



(12)发明专利申请

(10)申请公布号 CN 110675433 A

(43)申请公布日 2020.01.10

(21)申请号 201911049133.7

(22)申请日 2019.10.31

(71)申请人 北京达佳互联信息技术有限公司
地址 100085 北京市海淀区上地西路6号1
幢1层101D1-7

(72)发明人 张雷 李强

(74)专利代理机构 北京三高永信知识产权代理
有限责任公司 11138
代理人 宁立存

(51) Int. Cl.
G06T 7/246(2017.01)
G06K 9/00(2006.01)

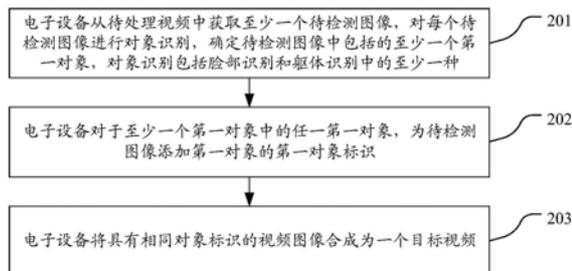
权利要求书2页 说明书15页 附图4页

(54)发明名称

视频处理方法、装置、电子设备及存储介质

(57)摘要

本公开关于一种视频处理方法、装置、电子设备和存储介质,涉及多媒体技术领域。方法包括:从待处理视频中获取至少一个待检测图像,对每个待检测图像进行对象识别,确定所述待检测图像中包括的至少一个第一对象,所述对象识别包括脸部识别和躯体识别中的至少一种;对于所述至少一个第一对象中的任一第一对象,为所述待检测图像添加所述第一对象的第一对象标识;将具有相同对象标识的视频图像合成为一个目标视频。根据本公开的技术方案,可以自动的将待处理视频中的视频图像按照对象的不同重新合成为单独的视频,视频处理的准确度和效率高。



1. 一种视频处理方法,其特征在于,所述方法包括:

从待处理视频中获取至少一个待检测图像,对每个待检测图像进行对象识别,确定所述待检测图像中包括的至少一个第一对象,所述对象识别包括脸部识别和躯体识别中的至少一种;

对于所述至少一个第一对象中的任一第一对象,为所述待检测图像添加所述第一对象的第一对象标识;

将具有相同对象标识的视频图像合成为一个目标视频。

2. 根据权利要求1所述的方法,其特征在于,所述对待处理视频中的每个待检测图像进行对象识别,确定所述待检测图像中包括的至少一个第一对象,包括:

对于所述待处理视频中的每个待检测图像,对所述待检测图像进行脸部识别,得到至少一个第二对象,对于所述至少一个第二对象中的每个第二对象,当所述第二对象的脸部与脸部数据库中的脸部的相似度大于第一相似阈值时,将所述第二对象作为所述第一对象;或者,

对于所述待处理视频中的每个待检测图像,对所述待检测图像进行躯体识别,得到至少一个第三对象,对于所述至少一个第三对象中的每个第三对象,当所述第三对象的躯体与躯体数据库中的躯体的相似度大于第二相似阈值时,将所述第三对象作为所述第一对象;或者,

对于所述待处理视频中的每个待检测图像,对所述待检测图像进行脸部识别和躯体识别,得到至少一个第四对象,对于所述至少一个第四对象中的每个第四对象,当所述第四对象的脸部与脸部数据库中的脸部的相似度大于第一相似阈值,且躯体与躯体数据库中的躯体的相似度大于第二相似阈值时,将所述第四对象作为所述第一对象。

3. 根据权利要求1所述的方法,其特征在于,所述将具有相同对象标识的视频图像合成为一个目标视频之前,所述方法还包括:

对于所述待处理视频中的每个视频图像,对所述视频图像进行对象跟踪,得到至少一个对象跟踪框,所述对象跟踪包括脸部跟踪和躯体跟踪中的至少一种;

当任一对象跟踪框对应的对象的对象标识与所述对象库中任一对象的对象标识相同时,将所述对象跟踪框作为所述目标跟踪框;

为所述视频图像添加至少一个目标跟踪框对应的至少一个第五对象的第二对象标识。

4. 根据权利要求3所述的方法,其特征在于,所述对于所述待处理视频中的每个视频图像,对所述视频图像进行对象跟踪之后,所述方法还包括:

当确定任一对象跟踪失败时,确定所述对象的第三对象标识,开始缓存视频图像;

当再次检测到所述第三对象标识时,停止缓存视频图像,确定缓存的视频图像的缓存数量;

当所述缓存数量小于缓存数量阈值时,为所述缓存的视频图像添加所述第三对象标识。

5. 根据权利要求3所述的方法,其特征在于,所述为所述视频图像添加所述至少一个目标跟踪框对应的至少一个第五对象的第二对象标识之前,所述方法还包括:

对所述至少一个目标跟踪框进行表情识别,得到至少一个表情标签;

根据表情标签与第一合成元素的对应关系,获取至少一个第一合成元素,所述第一合

成元素包括音频元素、视频元素、图片元素以及特效元素中的至少一种；

为所述视频图像添加所述第一合成元素的元素标识。

6. 根据权利要求3所述的方法,其特征在於,所述为所述视频图像添加所述至少一个目标跟踪框对应的至少一个第五对象的第二对象标识之前,所述方法还包括:

对所述至少一个目标跟踪框进行动作识别,得到至少一个动作标签;

根据动作标签与第二合成元素的对应关系,获取至少一个第二合成元素,所述第二合成元素包括音频元素、视频元素、图片元素以及特效元素中的至少一种;

为所述视频图像添加所述第二合成元素的元素标识。

7. 根据权利要求5或6任一项权利要求所述的方法,其特征在於,所述将具有相同对象标识的视频图像合成为一个目标视频,包括:

获取具有相同对象标识的至少一个视频图像,根据每个视频图像对应的元素标识,为所述视频图像添加所述第一合成元素或所述第二合成元素中的至少一种;

将至少一个经过添加处理后的视频图像按照时间顺序进行排序,合成为一个目标视频。

8. 一种视频处理装置,其特征在於,包括:

识别单元,被配置为从待处理视频中获取至少一个待检测图像,对每个待检测图像进行对象识别,确定所述待检测图像中包括的至少一个第一对象,所述对象识别包括脸部识别和躯体识别中的至少一种;

添加单元,被配置为对于所述至少一个第一对象中的任一第一对象,为所述待检测图像添加所述第一对象的第一对象标识;

合成单元,被配置为将具有相同对象标识的视频图像合成为一个目标视频。

9. 一种电子设备,其特征在於,包括:

处理器;

用于存储所述处理器可执行指令的存储器;

其中,所述处理器被配置为执行所述指令,以实现如权利要求1至7中任一项所述的视频处理方法。

10. 一种存储介质,当所述存储介质中的指令由电子设备的处理器执行时,使得所述电子设备能够执行如权利要求1至7中任一项所述的视频处理方法。

视频处理方法、装置、电子设备及存储介质

技术领域

[0001] 本公开涉及多媒体技术领域,尤其涉及一种视频处理方法、装置、电子设备及存储介质。

背景技术

[0002] 随着多媒体技术的发展,用户可以随时随地的拍摄和分享视频。用户为了使自己分享的视频能够吸引更多的人观看,往往会对视频进行一定的处理。而视频剪辑是一种常见的视频处理方法。

[0003] 目前,用户通常是使用视频剪辑程序来处理视频,手动从视频中提取视频图像,然后将提取到的视频图像进行拼接,来生成新的视频。用户还可以通过视频剪辑程序在生成的视频中加入图片、背景音乐、特效以及场景等元素,以提高生成的视频的表现力。

[0004] 上述技术方案存在的问题是,对视频的剪辑需要由用户依靠视频剪辑程序来完成,对用户的专业技能要求较高,且视频处理的耗时长、效率低。

发明内容

[0005] 本公开提供一种视频处理方法、装置、电子设备及存储介质,用于解决对视频的剪辑需要由用户依靠视频剪辑程序来完成,对用户的专业技能要求较高,且视频处理的耗时长、效率低的问题。本公开的技术方案如下:

[0006] 根据本公开实施例的一方面,提供一种视频处理方法,包括:

[0007] 从待处理视频中获取至少一个待检测图像,对每个待检测图像进行对象识别,确定所述待检测图像中包括的至少一个第一对象,所述对象识别包括脸部识别和躯体识别中的至少一种;

[0008] 对于所述至少一个第一对象中的任一第一对象,为所述待检测图像添加所述第一对象的第一对象标识;

[0009] 将具有相同对象标识的视频图像合成为一个目标视频。

[0010] 在一种可能的实现方式中,所述对待处理视频中的每个待检测图像进行对象识别,确定所述待检测图像中包括的至少一个第一对象,包括:

[0011] 对于所述待处理视频中的每个待检测图像,对所述待检测图像进行脸部识别,得到至少一个第二对象,对于所述至少一个第二对象中的每个第二对象,当所述第二对象的脸部与脸部数据库中的脸部的相似度大于第一相似阈值时,将所述第二对象作为所述第一对象;或者,

[0012] 对于所述待处理视频中的每个待检测图像,对所述待检测图像进行躯体识别,得到至少一个第三对象,对于所述至少一个第三对象中的每个第三对象,当所述第三对象的躯体与躯体数据库中的躯体的相似度大于第二相似阈值时,将所述第三对象作为所述第一对象;或者,

[0013] 对于所述待处理视频中的每个待检测图像,对所述待检测图像进行脸部识别和躯

体识别,得到至少一个第四对象,对于所述至少一个第四对象中的每个第四对象,当所述第四对象的脸部与脸部数据库中的脸部的相似度大于第一相似阈值,且躯体与躯体数据库中的躯体的相似度大于第二相似阈值时,将所述第四对象作为所述第一对象。

[0014] 在一种可能的实现方式中,所述对待处理视频中的每个待检测图像进行对象识别,确定所述待检测图像中包括的至少一个第一对象之前,所述方法还包括:

[0015] 对待检测图像进行脸部检测,当检测到的脸部为正脸且占所述待检测图像的比例大于第一比例阈值时,执行对所述待检测图像进行脸部识别的步骤;或者,

[0016] 对待检测图像进行躯体检测,当检测到的躯体为直立且占所述待检测图像的比例大于第二比例阈值时,执行对所述待检测图像进行躯体识别的步骤;或者,

[0017] 对待检测图像进行脸部检测和躯体检测,当检测到的脸部为正脸、所述脸部占所述待检测图像的比例大于第一比例阈值、检测到的躯体为直立且所述躯体占所述待检测图像的比例大于第二比例阈值时,执行对所述待检测图像进行脸部识别和躯体识别的步骤。

[0018] 在一种可能的实现方式中,所述对所述待检测图像进行脸部识别和躯体识别,包括:

[0019] 对于具有重合部分的脸部检测框和躯体检测框,确定所述脸部检测框和所述躯体检测框的交并比,当所述交并比大于目标比例阈值时,确定所述脸部检测框和所述躯体检测框属于同一第四对象,为所述脸部检测框和所述躯体检测框设置相同的对象标识,当所述交并比不大于所述目标比例阈值时,确定所述脸部检测框和所述躯体检测框不属于同一第四对象,为所述脸部检测框和所述躯体检测框分别设置对象标识,所述交并比为具有重合部分的两个检测框之间交集和并集的比值;

[0020] 对于未重合的脸部检测框,为所述脸部检测框设置对象标识;

[0021] 对于未重合的躯体检测框,为所述躯体检测框设置对象标识。

[0022] 在一种可能的实现方式中,所述对所述待检测图像进行对象识别之后,所述方法还包括:

[0023] 当通过脸部识别的方式识别出新对象时,将所述新对象的脸部信息存储至所述脸部数据库;或者,

[0024] 当通过躯体识别的方式识别出新对象时,将所述新对象的躯体信息存储至所述躯体数据库;或者,

[0025] 当同时通过脸部识别和躯体识别的方式识别出新对象时,将所述新对象的脸部信息存储至所述脸部数据库,将所述新对象的躯体信息存储至所述躯体数据库。

[0026] 在一种可能的实现方式中,所述将具有相同对象标识的视频图像合成为一个目标视频之前,所述方法还包括:

[0027] 对于所述待处理视频中的每个视频图像,对所述视频图像进行对象跟踪,得到至少一个对象跟踪框,所述对象跟踪包括脸部跟踪和躯体跟踪中的至少一种;

[0028] 当任一对象跟踪框对应的对象的对象标识与所述对象库中任一对象的对象标识相同时,将所述对象跟踪框作为所述目标跟踪框;

[0029] 为所述视频图像添加至少一个目标跟踪框对应的至少一个第五对象的第二对象标识。

[0030] 在一种可能的实现方式中,所述对于所述待处理视频中的每个视频图像,对所述

视频图像进行对象跟踪之后,所述方法还包括:

- [0031] 当确定任一对象跟踪失败时,确定所述对象的第三对象标识,开始缓存视频图像;
- [0032] 当再次检测到所述第三对象标识时,停止缓存视频图像,确定缓存的视频图像的缓存数量;
- [0033] 当所述缓存数量小于缓存数量阈值时,为所述缓存的视频图像添加所述第三对象标识。
- [0034] 在一种可能的实现方式中,所述方法还包括:
- [0035] 当所述缓存数量不小于所述缓存数量阈值时,清除所述缓存的视频图像。
- [0036] 在一种可能的实现方式中,所述为所述视频图像添加所述至少一个目标跟踪框对应的至少一个第五对象的第二对象标识之前,所述方法还包括:
- [0037] 对所述至少一个目标跟踪框进行表情识别,得到至少一个表情标签;
- [0038] 根据表情标签与第一合成元素的对应关系,获取至少一个第一合成元素,所述第一合成元素包括音频元素、视频元素、图片元素以及特效元素中的至少一种;
- [0039] 为所述视频图像添加所述第一合成元素的元素标识。
- [0040] 在一种可能的实现方式中,所述为所述视频图像添加所述至少一个目标跟踪框对应的至少一个第五对象的第二对象标识之前,所述方法还包括:
- [0041] 对所述至少一个目标跟踪框进行动作识别,得到至少一个动作标签;
- [0042] 根据动作标签与第二合成元素的对应关系,获取至少一个第二合成元素,所述第二合成元素包括音频元素、视频元素、图片元素以及特效元素中的至少一种;
- [0043] 为所述视频图像添加所述第二合成元素的元素标识。
- [0044] 在一种可能的实现方式中,所述将具有相同对象标识的视频图像合成为一个目标视频,包括:
- [0045] 获取具有相同对象标识的至少一个视频图像,根据每个视频图像对应的元素标识,为所述视频图像添加所述第一合成元素或所述第二合成元素中的至少一种;
- [0046] 将至少一个经过添加处理后的视频图像按照时间顺序进行排序,合成为一个目标视频。
- [0047] 根据本公开实施例的另一方面,提供一种视频处理装置,包括:
- [0048] 识别单元,被配置为从待处理视频中获取至少一个待检测图像,对每个待检测图像进行对象识别,确定所述待检测图像中包括的至少一个第一对象,所述对象识别包括脸部识别和躯体识别中的至少一种;
- [0049] 添加单元,被配置为对于所述至少一个第一对象中的任一第一对象,为所述待检测图像添加所述第一对象的第一对象标识;
- [0050] 合成单元,被配置为将具有相同对象标识的视频图像合成为一个目标视频。
- [0051] 在一种可能的实现方式中,所述识别单元,还被配置为实现:
- [0052] 对于所述待处理视频中的每个待检测图像,对所述待检测图像进行脸部识别,得到至少一个第二对象,对于所述至少一个第二对象中的每个第二对象,当所述第二对象的脸部与脸部数据库中的脸部的相似度大于第一相似阈值时,将所述第二对象作为所述第一对象;或者,对于所述待处理视频中的每个待检测图像,对所述待检测图像进行躯体识别,得到至少一个第三对象,对于所述至少一个第三对象中的每个第三对象,当所述第三对象

的躯体与躯体数据库中的躯体的相似度大于第二相似阈值时,将所述第三对象作为所述第一对象;或者,对于所述待处理视频中的每个待检测图像,对所述待检测图像进行脸部识别和躯体识别,得到至少一个第四对象,对于所述至少一个第四对象中的每个第四对象,当所述第四对象的脸部与脸部数据库中的脸部的相似度大于第一相似阈值,且躯体与躯体数据库中的躯体的相似度大于第二相似阈值时,将所述第四对象作为所述第一对象。

[0053] 在一种可能的实现方式中,所述装置还包括:

[0054] 检测单元,被配置为对待检测图像进行脸部检测,当检测到的脸部为正脸且占所述待检测图像的比例大于第一比例阈值时,对所述待检测图像进行脸部识别;或者,

[0055] 所述检测单元,还被配置为对待检测图像进行躯体检测,当检测到的躯体为直立且占所述待检测图像的比例大于第二比例阈值时,对所述待检测图像进行躯体识别;或者,

[0056] 所述检测单元,还被配置为对待检测图像进行脸部检测和躯体检测,当检测到的脸部为正脸、所述脸部占所述待检测图像的比例大于第一比例阈值,并且检测到的躯体为直立且所述躯体占所述待检测图像的比例大于第二比例阈值时,对所述待检测图像进行脸部识别和躯体识别。

[0057] 在一种可能的实现方式中,所述识别单元,还被配置为对于具有重合部分的脸部检测框和躯体检测框,确定所述脸部检测框和所述躯体检测框的交并比,当所述交并比大于目标比例阈值时,确定所述脸部检测框和所述躯体检测框属于同一第四对象,为所述脸部检测框和所述躯体检测框设置相同的对象标识,当所述交并比不大于所述目标比例阈值时,确定所述脸部检测框和所述躯体检测框不属于同一第四对象,为所述脸部检测框设置对象标识,所述交并比为具有重合部分的两个检测框之间交集和并集的比值;对于未重合的脸部检测框,为所述脸部检测框设置对象标识;对于未重合的躯体检测框,为所述躯体检测框设置对象标识。

[0058] 在一种可能的实现方式中,所述装置还包括:

[0059] 存储单元,被配置为当通过脸部识别的方式识别出新对象时,将所述新对象的脸部信息存储至所述脸部数据库;或者,

[0060] 所述存储单元,还被配置为当通过躯体识别的方式识别出新对象时,将所述新对象的躯体信息存储至所述躯体数据库;或者,

[0061] 所述存储单元,还被配置为当同时通过脸部识别和躯体识别的方式识别出新对象时,将所述新对象的脸部信息存储至所述脸部数据库,将所述新对象的躯体信息存储至所述躯体数据库。

[0062] 在一种可能的实现方式中,所述装置还包括:

[0063] 跟踪单元,被配置为对于所述待处理视频中的每个视频图像,对所述视频图像进行对象跟踪,得到至少一个对象跟踪框,所述对象跟踪包括脸部跟踪和躯体跟踪中的至少一种;

[0064] 所述跟踪单元,还被配置为当任一对象跟踪框对应的对象的对象标识与所述对象库中任一对象的对象标识相同时,将所述对象跟踪框作为所述目标跟踪框;

[0065] 所述添加单元,还被配置为为所述视频图像添加至少一个目标跟踪框对应的至少一个第五对象的第二对象标识。

[0066] 在一种可能的实现方式中,所述装置还包括:

[0067] 缓存单元,被配置为当确定任一对象跟踪失败时,确定所述对象的第三对象标识,开始缓存视频图像;

[0068] 所述缓存单元,还被配置为当再次检测到所述第三对象标识时,停止缓存视频图像,确定缓存的视频图像的缓存数量;

[0069] 所述添加单元,还被配置为当所述缓存数量小于缓存数量阈值时,为所述缓存的视频图像添加所述第三对象标识。

[0070] 在一种可能的实现方式中,所述装置还包括:

[0071] 清除单元,被配置为当所述缓存数量不小于所述缓存数量阈值时,清除所述缓存的视频图像。

[0072] 在一种可能的实现方式中,所述装置还包括:

[0073] 所述识别单元,还被配置为对所述至少一个目标跟踪框进行表情识别,得到至少一个表情标签;

[0074] 获取单元,被配置为根据表情标签与第一合成元素的对应关系,获取至少一个第一合成元素,所述第一合成元素包括音频元素、视频元素、图片元素以及特效元素中的至少一种;

[0075] 所述添加单元,还被配置为为所述视频图像添加所述第一合成元素的元素标识。

[0076] 在一种可能的实现方式中,所述装置还包括:

[0077] 所述识别单元,还被配置为对所述至少一个目标跟踪框进行动作识别,得到至少一个动作标签;

[0078] 获取单元,被配置为根据动作标签与第二合成元素的对应关系,获取至少一个第二合成元素,所述第二合成元素包括音频元素、视频元素、图片元素以及特效元素中的至少一种;

[0079] 所述添加单元,还被配置为为所述视频图像添加所述第二合成元素的元素标识。

[0080] 在一种可能的实现方式中,所述合成单元,还被配置为获取具有相同对象标识的至少一个视频图像,根据每个视频图像对应的元素标识,为所述视频图像添加所述第一合成元素或所述第二合成元素中的至少一种;将至少一个经过添加处理后的视频图像按照时间顺序进行排序,合成为一个目标视频。

[0081] 根据本公开实施例的另一方面,提供一种电子设备,包括处理器;用于存储所述处理器可执行指令的存储器;其中,所述处理器被配置为执行所述指令,以实现上述视频处理方法。

[0082] 根据本公开实施例的另一方面,提供一种存储介质,当所述存储介质中的指令由电子设备的处理器执行时,使得所述电子设备能够执行上述视频处理方法。

[0083] 根据本公开实施例的另一方面,提供一种计算机程序产品,当所述计算机程序产品中的指令由电子设备的处理器执行时,使得所述电子设备能够执行上述视频处理方法。

[0084] 本公开的实施例提供的技术方案至少带来以下有益效果:

[0085] 通过对待处理视频中的每个待检测图像进行对象识别,确定该待检测图像中包括的至少一个第一对象,然后对于上述至少一个第一对象中的任一第一对象,为该待检测图像添加该第一对象的第一对象标识,最后将具有相同对象标识的视频图像合成为一个目标视频。根据本公开的技术方案,可以自动的将待处理视频中的视频图像按照对象的不同重

新合成为单独的视频,视频处理的准确度和效率高。

[0086] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,并不能限制本公开。

附图说明

[0087] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本公开的实施例,并与说明书一起用于解释本公开的原理,并不构成对本公开的不当限定。

[0088] 图1是本公开实施例示出的一种视频服务系统的结构框图;

[0089] 图2是本公开实施例示出的一种视频处理方法的流程图;

[0090] 图3是本公开实施例示出的另一种视频处理方法的流程图;

[0091] 图4是本公开实施例示出的另一种视频处理方法的流程图;

[0092] 图5是本公开实施例示出的一种人体识别跟踪流程示意图;

[0093] 图6是本公开实施例示出的一种视频处理装置的框图;

[0094] 图7是本公开实施例提供的一种电子设备的结构框图。

具体实施方式

[0095] 为了使本领域普通人员更好地理解本公开的技术方案,下面将结合附图,对本公开实施例中的技术方案进行清楚、完整地描述。

[0096] 需要说明的是,本公开的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本公开的实施例能够以除了在这里图示或描述的那些以外的顺序实施。以下示例性实施例中所描述的实施方式并不代表与本公开相一致的所有实施方式。相反,它们仅是与如所附权利要求书中所详述的本公开的一些方面相一致的装置和方法的例子。

[0097] 本公开提供的实施例的应用场景为视频处理的场景。主要涉及的是对视频进行剪辑处理的场景,例如用户期望从一个多人视频中,将包含某一个人的视频图像分离出来,将分离得到的视频图像组成新的视频。

[0098] 本公开实施例提供的视频处理方法,可以应用与上述场景中。用户在使用本公开实施例提供的视频处理方法进行视频剪辑时,可以提供包括目标对象的图像,也可以不提供包括目标对象的图像。当用户提供包括目标对象的图像时,本公开实施例提供的视频处理方法可以从待处理视频中提取包括该目标对象的视频图像,然后合成得到目标视频;当用户未提供包括目标对象的图像时,本公开实施例提供的视频处理方法可以根据待处理视频中包括的多个对象生成多个目标视频,每个目标视频对应一个对象。

[0099] 下面简单介绍一下本公开实施例中可能用到的技术:

[0100] 脸部识别技术,以人脸识别技术为例,人脸识别技术是指基于人的脸部特征信息进行身份识别的一种生物识别技术。

[0101] 躯体识别技术,以人体识别技术为例,人体识别技术是指基于人的身体特征信息进行身份识别的一种生物识别技术。

[0102] 目标跟踪技术,是在给定某个视频图像中目标的大小和位置的情况下,预测后续

视频图像中该目标的大小和位置。

[0103] 图1是本公开实施例示出的一种视频服务系统100的结构框图。该视频服务系统100可以用于视频处理,包括:电子设备110和视频服务平台120。

[0104] 电子设备110可以通过无线网络或有线网络与视频服务平台120相连。电子设备110可以是智能手机、摄像机、台式计算机、平板电脑、MP4播放器和膝上型便携计算机中的至少一种。电子设备110安装和运行有支持视频处理的应用程序。示意性的,电子设备110可以是用户使用的电子设备,电子设备运行的应用程序内登录有该用户的账号。

[0105] 视频服务平台120包括一台服务器、多台服务器和云计算平台中的至少一种。视频服务平台120用于提供客户端的后台服务,如视频处理、视频发布、账号关联等。可选的,可以由视频服务器120承担主要的视频处理工作,电子设备110承担次要的视频处理工作;或者,可以由视频服务器120承担次要的视频处理工作,电子设备110承担主要的视频处理工作;或者,可以由视频服务器120和电子设备110分别单独承担视频处理工作。

[0106] 可选的,视频服务平台120包括:接入服务器、视频处理服务器、视频发布服务器、用户管理服务器和数据库。接入服务器用于提供电子设备110的接入服务。视频处理服务器用于对上传的视频进行处理。视频处理服务器可以是一台或多台,当视频处理服务器是多台时,存在至少两台视频处理服务器用于提供不同的服务,和/或,存在至少两台视频处理服务器用于提供相同的服务,比如以负载均衡方式提供同一种服务或者以主服务器和镜像服务器的方式提供同一种服务,本公开实施例对此不加以限定。数据库用于存储上传的视频。

[0107] 电子设备110可以泛指多个电子设备中的一个,本实施例仅以电子设备110来举例说明。本领域技术人员可以知晓,上述电子设备的数量可以更多或更少。比如上述电子设备可以仅为一个,或者上述电子设备为几十个或几百个,或者更多数量,此时上述视频服务系统还包括其他电子设备。本公开实施例对电子设备的数量和类型不加以限定。

[0108] 图2是本公开实施例示出的一种视频处理方法的流程图,如图2所示,包括以下步骤:

[0109] 201、电子设备从待处理视频中获取至少一个待检测图像,对每个待检测图像进行对象识别,确定待检测图像中包括的至少一个第一对象,对象识别包括脸部识别和躯体识别中的至少一种。

[0110] 202、电子设备对于至少一个第一对象中的任一第一对象,为待检测图像添加第一对象的第一对象标识。

[0111] 203、电子设备将具有相同对象标识的视频图像合成为一个目标视频。

[0112] 在本公开实施例中,通过对待处理视频中的每个待检测图像进行对象识别,确定该待检测图像中包括的至少一个第一对象,然后对于上述至少一个第一对象中的任一第一对象,为该待检测图像添加该第一对象的第一对象标识,最后将具有相同对象标识的视频图像合成为一个目标视频。根据本公开的技术方案,可以自动的将待处理视频中的视频图像按照对象的不同重新合成为单独的视频,视频处理的准确度和效率高。

[0113] 图3是本公开实施例示出的另一种视频处理方法的流程图,如图3所示,包括以下步骤:

[0114] 301、电子设备对待处理视频进行解码,得到多个视频图像,每个视频图像对应该

待处理视频的一个视频帧。

[0115] 在本公开实施例中,待处理视频可以为用户刚拍摄完成的视频,也可以为用户在历史时间拍摄完成的视频,还可以为用户获取的其他视频,本公开实施例对此不进行具体限制。电子设备可以对该待处理视频进行解码操作,从该待处理视频中提取多个视频帧,每个视频帧即为一个视频图像,从而得到多个视频图像。

[0116] 需要说明的是,由于视频在拍摄时可能会受到环境或者设备的影响,使得摄像头发生抖动,导致拍摄得到的视频画面晃动,从而不利于对象的检测和跟踪。因此,在电子设备从待处理视频中获取至少一个待检测图像之前,可以对该待处理视频进行视频稳像处理。本公开实施例对视频稳像的方式不进行具体限定。

[0117] 302、电子设备从待处理视频中获取至少一个待检测图像,对每个待检测图像进行对象识别,确定该待检测图像中包括的至少一个第一对象,该对象识别包括脸部识别和躯体识别中的至少一种。

[0118] 在本公开实施例中,在对待识别视频中包括的至少一个对象进行对象识别时,电子设备可以对上述得到的多个视频图像中的每个视频图像进行对象识别,此时上述多个视频图像为终端从待处理视频中获取的至少一个待检测图像;电子设备还可以从上述多个视频图像中每间隔预设帧数获取一个视频图像,从而得到至少一个待检测图像。其中,预设帧数可以为5帧、10帧等,本公开实施例对此不进行具体限制。相较于对每个视频图像都进行对象识别,每间隔预设帧数进行一次对象识别可以在对象识别算法较为复杂时,降低电子设备的负载,也即降低电子设备的占用率。

[0119] 在一种可选的实现方式中,以电子设备每隔预设帧数,获取一个待检测图像为例进行说明。电子设备可以对每个待检测图像进行对象识别,该对象识别包括脸部识别和躯体识别中的至少一种。

[0120] 相应的,当该对象识别为脸部识别时,电子设备对每个待检测图像进行对象识别,确定该待检测图像中包括的至少一个第一对象的步骤可以为:对于待处理视频中的每个待检测图像,电子设备可以对该待检测图像进行脸部识别,得到至少一个第二对象,对于该至少一个第二对象中的每个第二对象,当该第二对象的脸部与脸部数据库中的脸部的相似度大于第一相似阈值时,电子设备可以将该第二对象作为上述第一对象。其中,电子设备在进行脸部识别时,可以先从待检测视频中检测对象的脸部,得到至少一个脸部图像,对该至少一个脸部图像进行脸部关键点检测和校正,通过脸部验证网络从经过校正的脸部图像以及脸部数据库中存储的脸部图像中分别提取脸部特征。电子设备可以计算提取到的脸部特征之间的相似度。上述脸部验证网络可以基于卷积神经网络得到。

[0121] 相应的,当该对象识别为躯体识别时,电子设备对每个待检测图像进行对象识别,确定该待检测图像中包括的至少一个第一对象的步骤可以为:对于待处理视频中的每个待检测图像,电子设备可以对该待检测图像进行躯体识别,得到至少一个第三对象,对于该至少一个第三对象中的每个第三对象,当该第三对象的躯体与躯体数据库中的躯体的相似度大于第二相似阈值时,电子设备可以将该第三对象作为上述第一对象。其中,电子设备在进行躯体识别时,可以先从待检测视频中检测对象的躯体,得到至少一个躯体图像,对该至少一个躯体图像进行人体姿态和局部匹配,然后分别提取躯体特征。电子设备可以计算提取到的躯体特征之间的相似度。

[0122] 相应的,当该对象识别为脸部识别和躯体识别时,电子设备对每个待检测图像进行对象识别,确定该待检测图像中包括的至少一个第一对象的步骤可以为:对于待处理视频中的每个待检测图像,电子设备可以对待检测图像进行脸部识别和躯体识别,得到至少一个第四对象,对于该至少一个第四对象中的每个第四对象,当该第四对象的脸部与脸部数据库中的脸部的相似度大于第一相似阈值,且躯体与躯体数据库中的躯体的相似度大于第二相似阈值时,电子设备可以将该第四对象作为上述第一对象。

[0123] 在一种可选的实现方式中,电子设备在对待处理视频中的每个待检测图像进行对象识别,确定待检测图像中包括的至少一个第一对象之前,可以根据检测到的脸部或者躯体中至少一个的检测结果来确定是否要继续执行上述对象识别的步骤。具体的,电子设备对待检测图像进行脸部检测,当检测到的脸部为正脸且占待检测图像的比例大于第一比例阈值时,电子设备可以执行对待检测图像进行脸部识别的步骤;或者,电子设备对待检测图像进行躯体检测,当检测到的躯体为直立且占待检测图像的比例大于第二比例阈值时,电子设备可以执行对待检测图像进行躯体识别的步骤;或者,电子设备对待检测图像进行脸部检测和躯体检测,当检测到的脸部为正脸、脸部占待检测图像的比例大于第一比例阈值,并且检测到的躯体为直立且躯体占待检测图像的比例大于第二比例阈值时,电子设备可以执行对待检测图像进行脸部识别和躯体识别的步骤。由于在进行对象识别之前,根据脸部检测或者躯体检测中至少一种的检测结果,来确定是否执行后续对象识别的步骤,实现了对低质量的待检测图像的过滤,提高了对象识别的效率。

[0124] 在一种可选的实现方式中,电子设备同时对待检测图像进行脸部识别和躯体识别时,同一个对象的脸部检测框和躯体检测框可能会出现重合部分,电子设备可以根据重合部分的大小,确定具有重合部分的脸部检测框和躯体检测框是否属于同一个对象。当属于同一个对象时,电子设备可以为该脸部检测框和躯体检测框设置相同的对象标识;当不属于同一个对象时,电子设备可以分别为该脸部检测框和该躯体检测框设置对象标识,以区分不同的对象。

[0125] 相应的,电子设备对待检测图像进行脸部识别和躯体识别的步骤可以为:对于具有重合部分的脸部检测框和躯体检测框,电子设备可以确定脸部检测框和躯体检测框的交并比,当交并比大于目标比例阈值时,确定脸部检测框和躯体检测框属于同一第四对象,电子设备可以为脸部检测框和躯体检测框设置相同的对象标识;当交并比不大于目标比例阈值时,确定脸部检测框和躯体检测框不属于同一第四对象,电子设备为脸部检测框设置对象标识。这是由于,交并比小于目标比例阈值时,可以认为躯体检测框的置信度较低,而脸部检测框的置信度较高,也即躯体检测的结果不可信,因此终端之给脸部检测框设置对象标识。而对于未重合的脸部检测框,电子设备可以为该脸部检测框设置对象标识;对于未重合的躯体检测框,电子设备可以为该躯体检测框设置对象标识。其中,交并比用于反映两个检测框的重叠程度(IoU, Intersection over Union),可以由具有重合部分的两个检测框之间交集和并集的比值来表示。

[0126] 需要说明的是,当对象为人时,由于人的头部在躯体的上方,因此,电子设备在确定人脸检测框和人体检测框是否属于同一个人时,在交并比大于目标比例阈值的同时,还需要满足人脸检测框处于人体检测框的靠上的位置这一条件,电子设备才可以确定该人脸检测框和人体检测框属于同一个人。

[0127] 在一种可选的实现方式中,用户可以提供包括目标对象的图像,电子设备可以对该目标对象的脸部进行特征提取,将该目标对象的脸部信息存储至脸部数据库,电子设备还可以对该目标对象的躯体进行特征提取,将该目标对象的躯体信息存储至躯体数据库。此时,电子设备在对每个待检测图像进行对象识别时,仅需要确定该待检测图像中是否包括该目标对象,也即该目标对象即为上述第一对象。其中,该目标对象可以为一个或者多个。

[0128] 在一种可选的实现方式中,如果用户未提供包括目标对象的图像,电子设备可以对每个待检测图像进行对象识别,当电子设备通过脸部识别的方式识别出新对象时,可以将该新对象的脸部信息存储至脸部数据库,或者,当电子设备通过躯体识别的方式识别出新对象时,可以将新对象的躯体信息存储至躯体数据库,或者,当电子设备同时通过脸部识别和躯体识别的方式识别出新对象时,可以将新对象的脸部信息存储至脸部数据库,将新对象的躯体信息存储至躯体数据库。当电子设备识别出已在脸部数据库或者躯体数据库中存储的对象时,电子设备可以记录该对象的对象标识。该对象标识可以为编号、ID (Identity document,唯一编码) 或者名称等。

[0129] 需要说明的是,电子设备可以通过任意的脸部识别算法实现对象的脸部识别,也可以通过任意的躯体识别算法实现对象的躯体识别。本公开实施例对此不进行具体限制。例如,通过人脸识别算法识别人脸,通过人体识别算法识别人体。

[0130] 303、对于上述至少一个第一对象中的任一第一对象,电子设备为上述待检测图像添加该第一对象的第一对象标识。

[0131] 在本公开实施例中,对于任一待检测图像,电子设备可以为该待检测图像添加在该待检测图像中确定的至少一个第一对象中的每个第一对象的对象标识。

[0132] 在一种可选的实现方式中,如果用户提供了包括目标对象的图像,也即用户给定了目标对象,则电子设备确定的第一对象为该目标对象。若待检测图像中包括该目标对象,则电子设备可以为该待检测图像添加该目标对象的对象标识;若待检测图像中不包括该目标对象,则电子设备没有获取到第一对象,不需要为该待检测图像添加对象标识。

[0133] 在一种可选的实现方式中,如果用户未给定目标对象,当该待检测图像中仅有一个对象时,电子设备可以将该对象作为第一对象,为该待检测图像添加该第一对象的对象标识;当该待检测图像中有多个对象时,电子设备可以将该多个对象作为第一对象,为该待检测图像添加上述多个第一对象的对象标识。

[0134] 304、对于待处理视频中的每个视频图像,电子设备对该视频图像进行对象跟踪,得到至少一个对象跟踪框,该对象跟踪包括脸部跟踪和躯体跟踪中的至少一种;当任一对象跟踪框对应的对象的对象标识与对象库中任一对象的对象标识相同时,电子设备将该对象跟踪框作为目标跟踪框。

[0135] 在本公开实施例中,电子设备可以对待处理视频中包括的至少一个对象进行对象跟踪。电子设备在进行对象跟踪时可以与对象识别相结合,当电子设备检测到新对象时,电子设备可以基于该新对象的检测框对该新对象进行跟踪。在对新对象开始进行跟踪时,电子设备可以通过该新对象的检测框和该视频图像进行跟踪初始化,预测出该对象在下一帧的位置,也即该对象在下一个视频图像中可能出现的位置,电子设备可以通过跟踪框来标识该对象在下一个视频图像中的位置。

[0136] 在一种可选的实现方式中,如果用户给定了目标对象,则对象库中存储的为目标对象的对象标识,电子设备可以仅跟踪该目标对象。当在视频图像中跟踪到对象时,电子设备得到至少一个对象跟踪框的位置信息,电子设备可以将每个对象跟踪框对应的对象的对象标识分别与对象库中存储的目标对象的对象标识进行比对,当任一对象的对象标识与目标对象的对象标识相同时,表示该对象跟踪框跟踪的对象为目标对象,电子设备可以将该对象跟踪框作为目标跟踪框。该目标对象可以为一个或多个。

[0137] 在一种可选的实现方式中,如果用户未给定目标对象时,电子设备可以对待处理视频中检测到的所有对象进行跟踪。对于每个视频图像,电子设备可以获取至少一个对象跟踪框,电子设备可以直接将该至少一个对象跟踪框作为目标跟踪框。需要说明的是,当电子设备在该视频图像中检测到新对象时,电子设备可以为该新对象设置对象标识,以及对该新对象进行跟踪。

[0138] 在一种可选的实现方式中,当检测到新对象的脸部时,电子设备可以根据脸部检测框预测该新对象在下一个视频图像中脸部可能的位置,在下一个视频图像中用脸部跟踪框表示该位置。

[0139] 在一种可选的实现方式中,当检测到新对象的躯体时,电子设备可以根据躯体检测框预测该新对象在下一个视频图像中脸部可能的位置,在下一个视频图像中用躯体跟踪框表示该位置。

[0140] 需要说明的是,当新对象的脸部检测框和躯体检测框的交并比大于目标比例阈值时,该新对象的脸部跟踪框和躯体跟踪框的对应的对象标识为相同的对象标识。

[0141] 在一种可选的实现方式中,电子设备在对任一视频图像中的对象进行对象跟踪,得到至少一个目标跟踪框时,电子设备可以对至少一个目标跟踪框进行表情识别,得到至少一个表情标签。电子设备可以根据表情标签与第一合成元素的对应关系,获取至少一个第一合成元素,该第一合成元素可以包括音频元素、视频元素、图片元素以及特效元素中的至少一种。电子设备可以为该视频图像添加该第一合成元素的元素标识。

[0142] 在一种可选的实现方式中,电子设备在对任一视频图像中的对象进行对象跟踪,得到至少一个目标跟踪框时,电子设备可以对至少一个目标跟踪框进行动作识别,得到至少一个动作标签。电子设备可以根据动作标签与第二合成元素的对应关系,获取至少一个第二合成元素,该第二合成元素可以包括音频元素、视频元素、图片元素以及特效元素中的至少一种。电子设备可以为该视频图像添加该第二合成元素的元素标识。

[0143] 还需要说明的是,由于目标跟踪算法有置信度输出,当任一对象的置信度小于目标置信度时,电子设备对该对象的跟踪失败。当电子设备确定任一对象跟踪失败时,电子设备可以确定该对象的第三对象标识,电子设备可以从跟踪失败的视频图像开始,缓存之后出现的每个视频图像。当电子设备再次检测到该第三对象标识时,表示该对象再次出现,电子设备可以停止缓存视频图像,确定缓存的视频图像的缓存数量。当该缓存数量小于缓存数量阈值时,表示该对象仍存在于缓存的视频图像中,电子设备可以为缓存的视频图像添加第三对象标识;当该缓存数量不小于缓存数量阈值时,表示该对象可能在缓存的视频图像中未出现,电子设备可以清除上述缓存的视频图像。

[0144] 还需要说明的是,由于跟踪算法可能存在一定的误差,电子设备可以根据每隔预设帧数进行一次的对象识别时生成的对象检测框来校正跟踪框。也即当电子设备进行对象

识别时,可以通过检测框来标识当前视频图像中的对象,当任一对象的检测框和跟踪框的重合程度小于目标重合阈值时,表示跟踪框出现了较大的跟踪误差,电子设备可以根据该检测框和该视频图像来初始化该对象的跟踪框。

[0145] 305、对于待处理视频中的每个视频图像,电子设备为该视频图像添加至少一个目标跟踪框对应的至少一个第五对象的第二对象标识。

[0146] 在本发明实施例中,对于待处理视频中的每个视频图像,电子设备可以根据上述步骤确定的至少一个目标跟踪框来为该视频图像添加第二对象标识。在用户已给定了目标对象时,电子设备可以确定上述至少一个目标跟踪框对应的至少一个第五对象,该至少一个第五对象即为用户给定的至少一个目标对象。在用户未给定目标对象时,电子设备可以确定上述至少一个目标跟踪框对应的至少一个第五对象,该至少一个第五对象为该视频图像中包括的所有对象。

[0147] 306、电子设备将具有相同对象标识的视频图像合成为一个目标视频。

[0148] 在本公开实施例中,电子设备可以根据对象标识对上述待处理视频中的多个视频图像进行区分,每个对象标识对应至少一个视频图像,每个视频图像也对应至少一个对象标识。也即是一个对象可以出现在一个或者多个视频图像中,一个视频图像中可以包括一个或者多个对象。电子设备可以将具有相同对象标识的视频图像按照时间顺序合成为一个目标视频。

[0149] 在一种可选的实现方式中,电子设备还可以在合成目标视频时添加合成元素。相应的,电子设备将具有相同对象标识的视频图像合成为一个目标视频的步骤可以为:电子设备获取具有相同对象标识的至少一个视频图像,根据每个视频图像对应的元素标识,为该视频图像添加第一合成元素或第二合成元素中的至少一种。电子设备可以将至少一个经过添加处理后的视频图像按照时间顺序进行排序,合成为一个目标视频。其中,电子设备可以通过对上述至少一个视频图像进行编码,以实现视频图像的合成,得到完整的视频。

[0150] 需要说明的是,上述步骤301至步骤306是本公开实施例示例性示出的可能的实现方式,在上述步骤中,步骤302和步骤304可以同时执行。在一种可选的实现方式中,电子设备可以先执行步骤302和步骤303,再执行步骤304和步骤305,本公开实施例对此不进行具体限制。

[0151] 为了使上述步骤301至步骤306描述的过程更清晰,可以参见图4所示,图4是本公开实施例示出的另一种视频处理方法的流程图。在图4中以对象为人为例,示出了电子设备对视频处理的主要流程,包括:获取待处理视频(可选:包括目标对象的图像),视频解码,视频稳像,人脸检测及识别,人体检测及识别,人脸跟踪,人体跟踪,视频图像添加对象标识,人脸表情识别,人体动作识别,添加音频/视频特效等合成元素,视频编码。

[0152] 另外,为了使上述步骤302至步骤305中电子设备对对象进行识别和跟踪的过程更清晰,可以参见图5所示,图5是本公开实施例示出的一种人体识别跟踪流程示意图。在图5中以对象为人为例,示出了电子设备对人脸的识别和跟踪流程包括:人脸检测,人脸图像,将人脸关键点和校正后的图像输入人脸验证网络,将人脸特征和人脸部数据库中的特征进行匹配,计算相似度,根据相似度确定该人脸对应的对象的对象标识,可选的可以从备选框中确定得分最高的框作为人脸检测框,根据当前视频图像检测框和各对象的对象标识为视频图像添加对象标识,根据人脸检测框对至少一个对象进行人脸跟踪。电子设备对人体的

识别和跟踪流程包括：人体检测，人体图像，通过姿态和局部匹配提取人体特征，将人体特征和人体库中的特征进行匹配，计算相似度，根据相似度确定该人体对应的对象的对象标识，可选的可以从备选框中确定得分最高的框作为人体检测框，根据当前视频图像检测框和各对象的对象标识为视频图像添加对象标识，根据人体检测框对至少一个对象进行人体跟踪。

[0153] 在本公开实施例中，通过对待处理视频中的每个待检测图像进行对象识别，确定该待检测图像中包括的至少一个第一对象，然后对于上述至少一个第一对象中的任一第一对象，为该待检测图像添加该第一对象的第一对象标识，最后将具有相同对象标识的视频图像合成为一个目标视频。根据本公开的技术方案，可以自动的将待处理视频中的视频图像按照对象的不同重新合成为单独的视频，视频处理的准确度和效率高。

[0154] 图6是本公开实施例示出的一种视频处理装置的框图，如图6所示，包括：识别单元601、添加单元602和合成单元603。

[0155] 识别单元601，被配置为从待处理视频中获取至少一个待检测图像，对每个待检测图像进行对象识别，确定待检测图像中包括的至少一个第一对象，对象识别包括脸部识别和躯体识别中的至少一种；

[0156] 添加单元602，被配置为对于至少一个第一对象中的任一第一对象，为待检测图像添加第一对象的第一对象标识；

[0157] 合成单元603，被配置为将具有相同对象标识的视频图像合成为一个目标视频。

[0158] 在一种可能的实现方式中，识别单元601，还被配置为实现：

[0159] 对于待处理视频中的每个待检测图像，对待检测图像进行脸部识别，得到至少一个第二对象，对于至少一个第二对象中的每个第二对象，当第二对象的脸部与脸部数据库中的脸部的相似度大于第一相似阈值时，将第二对象作为第一对象；或者，对于待处理视频中的每个待检测图像，对待检测图像进行躯体识别，得到至少一个第三对象，对于至少一个第三对象中的每个第三对象，当第三对象的躯体与躯体数据库中的躯体的相似度大于第二相似阈值时，将第三对象作为第一对象；或者，对于待处理视频中的每个待检测图像，对待检测图像进行脸部识别和躯体识别，得到至少一个第四对象，对于至少一个第四对象中的每个第四对象，当第四对象的脸部与脸部数据库中的脸部的相似度大于第一相似阈值，且躯体与躯体数据库中的躯体的相似度大于第二相似阈值时，将第四对象作为第一对象。

[0160] 在一种可能的实现方式中，装置还包括：

[0161] 检测单元，被配置为对待检测图像进行脸部检测，当检测到的脸部为正脸且占待检测图像的比例大于第一比例阈值时，对待检测图像进行脸部识别；或者，

[0162] 检测单元，还被配置为对待检测图像进行躯体检测，当检测到的躯体为直立且占待检测图像的比例大于第二比例阈值时，对待检测图像进行躯体识别；或者，

[0163] 检测单元，还被配置为对待检测图像进行脸部检测和躯体检测，当检测到的脸部为正脸、脸部占待检测图像的比例大于第一比例阈值，并且检测到的躯体为直立且躯体占待检测图像的比例大于第二比例阈值时，对待检测图像进行脸部识别和躯体识别。

[0164] 在一种可能的实现方式中，识别单元601，还被配置为对于具有重合部分的脸部检测框和躯体检测框，确定脸部检测框和躯体检测框的交并比，当交并比大于目标比例阈值时，确定脸部检测框和躯体检测框属于同一第四对象，为脸部检测框和躯体检测框设置相

同的对象标识,当交并比不大于目标比例阈值时,确定脸部检测框和躯体检测框不属于同一第四对象,为脸部检测框设置对象标识,交并比为具有重合部分的两个检测框之间交集和并集的比值;对于未重合的脸部检测框,为脸部检测框设置对象标识;对于未重合的躯体检测框,为躯体检测框设置对象标识。

[0165] 在一种可能的实现方式中,装置还包括:

[0166] 存储单元,被配置为当通过脸部识别的方式识别出新对象时,将新对象的脸部信息存储至脸部数据库;或者,

[0167] 存储单元,还被配置为当通过躯体识别的方式识别出新对象时,将新对象的躯体信息存储至躯体数据库;或者,

[0168] 存储单元,还被配置为当同时通过脸部识别和躯体识别的方式识别出新对象时,将新对象的脸部信息存储至脸部数据库,将新对象的躯体信息存储至躯体数据库。

[0169] 在一种可能的实现方式中,装置还包括:

[0170] 跟踪单元,被配置为对于待处理视频中的每个视频图像,对视频图像进行对象跟踪,得到至少一个对象跟踪框,对象跟踪包括脸部跟踪和躯体跟踪中的至少一种;

[0171] 跟踪单元,还被配置为当任一对象跟踪框对应的对象的对象标识与对象库中任一对象的对象标识相同时,将对象跟踪框作为目标跟踪框;

[0172] 添加单元602,还被配置为为视频图像添加至少一个目标跟踪框对应的至少一个第五对象的第二对象标识。

[0173] 在一种可能的实现方式中,装置还包括:

[0174] 缓存单元,被配置为当确定任一对象跟踪失败时,确定对象的第三对象标识,开始缓存视频图像;

[0175] 缓存单元,还被配置为当再次检测到第三对象标识时,停止缓存视频图像,确定缓存的视频图像的缓存数量;

[0176] 添加单元602,还被配置为当缓存数量小于缓存数量阈值时,为缓存的视频图像添加第三对象标识。

[0177] 在一种可能的实现方式中,装置还包括:

[0178] 清除单元,被配置为当缓存数量不小于缓存数量阈值时,清除缓存的视频图像。

[0179] 在一种可能的实现方式中,装置还包括:

[0180] 识别单元601,还被配置为对至少一个目标跟踪框进行表情识别,得到至少一个表情标签;

[0181] 获取单元,被配置为根据表情标签与第一合成元素的对应关系,获取至少一个第一合成元素,第一合成元素包括音频元素、视频元素、图片元素以及特效元素中的至少一种;

[0182] 添加单元602,还被配置为为视频图像添加第一合成元素的元素标识。

[0183] 在一种可能的实现方式中,方法还包括:

[0184] 识别单元601,还被配置为对至少一个目标跟踪框进行动作识别,得到至少一个动作标签;

[0185] 获取单元,还被配置为根据动作标签与第二合成元素的对应关系,获取至少一个第二合成元素,第二合成元素包括音频元素、视频元素、图片元素以及特效元素中的至少一

种;

[0186] 添加单元602,还被配置为为视频图像添加第二合成元素的元素标识。

[0187] 在一种可能的实现方式中,合成单元603,还被配置为获取具有相同对象标识的至少一个视频图像,根据每个视频图像对应的元素标识,为视频图像添加第一合成元素或第二合成元素中的至少一种;将至少一个经过添加处理后的视频图像按照时间顺序进行排序,合成为一个目标视频。

[0188] 通过对待处理视频中的每个待检测图像进行对象识别,确定该待检测图像中包括的至少一个第一对象,然后对于上述至少一个第一对象中的任一第一对象,为该待检测图像添加该第一对象的第一对象标识,最后将具有相同对象标识的视频图像合成为一个目标视频。根据本公开的技术方案,可以自动的将待处理视频中的视频图像按照对象的不同重新合成为单独的视频,视频处理的准确度和效率高。

[0189] 图7是本公开实施例提供的一种电子设备的结构框图,该电子设备700可因配置或性能不同而产生比较大的差异,可以包括一个或一个以上处理器(central processing units,CPU)701和一个或一个以上的存储器702,其中,该存储器702中存储有至少一条指令,该至少一条指令由处理器701加载并执行以实现上述各个方法实施例提供的方法。当然,该电子设备还可以具有有线或无线网络接口、键盘以及输入输出接口等部件,以便进行输入输出,该电子设备还可以包括其他用于实现设备功能的部件,在此不做赘述。

[0190] 本公开实施例还提供了一种存储介质,用于服务器,当该存储介质中存储的指令由处理器执行时,使得服务器能够执行上述实施例的视频处理方法。

[0191] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本公开的其它实施方案。本公开旨在涵盖本公开的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本公开的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本公开的真正范围和精神由权利要求指出。

[0192] 应当理解的是,本公开并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本公开的范围仅由所附的权利要求来限制。

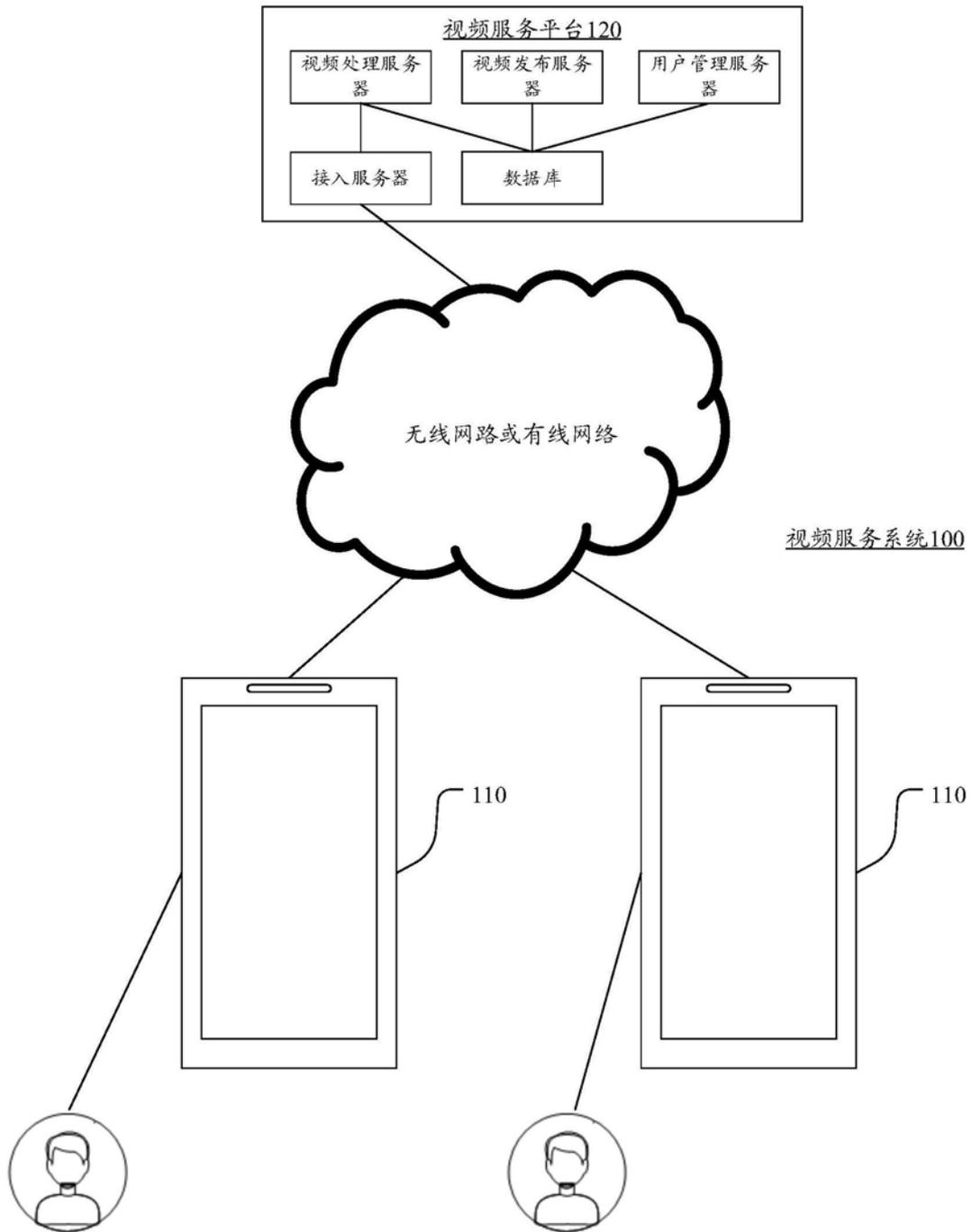


图1

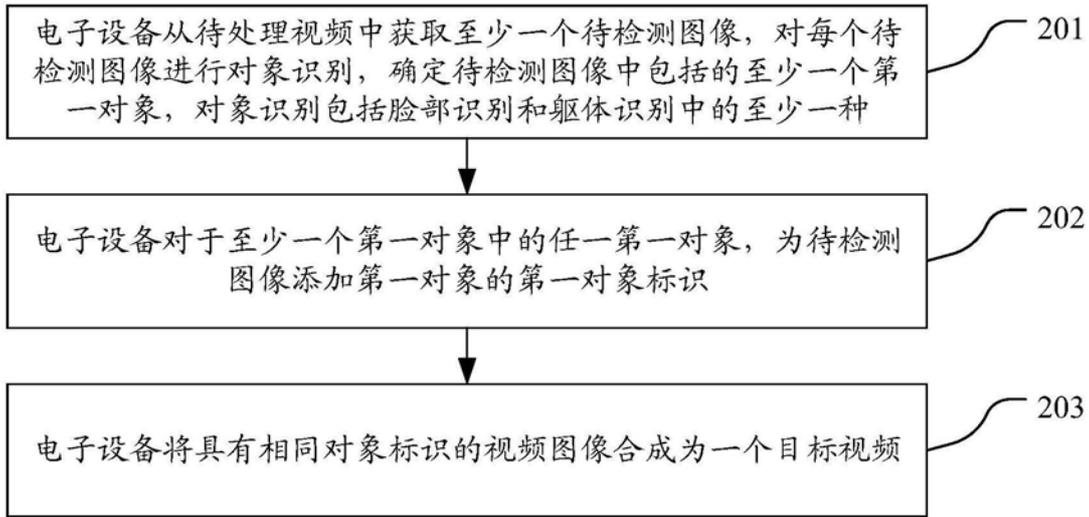


图2

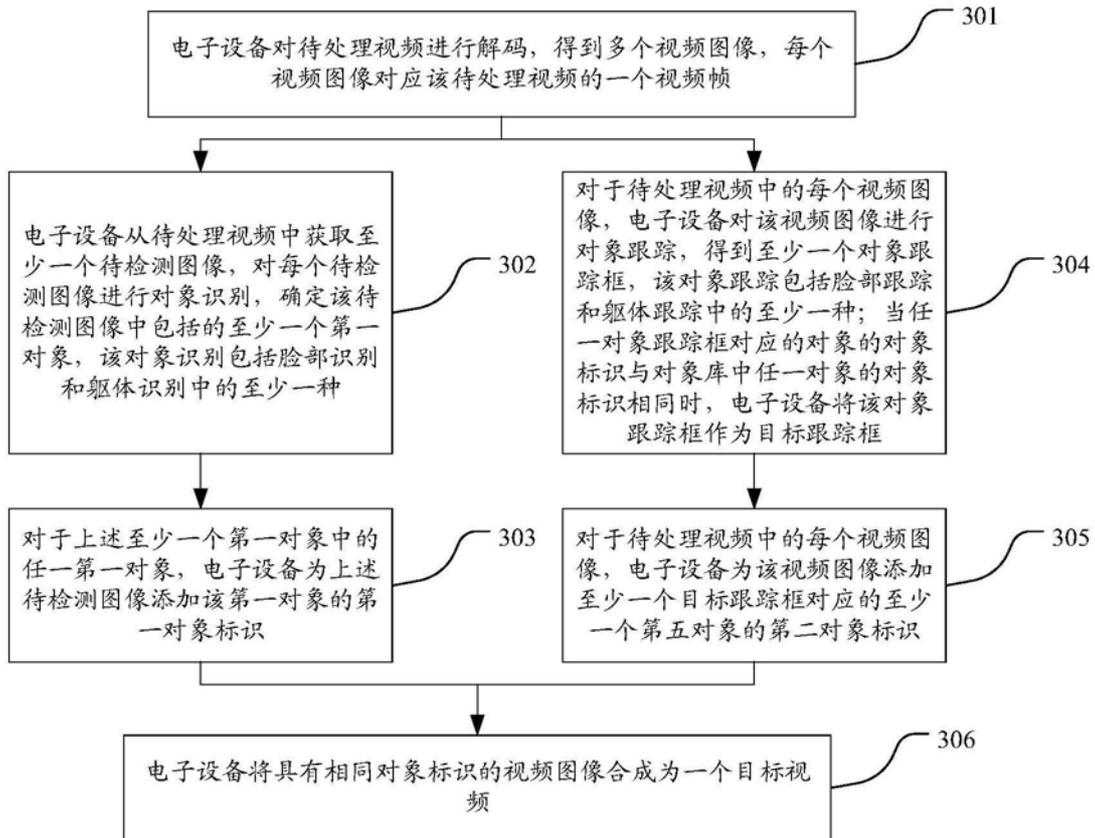


图3

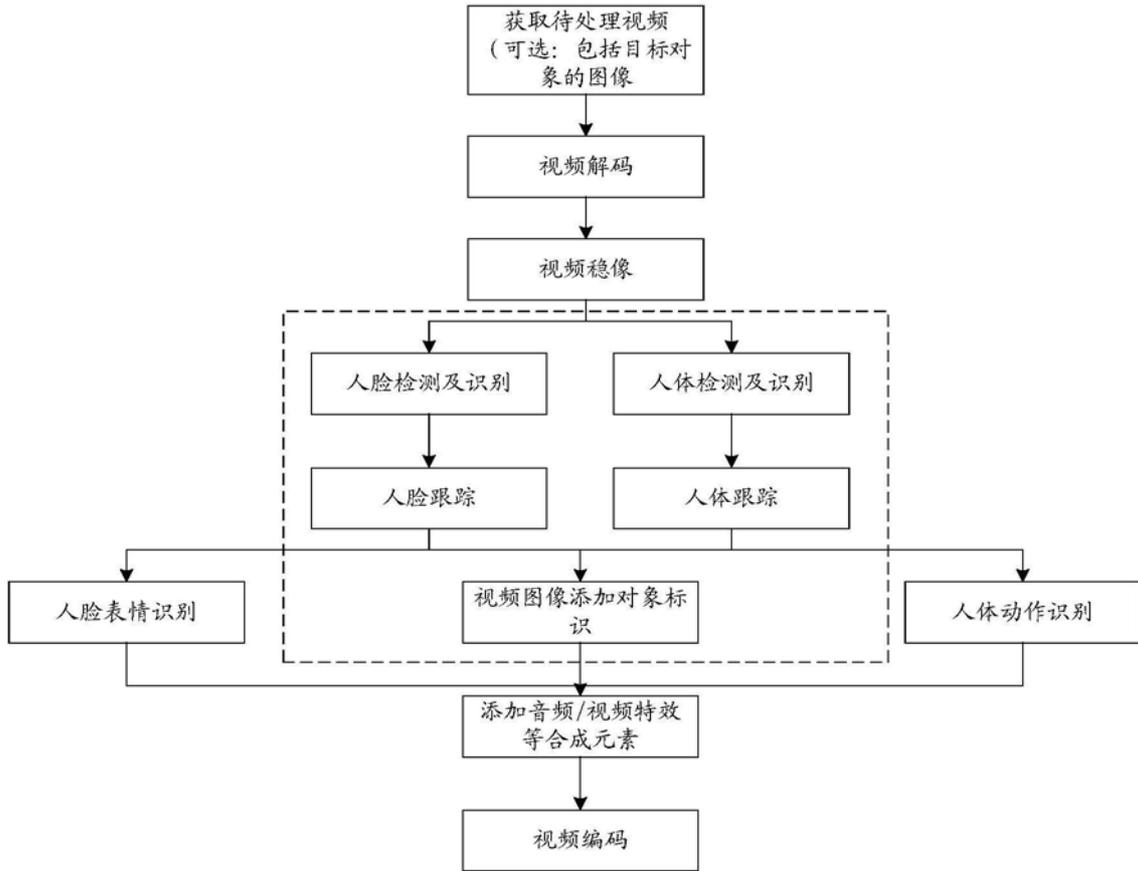


图4

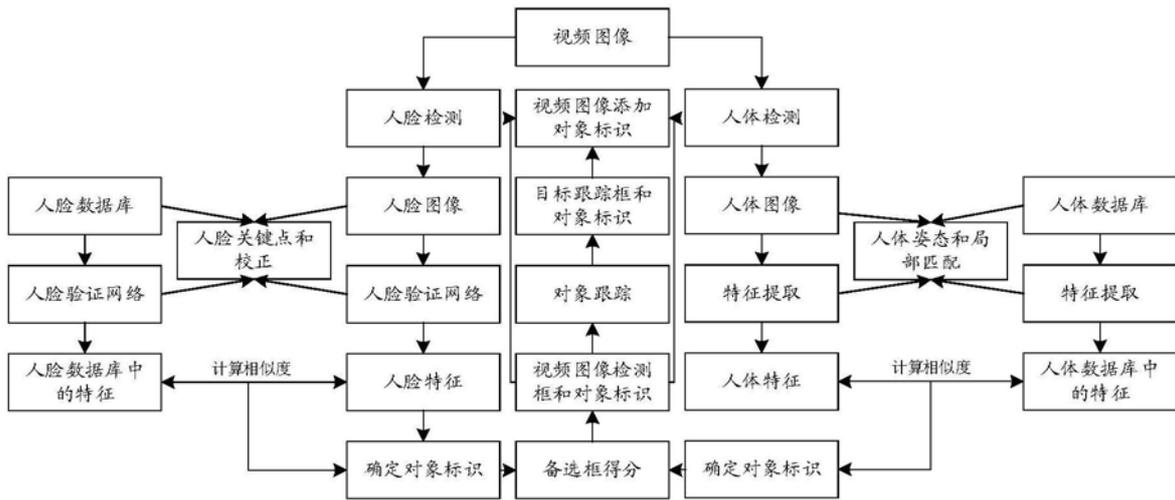


图5

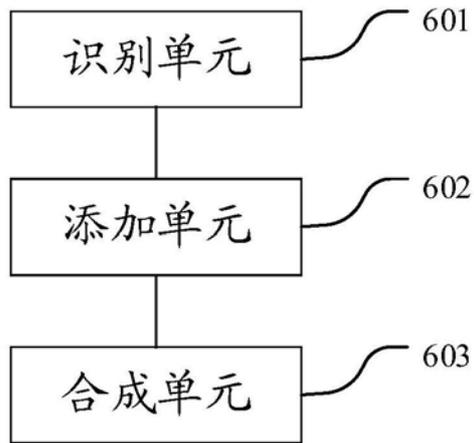


图6

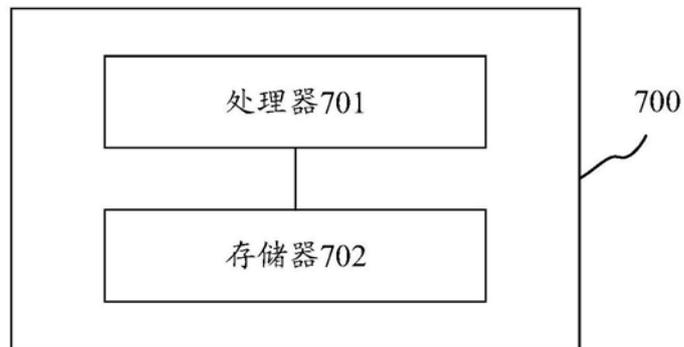


图7