中，获取对应的身份标识；

25 将身份标识确定为视频图像包括的对象所属的关键词。

在本公开的另一实施例中，基于关键词和监控视频，生成索引库，包括：

从监控视频中，确定视频图像所在的监控时间点；

将关键词和监控时间点存储在索引库包括的关键词与监控时间点之间

5 的对应关系中。

在本公开的另一实施例中，基于关键词和监控视频，生成索引库，包括：

将关键词和视频图像存储在索引库包括的关键词与视频图像之间的对应关系中。

10 在本公开实施例中，服务器接收播放请求，该播放请求中携带目标对象信息，服务器基于目标对象信息确定目标对象在监控视频中所处的视频段，并将该视频段发送给终端，使终端可以直接播放该监控视频中存在目标对象的视频段，无需播放该监控视频中除该目标对象之外的其他对象所在的视频段，进而避免用户手动对该监控视频的播放进行调整来观看目标
15 对象所处的视频，操作简单，且提高了视频播放效率。

本领域技术人员在考虑说明书及实践这里公开的发明后，将容易想到本发明的其它实施方案。本申请旨在涵盖本发明的任何变型、用途或者适应性变化，这些变型、用途或者适应性变化遵循本发明的一般性原理并包括本公开实施例未公开的本技术领域中的公知常识或惯用技术手段。说明书
20 书和实施例仅被视为示例性的，本发明的真正范围和精神由下面的权利要求指出。

应当理解的是，本发明并不局限于上面已经描述并在附图中示出的精确结构，并且可以在不脱离其范围进行各种修改和改变。本发明的范围仅由所附的权利要求来限制。

工业实用性

在本公开实施例中，服务器接收播放请求，该播放请求中携带目标对象信息，服务器基于目标对象信息确定目标对象在监控视频中所处的视频段，并将该视频段发送给终端，使终端可以直接播放该监控视频中存在目标对象的视频段，无需播放该监控视频中除该目标对象之外的其他对象所在的视频段，进而避免用户手动对该监控视频的播放进行调整来观看目标对象所处的视频，操作简单，且提高了视频播放效率。

权利要求书

1、一种用于视频播放的方法，所述方法包括：

接收播放请求，所述播放请求中携带目标对象信息，所述目标对象信息包括目标对象所在的目标图像或者所述目标对象所属的目标关键词；

基于所述目标对象信息，确定所述目标对象在监控视频中所处的视频段；

将所述视频段发送给终端，使所述终端播放所述视频段。

2、如权利要求 1 所述的方法，其中，所述基于所述目标对象信息，确定所述目标对象在监控视频中所处的视频段，包括：

当所述目标对象信息包括目标对象所在的目标图像时，基于指定分类模型和所述目标图像，确定所述目标对象所属的目标类别；

基于所述目标类别，确定所述目标对象所属的目标关键词；

基于所述目标关键词，确定所述目标对象在所述监控视频中所处的视频段。

3、如权利要求 1 所述的方法，其中，所述基于所述目标对象信息，确定所述目标对象在监控视频中所处的视频段，包括：

基于所述目标对象信息对应的所述目标关键词和存储的索引库，获取所述目标对象在所述监控视频中所处的至少一帧视频图像；

将所述至少一帧视频图像组成所述目标对象在所述监控视频中所处的视频段。

4、如权利要求 3 所述的方法，其中，所述基于所述目标对象信息对应的所述目标关键词和存储的索引库，获取所述目标对象在所述监控视频中所处的至少一帧视频图像，包括：

当所述索引库中存储关键词与监控时间点之间的对应关系时，基于

所述目标对象信息对应的所述目标关键词，从所述关键词与监控时间点之间的对应关系中，获取至少一个监控时间点；

基于所述至少一个监控时间点，从所述监控视频中，获取至少一帧视频图像。

- 5 5、如权利要求3所述的方法，其中，所述基于所述目标对象信息对应的所述目标关键词和存储的索引库，获取所述目标对象在所述监控视频中所处的至少一帧视频图像，包括：

10 当所述索引库中存储关键词与视频图像之间的对应关系时，基于所述目标对象信息对应的所述目标关键词，从所述关键词与视频图像之间的对应关系中，获取至少一帧频图像。

6、如权利要求3-5任一权利要求所述的方法，其中，所述基于所述目标对象信息对应的所述目标关键词和存储的索引库，获取所述目标对象在所述监控视频中的至少一帧视频图像之前，还包括：

获取监控视频；

- 15 对于所述监控视频中每帧视频图像，基于指定分类模型，确定所述视频图像包括的对象所属的对象类别；

基于所述对象类别，确定所述视频图像包括的对象所属的关键词；

基于所述关键词和所述监控视频，生成所述索引库。

- 20 7、如权利要求6所述的方法，其中，所述基于所述对象类别，确定所述视频图像包括的对象所属的关键词，包括：

当所述对象类别为人时，对所述视频图像包括的对象进行人脸识别，得到人脸特征；

基于所述人脸特征，从存储的人脸特征与身份标识之间的对应关系中，获取对应的身份标识；

- 25 将所述身份标识确定为所述视频图像包括的对象所属的关键词。

8、如权利要求 6 所述的方法，其中，所述基于所述关键词和所述监控视频，生成所述索引库，包括：

从所述监控视频中，确定所述视频图像所在的监控时间点；

5 将所述关键词和所述监控时间点存储在所述索引库包括的关键词与监控时间点之间的对应关系中。

9、如权利要求 6 所述的方法，其中，所述基于所述关键词和所述监控视频，生成所述索引库，包括：

将所述关键词和所述视频图像存储在所述索引库包括的关键词与视频图像之间的对应关系中。

10 10 一种用于视频播放的装置，所述装置包括：

接收模块，配置为接收播放请求，所述播放请求中携带目标对象信息，所述目标对象信息包括目标对象所在的目标图像或者所述目标对象所属的目标关键词；

15 确定模块，配置为基于所述目标对象信息，确定所述目标对象在监控视频中所处的视频段；

发送模块，配置为将所述视频段发送给终端，使所述终端播放所述视频段。

11、如权利要求 10 所述的装置，其中，所述确定模块包括：

20 第一确定单元，配置为当所述目标对象信息包括目标对象所在的目标图像时，基于指定分类模型和所述目标图像，确定所述目标对象所属的目标类别；

第二确定单元，配置为基于所述目标类别，确定所述目标对象所属的目标关键词；

25 第三确定单元，配置为基于所述目标关键词，确定所述目标对象在所述监控视频中所处的视频段。

12、如权利要求 10 所述的装置，其中，所述确定模块包括：

第一获取单元，配置为基于所述目标对象信息对应的所述目标关键词和存储的索引库，获取所述目标对象在所述监控视频中所处的至少一帧视频图像；

5 组成单元，配置为将所述至少一帧视频图像组成所述目标对象在所述监控视频中所处的视频段。

13、如权利要求 12 所述的装置，其中，所述第一获取单元包括：

10 第一获取子单元，配置为当所述索引库中存储关键词与监控时间点之间的对应关系时，基于所述目标对象信息对应的所述目标关键词，从所述关键词与监控时间点之间的对应关系中，获取至少一个监控时间点；

第二获取子单元，配置为基于所述至少一个监控时间点，从所述监控视频中，获取至少一帧视频图像。

14、如权利要求 12 所述的装置，其中，所述第一获取单元包括：

15 第三获取子单元，配置为当所述索引库中存储关键词与视频图像之间的对应关系时，基于所述目标对象信息对应的所述目标关键词，从所述关键词与视频图像之间的对应关系中，获取至少一帧视频图像。

15、如权利要求 12-14 任一权利要求所述的装置，其中，所述确定模块还包括：

第二获取单元，配置为获取监控视频；

20 第四确定单元，配置为对于所述监控视频中每帧视频图像，基于指定分类模型，确定所述视频图像包括的对象所属的对象类别；

第五确定单元，配置为基于所述对象类别，确定所述视频图像包括的对象所属的关键词；

25 生成单元，配置为基于所述关键词和所述监控视频，生成所述索引库。

16、如权利要求 15 所述的装置，其中，所述第五确定单元包括：

识别子单元，配置为当所述对象类别为人时，对所述视频图像包括的对象进行人脸识别，得到人脸特征；

5 第四获取子单元，配置为基于所述人脸特征，从存储的人脸特征与身份标识之间的对应关系中，获取对应的身份标识；

第一确定子单元，配置为将所述身份标识确定为所述视频图像包括的对象所属的关键词。

17、如权利要求 15 所述的装置，其中，所述生成单元包括：

10 第二确定子单元，配置为从所述监控视频中，确定所述视频图像所在的监控时间点；

第一存储子单元，配置为将所述关键词和所述监控时间点存储在所述索引库包括的关键词与监控时间点之间的对应关系中。

18、如权利要求 15 所述的装置，其中，所述生成单元包括：

15 第二存储子单元，配置为将所述关键词和所述视频图像存储在所述索引库包括的关键词与视频图像之间的对应关系中。

19、一种用于视频播放的装置，所述装置包括：

处理器；

配置为存储处理器可执行指令的存储器；

其中，所述处理器被配置为：

20 接收播放请求，所述播放请求中携带目标对象信息，所述目标对象信息包括目标对象所在的目标图像或者所述目标对象所属的目标关键词所属的目标关键词；

基于所述目标对象信息，确定所述目标对象在监控视频中所处的视频段；

25 将所述视频段发送给终端，使所述终端播放所述视频段。

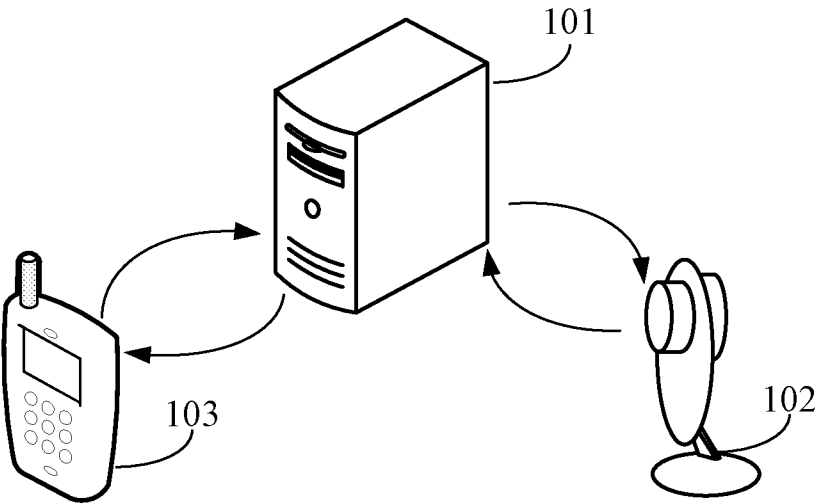


图 1

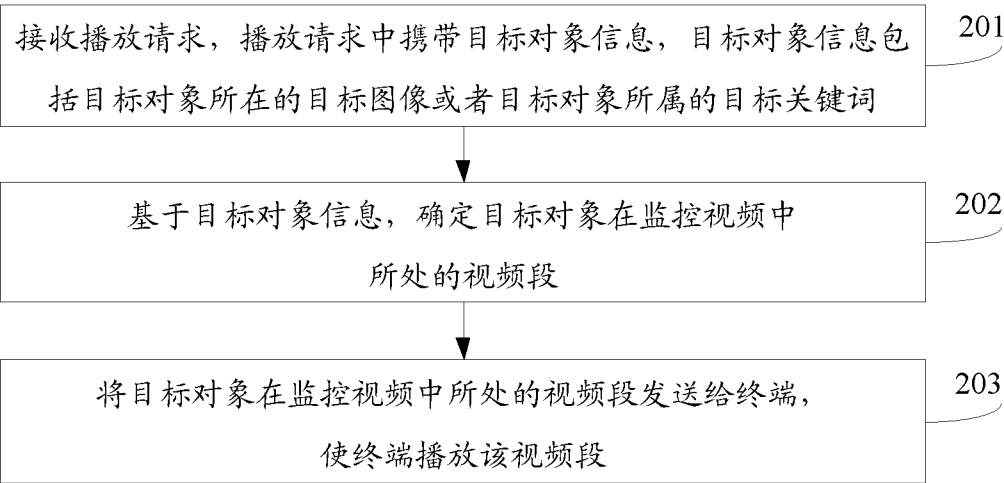


图 2

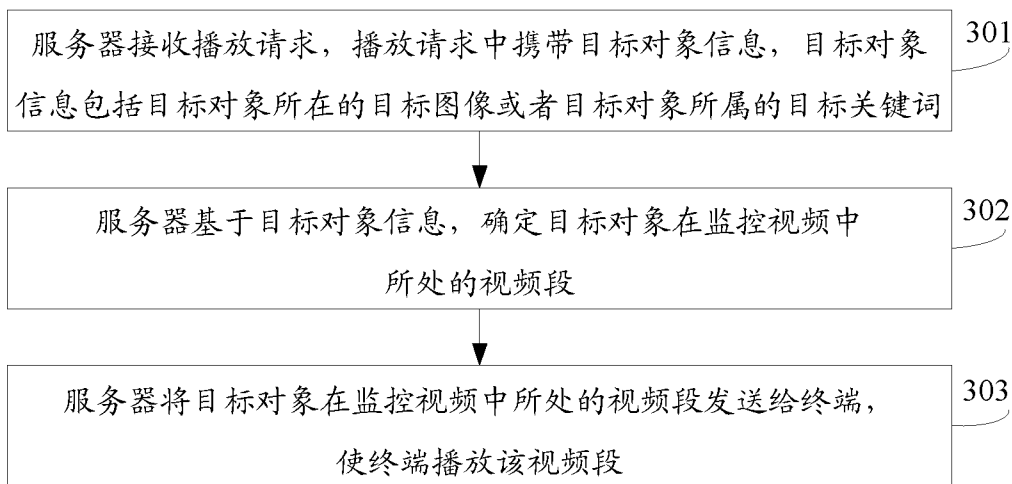


图 3

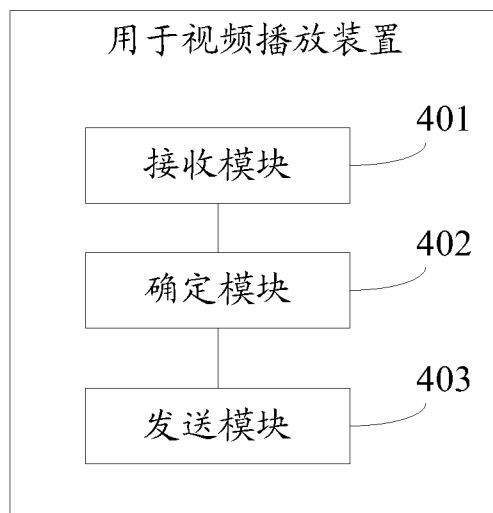


图 4



图 5

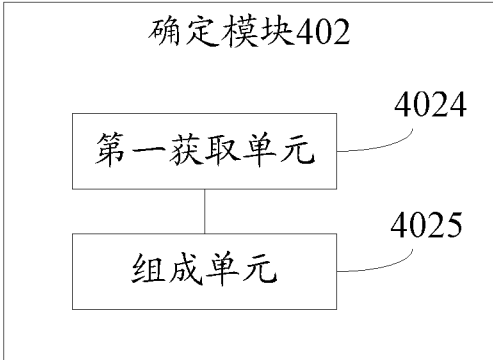


图 6

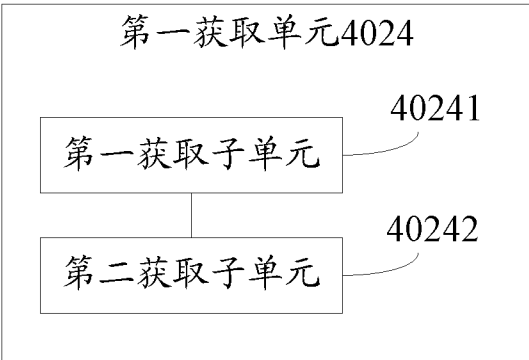


图 7

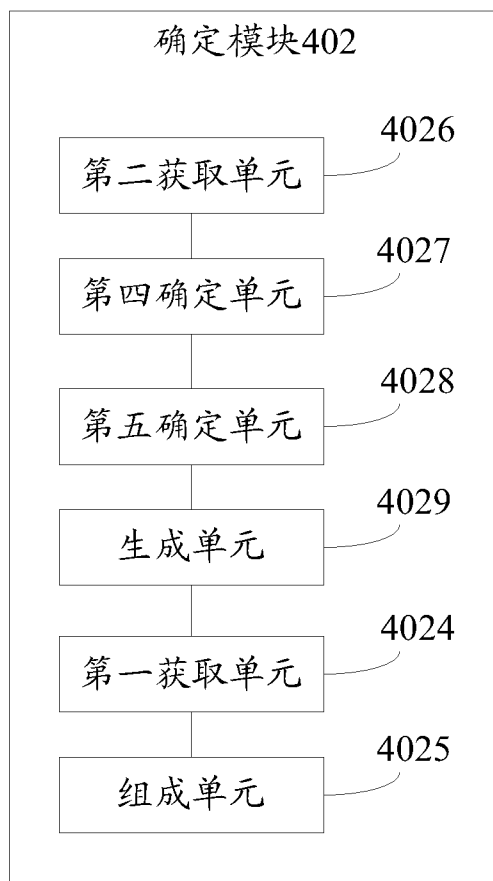


图 8

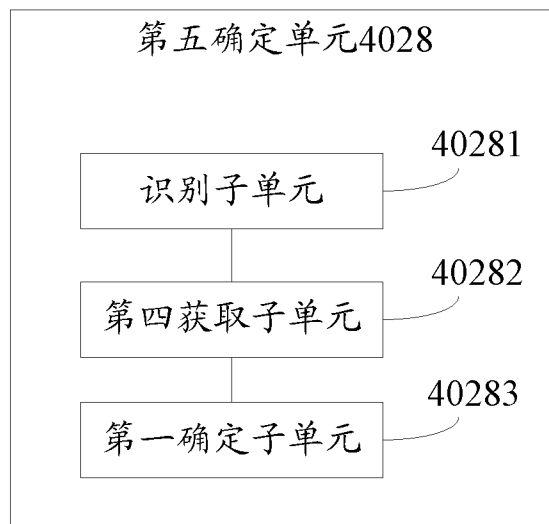


图 9

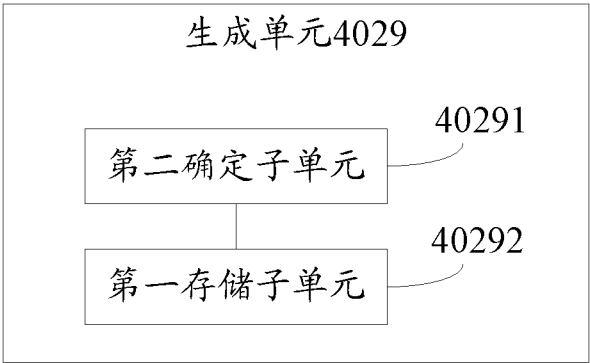


图 10

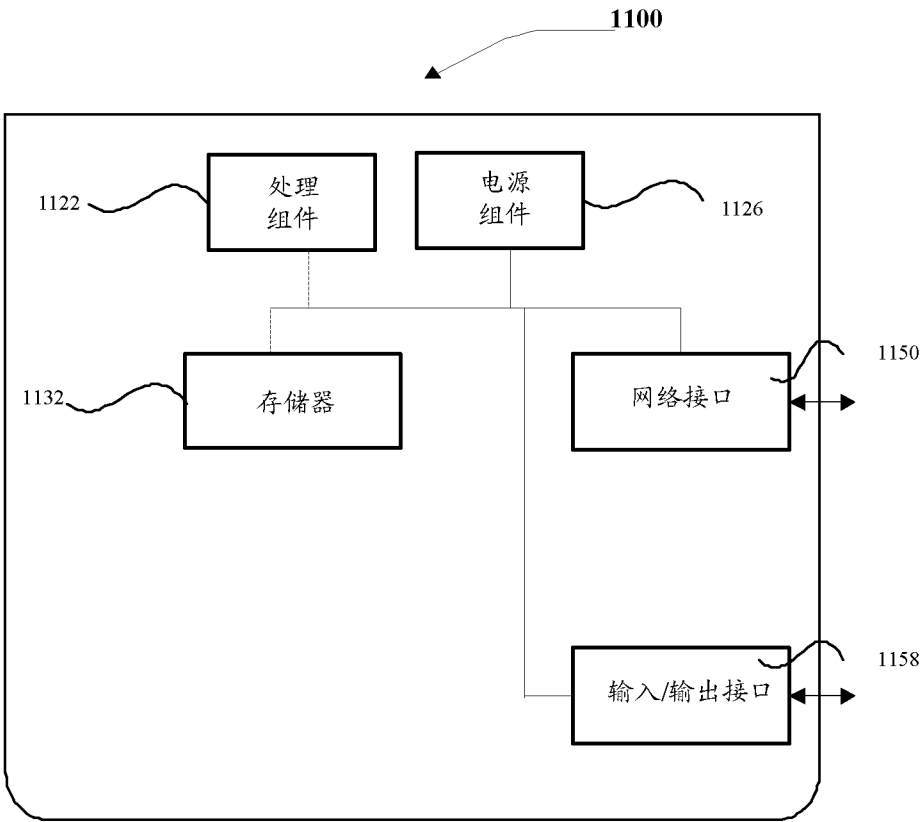


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/099593

A. CLASSIFICATION OF SUBJECT MATTER

H04N 7/18 (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)

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, CNKI, VEN: determine, save, video, play, target, object, choose, monitor, search, database, memory, storage

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 102129474 A (H3C TECHNOLOGIES CO., LIMITED), 20 July 2011 (20.07.2011), description, paragraphs [0117]-[0150], and figures 5-8	1-19
Y	CN 101299812 A (BEIJING VIMICRO CORP.), 05 November 2008 (05.11.2008), description, page 6, line 1 to page 11, line 9	1-19
Y	CN 104754267 A (XIAOMI TECHNOLOGY CO., LTD.), 01 July 2015 (01.07.2015), description, paragraphs [0055]-[0150]	2-9, 11-18
PX	CN 105357475 A (XIAOMI TECHNOLOGY CO., LTD.), 24 February 2016 (24.02.2016), the whole document	1-19

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 08 July 2016 (08.07.2016)	Date of mailing of the international search report 19 July 2016 (19.07.2016)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer MA, Guili Telephone No.: (86-10) 62411464

International application No.
PCT/CN2015/099593

Form PCT/ISA/210 (patent family annex) (July 2009)

国际检索报告

国际申请号

PCT/CN2015/099593

A. 主题的分类 H04N 7/18(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) 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, CNKI, VEN:视频, 播放, 目标, 对象, 确定, 选择, 监控, 检索, 搜索, 数据库, 存储, 保存, video, play, target, object, choose, monitor, search, database, memory, storage		
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	CN 102129474 A (杭州华三通信技术有限公司) 2011年 7月 20日 (2011 - 07 - 20) 说明书第[0117]段-第[0150]段及图5-8	1-19
Y	CN 101299812 A (北京中星微电子有限公司) 2008年 11月 5日 (2008 - 11 - 05) 说明书第6页第1行-第11页第9行	1-19
Y	CN 104754267 A (小米科技有限责任公司) 2015年 7月 1日 (2015 - 07 - 01) 说明书第[0055]段-第[0150]段	2-9、11-18
PX	CN 105357475 A (小米科技有限责任公司) 2016年 2月 24日 (2016 - 02 - 24) 全文	1-19
☐ 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 2016年 7月 8日		国际检索报告邮寄日期 2016年 7月 19日
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(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 马桂丽 电话号码 (86-10)62411464

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2015/099593

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	102129474	A	2011年 7月 20日	CN 102129474 B	2015年 2月 11日
CN	101299812	A	2008年 11月 5日	CN 101299812 B	2012年 12月 5日
CN	104754267	A	2015年 7月 1日	无	
CN	105357475	A	2016年 2月 24日	无	