



(12) 发明专利

(10) 授权公告号 CN 112954450 B

(45) 授权公告日 2022.06.17

(21) 申请号 202110145223.7

(56) 对比文件

(22) 申请日 2021.02.02

CN 108509941 A, 2018.09.07

(65) 同一申请的已公布的文献号

审查员 宫贺

申请公布号 CN 112954450 A

(43) 申请公布日 2021.06.11

(73) 专利权人 北京字跳网络技术有限公司

地址 100190 北京市海淀区紫金数码园4号楼2层0207

(72) 发明人 宗博文 杨晶生 陈可蓉

(74) 专利代理机构 北京植德律师事务所 11780

专利代理师 唐华东

(51) Int. Cl.

H04N 21/44 (2011.01)

H04N 21/845 (2011.01)

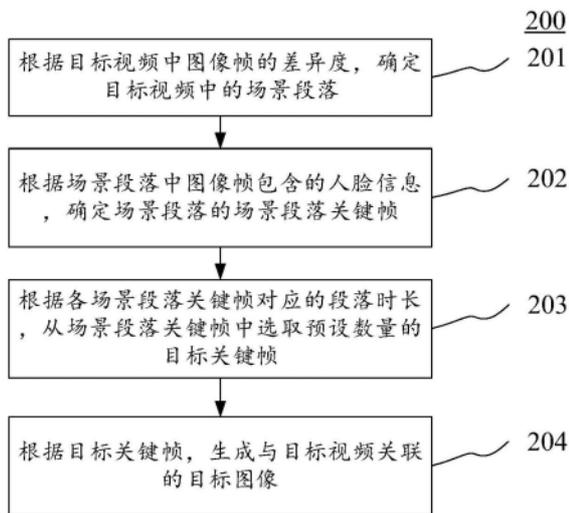
权利要求书4页 说明书16页 附图3页

(54) 发明名称

视频处理方法、装置、电子设备和存储介质

(57) 摘要

本公开提供一种视频处理方法、装置、电子设备和存储介质。该方法的一具体实施方式包括：根据目标视频中图像帧的差异度，确定目标视频中的场景段落；根据场景段落中图像帧包含的人脸信息，确定场景段落的场景段落关键帧；根据各场景段落关键帧对应的段落时长，从场景段落关键帧中选取预设数量的目标关键帧；根据目标关键帧，生成与目标视频关联的目标图像。该实施方式能够获得在人脸信息和段落时长上具有代表性的目标图像，有利于准确合理地体现视频内容。



1. 一种视频处理方法,包括:

根据目标视频中图像帧的差异度,确定所述目标视频中的场景段落,其中,所述目标视频为会议记录视频;

根据所述场景段落中图像帧包含的人脸信息,确定所述场景段落的场景段落关键帧;

根据各所述场景段落关键帧对应的段落时长,从所述场景段落关键帧中选取预设数量的目标关键帧;

根据所述目标关键帧,生成与所述目标视频关联的目标图像,包括:根据各所述目标关键帧包含的人脸信息,对各所述目标关键帧进行截取和拼接处理,得到与所述目标视频关联的目标图像;

所述对各所述目标关键帧进行截取和拼接处理,得到与所述目标视频关联的目标图像,包括:

对于每个所述目标关键帧,基于该目标关键帧中的人脸中心,截取相应的图像区域,并对所述图像区域进行拼接,得到所述目标图像。

2. 根据权利要求1所述的方法,其中,在所述根据各所述场景段落关键帧对应的段落时长,从所述场景段落关键帧中选取预设数量的目标关键帧之后,所述方法还包括:

根据所述目标关键帧,生成所述目标视频的预览视频。

3. 根据权利要求1或2所述的方法,其中,所述根据所述目标视频中图像帧的差异度,确定所述目标视频中的场景段落,包括:

根据所述目标视频中图像帧的差异度,确定所述目标视频的场景切换关键帧;

根据所述场景切换关键帧,确定所述目标视频中的场景段落。

4. 根据权利要求3所述的方法,其中,所述根据所述目标视频中图像帧的差异度,确定所述目标视频的场景切换关键帧,包括:

按照预设时间间隔从所述目标视频中抽取图像帧,得到多个第一图像帧;

计算相邻所述第一图像帧之间的差异度,得到差异度序列;

根据所述差异度序列中的极大值和预设的极大值阈值,确定所述目标视频的场景切换关键帧。

5. 根据权利要求4所述的方法,其中,所述计算相邻所述第一图像帧之间的差异度,包括:

分别计算相邻所述第一图像帧之间的差值哈希距离和结构距离;

根据所述差值哈希距离和所述结构距离,确定所述相邻所述第一图像帧之间的差异度。

6. 根据权利要求1所述的方法,其中,所述根据所述场景段落中图像帧包含的人脸信息,确定所述场景段落的场景段落关键帧,包括:

从所述场景段落中抽取预设数量的图像帧,得到多个第二图像帧;

基于预先训练的机器学习模型,确定所述第二图像帧包含的人脸信息;

根据所述第二图像帧包含的人脸信息,从多个所述第二图像帧中选出所述场景段落关键帧。

7. 根据权利要求6所述的方法,其中,所述根据所述第二图像帧包含的人脸信息,从多个所述第二图像帧中选出所述场景段落关键帧,包括:

根据预设的第一指标体系对所述多个第二图像帧进行排序,得到相应的第一排序结果,其中,所述第一指标体系与所述第二图像帧包含的人脸信息关联;

根据所述第一排序结果,确定所述场景段落关键帧。

8. 根据权利要求7所述的方法,其中,所述第二图像帧包含的人脸信息包括所述第二图像帧的人脸数量信息和所述第二图像帧的正面人脸数量信息中的至少一项。

9. 根据权利要求8所述的方法,其中,所述第二图像帧包含的人脸信息包括所述第二图像帧的人脸数量信息和所述第二图像帧的正面人脸数量信息;以及

所述根据预设的第一指标体系对所述多个第二图像帧进行排序,得到相应的第一排序结果,包括:

根据所述第二图像帧包含的人脸数量信息对所述多个第二图像帧进行排序,得到第一层次排序结果;

在所述第二图像帧包含的人脸数量信息相同的情况下,根据所述正面人脸数量信息对相关的所述第二图像帧进行排序,得到第二层次排序结果;

根据所述第一层次排序结果和所述第二层次排序结果,确定所述第一排序结果。

10. 根据权利要求7所述的方法,其中,所述第一指标体系还与所述第二图像帧的图像质量关联。

11. 根据权利要求10所述的方法,其中,所述第二图像帧的图像质量根据所述第二图像帧的图像清晰度和所述第二图像帧的色彩丰富度中的至少一项确定。

12. 根据权利要求11所述的方法,其中,所述第二图像帧的色彩丰富度通过以下方式确定:

对于所述第二图像帧中的每个像素点,根据该像素点的红色分量与绿色分量之差的绝对值,确定该像素点的第一参数;

对于所述第二图像帧中的每个像素点,根据该像素点的红色分量和绿色分量的均值与蓝色分量之差的绝对值,确定该像素点的第二参数;

根据所述第二图像帧中各像素点的第一参数的标准差和均值,以及所述第二图像帧中各像素点的第二参数的标准差和均值,确定所述第二图像帧的色彩丰富度。

13. 根据权利要求1所述的方法,其中,在所述根据各所述场景段落关键帧对应的段落时长,从所述场景段落关键帧中选取预设数量的目标关键帧之前,所述方法还包括:

根据所述目标视频中不同场景段落的场景段落关键帧之间的相似度预设的相似度阈值,对所述目标视频中不同的场景段落进行合并,得到相应的合并场景段落;

根据合并前各所述场景段落的场景段落关键帧确定所述合并场景段落的场景段落关键帧,以及根据合并前各所述场景段落的段落时长确定所述合并场景段落的段落时长。

14. 根据权利要求1所述的方法,其中,所述根据各所述场景段落关键帧对应的段落时长,从所述场景段落关键帧中选取预设数量的目标关键帧,包括:

根据预设的第二指标体系对各所述场景段落关键帧进行排序,得到相应的第二排序结果,其中,所述第二指标体系与所述场景段落关键帧的段落时长关联;

根据所述第二排序结果和所述预设数量,确定所述目标关键帧。

15. 根据权利要求14所述的方法,其中,所述第二指标体系还与所述场景段落关键帧包含的人脸信息和所述场景段落关键帧的图像质量中的至少一项关联。

16. 根据权利要求1、2和4-13中任一项所述的方法,其中,所述与所述目标视频关联的目标图像为所述目标视频的封面图像。

17. 根据权利要求2所述的方法,其中,所述根据所述目标关键帧,生成所述目标视频的预览视频,包括:

截取所述目标视频中至少一个目标关键帧附近预设时间范围内的视频片段并进行拼接,得到所述目标视频的预览视频。

18. 一种视频处理方法,包括:

根据目标视频中图像帧的差异度,确定所述目标视频中的场景段落,其中,所述目标视频为会议记录视频;

根据所述场景段落中图像帧包含的人脸信息,确定所述场景段落的场景段落关键帧;

根据各所述场景段落关键帧对应的段落时长,从所述场景段落关键帧中选取预设数量的目标关键帧;

根据所述目标关键帧,生成与所述目标视频关联的目标图像,包括:根据各所述目标关键帧包含的人脸信息,对各所述目标关键帧进行截取和拼接处理,得到与所述目标视频关联的目标图像;

所述对各所述目标关键帧进行截取和拼接处理,得到与所述目标视频关联的目标图像,包括:

对于每个所述目标关键帧,基于该目标关键帧中的人脸中心,截取相应的图像区域,并对所述图像区域进行拼接,得到所述目标图像;

根据所述目标关键帧,生成所述目标视频的预览视频。

19. 一种视频处理装置,包括:

第一处理单元,用于根据目标视频中图像帧的差异度,确定所述目标视频中的场景段落,其中,所述目标视频为会议记录视频;

第二处理单元,用于根据所述场景段落中图像帧包含的人脸信息,确定所述场景段落的场景段落关键帧;

第三处理单元,用于根据各所述场景段落关键帧对应的段落时长,从所述场景段落关键帧中选取预设数量的目标关键帧;

目标图像生成单元,用于根据所述目标关键帧,生成与所述目标视频关联的目标图像,包括:根据各所述目标关键帧包含的人脸信息,对各所述目标关键帧进行截取和拼接处理,得到与所述目标视频关联的目标图像;

所述对各所述目标关键帧进行截取和拼接处理,得到与所述目标视频关联的目标图像,包括:

对于每个所述目标关键帧,基于该目标关键帧中的人脸中心,截取相应的图像区域,并对所述图像区域进行拼接,得到所述目标图像。

20. 一种电子设备,包括:

一个或多个处理器;

存储装置,其上存储有一个或多个程序,

当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现如权利要求1-18中任一项所述的方法。

21. 一种计算机可读存储介质,其上存储有计算机程序,其中,所述计算机程序被一个或多个处理器执行时实现如权利要求1-18中任一项所述的方法。

## 视频处理方法、装置、电子设备和存储介质

### 技术领域

[0001] 本公开的实施例涉及视频处理技术领域,具体涉及视频处理方法、装置、电子设备和存储介质。

### 背景技术

[0002] 在对视频文件进行展示时,为了使用户快速了解视频文件的内容,通常会将能够代表视频文件内容的图片或片段进行展示,即为视频文件的封面或预览。

[0003] 现有方式中,通常由用户自行指定视频的封面或预览,或者直接提取视频的第一帧作为视频封面,又或者直接提取视频的开头片段作为视频预览。通过上述方式获得的视频封面或视频预览,存在缺乏代表性甚至与视频实际内容相差甚远的问题。

[0004] 因此,有必要提出一种新的生成视频封面或视频预览的技术方案。

### 发明内容

[0005] 本公开的实施例提出了视频处理方法、装置、电子设备和存储介质。

[0006] 第一方面,本公开提供了一种视频处理方法,包括:

[0007] 根据目标视频中图像帧的差异度,确定上述目标视频中的场景段落;

[0008] 根据上述场景段落中图像帧包含的人脸信息,确定上述场景段落的场景段落关键帧;

[0009] 根据各上述场景段落关键帧对应的段落时长,从上述场景段落关键帧中选取预设数量的目标关键帧;

[0010] 根据上述目标关键帧,生成与上述目标视频关联的目标图像。

[0011] 在一些可选的实施方式中,在上述根据各上述场景段落关键帧对应的段落时长,从上述场景段落关键帧中选取预设数量的目标关键帧之后,上述方法还包括:

[0012] 根据上述目标关键帧,生成上述目标视频的预览视频。

[0013] 在一些可选的实施方式中,上述根据上述目标视频中图像帧的差异度,确定上述目标视频中的场景段落,包括:

[0014] 根据上述目标视频中图像帧的差异度,确定上述目标视频的场景切换关键帧;

[0015] 根据上述场景切换关键帧,确定上述目标视频中的场景段落。

[0016] 在一些可选的实施方式中,上述根据上述目标视频中图像帧的差异度,确定上述目标视频的场景切换关键帧,包括:

[0017] 按照预设时间间隔从上述目标视频中抽取图像帧,得到多个第一图像帧;

[0018] 计算相邻上述第一图像帧之间的差异度,得到差异度序列;

[0019] 根据上述差异度序列中的极大值和预设的极大值阈值,确定上述目标视频的场景切换关键帧。

[0020] 在一些可选的实施方式中,上述计算相邻上述第一图像帧之间的差异度,包括:

[0021] 分别计算相邻上述第一图像帧之间的差值哈希距离和结构距离;

[0022] 根据上述差值哈希距离和上述结构距离,确定上述相邻上述第一图像帧之间的差异度。

[0023] 在一些可选的实施方式中,上述根据上述场景段落中图像帧包含的人脸信息,确定上述场景段落的场景段落关键帧,包括:

[0024] 从上述场景段落中抽取预设数量的图像帧,得到多个第二图像帧;

[0025] 基于预先训练的机器学习模型,确定上述第二图像帧包含的人脸信息;

[0026] 根据上述第二图像帧包含的人脸信息,从多个上述第二图像帧中选出上述场景段落关键帧。

[0027] 在一些可选的实施方式中,上述根据上述第二图像帧包含的人脸信息,从多个上述第二图像帧中选出上述场景段落关键帧,包括:

[0028] 根据预设的第一指标体系对上述多个第二图像帧进行排序,得到相应的第一排序结果,其中,上述第一指标体系与上述第二图像帧包含的人脸信息关联;

[0029] 根据上述第一排序结果,确定上述场景段落关键帧。

[0030] 在一些可选的实施方式中,上述第二图像帧包含的人脸信息包括上述第二图像帧的人脸数量信息和上述第二图像帧的正面人脸数量信息中的至少一项。

[0031] 在一些可选的实施方式中,上述第二图像帧包含的人脸信息包括上述第二图像帧的人脸数量信息和上述第二图像帧的正面人脸数量信息;以及

[0032] 上述根据预设的第一指标体系对上述多个第二图像帧进行排序,得到相应的第一排序结果,包括:

[0033] 根据上述第二图像帧包含的人脸数量信息对上述多个第二图像帧进行排序,得到第一层次排序结果;

[0034] 在上述第二图像帧包含的人脸数量信息相同的情况下,根据上述正面人脸数量信息对相关的上述第二图像帧进行排序,得到第二层次排序结果;

[0035] 根据上述第一层次排序结果和上述第二层次排序结果,确定上述第一排序结果。

[0036] 在一些可选的实施方式中,上述第一指标体系还与上述第二图像帧的图像质量关联。

[0037] 在一些可选的实施方式中,上述第二图像帧的图像质量根据上述第二图像帧的图像清晰度和上述第二图像帧的色彩丰富度中的至少一项确定。

[0038] 在一些可选的实施方式中,上述第二图像帧的色彩丰富度通过以下方式确定:

[0039] 对于上述第二图像帧中的每个像素点,根据该像素点的红色分量与绿色分量之差的绝对值,确定该像素点的第一参数;

[0040] 对于上述第二图像帧中的每个像素点,根据该像素点的红色分量和绿色分量的均值与蓝色分量之差的绝对值,确定该像素点的第二参数;

[0041] 根据上述第二图像帧中各像素点的第一参数的标准差和均值,以及上述第二图像帧中各像素点的第二参数的标准差和均值,确定上述第二图像帧的色彩丰富度。

[0042] 在一些可选的实施方式中,在上述根据各上述场景段落关键帧对应的段落时长,从上述场景段落关键帧中选取预设数量的目标关键帧之前,上述方法还包括:

[0043] 根据上述目标视频中不同场景段落的场景段落关键帧之间的相似度预设的相似度阈值,对上述目标视频中不同的场景段落进行合并,得到相应的合并场景段落;

[0044] 根据合并前各上述场景段落的场景段落关键帧确定上述合并场景段落的场景段落关键帧,以及根据合并前各上述场景段落的段落时长确定上述合并场景段落的段落时长。

[0045] 在一些可选的实施方式中,上述根据各上述场景段落关键帧对应的段落时长,从上述场景段落关键帧中选取预设数量的目标关键帧,包括:

[0046] 根据预设的第二指标体系对各上述场景段落关键帧进行排序,得到相应的第二排序结果,其中,上述第二指标体系与上述场景段落关键帧的段落时长关联;

[0047] 根据上述第二排序结果和上述预设数量,确定上述目标关键帧。

[0048] 在一些可选的实施方式中,上述第二指标体系还与上述场景关键帧包含的人脸信息和上述场景段落关键帧的图像质量中的至少一项关联。

[0049] 在一些可选的实施方式中,上述根据上述目标关键帧,生成与上述目标视频关联的目标图像,包括:

[0050] 根据各上述目标关键帧包含的人脸信息,对各上述目标关键帧进行截取和拼接处理,得到与上述目标视频关联的目标图像。

[0051] 在一些可选的实施方式中,上述与上述目标视频关联的目标图像为上述目标视频的封面图像。

[0052] 在一些可选的实施方式中,上述根据上述目标关键帧,生成上述目标视频的预览视频,包括:

[0053] 截取上述目标视频中至少一个目标关键帧附近预设时间范围内的视频片段并进行拼接,得到上述目标视频的预览视频。

[0054] 第二方面,本公开提供了一种视频处理装置,包括:

[0055] 第一处理单元,用于根据目标视频中图像帧的差异度,确定上述目标视频中的场景段落;

[0056] 第二处理单元,用于根据上述场景段落中图像帧包含的人脸信息,确定上述场景段落的场景段落关键帧;

[0057] 第三处理单元,用于根据各上述场景段落关键帧对应的段落时长,从上述场景段落关键帧中选取预设数量的目标关键帧;

[0058] 目标图像生成单元,用于根据上述目标关键帧,生成与上述目标视频关联的目标图像。

[0059] 在一些可选的实施方式中,上述装置还包括预览生成单元,上述预览生成单元用于:

[0060] 根据上述目标关键帧,生成上述目标视频的预览视频。

[0061] 在一些可选的实施方式中,上述第一处理单元进一步用于:

[0062] 根据上述目标视频中图像帧的差异度,确定上述目标视频的场景切换关键帧;

[0063] 根据上述场景切换关键帧,确定上述目标视频中的场景段落。

[0064] 在一些可选的实施方式中,上述第一处理单元进一步用于:

[0065] 按照预设时间间隔从上述目标视频中抽取图像帧,得到多个第一图像帧;

[0066] 计算相邻上述第一图像帧之间的差异度,得到差异度序列;

[0067] 根据上述差异度序列中的极大值和预设的极大值阈值,确定上述目标视频的场景

切换关键帧。

[0068] 在一些可选的实施方式中,上述第一处理单元进一步用于:

[0069] 分别计算相邻上述第一图像帧之间的差值哈希距离和结构距离;

[0070] 根据上述差值哈希距离和上述结构距离,确定上述相邻上述第一图像帧之间的差异度。

[0071] 在一些可选的实施方式中,上述第二处理单元进一步用于:

[0072] 从上述场景段落中抽取预设数量的图像帧,得到多个第二图像帧;

[0073] 基于预先训练的机器学习模型,确定上述第二图像帧包含的人脸信息;

[0074] 根据上述第二图像帧包含的人脸信息,从多个上述第二图像帧中选出上述场景段落关键帧。

[0075] 在一些可选的实施方式中,上述第二处理单元进一步用于:

[0076] 根据预设的第一指标体系对上述多个第二图像帧进行排序,得到相应的第一排序结果,其中,上述第一指标体系与上述第二图像帧包含的人脸信息关联;

[0077] 根据上述第一排序结果,确定上述场景段落关键帧。

[0078] 在一些可选的实施方式中,上述第二图像帧包含的人脸信息包括上述第二图像帧的人脸数量信息和上述第二图像帧的正面人脸数量信息中的至少一项。

[0079] 在一些可选的实施方式中,上述第二图像帧包含的人脸信息包括上述第二图像帧的人脸数量信息和上述第二图像帧的正面人脸数量信息;以及

[0080] 上述第二处理单元进一步用于:

[0081] 根据上述第二图像帧包含的人脸数量信息对上述多个第二图像帧进行排序,得到第一层次排序结果;

[0082] 在上述第二图像帧包含的人脸数量信息相同的情况下,根据上述正面人脸数量信息对相关的上述第二图像帧进行排序,得到第二层次排序结果;

[0083] 根据上述第一层次排序结果和上述第二层次排序结果,确定上述第一排序结果。

[0084] 在一些可选的实施方式中,上述第一指标体系还与上述第二图像帧的图像质量关联。

[0085] 在一些可选的实施方式中,上述第二图像帧的图像质量根据上述第二图像帧的图像清晰度和上述第二图像帧的色彩丰富度中的至少一项确定。

[0086] 在一些可选的实施方式中,上述第二图像帧的色彩丰富度通过以下方式确定:

[0087] 对于上述第二图像帧中的每个像素点,根据该像素点的红色分量与绿色分量之差的绝对值,确定该像素点的第一参数;

[0088] 对于上述第二图像帧中的每个像素点,根据该像素点的红色分量和绿色分量的均值与蓝色分量之差的绝对值,确定该像素点的第二参数;

[0089] 根据上述第二图像帧中各像素点的第一参数的标准差和均值,以及上述第二图像帧中各像素点的第二参数的标准差和均值,确定上述第二图像帧的色彩丰富度。

[0090] 在一些可选的实施方式中,上述装置还包括合并单元,上述合并单元用于:

[0091] 根据上述目标视频中不同场景段落的场景段落关键帧之间的相似度预设的相似度阈值,对上述目标视频中不同的场景段落进行合并,得到相应的合并场景段落;

[0092] 根据合并前各上述场景段落的场景段落关键帧确定上述合并场景段落的场景段

落关键帧,以及根据合并前各上述场景段落的段落时长确定上述合并场景段落的段落时长。

[0093] 在一些可选的实施方式中,上述第三处理单元进一步用于:

[0094] 根据预设的第二指标体系对各上述场景段落关键帧进行排序,得到相应的第二排序结果,其中,上述第二指标体系与上述场景段落关键帧的段落时长关联;

[0095] 根据上述第二排序结果和上述预设数量,确定上述目标关键帧。

[0096] 在一些可选的实施方式中,上述第二指标体系还与上述场景关键帧包含的人脸信息和上述场景段落关键帧的图像质量中的至少一项关联。

[0097] 在一些可选的实施方式中,上述目标图像生成单元进一步用于:

[0098] 根据各上述目标关键帧包含的人脸信息,对各上述目标关键帧进行截取和拼接处理,得到与上述目标视频关联的目标图像。

[0099] 在一些可选的实施方式中,上述与上述目标视频关联的目标图像为上述目标视频的封面图像。

[0100] 在一些可选的实施方式中,上述目标图像生成单元进一步用于:

[0101] 截取上述目标视频中至少一个目标关键帧附近预设时间范围内的视频片段并进行拼接,得到上述目标视频的预览视频。

[0102] 第三方面,本公开提供了一种电子设备,包括:

[0103] 一个或多个处理器;

[0104] 存储装置,其上存储有一个或多个程序,

[0105] 当上述一个或多个程序被上述一个或多个处理器执行时,使得上述一个或多个处理器实现如本公开第一方面任一实施方式描述的方法。

[0106] 第四方面,本公开提供了一种计算机可读存储介质,其上存储有计算机程序,其中,上述计算机程序被一个或多个处理器执行时实现如本公开第一方面任一实施方式描述的方法。

[0107] 本公开的实施例提供的视频处理方法、装置、电子设备和存储介质,根据场景段落中图像帧包含的人脸信息确定场景段落关键帧,根据段落时长从场景段落关键帧中选取目标关键帧进而生成与目标视频关联的目标图像,能够获得在人脸信息和段落时长上具有代表性的目标图像,有利于准确合理地体现视频内容,进而可以提高生成视频封面的效果和效率以及所生成视频封面的质量。

[0108] 此外,本公开的实施例中,根据目标视频中图像帧的差异度确定目标视频中的场景段落,仅利用图像信息完成视频内容分段,实施方便,简便可靠。

[0109] 此外,本实施例中各个处理步骤可以基于轻量化计算规则或轻量化神经网络实现,算法轻量快速,使得该视频处理方案能够进行实时检测,还能够部署在纯CPU (Central Processing Unit,中央处理器) 硬件环境中。

## 附图说明

[0110] 通过阅读参照以下附图所作的对非限制性实施例所作的详细描述,本公开的其它特征、目的和优点将会变得更明显。附图仅用于示出具体实施方式的目的,而并不认为是对本发明的限制。在附图中:

- [0111] 图1是根据本公开的视频处理系统的一个实施例的系统架构图；
- [0112] 图2是根据本公开的视频处理方法的一个实施例的流程图；
- [0113] 图3是根据本公开的视频处理方法的一个具体例子的示意图；
- [0114] 图4是根据本公开的视频处理装置的一个实施例的结构示意图；
- [0115] 图5是适于用来实现本公开的实施例的电子设备的计算机系统的结构示意图。

### 具体实施方式

[0116] 下面结合附图和实施例对本公开作进一步的详细说明。可以理解的是，此处所描述的具体实施例仅仅用于解释相关发明，而非对该发明的限定。另外还需要说明的是，为了便于描述，附图中仅示出了与有关发明相关的部分。

[0117] 需要说明的是，在不冲突的情况下，本公开中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本公开。

[0118] 图1示出了可以应用本公开的视频处理方法、装置、终端设备和存储介质的实施例的示例性系统架构100。

[0119] 如图1所示，系统架构100可以包括终端设备101、102、103，网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

[0120] 用户可以使用终端设备101、102、103通过网络104与服务器105交互，以接收或发送消息等。终端设备101、102、103上可以安装有各种通讯客户端应用，例如语音交互类应用、视频会议类应用、短视频社交类应用、网页浏览器应用、购物类应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

[0121] 终端设备101、102、103可以是硬件，也可以是软件。当终端设备101、102、103为硬件时，可以是具有麦克风和扬声器的各种电子设备，包括但不限于智能手机、平板电脑、电子书阅读器、MP3播放器 (Moving Picture Experts Group Audio Layer III, 动态影像专家压缩标准音频层面3)、MP4 (Moving Picture Experts Group Audio Layer IV, 动态影像专家压缩标准音频层面4) 播放器、膝上型便携计算机和台式计算机等等。当终端设备101、102、103为软件时，可以安装在上述所列举的电子设备中。其可以实现成多个软件或软件模块 (例如用来视频处理服务)，也可以实现成单个软件或软件模块。在此不做具体限定。

[0122] 服务器105可以是提供各种服务的服务器，例如对终端设备101、102、103上采集的视频提供处理服务的后台服务器。后台服务器可以对接收到的视频等进行相应处理。

[0123] 在一些情况下，本公开所提供的视频处理方法可以由终端设备101、102、103和服务器105共同执行，例如，“根据目标视频中图像帧的差异度，确定目标视频中的场景段落”的步骤可以由终端设备101、102、103执行，“根据场景段落中图像帧包含的人脸信息，确定场景段落的场景段落关键帧”的步骤可以由服务器105执行。本公开对此不做限定。相应地，视频处理装置也可以分别设置于终端设备101、102、103和服务器105中。

[0124] 在一些情况下，本公开所提供的视频处理方法可以由终端设备101、102、103执行，相应地，视频处理装置也可以设置于终端设备101、102、103中，这时，系统架构100也可以不包括服务器105。

[0125] 在一些情况下，本公开所提供的视频处理方法可以由服务器105执行，相应地，视

频处理装置也可以设置于服务器105中,这时,系统架构100也可以不包括终端设备101、102、103。

[0126] 需要说明的是,服务器105可以是硬件,也可以是软件。当服务器105为硬件时,可以实现成多个服务器组成的分布式服务器集群,也可以实现成单个服务器。当服务器105为软件时,可以实现成多个软件或软件模块(例如用来提供分布式服务),也可以实现成单个软件或软件模块。在此不做具体限定。

[0127] 应该理解,图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器的。

[0128] 继续参考图2,其示出了根据本公开的视频处理方法的一个实施例的流程200,应用于图1中的终端设备或服务器,该流程200包括以下步骤:

[0129] 步骤201,根据目标视频中图像帧的差异度,确定目标视频中的场景段落。

[0130] 本实施例的目标视频可以通过以下方式获取:在执行主体为终端设备的情况下,可以由终端设备中的摄像模块进行拍摄,从而获得目标视频。还可以由终端设备通过网络接收其他终端设备或者服务器发送的视频,从而获得目标视频。在执行主体为服务器的情况下,可以由服务器通过网络接收终端设备或者其他服务器发送的视频,从而获得目标视频。在执行主体为终端设备和服务器的情况下,可以由终端设备通过摄像头拍摄视频并将其发送至服务器进行处理。

[0131] 在本实施例中,目标视频可以包括一系列的图像帧。其中,图像帧是组成视频的最小单位。一个图像帧是一幅静止的图像或画面。

[0132] 在一个视频会议场景的例子中,目标视频可以是会议记录视频,由各个参会方的终端设备所拍摄的画面组成。

[0133] 在本实施例中,图像帧的差异度可以表示不同图像帧所包含的画面信息的差异大小。

[0134] 在本实施例中,场景可以包括环境背景、人物等要素。一个场景段落可以是对应某个特定场景的视频片段。例如,在视频会议的例子中,某个参会者发言的视频片段可形成一个场景段落。

[0135] 在一个例子中,步骤201可以按照如下方式实施:

[0136] 第一步,根据目标视频中图像帧的差异度,确定目标视频的场景切换关键帧。

[0137] 在本实施例中,场景切换关键帧是指用于区分不同场景的图像帧。在一个例子中,可以按照如下方式确定场景切换关键帧:

[0138] 首先,按照预设时间间隔从目标视频中抽取图像帧,得到多个第一图像帧。例如,可以按照1秒的时间间隔从目标视频中抽取图像帧。

[0139] 其次,计算相邻第一图像帧之间的差异度,得到差异度序列。

[0140] 这里,相邻第一图像帧之间的差异度可以按照如下方式确定:分别计算相邻第一图像帧之间的差值哈希距离和结构距离;根据差值哈希距离和结构距离,确定相邻第一图像帧之间的差异度。

[0141] 差异哈希算法(Different hash algorithm,DHA)是哈希算法的一类,可以用来做相似图片的搜索工作。差值哈希的处理过程可以包括缩小图片、灰度化、计算灰度差值和计算哈希值等步骤。

[0142] 在本实施例中,可以计算两个第一图像帧的差值哈希结果之间的汉明距离,并将其作为两个第一图像帧之间的差值哈希距离。

[0143] 结构相似度(Structural Similarity,SSIM)是一种衡量两幅图像相似度的指标。结构相似度的计算过程可以包括修改尺寸、灰度化、加窗和计算等步骤。

[0144] 在本实施例中,可以利用预设值(例如1)减去两个第一图像帧的结构相似度,并将得到的差作为两个第一图像帧的结构距离。

[0145] 在本实施例中,可以计算两个第一图像帧之间的差值哈希距离和结构距离的和,作为两个第一图像帧之间的差异度。

[0146] 在本实施例中,可以将各个差异度按其对应时间的先后顺序排列,形成差异度序列。这里,某个差异度对应的时间,可以是该差异度对应的两个第一图像帧之中在先图像帧的时间,也可以是该差异度对应的两个第一图像帧之中在后图像帧的时间,还可以是该差异度对应的两个第一图像帧的时间均值,等等。

[0147] 最后,根据差异度序列中的极大值和预设的极大值阈值,确定目标视频的场景切换关键帧。

[0148] 在本实施例中,可以计算差异度序列中的极大值,并将其与预设的极大值阈值进行比较,选出大于差异度阈值的极大值。对于选出的极大值,可以将该极大值对应的两个第一图像帧中的一个(在先图像帧或在后图像帧)作为场景切换关键帧。

[0149] 第二步,根据场景切换关键帧,确定目标视频中的场景段落。

[0150] 在本实施例中,可以将相邻两个场景切换关键帧之间的视频段落,确定为相应的场景段落。

[0151] 在本实施例中,根据目标视频中图像帧的差异度确定目标视频中的场景段落,仅利用图像信息完成视频内容分段,实施方便,简便可靠。

[0152] 步骤202,根据场景段落中图像帧包含的人脸信息,确定场景段落的场景段落关键帧。

[0153] 在本实施例中,场景段落关键帧是能够反映场景段落主要信息的图像帧。

[0154] 在一个例子中,第二图像帧包含的人脸信息包括第二图像帧的人脸数量信息和第二图像帧的正面人脸数量信息中的至少一项。

[0155] 在一个例子中,步骤202可以按照如下方式实施:

[0156] 第一步,从场景段落中抽取预设数量的图像帧,得到多个第二图像帧。例如,可以抽取场景段落中前若干帧(例如6-10帧)图像帧作为第二图像帧。

[0157] 第二步,基于预先训练的机器学习模型,确定第二图像帧包含的人脸信息。

[0158] 在一个例子中,上述人脸信息可以通过以下方式获取:

[0159] 首先,将第二图像帧输入机器学习模型,得到相应的人脸检测结果,其中,人脸检测结果包括第二图像帧中人脸的人脸置信度、人脸区域范围和人脸五官关键点。

[0160] 这里,机器学习模型可以是轻量卷积神经网络模型。

[0161] 这里,人脸五官关键点可以包括鼻子、双眼和双嘴角的位置。

[0162] 其次,根据人脸五官关键点,确定人脸检测结果中的人脸是否为正面人脸。

[0163] 这里,可以通过判断鼻子位置是否位于双眼和双嘴角围成的四边形内部来确定人脸检测结果中的人脸是否为正面人脸。

[0164] 最后,根据人脸检测结果中的人脸数量和正面人脸数量,得到第二图像帧包含的人脸信息。

[0165] 第三步,根据第二图像帧包含的人脸信息,从多个第二图像帧中选出场景段落关键帧。

[0166] 在一个例子中,场景段落关键帧的选择过程可以包括如下步骤:

[0167] 首先,根据预设的第一指标体系对多个第二图像帧进行排序,得到相应的第一排序结果,其中,第一指标体系与第二图像帧包含的人脸信息关联。

[0168] 在一个例子中,第一指标体系可以包括正面人脸数量指标和人脸数量指标,并且正面人脸数量指标的优先级高于人脸数量指标。相应地,第一排序结果可以通过以下方式获取:根据第二图像帧包含的人脸数量信息对多个第二图像帧进行排序,得到第一层次排序结果;在第二图像帧包含的人脸数量信息相同的情况下,根据正面人脸数量信息对相关的第二图像帧进行排序,得到第二层次排序结果;根据第一层次排序结果和第二层次排序结果,确定第一排序结果。

[0169] 例如,假设多个第二图像帧分别为a1、a2、a3和a4,其中a1包含6个正面人脸和7个人脸,a2包含5个正面人脸和5个人脸,a3包含5个正面人脸和7个人脸,a4包含3个正面人脸和7个人脸。由于正面人脸数量这一指标的优先级最高,因此先按正面人脸数量排序,得到第一层次排序结果为 $a1 > a2 = a3 > a4$ 。由于 $a2 = a3$ ,二者之间再根据人脸数量排序,得到的第二层次排序结果为 $a3 > a2$ 。因此,最终得到的第一排序结果为 $a1 > a3 > a2 > a4$ 。

[0170] 其次,根据第一排序结果,确定场景段落关键帧。

[0171] 在本实施例中,可以根据第一排序结果,选择排序最靠前的第二图像帧作为场景段落关键帧。

[0172] 在一个例子中,上述第一指标体系还与第二图像帧的图像质量关联,并且各指标优先级由高到低依次为正面人脸数量指标、人脸数量指标和图像质量指标。即,在两个第二图像帧对应的正面人脸数量和人脸数量均相同的情况下,根据第二图像帧的图像质量确定这两个第二图像帧的排序。

[0173] 第二图像帧的图像质量根据第二图像帧的图像清晰度和第二图像帧的色彩丰富度中的至少一项确定。

[0174] 在一个例子中,第二图像帧的图像质量可以通过以下方式获取:

[0175] 首先,计算第二图像帧的图像清晰度。

[0176] 例如,可以先对第二图像帧进行拉普拉斯运算得到相应的边缘特征图,再对边缘特征图中的像素求方差得到图像清晰度,并对得到的图像清晰度行标准化处理。

[0177] 对于多个第二图像帧中的每个第二图像帧,该第二图像帧的图像清晰度的标准化处理过程为:先计算多个第二图像帧的图像清晰度的清晰度均值和标准差,再计算该第二图像帧的图像清晰度与上述清晰度均值的差值,最后计算该差值与上述标准差的比值,得到的结果即为该第二图像帧的图像清晰度的标准化处理结果。

[0178] 其次,计算第二图像帧的色彩丰富度。

[0179] 在一个例子中,可以通过如下方式计算第二图像帧的色彩丰富:

[0180] 第一步,对于第二图像帧中的每个像素点,根据该像素点的红色分量与绿色分量之差的绝对值,确定该像素点的第一参数。例如,可以通过下式(1),计算第二图像帧中每个

像素点的第一参数:

$$[0181] \quad rg = |R - G| \quad \text{式 (1)}$$

[0182] 其中,rg为第二图像帧中像素点的第一参数,R为第二图像帧中像素点的红色通道数值,G为第二图像帧中像素点的绿色通道数值;

[0183] 第二步,对于第二图像帧中的每个像素点,根据该像素点的红色分量和绿色分量的均值与蓝色分量之差的绝对值,确定该像素点的第二参数。例如,可以通过下式(2),计算第二图像帧中每个像素点的第二参数:

$$[0184] \quad yb = |0.5 \times (R + G) - B| \quad \text{式 (2)}$$

[0185] 其中,yb为第二图像帧中像素点的第二参数,B为第二图像帧中像素点的蓝色通道数值;

[0186] 第三步,根据第二图像帧中各像素点的第一参数的标准差和均值,以及第二图像帧中各像素点的第二参数的标准差和均值,确定第二图像帧的色彩丰富度。例如,可以通过下式(3),计算第二图像帧中的色彩丰富度:

$$[0187] \quad c = \sqrt{\text{std}(rg)^2 + \text{std}(yb)^2} + 0.3 \times \sqrt{\text{mean}(rg)^2 + \text{mean}(yb)^2} \quad \text{式 (3)}$$

[0188] 其中,c为第二图像帧中的色彩丰富度,std(rg)为第二图像帧中各像素点的第一参数的标准差,std(yb)为第二图像帧中各像素点的第二参数的标准差,mean(rg)为第二图像帧中各像素点的第一参数的均值,mean(yb)为第二图像帧中各像素点的第二参数的均值。

[0189] 在本实施例中,对色彩丰富度也进行标准化处理。该标准化处理过程与图像清晰度的标准化处理过程类似,这里不再赘述。

[0190] 通过对色彩丰富度和图像清晰度进行标准化处理,能够使二者的数值具有可比性,保证后续求和结果具有参考价值。

[0191] 最后,根据图像清晰度和色彩丰富度,确定第二图像帧的图像质量。

[0192] 例如,计算标准化处理后的图像清晰度和标准化处理后的图像色彩丰富度之和,得到第二图像帧的图像质量。

[0193] 在其他例子中,可以只根据标准化处理后的图像清晰度确定第二图像帧的图像质量,或者只根据标准化处理后的图像色彩丰富度确定第二图像帧的图像质量,还可以通过其他方式确定第二图像帧的图像质量,本实施例对此不作限定。

[0194] 步骤203,根据各场景段落关键帧对应的段落时长,从场景段落关键帧中选取预设数量的目标关键帧。

[0195] 在本实施例中,目标关键帧可以用于生成目标视频的封面图像或预览视频。

[0196] 在一个例子中,步骤203可以按照如下方式实施:

[0197] 首先,根据预设的第二指标体系对各场景段落关键帧进行排序,得到相应的第二排序结果,其中,第二指标体系与场景段落关键帧的段落时长关联。

[0198] 这里,可以按照段落时长由大到小的顺序对各场景段落关键帧进行排序。

[0199] 其次,根据第二排序结果和预设数量,确定目标关键帧。

[0200] 在一个例子中,预设数量为3,则可以根据第二排序结果,选择排序前三位的场景

段落关键帧作为目标关键帧。

[0201] 第二指标体系还可以与场景关键帧包含的人脸信息和场景段落关键帧的图像质量中的至少一项关联。

[0202] 在一个例子中,除了段落时长外,第二指标体系还包括正面人脸数量指标、人脸数量指标和图像质量指标,其中,各指标按照优先级由高到低的顺序依次为段落时长指标、正面人脸数量指标、人脸数量指标和图像质量指标。

[0203] 基于第二指标体系的排序步骤与基于第一指标体系的排序步骤类似,这里不再赘述。

[0204] 在一个例子中,在步骤203之前,可以先对各个场景段落执行如下合并操作:

[0205] 首先,根据目标视频中不同场景段落的场景段落关键帧之间的相似度预设的相似度阈值,对目标视频中不同的场景段落进行合并,得到相应的合并场景段落。

[0206] 这里,可以分别计算两个场景段落关键帧之间结构相似度和梯度直方相似度,并用二者之和作为两个场景段落关键帧之间的相似度。在两个场景段落关键帧之间的相似度大于预设的相似度阈值,对这两个场景段落关键帧对应的两个场景段落进行合并,得到相应的合并场景段落。

[0207] 可以将符合条件的两个场景段落按照先后顺序拼接,得到相应的合并场景段落。

[0208] 其次,根据合并前各场景段落的场景段落关键帧确定合并场景段落的场景段落关键帧,以及根据合并前各场景段落的段落时长确定合并场景段落的段落时长。

[0209] 这里,合并前各场景段落的段落时长之和,即为合并场景段落的段落时长。

[0210] 这里,可以将合并前各场景段落的场景段落关键帧,依次按照正面人脸数量、人脸数量、图像质量倒序排列,取排序第一位的场景段落关键帧作为合并场景段落的场景段落关键帧。

[0211] 在本实施例中,可以基于上述合并操作的结果执行步骤203,避免选出过于相近的目标关键帧,导致目标图像中出现重复信息。

[0212] 步骤204,根据目标关键帧,生成与目标视频关联的目标图像。

[0213] 在一个例子中,可以根据各目标关键帧包含的人脸信息,对各目标关键帧进行截取和拼接处理,生成与目标视频关联的目标图像。

[0214] 例如,假设存在3个目标关键帧,对于每个目标关键帧,可以基于该目标关键帧中的人脸中心,截取高度与原高度相同、宽度为原宽度三分之一的图像区域,并将截取的各个图像区域横向拼接,形成相应的目标图像。

[0215] 这里,目标图像可以作为目标视频的封面图像。

[0216] 本公开的实施例提供的视频处理方法,根据场景段落中图像帧包含的人脸信息确定场景段落关键帧,根据段落时长从场景段落关键帧中选取目标关键帧进而生成与目标视频关联的目标图像,能够获得在人脸信息和段落时长上具有代表性的目标图像,有利于准确合理地体现视频内容,进而可以提高生成视频封面的效果和效率以及所生成视频封面的质量。

[0217] 此外,本实施例中对各个处理步骤可以基于轻量化计算规则或轻量化神经网络实现,算法轻量快速,使得该视频处理方案能够进行实时检测,还能够部署在纯CPU (Central Processing Unit,中央处理器) 硬件环境中。

[0218] 在一个例子中,在步骤204之后,还可以根据目标关键帧,生成目标视频的预览视频。

[0219] 例如,截取各目标关键帧附近预设时间范围内的视频片段并进行拼接,得到目标视频的预览视频。

[0220] 上述预览视频的生成方式,使得预览视频在人脸信息和段落时长上具有代表性,能够准确合理地体现视频内容,提高了生成预览视频的效果和效率以及所生成预览视频的质量。

[0221] 图3是根据本公开的视频处理方法的一个具体例子的示意图。图3中的状态1为场景切换关键帧的检测结果,其中水平实线代表目标视频(部分),垂直虚线代表场景切换关键帧,包括切换帧1至切换帧6。其中,相邻两个场景切换关键帧之间的视频段落为相应的场景段落,例如切换帧1和切换帧2之间的实线段为场景段落12,切换帧3和切换帧4之间的实线段为场景段落34。

[0222] 图3中的状态2为提取场景段落关键帧的结果。其中,垂直实线代表场景段落关键帧,包括段落帧1至段落帧5。从图中可以看出,段落帧1对应于场景段落12,段落帧3对应于场景段落34,等等。

[0223] 图3中的状态3为场景段落合并的结果。这里,假设段落帧1和段落2之间的相似度大于预设的相似度阈值,段落帧4和段落帧5之间的相似度大于预设的相似度阈值,因此可以将场景段落12和场景段落23合并得到场景段落13,可以将场景段落45和场景段落56合并得到场景段落46。

[0224] 图3中的状态4为合并操作后的场景段落关键帧。由图3可见,合并操作后得到的场景段落为场景段落13、场景段落34和场景段落46,其对应的场景段落关键帧依次为段落帧1、段落帧3和段落帧4。

[0225] 在图3所示的例子中,假设目标关键帧的数目为2,由于段落帧1对应最大段落时长,段落帧4对应第二大段落时长,因此最终的目标关键帧为段落帧1和段落帧4。

[0226] 进一步参考图4,作为对上述各图所示方法的实现,本公开提供了一种视频处理装置的一个实施例,该装置实施例与图2所示的方法实施例相对应,该装置具体可以应用于各种终端设备中。

[0227] 如图4所示,本实施例的视频处理装置400包括:第一处理单元401、第二处理单元402、第三处理单元403和目标图像生成单元404。其中,第一处理单元401,用于根据目标视频中图像帧的差异度,确定目标视频中的场景段落;第二处理单元402,用于根据场景段落中图像帧包含的人脸信息,确定场景段落的场景段落关键帧;第三处理单元403,用于根据各场景段落关键帧对应的段落时长,从场景段落关键帧中选取预设数量的目标关键帧;目标图像生成单元404,用于根据目标关键帧,生成与目标视频关联的目标图像。

[0228] 在本实施例中,视频处理装置400的第一处理单元401、第二处理单元402、第三处理单元403和目标图像生成单元404的具体处理及其所带来的技术效果可分别参考图2对应实施例中步骤201、步骤202、步骤203和步骤204的相关说明,在此不再赘述。

[0229] 在一些可选的实施方式中,上述装置还可以包括预览生成单元(图4中未示出),上述预览生成单元用于:根据上述目标关键帧,生成上述目标视频的预览视频。

[0230] 在一些可选的实施方式中,上述第一处理单元401可以进一步用于:根据上述目标

视频中图像帧的差异度,确定上述目标视频的场景切换关键帧;根据上述场景切换关键帧,确定上述目标视频中的场景段落。

[0231] 在一些可选的实施方式中,上述第一处理单元401可以进一步用于:按照预设时间间隔从上述目标视频中抽取图像帧,得到多个第一图像帧;计算相邻上述第一图像帧之间的差异度,得到差异度序列;根据上述差异度序列中的极大值和预设的极大值阈值,确定上述目标视频的场景切换关键帧。

[0232] 在一些可选的实施方式中,上述第一处理单元401可以进一步用于:分别计算相邻上述第一图像帧之间的差值哈希距离和结构距离;根据上述差值哈希距离和上述结构距离,确定上述相邻上述第一图像帧之间的差异度。

[0233] 在一些可选的实施方式中,上述第二处理单元402进一步用于:从上述场景段落中抽取预设数量的图像帧,得到多个第二图像帧;基于预先训练的机器学习模型,确定上述第二图像帧包含的人脸信息;根据上述第二图像帧包含的人脸信息,从多个上述第二图像帧中选出上述场景段落关键帧。

[0234] 在一些可选的实施方式中,上述第二处理单元402可以进一步用于:根据预设的第一指标体系对上述多个第二图像帧进行排序,得到相应的第一排序结果,其中,上述第一指标体系与上述第二图像帧包含的人脸信息关联;根据上述第一排序结果,确定上述场景段落关键帧。

[0235] 在一些可选的实施方式中,上述第二图像帧包含的人脸信息可以包括上述第二图像帧的人脸数量信息和上述第二图像帧的正面人脸数量信息中的至少一项。

[0236] 在一些可选的实施方式中,上述第二图像帧包含的人脸信息可以包括上述第二图像帧的人脸数量信息和上述第二图像帧的正面人脸数量信息;以及上述第二处理单元402可以进一步用于:根据上述第二图像帧包含的人脸数量信息对上述多个第二图像帧进行排序,得到第一层次排序结果;在上述第二图像帧包含的人脸数量信息相同的情况下,根据上述正面人脸数量信息对相关的上述第二图像帧进行排序,得到第二层次排序结果;根据上述第一层次排序结果和上述第二层次排序结果,确定上述第一排序结果。

[0237] 在一些可选的实施方式中,上述第一指标体系还可以与上述第二图像帧的图像质量关联。

[0238] 在一些可选的实施方式中,上述第二图像帧的图像质量可以根据上述第二图像帧的图像清晰度和上述第二图像帧的色彩丰富度中的至少一项确定。

[0239] 在一些可选的实施方式中,上述第二图像帧的色彩丰富度可以通过以下方式确定:对于上述第二图像帧中的每个像素点,根据该像素点的红色分量与绿色分量之差的绝对值,确定该像素点的第一参数;对于上述第二图像帧中的每个像素点,根据该像素点的红色分量和绿色分量的均值与蓝色分量之差的绝对值,确定该像素点的第二参数;根据上述第二图像帧中各像素点的第一参数的标准差和均值,以及上述第二图像帧中各像素点的第二参数的标准差和均值,确定上述第二图像帧的色彩丰富度。

[0240] 在一些可选的实施方式中,上述装置还可以包括合并单元(图4中未示出),上述合并单元可以用于:根据上述目标视频中不同场景段落的关键帧之间的相似度预设的相似度阈值,对上述目标视频中不同的场景段落进行合并,得到相应的合并场景段落;根据合并前各上述场景段落的关键帧确定上述合并场景段落的关键帧,以

及根据合并前各上述场景段落的段落时长确定上述合并场景段落的段落时长。

[0241] 在一些可选的实施方式中,上述第三处理单元403可以进一步用于:根据预设的第二指标体系对各上述场景段落关键帧进行排序,得到相应的第二排序结果,其中,上述第二指标体系与上述场景段落关键帧的段落时长关联;根据上述第二排序结果和上述预设数量,确定上述目标关键帧。

[0242] 在一些可选的实施方式中,上述第二指标体系还可以与上述场景关键帧包含的人脸信息和上述场景段落关键帧的图像质量中的至少一项关联。

[0243] 在一些可选的实施方式中,上述目标图像生成单元404可以进一步用于:根据各上述目标关键帧包含的人脸信息,对各上述目标关键帧进行截取和拼接处理,得到与上述目标视频关联的目标图像。

[0244] 在一些可选的实施方式中,上述与上述目标视频关联的目标图像可以为上述目标视频的封面图像。

[0245] 在一些可选的实施方式中,上述目标图像生成单元404可以进一步用于:截取上述目标视频中至少一个目标关键帧附近预设时间范围内的视频片段并进行拼接,得到上述目标视频的预览视频。

[0246] 需要说明的是,本公开的实施例提供的视频处理装置中各单元的实现细节和技术效果可以参考本公开中其它实施例的说明,在此不再赘述。

[0247] 下面参考图5,其示出了适于用来实现本公开的终端设备的计算机系统500的结构示意图。图5示出的计算机系统500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。

[0248] 如图5所示,计算机系统500可以包括处理装置(例如中央处理器、图形处理器等)501,其可以根据存储在只读存储器(ROM)502中的程序或者从存储装置508加载到随机访问存储器(RAM)503中的程序而执行各种适当的动作和处理。在RAM 503中,还存储有计算机系统500操作所需的各种程序和数据。处理装置501、ROM 502以及RAM 503通过总线504彼此相连。输入/输出(I/O)接口505也连接至总线504。

[0249] 通常,以下装置可以连接至I/O接口505:包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风等的输入装置506;包括例如液晶显示器(LCD)、扬声器、振动器等的输出装置507;包括例如磁带、硬盘等的存储装置508;以及通信装置509。通信装置509可以允许计算机系统500与其他设备进行无线或有线通信以交换数据。虽然图5示出了具有各种装置的电子设备的计算机系统500,但是应理解的是,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。

[0250] 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509从网络上被下载和安装,或者从存储装置508被安装,或者从ROM 502被安装。在该计算机程序被处理装置501执行时,执行本公开实施例的方法中限定的上述功能。

[0251] 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。

[0252] 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。

[0253] 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，使得该电子设备实现如图2所示的实施例及其可选实施方式示出的视频处理方法。

[0254] 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对象的程序设计语言——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 (LAN) 或广域网 (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)。

[0255] 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。

[0256] 描述于本公开的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，获取单元还可以被描述为“获取目标视频的单元”。

[0257] 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。

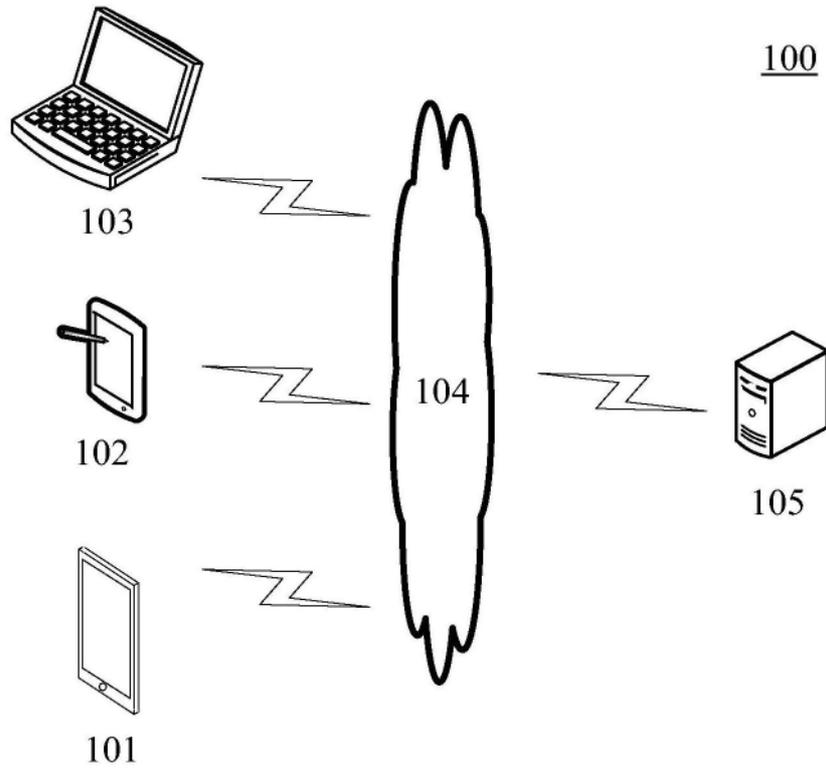


图1

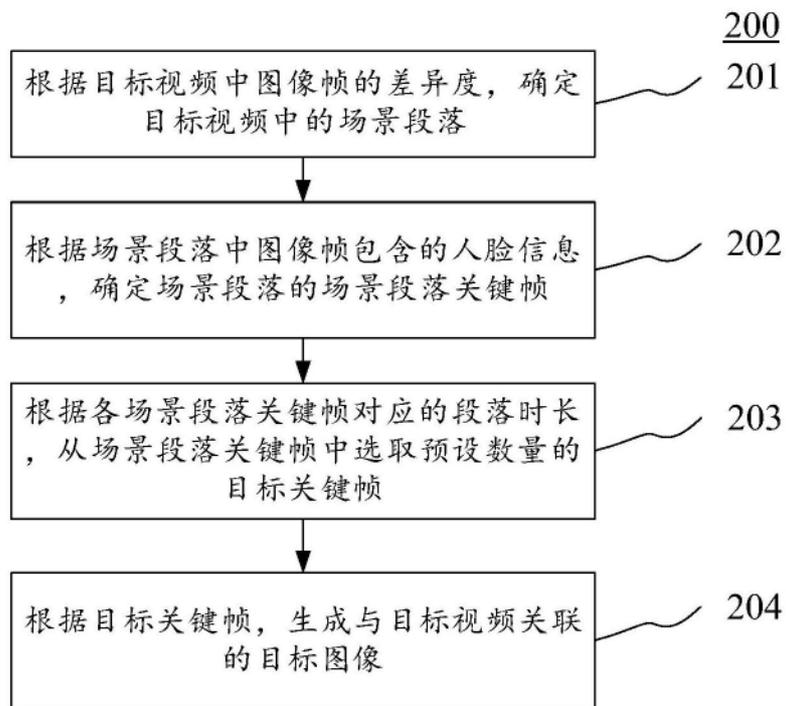


图2

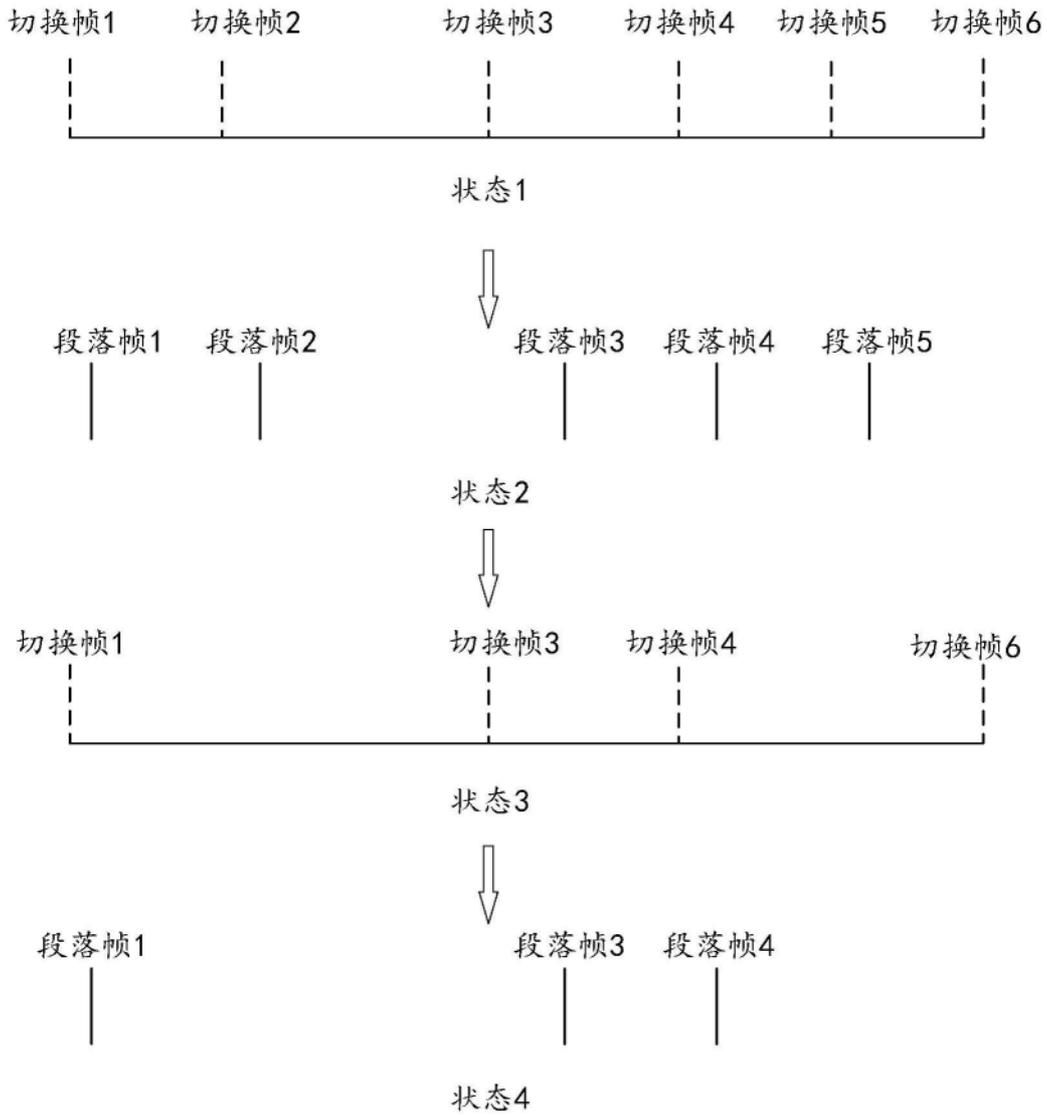


图3

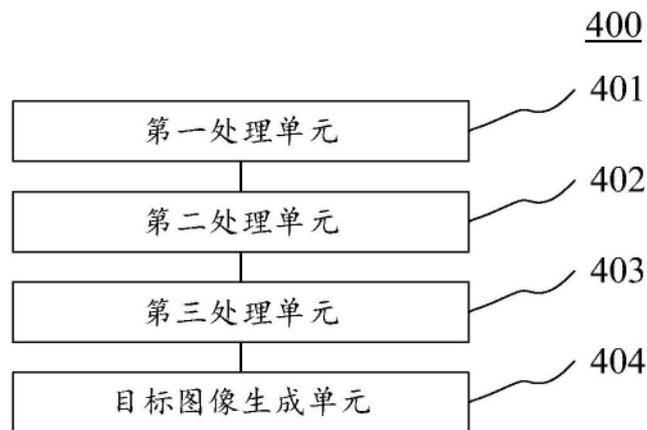


图4

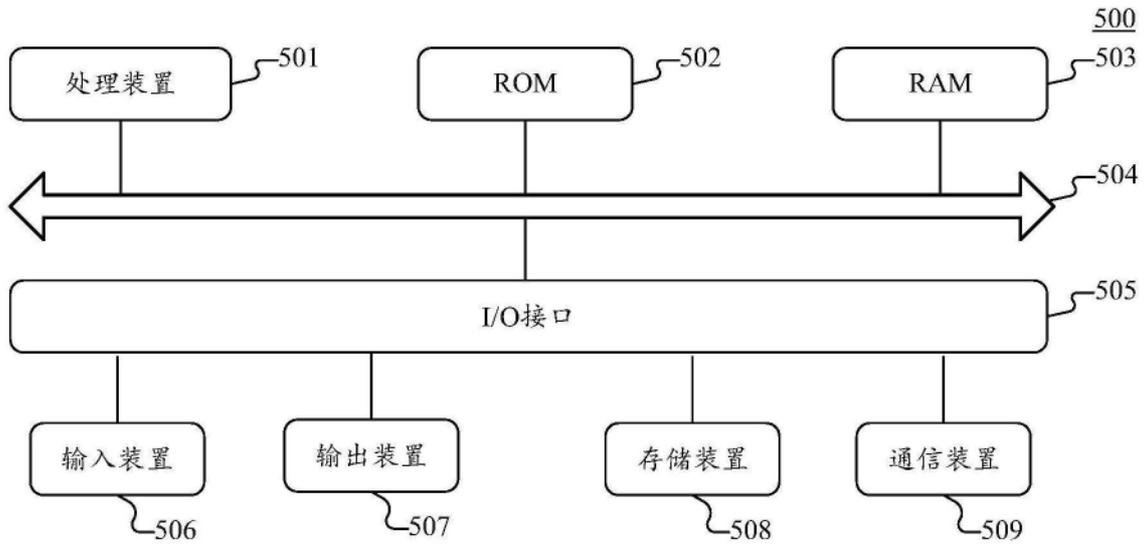


图5