



(12)发明专利申请

(10)申请公布号 CN 110969640 A

(43)申请公布日 2020.04.07

(21)申请号 201811148953.7

(22)申请日 2018.09.29

(71)申请人 TCL集团股份有限公司

地址 516006 广东省惠州市仲恺高新技术
开发区十九号小区

(72)发明人 俞大海 张鼎文

(74)专利代理机构 深圳中一专利商标事务所
44237

代理人 官建红

(51)Int.Cl.

G06T 7/194(2017.01)

权利要求书3页 说明书10页 附图3页

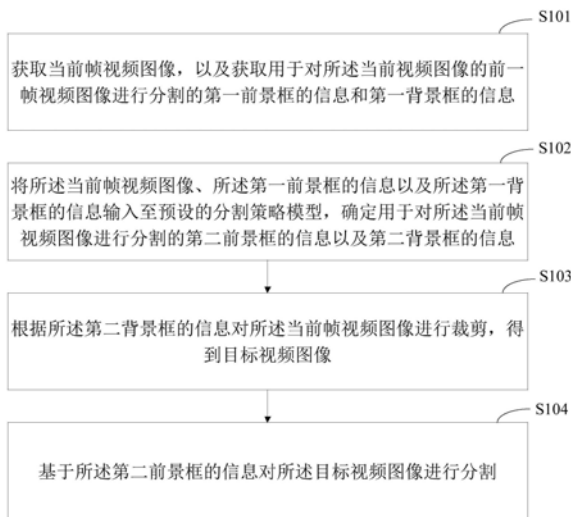
(54)发明名称

视频图像的分割方法、终端设备以及计算机
可读存储介质

(57)摘要

本发明适用于计算机技术领域,提供了视频
图像的分割方法、终端设备及计算机可读存储介
质,包括:获取当前帧视频图像,以及获取用于对
所述当前帧视频图像的前一帧视频图像进行分割
的第一前景框的信息和第一背景框的信息;所述
信息包括位置信息和尺寸信息;将所述当前帧视
频图像、所述第一前景框的信息以及所述第一背
景框的信息输入至预设的分割策略模型,确定用
于对所述当前帧视频图像进行分割的第二前景
框的信息以及第二背景框的信息;根据所述第二
背景框的信息对所述当前帧视频图像进行裁剪,
得到目标视频图像;基于所述第二前景框的信息
对所述目标视频图像进行分割,从而能够减少视
频图像分割过程中的计算量,且提高对视频图像
进行分割的效率。

CN 110969640 A



1. 一种视频图像的分割方法,其特征在于,包括:

获取当前帧视频图像,以及获取用于对所述当前视频图像的前一帧视频图像进行分割的第一前景框的信息和第一背景框的信息;所述信息包括位置信息和尺寸信息;

将所述当前帧视频图像、所述第一前景框的信息以及所述第一背景框的信息输入至预设的分割策略模型,确定用于对所述当前帧视频图像进行分割的第二前景框的信息以及第二背景框的信息;

根据所述第二背景框的信息对所述当前帧视频图像进行裁剪,得到目标视频图像;

基于所述第二前景框的信息对所述目标视频图像进行分割。

2. 根据权利要求1所述的视频图像的分割方法,其特征在于,所述将所述当前帧视频图像、所述第一前景框的信息以及所述第一背景框的信息输入至预设的分割策略模型,确定用于对所述当前帧视频图像进行分割的第二前景框的信息以及第二背景框的信息之前,还包括:

获取用于对第一训练模型进行训练的第一样本数据,并将所述第一样本数据分成第一训练集和第一测试集;其中,所述第一样本数据中的每条样本数据均包括同一视频对应的所有帧的视频图像及所述视频图像对应的前景框的信息以及背景框的信息;

将所述第一训练集中的每条样本数据包含的同一视频对应的所有帧的视频图像、每一帧所述视频图像的前一帧视频图像对应的前景框的信息及背景框的信息作为所述第一训练模型的输入,将每一帧所述视频图像的前景框的信息及背景框的信息作为所述第一训练模型的输出,对所述第一训练模型进行训练;

将所述第一测试集中的每条样本数据包含的同一视频对应的所有帧的视频图像输入至训练好的所述第一训练模型,得到每一帧所述视频图像对应的预测前景框的信息以及预测背景框的信息;

基于每一帧所述视频图像对应的前景框的信息、预测前景框的信息对训练好的所述第一训练模型进行校验;

若校验通过,则将训练好的所述第一训练模型识别为分割策略模型。

3. 根据权利要求2所述的视频图像的分割方法,其特征在于,所述基于每一帧所述视频图像对应的前景框的信息、预测前景框的信息对训练好的所述第一训练模型进行校验,包括:

基于以下公式计算每一帧所述视频图像对应的前景框与所述每一帧所述视频图像对应的预测前景框之间的交并比:

$$IoU = \frac{area(m) \cap area(y)}{area(m) \cup area(y)};$$

其中,所述m为所述每一帧所述视频图像对应的前景框的位置信息和尺寸信息,所述y为所述每一帧所述视频图像对应的预测前景框的位置信息和尺寸信息,所述IoU为所述每一帧所述视频图像对应的前景框与所述每一帧所述视频图像对应的预测前景框之间的交并比,所述area(m)为所述每一帧所述视频图像对应的前景框的面积,所述area(y)为所述每一帧所述视频图像对应的预测前景框,所述area(m) ∩ area(y)为所述每一帧所述视频图像对应的前景框的面积与所述每一帧所述视频图像对应的预测前景框的面积之间的交集,所述area(m) ∪ area(y)为所述每一帧所述视频图像对应的前景框的面积与所述每一帧所

述视频图像对应的预测前景框的面积之间的交集；

基于所述交并比对所述第一训练模型进行校验。

4. 根据权利要求3所述的视频图像的分割方法,其特征在於,所述若校验通过,则将训练好的所述第一训练模型识别为分割策略模型,包括:

若所述交并比均大于或等于预设阈值,则识别为校验通过,将训练好的所述第一训练模型识别为分割策略模型。

5. 根据权利要求1所述的视频图像的分割方法,其特征在於,所述将所述当前帧视频图像、所述第一前景框的信息以及所述第一背景框的信息输入至预设的分割策略模型,确定用于对所述当前帧视频图像进行分割的第二前景框的信息以及第二背景框的信息,包括:

对所述当前帧视频图像进行特征提取,得到所述当前帧视频图像的特征图;

对所述特征图进行卷积计算,得到所述当前帧视频图像的特征向量;

根据所述特征向量和所述第一前景框的信息确定用于对所述当前帧视频图像进行分割的第二前景框的信息;

根据所述特征向量和所述第一背景框的信息确定用于对所述当前帧视频图像进行分割的第二背景框的信息。

6. 一种终端设备,其特征在於,包括:

第一获取单元,用于获取当前帧视频图像,以及获取用于对所述当前帧视频图像的前一帧视频图像进行分割的第一前景框的信息和第一背景框的信息;所述信息包括位置信息和尺寸信息;

确定单元,用于将所述当前帧视频图像、所述第一前景框的信息以及所述第一背景框的信息输入至预设的分割策略模型,确定用于对所述当前帧视频图像进行分割的第二前景框的信息以及第二背景框的信息;

裁剪单元,用于根据所述第二背景框的信息对所述当前帧视频图像进行裁剪,得到目标视频图像;

分割单元,用于基于所述第二前景框的信息对所述目标视频图像进行分割。

7. 根据权利要求4所述的终端设备,其特征在於,还包括:

第二获取单元,用于获取用于对第一训练模型进行训练的第一样本数据,并将所述第一样本数据分成第一训练集和第一测试集;其中,所述第一样本数据中的每条样本数据均包括同一视频对应的所有帧的视频图像及所述视频图像对应的前景框的信息以及背景框的信息;

训练单元,用于将所述第一训练集中的每条样本数据包含的同一视频对应的所有帧的视频图像、每一帧所述视频图像的前一帧视频图像对应的前景框的信息及背景框的信息作为所述第一训练模型的输入,将每一帧所述视频图像的前景框的信息及背景框的信息作为所述第一训练模型的输出,对所述第一训练模型进行训练;

预测单元,用于将所述第一测试集中的每条样本数据包含的同一视频对应的所有帧的视频图像输入至训练好的所述第一训练模型,得到每一帧所述视频图像对应的预测前景框的信息以及预测背景框的信息;

校验单元,用于基于每一帧所述视频图像对应的前景框的信息、预测前景框的信息对训练好的所述第一训练模型进行校验;

识别单元,用于若校验通过,则将训练好的所述第一训练模型识别为分割策略模型。

8. 根据权利要求7所述的终端设备,其特征在于,所述校验单元,包括:

第一计算子单元,用于基于以下公式计算每一帧所述视频图像对应的前景框与所述每一帧所述视频图像对应的预测前景框之间的交并比:

$$IoU = \frac{area(m) \cap area(y)}{area(m) \cup area(y)};$$

其中,所述m为所述每一帧所述视频图像对应的前景框的位置信息和尺寸信息,所述y为所述每一帧所述视频图像对应的预测前景框的位置信息和尺寸信息,所述IoU为所述每一帧所述视频图像对应的前景框与所述每一帧所述视频图像对应的预测前景框之间的交并比,所述area(m)为所述每一帧所述视频图像对应的前景框的面积,所述area(y)为所述每一帧所述视频图像对应的预测前景框,所述area(m) ∩ area(y)为所述每一帧所述视频图像对应的前景框的面积与所述每一帧所述视频图像对应的预测前景框的面积之间的交集,所述area(m) ∪ area(y)为所述每一帧所述视频图像对应的前景框的面积与所述每一帧所述视频图像对应的预测前景框的面积之间的并集;

基于所述交并比对所述第一训练模型进行校验。

9. 一种终端设备,包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现如权利要求1至5任意一项所述方法的步骤。

10. 一种计算机可读存储介质,所述计算机可读存储介质存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现如权利要求1至5任意一项所述方法的步骤。

视频图像的分割方法、终端设备以及计算机可读存储介质

技术领域

[0001] 本发明属于计算机技术领域,尤其涉及视频图像的分割方法、终端设备以及计算机可读存储介质。

背景技术

[0002] 图像分割是指将图像中的特定目标作为前景,对其进行检测、定位,并将其从图像的背景中准确地分割出来的过程。现有技术通常采用图像分割算法来对单帧视频图像进行分割。在对视频图像进行分割时,现有技术是针对视频中的每一帧视频图像单独进行图像分割,即采用相同的图像分割算法对视频中的每一帧图像分别进行分割,从而达到对整体视频的分割结果。这种分割方法的缺点在于计算量大,且效率低,因此无法满足产业的实际应用需求。

发明内容

[0003] 有鉴于此,本发明实施例提供了视频图像的分割方法、终端设备以及计算机可读存储介质,以解决现有的视频图像分割方法存在的效率低,计算量大的问题。

[0004] 本发明实施例的第一方面提供了一种视频图像的分割方法,包括:

[0005] 获取当前帧视频图像,以及获取用于对所述当前视频图像的前一帧视频图像进行分割的第一前景框的信息和第一背景框的信息;所述信息包括位置信息和尺寸信息;

[0006] 将所述当前帧视频图像、所述第一前景框的信息以及所述第一背景框的信息输入至预设的分割策略模型,确定用于对所述当前帧视频图像进行分割的第二前景框的信息以及第二背景框的信息;

[0007] 根据所述第二背景框的信息对所述当前帧视频图像进行裁剪,得到目标视频图像;

[0008] 基于所述第二前景框的信息对所述目标视频图像进行分割。

[0009] 本发明实施例的第二方面提供了一种终端设备,包括:

[0010] 第一获取单元,用于获取当前帧视频图像,以及获取用于对所述当前视频图像的前一帧视频图像进行分割的第一前景框的信息和第一背景框的信息;所述信息包括位置信息和尺寸信息;

[0011] 确定单元,用于将所述当前帧视频图像、所述第一前景框的信息以及所述第一背景框的信息输入至预设的分割策略模型,确定用于对所述当前帧视频图像进行分割的第二前景框的信息以及第二背景框的信息;

[0012] 裁剪单元,用于根据所述第二背景框的信息对所述当前帧视频图像进行裁剪,得到目标视频图像;

[0013] 分割单元,用于基于所述第二前景框的信息对所述目标视频图像进行分割。

[0014] 本发明实施例的第三方面提供了一种终端设备,包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时

实现上述方法的步骤。

[0015] 本发明实施例的第四方面提供了一种计算机可读存储介质,所述计算机可读存储介质存储有计算机程序,所述计算机程序被处理器执行时实现上述方法的步骤。

[0016] 本发明实施例与现有技术相比存在的有益效果是:将所述当前帧视频图像、所述第一前景框的信息以及所述第一背景框的信息输入至预设的分割策略模型,该分割策略模型将以前一帧视频图像进行分割的前景框以及背景框作为参考,能快速确定得到用于对当前帧视频图像进行分割的第二前景框的信息以及第二背景框的信息,即得到对当前帧视频图像进行分割的前景框的位置和尺寸以及背景框的位置和尺寸,从而可以实现提高分割的效率且可以减少在分割过程中的计算量;便于终端设备确定对当前帧视频图像进行分割的前景框和背景框在当前帧视频图像中的具体位置的信息。

附图说明

[0017] 为了更清楚地说明本发明实施例中的技术方案,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发明的一些实施例,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。

[0018] 图1是本发明第一实施例提供的视频图像的分割方法的实现流程图;

[0019] 图2是本发明第二实施例提供的视频图像的分割方法的实现流程图;

[0020] 图3是本发明第三实施例提供的终端设备的示例图;

[0021] 图4是本发明第四实施例提供的终端设备的示意图。

具体实施方式

[0022] 以下描述中,为了说明而不是为了限定,提出了诸如特定系统结构、技术之类的具体细节,以便透彻理解本发明实施例。然而,本领域的技术人员应当清楚,在没有这些具体细节的其它实施例中也可以实现本发明。在其它情况中,省略对众所周知的系统、装置、电路以及方法的详细说明,以免不必要的细节妨碍本发明的描述。

[0023] 为了说明本发明所述的技术方案,下面通过具体实施例来进行说明。

[0024] 图像分割是指将图像中的特定目标作为前景,对其进行检测、定位,并将其从图像的背景中准确地分割出来的过程,而在对视频进行分割的时候,是指对整体视频的前景进行分割的过程。

[0025] 参阅图1,图1是本发明第一实施例提供的一种视频图像的分割方法的实现流程图,本实施例中视频图像的分割方法的执行主体为用终端设备。

[0026] 如图1所示的视频图像的分割方法可包括以下步骤:

[0027] 步骤S101,获取当前帧视频图像,以及获取用于对所述当前视频图像的前一帧视频图像进行分割的第一前景框的信息和第一背景框的信息;所述信息包括位置信息和尺寸信息。

[0028] 在需要对视频进行整体分割时,可以将需要进行分割处理的原始视频导入终端设备中,终端设备会对需要进行分割处理的原始视频进行帧提取操作,得到该视频对应的所有帧的视频图像。

[0029] 对视频进行整体分割时,需要先将视频中的每一帧视频图像进行前景分割得到前景目标图像,根据每一帧视频图像进行前景分割得到前景目标图像实现对视频进行整体分割前景目标图像的目的。而在对该视频中的某一帧视频图像进行分割时,终端设备会将需要进行分割的该帧视频图像作为当前帧视频图像,本实施例中,终端设备还获取用于对当前帧视频图像的前一帧视频图像进行分割的第一前景框的信息和第一背景框的信息,其中前一帧视频图像指的是在该视频中与当前帧视频图像相邻的前一帧视频图像;而上述信息包括位置信息和尺寸信息,用于作为确定对前一帧视频图像进行分割的前景框在前一帧视频图像中的具体位置信息以及对应的前景框大小,以及作为确定对前一帧视频图像进行分割的背景框在前一帧视频图像中的具体位置信息以及对应的背景框大小。而再对当前帧视频图像进行分割时,会使用对前一帧视频图像进行分割的前景框和背景框作为参考。

[0030] 步骤S102,将所述当前帧视频图像、所述第一前景框的信息以及所述第一背景框的信息输入至预设的分割策略模型,确定用于对所述当前帧视频图像进行分割的第二前景框的信息以及第二背景框的信息。

[0031] 终端设备会将当前帧视频图像、第一前景框的信息以及第一背景框的信息输入至预设的分割策略模型中,其中预设的分割策略模型作为训练好的模型存储于终端设备中;该分割策略模型将以前一帧视频图像进行分割的前景框以及背景框作为参考,能快速确定得到用于对当前帧视频图像进行分割的第二前景框的信息以及第二背景框的信息,即得到对当前帧视频图像进行分割的前景框的位置和尺寸以及背景框的位置和尺寸,从而实现可以减少在分割过程中的计算量且提高分割的效率;便于终端设备确定对当前帧视频图像进行分割的前景框和背景框在当前帧视频图像中的具体位置以及对应的框大小。

[0032] 作为本发明一实施例,步骤S102可以包括以下步骤:

[0033] 对所述当前帧视频图像进行特征提取,得到所述当前帧视频图像的特征图;

[0034] 对所述特征图进行卷积计算,得到所述当前帧视频图像的特征向量;

[0035] 根据所述特征向量和所述第一前景框的信息确定用于对所述当前帧视频图像进行分割的第二前景框的信息;

[0036] 根据所述特征向量和所述第一背景框的信息确定用于对所述当前帧视频图像进行分割的第二背景框的信息。

[0037] 在对当前帧视频图像进行分割时,终端设备需要对当前帧视频图像进行特征提取得到当前帧视频图像的特征图,便于作为输入数据输入至训练好的分割策略模型中进行处理,其中分割策略模型的输入数据还需要包含对前一帧视频图像进行分割的前景框的位置和尺寸信息以及背景框的位置和尺寸信息。

[0038] 分割策略模型包括卷积层和全连接层,其中分割策略模型的卷积层具体用于对当前帧视频图像的特征图进行卷积计算,得到当前帧视频图像的特征向量。可选的,分割策略模型的卷积层具体可以包括第一卷积层、第二卷积层以及第三卷积层;分割策略模型的卷积层对输出的当前帧视频图像的特征图进行卷积计算时,第一卷积层用于对当前帧视频图像的特征图进行卷积计算得到 $288 \times 7 \times 7$ 的第一特征图;第二卷积层用于对第一特征图进行卷积计算得到 $112 \times 5 \times 5$ 的第二特征图;第三卷积层用于对第二特征图进行卷积计算得到 $56 \times 3 \times 3$ 的第三特征图,并对 $56 \times 3 \times 3$ 的第三特征图进行变形处理得到对应的特征向量,上述特征向量具体为504维的向量。

[0039] 分割策略模型中的全连接层具体包括两个全连接层结构,其中一个全连接层用于根据当前帧视频图像的特征向量,以及参考对前一帧视频图像进行分割的前景框的位置和尺寸,通过计算确定得到用于对当前帧视频图像进行分割的第二前景框的位置和尺寸。另外一个全连接层则用于根据当前帧视频图像的特征向量,以及参考对前一帧视频图像进行分割的背景框的位置和尺寸,通过计算确定得到用于对当前帧视频图像进行分割的第二背景框的位置和尺寸。由于分割策略模型会参考前一帧视频图像进行分割的前景框的位置和尺寸来确定对当前帧视频图像进行分割的背景框的位置和尺寸,以及参考前一帧视频图像进行分割的背景框的位置和尺寸来确定对当前帧视频图像进行分割的背景框的位置和尺寸,因此可以极大的减少在确定当前帧视频图像进行分割的前景框以及背景框时的计算量,从而能显著提高分割的效率。

[0040] 步骤S103,根据所述第二背景框的信息对所述当前帧视频图像进行裁剪,得到目标视频图像。

[0041] 通过分割策略模型确定得到对当前帧视频图像进行分割前景框的位置和尺寸以及背景框的位置和尺寸后,终端设备将根据背景框的位置和尺寸对当前帧视频图像进行裁剪,得到裁剪后的目标视频图像,实现对当前帧视频图像进行背景裁剪的目的。

[0042] 步骤S104,基于所述第二前景框的信息对所述目标视频图像进行分割。

[0043] 对于经过裁剪后的目标视频图像,终端设备将根据前景框的位置和尺寸对当前帧视频图像进行分割得到前景目标图像。需要指出的是,本实施例中根据前景框的位置和尺寸对当前帧视频图像进行分割得到前景目标图像的方法具体采用的是通过预设的分割执行模型来进行分割,其中上述分割执行模型需要经过训练得到,其存储于终端设备中。

[0044] 其中对上述分割执行模型进行训练的方法为获取用于对分割执行模型进行训练的样本数据,并将样本数据分成训练集和测试集;其中,样本数据中的每条样本数据均包括视频图像及视频图像对应的实际需要分割的前景框的位置和尺寸;将训练集中的每条样本数据中包含视频图像以及视频图像对应的实际需要分割的前景框的位置和尺寸作为分割执行模型的输入,将视频图像对应的实际需要分割的前景框的位置和尺寸作为原始训练模型的输出,对原始训练模型进行训练。

[0045] 在校验训练后的原始训练模型是否训练完成时,通过将测试集中的每条样本数据包含的视频图像输入至训练后的原始训练模型,得到视频图像对应的预测前景框的位置和尺寸;通过将视频图像对应的前景框的位置和尺寸与预测的前景框的位置和尺寸进行比对标校。

[0046] 其中比对标校的方法具体是,判断预测的前景框与视频图像对应的实际需要分割的前景框之间计算得到的交并比是否满足大于或等于预设阈值的条件;当预测的前景框与视频图像对应的实际需要分割的前景框之间计算得到的交并比满足大于或等于预设阈值的条件时,则说明校验通过,此时则可以将原始训练模型识别为分割策略模型,否则,则需要继续对原始训练模型进行训练,通过上述分割执行模型能快速、准确地执行根据前景框的位置和尺寸对当前帧视频图像进行分割得到前景目标。

[0047] 参阅图2,图2是本发明第二实施例提供的视频图像的分割方法的实现流程图。本实施例与上一实施例的区别在于,本实施例中在步骤S102之前还包括步骤S201~S205。其中步骤S101~S104的内容具体请参阅上一实施例中步骤S101~S104的相关描述,此处不赘

述。S201~S205具体如下：

[0048] S201, 获取用于对第一训练模型进行训练的第一样本数据, 并将所述第一样本数据分成第一训练集和第一测试集; 其中, 所述第一样本数据中的每条样本数据均包括同一视频对应的所有帧的视频图像及所述视频图像对应的前景框的信息以及背景框的信息。

[0049] 终端设备在将当前帧视频图像、第一前景框的信息以及第一背景框的信息输入至预设的分割策略模型之前, 需要对分割策略模型进行训练。其中终端设备通过获取大量的第一样本数据来对第一训练模型进行训练, 以期望得到分割策略模型, 其中第一样本数据中的每条样本数据均包括同一视频对应的所有帧的视频图像及视频图像对应的前景框的位置信息和尺寸信息, 以及背景框的位置信息和尺寸信息。终端设备会将第一样本数据随机分成第一训练集和第一测试集, 其中第一训练集的样本数据作为对第一训练模型进行训练的样本数据, 而第一测试集的样本数据作为对第一训练模型进行校验的样本数据。

[0050] S202, 将所述第一训练集中的每条样本数据包含的同一视频对应的所有帧的视频图像、每一帧所述视频图像的前一帧视频图像对应的前景框的位置信息和尺寸信息及背景框的位置信息和尺寸信息作为第一训练模型的输入, 将每一帧所述视频图像的前景框的位置信息和尺寸信息及背景框的位置信息和尺寸信息作为第一训练模型的输出, 对第一训练模型进行训练。

[0051] 将第一训练集中的每条样本数据包含的同一视频对应的所有帧的视频图像、每一帧视频图像的前一帧视频图像对应的前景框的信息及背景框的信息作为所述第一训练模型的输入, 将每一帧所述视频图像的前景框的信息及背景框的信息作为所述第一训练模型的输出, 对所述第一训练模型进行训练。

[0052] 需要说明的是, 第一训练模型需要学习的是如何根据前一帧视频图像对应的前景框的位置信息和尺寸信息及背景框的位置信息和尺寸信息来对当前帧视频图像进行分割, 因为视频图像中相邻的两帧视频图像之间存在较大的关联, 对应的前景框和背景框的位置和尺寸也存在较大的关联, 因此将每一帧视频图像的前一帧视频图像对应的前景框的位置和尺寸及背景框的位置和尺寸作为输入数据。

[0053] S203, 将所述第一测试集中的每条样本数据包含的同一视频对应的所有帧的视频图像输入至训练好的所述第一训练模型, 得到每一帧所述视频图像对应的预测前景框的信息以及预测背景框的信息。

[0054] 对于训练完成的第一训练模型, 终端设备通过第一测试集中的样本数据来进行验证, 其中具体是将第一测试集中的每条样本数据包含的同一视频对应的所有帧的视频图像输入至训练好的第一训练模型中, 得到每一帧视频图像对应的预测前景框的信息以及预测背景框的信息, 得到进行校验的数据。

[0055] S204, 基于每一帧所述视频图像对应的前景框的信息、预测前景框的信息对训练好的所述第一训练模型进行校验。

[0056] 终端设备对于每一帧视频图像对应的前景框的信息、预测前景框的信息对训练好的第一训练模型进行校验, 验证第一训练模型是否达到预设的校验条件。

[0057] 进一步地, S204包括:

[0058] 基于以下公式计算交并比:
$$IoU = \frac{area(m) \cap area(y)}{area(m) \cup area(y)};$$

[0059] 其中,所述m为所述每一帧所述视频图像对应的前景框的位置信息和尺寸信息,所述y为所述每一帧所述视频图像对应的预测前景框的位置信息和尺寸信息,所述IoU为所述每一帧所述视频图像对应的前景框与所述每一帧所述视频图像对应的预测前景框之间的交并比;

[0060] 基于所述交并比对所述第一训练模型进行校验。

[0061] 在训练好的第一训练模型进行校验时,终端设备具体通过公式 $IoU = \frac{area(m) \cap area(y)}{area(m) \cup area(y)}$

计算出每一帧视频图像对应的前景框与每一帧视频图像对应的预测前景框之间的交并比IoU,其中m为每一帧视频图像对应的前景框的位置信息和尺寸信息,y为每一帧视频图像对应的预测前景框的位置信息和尺寸信息;area(m)为每一帧视频图像对应的前景框的面积,area(y)为每一帧视频图像对应的预测前景框的面积,area(m) ∩ area(y)为每一帧所述视频图像对应的前景框的面积与每一帧所述视频图像对应的预测前景框的面积之间的交集,area(m) ∪ area(y)为每一帧视频图像对应的前景框的面积与每一帧视频图像对应的预测前景框的面积之间的并集。通过交并比可以得到每一帧视频图像对应的前景框与每一帧视频图像对应的预测前景框之间的交叠率情况,当交叠率较大时,说明第一训练模型已经寻找到较好的分割位置,终端设备将基于每一帧视频图像对应的前景框与每一帧视频图像对应的预测前景框之间的交并比IoU对第一训练模型进行校验。

[0062] S205,若校验通过,则将训练好的所述第一训练模型识别为分割策略模型。

[0063] 当终端设备只有在校验通过时,则将训练好的第一训练模型识别为分割策略模型。

[0064] 进一步地,所述S205,包括:

[0065] 若所述交并比均大于或等于预设阈值,则识别为校验通过,将训练好的所述第一训练模型识别为分割策略模型。

[0066] 终端设备需要判断基于每一帧视频图像对应的前景框与每一帧视频图像对应的预测前景框之间的交并比是否均大于或等于预设阈值,当基于每一帧视频图像对应的前景框与每一帧视频图像对应的预测前景框之间的交并比均大于或等于预设阈值时,则终端设备识别为校验通过,终端设备则将训练好的第一训练模型识别为分割策略模型。

[0067] 可选地,在对第一训练模型进行训练时,终端设备还可以通过公式 $IoU = \frac{area(m) \cap area(y)}{area(m) \cup area(y)}$

计算出每一帧视频图像对应的前景框与第一训练模型确定的每一帧视频图像对应的预测前景框之间的交并比IoU。

[0068] 此外,终端设备还通过公式 $IoU' = \frac{area(m') \cap area(y)}{area(m') \cup area(y)}$ 计算出每一帧视频图像对

应的前景框与不经过第一训练模型进行预测而直接进行裁剪得到的每一帧视频图像前景框之间的交并比IoU',其中,area(m')为不经过第一训练模型进行预测而直接进行裁剪得到的每一帧视频图像前景框的面积,area(m') ∩ area(y)为每一帧所述视频图像对应的前景框的面积与不经过第一训练模型进行预测而直接进行裁剪得到的前景框的面积之间的交集,area(m') ∪ area(y)为每一帧视频图像对应的前景框的面积与不经过第一训练模型进行预测而直接进行裁剪得到的前景框的面积之间的并集。

[0069] 终端设备基于公式 $\Delta = IoU - IoU'$ 计算出对应的IoU和IoU'之间的交并比差值Δ,

根据对应的交并比差值 Δ 以及公式 $r = \begin{cases} \alpha, & \Delta > 0.1 \\ 10 \cdot \alpha \cdot \Delta, & -0.1 \leq \Delta \leq 0.1 \\ -\alpha, & \Delta \leq -0.1 \end{cases}$ 计算分割策略模型

对第二前景框进行预测得到的奖励值函数 r 。

[0070] 需要说明的是,第一训练模型可以通过采用前景框搜索动作进行移动的方式来确定每一帧视频图像对应的预测前景框的位置信息和尺寸信息,前景框搜索动作包含向上平移、向下平移、向左平移和向右平移对应的四个平移动作,还包括水平扩张、水平压缩、竖直扩张和竖直压缩对应的四个缩放动作,以及包括静止不动的动作,当前景框收缩动作是停止动作时, α 取值为3,当前景框收缩动作不是停止动作时, α 取值为1。当奖励值函数 r 小于指定阈值,终端设备将确定第一训练模型已经训练完成,并将训练好的第一训练模型识别为分割策略模型。

[0071] 应理解,上述实施例中各步骤的序号的大小并不意味着执行顺序的先后,各过程的执行顺序应以其功能和内在逻辑确定,而不对本发明实施例的实施过程构成任何限定。

[0072] 参照图3,图3是本发明第三实施例提供的一种终端设备的示意图。本实施例的终端设备100包括的各单元用于执行图1对应的实施例中的各步骤,具体请参阅图1以及图1对应的实施例中的相关描述,此处不赘述。本实施例的终端设备100包括:第一获取单元101、确定单元102、裁剪单元103以及分割单元104。

[0073] 第一获取单元101,用于获取当前帧视频图像,以及获取用于对所述当前视频图像的前一帧视频图像进行分割的第一前景框的信息和第一背景框的信息;所述信息包括位置信息和尺寸信息。

[0074] 确定单元102,用于将所述当前帧视频图像、所述第一前景框的信息以及所述第一背景框的信息输入至预设的分割策略模型,确定用于对所述当前帧视频图像进行分割的第二前景框的信息以及第二背景框的信息。

[0075] 裁剪单元103,用于根据所述第二背景框的信息对所述当前帧视频图像进行裁剪,得到目标视频图像;

[0076] 分割单元104,用于基于所述第二前景框的信息对所述目标视频图像进行分割。

[0077] 作为本发明一实施例,所述的终端设备,还包括:

[0078] 第二获取单元201,用于获取用于对第一训练模型进行训练的第一样本数据,并将所述第一样本数据分成第一训练集和第一测试集;其中,所述第一样本数据中的每条样本数据均包括同一视频对应的所有帧的视频图像及所述视频图像对应的前景框的信息以及背景框的信息;

[0079] 训练单元202,用于将所述第一训练集中的每条样本数据包含的同一视频对应的所有帧的视频图像、每一帧所述视频图像的前一帧视频图像对应的前景框的信息及背景框的信息作为所述第一训练模型的输入,将每一帧所述视频图像的前景框的信息及背景框的信息作为所述第一训练模型的输出,对所述第一训练模型进行训练;

[0080] 预测单元203,用于将所述第一测试集中的每条样本数据包含的同一视频对应的所有帧的视频图像输入至训练好的所述第一训练模型,得到每一帧所述视频图像对应的预测前景框的信息以及预测背景框的信息;

[0081] 校验单元204,用于基于每一帧所述视频图像对应的前景框的信息、预测前景框的信息对训练好的所述第一训练模型进行校验;

[0082] 识别单元205,用于若校验通过,则将训练好的所述第一训练模型识别为分割策略模型。

[0083] 作为本发明一实施例,所述校验单元204,包括:

[0084] 第一计算子单元,用于基于以下公式计算每一帧所述视频图像对应的前景框与所述每一帧所述视频图像对应的预测前景框之间的交并比:
$$IoU = \frac{area(m) \cap area(y)}{area(m) \cup area(y)};$$

[0085] 其中,所述m为所述每一帧所述视频图像对应的前景框的位置信息和尺寸信息,所述y为所述每一帧所述视频图像对应的预测前景框的位置信息和尺寸信息,所述IoU为所述每一帧所述视频图像对应的前景框与所述每一帧所述视频图像对应的预测前景框之间的交并比,所述area(m)为所述每一帧所述视频图像对应的前景框的面积,所述area(y)为所述每一帧所述视频图像对应的预测前景框,所述area(m) ∩ area(y)为所述每一帧所述视频图像对应的前景框的面积与所述每一帧所述视频图像对应的预测前景框的面积之间的交集,所述area(m) ∪ area(y)为所述每一帧所述视频图像对应的前景框的面积与所述每一帧所述视频图像对应的预测前景框的面积之间的并集;

[0086] 校验子单元,用于基于所述交并比对所述第一训练模型进行校验。

[0087] 作为本发明一实施例,所述识别单元205,包括:

[0088] 若所述交并比均大于或等于预设阈值,则识别为校验通过,将训练好的所述第一训练模型识别为分割策略模型。

[0089] 作为本发明一实施例,所述确定单元102,包括:

[0090] 提取子单元,用于对所述当前帧视频图像进行特征提取,得到所述当前帧视频图像的特征图;

[0091] 第二计算子单元,用于对所述特征图进行卷积计算,得到所述当前帧视频图像的特征向量;

[0092] 分割子单元,用于根据所述特征向量和所述第一前景框的信息确定用于对所述当前帧视频图像进行分割的第二前景框的信息;

[0093] 确定子单元,用于根据所述特征向量和所述第一背景框的信息确定用于对所述当前帧视频图像进行分割的第二背景框的信息。

[0094] 上述方案,终端设备获取当前帧视频图像,以及获取用于对所述当前帧视频图像的前一帧视频图像进行分割的第一前景框的信息和第一背景框的信息;所述信息包括位置信息和尺寸信息;将所述当前帧视频图像、所述第一前景框的信息以及所述第一背景框的信息输入至预设的分割策略模型,能快速确定得到用于对当前帧视频图像进行分割的第二前景框的信息以及第二背景框的信息,即得到对当前帧视频图像进行分割的前景框的位置和尺寸以及背景框的位置和尺寸,从而可以实现提高分割的效率且可以减少在分割过程中的计算量;便于终端设备确定对当前帧视频图像进行分割的前景框和背景框在当前帧视频图像中的具体位置的信息。

[0095] 图4是本发明第四实施例提供的终端设备的示意图。如图4所示,该实施例的终端设备4包括:处理器40、存储器41以及存储在所述存储器41中并可在所述处理器40上运行的

计算机程序42,例如视频图像的分割方法的程序。所述处理器40执行所述计算机程序42时实现上述各个视频图像的分割方法的实施例中的步骤,例如图1所示的步骤S101至S104。或者,所述处理器40执行所述计算机程序42时实现上述各装置实施例中各单元的功能,例如图3所示单元101至104的功能。

[0096] 示例性的,所述计算机程序42可以被分割成一个或多个单元,所述一个或者多个单元被存储在所述存储器41中,并由所述处理器40执行,以完成本发明。所述一个或多个单元可以是能够完成特定功能的一系列计算机程序指令段,该指令段用于描述所述计算机程序42在所述终端设备4中的执行过程。例如,所述计算机程序42可以被分割成第一获取单元、确定单元、裁剪单元及以分割单元,各模块具体功能如上所述。

[0097] 所述终端设备4可包括,但不仅限于,处理器40、存储器41。本领域技术人员可以理解,图4仅仅是终端设备4的示例,并不构成用于终端设备4的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件,例如所述终端设备还可以包括输入输出设备、网络接入设备、总线等。

[0098] 所称处理器40可以是中央处理单元(Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0099] 所述存储器41可以是终端设备4的内部存储单元,例如用于终端设备4的硬盘或内存。所述存储器41也可以是所述终端设备4的外部存储设备,例如所述终端设备4上配备的插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)等。进一步地,所述存储器41还可以既包括所述终端设备4的内部存储单元也包括外部存储设备。所述存储器41用于存储所述计算机程序以及所述终端设备所需的其他程序和数据。所述存储器41还可以用于暂时地存储已经输出或者将要输出的数据。

[0100] 所属领域的技术人员可以清楚地了解到,为了描述的方便和简洁,仅以上述各功能单元、模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能单元、模块完成,即将所述装置的内部结构划分成不同的功能单元或模块,以完成以上描述的全部或者部分功能。实施例中的各功能单元、模块可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中,上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。另外,各功能单元、模块的具体名称也只是为了便于相互区分,并不用于限制本申请的保护范围。上述系统中单元、模块的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0101] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述或记载的部分,可以参见其它实施例的相关描述。

[0102] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出

本发明的范围。

[0103] 在本发明所提供的实施例中,应该理解到,所揭露的终端设备和方法,可以通过其它的方式实现。例如,以上所描述的终端设备的实施例仅仅是示意性的,例如,所述模块或单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通讯连接可以是通过一些接口,装置或单元的间接耦合或通讯连接,可以是电性,机械或其它的形式。

[0104] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0105] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0106] 所述集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明实现上述实施例方法中的全部或部分流程,也可以通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于计算机可读存储介质中,该计算机程序在被处理器执行时,可实现上述各个方法实施例的步骤。。其中,所述计算机程序包括计算机程序代码,所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括:能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、电载波信号、电信信号以及软件分发介质等。需要说明的是,所述计算机可读介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减,例如在某些司法管辖区,根据立法和专利实践,计算机可读介质不包括电载波信号和电信信号。

[0107] 以上所述实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围,均应包含在本发明的保护范围之内。

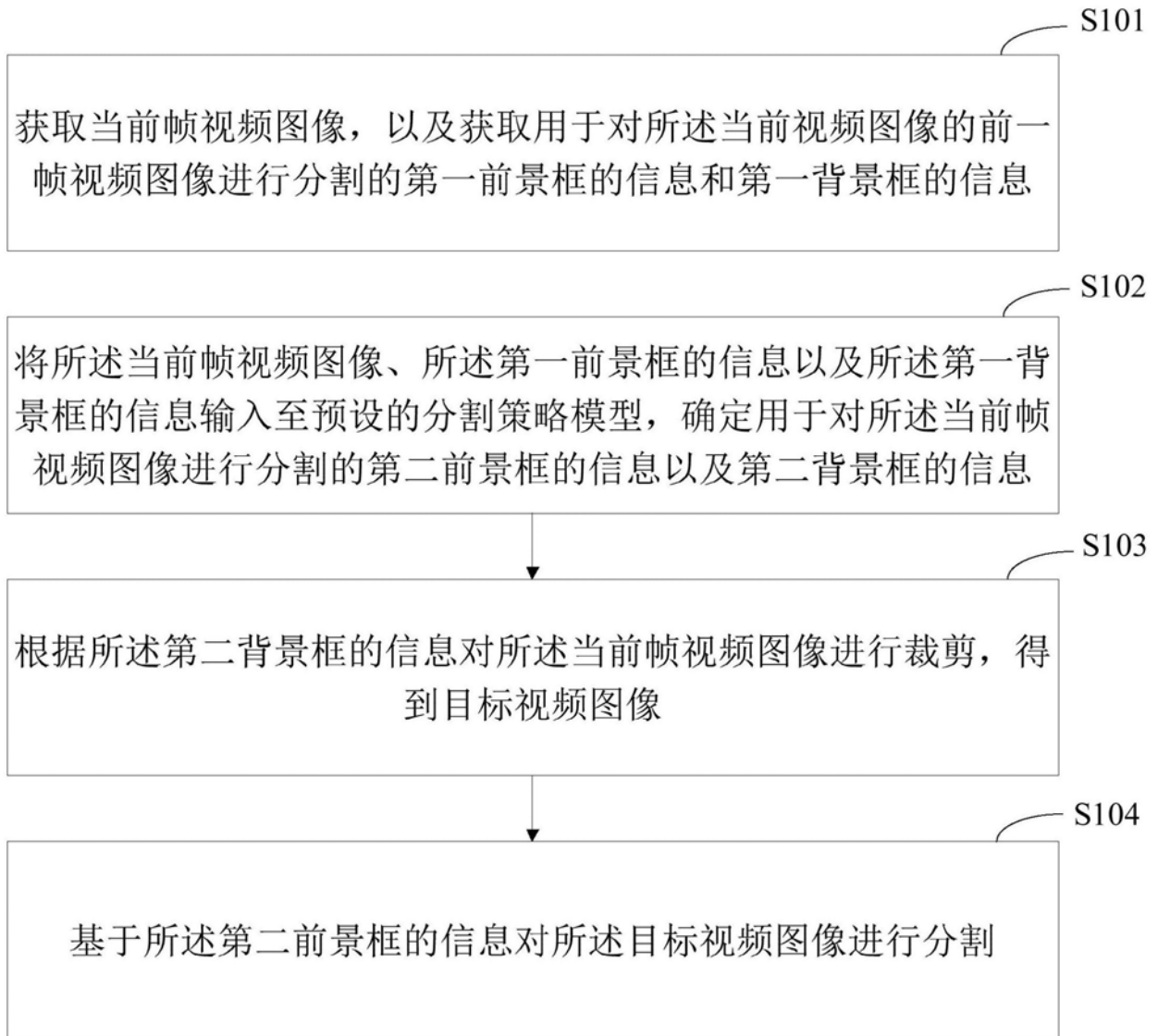


图1

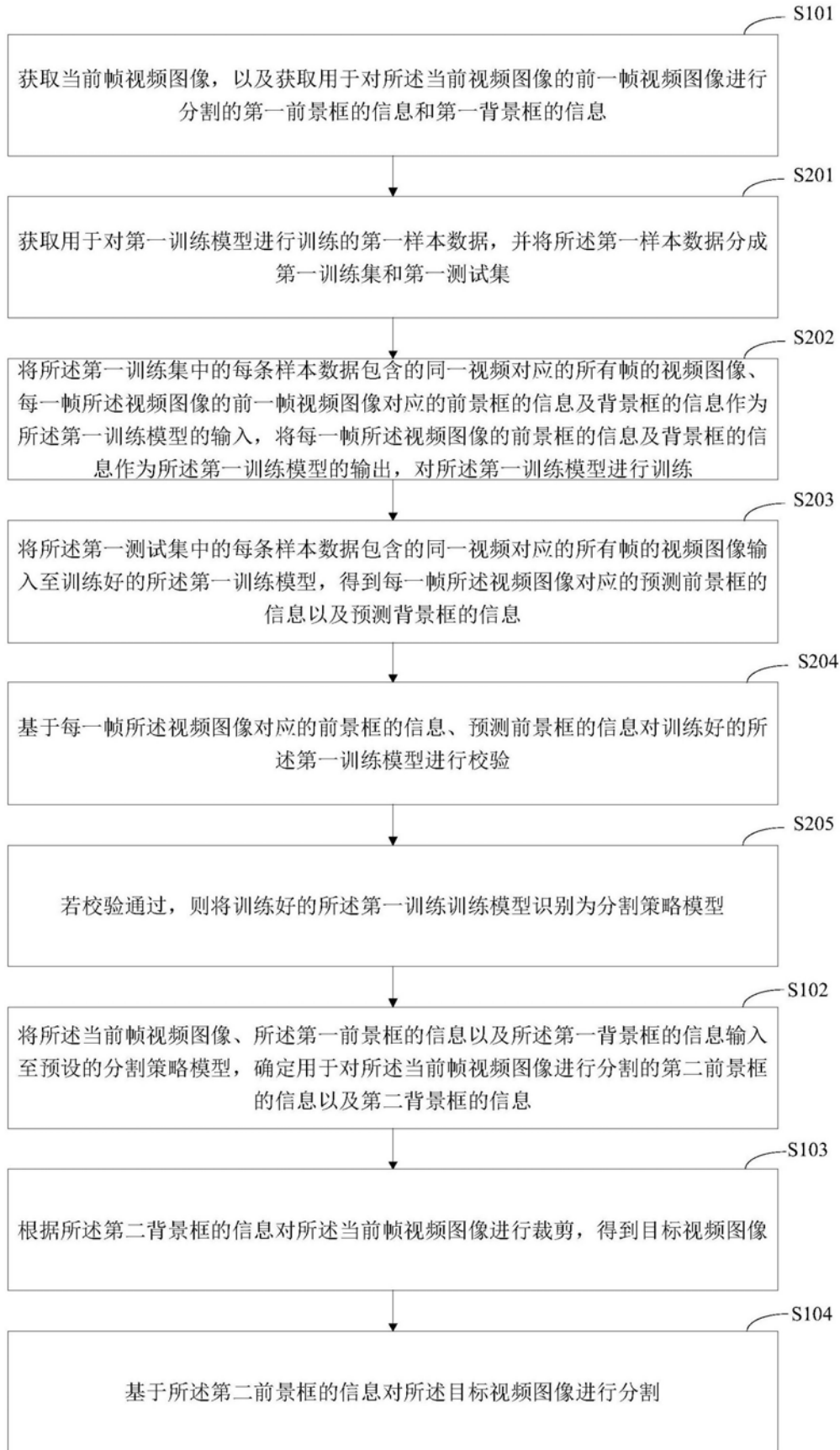


图2

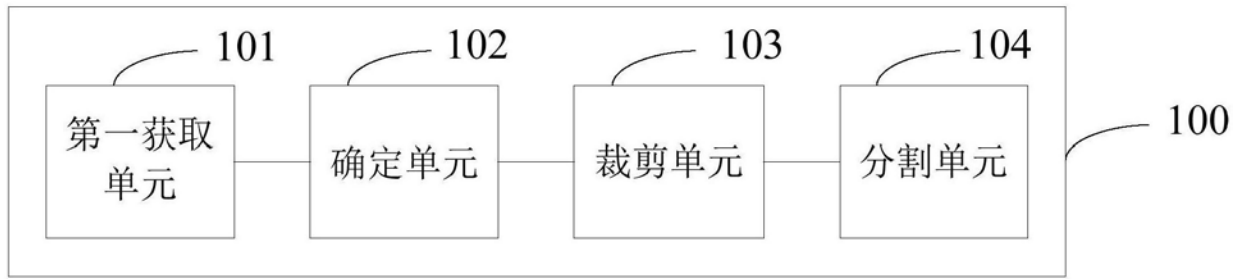


图3

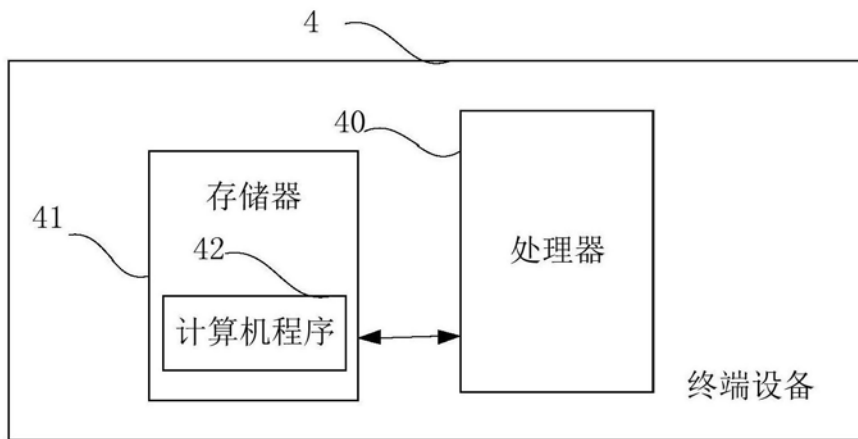


图4