

(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 11 月 5 日 (05.11.2020)



(10) 国际公布号
WO 2020/220968 A1

(51) 国际专利分类号:
H04N 21/431 (2011.01) *H04N 21/488* (2011.01)
H04N 21/44 (2011.01)

(21) 国际申请号: PCT/CN2020/084112

(22) 国际申请日: 2020 年 4 月 10 日 (10.04.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910358569.8 2019年4月30日 (30.04.2019) 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: 郑远力 (ZHENG, Yuanli); 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)。殷泽龙 (YIN, Zelong); 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)。谢年华 (XIE, Nianhua); 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)。

(54) Title: VIDEO DATA PROCESSING METHOD AND RELATED DEVICE

(54) 发明名称: 一种视频数据处理方法和相关装置

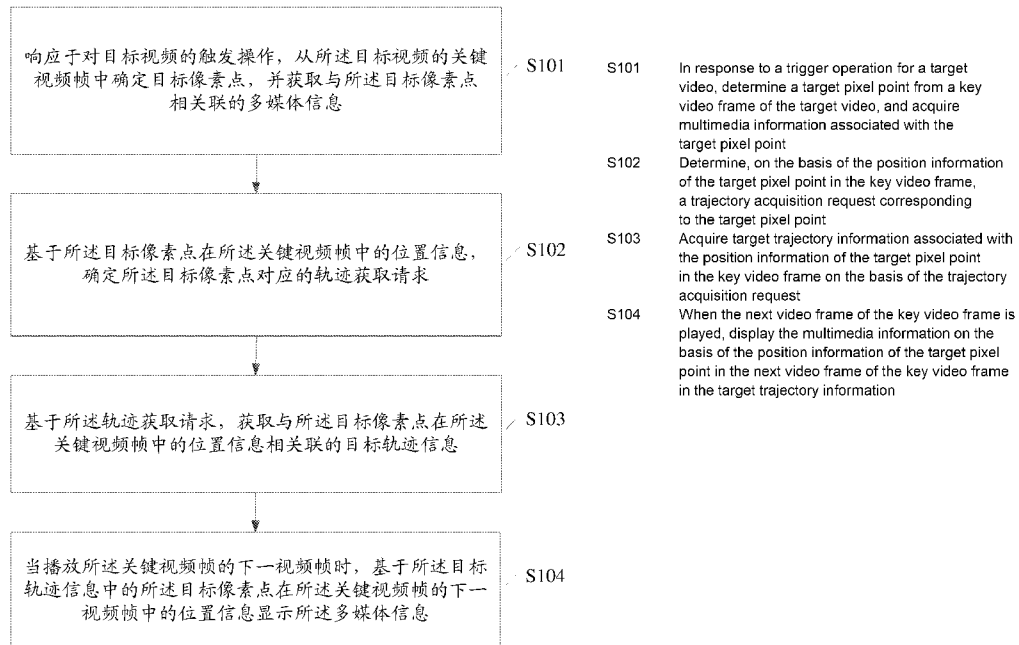


图 4

(57) Abstract: Disclosed in the embodiments of the present application are a video data processing method and device, and a storage medium. The method comprises: in response to a trigger operation for a target video, determining a target object from a key video frame of the target video, and acquiring multimedia information associated with the target object; determining, on the basis of the position information of a target pixel point of the target object in the key video frame, a trajectory acquisition request corresponding to the target pixel point; acquiring target trajectory information associated with the position information of the target pixel point in the key video

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, 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, 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))。

frame on the basis of the trajectory acquisition request; and when the next video frame of the key video frame is played, displaying the multimedia information on the basis of the position information of the target pixel point in the next video frame of the key video frame in the target trajectory information.

(57) 摘要: 本申请实施例公开了一种视频数据处理方法、装置以及存储介质, 该方法包括: 响应针对目标视频的触发操作, 从所述目标视频的关键视频帧中确定目标对象, 并获取与所述目标对象相关联的多媒体信息; 基于所述目标对象中的目标像素点在所述关键视频帧中的位置信息, 确定所述目标像素点对应的轨迹获取请求; 基于所述轨迹获取请求获取与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息; 当播放所述关键视频帧的下一视频帧时, 基于所述目标轨迹信息中的所述目标像素点在所述关键视频帧的下一视频帧中的位置信息显示所述多媒体信息。

一种视频数据处理方法和相关装置

本申请要求于 2019 年 04 月 30 日提交的申请号为 201910358569.8、发明名称为“一种视频数据处理方法和相关装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及互联网技术领域，尤其涉及一种视频数据处理方法和相关装置。

背景技术

用户在通过用户终端观看网络视频的过程中，可以在视频播放界面上看到该用户或其它用户发表的用户文字或用户评论。在现有的用户文字显示方式中，输出至视频播放界面中的用户文字，通常是通过该视频播放界面中固定的文字显示轨道进行输出和显示。

发明内容

本申请实施例提供一种视频数据处理方法和相关装置。

本申请实施例一方面提供了一种视频数据处理方法，所述方法应用于计算机设备，包括：

响应于对目标视频的触发操作，从所述目标视频的关键视频帧中确定目标像素点，并获取与所述目标像素点相关联的多媒体信息，其中，所述关键视频帧是所述触发操作所在的视频帧，所述目标像素点是所述关键视频帧中与所述触发操作对应的像素点；

基于所述目标像素点在所述关键视频帧中的位置信息，确定所述目标像素点对应的轨迹获取请求；

基于所述轨迹获取请求，获取与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，其中，所述目标轨迹信息包含所述目标像素点在所述关键视频帧的下一视频帧中的位置信息，所述目标像素点在所述关键视频帧的下一视频帧中的位置信息是通过跟踪所述目标像素点得到的；

当播放所述关键视频帧的下一视频帧时，基于所述目标轨迹信息中的所述目标像素点在所述关键视频帧的下一视频帧中的位置信息，显示所述多媒体信息。

本申请实施例一方面提供了一种视频数据处理方法，所述方法应用于业务服务器，包括：

响应针对关键视频帧中的目标像素点的轨迹获取请求，获取与目标视频相关联的轨迹信息，其中，所述关键视频帧是所述目标视频中的视频帧，所述目标像素点是所述关键视频帧中的像素点，所述轨迹信息是由所述目标视频的每个视频帧中的像素点的位置信息所确定的；

从所述目标视频相关联的轨迹信息中，筛选与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，并返回所述目标轨迹信息，其中，所述目标轨迹信息包含目标位置信息，所述目标位置信息用于触发在所述关键视频帧的下一视频帧中，显示与所述目标像素点相关联的多媒体信息。

本申请实施例一方面提供了一种视频数据处理方法，所述方法包括：

从目标视频中获取相邻的第一视频帧和第二视频帧；

基于所述目标视频对应的光流追踪规则、所述第一视频帧中的像素点、所述第二视频帧中的像素点，确定所述第一视频帧对应的平均位移矩阵；

基于所述平均位移矩阵，对所述第一视频帧中的像素点的位置信息进行跟踪，并在所述第二视频帧中确定所跟踪得到的像素点的位置信息；

基于所述第一视频帧中的像素点的位置信息、所述跟踪得到的像素点在所述第二视频帧中的位置信息，生成与所述目标视频相关联的轨迹信息，其中，所述轨迹信息中包含用于对目标视频中的目标像素点所关联的多媒体信息进行跟踪显示的目标轨迹信息。

本申请实施例一方面提供了一种视频数据处理装置，所述装置应用于计算机设备，包括：

对象确定模块，用于响应于对目标视频的触发操作，从所述目标视频的关键视频帧中确定目标像素点，并获取与所述目标像素点相关联的多媒体信息，其中，所述关键视频帧是所述触发操作所在的视频帧，所述目标像素点是所述关键视频帧中与所述触发操作对应的像素点；

请求确定模块，用于基于所述目标像素点在所述关键视频帧中的位置信息，

确定所述目标像素点对应的轨迹获取请求；

轨迹获取模块，用于基于所述轨迹获取请求，获取与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，其中，所述目标轨迹信息包含所述目标像素点在所述关键视频帧的下一视频帧中的位置信息，所述目标像素点在所述关键视频帧的下一视频帧中的位置信息是通过跟踪所述目标像素点得到的；

文本显示模块，用于当播放所述关键视频帧的下一视频帧时，基于所述目标轨迹信息中的所述目标像素点在所述关键视频帧的下一视频帧中的位置信息显示所述多媒体信息。

本申请实施例一方面提供了一种视频数据处理装置，所述装置应用于业务服务器，包括：

请求响应模块，用于响应于对关键视频帧中的目标像素点的轨迹获取请求，获取与目标视频相关联的轨迹信息，其中，所述关键视频帧是所述目标视频中的视频帧，所述目标像素点是所述关键视频帧中的像素点，所述轨迹信息是由所述目标视频的每个视频帧中的像素点的位置信息所确定的；

轨迹筛选模块，用于从所述目标视频相关联的轨迹信息中，筛选与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，并返回所述目标轨迹信息，其中，所述目标轨迹信息包含目标位置信息，所述目标位置信息用于触发在所述关键视频帧的下一视频帧中，显示与所述目标像素点相关联的多媒体信息。

本申请实施例一方面提供了一种视频数据处理装置，所述装置包括：

第一获取模块，用于从目标视频中获取相邻的第一视频帧和第二视频帧；

矩阵获取模块，用于基于所述目标视频对应的光流追踪规则、所述第一视频帧中的像素点、所述第二视频帧中的像素点，确定所述第一视频帧对应的平均位移矩阵；

位置跟踪模块，用于基于所述平均位移矩阵对所述第一视频帧中的像素点的位置信息进行跟踪，并在所述第二视频帧中确定所跟踪得到的像素点的位置信息；

轨迹生成模块，用于基于所述第一视频帧中的像素点的位置信息、所述跟踪得到的像素点在所述第二视频帧中的位置信息，生成与所述目标视频相关联

的轨迹信息，其中，所述轨迹信息中包含用于对目标视频中的目标像素点所关联的多媒体信息进行跟踪显示的目标轨迹信息。

本申请实施例一方面提供了一种计算机设备，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，以执行如本申请实施例中一方面中的方法。

本申请实施例一方面提供了一种计算机存储介质，所述计算机存储介质存储有计算机程序，所述计算机程序包括程序指令，所述程序指令当被处理器执行时，执行如本申请实施例中一方面中的方法。

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目标视频中的多个视频帧的示意图；

图 3 是本申请实施例提供的一种获取目标视频的场景示意图；

图 4 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 5 是本申请实施例提供的一种获取多媒体信息的示意图；

图 6 是本申请实施例提供的一种全图像素追踪的示意图；

图 7 是本申请实施例提供的一种在连续多个视频帧中跟踪弹幕数据的示意图；

图 8 是本申请实施例提供的另一种视频数据处理方法的示意图；

图 9 是本申请实施例提供的一种确定有效像素点的方法；

图 10 是本申请实施例提供的一种基于轨迹信息显示弹幕数据的示意图；

图 11 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 12 是本申请实施例提供的一种计算机设备的结构示意图；

图 13 是本申请实施例提供的另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另一种计算机设备的结构示意图；

图 15 是本申请实施例提供的又一种视频数据处理装置的结构示意图；

图 16 是本申请实施例提供的又一种计算机设备的结构示意图。

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。

在现有的网络视频的播放过程中，视频播放界面中显示的用户文字是独立于视频播放界面上所播放的视频内容，以至于所显示的用户文字与视频内容之间缺乏一定的相关性。此外，当用户需要发送用户文字时，用户终端是将获取到的用户文字，通过预定的文字显示轨道进行输出。因此，对于每个用户所发送的用户文字而言，均是通过相同的文字显示轨道进行输出，从而不能有针对性地对视频内容进行评论。

请参见图 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，所述网络架构可以包括业务服务器 2000（或应用服务器 2000）和用户终端集群。所述业务服务器 2000 可以是由大量服务器组成的服务器集群，例如，云端服务器，或简称为云端。所述用户终端集群可以包括多个用户终端，如图 1 所示，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。如图 1 所示，用户终端 3000a、用户终端 3000b、用户终端 3000c、...、用户终端 3000n 可以分别与所述业务服务器 2000 进行网络连接，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。

如图 1 所示，用户终端集群中的每个用户终端均可以集成安装有目标应用，当该目标应用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，目标应用可以包含多媒体应用、社交应用、娱乐应用等具有视频播放功能的应用。为便于理解，本申请实施例以所述多个用户终端中的一个用户终端作为目标用户终端为例，以阐述集成有该目标应用的

目标用户终端通过业务数据展示平台与所述业务服务器 2000 之间实现数据交互的具体过程。其中，本申请实施例中的目标用户终端可以包含个人电脑、平板电脑、笔记本电脑、智能手机等集成有上述目标应用的移动终端。其中，该业务服务器 2000 可以为该目标应用的后台服务器，该后台服务器对应的业务数据库可以用于存储展示在该业务数据展示平台上的每个业务数据信息，该业务数据信息可以包含视频数据等互联网信息。应当理解，该业务数据展示平台上可以显示多个视频，当目标用户在该目标用户终端中通过该业务数据展示平台触发多个视频中的一个视频时，可以获取该视频对应的视频数据，进而可以在该目标用户终端中播放该视频数据，并可以进一步将当前正在该目标用户终端中播放的视频数据称之为目标视频。该目标视频是由该业务服务器 2000 基于目标用户终端所发送的数据加载指令所返回的视频数据。

其中，所述目标视频中可以包含多个视频帧，每个视频帧均可以称之为一个图像数据，且每个视频帧均对应于该目标视频的播放时长中的一个播放时间戳（即一个时刻），以便于后续目标用户终端在加载并播放该目标视频时，可以基于该目标视频中的每个视频帧对应的播放时间戳在播放显示界面中显示相应的视频帧。

其中，业务服务器 2000 可以在视频的预处理阶段，将业务数据库中所存储的视频集中的每个视频进行分帧处理，从而可以将每个视频所包含的多个视频帧拆分成一张张图片。为便于理解，进一步地，请参见图 2，是本申请实施例提供的一种目标视频中的多个视频帧的示意图。其中，该目标视频可以为前述业务数据库中的视频 A，如图 2 所示，该视频 A 可以包含 n 个视频帧， n 为大于 0 的正整数。业务服务器 2000 可以预先将该视频 A 中的 n 个视频帧拆分为 n 个图片。可以将这 n 个图片中的每两个前后相邻的图片称之为一个图像对。比如，如图 2 所示，本申请实施例可以将图 2 所示的第一时刻对应的视频帧与第二时刻对应的视频帧称之为一个图像对，并可以将第二时刻对应的视频帧与第三时刻对应的视频帧称之为一个图像对，...，并可以将第 $n-1$ 时刻的视频帧与第 n 时刻对应的视频帧称之为第一图像对。换言之，对于一个目标视频而言，可以从该目标视频的多个视频帧中确定出多个图像对，每个图像对中均可以包含前后相邻的两个时刻对应的视频帧，即每个图像对中均可以包含两个相邻的视频帧。

为便于理解，本申请实施例以所述多个图像对中的首个图像对为例。在视频预处理阶段，本申请实施例可以将该首个图像对中的一个视频帧（例如，图 2 所示的第一时刻对应的视频帧）称之为第一视频帧，并可以将该图像对中的另一个视频帧（即第二时刻对应的视频帧）称之为第二视频帧，进而可以基于光流追踪规则，对该图像对中第一视频帧中的所有像素点的位置信息进行跟踪，以追踪得到该第一视频帧中的每个像素点出现在第二视频帧中的位置信息。由于每个图像对中均包含两个相邻的视频帧，从而可以计算得到每个图像对中第一视频帧中的像素点出现在下一视频帧中的位置信息，最后可以在该业务服务器 2000 中确定出该视频 A 中的所有像素点在所有视频帧中的运动轨迹，并可以将这些像素点的运动轨迹统称为像素点的轨迹信息。

由于业务服务器 2000 可以预先计算好该视频 A 中的所有像素点的轨迹信息，因此，当目标用户在目标用户终端中播放该视频 A 时，可以将该当前播放的视频 A 称之为一个目标视频。在播放该目标视频的过程中，若目标用户需要跟踪某个对象，则可以在目标用户终端中对需要跟踪的这个对象（即目标对象）执行触发操作，与触发操作对应的像素点称为目标像素点，即该目标像素点是由该目标用户针对当前播放的视频帧中的目标对象所执行的触发操作而确定的，该触发操作可以用于在当前所播放的视频帧中选择需要进行跟踪的目标对象。其中，本申请实施例可以将该触发操作对应的视频帧称之为关键视频帧，换言之，本申请实施例可以将当前包含该目标像素点的视频帧称之为该目标视频中的关键视频帧。可以理解的是，该关键视频帧可以为上述图 2 所对应实施例中的第一时刻对应的视频帧，可选地，该关键视频帧也可以为上述图 2 所对应实施例中第二时刻对应的视频帧，这里将不一一列举。

应当理解，本申请实施例可以将该关键视频帧、该关键视频帧中的目标像素点以及该目标像素点的位置信息给到上述图 1 所对应实施例中的业务服务器 2000，以使该业务服务器 2000 可以基于该目标像素点在该关键视频帧中的位置信息，从预先计算得到的该目标视频中的所有像素点的轨迹信息中，筛选与该目标像素点的位置信息相匹配的轨迹信息，作为目标轨迹信息。其中，所述目标轨迹信息可以包括目标像素点在所述关键视频帧之后的视频帧中的位置坐标。进而，业务服务器 2000 可以将该目标轨迹信息返回给该目标用户终端，以使目标用户终端可以在播放该关键视频帧的下一视频帧时，进一步根据该目标

轨迹信息确定目标像素点在该关键视频帧的下一视频帧中的位置信息，即可以得到该目标像素点的目标位置信息，进而可以基于该目标位置信息显示该目标对象对应的多媒体信息。

其中，进一步地，请参见图 3，是本申请实施例提供的一种获取目标视频的场景示意图。如图 3 所示的目标用户终端可以为上述图 1 所对应实施例中的用户终端 3000a。如图 3 所示，目标用户可以在进入该目标应用之后，在该目标用户终端（例如，智能手机）中显示该目标应用的业务数据展示平台，该业务数据展示平台上可以显示图 3 所示的视频 10a、视频 20a、视频 30a 和视频 40a。当目标用户需要在目标用户终端中播放图 3 所示的视频 30a（该视频 30a 可以为上述图 3 所对应实施例中的视频 A）时，可以对该视频 30a 所在区域执行播放操作（例如，目标用户可以对该视频 30a 执行点击操作），进而可以将该视频 30a 对应的目标标识信息添加到图 3 所示的数据加载指令中，以将该数据加载指令进一步给到与该目标用户终端具有网络连接关系的应用服务器，该应用服务器可以为上述图 1 所对应实施例中的业务服务器 2000。可以理解的是，该应用服务器可以在获取到数据加载指令时，从业务数据库中查找该目标标识信息对应的视频数据，并可以将查找到的视频数据统称为目标数据，以便于能够将该目标数据给到图 3 所示的目标用户终端，以使该目标用户终端可以在图 3 所示的视频播放界面中播放该视频数据，此时，该目标用户终端可以将该目标用户所选取并播放的视频 30a 称之为目标视频，即此时，该目标用户终端可以按照上述图 3 所示的播放时间戳播放视频 A 中的每个视频帧。

其中，上述目标用户终端获取目标对象以及获取目标轨迹信息的具体过程，可以参见如下图 4 至图 7 所对应的实施例所提供的实现方式。另外，上述业务服务器 2000 获取所述像素点在所述第二视频帧中的位置信息。以及筛选所述目标像素点对应的目标轨迹信息的具体过程，可以参见如下图 8 至图 10 所对应的实施例所提供的实现方式。

请参见图 4，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 4 所示，该方法可以应用于上述图 1 所对应实施例中的目标用户终端，该方法可以包括：

步骤 S101，响应于对目标视频的触发操作，从所述目标视频的关键视频帧中确定目标像素点，并获取与所述目标像素点相关联的多媒体信息，其中，所

述关键视频帧是所述触发操作所在的视频帧，所述目标像素点是所述关键视频帧中与所述触发操作对应的像素点。

具体地，目标用户终端可以在访问目标应用时，在该目标应用的显示界面上显示用于承载多个业务数据信息的业务数据展示平台，例如，该业务数据展示平台上的每个业务数据信息可以为一个视频。在该业务数据展示平台上所展示的业务数据信息可以是由与该目标用户终端具有网络连接关系的应用服务器基于目标用户的用户画像数据（例如，该目标用户的历史行为数据）进行筛选后所确定的。当目标用户针对该业务数据展示平台上的一个业务数据信息（例如，一个视频）执行播放操作时，可以从该应用服务器对应的业务数据库中加载得到该视频对应的视频数据，进而可以在该目标用户终端的视频播放界面中播放该加载到的视频数据。进一步地，目标用户终端可以在播放该视频数据的过程中，获取该目标用户针对该视频播放界面中的目标对象（即需要跟踪的对象）执行的触发操作。所述触发操作例如是用鼠标点击或触摸目标用户终端显示屏上显示的视频帧中的目标对象中的某一点。可以将该触发操作对应的视频帧称之为关键视频帧，将在该关键视频帧中的触发操作对应的像素点称之为目标像素点。像素点是图像（例如，视频帧）中的一个点。如果图像是 640×480 分辨率的图片，则上面分布着 640×480 个像素点。通常，图像中的像素点具有空间位置和颜色（或灰度）属性。与此同时，该目标用户终端还可以在独立于该视频播放界面的子窗口中创建一个文本框，以便于该目标用户可以在该文本框中输入与该目标对象具有关联关系的多媒体信息。当目标用户在该文本框中输入多媒体信息之后，该目标用户终端可以获取与该目标对象相关联的多媒体信息，即与该目标对象相关联的多媒体信息可以统称为该目标用户所输入的用户文字，或用户评论。

其中，所述目标用户终端可以为具有视频数据播放功能的终端设备，所述目标用户终端可以为上述图 1 所对应实施例中的用户终端 3000a，该目标用户终端可以理解的一种移动终端。其中，所述应用服务器可以为上述图 1 所对应实施例中的业务服务器 2000。

为便于理解，进一步地，请参见图 5，是本申请实施例提供的一种获取多媒体信息的示意图。如图 5 所示，当目标用户终端在播放上述图 3 所对应实施例中视频 30a 的过程中，该目标用户终端可以将当前播放的视频 30a 作为目标视频。

可以理解的是，目标用户可以在播放该视频 30a 的任意一个时刻，对该视频 30a 所包含的多个视频帧中的某个视频帧执行触发操作。目标用户终端可以将该触发操作对应的视频帧作为关键视频帧。例如，如图 5 所示，目标用户可以在图 5 所示的视频播放界面 100a 中选取对象 A 作为目标对象，则目标用户终端可以将视频播放界面 100a 当前所播放的视频帧称之为关键视频帧。换言之，该目标用户终端可以将该选取操作（即，触发操作）对应的视频帧作为关键视频帧，并可以将该关键视频帧中选取操作对应的像素点作为目标像素点。此时，该目标像素点为该目标用户终端所获取到的目标视频内的关键视频帧中的像素点。

如图 5 所示，当目标用户在图 5 所示的视频播放界面 100a 中针对对象 A 执行触发操作之后，可以在图 5 所示的视频播放界面 200a 中弹出图 5 所示的文本框，该文本框也可以称之为一个对话框。如图 5 所示的文本框可以理解为独立于该视频播放界面 200a 的一个浮窗，且图 5 所示的文本框可以与图 5 所示的对象 A 之间存在关联关系（例如，可以与该对象 A 中的目标像素点之间存在显示位置上的相对关系，以构建得到该视频 30a 中的目标对象的目标像素点与该目标对象相关联的多媒体信息之间的关联性）。所述浮窗的实现可以与所述视频播放界面的实现类似或相同。应当理解，在本申请实施例的对话框中所输入的多媒体信息可以包含用户文字、用户图片和用户表情等数据，并可以将该目标用户在该对话框中所输入的用户文字（即文本信息）、用户图片（即图片信息）和用户表情（即表情信息）等统称为弹幕数据。所述弹幕数据的显示可以和字幕类似。

因此，当目标用户在图 5 所示的视频播放界面 200a 中的文本框中输入文本信息 A 之后，可以在图 5 所示的视频播放界面 300a 中显示该输入的文本信息 A。该输入的文本信息 A 可以为图 5 所示的与该对象 A 中的目标像素点之间存在一定位置间隔距离的文本信息。显示在该视频播放界面 300a 中的文本信息 A 可以称之为与该目标对象相关联的弹幕数据。

步骤 S102，基于所述目标像素点在所述关键视频帧中的位置信息，确定所述目标像素点对应的轨迹获取请求。

具体地，目标用户终端可以在该关键视频帧中确定该目标像素点的位置信息，并可以基于该关键视频帧在目标视频中的帧号以及该目标像素点在该关键视频帧中的位置信息，生成该目标像素点对应的轨迹获取请求，以便于可以进

一步执行步骤 S103。

其中，该轨迹获取请求可以用于指示应用服务器从预先计算得到的该目标视频中的所有像素点对应的轨迹信息中筛选与该目标像素点相匹配的轨迹信息。

步骤 S103，基于所述轨迹获取请求，获取与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息。

其中，所述目标轨迹信息包含所述目标像素点在所述关键视频帧的下一视频帧中的位置信息，所述目标像素点在所述关键视频帧的下一视频帧中的位置信息是通过跟踪所述目标像素点得到的。

在本申请实施例中，目标用户终端可以基于应用服务器预先计算得到的目标视频中的所有像素点在所有视频帧中的运动轨迹（每个像素点的运动轨迹可以统称为一个轨迹信息），从这些像素点对应的运动轨迹中筛选与目标像素点相匹配的像素点的运动轨迹，作为与该目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息。换言之，该目标用户终端在获取到该目标轨迹信息时，可以基于该目标轨迹信息中所包含的所述目标像素点在该关键视频帧的下一视频帧的位置信息，快速确定上述图 5 所对应实施例中目标对象中的目标像素点与该目标对象相关联的多媒体信息之间的位置间隔距离，或者目标像素点与多媒体信息之间的位置间隔，即可以快速得到该多媒体信息出现在该关键视频帧的下一视频帧中的位置信息。

可以理解的是，该位置间隔距离可以理解为所述关键视频帧中的目标像素点与对应的弹幕数据之间的相对位置间隔距离。即该位置间隔距离可以包含水平（即横向）方向上的相对位置间隔距离，也可以包含垂直（即纵向）方向上的相对位置间隔距离，从而可以确保目标用户终端在得到目标像素点在关键视频帧的下一个视频帧中的位置信息时，可以基于所述相对位置间隔距离，快速计算得到文本信息 A 在关键视频帧的下一视频帧中的位置信息。即此时，显示在上述图 5 所对应实施例中的视频播放界面 300a 中的文本信息 A 的位置信息，取决于目标像素点在该关键视频帧之后的其他视频帧的位置信息。在目标用户终端的计算性能比较一般的情况下，该目标用户终端可以从与该目标用户终端具有网络连接关系的应用服务器中，获取与该关键视频帧中的目标像素点的位置信息相符的轨迹信息，作为目标轨迹信息，以使该目标用户终端能够在获取

到该应用服务器预先计算好的该目标像素点的目标轨迹信息时，进一步基于该目标轨迹信息中的目标像素点出现在该关键视频帧的下一视频帧的位置信息，在有效时长内快速且准确地实现对该弹幕数据的快速跟踪，从而可以有效地降低该目标用户终端的计算量，以确保在该目标用户终端的计算性能比较一般的情况下，也可以快速对该弹幕数据进行跟踪。

其中，该有效时长可以为弹幕数据对应的显示时长，即该目标用户终端可以在该显示时长内对该目标对象所关联的弹幕数据进行跟踪。

应当理解，由于该目标视频内的每个像素点的运动轨迹（即每个像素点的轨迹信息）是由每个像素点在该目标视频的每个视频帧中的位置信息所确定的。其中，对于包含多个视频帧的目标视频而言，本申请实施例可以将该多个视频帧中任意相邻的两个视频帧确定为一个图像对。应当理解，从该多个视频帧中所确定出的每个图像对所包含的两个视频帧中的一个视频帧，可以称之为第一视频帧，另一个视频帧可以称之为第二视频帧。对于上述图 2 所对应实施例中的第 1 时刻所对应的视频帧和第 2 时刻所对应的视频帧所构成的图像对 1 而言，可以将该图像对 1 中的第 1 时刻所对应的视频帧称之为第一视频帧，并可以将该第 2 时刻所对应的视频帧称之为第二视频帧，进而可以基于预先计算出的该图像对 1 中的这两个视频帧之间的平均位移矩阵，对该第一视频帧中所有像素点进行追踪，以确定出该第一视频帧中的所有像素点出现在第二视频帧中的位置信息。同理，对于上述图 2 所对应实施例中的第 2 时刻所对应的视频帧和第 3 时刻所对应的视频帧所构成的图像对 2 而言，也可以将该第 2 时刻所对应的视频帧称之为第一视频帧，并将该第 3 时刻所对应的视频帧称之为第二视频帧，从而可以基于预先计算好的该图像对 2 中的这两个视频帧之间的平均位移矩阵，对该第一视频帧中所有像素点进行追踪，以确定出该第一视频帧中的所有像素点出现在第二视频帧中的位置信息。以此类推，本申请实施例可以得到每个图像对对应的平均位移矩阵，每个图像对对应的平均位移矩阵可以称之为每个图像对中的第一视频帧对应的平均位移矩阵，每个第一视频帧对应的平均位移矩阵可以用于将第一视频帧中的所有像素点映射到第二视频帧中，以在第二视频帧中准确得到这些映射得到的像素点的位置信息。应当理解，本申请实施例中的平均位移矩阵可以包含纵向平均位移矩阵和横向平均位移矩阵。通过纵向平均位移矩阵可以将第一视频帧中的每个像素点的第一纵向坐标值（例如，y 值）

进行纵向坐标变换，以得到相应像素点映射到第二视频帧中的第二纵向坐标；同理，通过横向平均位移矩阵可以将第一视频帧中的每个像素点的第一横向坐标值（例如， x 值）进行横向坐标变换，以得到相应像素点映射到第二视频帧中的第二横向坐标。应当理解，本申请实施例可以将每个像素点在第一视频帧中的第一横向坐标和第一纵向坐标值称之为该第一视频帧中的每个像素点的第一位置信息，并可以将每个像素点映射在第一视频帧中的第二横向坐标和第二纵向坐标值称之为该第二视频帧中的每个映射得到的像素点的第二位置信息。由于每个图像对均对应一个平均位移矩阵，从而可以基于在第一视频帧中像素点的第一位置信息计算得到相应的第二位置信息，并可以将计算所得到的每个第二视频帧中所映射得到的像素点的第二位置信息进行保留，进而可以将同一像素点在每个视频帧中的位置信息进行整合，以得到该视频帧中的所有像素点的运动轨迹，从而可以实现对该目标视频的所有视频帧内的所有像素点的跟踪。

应当理解，上述图 2 所对应实施例所显示的该目标视频中的多个视频可以为多个连续的图像帧，因此，通过将上述图 2 所示的目标视频进行拆分后，可以对每个拆分所得到的图像帧（即视频帧）按照播放顺序设置相应的视频帧号，例如，上述第 1 时刻所得到的视频帧的视频帧号可以为 1，该视频帧号 1 则可以用于表示该第 1 时刻所得到的视频帧为该目标视频中的第一帧；同理，上述第 2 时刻所得到的视频帧的视频帧号可以为 2，该视频帧号 2 则可以用于表示该第 2 时刻所得到的视频帧为该目标视频中的第二帧。依次类推，上述第 $n-1$ 时刻所得到的视频帧的视频帧号可以为 $n-1$ ，该视频帧号 $n-1$ 则可以用于表示该第 $n-1$ 时刻所得到的视频帧为该目标视频中的第 $n-1$ 帧；上述第 n 时刻所得到的视频帧的视频帧号可以为 n ，该视频帧号 n 则可以用于表示该第 n 时刻所得到的视频帧为该目标视频中的第 n 帧，即该目标视频中的最后一帧。

为便于理解，本申请实施例可以将上述图 2 所示的多个视频帧的第一帧和第二帧所构成的图像对称之为首个图像对，以阐述通过平均位移矩阵将第一帧中的像素点平移变换到第二帧中，实现像素跟踪的具体过程。其中，该首个图像对中的第一帧即为上述图 2 所对应实施例中的第 1 时刻对应的视频帧，该首个图像对中的第二帧即为上述图 2 所对应实施例中的第 2 时刻对应的视频帧。进一步地，请参见图 6，是本申请实施例提供的一种全图像素追踪的示意图。如图 6 所示的图像对（1，2）可以为前述首个图像对。该首个图像对中的第一视

视频帧可以为前述第 1 时刻对应的视频帧（即第一帧），该首个图像对中的第二视频帧可以为前述第 2 时刻对应的视频帧（即第二帧）。其中，应当理解，该图像对（1，2）中的数值 1 即为第一帧的视频帧号，数值 2 即为第二帧的视频帧号。因此，可以用该目标视频中的每一个视频帧的视频帧号来表征该目标视频中的任意两个前后相邻的视频帧。如图 6 所示的像素点显示区域 600a 可以包括从该图像对的第一视频帧中所提取到的所有像素点，比如，该像素点显示区域 600a 中的每个像素点均可以对应一个区域标识。图 6 的像素点显示区域 600a 仅是用于示例，像素点显示区域 600a 也可称为像素点区域等。应当理解，本申请实施例仅以从该第一视频帧中所获取到的像素点为 20 个像素点为例，在实际情况下，从该第一视频帧中所获取到的像素点的数量会远远多于本申请实施例所列举的 20 个。应当理解，由于同一视频内的多个视频帧是由同一终端进行图像采集后所得到的，因此，对于同一视频内所包含的每个视频帧中的像素点的数量是相同的。

如图 6 所示，在获取到该第一视频帧中的每个像素点之后，可以将这些获取到的所有像素点统称为像素点，进而可以通过图 6 所示的平均位移矩阵对该像素点显示区域 600a 中的所有像素点进行跟踪，进而可以在第二视频帧对应的像素点显示区域 700a 中确定所映射得到的像素点的位置信息。比如，以图 6 所示的像素点 A 为例，该像素点 A 在图 6 所示的像素点显示区域 600a 中的位置信息可以为区域标识 5 的坐标位置信息，通过该平均位移矩阵可以将该像素点 A 映射到图 6 所示的像素点显示区域 700a 中，该像素点 A 在图 6 所示的像素点显示区域 700a 中的位置信息可以为区域标识 10 的坐标位置信息。本申请实施例的计算得到该像素点 A 在该第二视频帧中的位置信息之后，可以对其进行存储。由于该目标视频中的每个图像对均可以对应一个平均位移矩阵，因此，可以计算出每个第一视频帧中的像素点映射到第二视频帧中的位置信息。通过将每一个图像对中同一像素点出现在连续视频帧内的位置信息进行整合，可以得到该像素点 A 出现在该目标视频中的每个视频帧中的位置信息，进而可以基于该像素点 A 在该目标视频中的每个视频帧中的位置信息，得到该像素点 A 的运动轨迹。

同理，对于该视频帧中的所有像素点中的其他像素点而言，也可以通过每个图像对对应的平均位移矩阵（即每个图像对中的第一视频帧对应的平均位移

矩阵), 确定出其他像素点中的每个像素点在该目标视频的每个视频帧中的位置信息, 进而可以得到其他像素点的运动轨迹, 以实现对该每个图像对中的第一视频帧中的所有像素点的全图像素追踪, 进而可以得到该目标视频的所有像素点在每个视频帧中的位置信息。应当理解, 本申请实施例可以将该目标视频中所有像素点对应的轨迹信息统称为像素点对应的轨迹信息。

其中, 若该目标用户终端的计算性能难以满足跟踪大量像素点的计算要求时, 为了减小该目标用户终端的计算量, 可以通过与该目标用户终端具有网络连接关系的应用服务器预先计算好该目标视频内的所有像素点的运动轨迹。从而可以在该目标用户终端实际播放该目标视频时, 应用服务器接收目标用户终端所发送的目标像素点在关键视频帧中的位置信息, 从预先计算好的像素点对应的轨迹信息中筛选与该目标像素点匹配的轨迹信息, 作为目标轨迹信息, 进而将该目标轨迹信息返回给目标用户终端, 以使目标用户终端可以基于该获取到的目标轨迹信息进一步执行步骤 S104。其中, 该目标像素点为目标用户所选取的关键视频帧内的像素点。

可选地, 若该目标用户终端具有良好的计算性能, 则可以在该目标用户终端中预先计算好该目标视频内的所有像素点的运动轨迹, 从而可以在该目标用户终端实际播放该目标视频时, 进一步基于目标用户所选取的目标对象内的目标像素点在这些像素点对应的轨迹信息中筛选与该目标像素点匹配的轨迹信息作为目标轨迹信息, 以便于可以进一步执行步骤 S104。

步骤 S104, 当播放所述关键视频帧的下一视频帧时, 基于所述目标轨迹信息中的所述目标像素点在所述关键视频帧的下一视频帧中的位置信息, 显示所述多媒体信息。

为便于理解, 进一步地, 请参见图 7, 是本申请实施例提供的一种在连续多个视频帧中跟踪弹幕数据的示意图。应当理解, 本申请实施例中用于进行弹幕跟踪的连续多个视频帧可以包含当前正在播放的关键视频帧, 以及尚未播放的位于该目标视频内的关键视频帧之后的视频帧。比如, 当如图 7 所示的视频帧 10 作为关键视频帧时, 可以在该关键视频帧后的每个视频帧 (例如, 视频帧 20、视频帧 30 等视频帧) 中, 对出现在视频帧 10 中的弹幕数据进行弹幕跟踪。其中, 图 7 所示的视频帧 10 可以为上述图 5 所对应实施例中的显示在视频播放界面 300a 中的视频帧, 即图 7 所示的视频帧 10 可以为当前正在目标用户终端中

播放的目标视频中与该目标用户相关联的关键视频帧。换言之，本申请实施例中的关键视频帧可以理解为目标用户选取目标对象时所执行的触发操作对应的视频帧。应当理解，本申请实施例中的目标对象可以包含目标用户在正在播放的视频帧通过点击操作所选取的人物、动物、植物等对象。换言之，该目标用户终端可以将该目标用户所选取的对象称之为目标对象，并可以将该关键视频帧中目标对象内与触发操作对应的像素点作为目标像素点，进而可以从应用服务器预先计算好的该目标视频内的所有像素点的轨迹信息中，获取与该目标对象中目标像素点相关联的轨迹信息，从而可以将获取到的轨迹信息作为该目标像素点对应的目标轨迹信息。该目标轨迹信息中可以包含该目标像素点在该关键视频帧中的位置信息，还可以包含该目标像素点在该关键视频帧之后的每个视频帧（例如，该关键视频帧的下一视频帧）中的位置信息。应当理解，基于该目标像素点在该关键视频帧后的每个视频帧中的位置信息，可以快速计算得到与目标对象关联的（也是与目标像素点关联的）多媒体信息（即上述图 5 所对应实施例中的弹幕数据）在该关键视频帧后的每个视频帧中的位置信息，以实现对该目标对象相关联的弹幕数据的快速跟踪，从而可以在该目标用户终端播放该关键视频帧的下一视频帧时，基于计算得到的所述弹幕数据在所述关键视频帧中的位置信息，在所述下一视频帧中实时显示该弹幕数据。所述弹幕数据的显示可以类似于字幕的显示。

应当理解，本申请实施例通过将弹幕数据与该关键视频帧中的目标对象进行关联，可以实现弹幕数据与目标对象之间的如影随行，即用户所输入的弹幕可以在有效跟踪时长内，一直跟着这个要跟踪的目标对象进行相对运动。比如，在该目标视频内，若该关键视频帧后的连续多个视频帧中均存在目标对象，则可以基于该目标对象中的目标像素点在这几个连续视频帧中的位置信息，显示与该目标对象相关联的弹幕数据（即前述文本信息 A）。

目标用户终端也可以将用户输入的弹幕数据（多媒体信息）以及计算得到的弹幕数据在目标视频各视频帧中的位置信息传输给服务器。或者，服务器可以接收目标用户终端发送的用户点击的目标视频中关键视频帧的帧号、目标像素点坐标、输入的弹幕数据（多媒体信息），计算目标像素点在目标视频各视频帧中的目标轨迹信息，根据该目标轨迹信息，计算弹幕数据在目标视频各视频帧中的位置信息，并将所述弹幕数据的位置信息进行保存。服务器在接收目标

用户终端发送的信息时，还可以接收目标用户终端的标识和/或目标用户终端上用户登录目标应用的用户标识等信息。然后在其他用户终端播放所述目标视频时，服务器可以将所述弹幕数据及其在目标视频各视频帧中的位置信息、用户标识发送给其他用户终端，由其他用户终端根据弹幕数据的位置信息，在目标视频的各视频帧中显示弹幕数据。

应当理解，对于在目标用户终端中播放的任意一个视频而言，目标用户可以在当前时刻为 T1 时刻时，从当前播放的视频帧中选择出该目标用户认为需要进行跟踪的对象。可以将该选择的对象称为目标对象。进而，目标用户终端可以基于从预先计算好的该视频内的所有像素点对应的轨迹信息，筛选出与该目标对象中的目标像素点相关联的轨迹信息，快速得到该目标对象中的目标像素点对应的目标轨迹信息。其中，应当理解，预先计算好的该视频内的每个像素点的任意一个像素点对应的轨迹信息，均可以用于描述该像素点在该视频内的每个视频帧中的位置信息。因此，当目标用户终端将在 T1 时刻播放的视频帧作为关键视频帧时，可以在该关键视频帧内得到该目标对象中的目标像素点，进而可以从该目标像素点对应的目标轨迹信息中，快速得到该目标像素点像素点在该关键视频帧后的每个视频帧中的位置信息，从而可以基于该目标轨迹信息显示该目标对象所关联的多媒体信息。其中，可以理解的是，若目标像素点在每个视频帧中所构成的轨迹信息为一个圆，则与该目标对象相关联的多媒体信息则可以同步跟着该轨迹信息转圈。由此，通过预先对该目标视频内的所有像素点进行跟踪，可以预先得到每个像素点对应的轨迹信息，从而可以在目标用户终端中播放该目标视频时，可以基于目标用户所执行的触发操作，将触发操作所对应的目标对象中的像素点作为目标像素点，以获取与该目标像素点相关联的轨迹信息作为目标轨迹信息，进而可以快速基于该获取到的目标轨迹信息，快速实现对该目标对象所关联的多媒体信息的准确跟踪。

那么，对于在该关键视频帧中的不同对象而言，可以得到不同对象中的目标像素点分别对应的运动轨迹，从而使得与不同目标对象相关联的弹幕数据可以绕着不同的轨迹进行运动，使得弹幕数据和其所针对的对象之间的关联更强，进而可以丰富弹幕数据的视觉展示效果，还可以提升弹幕数据的显示方式的灵活度。

本申请实施例可以在获取到目标用户针对目标视频的触发操作时，将该目

标视频中该触发操作对应的视频帧作为关键视频帧，从而可以从该关键视频帧中确定目标像素点，并获取与该目标像素点以及该目标像素点所在目标对象关联的多媒体信息（例如，该多媒体信息可以为该目标视频中的用户文字、图片、表情等弹幕数据）。进一步地，基于该目标像素点在该关键视频帧中的位置信息，确定所述目标像素点对应的轨迹获取请求，进而可以基于该轨迹获取请求，获取所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，以便于在播放该关键视频帧的下一视频帧时，能够基于该目标轨迹信息，显示与该目标像素点以及该目标像素点所在目标对象相关联的弹幕数据。由此可见，本申请实施例可以在确定出关键视频帧时，进一步从该关键视频帧中的所有像素点的轨迹信息中筛选出目标像素点的轨迹信息，并将筛选出的目标像素点的轨迹信息作为目标轨迹信息，以便于能够基于得到的目标轨迹信息丰富弹幕数据的展示效果。比如，对于不同目标对象中的目标像素点，所得到的目标轨迹信息可能会不同，进而使得弹幕数据的展示效果会不同。此外，基于目标对象与弹幕数据之间的关联关系，可以快速确定出该弹幕数据在该关键视频帧之后的每个视频帧中的位置信息，换言之，该弹幕数据会在该目标视频中一直跟着该目标对象进行变动，进而可以丰富视频中的用户文字的视觉展示效果，并可以使得弹幕数据与目标对象或评论的视频中的对象之间的关联性更强。

为便于理解，进一步地，请参见图 8，是本申请实施例另一种视频数据处理方法的示意图。该方法主要用于阐述目标用户终端与应用服务器之间的数据交互过程，该方法可以包含以下步骤：

步骤 S201，从目标视频中获取相邻的第一视频帧和第二视频帧。

具体地，应用服务器可以从目标视频所包含的多个视频帧中确定多个图像对，该多个图像对中的每个图像对均是由该目标视频中的两个相邻的视频帧所构成的。

换言之，应用服务器可以在对该目标视频进行视频预处理的阶段中，先把该目标视频进行分帧处理，以将该目标视频中的多个视频帧按照播放时间序列分成一张张的图片，即可以得到上述图 2 所示的基于播放时间序列进行排布的多个视频帧。通过对该目标视频中的每个视频帧进行拆分，可以得到每个视频帧对应的图片，即一个图像可以视为一个图像帧。进一步地，该应用服务器可以通过前向后向光流法，对每个图像对中的两个视频帧内的像素点进行像素追

踪。比如，以包含 n 个视频帧的目标视频而言，该应用服务器可以根据该目标视频内的每个视频帧的视频帧号，将具有相邻帧号的两个视频帧确定为一个图像对。换言之，该应用服务器可以将视频帧号为 1 的视频帧和视频帧号为 2 的视频帧确定为一个图像对。同理，该应用服务器可以将视频帧号为 2 的视频帧和视频帧号为 3 的视频帧确定为一个图像对。以此类推，该应用服务器可以将视频帧号为 $n-1$ 的视频帧和视频帧号为 n 的视频帧确定为一个图像对。

对于该目标视频内的每个图像对而言，可以用每个视频帧的视频帧号来描述相应图像对中的两个视频帧。因此，对于前述包含 n 个视频帧的目标视频而言，可以得到 $n-1$ 个图像对，这 $n-1$ 个图像对可以表示为：(1, 2), (2, 3), (3, 4), ..., ($n-1$, n)。其中，图像对中的视频帧号为 1 的视频帧可以称之为该目标视频的第一帧，视频帧号为 2 的视频帧可以称之为该目标视频的第二帧，以此类推，可以将图像对中的视频帧号为 $n-1$ 的视频帧可以称之为该目标视频的第 $n-1$ 帧，视频帧号为 n 的视频帧可以称之为该目标视频的第 n 帧。进一步地，该应用服务器可以通过云端前向后向光流法，对该目标视频的每个图像对中的像素点进行跟踪。其中，该云端前向后向光流法可以统称为光流法，该光流法可以用于对每个图像对中的两个视频帧之间的像素点位移进行计算。

可以理解的是，由于每个图像对均是由前后相邻的两个视频帧所构成的，因此，可以将每个图像对中的一个视频帧称之为第一视频帧，并将每个图像对中的另一个视频帧称之为第二视频帧，以便于可以进一步执行步骤 S202。

其中，为便于理解，本申请实施例可以将目标视频中所获取到的每个图像对中的两个视频帧，统称为第一视频帧和第二视频帧，即该应用服务器可以从目标视频中获取相邻的第一视频帧和第二视频帧。

步骤 S202，基于所述目标视频对应的光流追踪规则、所述第一视频帧中的像素点、所述第二视频帧中的像素点，确定所述第一视频帧对应的平均位移矩阵。

具体地，应用服务器可以对第一视频帧的所有像素点进行提取。可以将提取到的所有像素点统称为像素点。其中，该目标视频对应的光流追踪规则可以包含前述云端前向后向光流法，还可以包含云端位移积分法和云端位移差分法。应当理解，通过该光流追踪规则可以对每个图像对中的第一视频帧中的像素点、第二视频帧中的像素点进行光流运算，以得到每个图像对对应的光流追踪结果，

进而可以基于该光流追踪结果确定每个图像对对应的目标状态矩阵和目标位移矩阵。如果直接使用目标状态矩阵和目标位移矩阵对视频帧中的像素点进行跟踪也是可以的，但可能跟踪结果不是很准确。为了提高跟踪的准确性，可以在计算像素点在第一视频帧与第二视频帧中的位移时，考虑像素点的周围像素点位移的信息。例如，应用服务器可以针对所述第一视频帧中的每个像素点，以其为中心，选取该像素点周围的图块（包括该像素点和该像素点周围的像素点），计算所述图块中所有像素点的平均位移，作为该像素点的位移。这种处理方式的运算量可能会比较大。根据本申请实施例，通过该光流追踪规则还可以进一步对每个图像对对应的目标状态矩阵和目标位移矩阵进行位移积分运算，以得到每个图像对对应的状态积分矩阵和位移积分矩阵。进一步地，通过该光流追踪规则还可以对每个图像对对应的状态积分矩阵和位移积分矩阵进行位移差分运算，以得到每个图像对对应的平均位移矩阵。换言之，通过该光流追踪规则可以准确得到能够用于对每个图像对中的第一视频帧中的像素点的位置信息进行准确跟踪的平均位移矩阵。通过上述的位移积分运算和位移差分运算，应用服务器能够批量计算第一视频帧中的像素点和第二视频帧中像素点的平均位移，从而可以提高运算的速度，提高像素点和视频帧的处理的效率。

其中，云端前向后向光流法可以用于同步对每个图像对中的第一视频帧和第二视频帧进行正向反向光流法计算，以得到每个图像对对应的光流追踪结果。换言之，应用服务器所得到的光流追踪结果可以包含每个图像对中的第一视频帧对应的正向位移矩阵，还可以包含每个图像对中的第二视频帧对应的反向位移矩阵。在本申请实施例中，正向位移矩阵和反向位移矩阵的中的每个矩阵元素均可以包含两个维度上的位移（比如， $(\Delta x, \Delta y)$ ）。其中，这两个维度上的位移可以理解为同一像素点分别在水平方向上的位移（即 Δx ）和在垂直方向上的位移（即 Δy ）。应当理解，对于该目标视频中的每个图像对而言，通过该光流法进行计算后，均可以得到一个正向水平方向位移矩阵、一个正向垂直方向位移矩阵、一个反向水平方向位移矩阵、一个反向垂直方向位移矩阵，并可以将得到的四个矩阵称之为光流结果。进一步地，应用服务器可以为每个图像对中的第一视频帧设置一个初始状态矩阵，进而可以根据前面所得到的正向位移矩阵和反向位移矩阵，判断每个图像对中的第一视频帧中的像素点是否满足目标筛选条件。若是第一视频帧中存在满足目标筛选条件的像素点，则应用服务器可

以将满足目标筛选条件的像素点确定为有效像素点，进而可以根据确定出的有效像素点，对所述第一视频帧对应的初始状态矩阵和所述正向位移矩阵进行修正，以得到每个图像对中的第一视频帧对应的目标状态矩阵和目标位移矩阵。进一步地，应用服务器可以通过上述云端位移积分法和云端位移差分法，以及所得到的目标状态矩阵和目标位移矩阵，确定得到每个图像对中的第一视频帧对应的平均位移矩阵。

其中，本申请实施例可以将该正向水平方向位移矩阵和正向垂直方向位移矩阵统称为正向位移矩阵，并可以将反向水平方向位移矩阵、反向垂直方向位移矩阵统称为反向位移矩阵。为便于理解，本申请实施例以多个图像对中的一个图像对为例，以阐述通过该图像对中的第一视频帧和第二视频帧，得到该图像对对应的平均位移矩阵的过程。其中，该图像对中的第一视频帧可以为上述视频帧号为1的视频帧，第二视频帧可以为上述视频帧号为2的视频帧。因此，由该视频帧号为1的视频帧和视频帧号为2的视频帧所构成的图像对称之为图像对1，且该图像对1可以表示为(1,2)。

其中，通过光流法计算后所得到的该图像对1对应的正向位移矩阵可以包含正向水平位移矩阵（例如，该正向水平位移矩阵可以为矩阵 $Q_{1,2,x}$ ）和正向垂直位移矩阵（例如，该正向垂直位移矩阵可以为矩阵 $Q_{1,2,y}$ ）。其中，应当理解，矩阵 $Q_{1,2,x}$ 中的每个矩阵元素可以理解为第一视频帧中的像素点在第二视频帧中的水平方向上的位移。即该正向水平位移矩阵中的每个矩阵元素可以称之为第一视频帧中的像素点对应的第一横向位移。同理，矩阵 $Q_{1,2,y}$ 中的每个矩阵元素可以理解为第一视频帧中的像素点在第二视频帧中的垂直方向上的位移。即该正向垂直位移矩阵中的每个矩阵元素可以称之为第一视频帧中的像素点对应的第一纵向位移。换言之，通过光流法计算后所得到的这两个矩阵（即矩阵 $Q_{1,2,x}$ 和矩阵 $Q_{1,2,y}$ ）的矩阵大小与第一视频帧的大小一样，即一个矩阵元素可以对应第一视频帧中的一个像素点。

同理，通过光流法计算后所得到的该图像对1对应的反向位移矩阵，可以包含反向水平位移矩阵（即该反向水平位移矩阵可以为矩阵 $Q_{2,1,x}$ ）和反向垂直位移矩阵（即该反向垂直位移矩阵可以为矩阵 $Q_{2,1,y}$ ）。其中，应当理解，矩阵 $Q_{2,1,x}$ 中的每个矩阵元素可以理解为第二视频帧中的像素点在第一视频帧中的水平方向上的位移。即该反向水平位移矩阵中的每个矩阵元素可以称之为第二视

视频帧中的像素点对应的第二横向位移。同理，矩阵 $Q_{2,1,y}$ 中的每个矩阵元素可以理解为第二视频帧中的像素点在第一视频帧中的垂直方向上的位移。即该反向垂直位移矩阵中的每个矩阵元素可以称之为第二视频帧中的像素点对应的第二纵向位移。换言之，通过光流计算法所得到的这两个矩阵（即矩阵 $Q_{2,1,x}$ 和矩阵 $Q_{2,1,y}$ ）的矩阵大小与第二视频帧的大小一样，即一个矩阵元素可以对应第二视频帧中的一个像素点。

应当理解，对于目标视频内的每个视频帧而言，每个视频帧中的像素点的数量是相同的，因此，通过光流计算法所得到的与该图像对 1 对应的这四个矩阵（即矩阵 $Q_{1,2,x}$ 、矩阵 $Q_{1,2,y}$ 、矩阵 $Q_{2,1,x}$ 、矩阵 $Q_{2,1,y}$ ）的矩阵大小是相同的。比如，若每个视频帧中的像素点数为 $m \times n$ 个，则所得到的这四个矩阵的矩阵大小均可以为 $m \times n$ 。由此可见，正向水平位移矩阵和正向垂直位移矩阵中的每个矩阵元素，均能够与第一视频帧中的相应像素点进行对应。因此，该图像对 1 对应的正向位移矩阵中的每个矩阵元素可以表示第一视频帧中的像素点在第二视频帧中的两个维度上的位移。可以将该图像对 1 对应的正向位移矩阵统称为第一视频帧对应的正向位移矩阵。同理，图像对 1 对应的反向位移矩阵中的每个矩阵元素可以表示第二视频帧中的像素点在第一视频帧中的两个维度上的位移。可以将该图像对 1 对应的反向位移矩阵统称为第二视频帧对应的反向位移矩阵。

由此可见，应用服务器可以基于所述第一视频帧中的像素点的第一位置信息和所述光流追踪规则，将所述第一视频帧中的像素点正向映射到所述第二视频帧，并在所述第二视频帧中确定所映射得到的第一映射点的第二位置信息，并可以进一步基于所述像素点的第一位置信息、所述第一映射点的第二位置信息，确定所述第一视频帧对应的正向位移矩阵。进一步地，应用服务器可以基于所述第二视频帧中的像素点的第二位置信息和所述光流追踪规则，将所述第二视频帧中的像素点反向映射到所述第一视频帧，并在所述第一视频帧中确定所映射得到第二映射点的第三位置信息，并进一步基于所述第一映射点的第二位置信息、所述第二映射点的第三位置信息，确定所述第二视频帧对应的反向位移矩阵。其中，所述第一映射点和第二映射点均为通过光流法将图像对中的一个视频帧中的像素点映射到另一个视频帧中所得到的像素点。

进一步地，应用服务器可以基于所述第一视频帧中的像素点的第一位置信

息、所述正向位移矩阵、所述反向位移矩阵，将所述像素点中满足目标筛选条件的像素点确定为有效像素点。其中，该应用服务器确定有效像素点的具体过程可以描述为：

应用服务器可以从所述第一视频帧中的像素点中获取第一像素点，并在所述第一视频帧中确定所述第一像素点的第一位置信息，并从所述正向位移矩阵中确定所述第一像素点对应的第一横向位移和第一纵向位移；进一步地，应用服务器可以基于所述第一像素点的第一位置信息、所述第一像素点对应的第一横向位移和第一纵向位移，将所述第一像素点正向映射到所述第二视频帧，并在所述第二视频帧中确定所映射得到的第二像素点的第二位置信息；进一步地，应用服务器可以从所述反向位移矩阵中确定所述第二像素点对应的第二横向位移和第二纵向位移，并基于所述第二像素点的第二位置信息、所述第二像素点对应的第二横向位移和第二纵向位移，将所述第二像素点反向映射到所述第一视频帧，并在所述第一视频帧中确定所映射得到的第三像素点的第三位置信息；进一步地，应用服务器可以基于所述第一像素点的第一位置信息、所述第三像素点的第三位置信息，确定所述第一像素点与所述第三像素点之间的误差距离，并根据所述第一像素点的第一位置信息、所述第二像素点的第二位置信息，确定包含第一像素点的图像块与包含所述第二像素点的图像块之间的相关系数；进一步地，应用服务器可以将所述像素点中误差距离小于误差距离阈值、且所述相关系数大于相关系数阈值的像素点确定为有效像素点。

为了验证通过上述得到的光流跟踪结果中的正向位移矩阵和反向位移矩阵中的矩阵元素的正确性，本申请实施例可以通过矩阵变换的方式对上述四个位移矩阵中的矩阵元素进行筛选，即可以通过所构建的初始状态矩阵中的相应像素点所在位置上的矩阵元素的变化情况，从这四个矩阵中去除相应像素点对应位置上的、存在较大位移误差的矩阵元素，即可以从第一视频帧的像素点中确定出有效像素点。

为便于理解，进一步地，请参见图 9，是本申请实施例提供的一种确定有效像素点的方法。如图 9 所示，该应用服务器在对这四个矩阵中的矩阵元素进行筛选之前，可以首先初始化一个与该第一视频具有相同大小的状态矩阵 S_1 。此时，该应用服务器可以将该状态矩阵 S_1 称为初始状态矩阵。其中，该初始状态矩阵中，与各像素点对应的矩阵元素的值可以称为第一数值。此时，该初始状

态矩阵中的第一数值均为零。该初始状态矩阵中的矩阵元素的值的变化情况可以用来表示第一视频帧中的像素点是否满足目标筛选条件，从而可以将满足目标筛选条件的像素点作为有效跟踪像素点（即有效像素点）。

其中，如图 9 所示的第一图像帧可以为上述图像对 1 中的视频帧号为 1 的视频帧。该第一视频帧中的像素点中可以包含如图 9 所示的第一像素点 p_1 ，即该第一像素点 p_1 可以为该第一视频帧的所有像素点中的一个像素点，并可以将该第一像素点 p_1 在第一视频帧中的位置信息称为第一位置信息。进一步地，应用服务器可以从上述正向位移矩阵中的正向水平位移矩阵中找到该第一像素点 p_1 对应的第一横向位移，并从上述正向位移矩阵中的正向垂直位移矩阵中找到该第一像素点 p_1 对应的第一纵向位移，进而可以基于该第一像素点 p_1 的第一位置信息、该第一像素点 p_1 对应的第一横向位移和第一纵向位移，将该第一像素点 p_1 正向映射到图 9 所示的第二视频帧，并在第二视频帧中确定出所映射得到的第二像素点 p_2 的第二位置信息。可以理解的是，此时，该第二像素点 p_2 是通过将第一像素点 p_1 进行矩阵变换后所得到的像素点。进一步地，该应用服务器可以从上述反向位移矩阵中确定该第二像素点 p_2 对应的第二横向位移和第二纵向位移，并可以根据该第二像素点 p_2 的第二位置信息、该第二像素点 p_2 对应的第二横向位移和第二纵向位移，将该第二像素点 p_2 反向映射回图 9 所示的第一视频帧，并可以在该第一视频帧中确定所映射得到的第三像素点 p_1' 的第三位置信息。可以理解的是，此时，该第三像素点 p_1' 是通过将由第一像素点 p_1 映射所得到的第二像素点 p_2 进行矩阵变换后所得到的像素点。

进一步地，该应用服务器可以在第一视频帧中，确定出第一像素点 p_1 的第一位置信息和矩阵变换后所得到的第三像素点 p_1' 的第三位置信息这两个位置之间的位置误差 $t_{11'}$ 。进一步地，应用服务器可以在图 9 所示的第一视频帧中以第一像素点 p_1 的第一位置信息为中心，选取一个 $k \times k$ 像素（例如， 8×8 像素）大小的图像块 10。另外，如图 9 所示，应用服务器可以在图 9 所示的第二视频帧中以第二像素点 p_2 的第二位置信息为中心，同样选取一个 $k \times k$ 像素大小的图像块 20，进而可以计算这两个图像块之间的相关系数（该相关系数可以为 $N_{1,2}$ ）。

其中，相关系数 $N_{1,2}$ 的计算表达式如下：

$$N_{1,2} = \frac{\sum_{a=1}^k \sum_{b=1}^k (patch_1(a,b) - E(patch_1))(patch_2(a,b) - E(patch_2))}{\sqrt{\sum_{a=1}^k \sum_{b=1}^k (patch_1(a,b) - E(patch_1))^2 (patch_2(a,b) - E(patch_2))^2}} \quad \text{公式 (1)}$$

在公式(1)中的 $patch_1(a,b)$ 可以表示图9所示的图像块10的第 a 行 b 列位置上的像素点的像素值。所述像素值可以是像素的灰度值,在 0~255 之间。 $E(patch_1)$ 表示图9所示的图块10的平均像素值。 $patch_2(a,b)$ 表示图9所示的图像块20的第 a 行 b 列位置上的像素点的像素值。 $E(patch_2)$ 表示图9所示的图像块20的平均像素值。

其中,可以理解的是,本申请实施例在计算得到图9所示的第一像素点 $p1$ 和第三像素点 $p1'$ 之间的误差距离之后,可以将该计算得到的误差距离与预设的误差距离进行比较,如果 $t_{11'} < T_B$ 且 $N_{1,2} > T_A$ (其中, T_B 为设定的误差距离阈值, T_A 为设定的相关系数阈值),则表示该第一视频帧中的第一像素点 $p1$ 满足前述目标筛选条件,从而可以确定该第一像素点 $p1$ 是有效像素点。

进一步地,该应用服务器可以将与该第一像素点 $p1$ 对应的初始状态矩阵 S_1 中对应位置上的矩阵元素的值设置为第二数值。例如,可以将该初始状态矩阵 S_1 中与第一像素点 $p1$ 对应的元素的值由 0 切换为 1,以表示第一视频帧中的第一像素点 $p1$ 为有效像素点。反之,如果 $t_{11'} \geq T_B$ 和/或 $N_{1,2} < T_A$,则表示该第一视频帧中的第一像素点 $p1$ 不满足前述目标筛选条件。此时,该应用服务器可以判断上述图9所示的第一像素点 $p1$ 为无效跟踪像素点。即在该初始状态矩阵 S_1 中与第一像素点 $p1$ 对应的元素的值仍然为 0。与此同时,该应用服务器可以进一步对上述正向位移矩阵(即上述矩阵 $Q_{1,2,x}$ 和矩阵 $Q_{1,2,y}$)中与该第一像素点 $p1$ 对应位置上的矩阵元素的值设置为 0,从而可以将该包含所述第一数值的正向位移矩阵确定为目标位移矩阵(例如,正向水平位移矩阵 Q_{x1} 和正向水平位移矩阵 Q_{y1})。即该目标位移矩阵中的这些位置上的矩阵元素可以用于表示从上述正向位移矩阵中所筛选出、并过滤掉存在较大误差的误跟踪位移之后所确定的矩阵。

可以理解的是,对于上述图9所示的其他像素点而言,可以依次从图9所示的第一视频帧中选取像素点作为第一像素点,以重复上面确定有效像素点的步骤,直到该第一视频帧中的所有像素点都作为第一像素点之后,可以确定出该第一视频帧中的所有有效像素点。从而可以基于该有效像素点在该初始状态矩阵中的位置信息对该初始状态矩阵中的矩阵元素进行更新,进而可以将包含

第二数值的初始状态矩阵确定为所述第一视频帧对应的目标状态矩阵 S_1 。并可以得到该第一视频帧对应的目标位移矩阵（即目标水平位移矩阵 $Q_{x,1}$ 和目标垂直位移矩阵 $Q_{y,1}$ ）。同理，对于多个图像对中的其他图像对而言，通过重复上面从图像对 1 中确定有效像素点的步骤，也可以得到剩余图像对中的每一个图像对中的第一视频帧对应的目标状态矩阵、目标位移矩阵。比如，以目标视频内的视频帧号分别为 1、2、3、4、 $\dots$ 、n 的连续多个视频帧为例，则组成的多个图像对可以分别表示为 (1, 2)、(2, 3)、(3, 4)、 $\dots$ 、(n-1, n)。那么，根据每个图像对对应的光流追踪结果而言，可以通过上述有效像素点的判断方式最后得到图像对 (1, 2) 对应的目标状态矩阵 S_1 和图像对 (1, 2) 对应的目标位移矩阵 Q_1 （即前述目标水平位移矩阵 $Q_{x,1}$ 和目标垂直位移矩阵 $Q_{y,1}$ ）。依次类推，可以得到图像对 (n-1, n) 对应的目标状态矩阵 S_{n-1} 和图像对 (1, 2) 对应的目标位移矩阵 Q_{n-1} （即前述目标水平位移矩阵 $Q_{x,n-1}$ 和目标垂直位移矩阵 $Q_{y,n-1}$ ）。

进一步地，应用服务器可以通过云端位移积分法对上述图像对 1 对应的目标状态矩阵 S_1 和所述目标位移矩阵 Q_1 进行积分运算，以得到该第一视频帧中的像素点对应的状态积分矩阵 $S_{in}(x, y)$ 和位移积分矩阵 $Q_{in}(x, y)$ 。其中，位移积分矩阵 $Q_{in}(x, y)$ 可以包含横向位移积分矩阵 $Q_{x,in}(x, y)$ 和纵向位移积分矩阵 $Q_{y,in}(x, y)$ 。且该状态积分矩阵 $S_{in}(x, y)$ 、横向位移积分矩阵 $Q_{x,in}(x, y)$ 和纵向位移积分矩阵 $Q_{y,in}(x, y)$ 可以通过如下矩阵积分公式得到：

$$S_{in}(x, y) = \sum_{x' \leq x, y' \leq y} S(x', y') \quad \text{公式 (2)}$$

$$Q_{x,in}(x, y) = \sum_{x' \leq x, y' \leq y} Q_x(x', y') \quad \text{公式 (3)}$$

$$Q_{y,in}(x, y) = \sum_{x' \leq x, y' \leq y} Q_y(x', y') \quad \text{公式 (4)}$$

在公式 (2)、公式 (3)、公式 (4) 中的 x 和 y 可以用于表示与第一视频帧对应的状态积分矩阵、位移积分矩阵中的所有矩阵元素的坐标，比如 $S_{in}(x, y)$ 可以表示状态积分矩阵的第 x 行 y 列的矩阵元素的值。另外，在公式 (2)、公式 (3)、公式 (4) 中的 x' 和 y' 可以表示目标状态矩阵和目标位移矩阵中的矩阵元素的坐标，比如 $S(x', y')$ 可以表示目标状态矩阵的第 x' 行 y' 列的矩阵元素的值。

进一步地，应用服务器可以通过云端位移差分方法，在第一视频帧中选取一个高度为 M 和宽度为 N 的目标框，进而可以在该目标框内进一步对由公式

(2)、公式(3)、公式(4)所得到的这三个积分矩阵进行位移差分运算,以分别得到状态差分矩阵 $S_{\text{dif}}(x, y)$ 和位移差分矩阵 $Q_{\text{dif}}(x, y)$ 。所述目标框,是为了选取像素点周围一定区域的所有像素点,来计算平均位移。例如为 80×80 像素的大小。

其中,位移差分矩阵 $Q_{\text{dif}}(x, y)$ 可以包含横向位移差分矩阵 $Q_{x,\text{dif}}(x, y)$ 和纵向位移积分矩阵 $Q_{y,\text{dif}}(x, y)$ 。该状态积分矩阵 $S_{\text{dif}}(x, y)$ 、横向位移差分矩阵 $Q_{x,\text{dif}}(x, y)$ 和纵向位移积分矩阵 $Q_{y,\text{dif}}(x, y)$ 可以通过如下矩阵差分公式(5)得到。

$$\begin{aligned} S_{\text{dif}}(x, y) &= S_{\text{in}}\left(x + \frac{N}{2}, y + \frac{M}{2}\right) - S_{\text{in}}\left(x - \frac{N}{2}, y + \frac{M}{2}\right) - S_{\text{in}}\left(x + \frac{N}{2}, y - \frac{M}{2}\right) + S_{\text{in}}\left(x - \frac{N}{2}, y - \frac{M}{2}\right) \\ Q_{x,\text{dif}}(x, y) &= Q_{x,\text{in}}\left(x + \frac{N}{2}, y + \frac{M}{2}\right) - Q_{x,\text{in}}\left(x - \frac{N}{2}, y + \frac{M}{2}\right) - Q_{x,\text{in}}\left(x + \frac{N}{2}, y - \frac{M}{2}\right) + Q_{x,\text{in}}\left(x - \frac{N}{2}, y - \frac{M}{2}\right) \\ Q_{y,\text{dif}}(x, y) &= Q_{y,\text{in}}\left(x + \frac{N}{2}, y + \frac{M}{2}\right) - Q_{y,\text{in}}\left(x - \frac{N}{2}, y + \frac{M}{2}\right) - Q_{y,\text{in}}\left(x + \frac{N}{2}, y - \frac{M}{2}\right) + Q_{y,\text{in}}\left(x - \frac{N}{2}, y - \frac{M}{2}\right) \end{aligned}$$

其中,可以理解的是,本申请实施例可以将该第一视频帧中该目标框所在的区域称之为差分区域,从而可以基于该差分区域的尺寸信息,状态积分矩阵、横向位移积分矩阵和纵向位移积分矩阵,确定所述第一视频帧对应的平均位移矩阵。其中,即该位移差分运算公式中的 M 和 N 即为该差分区域的长度值和宽度值。其中,该位移差分运算公式中的 x 和 y 分别为第一视频帧中的每个像素点的位置信息。通过该位移差分运算公式,可以快速得到该该差分区域内的所有像素点的平均值。比如,对于状态积分矩阵而言,可以得到该状态积分矩阵 $S_{\text{in}}(x, y)$ 对应的状态差分矩阵 $S_{\text{dif}}(x, y)$ 。而对于横向位移积分矩阵 $Q_{x,\text{in}}(x, y)$ 和纵向位移积分矩阵 $Q_{y,\text{in}}(x, y)$ 而言,可以得到横向位移差分矩阵 $Q_{x,\text{dif}}(x, y)$ 和纵向位移差分矩阵 $Q_{y,\text{dif}}(x, y)$ 。

进一步地,该应用服务器可以将横向位移差分矩阵 $Q_{x,\text{dif}}(x, y)$ 与所述状态差分矩阵 $S_{\text{dif}}(x, y)$ 之间的比值确定为横向平均位移矩阵 $Q_{x,\text{F}}(x, y)$,并将所述纵向位移差分矩阵 $Q_{y,\text{dif}}(x, y)$ 与所述状态差分矩阵 $S_{\text{dif}}(x, y)$ 之间的比值确定为纵向平均位移矩阵 $Q_{y,\text{F}}(x, y)$ 。

$$Q_{x,\text{F}}(x, y) = \frac{Q_{x,\text{dif}}(x, y)}{S_{\text{dif}}(x, y) + e} \quad \text{公式(6)}$$

$$Q_{y,\text{F}}(x, y) = \frac{Q_{y,\text{dif}}(x, y)}{S_{\text{dif}}(x, y) + e} \quad \text{公式(7)}$$

在公式(6)和公式(7)中的 e 用于表示一个人为设定的比较小的数字,

比如 0.001。即在公式 (6) 和公式 (7) 中的 e 是为了避免当状态差分矩阵 $S_{dif}(x, y)$ 中的所有矩阵元素的值为 0 时, 从而可以避免直接除以 0, 进而可以进一步执行步骤 S203, 以在该目标用户终端中预先计算出该第一视频帧中的像素点出现在第二视频帧中的位置信息。

步骤 S203, 基于所述平均位移矩阵, 对所述第一视频帧中的像素点的位置信息进行跟踪, 并在所述第二视频帧中确定所跟踪得到的像素点的位置信息。

具体地, 应用服务器可以基于前述步骤 S203 所得到的平均位移矩阵 (该平均位移矩阵可以包含横向平均位移矩阵 $Q_{x,F}(x, y)$ 和纵向平均位移矩阵 $Q_{y,F}(x, y)$), 进一步对该第一视频帧中的像素点出现在下一视频帧 (即上述图像对 1 中的第二视频帧) 中的位置信息进行快速、且准确地跟踪, 即通过进行位移变换, 可以在所述第二视频帧中确定出对该第一视频帧中的像素点进行跟踪所得到的像素点的位置信息。

$$C_x(x, y) = x + Q_{x,F}(x, y) \quad \text{公式 (8)}$$

$$C_y(x, y) = y + Q_{y,F}(x, y) \quad \text{公式 (9)}$$

其中, 公式 (8) 中的 x 为该第一视频帧中的像素点的横向位置坐标, $Q_{x,F}(x, y)$ 为该第一视频帧对应的横向平均位移矩阵, 通过该公式 (8) 可以将第一视频帧中的像素点的横向位置坐标进行坐标变换, 以得到该第一视频帧中的像素点的下一视频帧中的横向位置坐标。同理, 公式 (9) 中的 y 为该第一视频帧中的像素点的纵向位置坐标, $Q_{y,F}(x, y)$ 为该第一视频帧对应的纵向平均位移矩阵, 通过该公式 (9) 可以将第一视频帧中的像素点的纵向位置坐标进行坐标变换, 以得到该第一视频帧中的像素点的下一视频帧中的纵向位置坐标。

可以理解的是, 对于每个图像对而言, 通过每个图像对中的第一视频帧对应的平均位移矩阵, 都可以将相应图像对中的第一视频帧中的像素点进行快速追踪, 从而可以在相应图像对中的第二视频帧中确定所追踪得到的像素点的位置坐标, 即可以在每个图像对的第二视频帧中确定所跟踪得到的像素点的位置信息。该应用服务器可以进一步对将每个图像对中所跟踪得到的像素点的位置信息进行存储, 以便于可以进一步执行步骤 S204。

步骤 S204, 基于所述第一视频帧中的像素点的位置信息、所述跟踪得到的像素点在所述第二视频帧中的位置信息, 生成所述目标视频相关联的轨迹信息。

其中, 所述轨迹信息中包含用于对目标视频中的目标对象所关联的多媒体

信息进行跟踪显示的目标轨迹信息。

步骤 S205，响应于对目标视频的触发操作，从所述目标视频的关键视频帧中确定目标像素点，并获取与所述目标像素点相关联的多媒体信息。

步骤 S206，基于所述目标像素点在所述关键视频帧中的位置信息，确定所述目标像素点对应的轨迹获取请求。

其中，步骤 S205 和步骤 S206 的具体实现方式可以参见上述图 4 所对应实施例中目标用户终端的描述，这里将不再继续进行赘述。

步骤 S207，响应于对关键视频帧中的目标像素点的轨迹获取请求，获取与目标视频相关联的轨迹信息。

具体地，应用服务器可以接收目标用户终端基于关键视频帧中的目标像素点发送的轨迹获取请求，并可以进一步获取上述应用服务器所预先计算得到的与该目标视频中的所有像素点相关联的轨迹信息，以便于进一步执行步骤 S208。

步骤 S208，从所述目标视频相关联的轨迹信息中，筛选与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，并返回所述目标轨迹信息。

具体地，应用服务器可以从该轨迹获取请求中得到该关键视频帧在目标视频中的视频帧号和目标像素点在该关键视频帧中的位置信息，从而可以进一步从该应用服务器预先得到的与该目标视频相关联的轨迹信息中筛选出与该目标像素点相关联的轨迹信息，并可以将筛选得到的轨迹信息称之为目标轨迹信息，从而可以进一步将该目标轨迹信息返回给目标用户终端，以使该目标用户终端可以基于该关键视频帧的帧号，快速从该接收到的目标轨迹信息中找出目标像素点出现在该关键视频帧的下一视频中的位置信息，直到得到该目标像素点出现在该关键视频帧后的每个视频帧中的位置信息，此时，该目标用户终端可以将该目标像素点出现在该关键视频帧后的每个视频帧中的位置信息所构成的新的轨迹信息。可选地，应当理解，该应用服务器在得到该关键视频帧的帧号时，可以快速从该筛选出的轨迹信息中找出目标像素点出现在该关键视频帧的下一视频中的位置信息，直到得到该目标像素点出现在该关键视频帧后的每个视频帧中的位置信息，此时，该应用服务器可以将该目标像素点出现在该关键视频帧后的每个视频帧中的位置信息所构成的新的轨迹信息称之为目标轨迹信息。

其中，可以理解的是，若目标用户终端的计算性能无法满足对象跟踪需求，

则该目标用户终端可以在生成该目标像素点对应的轨迹获取请求时，可以将该轨迹获取请求发送至应用服务器，以使应用服务器可以基于所述轨迹获取请求获取与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，并可以将获取到的目标轨迹信息返回给该目标用户终端；

可选地，若目标用户终端的计算性能能够满足对象跟踪需求，则该目标用户终端可以在该目标用户终端中执行上述步骤 S201-步骤 S204，以在该目标用户终端中预先对该目标视频内的所有像素点的进行全图像素跟踪，以预先得到该目标视频内的所有像素点在每个视频帧中的位置信息，进而可以对该目标视频内的每个像素点在每个视频帧中的位置信息进行位置整合，以得到该目标视频内的每个像素点对应的轨迹信息。此时，该目标用户终端可以直接基于所述目标对象中的目标像素点在所述关键视频帧中的位置信息，在该目标用户终端中获取与该目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，以便于可以进一步执行步骤 S209。

其中，所述目标轨迹信息包含所述目标像素点在所述关键视频帧的下一视频帧中的位置信息；所述关键视频帧的下一视频帧中的位置信息是通过跟踪所述目标对象得到的。

比如，以当前正在播放的视频（即目标视频）所包含的多个连续视频帧为以下 6 个视频帧为例，这 6 个视频帧可以为视频帧 a、视频帧 b、视频帧 c、视频帧 d、视频帧 e、视频帧 f。因此，应用服务器在视频预处理阶段，可以预先对该目标视频内的每个视频帧进行预处理，即可以基于上述光流追踪规则确定该目标视频内的每两个相邻视频帧所构成的图像对对应的平均位移矩阵，进而可以根据每个图像对对应的平均位移矩阵（也可以称之为根据每个图像对中第一视频帧对应的平均位移矩阵）对该第一视频帧中的所有像素点进行跟踪，以得到该第一视频帧中的所有像素点出现在第二视频帧中的位置信息，进而可以得到该目标视频的所有像素点在每个视频帧（即可以得到该目标视频的所有像素点在上述视频帧 a、视频帧 b、视频帧 c、视频帧 d、视频帧 e、视频帧 f）中的位置信息，即可以基于该目标视频的所有像素点在每个视频帧中的位置信息，得到该目标视频的所有像素点对应的轨迹信息。该目标视频的所有像素点对应的轨迹信息称之为与该目标视频相关联的轨迹信息。

又比如，应用服务器在对该目标视频进行预处理时，该应用服务器可以预

先计算得到的该目标视频中的像素点 A（例如，该像素点 A 可以为该该目标视频中的所有像素点中的一个像素点）对应的轨迹信息，若该像素点 A 对应的轨迹信息中包含该像素点 A 在该目标视频的每个视频帧（即在上述视频帧 a、视频帧 b、视频帧 c、视频帧 d、视频帧 e、视频帧 f）中的位置信息，则在目标用户终端中的目标像素点对应的关键视频帧为该目标视频的视频帧 c 时，可以进一步将该视频帧 c 中的目标对象内的像素点 A 作为目标像素点，进而可以在该目标像素点为像素点 A 时，可以从应用服务器中筛选出该像素点 A 的轨迹信息，进而可以基于该筛选出的像素点 A 的轨迹信息得到该像素点 A 在该关键视频帧后的每个视频帧（即视频帧 d、视频帧 e、视频帧 f）中的位置信息。

在本申请实施例中，若关键视频帧为该目标视频内的首个视频帧，则目标用户终端所获取到的目标轨迹信息可以为预先计算得到的轨迹信息。例如，以前述关键视频帧为目标视频内的视频帧 a（首个视频帧）为例，则该目标用户终端所获取到的目标轨迹信息可以包含该目标像素点在上述视频帧 a、视频帧 b、视频帧 c、视频帧 d、视频帧 e、视频帧 f 的位置信息。可选地，若关键视频帧为该目标视频内的非首个视频帧，则目标用户终端所获取到的目标轨迹信息可以为从预先计算得到的轨迹信息中所确定出的部分位置信息所构成的轨迹信息，例如，以前述关键视频帧为目标视频内的视频帧 c（即非首个视频帧）为例，则该目标用户终端所获取到的目标轨迹信息可以包含该目标像素点在视频帧 d、视频帧 e、视频帧 f 中的位置信息，并可以将目标像素点在视频帧 d、视频帧 e、视频帧 f 中的位置信息称之为部分位置信息。

可以理解的是，可选地，该目标用户终端还可以将该应用服务器所查找到的包含该目标像素点的在关键视频帧中的位置信息的轨迹信息（即上述像素点 A 对应的轨迹信息）统称为目标轨迹信息，此时，该目标轨迹信息可以视为从该目标视频的所有像素点中所查找到与目标像素点相匹配的像素点 A 对应的轨迹信息。由于该轨迹信息中可以包含该像素点 A 在该目标视频的每个视频帧中的位置信息，自然，也可以快速从该轨迹信息中得到该目标像素点在该关键视频后的每个视频帧中的位置信息。

步骤 S209，当播放所述关键视频帧的下一视频帧时，基于所述目标轨迹信息中的所述目标像素点在所述关键视频帧的下一视频帧中的位置信息显示所述多媒体信息。

其中，可以理解的是，本申请实施例可以在获取到目标用户所选取的目标对象中的目标像素点时，可以从这些预先计算好的所有像素点对应的轨迹信息中筛选出该关键视频帧中的目标像素点的位置信息相关联的轨迹信息，进而可以将筛选出的轨迹信息称之为目标轨迹信息。由于本申请实施例可以预先对该视频内的每个视频帧中的像素点进行像素跟踪，因此，在获取到每个图像对中的第一视频帧对应的平均位移矩阵时，可以快速得到该视频内的每个像素点在相应视频帧中的位置信息。应当理解，预先计算得到的每个像素点在相应视频帧中的位置信息，可以用于表征在当前视频播放界面中所播放的该视频内的每个像素点在相应视频帧中的位置信息。因此，当在目标用户终端获取到目标对象中的目标像素点以及目标对象相关联的多媒体信息时，可以快速将从所有像素点对应的轨迹信息中所筛选出的该目标像素点对应的轨迹信息称之为目标轨迹信息，进而可以将该目标轨迹信息返回给目标用户终端，以使该目标用户终端可以基于该目标轨迹信息中所携带的该目标像素点在关键视频帧之后的每个视频帧中的位置信息，对该目标对象所关联的多媒体信息（例如，弹幕数据）进行跟踪显示。

可以理解的是，若该目标像素点在该关键视频帧后的连续多个视频帧中的位置信息所构成的轨迹为一个圈，则该弹幕数据在该目标用户终端中可以跟踪这个轨迹进行转圈显示。

为便于理解，请参见图 10，是本申请实施例提供的一种基于轨迹信息显示弹幕数据的示意图。如图 10 所示的视频帧 100 中可以包含多个对象，例如，可以包含图 10 所示的对象 1、对象 2 和对象 3。若目标用户在目标用户终端中将图 10 所示的对象 1 作为目标对象时，可以将该视频帧 100 称之为关键视频帧，并可以将该目标对象中与该目标用户所执行的触发操作对应的像素点称为目标像素点。若该目标用户终端具有较强的计算性能时，则可以在该目标用户终端中预先计算好该目标视频中的每个像素点在每个视频帧中的位置信息，从而可以在该目标用户终端中得到与该目标视频相关联的轨迹信息。例如，可以得到预先计算得到图 10 所示的轨迹信息 1，即该轨迹信息 1 中的位置信息均是由该目标视频中的像素点在该目标视频中的每个视频帧中的位置信息所确定的。因此，该目标用户终端可以快速从基于该对象 1 中的目标像素点的位置信息，将图 10 所示的轨迹信息 1 视为目标轨迹信息，从而可以基于该轨迹信息 1 中的对

象 1 在该关键视频帧后的每个视频帧(即图 10 所示的视频帧 200 和视频帧 300) 中的位置信息,快速对该目标对象(即对象 1)所关联的多媒体信息(即图 10 所示的弹幕数据 1 为 BBBBB)进行跟踪和显示。即显示在图 10 所示的视频帧 200 和视频帧 300 中的弹幕数据均是由图 10 所示的轨迹信息系 1 中的位置信息所决定的。

可以理解的是,图 10 所示的目标视频相关联的轨迹信息还可以为上述应用服务器所预先计算好的,从而可以在应用服务器接收到上述针对对象 1 中的目标像素点轨迹获取请求时,也可以快速从图 10 所示的目标视频相关联的轨迹信息中获取该关键视频帧中的目标像素点的位置信息相关联的轨迹信息,即通过对目标视频中的所有像素点的全图像素跟踪放在应用服务器中执行,可以有效地减小目标用户终端的计算量,从而可以确保目标用户终端在得到图 10 所示的轨迹信息 1 时,可以基于该轨迹信息 1 中的位置信息对图 10 所示的弹幕数据 1 进行快速跟踪和显示,从而可以提升弹幕数据的显示的灵活度。应当理解,视频播放界面中可以存在与该目标对象相关联的多个弹幕,但是当目标用户终端检测到该视频播放界面中所存在的多条弹幕存在重合时,可以对这些存在重合的弹幕进行合并,以在这些存在重合的弹幕中保留该目标用户终端最新得到的弹幕。

本申请实施例可以在获取到目标用户针对目标视频的触发操作时,将该目标视频中该触发操作对应的视频帧称之为关键视频帧,从而可以从该关键视频帧中确定目标像素点,并获取与该目标像素点以及该目标像素点所在目标对象关联的多媒体信息(例如,该多媒体信息可以为该目标视频中的用户文字、图片、表情等弹幕数据);进一步地,基于目标像素点在该关键视频帧中的位置信息确定所述目标像素点对应的轨迹获取请求,进而可以基于该轨迹获取请求获取所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息,以便于在播放该关键视频帧的下一视频帧时,能够基于该目标轨迹信息显示与该目标像素点以及该目标像素点所在目标对象相关联的弹幕数据。由此可见,本申请实施例可以在确定出关键视频帧时,进一步从该关键视频帧中的所有像素点的轨迹信息中筛选出目标像素点的轨迹信息,并将筛选出的目标像素点的轨迹信息称之为目标轨迹信息,以便于能够基于得到的目标轨迹信息丰富弹幕数据的展示效果,比如,对于不同目标对象中的目标像素点,所得到的目标轨迹

信息可能会不同，进而使得弹幕数据的展示效果会不同。此外，基于目标对象与弹幕数据之间的关联关系，可以快速确定出该弹幕数据在该关键视频帧之后的每个视频帧中的位置信息，换言之，该弹幕数据会在该目标视频中一直跟着该目标对象进行变动，进而可以丰富视频中的用户文字的视觉展示效果，并可以使得弹幕数据与目标对象或评论的视频中的对象之间的关联性更强。

进一步地，请参见图 11，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。如图 11 所示，该视频数据处理装置 1 可以应用于上述图 1 所对应实施例中的目标用户终端。该视频数据处理装置 1 可以包含：对象确定模块 1101，请求确定模块 1102，轨迹获取模块 1103，文本显示模块 1104；

对象确定模块 1101，用于响应于对目标视频的触发操作，从所述目标视频的关键视频帧中确定目标像素点，并获取与所述目标像素点相关联的多媒体信息，其中，所述关键视频帧是所述触发操作所在的视频帧，所述目标像素点是所述关键视频帧中与所述触发操作对应的像素点；

请求确定模块 1102，用于基于所述目标像素点在所述关键视频帧中的位置信息，确定所述目标像素点对应的轨迹获取请求；

轨迹获取模块 1103，用于基于所述轨迹获取请求，获取与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，其中，所述目标轨迹信息包含所述目标像素点在所述关键视频帧的下一视频帧中的位置信息，所述目标像素点在所述关键视频帧的下一视频帧中的位置信息是通过跟踪所述目标像素点得到的；

文本显示模块 1104，用于当播放所述关键视频帧的下一视频帧时，基于所述目标轨迹信息中的所述目标像素点在所述关键视频帧的下一视频帧中的位置信息，显示所述多媒体信息。

其中，对象确定模块 1101，请求确定模块 1102，轨迹获取模块 1103，文本显示模块 1104 的具体执行方式可参见上述图 4 所对应实施例中对步骤 S101-步骤 S104 的描述，这里将不再继续进行赘述。

本申请实施例可以在获取到目标用户针对目标视频的触发操作时，将该目标视频中该触发操作对应的视频帧作为关键视频帧，从而可以从该关键视频帧中确定目标像素点，并获取与该目标像素点以及该目标像素点所在目标对象关联的多媒体信息（例如，该多媒体信息可以为该目标视频中的用户文字、图片、

表情等弹幕数据)。进一步地,基于目标像素点在该关键视频帧中的位置信息确定所述目标像素点对应的轨迹获取请求,进而可以基于该轨迹获取请求获取所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息,以便于在播放该关键视频帧的下一视频帧时,能够基于该目标轨迹信息显示与该目标像素点以及该目标像素点所在目标对象相关联的弹幕数据。由此可见,本申请实施例可以在确定出关键视频帧时,进一步从该关键视频帧中的所有像素点的轨迹信息中筛选出目标像素点的轨迹信息,并将筛选出的目标像素点的轨迹信息作为目标轨迹信息,以便于能够基于得到的目标轨迹信息丰富弹幕数据的展示效果。比如,对于不同目标对象中的目标像素点,所得到的目标轨迹信息可能会不同,进而使得弹幕数据的展示效果会不同。此外,基于目标对象与弹幕数据之间的关联关系,可以快速确定出该弹幕数据在该关键视频帧之后的每个视频帧中的位置信息,换言之,该弹幕数据会在该目标视频中一直跟着该目标对象进行变动,进而可以丰富视频中的用户文字的视觉展示效果,并可以使得弹幕数据与目标对象或评论的视频中的对象之间的关联性更强。

进一步地,请参见图 12,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 12 所示,计算机设备 1000 可以为上述图 1 对应实施例中的目标用户终端。上述计算机设备 1000 可以包括:处理器 1001,网络接口 1004 和存储器 1005。此外,上述计算机设备 1000 还可以包括:用户接口 1003,和至少一个通信总线 1002。其中,通信总线 1002 用于实现这些组件之间的连接通信。其中,用户接口 1003 可以包括显示屏(Display)、键盘(Keyboard)。可选用户接口 1003 还可以包括标准的有线接口、无线接口。网络接口 1004 可选的可以包括标准的有线接口、无线接口(如 WI-FI 接口)。存储器 1004 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。存储器 1005 可选的还可以是至少一个位于远离前述处理器 1001 的存储装置。如图 12 所示,作为一种计算机存储介质的存储器 1005 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。

在图 12 所示的计算机设备 1000 中,网络接口 1004 可提供网络通讯功能;而用户接口 1003 主要用于为用户提供输入的接口;而处理器 1001 可以用于调用存储器 1005 中存储的设备控制应用程序,以实现:

响应于对目标视频的触发操作,从所述目标视频的关键视频帧中确定目标

像素点，并获取与所述目标像素点相关联的多媒体信息，其中，所述关键视频帧是所述触发操作所在的视频帧，所述目标像素点是所述关键视频帧中与所述触发操作对应的像素点；

基于所述目标像素点在所述关键视频帧中的位置信息，确定所述目标像素点对应的轨迹获取请求；

基于所述轨迹获取请求，获取与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息；所述目标轨迹信息包含所述目标像素点在所述关键视频帧的下一视频帧中的位置信息；所述目标像素点在所述关键视频帧的下一视频帧中的位置信息是通过跟踪所述目标像素点得到的；

当播放所述关键视频帧的下一视频帧时，基于所述目标轨迹信息中的所述目标像素点在所述关键视频帧的下一视频帧中的位置信息，显示所述多媒体信息。

应当理解，本申请实施例中所描述的计算机设备 1000 可执行前文图 4 所对应实施例中对上述视频数据处理方法的描述，也可执行前文图 11 所对应实施例中对上述视频数据处理装置 1 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存储介质，且上述计算机存储介质中存储有前文提及的视频数据处理装置 1 所执行的计算机程序，且上述计算机程序包括程序指令，当上述处理器执行上述程序指令时，能够执行前文图 4 所对应实施例中对上述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。

进一步地，请参见图 13，是本申请实施例提供的另一种视频数据处理装置的结构示意图。如图 13 所示，该视频数据处理装置 2 可以应用于上述图 8 所对应实施例中的应用服务器，该应用服务器可以为上述图 1 所对应实施例中的业务服务器 2000。该视频数据处理装置 2 可以包含：请求响应模块 1301 和轨迹筛选模块 1302；

请求响应模块 1301，用于响应于对关键视频帧中的目标像素点的轨迹获取请求，获取与目标视频相关联的轨迹信息，其中，所述关键视频帧是所述目标

视频中的视频帧，所述目标像素点是所述关键视频帧中的像素点，所述轨迹信息是由所述目标视频中的每个视频帧中的像素点位置信息所确定的；

轨迹筛选模块 1302，用于从所述目标视频相关联的轨迹信息中，筛选与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，并返回所述目标轨迹信息；所述目标轨迹信息包含目标位置信息；所述目标位置信息用于触发在所述关键视频帧的下一视频帧中显示与所述目标像素点相关联的多媒体信息。

其中，请求响应模块 1301 和轨迹筛选模块 1302 的具体实现方式可以参见上述图 8 所对应实施例中对步骤 S207 和步骤 S208 的描述，这里将不再继续进行赘述。

进一步地，请参见图 14，是本申请实施例提供的另一种计算机设备的结构示意图。如图 14 所示，计算机设备 2000 可以为上述图 1 对应实施例中的目标业务服务器 2000。上述计算机设备 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。其中，用户接口 2003 可以包括显示屏（Display）、键盘（Keyboard），可选用户接口 2003 还可以包括标准的有线接口、无线接口。网络接口 2004 可选的可以包括标准的有线接口、无线接口（如 WI-FI 接口）。存储器 2004 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。存储器 2005 可选的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 14 所示，作为一种计算机存储介质的存储器 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。

在图 14 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，以实现：

响应于对关键视频帧中的目标像素点的轨迹获取请求，获取与目标视频相关联的轨迹信息，其中，所述关键视频帧是所述目标视频中的视频帧，所述目标像素点是所述关键视频帧中的像素点，所述轨迹信息是由所述目标视频的每个视频帧中的像素点位置信息所确定的；

从所述目标视频相关联的轨迹信息中，筛选与所述目标像素点在所述关键

视频帧中的位置信息相关联的目标轨迹信息，并返回所述目标轨迹信息；所述目标轨迹信息包含目标位置信息；所述目标位置信息用于触发在所述关键视频帧的下一视频帧中显示与所述目标像素点相关联的多媒体信息。

应当理解，本申请实施例中所描述的计算机设备 2000 可执行前文图 8 所对应实施例中对上述视频数据处理方法的描述，也可执行前文图 13 所对应实施例中对上述视频数据处理装置 2 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存储介质，且上述计算机存储介质中存储有前文提及的视频数据处理装置 2 所执行的计算机程序，且上述计算机程序包括程序指令，当上述处理器执行上述程序指令时，能够执行前文图 8 所对应实施例中对上述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。

进一步地，请参见图 15，是本申请实施例提供的又一种视频数据处理装置的结构示意图。如图 15 所示，该视频数据处理装置 3 可以应用于上述图 1 所对应实施例中的业务服务器 2000，也可以应用于上述图 1 所对应实施例中的目标用户终端。该视频数据处理装置 3 可以包含：第一获取模块 310、矩阵获取模块 410、位置跟踪模块 510、轨迹生成模块 610；

第一获取模块 310，用于从目标视频中获取相邻的第一视频帧和第二视频帧；

矩阵获取模块 410，用于基于所述目标视频对应的光流追踪规则、所述第一视频帧中的像素点、所述第二视频帧中的像素点，确定所述第一视频帧对应的平均位移矩阵；

其中，所述矩阵获取模块 410 包括：第一确定单元 4001、矩阵确定单元 4002、像素点筛选单元 4003、矩阵修正单元 4004、第二确定单元 4005；

第一确定单元 4001，用于获取所述目标视频对应的光流追踪规则，并将所述第一视频帧中的像素点的位置信息确定为第一位置信息，并将所述第二视频帧中的像素点的位置信息确定为第二位置信息；

矩阵确定单元 4002，用于基于所述光流追踪规则、所述第一视频帧中的像

素点的第一位置信息、所述第二视频帧中的像素点的第二位置信息，获取所述第一视频帧对应的正向位移矩阵，并获取所述第二视频帧对应的反向位移矩阵；

其中，其中，矩阵确定单元 4002 包括：第一追踪子单元 4021 和第二追踪子单元 4022；

第一追踪子单元 4021，用于基于所述第一视频帧中的像素点的第一位置信息和所述光流追踪规则，将所述第一视频帧中的像素点正向映射到所述第二视频帧，并在所述第二视频帧中确定所映射得到的第一映射点的第二位置信息，并基于所述像素点的第一位置信息、所述第一映射点的第二位置信息确定所述第一视频帧对应的正向位移矩阵；

第二追踪子单元 4022，基于所述第二视频帧中的像素点的第二位置信息和所述光流追踪规则，将所述第二视频帧中的第一映射点反向映射到所述第一视频帧，并在所述第一视频帧中确定所映射得到第二映射点的第三位置信息，并基于所述第一映射点的第二位置信息、所述第二映射点的第三位置信息确定所述第二视频帧对应的反向位移矩阵。

其中，第一追踪子单元 4021 和第二追踪子单元 4022 的具体实现方式可以参见上述图 8 所对应实施例中对云端前向后向光流法的描述，这里将不再继续进行赘述。

像素点筛选单元 4003，用于基于所述第一视频帧中的像素点的第一位置信息、所述正向位移矩阵、所述反向位移矩阵，将所述像素点中满足目标筛选条件的像素点确定为有效像素点；

其中，所述像素点筛选单元 4003 包括：第一位置确定子单元 4031、第二位置确定子单元 4032、第三位置确定子单元 4033、误差确定子单元 4034 和有效筛选子单元 4035；

第一位置确定子单元 4031，用于从所述第一视频帧中的像素点中获取第一像素点，并在所述第一视频帧中确定所述第一像素点的第一位置信息，并从所述正向位移矩阵中确定所述第一像素点对应的第一横向位移和第一纵向位移；

第二位置确定子单元 4032，用于基于所述第一像素点的第一位置信息、所述第一像素点对应的第一横向位移和第一纵向位移，将所述第一像素点正向映射到所述第二视频帧，并在所述第二视频帧中确定所映射得到的第二像素点的第二位置信息；

第三位置确定子单元 4033, 用于从所述反向位移矩阵中确定所述第二像素点对应的第二横向位移和第二纵向位移, 并基于所述第二像素点的第二位置信息、所述第二像素点对应的第二横向位移和第二纵向位移, 将所述第二像素点反向映射到所述第一视频帧, 并在所述第一视频帧中确定所映射得到的第三像素点的第三位置信息;

误差确定子单元 4034, 用于基于所述第一像素点的第一位置信息、所述第三像素点的第三位置信息, 确定所述第一像素点与所述第三像素点之间的误差距离, 并根据所述第一像素点的第一位置信息、所述第二像素点的第二位置信息, 确定包含第一像素点的图像块与包含所述第二像素点的图像块之间的相关系数;

有效筛选子单元 4035, 用于将所述像素点中误差距离小于误差距离阈值、且所述相关系数大于等于相关系数阈值的像素点确定为有效像素点。

其中, 第一位置确定子单元 4031、第二位置确定子单元 4032、第三位置确定子单元 4033、误差确定子单元 4034 和有效筛选子单元 4035 的具体实现方式可以参见上述图 8 所对应实施例中对确定有效像素点的具体过程的描述, 这里将不再继续进行赘述。

矩阵修正单元 4004, 用于基于所述有效像素点对所述第一视频帧对应的初始状态矩阵和所述正向位移矩阵进行修正, 得到所述第一视频帧对应的目标状态矩阵和目标位移矩阵;

其中, 所述矩阵修正单元 4004 包括: 初始获取子单元 4041、数值切换子单元 4042、位移设置子单元 4043;

初始获取子单元 4041, 获取第一视频帧对应的初始状态矩阵; 所述初始状态矩阵中的每个矩阵元素的状态值均为第一数值, 一个矩阵元素对应所述像素点中的一个像素点;

数值切换子单元 4042, 用于在所述初始状态矩阵中将与所述有效像素点对应的矩阵元素的状态值由第一数值切换为第二数值, 并将包含第二数值的初始状态矩阵确定为所述第一视频帧对应的目标状态矩阵;

位移设置子单元 4043, 用于在所述正向位移矩阵中将所述剩余像素点对应的矩阵元素的位移设置为所述第一数值, 并将包含所述第一数值的正向位移矩阵确定为目标位移矩阵; 所述剩余像素点为所述像素点中除所述有效像素点之

外的像素点。

其中，所述位移设置子单元 4043，具体用于若所述正向位移矩阵包含初始横向位移矩阵和初始纵向位移矩阵，则在所述初始横向位移矩阵中将所述剩余像素点对应的矩阵元素的第一横向位移设置为所述第一数值，并将包含所述第一数值的初始横向位移确定为所述第一视频帧对应的横向位移矩阵；

所述位移设置子单元 4043，还具体用于在所述初始纵向位移矩阵将所述剩余像素点对应的矩阵元素的第一纵向位移设置为所述第一数值，并将包含所述第一数值的初始纵向位移确定为所述第一视频帧对应的纵向位移矩阵；

所述位移设置子单元 4043，还具体用于将所述第一视频帧对应的横向位移矩阵和所述第一视频帧对应的纵向位移矩阵确定为目标位移矩阵。

初始获取子单元 4041、数值切换子单元 4042、位移设置子单元 4043 的具体实现方式可以参见上述图 8 所对应实施例中对修正初始状态矩阵和正向位移矩阵的描述，这里将不再继续进行赘述。

第二确定单元 4005，用于基于所述目标状态矩阵和所述目标位移矩阵，确定所述第一视频帧对应的平均位移矩阵。

其中，所述第二确定单元 4005 包括：第一积分子单元 4051、第二积分子单元 4052、第三积分子单元 4053、差分运算子单元 4054；

第一积分子单元 4051，用于在所述第一视频帧中对所述目标状态矩阵进行位移积分运算，得到所述第一视频帧中的像素点对应的状态积分矩阵；

第二积分子单元 4052，用于在所述第一视频帧中对所述目标状态矩阵中的横向位移矩阵进行位移积分运算，得到所述第一视频帧中的像素点对应的横向位移积分矩阵；

第三积分子单元 4053，用于在所述第一视频帧中对所述目标状态矩阵中的纵向位移矩阵进行位移积分运算，得到所述第一视频帧中的像素点对应的纵向位移积分矩阵；

差分运算子单元 4054，用于从所述第一视频帧中确定位移差分运算对应的差分区域，基于所述差分区域的尺寸信息、状态积分矩阵、横向位移积分矩阵和纵向位移积分矩阵，确定所述第一视频帧对应的平均位移矩阵。

其中，所述差分运算子单元 4054 包括：第一差分子单元 4055、第二差分子单元 4056、第三差分子单元 4057、平均确定子单元 4058；

第一差分子单元 4055, 用于基于所述差分区域对应的长度信息和宽度信息, 对所述状态积分矩阵进行位移差分运算, 得到所述第一图像帧对应的状态差分矩阵;

第二差分子单元 4056, 用于基于所述差分区域对应的长度信息和宽度信息, 分别对所述横向位移积分矩阵和纵向位移积分矩阵进行位移差分运算, 得到所述第一图像帧对应的横向位移差分矩阵和纵向位移差分矩阵;

第三差分子单元 4057, 用于将所述横向位移差分矩阵与所述状态差分矩阵之间的比值确定为横向平均位移矩阵, 并将所述纵向位移差分矩阵与所述状态差分矩阵之间的比值确定为纵向平均位移矩阵;

平均确定子单元 4058, 用于将所述纵向位移差分矩阵和所述纵向平均位移矩阵确定为所述第一视频帧对应的平均位移矩阵。

其中, 第一积分子单元 4051、第二积分子单元 4052、第三积分子单元 4053、差分运算符单元 4054 的具体实现方式可以参见上述图 8 所对应实施例中对云端位移积分方法和云端位移差分方法的描述, 这里将不再继续进行赘述。

其中, 第一确定单元 4001、矩阵确定单元 4002、像素点筛选单元 4003、矩阵修正单元 4004、第二确定单元 4005 的具体实现方式可以参见上述图 8 所对应实施例中对步骤 S202 的描述, 这里将不再继续进行赘述。

位置跟踪模块 510, 用于基于所述平均位移矩阵对所述第一视频帧中的像素点的位置信息进行跟踪, 并在所述第二视频帧中确定所跟踪得到的像素点的位置信息;

轨迹生成模块 610, 用于基于所述第一视频帧中的像素点的位置信息、所述跟踪得到的像素点在所述第二视频帧中的位置信息, 生成所述目标视频相关联的轨迹信息; 所述轨迹信息中包含用于对目标视频中的目标像素点所关联的多媒体信息进行跟踪显示的目标轨迹信息。

其中, 第一获取模块 310、矩阵获取模块 410、位置跟踪模块 510、轨迹生成模块 610 的具体实现方式可以参见上述图 8 所对应实施例中对步骤 S201-步骤 S204 的描述, 这里将不再继续进行赘述。

进一步地, 请参见图 16, 是本申请实施例提供的又一种计算机设备的结构示意图。如图 16 所示, 上述计算机设备 3000 可以应用于上述图 1 对应实施例中的业务服务器 2000。上述计算机设备 3000 可以包括: 处理器 3001, 网络接

口 3004 和存储器 3005, 此外, 上述视频数据处理装置 3000 还可以包括: 用户接口 3003, 和至少一个通信总线 3002。其中, 通信总线 3002 用于实现这些组件之间的连接通信。其中, 用户接口 3003 可以包括显示屏 (Display)、键盘 (Keyboard), 可选用户接口 3003 还可以包括标准的有线接口、无线接口。网络接口 3004 可选的可以包括标准的有线接口、无线接口 (如 WI-FI 接口)。存储器 3004 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。存储器 3005 可选的还可以是至少一个位于远离前述处理器 3001 的存储装置。如图 16 所示, 作为一种计算机存储介质的存储器 3005 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。

在图 16 所示的计算机设备 3000 中, 网络接口 3004 可提供网络通讯功能; 而用户接口 3003 主要用于为用户提供输入的接口; 而处理器 3001 可以用于调用存储器 3005 中存储的设备控制应用程序, 以实现:

从目标视频中获取相邻的第一视频帧和第二视频帧;

基于所述目标视频对应的光流追踪规则、所述第一视频帧中的像素点、所述第二视频帧中的像素点, 确定所述第一视频帧对应的平均位移矩阵;

基于所述平均位移矩阵对所述第一视频帧中的像素点的位置信息进行跟踪, 并在所述第二视频帧中确定所跟踪得到的像素点的位置信息;

基于所述第一视频帧中的像素点的位置信息、所述跟踪得到的像素点在所述第二视频帧中的位置信息, 生成所述目标视频相关联的轨迹信息; 所述轨迹信息中包含用于对目标视频中的目标像素点所关联的多媒体信息进行跟踪显示的目标轨迹信息。

应当理解, 本申请实施例中所描述的计算机设备 3000 可执行前文图 8 所对应实施例中对上述视频数据处理方法的描述, 也可执行前文图 15 所对应实施例中对上述视频数据处理装置 3 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存储介质, 且上述计算机存储介质中存储有前文提及的视频数据处理装置 3 所执行的计算机程序, 且上述计算机程序包括程序指令, 当上述处理器执行上述程序指令时, 能够执行前文图 8 所对应实施例中对上述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。

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，上述的程序可存储于一计算机可读取存储介质中，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，上述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory, ROM）或随机存储记忆体（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，所述方法应用于计算机设备，其特征在于，包括：

响应于对目标视频的触发操作，从所述目标视频的关键视频帧中确定目标像素点，并获取与所述目标像素点相关联的多媒体信息，其中，所述关键视频帧是所述触发操作所在的视频帧，所述目标像素点是所述关键视频帧中与所述触发操作对应的像素点；

基于所述目标像素点在所述关键视频帧中的位置信息，确定所述目标像素点对应的轨迹获取请求；

基于所述轨迹获取请求，获取与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，其中，所述目标轨迹信息包含所述目标像素点在所述关键视频帧的下一视频帧中的位置信息，所述目标像素点在所述关键视频帧的下一视频帧中的位置信息是通过跟踪所述目标像素点得到的；

当播放所述关键视频帧的下一视频帧时，基于所述目标轨迹信息中的所述目标像素点在所述关键视频帧的下一视频帧中的位置信息，显示所述多媒体信息。

2、一种视频数据处理方法，所述方法应用于业务服务器，其特征在于，包括：

响应于对关键视频帧中的目标像素点的轨迹获取请求，获取与目标视频相关联的轨迹信息，其中，所述关键视频帧是所述目标视频中的视频帧，所述目标像素点是所述关键视频帧中的像素点，所述轨迹信息是由所述目标视频的每个视频帧中的像素点的位置信息所确定的；

从所述目标视频相关联的轨迹信息中，筛选与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，并返回所述目标轨迹信息，其中，所述目标轨迹信息包含目标位置信息，所述目标位置信息用于触发在所述关键视频帧的下一视频帧中，显示与所述目标像素点相关联的多媒体信息。

3、一种视频数据处理方法，其特征在于，包括：

从目标视频中获取相邻的第一视频帧和第二视频帧；

基于所述目标视频对应的光流追踪规则、所述第一视频帧中的像素点、所述第二视频帧中的像素点，确定所述第一视频帧对应的平均位移矩阵；

基于所述平均位移矩阵，对所述第一视频帧中的像素点的位置信息进行跟踪，并在所述第二视频帧中确定所跟踪得到的像素点的位置信息；

基于所述第一视频帧中的像素点的位置信息、所述跟踪得到的像素点在所述第二视频帧中的位置信息，生成与所述目标视频相关联的轨迹信息，其中，所述轨迹信息中包含用于对目标视频中的目标像素点所关联的多媒体信息进行跟踪显示的目标轨迹信息。

4、根据权利要求3所述的方法，其特征在于，所述基于所述目标视频对应的光流追踪规则、所述第一视频帧中的像素点、所述第二视频帧中的像素点，确定所述第一视频帧对应的平均位移矩阵，包括：

获取所述目标视频对应的光流追踪规则，并将所述第一视频帧中的像素点的位置信息确定为第一位置信息，并将所述第二视频帧中的像素点的位置信息确定为第二位置信息；

基于所述光流追踪规则、所述第一视频帧中的像素点的第一位置信息、所述第二视频帧中的像素点的第二位置信息，获取所述第一视频帧对应的正向位移矩阵，并获取所述第二视频帧对应的反向位移矩阵；

基于所述第一视频帧中的像素点的第一位置信息、所述正向位移矩阵、所述反向位移矩阵，将所述像素点中满足目标筛选条件的像素点确定为有效像素点；

基于所述有效像素点对所述第一视频帧对应的初始状态矩阵和所述正向位移矩阵进行修正，得到所述第一视频帧对应的目标状态矩阵和目标位移矩阵；

基于所述目标状态矩阵和所述目标位移矩阵，确定所述第一视频帧对应的平均位移矩阵。

5、根据权利要求4所述的方法，其特征在于，所述基于所述光流追踪规则、所述第一视频帧中的像素点的第一位置信息、所述第二视频帧中的像素点的第二位置信息，获取所述第一视频帧对应的正向位移矩阵，并获取所述第二视频帧对应的反向位移矩阵，包括：

基于所述第一视频帧中的像素点的第一位置信息和所述光流追踪规则，将所述第一视频帧中的像素点正向映射到所述第二视频帧，并在所述第二视频帧中确定所映射得到的第一映射点的第二位置信息，并基于所述像素点的第一位置信息、所述第一映射点的第二位置信息确定所述第一视频帧对应的正向位移矩阵；

基于所述第二视频帧中的像素点的第二位置信息和所述光流追踪规则，将所述第二视频帧中的像素点反向映射到所述第一视频帧，并在所述第一视频帧中确定所映射得到第二映射点的第三位置信息，并基于所述第一映射点的第二位置信息、所述第二映射点的第三位置信息确定所述第二视频帧对应的反向位移矩阵。

6、根据权利要求4所述的方法，其特征在于，所述基于所述第一视频帧中的像素点的第一位置信息、所述正向位移矩阵、所述反向位移矩阵，将所述像素点中满足目标筛选条件的像素点确定为有效像素点，包括：

从所述第一视频帧中的像素点中获取第一像素点，并在所述第一视频帧中确定所述第一像素点的第一位置信息，并从所述正向位移矩阵中确定所述第一像素点对应的第一横向位移和第一纵向位移；

基于所述第一像素点的第一位置信息、所述第一像素点对应的第一横向位移和第一纵向位移，将所述第一像素点正向映射到所述第二视频帧，并在所述第二视频帧中确定所映射得到的第二像素点的第二位置信息；

从所述反向位移矩阵中确定所述第二像素点对应的第二横向位移和第二纵向位移，并基于所述第二像素点的第二位置信息、所述第二像素点对应的第二横向位移和第二纵向位移，将所述第二像素点反向映射到所述第一视频帧，并在所述第一视频帧中确定所映射得到的第三像素点的第三位置信息；

基于所述第一像素点的第一位置信息、所述第三像素点的第三位置信息，确定所述第一像素点与所述第三像素点之间的误差距离，并根据所述第一像素点的第一位置信息、所述第二像素点的第二位置信息，确定包含第一像素点的图像块与包含所述第二像素点的图像块之间的相关系数；

将所述像素点中误差距离小于误差距离阈值、且所述相关系数大于等于相关系数阈值的像素点确定为有效像素点。

7、根据权利要求4所述的方法，其特征在于，所述基于所述有效像素点对所述第一视频帧对应的初始状态矩阵和所述正向位移矩阵进行修正，得到所述第一视频帧对应的目标状态矩阵和目标位移矩阵，包括：

获取第一视频帧对应的初始状态矩阵；所述初始状态矩阵中的每个矩阵元素的状态值均为第一数值，一个矩阵元素对应所述像素点中的一个像素点；

在所述初始状态矩阵中将与所述有效像素点对应的矩阵元素的状态值由第一数值切换为第二数值，并将包含第二数值的初始状态矩阵确定为所述第一视频帧对应的目标状态矩阵；

在所述正向位移矩阵中将所述剩余像素点对应的矩阵元素的位移设置为所述第一数值，并将包含所述第一数值的正向位移矩阵确定为目标位移矩阵；所述剩余像素点为所述像素点中除所述有效像素点之外的像素点。

8、根据权利要求7所述的方法，其特征在于，所述在所述正向位移矩阵中将所述剩余像素点对应的矩阵元素的位移设置为所述第一数值，并将包含所述第一数值的正向位移矩阵确定为目标位移矩阵，包括：

若所述正向位移矩阵包含初始横向位移矩阵和初始纵向位移矩阵，则在所述初始横向位移矩阵中将所述剩余像素点对应的矩阵元素的第一横向位移设置为所述第一数值，并将包含所述第一数值的初始横向位移确定为所述第一视频帧对应的横向位移矩阵；

在所述初始纵向位移矩阵中将所述剩余像素点对应的矩阵元素的第一纵向位移设置为所述第一数值，并将包含所述第一数值的初始纵向位移确定为所述第一视频帧对应的纵向位移矩阵；

将所述第一视频帧对应的横向位移矩阵和所述第一视频帧对应的纵向位移矩阵确定为目标位移矩阵。

9、根据权利要求4所述的方法，其特征在于，所述基于所述目标状态矩阵和所述目标位移矩阵，确定所述第一视频帧对应的平均位移矩阵，包括：

在所述第一视频帧中对所述目标状态矩阵进行位移积分运算，得到所述第一视频帧中的像素点对应的状态积分矩阵；

在所述第一视频帧中对所述目标状态矩阵中的横向位移矩阵进行位移积分运算，得到所述第一视频帧中的像素点对应的横向位移积分矩阵；

在所述第一视频帧中对所述目标状态矩阵中的纵向位移矩阵进行位移积分运算，得到所述第一视频帧中的像素点对应的纵向位移积分矩阵；

从所述第一视频帧中确定位移差分运算对应的差分区域，基于所述差分区域的尺寸信息、状态积分矩阵、横向位移积分矩阵和纵向位移积分矩阵，确定所述第一视频帧对应的平均位移矩阵。

10、根据权利要求 9 所述的方法，其特征在于，所述基于所述差分区域的尺寸信息、状态积分矩阵、横向位移积分矩阵和纵向位移积分矩阵，确定所述第一视频帧对应的平均位移矩阵，包括：

基于所述差分区域对应的长度信息和宽度信息，对所述状态积分矩阵进行位移差分运算，得到所述第一图像帧对应的状态差分矩阵；

基于所述差分区域对应的长度信息和宽度信息，分别对所述横向位移积分矩阵和纵向位移积分矩阵进行位移差分运算，得到所述第一图像帧对应的横向位移差分矩阵和纵向位移差分矩阵；

将所述横向位移差分矩阵与所述状态差分矩阵之间的比值确定为横向平均位移矩阵，并将所述纵向位移差分矩阵与所述状态差分矩阵之间的比值确定为纵向平均位移矩阵；

将所述纵向位移差分矩阵和所述纵向平均位移矩阵确定为所述第一视频帧对应的平均位移矩阵。

11、一种视频数据处理装置，所述装置应用于计算机设备，其特征在于，包括：

对象确定模块，用于响应于对目标视频的触发操作，从所述目标视频的关键视频帧中确定目标像素点，并获取与所述目标像素点相关联的多媒体信息，其中，所述关键视频帧是所述触发操作所在的视频帧，所述目标像素点是所述关键视频帧中与所述触发操作对应的像素点；

请求确定模块，用于基于所述目标像素点在所述关键视频帧中的位置信息，确定所述目标像素点对应的轨迹获取请求；

轨迹获取模块，用于基于所述轨迹获取请求，获取与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，其中，所述目标轨迹信息包含所述目标像素点在所述关键视频帧的下一视频帧中的位置信息，所述目标像素点在所述关键视频帧的下一视频帧中的位置信息是通过跟踪所述目标像素点得到的；

文本显示模块，用于当播放所述关键视频帧的下一视频帧时，基于所述目标轨迹信息中的所述目标像素点在所述关键视频帧的下一视频帧中的位置信息显示所述多媒体信息。

12、一种视频数据处理装置，所述方法应用于业务服务器，其特征在于，包括：

请求响应模块，用于响应于对关键视频帧中的目标像素点的轨迹获取请求，获取与目标视频相关联的轨迹信息，其中，所述关键视频帧是所述目标视频中的视频帧，所述目标像素点是所述关键视频帧中的像素点，所述轨迹信息是由所述目标视频的每个视频帧中的像素点的位置信息所确定的；

轨迹筛选模块，用于从所述目标视频相关联的轨迹信息中，筛选与所述目标像素点在所述关键视频帧中的位置信息相关联的目标轨迹信息，并返回所述目标轨迹信息，其中，所述目标轨迹信息包含目标位置信息，所述目标位置信息用于触发在所述关键视频帧的下一视频帧中，显示与所述目标像素点相关联的多媒体信息。

13、一种视频数据处理装置，其特征在于，包括：

第一获取模块，用于从目标视频中获取相邻的第一视频帧和第二视频帧；

矩阵获取模块，用于基于所述目标视频对应的光流追踪规则、所述第一视频帧中的像素点、所述第二视频帧中的像素点，确定所述第一视频帧对应的平均位移矩阵；

位置跟踪模块，用于基于所述平均位移矩阵对所述第一视频帧中的像素点的位置信息进行跟踪，并在所述第二视频帧中确定所跟踪得到的像素点的位置信息；

轨迹生成模块，用于基于所述第一视频帧中的像素点的位置信息、所述跟

踪得到的像素点在所述第二视频帧中的位置信息，生成与所述目标视频相关联的轨迹信息，其中，所述轨迹信息中包含用于对目标视频中的目标像素点所关联的多媒体信息进行跟踪显示的目标轨迹信息。

14、一种计算机设备，其特征在于，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，以执行如权利要求 1、2、3-10 任一项所述的方法。

15、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有计算机程序，所述计算机程序包括程序指令，所述程序指令当被处理器执行时，执行如权利要求 1、2、3-10 任一项所述的方法。

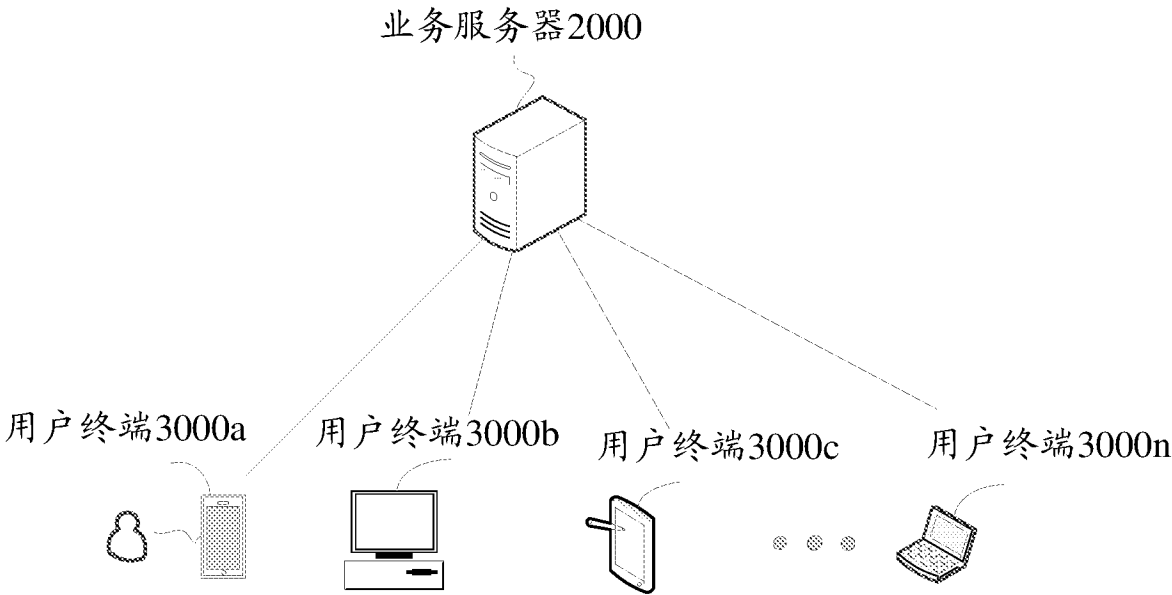


图 1

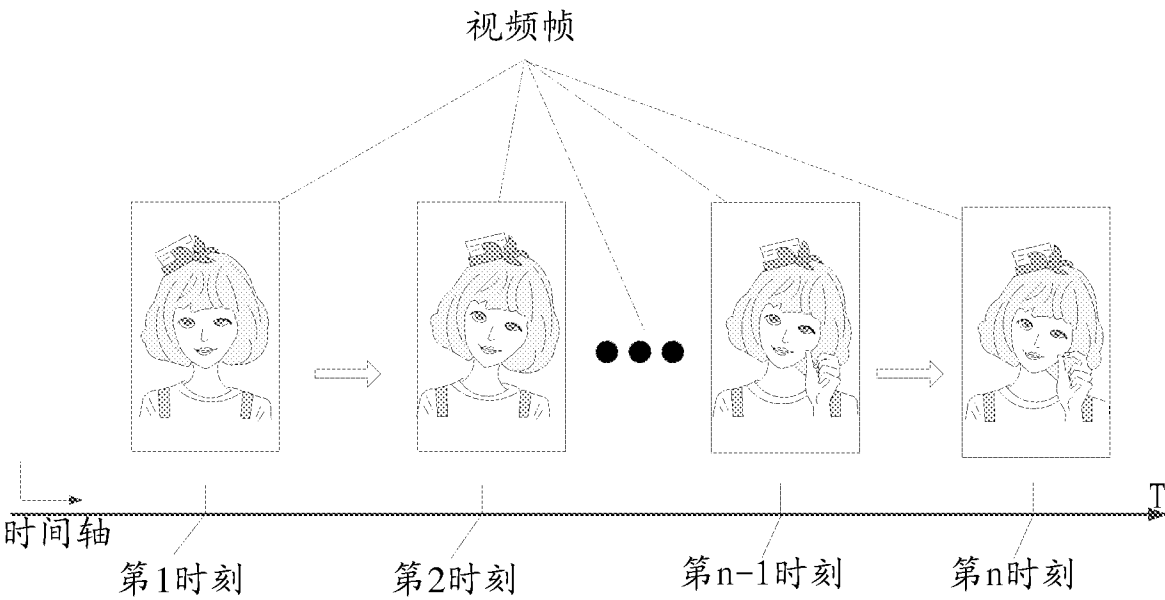


图 2

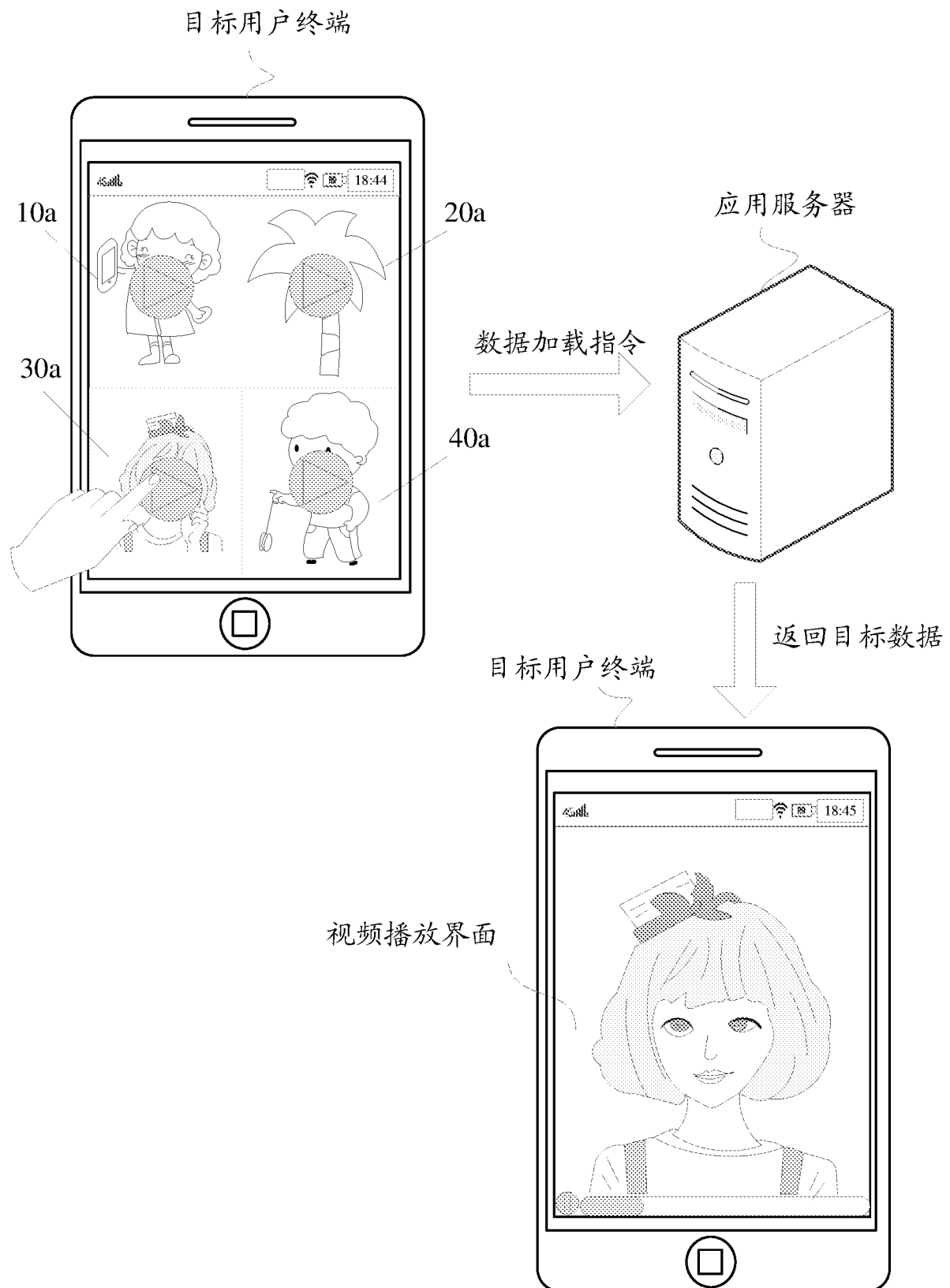


图 3

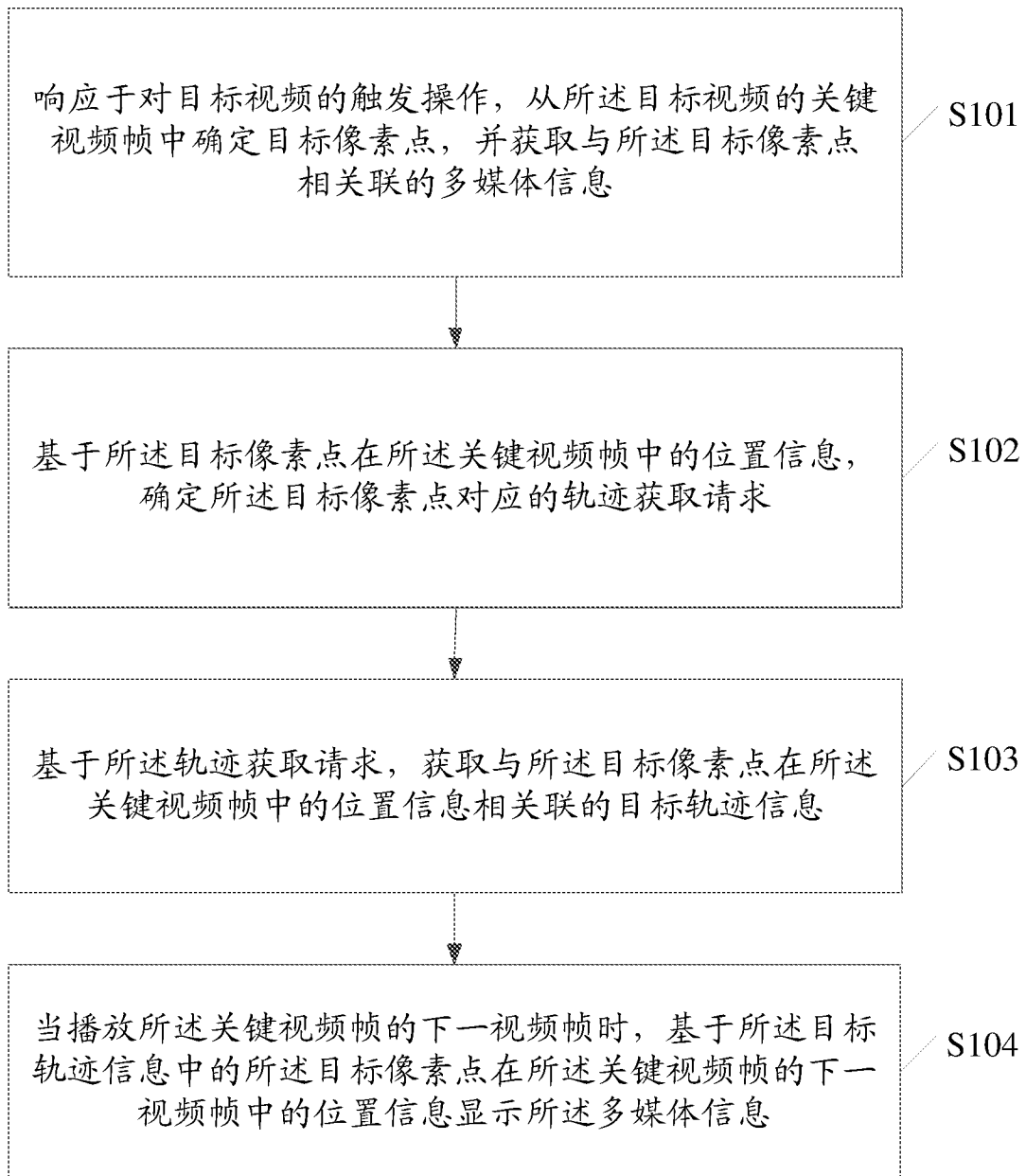


图 4

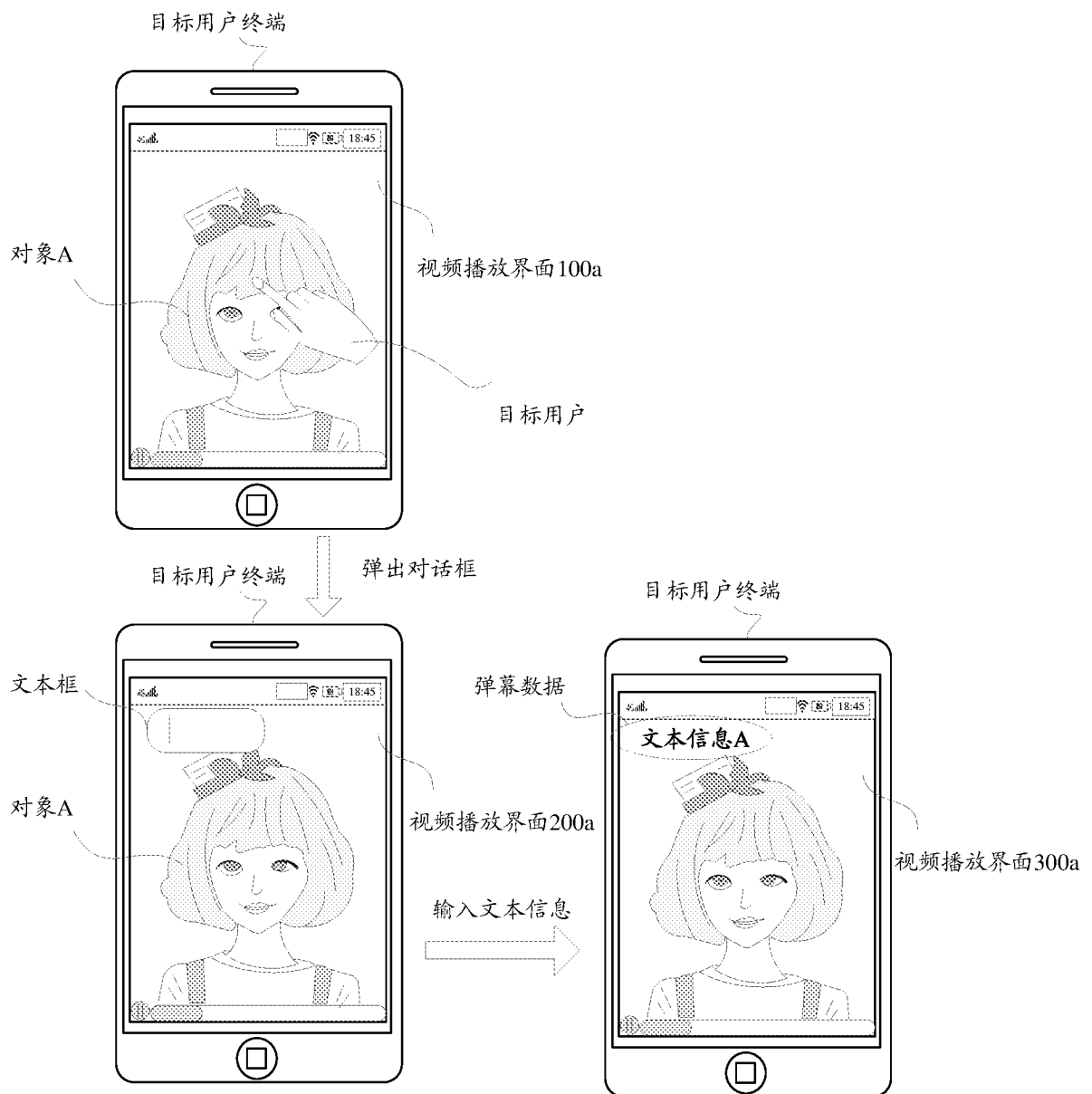


图 5

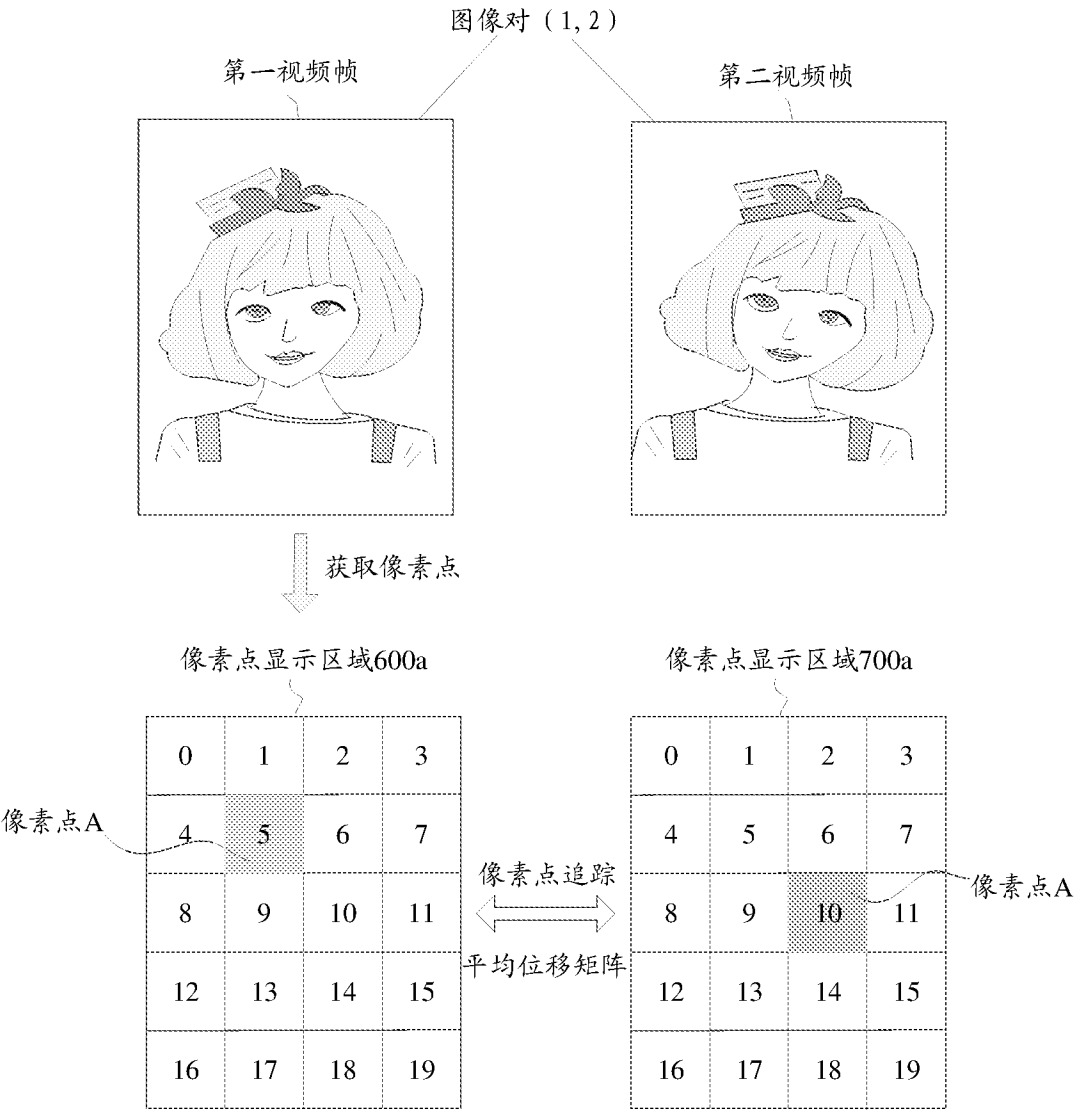


图 6

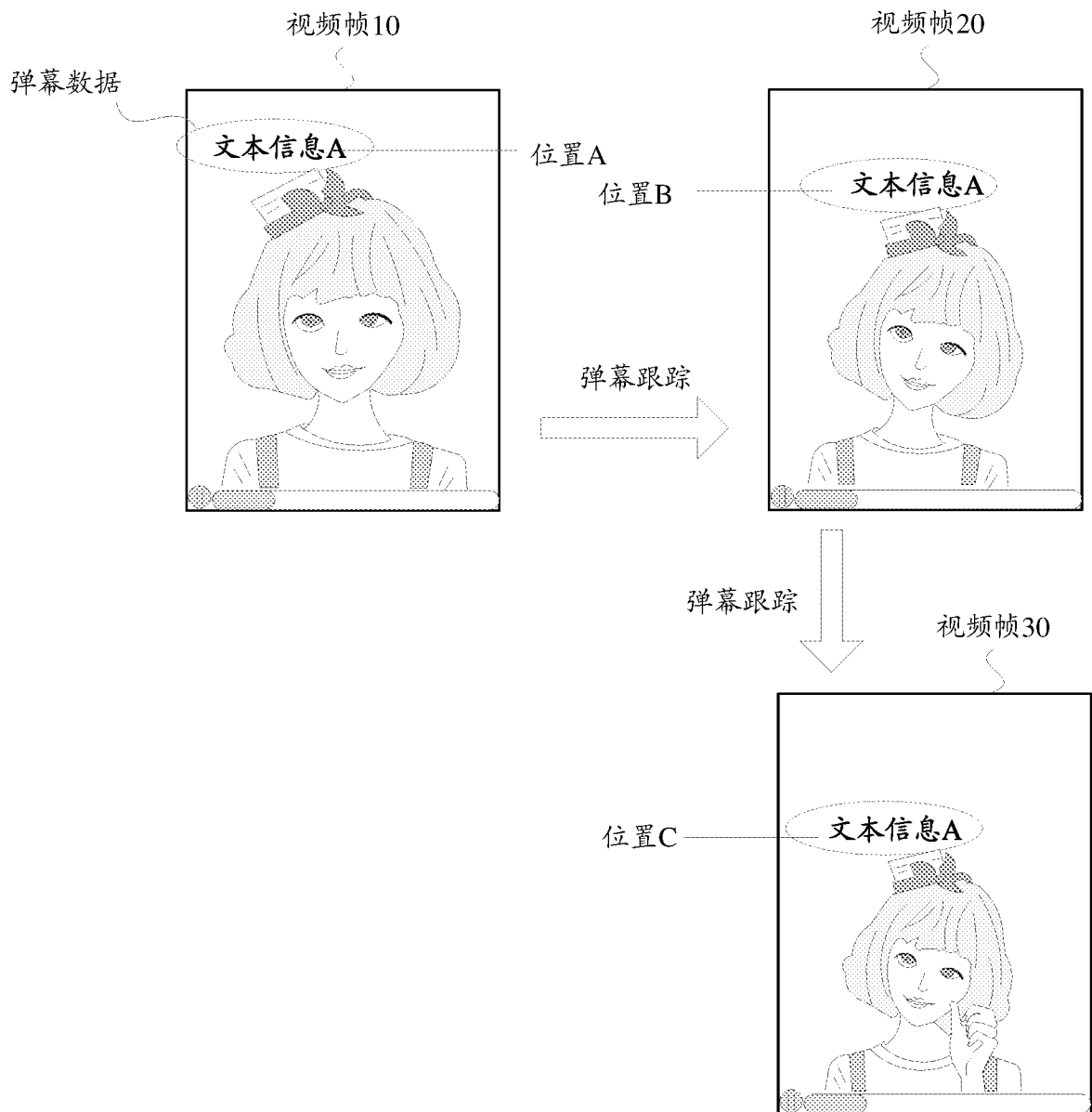


图 7

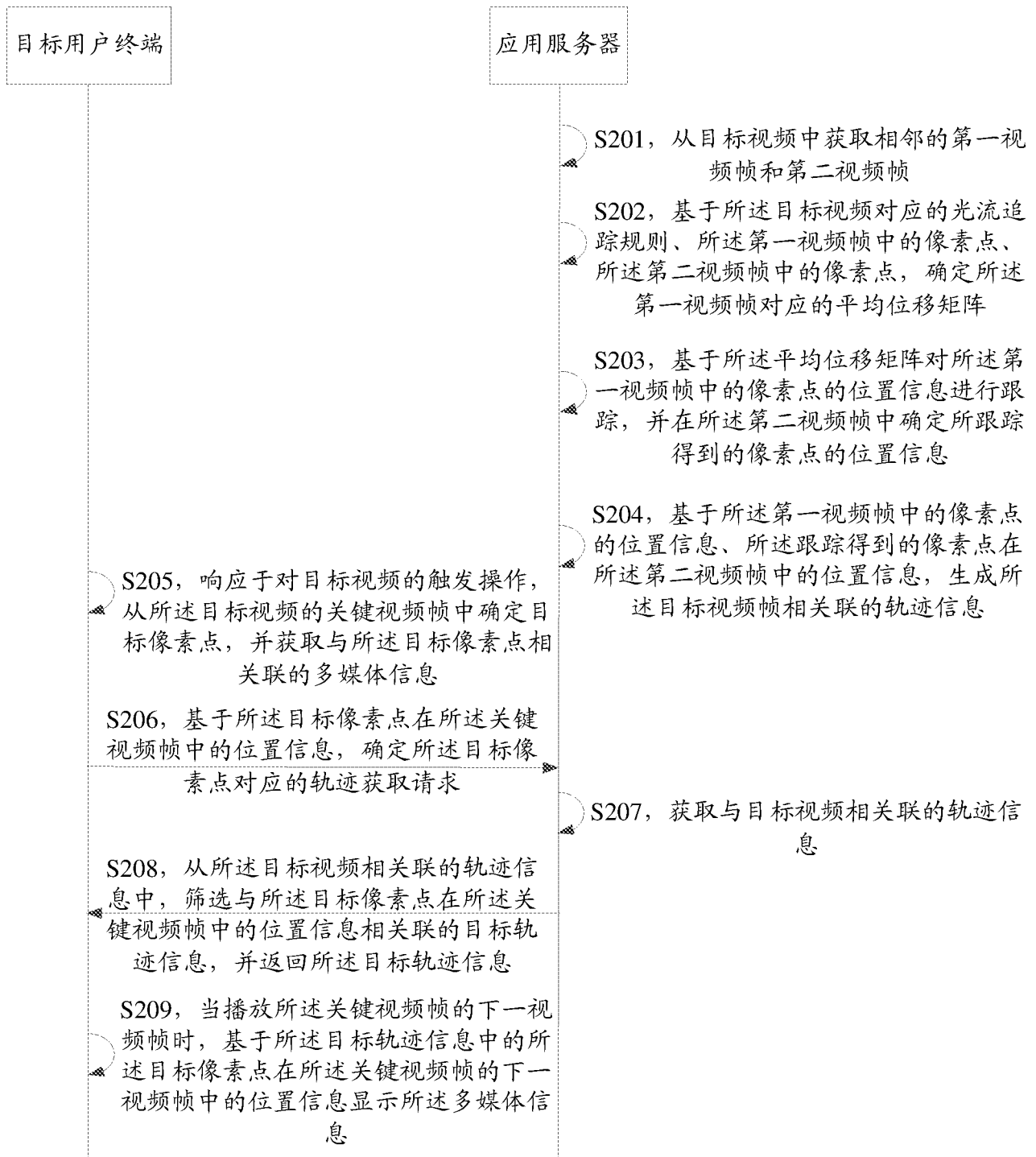


图 8

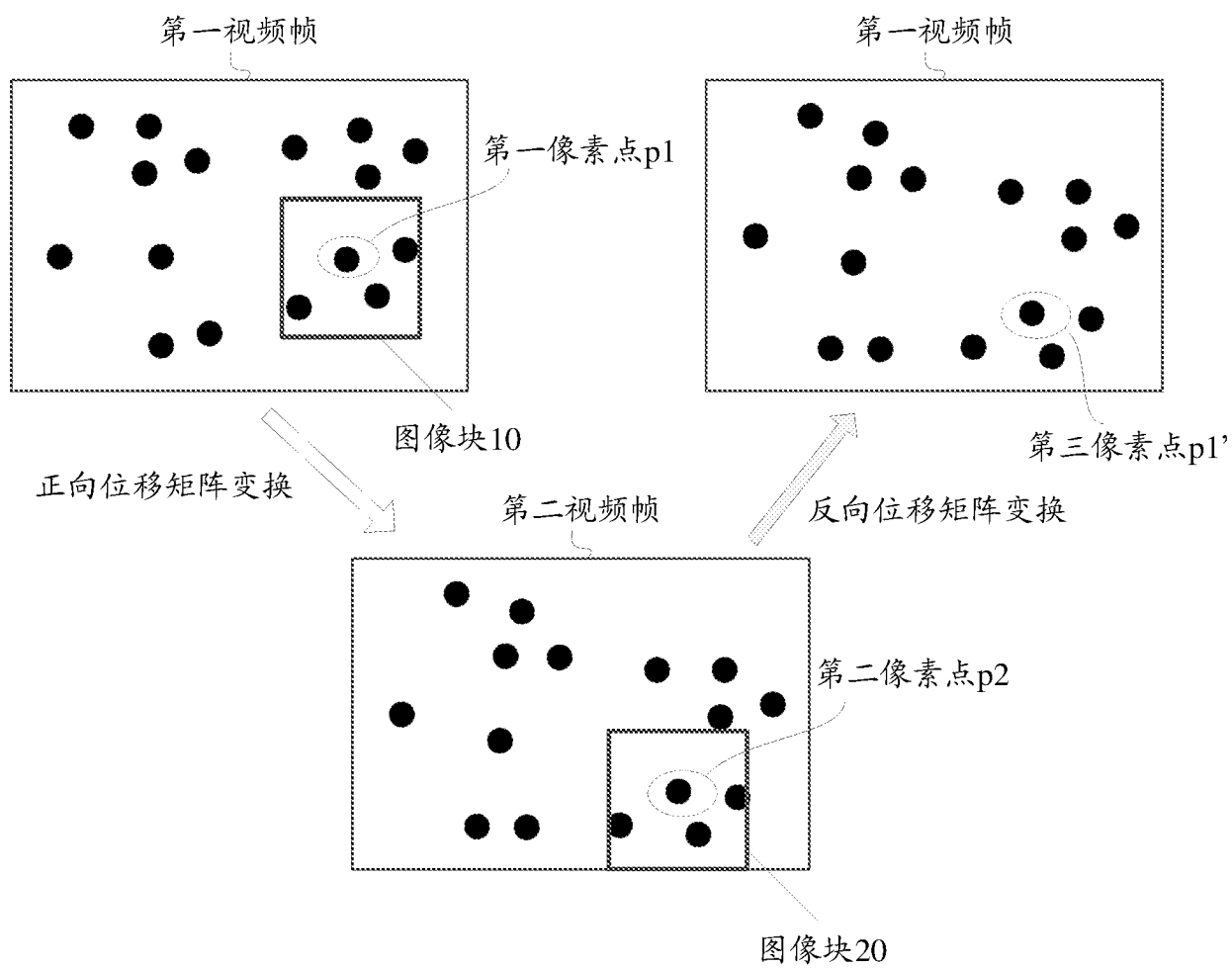


图 9

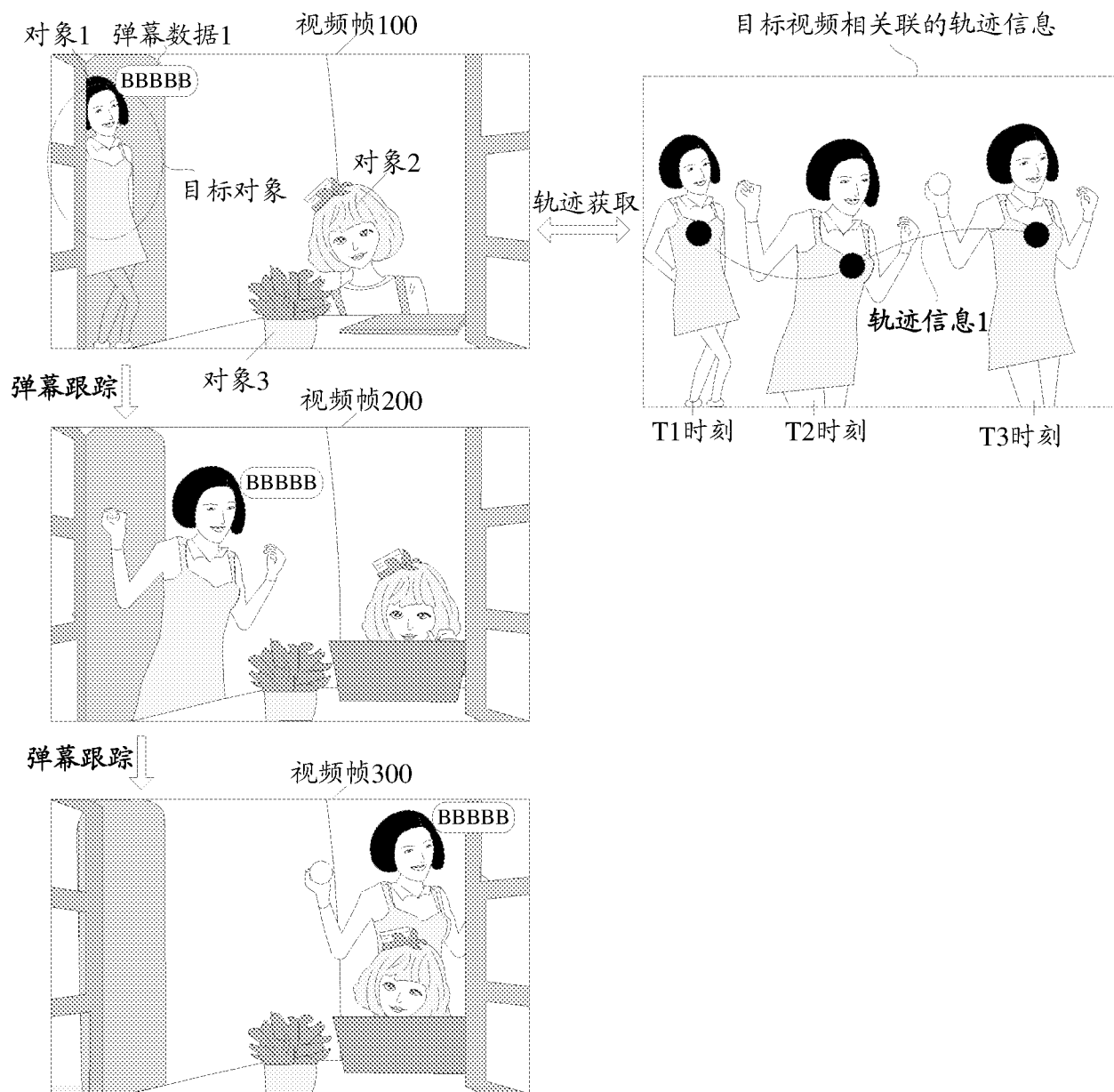


图 10

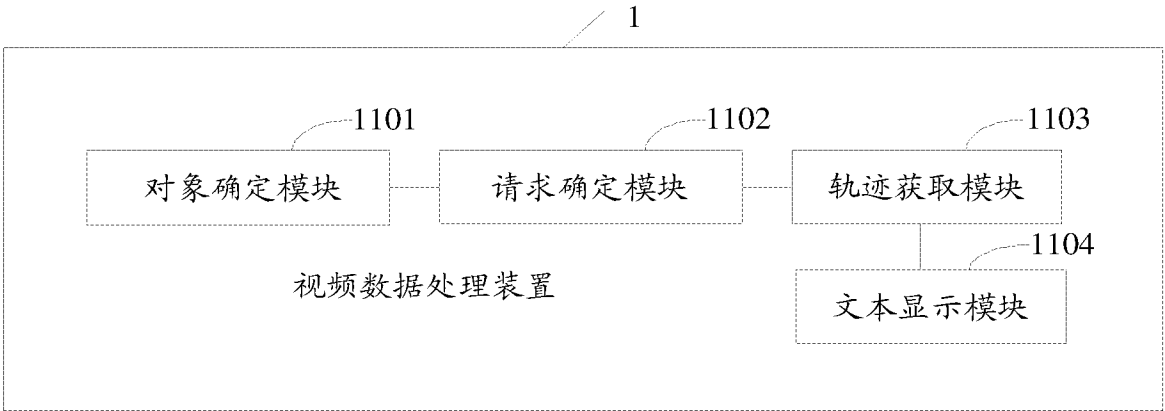


图 11

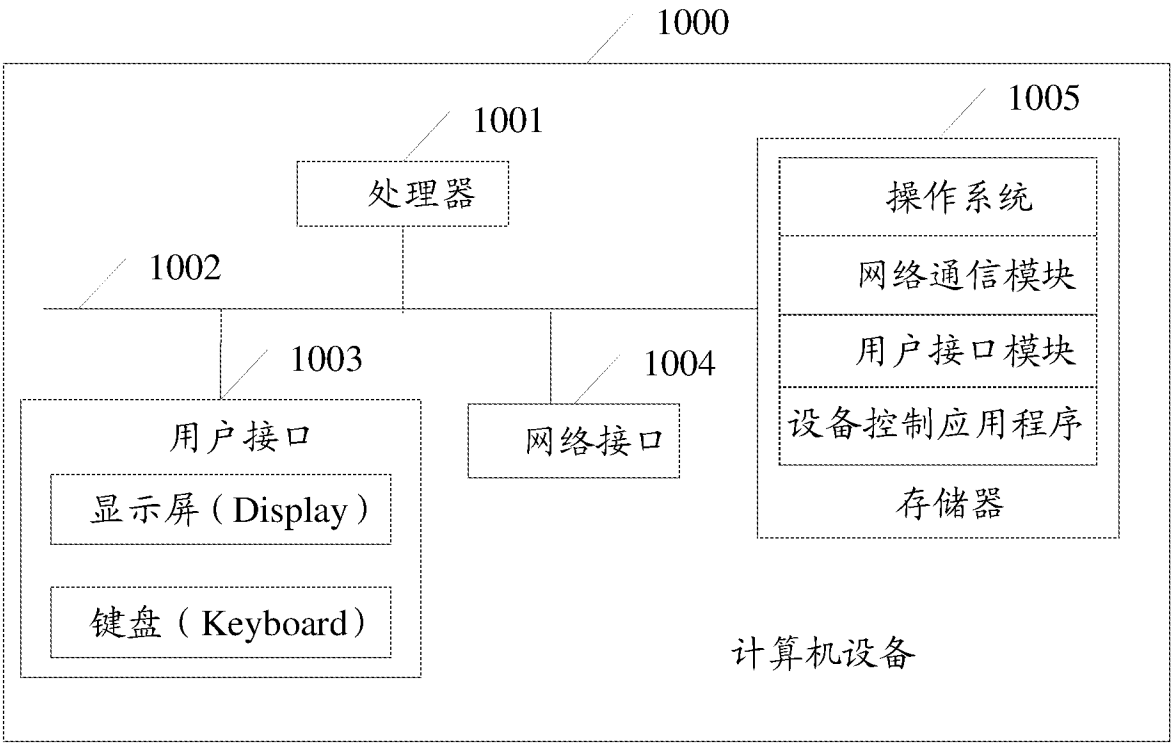


图 12

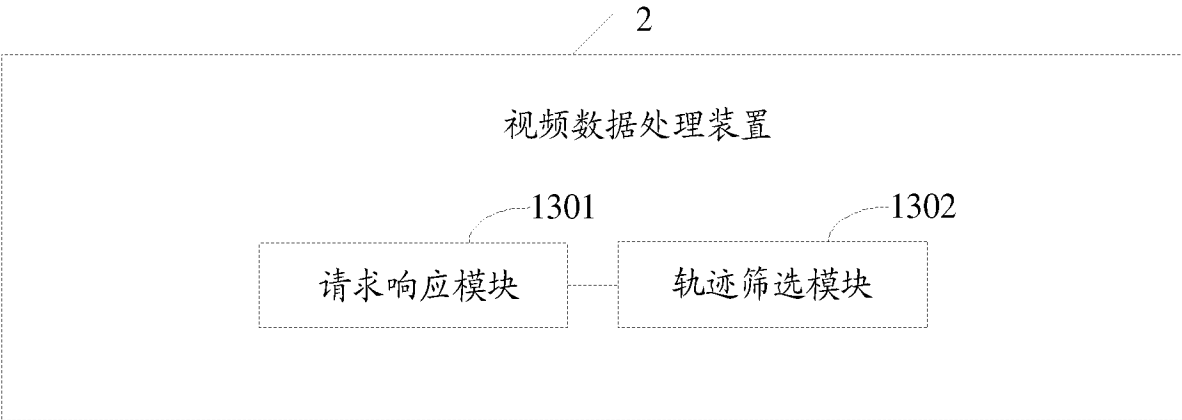


图 13

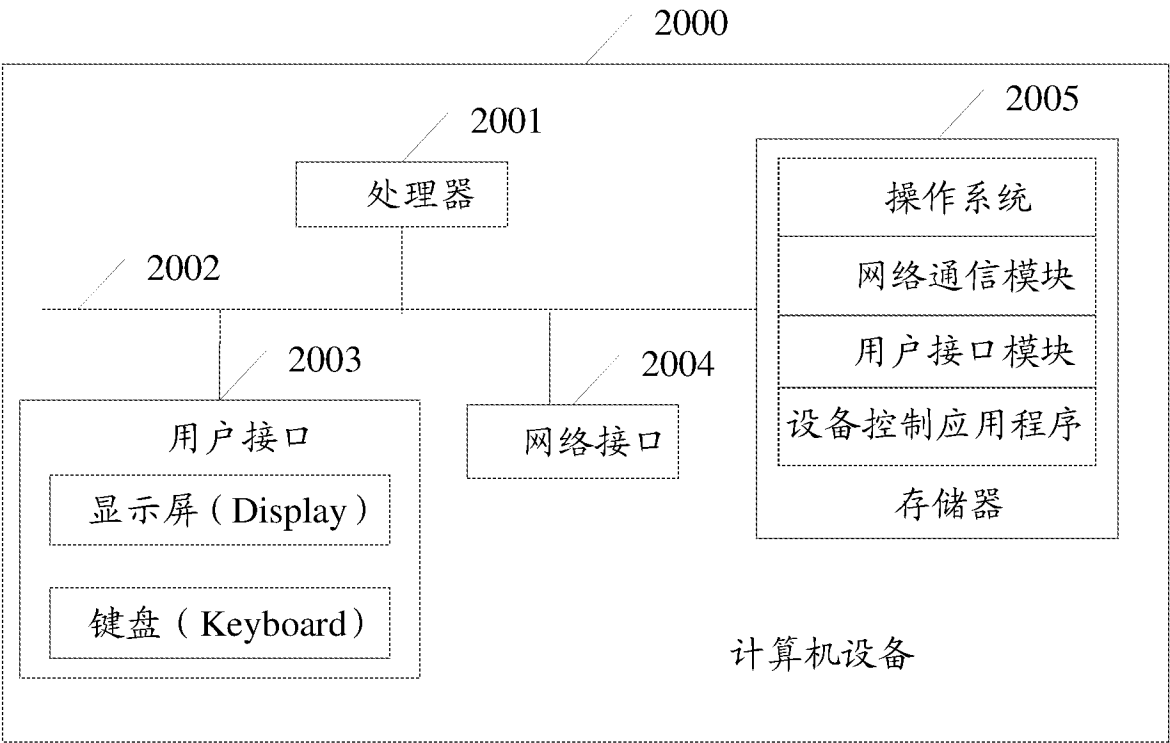
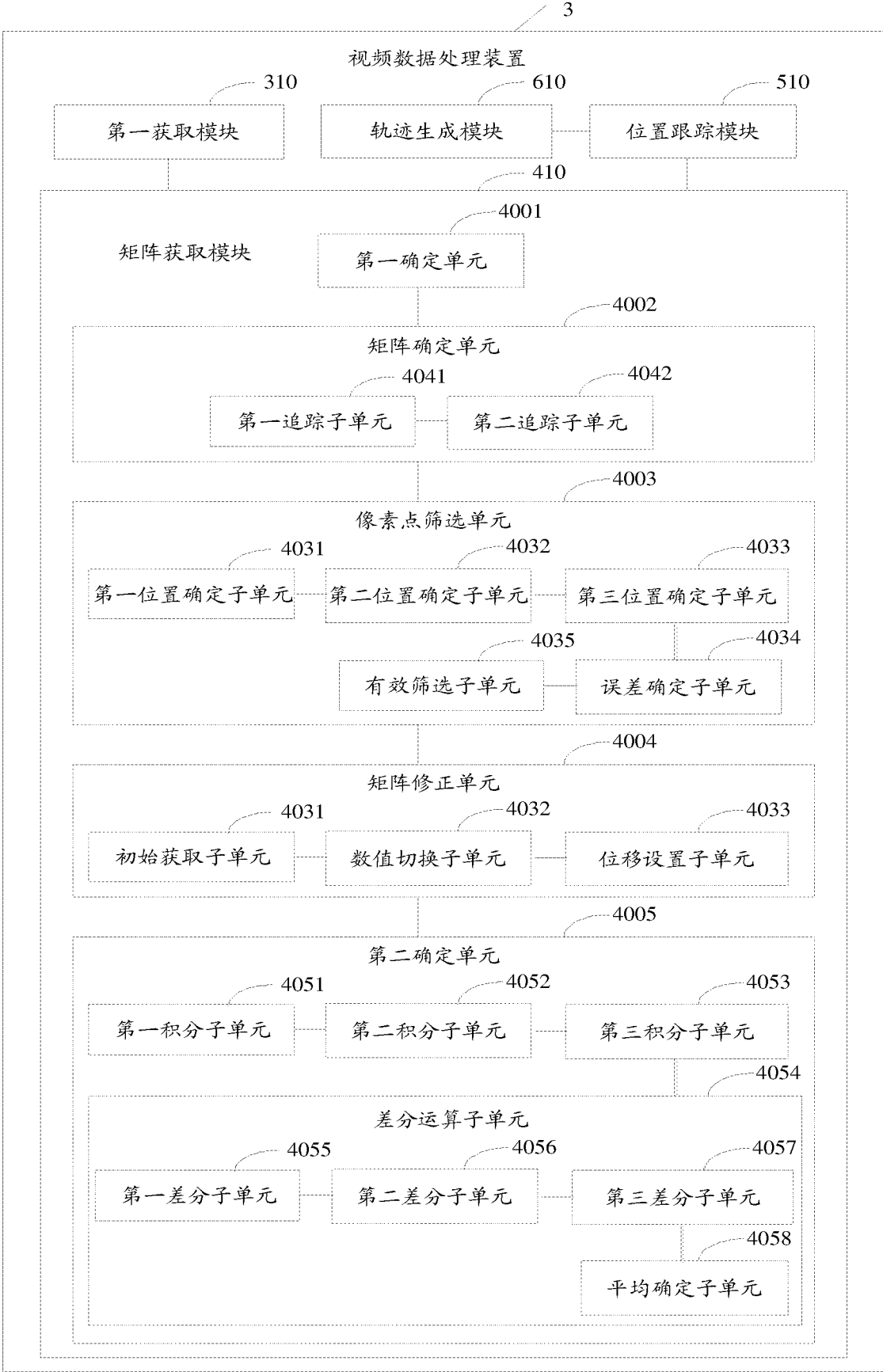


图 14



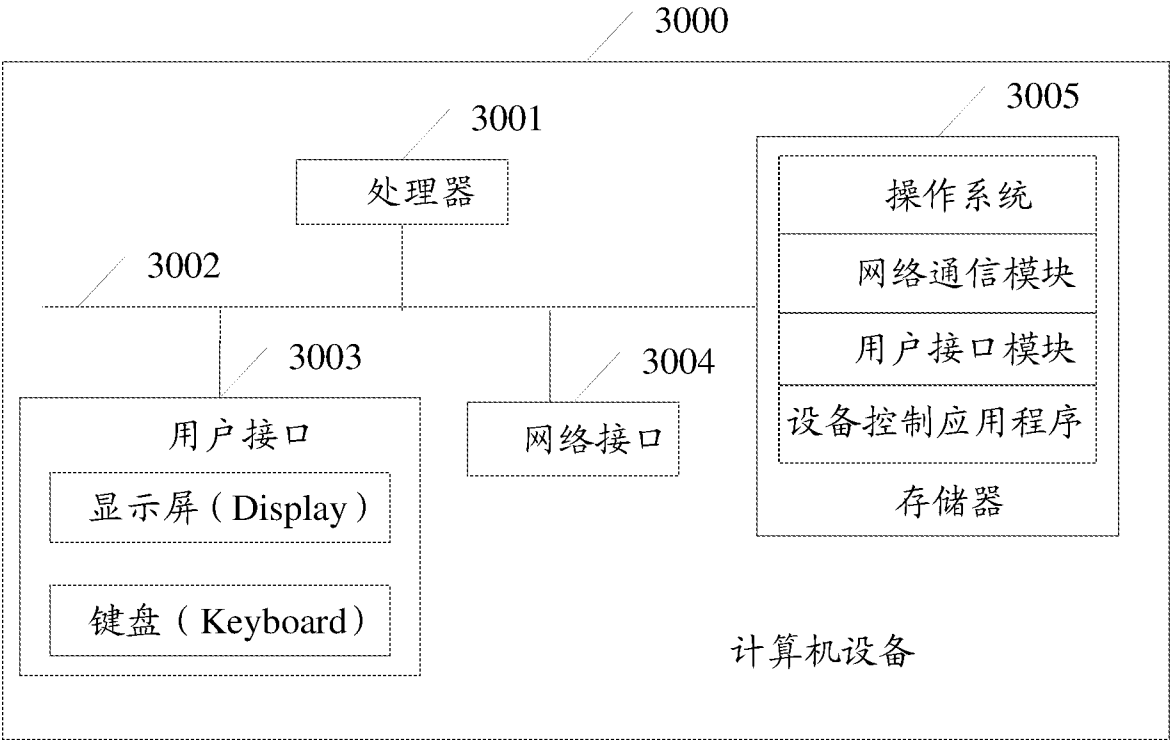


图 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/084112

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/431(2011.01)i; H04N 21/44(2011.01)i; H04N 21/488(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)

CNPAT, CNKI, WPI, EPODOC, IEEE: 视频, 帧, 目标, 像素点, 区域, 块, 轨迹, 移动, 运动, 位置, 坐标, 标识, 文字, 跟踪, 追踪, 光流, 矩阵, 播放, video, frame?, target, pixel?, point?, block?, object, track+, position, location, next, optical, flow, matrix, trajectory, character, play+, mov+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110062272 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 26 July 2019 (2019-07-26) entire document	1-15
X	CN 101930779 A (HUAWEI DEVICE CO., LTD.) 29 December 2010 (2010-12-29) description paragraphs 35-61, figures 1-8	1, 2, 11, 12, 14, 15
Y	CN 101930779 A (HUAWEI DEVICE CO., LTD.) 29 December 2010 (2010-12-29) description paragraphs 35-61, figures 1-8	3, 13
Y	CN 104881640 A (HUAWEI TECHNOLOGIES CO., LTD.) 02 September 2015 (2015-09-02) description paragraphs 140-153, 158-184, figures 1-4	3, 13
A	US 2014241573 A1 (BLENDAGRAM, INC.) 28 August 2014 (2014-08-28) entire document	1-15
A	CN 109558505 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 02 April 2019 (2019-04-02) entire document	1-15



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

30 June 2020

Date of mailing of the international search report

09 July 2020

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/CN2020/084112**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 108242062 A (BEIJING ZONGMU ANCHI INTELLIGENT TECHNOLOGY CO., LTD.) 03 July 2018 (2018-07-03) entire document	1-15
A	CN 109087335 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD. et al.) 25 December 2018 (2018-12-25) entire document	1-15
A	CN 105872442 A (NINGBO 3B ELECTRONIC TECH CO., LTD.) 17 August 2016 (2016-08-17) entire document	1-15

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/084112

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110062272	A	26 July 2019	None			
CN	101930779	A	29 December 2010	CN	101930779	B	29 February 2012
CN	104881640	A	02 September 2015	CN	104881640	B	15 June 2018
US	2014241573	A1	28 August 2014	WO	2014133559	A1	04 September 2014
				US	2014240353	A1	28 August 2014
				US	2014245350	A1	28 August 2014
CN	109558505	A	02 April 2019	None			
CN	108242062	A	03 July 2018	None			
CN	109087335	A	25 December 2018	WO	2020015498	A1	23 January 2020
CN	105872442	A	17 August 2016	None			

国际检索报告

国际申请号

PCT/CN2020/084112

A. 主题的分类 H04N 21/431(2011.01)i; H04N 21/44(2011.01)i; H04N 21/488(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)) CNPAT, CNKI, WPI, EPDOC, IEEE: 视频, 帧, 目标, 像素点, 区域, 块, 轨迹, 移动, 运动, 位置, 坐标, 标识, 文字, 跟踪, 追踪, 光流, 矩阵, 播放, video, frame?, target, pixel?, point?, block?, object, track+, position, location, next, optical, flow, matrix, trajectory, character, play+, mov+																										
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 110062272 A (腾讯科技深圳有限公司) 2019年 7月 26日 (2019 - 07 - 26) 全文</td> <td>1-15</td> </tr> <tr> <td>X</td> <td>CN 101930779 A (华为终端有限公司) 2010年 12月 29日 (2010 - 12 - 29) 说明书第35-61段, 图1-8</td> <td>1, 2, 11, 12, 14, 15</td> </tr> <tr> <td>Y</td> <td>CN 101930779 A (华为终端有限公司) 2010年 12月 29日 (2010 - 12 - 29) 说明书第35-61段, 图1-8</td> <td>3, 13</td> </tr> <tr> <td>Y</td> <td>CN 104881640 A (华为技术有限公司) 2015年 9月 2日 (2015 - 09 - 02) 说明书第140-153, 158-184段, 图1-4</td> <td>3, 13</td> </tr> <tr> <td>A</td> <td>US 2014241573 A1 (BLENDAGRAM, INC.) 2014年 8月 28日 (2014 - 08 - 28) 全文</td> <td>1-15</td> </tr> <tr> <td>A</td> <td>CN 109558505 A (百度在线网络技术北京有限公司) 2019年 4月 2日 (2019 - 04 - 02) 全文</td> <td>1-15</td> </tr> <tr> <td>A</td> <td>CN 108242062 A (北京纵目安驰智能科技有限公司) 2018年 7月 3日 (2018 - 07 - 03) 全文</td> <td>1-15</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 110062272 A (腾讯科技深圳有限公司) 2019年 7月 26日 (2019 - 07 - 26) 全文	1-15	X	CN 101930779 A (华为终端有限公司) 2010年 12月 29日 (2010 - 12 - 29) 说明书第35-61段, 图1-8	1, 2, 11, 12, 14, 15	Y	CN 101930779 A (华为终端有限公司) 2010年 12月 29日 (2010 - 12 - 29) 说明书第35-61段, 图1-8	3, 13	Y	CN 104881640 A (华为技术有限公司) 2015年 9月 2日 (2015 - 09 - 02) 说明书第140-153, 158-184段, 图1-4	3, 13	A	US 2014241573 A1 (BLENDAGRAM, INC.) 2014年 8月 28日 (2014 - 08 - 28) 全文	1-15	A	CN 109558505 A (百度在线网络技术北京有限公司) 2019年 4月 2日 (2019 - 04 - 02) 全文	1-15	A	CN 108242062 A (北京纵目安驰智能科技有限公司) 2018年 7月 3日 (2018 - 07 - 03) 全文	1-15
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																								
PX	CN 110062272 A (腾讯科技深圳有限公司) 2019年 7月 26日 (2019 - 07 - 26) 全文	1-15																								
X	CN 101930779 A (华为终端有限公司) 2010年 12月 29日 (2010 - 12 - 29) 说明书第35-61段, 图1-8	1, 2, 11, 12, 14, 15																								
Y	CN 101930779 A (华为终端有限公司) 2010年 12月 29日 (2010 - 12 - 29) 说明书第35-61段, 图1-8	3, 13																								
Y	CN 104881640 A (华为技术有限公司) 2015年 9月 2日 (2015 - 09 - 02) 说明书第140-153, 158-184段, 图1-4	3, 13																								
A	US 2014241573 A1 (BLENDAGRAM, INC.) 2014年 8月 28日 (2014 - 08 - 28) 全文	1-15																								
A	CN 109558505 A (百度在线网络技术北京有限公司) 2019年 4月 2日 (2019 - 04 - 02) 全文	1-15																								
A	CN 108242062 A (北京纵目安驰智能科技有限公司) 2018年 7月 3日 (2018 - 07 - 03) 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 2020年 6月 30日		国际检索报告邮寄日期 2020年 7月 9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 冀芊茜 电话号码 86-(10)-53961821																								

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
A	CN 109087335 A (腾讯科技深圳有限公司 等) 2018年 12月 25日 (2018 - 12 - 25) 全文	1-15
A	CN 105872442 A (宁波三博电子科技有限公司) 2016年 8月 17日 (2016 - 08 - 17) 全文	1-15

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/084112

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	110062272	A	2019年 7月 26日	无	
CN	101930779	A	2010年 12月 29日	CN	101930779 B 2012年 2月 29日
CN	104881640	A	2015年 9月 2日	CN	104881640 B 2018年 6月 15日
US	2014241573	A1	2014年 8月 28日	WO	2014133559 A1 2014年 9月 4日
				US	2014240353 A1 2014年 8月 28日
				US	2014245350 A1 2014年 8月 28日
CN	109558505	A	2019年 4月 2日	无	
CN	108242062	A	2018年 7月 3日	无	
CN	109087335	A	2018年 12月 25日	WO	2020015498 A1 2020年 1月 23日
CN	105872442	A	2016年 8月 17日	无	