

(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
2024 年 5 月 23 日 (23.05.2024)



(10) 国际公布号
WO 2024/104239 A1

(51) 国际专利分类号:
G06F 16/783 (2019.01) **G06F 16/78** (2019.01)

(21) 国际申请号: PCT/CN2023/130577

(22) 国际申请日: 2023 年 11 月 8 日 (08.11.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202211430306.1 2022 年 11 月 15 日 (15.11.2022) CN

(71) 申请人: 北京字跳网络技术有限公司
(**BEIJING ZITIAO NETWORK TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市海淀区紫金数码园 4 号楼 2 层 0207, Beijing 100190 (CN)。

(72) 发明人: 颜鹏翔 (**YAN, Pengxiang**); 中国北京市朝阳区七圣中街 12 号院融中心 B1 小邮局, Beijing 100028 (CN)。

(74) 代理人: 北京世辉律师事务所(**SHIHUI PARTNERS**); 中国北京市朝阳区建国门外大街 2 号北京银泰中心 C 座 42 层, Beijing 100022 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD,

(54) **Title:** VIDEO LABELING METHOD AND APPARATUS, AND DEVICE, MEDIUM AND PRODUCT

(54) 发明名称: 视频标注方法、装置、设备、介质及产品

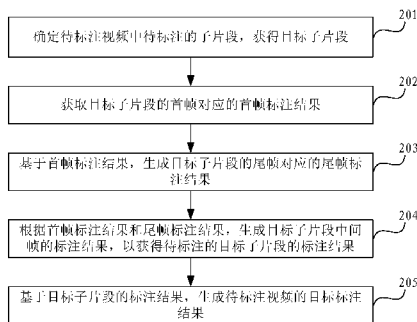


图 2

201 Determine sub-clips to be labeled that are in a video to be labeled, and obtain a target sub-clip

202 Acquire a first frame labeling result corresponding to a first frame of the target sub-clip

203 On the basis of the first frame labeling result, generate a tail frame labeling result corresponding to a tail frame of the target sub-clip

204 According to the first frame labeling result and the tail frame labeling result, generate a labeling result for an intermediate frame of the target sub-clip, so as to obtain a labeling result for the target sub-clip to be labeled

205 On the basis of the labeling result for the target sub-clip, generate a target labeling result for said video

(57) **Abstract:** Provided in the embodiments of the present disclosure are a video labeling method and apparatus, and a device, a medium and a product. The method comprises: determining sub-clips to be labeled that are in a video to be labeled, and obtaining a target sub-clip; acquiring a first frame labeling result corresponding to a first frame of the target sub-clip; on the basis of the first frame labeling result, generating a tail frame labeling result corresponding to a tail frame of the target sub-clip; according to the first frame labeling result and the tail frame labeling result, generating a labeling result for an intermediate frame of the target sub-clip, so as to obtain a labeling result for the target sub-clip to be labeled; and on the basis of the labeling result for the target sub-clip, generating a target labeling result for said video. The technical solution of the present disclosure improves the efficiency of video labeling.

(57) 摘要: 本公开实施例提供一种视频标注方法、装置、设备、介质及产品, 该方法包括: 确定待标注视频中待标注的子片段, 获得目标子片段; 获取所述目标子片段的首帧对应的首帧标注结果; 基于所述首帧标注结果, 生成所述目标子片段的尾帧对应的尾帧标注结果; 根据所述首帧标注结果和所述尾帧标注结果, 生成所述目标子片段中间帧的标注结果, 以获得所述待标注的目标子片段的标注结果; 基于所述目标子片段的标注结果, 生成所述待标注视频的目标标注结果。本公开的技术方案提高了视频标注效率。

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, CV, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SC, 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, ME, 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))。

视频标注方法、装置、设备、介质及产品

本申请要求 2022 年 11 月 15 日递交的申请号为 202211430306.1、标题为“视频标注方法、装置、设备、介质及产品”的中国发明专利申请的优先权，该申请的全部内容通过引用结合在本申请中。

技术领域

本公开实施例涉及计算机技术领域，尤其涉及一种视频标注方法、装置、设备、介质及产品。

10

背景技术

视频处理可以应用于诸多技术领域，例如，人工智能、智能交通、金融、内容推荐等多种技术领域，其具体涉及到的技术例如可以包括目标追踪、目标检测等。

15 相关技术中，视频的标注一般采用人工逐帧进行标注。但是，采用人工标注的方式，标注效率较低，标注成本过高。

发明内容

本公开实施例提供一种视频标注方法、装置、设备、介质及产品，以克服采用人工标注的方式，标注效率较低，标注成本过高的技术问题。

20 第一方面，本公开实施例提供一种视频标注方法，包括：

确定待标注视频中待标注的子片段，获得目标子片段；

获取所述目标子片段的首帧对应的首帧标注结果；

25 基于所述首帧标注结果，生成所述目标子片段的尾帧对应的尾帧标注结果；

根据所述首帧标注结果和所述尾帧标注结果，生成所述目标子片段中间帧的标注结果，以获得所述待标注的目标子片段的标注结果；

基于所述目标子片段的标注结果，生成所述待标注视频的目标标注结果。

第二方面，本公开实施例提供一种视频标注装置，包括：

第一确定单元，用于确定待标注视频中待标注的子片段，获得目标子片段；

首帧标注单元，用于获取所述目标子片段的首帧对应的首帧标注结果；

5 尾帧标注单元，用于基于所述首帧标注结果，生成所述目标子片段的尾帧对应的尾帧标注结果；

片段标注单元，用于根据所述首帧标注结果和所述尾帧标注结果，生成所述目标子片段中间帧的标注结果，以获得所述待标注的目标子片段的标注结果；

10 第二确定单元，用于基于所述目标子片段的标注结果，生成所述待标注视频的目标标注结果。

第三方面，本公开实施例提供一种电子设备，包括：处理器、存储器；
所述存储器存储计算机执行指令；

15 所述处理器执行所述存储器存储的计算机执行指令，使得所述处理器配置有如上第一方面以及第一方面各种可能的设计所述的视频标注方法。

第四方面，本公开实施例提供一种计算机可读存储介质，所述计算机可读存储介质中存储有计算机执行指令，当处理器执行所述计算机执行指令时，实现如上第一方面以及第一方面各种可能的设计所述的视频标注方法。

20 第五方面，本公开实施例提供一种计算机程序产品，包括计算机程序，所述计算机程序被处理器执行时实现如上第一方面以及第一方面各种可能的设计所述的视频标注方法。

本实施例提供的技术方案，针对待标注视频，可以从片段维度，确定待标注的目标子片段。对目标子片段进行详细标注时，可以先获取目标子片段的首帧对应的首帧标注结果，再基于首帧标注结果，生成目标子片段的尾帧对应的尾帧标注结果，可以利用首帧标注结果和尾帧标注结果对目标子片段中的中间帧进行标注，以实现目标子片段的中间帧的标注，获得目标子片段的标注结果。尾帧可通过首帧自动标注获得，而中间帧可以通过首帧标注结果和尾帧标注结果自动标注获得，实现中间帧的高效标注。获得目标子片段的标注结果之后，可以确定待标注视频的目标标注结果，通过时间维度更小的片段标注，可以提高片段标注准确性，相比于直接对待标注视频进行标

25

30

注，效率更高，准确度更高。

附图说明

为了更清楚地说明本公开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一简单地介绍，显而易见地，下面描述中的附图是本公开的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 为本公开实施例提供的一种视频标注方法的一个应用示例图；

图 2 为本公开实施例提供的一种视频标注方法的一个实施例的流程图；

10 图 3 为本公开实施例提供的一种视频标注方法的又一个实施例的流程图；

图 4 为本公开实施例提供的一个特征传播示例图；

图 5 为本公开实施例提供的一种视频标注方法的又一个实施例的流程图；

图 6 为本公开实施例提供的一个首帧标注结果的更新示例图；

图 7 为本公开实施例提供的一种视频标注方法的又一个实施例的流程图；

15 图 8 为本公开实施例提供的一种视频标注方法的又一个实施例的流程图；

图 9 为本公开实施例提供的一个视频子片段的划分示例图；

图 10 为本公开实施例提供的一个关键帧的提取示例图；

图 11 为本公开实施例提供的一种视频标注装置的一个实施例的结构示意图；

20 图 12 为本公开实施例提供的一种电子设备的硬件结构示意图。

具体实施方式

为使本公开实施例的目的、技术方案和优点更加清楚，下面将结合本公开实施例中的附图，对本公开实施例中的技术方案进行清楚、完整地描述，
25 显然，所描述的实施例是本公开一部分实施例，而不是全部的实施例。基于本公开中的实施例，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，都属于本公开保护的范围。

本公开的技术方案可以应用于视频标注场景中，通过获取首帧标注结果，并通过首帧标注结果自动标注尾帧，通过首帧标注结果和尾帧标注结果的获取
30 取可以对图像帧中的其他图像帧进行自动标注，提高视频的标注效率。

相关技术中，视频处理模型的训练需要大量的视频样本。视频样本可以包括视频本身以及视频的标签。视频的标签一般可以指视频中各图像帧的标签，各图像帧，也即视频中各图像帧的标注结果一般是人工标注获得。逐帧实现人工标注一般需要大量人工完成，标注效率较低，标注成本较高。

5 为了解决人工标注成本过高的问题，本公开考虑到自动完成对图像的标注。而图像的自动标注一般需要图像的区域识别模型，如果直接通过区域识别模型，获得的标注结果也不够准确。为了获得准确的标注结果，可以采用手动标注部分图像，再利用手动标注的图像，采用半监督标注方式，对剩余图像进行标注。通过此方式标注的图像准确度更高，标注效率也大大提升。

10 下面将以具体实施例对本公开的技术方案以及本公开的技术方案如何解决上述技术问题进行详细说明。下面几个具体实施例可以相互结合，对于相同或相似的概念或过程可能在某些实施例中不再赘述。下面将结合附图对本发明的实施例进行详细描述。

15 图 1 为本公开实施例提供的一种视频标注方法的一个应用示例图，该视频标注方法可以应用于电子设备 1 中，电子设备 1 可以包括显示装置 2。显示装置 2 可以显示待标注视频。待标注视频可以基于多个关键帧被划分为至少一个视频子片段。根据本公开的技术方案，可以对待标注视频按照各视频子片段进行标注，例如对目标子片段 3 进行片段标注，电子设备 1 可以在显示装置 2 中显示目标子片段 3 中任意图像 4 的片段标注结果，片段标注结果
20 例如可以为图 1 中的车辆所在区域 5，图像中的其他类型的对象，例如路灯 6 即可以不进行标注，获得该图像 4 的片段标注结果。其中，为了便于理解，图 1 所示的车辆区域 5 采用矩形框标注，该标注方式仅仅是示例性的，并不应构成对标注方式以及标注种类的具体限定，在实际应用中，还可以采用被标注对象的轮廓、圆形、多边形等其他形状进行标注。片段标注结果确定之后，
25 可以利用标注的目标子片段，确定待标注视频的目标标注结果。

 如图 2 所示，为本公开实施例提供的一种视频标注方法的一个实施例的流程图，该视频标注方法可以配置为一视频标注装置，视频标注装置可以位于电子设备中，视频标注方法可以包括以下几个步骤：

201：确定待标注视频中待标注的子片段，获得目标子片段。

30 可选地，确定待标注视频中待标注的子片段，获得目标子片段之前，还

可以包括：响应于视频标注请求，获取待标注视频。

目标子片段可以为待标注视频中至少一个视频子片段中待标注的子片段。可以将待标注视频划分为至少一个视频子片段，至少一个视频子片段可以通过待标注视频片段划分获得。

5 202：获取目标子片段的首帧对应的首帧标注结果。

首帧可以为目标子片段的第一个图像，也可以为目标子片段的任意图像。

首帧标注结果可以通过人工标注获得，也可以通过图像标注模型提取获得。为了提高首帧标注效率，还可以先通过图像标注模型自动标注，之后通过人工对图像标注模型的标注结果进行修正，获得最终的首帧标注结果。

10 尾帧可以为目标子片段的最后一个图像。

203：基于首帧标注结果，生成目标子片段的尾帧对应的尾帧标注结果。

尾帧标注结果可以通过目标子片段的第一个图像的标注结果结合半监督标注算法提取获得。半监督标注算法可以采用前向传播方式，将目标子片段的第一个图像的标注结果传播至尾帧，获得尾帧的尾帧标注结果。

15 204：根据首帧标注结果和尾帧标注结果，生成目标子片段中间帧的标注结果，以获得待标注的目标子片段的标注结果。

中间帧可以包括目标子片段中未标注的图像帧，中间帧可以通过首帧标注结果和尾帧标注结果标注获得。目标子片段中可以包括多个图像或者图像帧，每个图像均可以进行标注，获得每个图像的标注结果。目标子片段中的多个图像帧均标注结束，可以获得目标子片段的多个图像帧分别对应的标注结果所构成目标子片段的片段标注结果。

20 204：基于目标子片段的标注结果，生成待标注视频的目标标注结果。

待标注视频可以包括至少一个视频子片段，每个视频子片段标注过程中可以称为目标子片段，标注结束即可以获得目标子片段的标注结果。待标注视频的目标标注结果可以包括多个视频子片段分别对应的标注结果。

25 本公开实施例中，针对待标注视频，可以从片段维度，确定待标注的片段，获得目标子片段。对目标子片段的标注中，可以先获取目标子片段的首帧对应的首帧标注结果，并通过首帧的首帧标注结果生成尾帧对应的尾帧标注结果，可以利用首帧标注结果和尾帧标注结果对目标子片段中的中间帧分别进行标注，以实现目标子片段的自动标注，获得目标子片段的标注结果。

目标子片段中各图像可以通过其首帧标注结果和尾帧标注结果自动标注获得，获得高效的标注效果。获得目标子片段的片段标注结果之后，可以确定待标注视频的目标标注结果，通过时间维度更小的片段标注，可以提高片段标注效率，相比于直接对待标注视频进行标注，准确度更高。

5

在一般情况下，可以采用人工标注方式获得尾帧的尾帧标注结果。但是，为了提高尾帧的标注效率，可以利用前向传播算法确定尾帧的尾帧标注结果。

如图 3 所示，为本公开实施例提供的一种图像标注方法的一个实施例的流程图，该方法与上述实施例的不同之处在于，基于首帧的标注结果生成目标子片段尾帧的标注结果，包括：

10

301：获取目标子片段的首帧对应的首帧标注结果。

302：根据首帧标注结果，利用前向传播算法确定尾帧对应的尾帧标注结果。

15

本公开实施例中，可以根据首帧标注结果并结合前向传播算法，自动确定尾帧的尾帧标注结果。通过自动确定尾帧的尾帧标注结果，可以有效提升尾帧的标注效率。

在一种可能的设计中，根据首帧标注结果，利用前向传播算法确定尾帧对应的尾帧标注结果，包括：

20

利用前向传播算法，将首帧标注结果向目标子片段中未标注的图像帧进行标注结果的顺序传播，获得目标子片段中未标注图像帧的标注结果；

获取目标子片段的最后一个图像帧的标注结果作为尾帧对应的尾帧标注结果。

25

本公开实施例中，利用前向传播算法，将首帧标注结果向目标子片段中未标注的图像帧进行标注结果的顺序传播，获得未标注图像帧的标注结果，通过前向传播算法将首帧标注结果向未标注的图像帧进行传播，直至传播至目标视频片段的最后一个图像帧，获得尾帧对应的尾帧标注结果，通过标注结果的传播，使得尾帧的标注是不断传播获得，使得尾帧的标注参考到其附近，例如尾帧的前一个图像帧的标注结果，提高尾帧标注效率和准确性。

30

在实际应用中，可以采用双向传播方式，对目标子片段的中间帧进行标注处理。对于不同位置的中间帧，可以按照中间帧分别与首帧和尾帧的位置差异对图像自动标注，以提高图像的标注精度。

因此，如图 4 所示，为本公开实施例提供的一种视频标注方法的又一个
5 实施例的流程图，与前述实施例的不同之处在于，根据首帧标注结果和尾帧标注结果，生成目标子片段中间帧的标注结果，可以包括：

401：基于首帧标注结果，结合前向传播算法，提取目标子片段的中间帧的前向传播特征。

可选地，前向传播算法可以包括机器学习算法、神经网络算法等算法，
10 可以通过训练获得。前向传播算法可以用于将首帧的首帧标注结果向位于首帧之后的中间帧进行特征传播，获得中间帧的前向传播特征。

目标子片段可以包括 N 个图像帧，每个图像帧均可以作为中间帧进行标注。N 为大于 1 的正整数。可以先对首帧和尾帧进行标注，之后可以从目标子片段中的第二个图像帧开始，依次将每个图像帧作为中间帧，获得每个中
15 间帧的标注结果，直至获得目标子片段的尾帧的前一个图像的标注结果，此时目标子片段标注结束。

其中，前向传播特征可以指从首帧开始，将首帧的标签逐帧向位于其之后的其他图像进行特征传播，传播至图像序号对应的图像即停止传播，获得的图像特征。首帧标注结果作为特征传播 mask（掩码）参与特征计算。具体
20 可以使用下述实施例中的半监督分割算法进行特征传输。

402：基于尾帧标签结果，结合后向传播算法，提取目标子片段的中间帧的后向传播特征。

可选地，后向传播算法可以包括机器学习算法、神经网络算法等算法，
可以通过训练获得。后向传播算法可以将尾帧标注结果向尾帧之前的中间帧
25 进行传播，获得中间帧的后向传播特征。

其中，后向传播特征可以指从尾帧开始，将尾帧的标签逐帧向位于其之前的其他图像进行特征传播，传播至图像序号对应的图像即停止传播，获得的图像特征。同样，尾帧标注结果也可以作为特征传播掩码参与特征计算。

403：将前向传播特征和后向传播特征进行特征融合处理，获得中间帧的
30 目标图像特征。

404: 根据目标图像特征, 确定中间帧的标注结果。

目标子片段可以包括一个或多个中间帧, 每个中间帧均可以进行标注, 获得各中间帧的标注结果。目标子片段的片段标注结果可以包括多个中间帧各自的标注结果。

- 5 本公开实施例中, 利用前向传播算法可以获得中间帧的前向传播特征, 利用后向传播算法可以获得中间帧的后向传播特征。前向传播特征和后向传播特征的融合可以使得目标图像特征融合了首帧标注结果和尾帧标注结果, 通过目标图像特征可以更好地表征中间帧的标注特征, 提高中间帧的标注精度和准确性。

10

作为一个实施例, 前向传播特征的提取步骤可以包括: 可以根据首帧标注结果和图像序号, 利用前向传播算法, 确定中间帧的前向传播特征。后向传播特征的提取步骤可以包括: 根据尾帧标注结果和图像序号, 利用后向传播算法, 确定中间帧的后向传播特征。

- 15 在实际应用中, 对图像打标可以根据实际使用需求的不同设置不同的类别。在同一次打标中可以一次打多个类型的标签, 例如, 在自然图像处理场景中, 可以对视频中的车辆、行人进行目标追踪, 因此, 车辆和行人可以作为两个标签类别, 以分别进行打标。在图像特征提取过程中, 为了更好地表征不同类别的标签, 各类别的标签不受其他类别的影响, 可以为各标签类别
20 分别生成标签特征。标签特征的元素可以代表各像素点属于该标签类别的概率。对于同一个坐标, 该坐标的元素值具体可以包括该坐标在至少一个标签类别分别对应的概率值, 概率值最大的标签类别所代表的标签即为该坐标的标签。

- 在一种可能的设计中, 获得前向传播特征和后向传播特征之后, 可以对
25 前向传播特征和后向传播特征进行特征融合处理, 获得中间帧的目标图像特征; 根据目标图像特征通过特征识别即可以确定中间帧的标注结果。根据首帧标注结果和图像序号, 确定中间帧的前向传播特征包括: 根据首帧标注结果对首帧进行特征提取, 获得首帧在至少一个标签类别分别对应的标签特征, 将首帧对应的至少一个标签类别分别对应的标签特征向后传播, 获得中间帧
30 在至少一个标签类别分别对应的前向标签特征, 以获得至少一个标签类别分

别对应的前向传播特征。

可选地，根据尾帧标注结果和图像序号，确定中间帧的后向传播特征，包括：根据尾帧标对尾帧进行特征提取，获得尾帧在至少一个标签类别分别对应的标签特征，将尾帧对应的至少一个标签类别分别对应的标签特征向前传播，获得中间帧的至少一个标签类别分别对应的后向标签特征，以获得至少一个标签类别分别对应的前向传播特征。

本公开实施例中，前向传播特征基于首帧标注结果和图像序号获得，使得前向传播特征综合了首帧标注结果和图像序号的特性。而后向传播特征基于尾帧标注结果和图像序号获得，综合了尾帧标注结果和图像序号的特性。

前向传播特征和后向传播特征分别为图像特征从首帧传播和从尾帧传播获得的结果。利用前向传播特征和后向传播特征进行特征融合处理，获得中间帧的目标图像特征。目标图像特征综合了前向和后向两个方向的传播特性，利用目标图像特征获得的标注结果更准确，可以提高中间帧的标注效率和准确性。

在一种可能的设计中，前向传播算法可以包括：半监督分割算法。

后向传播算法可以包括：半监督分割算法。

可以根据首帧标注结果，利用半监督分割算法对目标子片段从首帧开始逐帧进行前向特征传播，直至获得图像序号处的前向传播特征。可以根据尾帧标注结果，利用半监督分割算法对目标子片段从尾帧开始逐帧进行后向特征传播处理，直至获得图像序号处的后向传播特征。

其中，半监督分割算法具体可以为半监督物体分割算法。可以通过半监督分割算法对目标子片段从首帧或尾帧开始，利用前一帧的图像特征计算当前帧的图像特征。直至获得图像序号对应的前向或后向传播特征。

本公开实施例中，可以通过半监督分割算法，对目标子片段从首帧开始逐帧进行前向特征传播，直至获得图像序号位置处的前向传播特征。以半监督的分割算法，可以完成图像特征的前向传播，使得对应的计算获得的前向传播特征综合首帧及其之前的图像的前向特征，特征的表现度更高通过半监督分割算法还可以从尾帧开始传播，也即从尾帧开始逐帧进行后向特征传播处理，直至获得图像序号处的后向传播特征。通过半监督分割算法，可以对图像特征进行前向或后向的传播，提高图像特征的计算准确度。

在获得前向传播特征和后向传播特征之后，可以根据前向传播特征和后向传播特征进行特征的融合计算，以使得中间帧的图像特征综合前向和后向两个方面的特征。在某些实施例中，将前向传播特征和后向传播特征进行特征融合处理，获得中间帧的目标图像特征，包括：

确定中间帧在目标子片段中的图像序号；

根据图像序号确定序号比值；

根据序号比值，确定前向传播权重和后向传播权重；

根据前向传播权重、后向传播权重、前向传播特征和后向传播特征，获得中间帧的目标图像特征。

可选地，中间帧的图像序号可以指中间帧在目标子片段的在目标子片段的出现顺序。例如，目标子片段中的第一个图像的图像序号可以为 1，第二个出现的图形的图像序号可以为 2。通过图像序号可以确定中间帧在目标子片段中的位置。每个图像帧可以按照其标注顺序，确定相应的图像序号，例如首帧的图像序号可以为 1，尾帧的图像序号可以为 $N+1$ 。

本公开实施例中，可以确定目标子片段中的中间帧和中间帧的图像序号，中间帧的序号可以代表其与首帧和尾帧的位置关系。通过首帧标注结果和尾帧标注结果并结合中间帧的图像序号，可以确定中间帧的标注结果。使得中间帧的标注效果与中间帧在目标子片段中的位置关联，提高标注准确性。

作为一个实施例，根据图像序号确定序号比值，可以包括：

计算中间帧的图像序号和目标子片段的尾帧对应的尾帧序号的序号比值。

作为又一个实施例，根据前向传播权重、后向传播权重、前向传播特征和后向传播特征，获得中间帧的目标图像特征可以包括：

根据前向传播权重和后向传播权重，对前向传播特征和后向传播特征进行特征融合处理加权求和，获得中间帧的目标图像特征。

图像序号为 K ，尾帧序号为 N ，则序号比值为 K/N 。根据序号比值，确定前向传播权重和后向传播权重，可以包括：确定序号比值 K/N 为后向传播权重，确定整数 1 和序号比值的差，也即 $1-K/N$ ，为前向传播权重。目标图像特征的加权求和步骤可以包括：

计算前向传播权重： $1-K/N$ 与前向传播特征 $F_{forward}$ 的乘积，获得第一特

征；计算后向传播权重： K/N 与后向传播特征 $F_{backward}$ 的乘积，获得第二特征；将第一特征和第二特征相加获得目标图像特征 $F_{current}$ 。

可选地，前向传播特征可以包括至少一个标签类别分别对应的前向标签特征。后向传播特征可以包括至少一个标签类别分别对应的后向标签特征。

- 5 根据前向传播权重和后向传播权重，将各标签类别的前向标签特征和后向标签特征加权求和，获得各标签类别分别对应的融合特征。而各标签类别分别对应的融合特征即为目标图像特征。

- 每个标签类别的前向标签特征和后向标签特征的加权求和可以包括：对于每个标签类别的前向标签特征和后向标签特征，将各像素坐标在前向标签特征的第一特征值和前向传播权重相乘，将在后向标签特征的第二取值和后向传播权重相乘，将两个乘积相加，获得各像素坐标在该标签类别的特征值，也即，获得该标签类别在各像素坐标的特征值。
- 10

目标图像特征可以表征为中间帧的各像素坐标在不同标签类别的特征值。

- 为了便于理解，如图 5 所示的特征传播示例图，假设首帧 501 的首帧标注结果为 5011，尾帧 502 的尾帧标注结果为 5021。其中，首帧 501 的首帧标注结果 5011 前向传播对应的前向传播特征，尾帧 502 的尾帧标注结果 5021 对应的后向传播特征。中间帧 503 可以基于其序号对前向传播特征和后向传播特征进行特征融合，获得相应的目标图像特征。目标图像特征经图像分类层识别可以获得中间帧的目标区域 5031。该目标区域 5031 即可以为中间帧的标注结果。
- 15
- 20

- 本公开实施例中，可以根据图像序号对图像与前向传播的特征和后向传播的特征的关联度进行计算，也即计算图像序号对应的序号比值。该序号比值可以用于确定前向传播权重和后向传播权重。通过前向传播和后向传播的相关特性的计算，可以对图像的传播效率准确提升，提高图像特征传播的准确度。
- 25

作为一个实施例，根据目标图像特征，确定中间帧的标注结果可以包括：根据图像分类层，识别目标图像特征的目标区域；以目标区域作为中间帧的标注结果。

- 30 可选地，根据图像分类层，识别目标图像特征的目标区域，可以包括：

确定目标图像特征中中间帧的各像素坐标在至少一个标签类别分别对应的特征值，获取各像素坐标在至少一个标签类别分别对应的特征值中的最大特征值，以获得各像素坐标的最大特征值。根据各像素坐标的最大特征值对应的标签类别，确定各标签类别对应的目标像素坐标，确定各标签类别的目标像素坐标所形成的标签区域，获得至少一个标签类别分别对应的标签区域构成的目标区域。也即，至少一个标签类别分别对应的标签区域可以为中间帧的标注结果。图像分类层可以为对图像特征进行特征分类的数学模型。

本公开实施例中，确定中间帧的标注结果之后，可以根据图像分类层识别目标图像特征的目标区域，以该目标区域作为中间帧的标注结果。通过图像分类层的使用可以对目标图像特征进行准确的标签提取。

如图 6 所示，为本公开实施例提供的一种图像标注方法的又一个实施例的流程图，与前述实施例的不同之处在于，在确定中间帧的标注结果之后，还包括：

601：输出中间帧的标注结果。

标注结果可以包括至少一个标签类别分别对应的标签区域。

602：检测用户针对中间帧的标注结果执行的标签确认操作，维持中间帧的标注结果不变。

603：检测用户针对中间帧的标注结果执行的标签修改操作，获得中间帧修改后的标注结果。

可以同时输出中间帧和其标注结果，对中间帧的自动标注结果进行输出，供用户查看。

本公开实施例中，输出中间帧的标注结果之后，用户可以查看中间帧的标注结果，对标注结果的打标效果进行查看，若打标不合格，可以对标注结果进行修改，若打标合格，可以直接确定中间帧的标注结果。通过与用户交互显示，可以使得中间帧的标注结果与用户的标注需求更匹配，标注准确度更高。

作为一个实施例，获取目标子片段的首帧对应的首帧标注结果，可以包括：

检测用户针对首帧执行的标注操作，获得标注操作对应的首帧标注结果。

或，获取目标子片段的前一个视频子片段，并确定前一个视频子片段的尾帧对应的尾帧标注结果为目标子片段的首帧对应的首帧标注结果。

可选地，在首帧为目标子片段的第一个图像且目标子片段为待标注视频的
5 第一个视频子片段，可以检测用户针对目标子片段的首帧执行的标签设置
操作，获得设置结束时的首帧标注结果。或者，目标子片段不为第一个视频
子片段时，获取目标子片段的前一个视频子片段的尾帧的尾帧标注结果作为
目标子片段的首帧对应的首帧标注结果。

本公开实施例中，通过检测用户针对首帧执行的标注操作，可以获得标
注操作对应的首帧标注结果，可以获得与用户标注需求更匹配的首帧标注结
10 果，或者还可以将前一个视频子片的尾帧标注结果作为首帧的标注结果，可
以提高首帧标注效率。

作为又一个实施例，首帧对应的首帧标注结果的获取方式除上述实施例
提供的技术方案之外，目标子片段的首帧以及首帧对应的首帧标注结果，还
可以通过下列方式获得：

15 若用户针对中间帧的标注结果执行的标签修改操作，更新修改标注结果
后的中间帧为首帧；

将中间帧修改后的标注结果作为首帧标注结果。

如图 7 所示，为本公开实施例提供的一种图像帧的标注提示示例图。参
考图 7，在获得中间帧 701 的标注结果 7011 之后，若检测到用户针对中间帧
20 修改其标注结果例如修改为标注结果 7012，可以将中间帧 701 作为首帧。而
原首帧 702 则可以不再作为首帧。当然，图 7 的图像帧的标注提示仅仅是示
例性的，并不具备限定作用。

本公开实施例中，在用户对中间帧执行标签修改操作时，可以说明标签
的传播精度降低，与用户的实际标注需求匹配度较低。将标签修改后的中间
25 帧作为首帧，中间帧修改后的图像标注作为首帧标注结果，可以提供更有效
的图像传播，提高图像的传播效率和准确度。

为了获得准确的视频子片段，如图 8 所示，为本公开实施例提供的一种
视频标注方法的又一个实施例的流程图，与前述实施例的不同之处在于，确
定待标注视频中待标注的子片段，获得目标子片段，包括：

30 801：提取待标注视频的关键帧。

802: 将关键帧中相邻的两个关键帧在待标注视频所围成的视频区间划分为一个视频子片段, 获得至少一个视频子片段。

803: 从至少一个视频子片段中确定待标注的目标子片段。

可选地, 可以以组的方式对待标注视频的关键帧进行分组, 相邻的两个
5 关键帧可以作为一组, 可以从至少一个关键帧中确定至少一组关键帧。一组
关键帧包括相邻的第一关键帧和第二关键帧, 第一关键帧位于第二关键帧之
前, 前一组关键帧的第二关键帧与后一组关键帧的第一关键帧相同。两个相
邻的关键帧所围成的视频区间可以作为视频子片段, 也即视频子片段可以包
括两个关键帧, 以及两个关键帧之间的中间帧, 当然中间帧可以按照预设采
10 样频率采样获得。

其中, 关键帧可以为与待标注视频中在其附近的图像差异较大的图像。
例如, 在 t_1 时间的图像不存在车辆, 在 t_2 时间的图像出现车辆, t_1 和 t_2
的时间差在时间约束内, 则确定 t_2 时间的图像为关键帧。

为了便于理解, 图 9 所为本公开实施例提供的一个视频子片段的划分示
15 例图。参考图 9, 待标注视频的关键帧分别为关键帧 1、关键帧 2、关键帧 4
以及关键帧 6。可以将相邻的两个关键帧作为一组。

其中, 关键帧 1 和关键帧 2 可以作为一组相邻的关键帧, 该组相邻的关
键帧之间所围成的图像帧可以为视频子片段 1。视频子片段 1 可以由关键帧 1、
关键帧 2 以及关键帧 1 和 2 之间的图像帧 3 组成。

20 关键帧 2 和关键帧 4 可以作为一组相邻的关键帧, 该组相邻的关键帧之
间所围成的图像帧可以为视频子片段 2。视频子片段 2 可以由关键帧 2、关键
帧 4 以及关键帧 2 和 4 之间的图像帧 5 组成。

关键帧 4 和关键帧 6 可以作为一组相邻的关键帧, 该组相邻的关键帧之
间所围成的图像帧可以为视频子片段 3。视频子片段 3 可以由关键帧 4、关键
25 帧 6 以及关键帧 4 和 6 之间的图像帧 7 组成。

相邻两组关键帧存在关键帧重叠, 参考图 9, 关键帧 2 可以为视频子片
段 1 的尾帧, 关键帧 2 还可以为视频子片段 2 的首帧。关键帧 4 可以为视频
子片段 2 的尾帧, 可以为视频子片段 3 的首帧。通过此关键帧的提取方式,
各个关键帧的提取

30 本公开实施例中, 通过提取待标注视频的关键帧, 可以基于关键帧完成

相邻两个关键帧的获取。而相邻的两个关键帧在待标注视频中所围成的视频区间可以为一个视频子片段，进而获得待标注视频对应的至少一个视频子片段，使得至少一个视频子片段中相邻的两个视频子片段的前一个视频子片段的最后一帧与后一个视频子片段的第一帧相同，完成待标注视频的全面而准确的分割，使得至少一个视频子片段的分割效率更高。

在某些实施例中，可以从待标注视频中，按照关键帧提取频率提取至少一个关键帧；或者，从待标注视频中，提取满足图像变化条件的至少一个关键帧。

关键帧提取频率可以根据使用需求设置，可以是预先设置获得的。关键帧提取频率的单位为帧/次。每间隔特征提取频率个图像帧，提取一个关键帧。例如，关键帧提取频率为 10 时，可以每 10 帧提取一个关键帧，第 1 帧、第 11 帧均可以为关键帧。

在一种可能的设计中，提取待标注视频的至少一个关键帧，包括：

针对待标注视频中的图像帧，计算各图像帧的运动幅度值；

根据运动幅度值获得图像帧中的至少一个关键帧。

图像变化条件可以包括：图像帧的运动幅度值大于指标阈值。

可选地，根据运动幅度值获得图像帧中的至少一个关键帧，可以包括：

若任意图像帧的运动幅度值大于指标阈值，则确定图像帧为关键帧，以获得多个图像帧中的至少一个关键帧。

运动幅度值可以指图像帧与其周围帧的幅度差异。可以将图像帧的幅度值与其周围帧的幅度值进行差值计算，获得运动幅度值。若运动幅度值大于指标阈值，则说明图像帧与周围帧差异较大，该图像帧可以作为关键帧。

为了便于理解，如图 10，为本公开实施例提供的一个关键帧的提取示例图。以纵轴各图像帧的运动幅值，横轴为待标注视频中的各个图像帧序号为例，从第一个图像帧 0 开始各个图像帧的幅值不断变化，各个图像帧的幅值连线形成曲线 1001。运动幅度值可以为各图像帧之间的幅值差异。由曲线 1001 可以的变化情况可以确定图像帧的幅值差异，也即前后运动幅值大于指标阈值的关键点 1002 对应的图像帧可以为关键帧。

本公开实施例中，可以针对待标注视频中的多个图像帧，计算各图像帧在运动幅度指标的指标数据，根据各图像帧的运动幅度可以进行关键帧的筛

选。关键帧可以用于获取视频子片段，以将运动幅度作为视频子片段的获取基础，使得同一视频子片段的运动幅度值作为划分基础，在进行图像自动标注时，可以有效提升图像的标注精度。

本公开实施例中，各图像帧的运动幅度值的计算步骤，可以包括：

- 5 计算各图像帧在帧间幅度差异指标对应的帧间差值，确定帧间差值为运动幅度值；

或者，计算各图像帧在帧间光流变化指标对应的帧间光流变化幅度值，确定帧间光流变化幅度值为运动幅度值。

- 10 或者，基于预训练的分割模型，计算各图像帧对应的分割结果的交并比，确定交并比为运动幅度值。

采用不同种类的运动幅度值，指标阈值可以根据运动幅度值的类型确定。

可选地，帧间差值可以指两个图像帧各自的像素均值的差值。

- 15 光流变化幅度值可以指两个或两个以上的图像帧的光流之间的差值，可以通过光流计算公式计算获得各图像帧对应的光流浮动阈值。各图像帧之间的分割结果进行交并比计算，交并比可以指图像帧和其周围帧分别进行图像分割处理，获得的图像帧的分割结果和其周围帧的分割结果之间的交集和并集的比值，如果二者重叠度较高，交并比的值较大，如果二者的重叠度较低，交并比的值较小。

- 20 本公开实施例中，通过计算图像帧对应的帧间差值、帧间光流变化幅度值或者分割结果的交并比，可以应用多种方式对各图像帧的运动幅度值进行准确计算。

作为一个实施例，确定待标注视频中待标注的子片段，获得目标子片段，包括：

- 25 根据至少一个视频子片段的时间先后顺序，确定至少一个视频子片段分别对应的片段顺序；

按照至少一个视频子片段分别对应的片段顺序，从第一个视频片段开始，依次选择一个视频片段作为待标注的子片段，获得目标子片段。

- 30 可选地，可以基于至少一个视频子片段分别对应的片段序号为各视频子片段的片段顺序。可以依次从至少一个视频子片段中确定目标子片段。在获得目标子片段之后可以执行上述实施例的标注方案，直至至少一个视频子片

段遍历结束，获得所有视频子片段的标注结果，将所有视频子片段的标注结果综合获得待标注视频的标注结果。

在视频进行片段分割时，每获得一个视频子片段，可以为该视频子片段设置片段序号。例如第一个获取的视频子片段的片段序号为 1，第二个视频子片段的片段序号为 2。

本公开实施例中，可以根据至少一个视频子片段分别对应的片段顺序，依次从至少一个视频子片段中选择目标子片段。利用片段顺序进行目标子片段的获取可以确保依次获得相应的目标子片段，进而依次完成各目标子片段的标注，实现对至少一个视频子片段的顺序、依次标注，提高视频子片段的标注全面性。

此外，本公开的技术方案还可以应用于游戏领域，具体例如可以包括三维游戏场景的设计、显示等应用领域。

如图 11 所示，为本公开实施例提供的一种视频标注装置的一个实施例的结构示意图，该装置可以位于电子设备中，可以配置有上述视频标注方法，该视频标注装置 1100 可以包括：

第一确定单元 1101，用于确定待标注视频中待标注的子片段，获得目标子片段；

首帧标注单元 1102，用于获取目标子片段的首帧对应的首帧标注结果；

尾帧标注单元 1103，用于基于首帧标注结果，生成目标子片段的尾帧对应的尾帧标注结果；

片段标注单元 1104，用于根据首帧标注结果和尾帧标注结果，生成目标子片段中间帧的标注结果，以获得待标注的目标子片段的标注结果；

第二确定单元 1105，用于基于目标子片段的标注结果，生成待标注视频的目标标注结果。

作为一个实施例，目标获取单元，包括：

关键提取模块，用于提取待标注视频的关键帧；

片段获得模块，用于将关键帧中相邻的两个关键帧在待标注视频所围成的视频区间划分为一个视频子片段，获得至少一个视频子片段；

目标确定模块，用于从至少一个视频子片段中确定待标注的目标子片段。

在某些实施例中，关键提取模块，包括：

幅值计算子模块，用于针对待标注视频中的图像帧，计算各图像帧的运动幅度值；

关键确定子模块，用于根据运动幅度值获得图像帧中的至少一个关键帧。

作为一个实施例，尾帧标注单元，可以包括：

5 首帧获取模块，用于获取目标子片段的首帧对应的首帧标注结果；

尾帧生成模块，用于根据首帧标注结果，利用前向传播算法确定尾帧对应的尾帧标注结果。

在一种可能的设计中，尾帧生成模块，可以包括：

10 标签传播子模块，用于利用前向传播算法，将首帧标注结果向目标子片段中未标注的图像帧进行标注结果的顺序传播，获得目标子片段中未标注图像帧的标注结果；

尾帧标注子模块，用于获取目标子片段的最后一个图像帧的标注结果作为尾帧对应的尾帧标注结果。

作为又一个实施例，片段标注单元，包括：

15 第一提取模块，用于基于首帧标注结果，结合前向传播算法，提取目标子片段的中间帧的前向传播特征。

第二提取模块，用于基于尾帧标注结果，结合后向传播算法，提取目标子片段的中间帧的后向传播特征；

20 特征融合模块，用于将前向传播特征和后向传播特征进行特征融合处理，获得中间帧的目标图像特征；

标签确定模块，用于根据目标图像特征，确定中间帧的标注结果。

在某些实施例中，特征融合模块，可以包括：

序号确定子模块，用于确定中间帧在目标子片段中的图像序号；

比值确定子模块，用于根据图像序号确定序号比值；

25 权重确定子模块，用于根据序号比值，确定前向传播权重和后向传播权重；

特征加权子模块，用于根据前向传播权重、后向传播权重、前向传播特征和后向传播特征获得中间帧的目标图像特征。

作为一个实施例，首帧标注单元，可以包括：

30 首帧标注模块，用于检测用户针对首帧执行的标注操作，获得标注操作

对应的首帧标注结果；

或者，首帧确定模块，用于获取目标子片段的前一个视频子片段，并确定前一个视频子片段的尾帧对应的尾帧标注结果为目标子片段的首帧对应的首帧标注结果。

5 作为一个实施例，第一确定单元，可以包括：

顺序确定模块，用于根据至少一个视频子片段的时间先后顺序，确定至少一个视频子片段分别对应的片段顺序；

片段遍历模块，用于按照至少一个视频子片段分别对应的片段顺序，从第一个视频片段开始，依次选择一个视频片段作为待标注的子片段，获得目标子片段。

10

本实施例提供的装置，可用于执行上述方法实施例的技术方案，其实现原理和技术效果类似，本实施例此处不再赘述。

为了实现上述实施例，本公开实施例还提供了一种电子设备。

参考图 12，其示出了适于用来实现本公开实施例的电子设备的结构示意图，该电子设备 1200 可以为终端设备或服务器。其中，终端设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、个人数字助理（Personal Digital Assistant，简称 PDA）、平板电脑（Portable Android Device，简称 PAD）、便携式多媒体播放器（Portable Media Player，简称 PMP）、车载终端（例如车载导航终端）等等的移动终端以及诸如数字 TV、台式计算机等等的固定终端。图 12 示出的电子设备仅仅是一个示例，不应对本公开实施例的功能和使用范围带来任何限制。

15

20

如图 12 所示，电子设备 1200 可以包括处理装置（例如中央处理器、图形处理器等）1201，其可以根据存储在只读存储器（Read Only Memory，简称 ROM）1202 中的程序或者从存储装置 1208 加载到随机访问存储器（Random Access Memory，简称 RAM）1203 中的程序而执行各种适当的动作和处理。在 RAM 1203 中，还存储有电子设备 1200 操作所需的各种程序和数据。处理装置 1201、ROM 1202 以及 RAM 1203 通过总线 1204 彼此相连。输入/输出（I/O）接口 1205 也连接至总线 1204。

25

通常，以下装置可以连接至 I/O 接口 1205：包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置 1206；包括

30

例如液晶显示器（Liquid Crystal Display，简称LCD）、扬声器、振动器等
的输出装置 1207；包括例如磁带、硬盘等的存储装置 1208；以及通信装置
1209。通信装置 1209 可以允许电子设备 1200 与其他设备进行无线或有线通
信以交换数据。虽然图 12 示出了具有各种装置的电子设备 1200，但是应理
5 解的是，并不要求实施或具备所有示出的装置。可以替代地实施或具备更多
或更少的装置。

特别地，根据本公开的实施例，上文参考流程图描述的过程可以被实现
为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包
括承载在计算机可读介质上的计算机程序，该计算机程序包含用于执行流程
10 图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信
装置 1209 从网络上被下载和安装，或者从存储装置 1208 被安装，或者从 ROM
1202 被安装。在该计算机程序被处理装置 1201 执行时，执行本公开实施例
的方法中限定的上述功能。

需要说明的是，本公开上述的计算机可读介质可以是计算机可读信号介
15 质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介
质例如可以是一——但不限于——电、磁、光、电磁、红外线、或半导体的系
统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例
子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、
硬盘、随机访问存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存
20 储器（EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（CD-ROM）、光
存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机
可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执
行系统、装置或者器件使用或者与其结合使用。而在本公开中，计算机可读
信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载
25 了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但
不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还
可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读信号
介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者
与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的
30 介质传输，包括但不限于：电线、光缆、RF（射频）等等，或者上述的任意

合适的组合。

上述计算机可读介质可以是上述电子设备中所包含的；也可以是单独存在，而未装配入该电子设备中。

上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被该电子设备执行时，使得该电子设备执行上述实施例所示的方法。

本公开还提供一种计算机可读存储介质，计算机可读存储介质中存储有计算机执行指令，当处理器执行计算机执行指令时，实现如上述任一实施例所提供的视频标注方法。

本公开还提供一种计算机程序产品，包括计算机程序，计算机程序被处理器执行，以配置上述任一实施例所提供的视频标注方法。

可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码，上述程序设计语言包括面向图像的程序设计语言—诸如Java、Smalltalk、C++，还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网（Local Area Network，简称LAN）或广域网（Wide Area Network，简称WAN）—连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

描述于本公开实施例中所涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现。其中，单元的名称在某种情况下并不构成对该单元本身的限定，例如，第一获取单元还可以被描述为“获取至少两个网际协议地址的单元”。

5 本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如，非限制性地，可以使用的示范类型的硬件逻辑部件包括：现场可编程门阵列（FPGA）、专用集成电路（ASIC）、专用标准产品（ASSP）、片上系统（SOC）、复杂可编程逻辑设备（CPLD）等等。

10 在本公开的上下文中，机器可读介质可以是有形的介质，其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备，或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦除可编程只读存储器（EPROM 或快闪存储器）、光纤、便捷式紧凑盘只读存储器（CD-ROM）、光学储存设备、磁储存设备、或上述内容的任何合适组合。

15 以上描述仅为本公开的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解，本公开中所涉及的公开范围，并不限于上述技术特征的特定组合而成的技术方案，同时也应涵盖在不脱离上述公开构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本公开中公开的（但不限于）具有类似功能的技术特征进行互相替换而形成的技术方案。

20 此外，虽然采用特定次序描绘了各操作，但是这不应理解为要求这些操作以所示出的特定次序或以顺序次序执行来执行。在一定环境下，多任务和并行处理可能是有利的。同样地，虽然在上面论述中包含了若干具体实现细节，但是这些不应被解释为对本公开的范围的限制。在单独的实施例的上下文中描述的某些特征还可以组合地实现在单个实施例中。相反地，在单个实施例的上下文中描述的各种特征也可以单独地或以任何合适的子组合的方式实现在多个实施例中。

25

30

尽管已经采用特定于结构特征和/或方法逻辑动作的语言描述了本主题，但是应当理解所附权利要求书中所限定的主题未必局限于上面描述的特定特征或动作。相反，上面所描述的特定特征和动作仅仅是实现权利要求书的示例形式。

5

权 利 要 求 书

1、一种视频标注方法，其特征在于，包括：

确定待标注视频中待标注的子片段，获得目标子片段；

获取所述目标子片段的首帧对应的首帧标注结果；

5 基于所述首帧标注结果，生成所述目标子片段的尾帧对应的尾帧标注结果；

根据所述首帧标注结果和所述尾帧标注结果，生成所述目标子片段中间帧的标注结果，以获得所述待标注的目标子片段的标注结果；

基于所述目标子片段的标注结果，生成所述待标注视频的目标标注结果。

10 2、根据权利要求 1 所述的方法，其特征在于，所述确定所述待标注视频中待标注的子片段，获得目标子片段，包括：

提取所述待标注视频的关键帧；

将所述关键帧中相邻的两个关键帧在所述待标注视频所围成的视频区间划分为一个视频子片段，获得至少一个视频子片段；

15 从所述至少一个视频子片段中确定待标注的所述目标子片段。

3、根据权利要求 2 所述的方法，其特征在于，所述提取待标注视频的至少一个关键帧，包括：

针对所述待标注视频中的图像帧，计算各图像帧的运动幅度值；

根据所述运动幅度值获得所述图像帧中的至少一个关键帧。

20 4、根据权利要求 1 所述的方法，其特征在于，所述基于所述首帧的标注结果生成所述目标子片段尾帧的标注结果，包括：

获取所述目标子片段的首帧对应的首帧标注结果；

根据所述首帧标注结果，利用前向传播算法确定所述尾帧对应的尾帧标注结果。

25 5、根据权利要求 4 所述的方法，其特征在于，所述根据所述首帧标注结果，利用前向传播算法确定所述尾帧对应的尾帧标注结果，包括：

利用所述前向传播算法，将所述首帧标注结果向所述目标子片段中未标注的图像帧进行标注结果的顺序传播，获得所述目标子片段中未标注图像帧的标注结果；

30 获取所述目标子片段的最后一个图像帧的标注结果作为所述尾帧对应的

尾帧标注结果。

6、根据权利要求1所述的方法，其特征在于，所述根据所述首帧标注结果和所述尾帧标注结果，生成所述目标子片段中间帧的标注结果，包括：

5 基于所述首帧标注结果，结合前向传播算法，提取所述目标子片段的中间帧的前向传播特征；

基于所述尾帧标注结果，结合后向传播算法，提取所述目标子片段的中间帧的后向传播特征；

将所述前向传播特征和所述后向传播特征进行特征融合处理，获得所述中间帧的目标图像特征；

10 根据所述目标图像特征，确定所述中间帧的标注结果。

7、根据权利要求6所述的方法，其特征在于，所述将所述前向传播特征和所述后向传播特征进行特征融合处理，获得所述中间帧的目标图像特征，包括：

确定所述中间帧在所述目标子片段中的图像序号；

15 根据所述图像序号确定序号比值；

根据所述序号比值，确定前向传播权重和后向传播权重；

根据所述前向传播权重、所述后向传播权重、前向传播特征和所述后向传播特征获得所述中间帧的目标图像特征。

20 8、根据权利要求1所述的方法，其特征在于，所述获取所述目标子片段的首帧对应的首帧标注结果，包括：

检测用户针对所述首帧执行的标注操作，获得所述标注操作对应的首帧标注结果；

25 或，获取所述目标子片段的前一个视频子片段，并确定所述前一个视频子片段的尾帧对应的尾帧标注结果为所述目标子片段的首帧对应的首帧标注结果。

9、根据权利要求1所述的方法，其特征在于，所述确定待标注视频中待标注的子片段，获得目标子片段，包括：

根据至少一个视频子片段的时间先后顺序，确定至少一个视频子片段分别对应的片段顺序；

30 按照至少一个视频子片段分别对应的片段顺序，从第一个视频片段开始，

依次选择一个视频片段作为待标注的子片段，获得所述目标子片段。

10、一种视频标注装置，其特征在于，包括：

第一确定单元，用于确定待标注视频中待标注的子片段，获得目标子片段；

5 首帧标注单元，用于获取所述目标子片段的首帧对应的首帧标注结果；

尾帧标注单元，用于基于所述首帧标注结果，生成所述目标子片段的尾帧对应的尾帧标注结果；

片段标注单元，用于根据所述首帧标注结果和所述尾帧标注结果，生成所述目标子片段中间帧的标注结果，以获得所述待标注的目标子片段的标注结果；

10

第二确定单元，用于基于所述目标子片段的标注结果，生成所述待标注视频的目标标注结果。

11、一种电子设备，其特征在于，包括：处理器、存储器；

所述存储器存储计算机执行指令；

15 所述处理器执行所述存储器存储的计算机执行指令，使得所述处理器配置有如权利要求 1 至 9 任一项所述的视频标注方法。

12、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中存储有计算机执行指令，当处理器执行所述计算机执行指令时，实现如权利要求 1 至 9 任一项所述的视频标注方法。

20 13、一种计算机程序产品，包括计算机程序，其特征在于，所述计算机程序被处理器执行，以配置有如权利要求 1 至 9 任一项所述的视频标注方法。

25

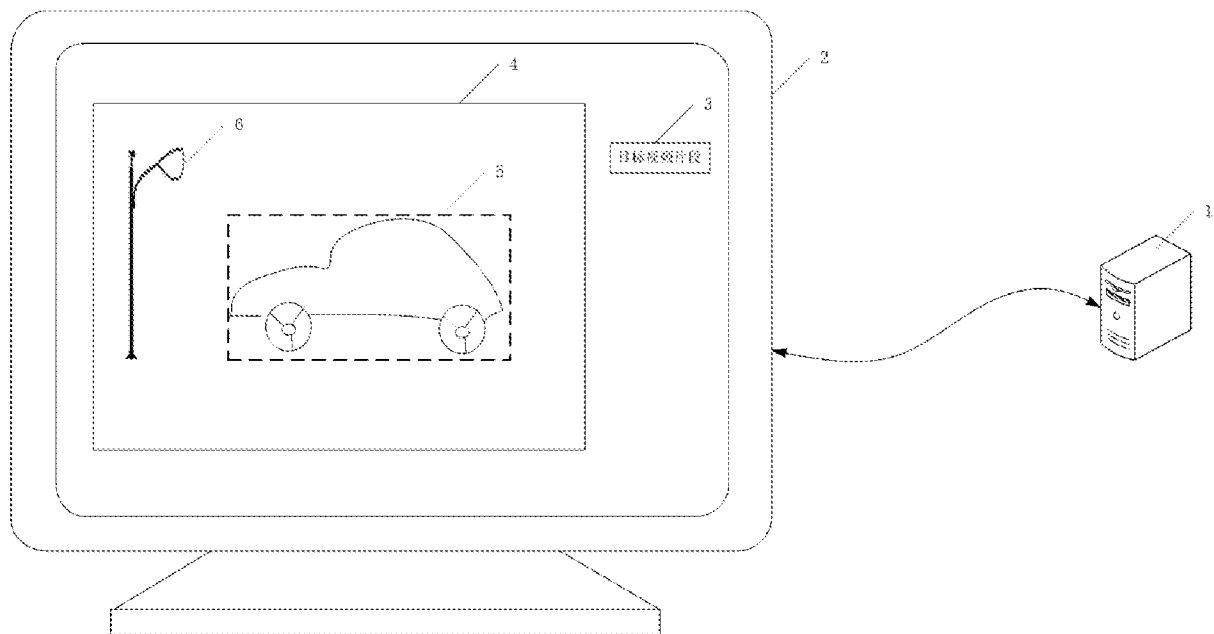


图 1

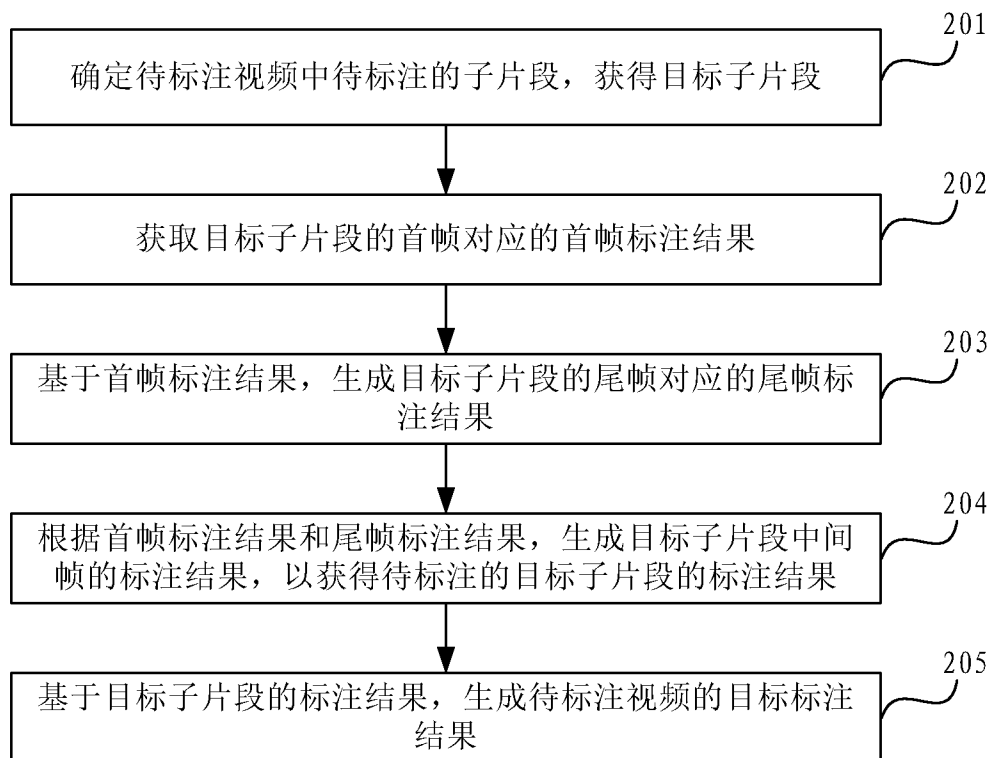


图 2

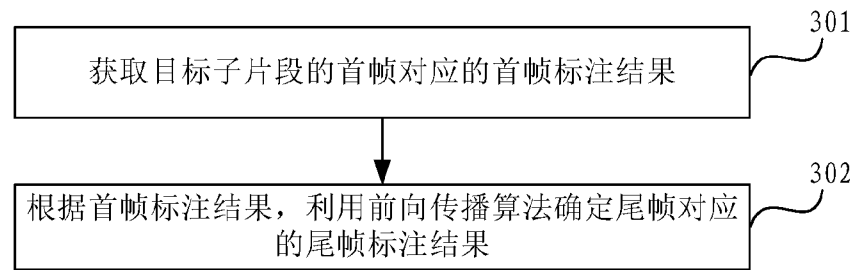


图 3

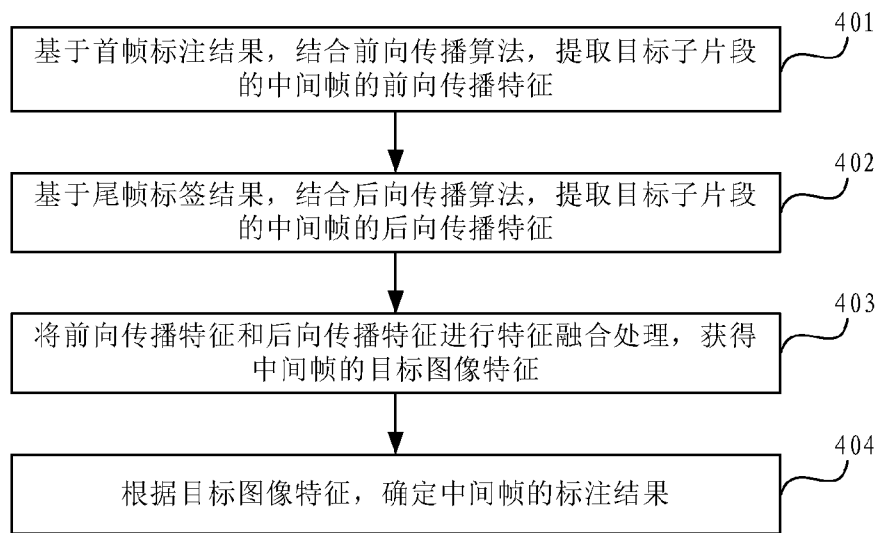


图 4

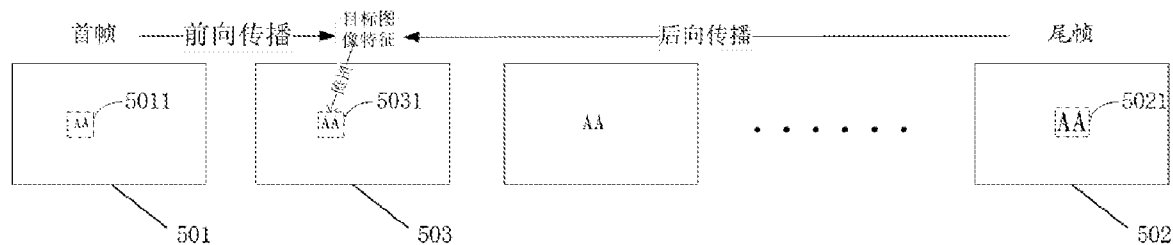


图 5

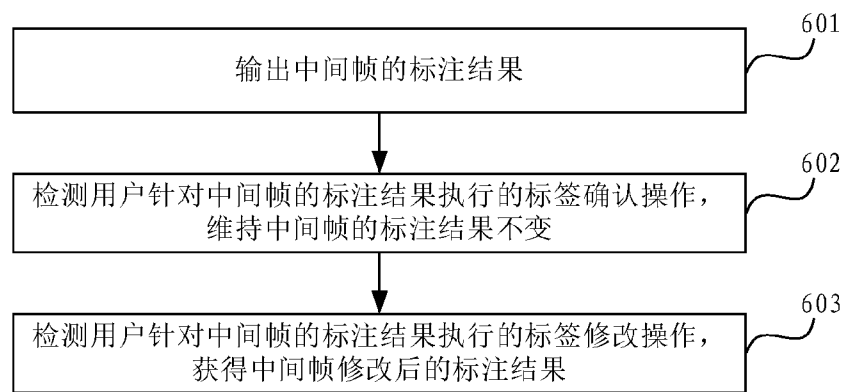


图 6

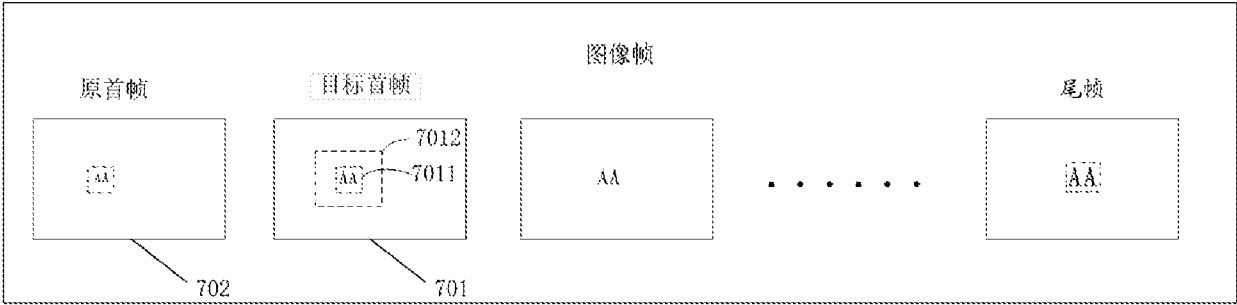


图 7

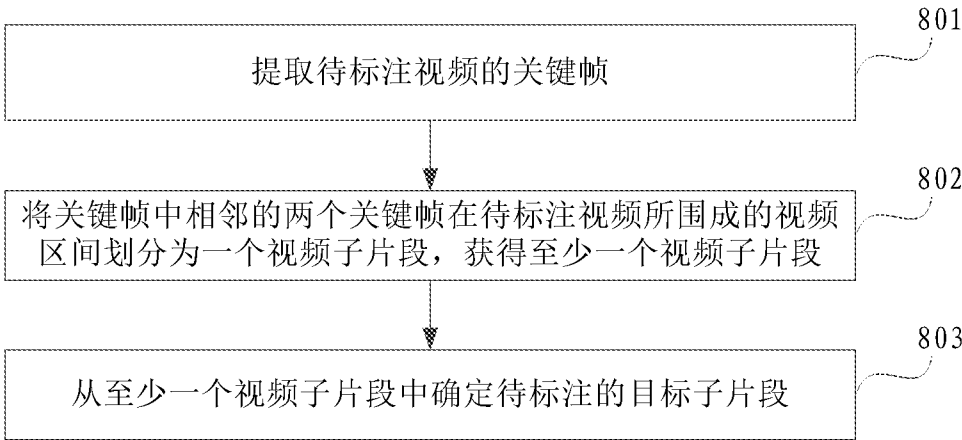


图 8

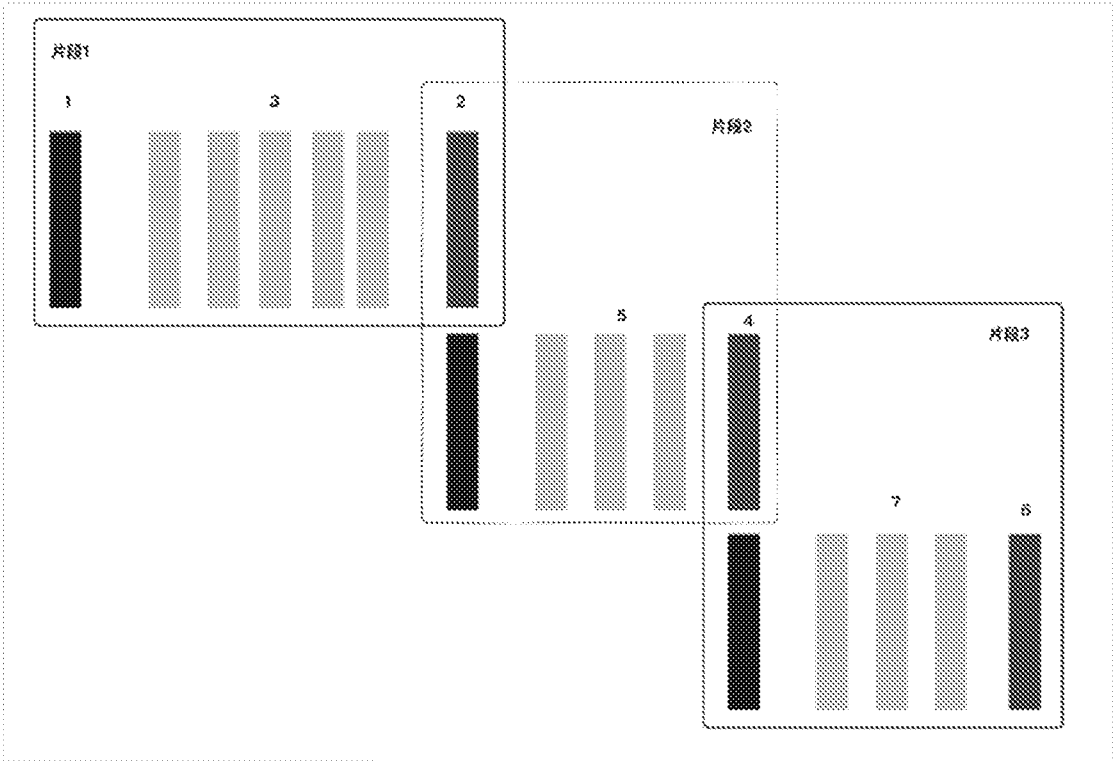


图 9

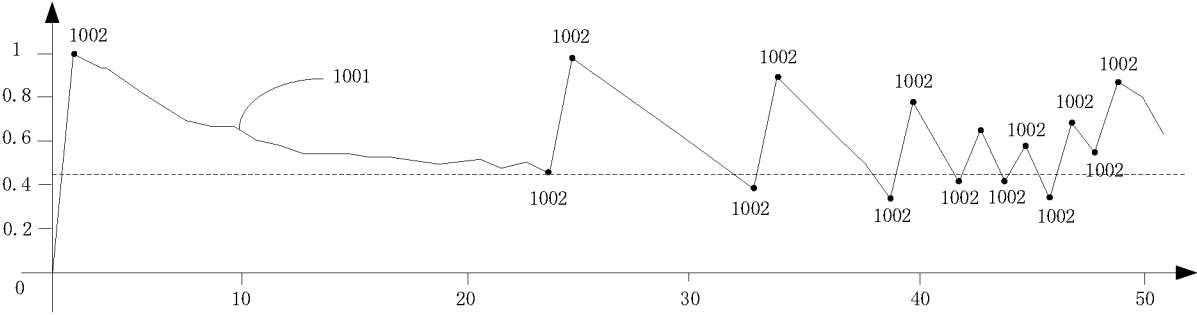


图 10

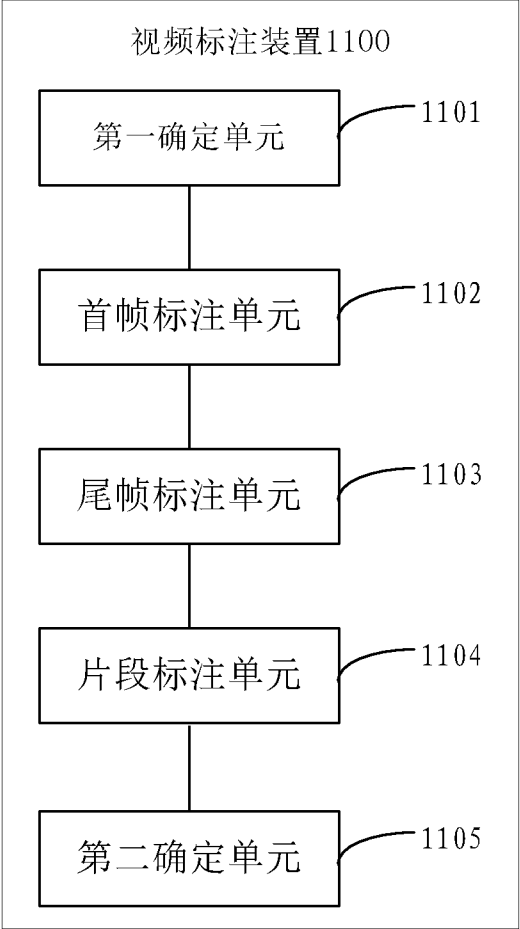


图 11

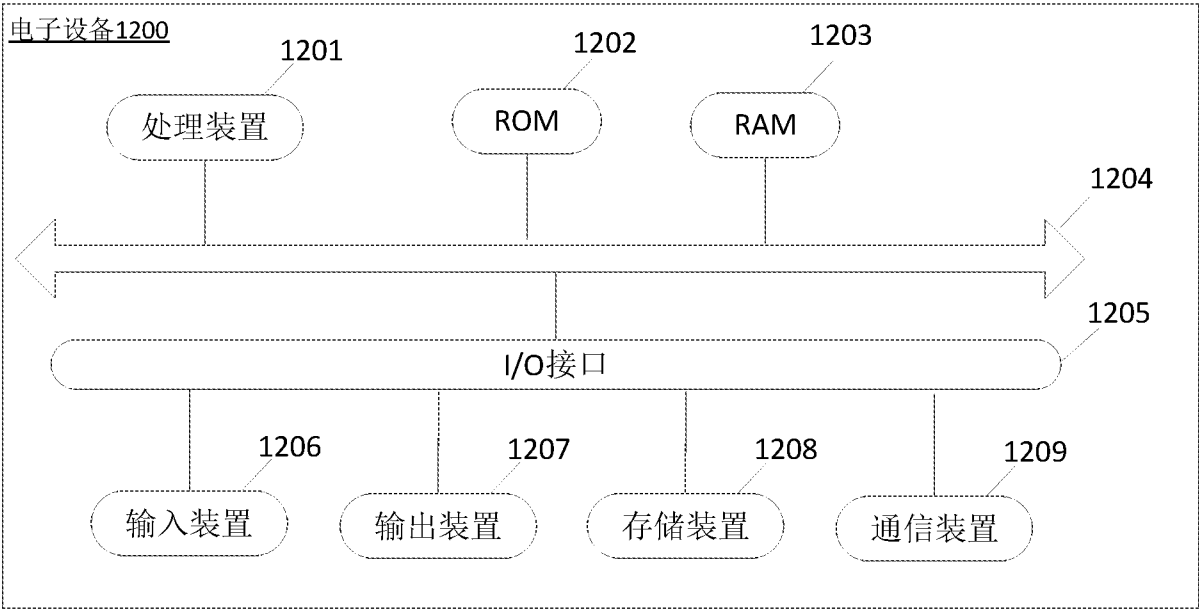


图 12

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/130577

A. CLASSIFICATION OF SUBJECT MATTER

G06F16/783(2019.01)i; G06F16/78(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G06F

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)

VEN, CNABS, CNTXT, WOTXT, EPTXT, USTXT, CNKI, IEEE: 视频, 图像, 标注, 片段, 首帧, 尾帧, 中间帧, 关键帧, 目标, 划分, 生成, 前, 后, video, image, label, fragment, frame, first, end, key, intermediate, target, divide, generate, front, back

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 115905622 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 04 April 2023 (2023-04-04) claims 1-13	1-13
Y	CN 112053323 A (TUSIMPLE, INC.) 08 December 2020 (2020-12-08) description, paragraphs [0003]-[0006] and [0025]-[0035]	1-13
Y	CN 114117128 A (HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.) 01 March 2022 (2022-03-01) description, paragraphs [0004]-[0048]	1-13
A	CN 111581433 A (GUANGDONG OPPO MOBILE COMMUNICATIONS CO., LTD.) 25 August 2020 (2020-08-25) entire document	1-13
A	CN 113378958 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 10 September 2021 (2021-09-10) entire document	1-13



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

“D” document cited by the applicant in the international application

“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

12 January 2024

Date of mailing of the international search report

18 January 2024

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/130577

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2021081671 A1 (BEIJING XIAOMI MOBILE SOFTWARE CO., LTD.) 18 March 2021 (2021-03-18) entire document	1-13

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/130577

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	115905622	A	04 April 2023	None			
CN	112053323	A	08 December 2020	None			
CN	114117128	A	01 March 2022	None			
CN	111581433	A	25 August 2020	None			
CN	113378958	A	10 September 2021	None			
US	2021081671	A1	18 March 2021	EP	3792818	A1	17 March 2021
				US	11288514	B2	29 March 2022

A. 主题的分类

G06F16/783(2019.01)i; G06F16/78(2019.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: G06F

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))

VEN, CNABS, CNTXT, WOTXT, EPTXT, USTXT, CNKI, IEEE: 视频, 图像, 标注, 片段, 首帧, 尾帧, 中间帧, 关键帧, 目标, 划分, 生成, 前, 后, video, image, label, fragment, frame, first, end, key, intermediate, target, divide, generate, front, back

C. 相关文件

类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 115905622 A (北京字跳网络技术有限公司) 2023年4月4日 (2023 - 04 - 04) 权利要求1-13	1-13
Y	CN 112053323 A (上海图森未来人工智能科技有限公司) 2020年12月8日 (2020 - 12 - 08) 说明书第[0003]-[0006]段、[0025]-[0035]段	1-13
Y	CN 114117128 A (华为云计算技术有限公司) 2022年3月1日 (2022 - 03 - 01) 说明书第[0004]-[0048]段	1-13
A	CN 111581433 A (OPPO广东移动通信有限公司) 2020年8月25日 (2020 - 08 - 25) 全文	1-13
A	CN 113378958 A (北京百度网讯科技有限公司) 2021年9月10日 (2021 - 09 - 10) 全文	1-13
A	US 2021081671 A1 (BEIJING XIAOMI MOBILE SOFTWARE CO., LTD.) 2021年3月18日 (2021 - 03 - 18) 全文	1-13

☐ 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

“D” 申请人在国际申请中引证的文件

“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

2024年1月12日

国际检索报告邮寄日期

2024年1月18日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

焦月

电话号码 (+86) 010-53961306

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2023/130577

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	115905622	A	2023年4月4日	无			
CN	112053323	A	2020年12月8日	无			
CN	114117128	A	2022年3月1日	无			
CN	111581433	A	2020年8月25日	无			
CN	113378958	A	2021年9月10日	无			
US	2021081671	A1	2021年3月18日	EP	3792818	A1	2021年3月17日
				US	11288514	B2	2022年3月29日