



(12) 发明专利

(10) 授权公告号 CN 109996056 B

(45) 授权公告日 2021.03.26

(21) 申请号 201910381407.6

H04N 19/132 (2014.01)

(22) 申请日 2019.05.08

H04N 19/597 (2014.01)

(65) 同一申请的已公布的文献号
申请公布号 CN 109996056 A

(56) 对比文件

CN 108921942 A, 2018.11.30

CN 107067452 A, 2017.08.18

(43) 申请公布日 2019.07.09

CN 109191369 A, 2019.01.11

(73) 专利权人 北京奇艺世纪科技有限公司
地址 100080 北京市海淀区海淀北一街2号
鸿城拓展大厦10、11层

CN 105979244 A, 2016.09.28

WO 2018046964 A1, 2018.03.15

US 2018286061 A1, 2018.10.04

(72) 发明人 陈平 吴方印 杨东

审查员 王敏

(74) 专利代理机构 北京柏杉松知识产权代理事
务所(普通合伙) 11413

代理人 丁芸 马敬

(51) Int. Cl.

H04N 13/261 (2018.01)

H04N 13/161 (2018.01)

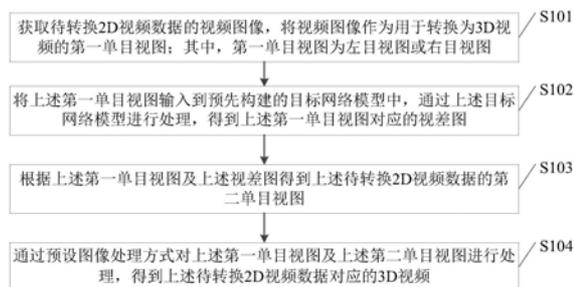
权利要求书5页 说明书22页 附图4页

(54) 发明名称

一种2D视频转3D视频的方法、装置及电子设备

(57) 摘要

本发明实施例提供了一种2D视频转3D视频的方法、装置及电子设备,其中,方法包括:获取待转换2D视频数据的视频图像,将所述视频图像作为用于转换为3D视频的第一单目视图,其中,所述第一单目视图为左目视图或右目视图;将所述第一单目视图输入到预先构建的目标网络模型中,通过所述目标网络模型进行处理,得到所述第一单目视图对应的视差图;根据所述第一单目视图及所述视差图得到所述待转换2D视频数据的第二单目视图;通过预设图像处理方式对所述第一单目视图及所述第二单目视图进行处理,得到所述待转换2D视频数据对应的3D视频,本发明实现了2D视频转换为3D视频。



1. 一种2D视频转3D视频的方法,其特征在于,所述方法包括:

获取待转换2D视频数据的视频图像,将所述视频图像作为用于转换为3D视频的第一单目视图;其中,所述第一单目视图为左目视图或右目视图;

将所述第一单目视图输入到预先构建的目标网络模型中,通过所述目标网络模型进行处理,得到所述第一单目视图对应的视差图;

根据所述第一单目视图及所述视差图得到所述待转换2D视频数据的第二单目视图;其中,所述第二单目视图为与所述第一单目视图对应的右目视图或左目视图;

通过预设图像处理方式对所述第一单目视图及所述第二单目视图进行处理,得到所述待转换2D视频数据对应的3D视频;

所述通过所述目标网络模型进行处理包括编码处理和解码处理;

通过如下步骤进行编码处理:将所述第一单目视图输入到预先构建的目标网络模型中,通过所述目标网络模型对所述第一单目视图进行多次下采样处理,并对每次下采样处理后的视图进行第一预设处理,依次得到尺寸由大到小的多个下采样图像;

通过如下步骤对所述下采样图像进行解码处理:对于所述多个下采样图像,按照尺寸由小到大的顺序进行排列;通过所述目标网络模型依次对排列后的各所述下采样图像进行上采样处理,并对每次上采样处理后的图像进行第二预设处理,得到所述第一单目视图对应的视差图;其中,所述视差图的尺寸与所述第一单目视图的尺寸相同;

当所述目标网络模型中包括第一级联上采样网络、第二级联上采样网络、第三级联上采样网络、第四级联上采样网络及第五级联上采样网络,且所述视差图为目标右视差图,所述第二预设处理为反卷积处理或双线性插值处理时,则通过如下步骤对所述下采样图像进行解码处理:

通过所述第一级联上采样网络对第五下采样图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第一上采样图像;其中,所述第一上采样图像的尺寸与第四下采样图像的尺寸相同;

对所述第一上采样图像进行卷积处理,得到第一右视差图;

将所述第一上采样图像、所述第一右视差图及所述第四下采样图像合并,得到第一目标图像;其中,所述第一目标图像的尺寸与所述第四下采样图像的尺寸相同;

通过所述第二级联上采样网络对所述第一目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第二上采样图像;其中,所述第二上采样图像的尺寸与第三下采样图像的尺寸相同;

对所述第二上采样图像进行卷积处理,得到第二右视差图;

将所述第二上采样图像、所述第二右视差图及所述第三下采样图像合并,得到第二目标图像;其中,所述第二目标图像的尺寸与所述第三下采样图像的尺寸相同;

通过所述第三级联上采样网络对所述第二目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第三上采样图像;其中,所述第三上采样图像的尺寸与第二下采样图像的尺寸相同;

对所述第三上采样图像进行卷积处理,得到第三右视差图;

将所述第三上采样图像、所述第三右视差图及所述第二下采样图像合并,得到第三目标图像;其中,所述第三目标图像的尺寸与所述第二下采样图像的尺寸相同;

通过所述第四级联上采样网络对所述第三目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第四上采样图像;其中,所述第四上采样图像的尺寸与第一下采样图像的尺寸相同;

对所述第四上采样图像进行卷积处理,得到第四右视差图;

将所述第四上采样图像、所述第四右视差图及所述第一下采样图像合并,得到第四目标图像;其中,所述第四目标图像的尺寸与所述第一下采样图像的尺寸相同;

通过所述第五级联上采样网络对所述第四目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第五上采样图像;其中,所述第五上采样图像的尺寸与所述第一单目视图的尺寸相同;

对所述第五上采样图像进行卷积处理,得到所述第一单目视图对应的目标右视差图。

2. 根据权利要求1所述的方法,其特征在于,当所述目标网络模型中包括第一级联下采样网络、第二级联下采样网络、第三级联下采样网络、第四级联下采样网络及第五级联下采样网络,且所述第一预设处理为卷积池化处理时,则通过如下步骤对所述第一单目视图进行编码处理:

将所述第一单目视图输入所述第一级联下采样网络,通过所述第一级联下采样网络对所述第一单目视图进行下采样及卷积池化处理,得到第一下采样图像;

通过所述第二级联下采样网络对所述第一下采样图像进行下采样及卷积池化处理,得到第二下采样图像;

通过所述第三级联下采样网络对所述第二下采样图像进行下采样及卷积池化处理,得到第三下采样图像;

通过所述第四级联下采样网络对所述第三下采样图像进行下采样及卷积池化处理,得到第四下采样图像;

通过所述第五级联下采样网络对所述第四下采样图像进行下采样及卷积池化处理,得到第五下采样图像。

3. 根据权利要求1所述的方法,其特征在于,预先构建并训练所述目标网络模型的过程,包括:

获取预设数量不同类型的3D片源作为训练样本;其中,每个训练样本包含第一单目视图及与所述第一单目视图对应的第二单目视图;

分别标定各第一单目视图对应的第二单目视图,得到所述训练样本中的各视图对;

分别将所述训练样本中的各所述视图对输入到预设网络模型中进行训练,得到所述目标网络模型。

4. 根据权利要求3所述的方法,其特征在于,所述分别将各所述视图对输入到预设网络模型中进行训练,得到所述目标网络模型,包括:

步骤A,在各所述视图对中选取一个视图对,将该视图对的第一单目视图输入到预设网络模型中,得到预测的左视差图和右视差图;

步骤B,根据已选取的视图对的第一单目视图、左视差图、第二单目视图及右视差图,计算预设损失函数的当前数值;

步骤C,按照所述预设损失函数的当前数值,调整所述预设网络模型中用于预测所述左视差图和右视差图各参数值;

步骤D,重复执行步骤A到步骤C,直至所述预设损失函数满足预设输出条件时得到所述目标网络模型。

5.根据权利要求1所述的方法,其特征在于,所述目标网络模型中还包括激活函数;所述方法还包括:获取拍摄所述3D片源的相机设备参数;根据所述相机设备参数,按照预设函数取值规则,确定所述激活函数的取值;

所述通过所述目标网络模型进行处理,得到所述第一单目视图对应的视差图,包括:

利用所述激活函数通过所述目标网络模型进行处理,得到所述第一单目视图对应的视差图;

其中,所述激活函数为双曲正切函数 \tanh ,所述激活函数 \tanh 的取值范围为 $[-1,1]$ 。

6.一种2D视频转3D视频的装置,其特征在于,所述装置包括:

第一单目视图获取模块,用于获取待转换2D视频数据的视频图像,将所述视频图像作为用于转换为3D视频的第一单目视图;其中,所述第一单目视图为左目视图或右目视图;

视差图确定模块,用于将所述第一单目视图输入到预先构建的目标网络模型中,通过所述目标网络模型进行处理,得到所述第一单目视图对应的视差图;

第二单目视图确定模块,用于根据所述第一单目视图及所述视差图得到所述待转换2D视频数据的第二单目视图;其中,所述第二单目视图为与所述第一单目视图对应的右目视图或左目视图;

3D视频确定模块,用于通过预设图像处理方式对所述第一单目视图及所述第二单目视图进行处理,得到所述待转换2D视频数据对应的3D视频;

所述通过所述目标网络模型进行处理包括编码处理和解码处理;所述视差图确定模块包括:

编码处理子模块,用于将所述第一单目视图输入到预先构建的目标网络模型中,通过所述目标网络模型对所述第一单目视图进行多次下采样处理,并对每次下采样处理后的视图进行第一预设处理,依次得到尺寸由大到小的多个下采样图像;

解码处理子模块,用于对于所述多个下采样图像,按照尺寸由小到大的顺序进行排列;通过所述目标网络模型依次对排列后的各所述下采样图像进行上采样处理,并对每次上采样处理后的图像进行第二预设处理,得到所述第一单目视图对应的视差图;其中,所述视差图的尺寸与所述第一单目视图的尺寸相同;

所述目标网络模型中包括第一级联上采样网络、第二级联上采样网络、第三级联上采样网络、第四级联上采样网络及第五级联上采样网络,所述视差图为目标右视差图;所述第二预设处理为预设处理;所述第二预设处理为反卷积处理或双线性插值处理时,所述解码处理子模块,包括:

第一上采样图像确定单元,用于通过所述第一级联上采样网络对第五下采样图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第一上采样图像;其中,所述第一上采样图像的尺寸与第四下采样图像的尺寸相同;

第一右视差图确定单元,用于对所述第一上采样图像进行卷积处理,得到第一右视差图;

第一目标图像确定单元,用于将所述第一上采样图像、所述第一右视差图及所述第四下采样图像合并,得到第一目标图像;其中,所述第一目标图像的尺寸与所述第四下采样图

像的尺寸相同；

第二上采样图像确定单元,用于通过所述第二级联上采样网络对所述第一目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第二上采样图像;其中,所述第二上采样图像的尺寸与第三下采样图像的尺寸相同;

第二右视差图确定单元,用于对所述第二上采样图像进行卷积处理,得到第二右视差图;

第二目标图像确定单元,用于将所述第二上采样图像、所述第二右视差图及所述第三下采样图像合并,得到第二目标图像;其中,所述第二目标图像的尺寸与所述第三下采样图像的尺寸相同;

第三上采样图像确定单元,用于通过所述第三级联上采样网络对所述第二目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第三上采样图像;其中,所述第三上采样图像的尺寸与第二下采样图像的尺寸相同;

第三右视差图确定单元,用于对所述第三上采样图像进行卷积处理,得到第三右视差图;

第三目标图像确定单元,用于将所述第三上采样图像、所述第三右视差图及所述第二下采样图像合并,得到第三目标图像;其中,所述第三目标图像的尺寸与所述第二下采样图像的尺寸相同;

第四上采样图像确定单元,用于通过所述第四级联上采样网络对所述第三目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第四上采样图像;其中,所述第四上采样图像的尺寸与第一下采样图像的尺寸相同;

第四右视差图确定单元,用于对所述第四上采样图像进行卷积处理,得到第四右视差图;

第四目标图像确定单元,用于将所述第四上采样图像、所述第四右视差图及所述第一下采样图像合并,得到第四目标图像;其中,所述第四目标图像的尺寸与所述第一下采样图像的尺寸相同;

第五上采样图像确定单元,用于通过所述第五级联上采样网络对所述第四目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第五上采样图像;其中,所述第五上采样图像的尺寸与所述第一单目视图的尺寸相同;

目标右视差图确定单元,用于对所述第五上采样图像进行卷积处理,得到所述第一单目视图对应的目标右视差图。

7. 根据权利要求6所述的装置,其特征在于,所述目标网络模型中包括第一级联下采样网络、第二级联下采样网络、第三级联下采样网络、第四级联下采样网络及第五级联下采样网络;且所述第一预设处理为卷积池化处理时,所述编码处理子模块,包括:

第一下采样图像确定单元,用于将所述第一单目视图输入所述第一级联下采样网络,通过所述第一级联下采样网络对所述第一单目视图进行下采样及卷积池化处理,得到第一下采样图像;

第二下采样图像确定单元,用于通过所述第二级联下采样网络对所述第一下采样图像进行下采样及卷积池化处理,得到第二下采样图像;

第三下采样图像确定单元,用于通过所述第三级联下采样网络对所述第二下采样图像

进行下采样及卷积池化处理,得到第三下采样图像;

第四下采样图像确定单元,用于通过所述第四级联下采样网络对所述第三下采样图像进行下采样及卷积池化处理,得到第四下采样图像;

第五下采样图像确定单元,用于通过所述第五级联下采样网络对所述第四下采样图像进行下采样及卷积池化处理,得到第五下采样图像。

8. 根据权利要求6所述的装置,其特征在于,所述装置还包括:

训练样本获取模块,用于获取预设数量不同类型的3D片源作为训练样本;其中,每个训练样本包含第一单目视图及与所述第一单目视图对应的第二单目视图;

视图对标定模块,用于分别标定各第一单目视图对应的第二单目视图,得到各视图对;

目标网络模型确定模块,用于分别将各所述视图对输入到预设网络模型中进行训练,得到所述目标网络模型。

9. 根据权利要求8所述的装置,其特征在于,所述目标网络模型确定模块,包括:

视差图确定单元,用于在各所述视图对中选取一个视图对,将该视图对的第一单目视图输入到预设网络模型中,得到预测的左视差图和右视差图;

损失函数值确定单元,用于根据已选取的视图对的第一单目视图、左视差图、第二单目视图及右视差图,计算预设损失函数的当前数值;

参数值调整单元,用于按照所述预设损失函数的当前数值,调整所述预设网络模型中用于预测所述左视差图和右视差图的各项参数;

目标网络模型确定单元,用于返回所述视差图确定子模块继续执行,直至所述预设损失函数满足预设输出条件时得到所述目标网络模型。

10. 根据权利要求6所述的装置,其特征在于,所述目标网络模型还包括激活函数,所述装置还包括:激活函数确定模块,用于获取拍摄所述3D片源的相机设备参数;根据所述相机设备参数,按照预设函数取值规则,确定所述激活函数的取值;

所述视差图确定模块,具体用于利用所述激活函数通过所述目标网络模型进行处理,得到所述第一单目视图对应的视差图;

其中,所述激活函数为双曲正切函数 \tanh ,所述激活函数 \tanh 的取值范围为 $[-1,1]$ 。

11. 一种电子设备,其特征在于,包括处理器、通信接口、存储器和通信总线,其中,所述处理器、所述通信接口、所述存储器通过所述通信总线完成相互间的通信;

所述存储器,用于存放计算机程序;

所述处理器,用于执行所述存储器上所存放的程序时,实现权利要求1-5任一所述的方法步骤。

一种2D视频转3D视频的方法、装置及电子设备

技术领域

[0001] 本发明涉及计算机技术领域,特别是一种2D视频转3D视频的方法、装置及电子设备。

背景技术

[0002] 视频源从视觉感官上分为2D和3D两种,2D视频由一张2D图片组成,3D视频由第一单目视图和第二单目视图通过图像处理软件形成。其中,第一单目视图为左眼单目看到的图像,第二单目视图为右眼单目看到的图像。因3D视频在成像时具有景深和立体感,故更符合人的主观视觉,深受人们喜爱。在某些场景下,3D视频的观影体验要明显优于2D视频,但是,3D视频主要来源于3D电影的大制作,其制作和拍摄极其困难和昂贵,因此除开这些3D电影,人们在日常观影时几乎无法体验3D视频的立体感。

[0003] 目前,2D视频到3D视频的转化技术主要通过人工实现,最经典的是泰坦尼克号2D版本到3D版本的制作,其制作不仅耗资巨大,而且还漫长,这明显不满足现实生活的需求。随着计算机视觉的发展,让计算机视觉实现2D视频到3D视频的转换变为一种可能。Deep3D通过3D电影源的视差估计,实现2D视频到3D视频的生成,但该方法最多只能估计出32个单位像素视差。可见,现有技术中,2D视频转3D视频的方法较为局限。因此,如何实现2D视频转换为3D视频仍然是亟待解决的技术问题。

发明内容

[0004] 本发明实施例的目的在于提供一种2D视频转3D视频的方法、装置及电子设备,以实现2D视频转换为3D视频。具体技术方案如下:

[0005] 第一方面,本发明实施例公开了一种2D视频转3D视频的方法,所述方法包括:

[0006] 获取待转换2D视频数据的视频图像,将所述视频图像作为用于转换为3D视频的第一单目视图;其中,所述第一单目视图为左目视图或右目视图;

[0007] 将所述第一单目视图输入到预先构建的目标网络模型中,通过所述目标网络模型进行处理,得到所述第一单目视图对应的视差图;

[0008] 根据所述第一单目视图及所述视差图得到所述待转换2D视频数据的第二单目视图;其中,所述第二单目视图为与所述第一单目视图对应的右目视图或左目视图;

[0009] 通过预设图像处理方式对所述第一单目视图及所述第二单目视图进行处理,得到所述待转换2D视频数据对应的3D视频。

[0010] 可选地,所述通过所述目标网络模型进行处理包括编码处理和解码处理;

[0011] 通过如下步骤进行编码处理:将所述第一单目视图输入到预先构建的目标网络模型中,通过所述目标网络模型对所述第一单目视图进行多次下采样处理,并对每次下采样处理后的视图进行第一预设处理,依次得到尺寸由大到小的多个下采样图像;

[0012] 通过如下步骤对所述下采样图像进行解码处理:对于所述多个下采样图像,按照尺寸由小到大的顺序进行排列;通过所述目标网络模型依次对排列后的各所述下采样图像

进行上采样处理,并对每次上采样处理后的图像进行第二预设处理,得到所述第一单目视图对应的视差图;其中,所述视差图的尺寸与所述第一单目视图的尺寸相同。

[0013] 可选地,当所述目标网络模型中包括第一级联下采样网络、第二级联下采样网络、第三级联下采样网络、第四级联下采样网络及第五级联下采样网络,且所述第一预设处理为卷积池化处理时,则通过如下步骤对所述第一单目视图进行编码处理:

[0014] 将所述第一单目视图输入所述第一级联下采样网络,通过所述第一级联下采样网络对所述第一单目视图进行下采样及卷积池化处理,得到第一下采样图像;

[0015] 通过所述第二级联下采样网络对所述第一下采样图像进行下采样及卷积池化处理,得到第二下采样图像;

[0016] 通过所述第三级联下采样网络对所述第二下采样图像进行下采样及卷积池化处理,得到第三下采样图像;

[0017] 通过所述第四级联下采样网络对所述第三下采样图像进行下采样及卷积池化处理,得到第四下采样图像;

[0018] 通过所述第五级联下采样网络对所述第四下采样图像进行下采样及卷积池化处理,得到第五下采样图像。

[0019] 可选地,当所述目标网络模型中包括第一级联上采样网络、第二级联上采样网络、第三级联上采样网络、第四级联上采样网络及第五级联上采样网络,且所述视差图为目标右视差图,所述第二预设处理为反卷积处理或双线性插值处理时,则通过如下步骤对所述下采样图像进行解码处理:

[0020] 通过所述第一级联上采样网络对所述第五下采样图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第一上采样图像;其中,所述第一上采样图像的尺寸与所述第四下采样图像的尺寸相同;

[0021] 对所述第一上采样图像进行卷积处理,得到第一右视差图;

[0022] 将所述第一上采样图像、所述第一右视差图及所述第四下采样图像合并,得到第一目标图像;其中,所述第一目标图像的尺寸与所述第四下采样图像的尺寸相同;

[0023] 通过所述第二级联上采样网络对所述第一目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第二上采样图像;其中,所述第二上采样图像的尺寸与所述第三下采样图像的尺寸相同;

[0024] 对所述第二上采样图像进行卷积处理,得到第二右视差图;

[0025] 将所述第二上采样图像、所述第二右视差图及所述第三下采样图像合并,得到第二目标图像;其中,所述第二目标图像的尺寸与所述第三下采样图像的尺寸相同;

[0026] 通过所述第三级联上采样网络对所述第二目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第三上采样图像;其中,所述第三上采样图像的尺寸与所述第二下采样图像的尺寸相同;

[0027] 对所述第三上采样图像进行卷积处理,得到第三右视差图;

[0028] 将所述第三上采样图像、所述第三右视差图及所述第二下采样图像合并,得到第三目标图像;其中,所述第三目标图像的尺寸与所述第三下采样图像的尺寸相同;

[0029] 通过所述第四级联上采样网络对所述第三目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第四上采样图像;其中,所述第

四上采样图像的尺寸与所述第一下采样图像的尺寸相同；

[0030] 对所述第四上采样图像进行卷积处理,得到第四右视差图；

[0031] 将所述第四上采样图像、所述第四右视差图及所述第一下采样图像合并,得到第四目标图像；其中,所述第四目标图像的尺寸与所述第一下采样图像的尺寸相同；

[0032] 通过所述第五级联上采样网络对所述第四目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第五上采样图像；其中,所述第五上采样图像的尺寸与所述第一单目视图的尺寸相同；

[0033] 对所述第五上采样图像进行卷积处理,得到所述第一单目视图对应的目标右视差图。

[0034] 可选地,预先构建并训练所述目标网络模型的过程,包括：

[0035] 获取预设数量不同类型的3D片源作为训练样本；其中,每个训练样本包含第一单目视图及与所述第一单目视图对应的第二单目视图；

[0036] 分别标定各第一单目视图对应的第二单目视图,得到所述训练样本中的各视图对；

[0037] 分别将所述训练样本中的各所述视图对输入到预设网络模型中进行训练,得到所述目标网络模型。

[0038] 可选地,所述分别将各所述视图对输入到预设网络模型中进行训练,得到所述目标网络模型,包括：

[0039] 步骤A,在各所述视图对中选取一个视图对,将该视图对的第一单目视图输入到预设网络模型中,得到预测的左视差图和右视差图；

[0040] 步骤B,根据已选取的视图对的第一单目视图、左视差图、第二单目视图及右视差图,计算预设损失函数的当前数值；

[0041] 步骤C,按照所述预设损失函数的当前数值,调整所述预设网络模型中用于预测所述左视差图和右视差图的各项参数；

[0042] 步骤D,重复执行步骤A到步骤C,直至所述预设损失函数满足预设输出条件时得到所述目标网络模型。

[0043] 可选地,所述目标网络模型中还包括激活函数；所述方法还包括：获取拍摄所述3D片源的相机设备参数；根据所述相机设备参数,按照预设函数取值规则,确定所述激活函数的取值；

[0044] 所述通过所述目标网络模型进行处理,得到所述第一单目视图对应的视差图,包括：

[0045] 利用所述激活函数通过所述目标网络模型进行处理,得到所述第一单目视图对应的视差图；

[0046] 其中,所述激活函数为双曲正切函数 \tanh ,所述激活函数 \tanh 的取值范围为 $[-1, 1]$ 。

[0047] 第二方面,本发明实施例还公开了一种2D视频转3D视频的装置,所述装置包括：

[0048] 第一单目视图获取模块,用于获取待转换2D视频数据的视频图像,将所述视频图像作为用于转换为3D视频的第一单目视图；其中,所述第一单目视图为左目视图或右目视图；

[0049] 视差图确定模块,用于将所述第一单目视图输入到预先构建的目标网络模型中,通过所述目标网络模型进行处理,得到所述第一单目视图对应的视差图;

[0050] 第二单目视图确定模块,用于根据所述第一单目视图及所述视差图得到所述待转换2D视频数据的第二单目视图;其中,所述第二单目视图为与所述第一单目视图对应的右目视图或左目视图;

[0051] 3D视频确定模块,用于通过预设图像处理方式对所述第一单目视图及所述第二单目视图进行处理,得到所述待转换2D视频数据对应的3D视频。

[0052] 可选地,所述通过所述目标网络模型进行处理包括编码处理和解码处理;所述视差图确定模块,包括:

[0053] 编码处理子模块,用于将所述第一单目视图输入到预先构建的目标网络模型中,通过所述目标网络模型对所述第一单目视图进行多次下采样处理,并对每次下采样处理后的视图进行第一预设处理,依次得到尺寸由大到小的多个下采样图像;

[0054] 解码处理子模块,用于对于所述多个下采样图像,按照尺寸由小到大的顺序进行排列;通过所述目标网络模型依次对排列后的各所述下采样图像进行上采样处理,并对每次上采样处理后的图像进行第二预设处理,得到所述第一单目视图对应的视差图;其中,所述视差图的尺寸与所述第一单目视图的尺寸相同。

[0055] 可选地,所述目标网络模型中包括第一级联下采样网络、第二级联下采样网络、第三级联下采样网络、第四级联下采样网络及第五级联下采样网络;且所述第一预设处理为卷积池化处理时,所述编码处理子模块,包括:

[0056] 第一下采样图像确定单元,用于将所述第一单目视图输入所述第一级联下采样网络,通过所述第一级联下采样网络对所述第一单目视图进行下采样及卷积池化处理,得到第一下采样图像;

[0057] 第二下采样图像确定单元,用于通过所述第二级联下采样网络对所述第一下采样图像进行下采样及卷积池化处理,得到第二下采样图像;

[0058] 第三下采样图像确定单元,用于通过所述第三级联下采样网络对所述第二下采样图像进行下采样及卷积池化处理,得到第三下采样图像;

[0059] 第四下采样图像确定单元,用于通过所述第四级联下采样网络对所述第三下采样图像进行下采样及卷积池化处理,得到第四下采样图像;

[0060] 第五下采样图像确定单元,用于通过所述第五级联下采样网络对所述第四下采样图像进行下采样及卷积池化处理,得到第五下采样图像。

[0061] 可选地,所述目标网络模型中包括第一级联上采样网络、第二级联上采样网络、第三级联上采样网络、第四级联上采样网络及第五级联上采样网络,所述视差图为目标右视差图;所述第二预设处理为预设处理;所述第二预设处理为反卷积处理或双线性插值处理时,所述解码处理子模块,包括:

[0062] 第一上采样图像确定单元,用于通过所述第一级联上采样网络对所述第五下采样图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第一上采样图像;其中,所述第一上采样图像的尺寸与所述第四下采样图像的尺寸相同;

[0063] 第一右视差图确定单元,用于对所述第一上采样图像进行卷积处理,得到第一右

视差图；

[0064] 第一目标图像确定单元,用于将所述第一上采样图像、所述第一右视差图及所述第四下采样图像合并,得到第一目标图像;其中,所述第一目标图像的尺寸与所述第四下采样图像的尺寸相同;

[0065] 第二上采样图像确定单元,用于通过所述第二级联上采样网络对所述第一目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第二上采样图像;其中,所述第二上采样图像的尺寸与所述第三下采样图像的尺寸相同;

[0066] 第二右视差图确定单元,用于对所述第二上采样图像进行卷积处理,得到第二右视差图;

[0067] 第二目标图像确定单元,用于将所述第二上采样图像、所述第二右视差图及所述第三下采样图像合并,得到第二目标图像;其中,所述第二目标图像的尺寸与所述第三下采样图像的尺寸相同;

[0068] 第三上采样图像确定单元,用于通过所述第三级联上采样网络对所述第二目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第三上采样图像;其中,所述第三上采样图像的尺寸与所述第二下采样图像的尺寸相同;

[0069] 第三右视差图确定单元,用于对所述第三上采样图像进行卷积处理,得到第三右视差图;

[0070] 第三目标图像确定单元,用于将所述第三上采样图像、所述第三右视差图及所述第二下采样图像合并,得到第三目标图像;其中,所述第三目标图像的尺寸与所述第三下采样图像的尺寸相同;

[0071] 第四上采样图像确定单元,用于通过所述第四级联上采样网络对所述第三目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第四上采样图像;其中,所述第四上采样图像的尺寸与所述第一下采样图像的尺寸相同;

[0072] 第四右视差图确定单元,用于对所述第四上采样图像进行卷积处理,得到第四右视差图;

[0073] 第四目标图像确定单元,用于将所述第四上采样图像、所述第四右视差图及所述第一下采样图像合并,得到第四目标图像;其中,所述第四目标图像的尺寸与所述第一下采样图像的尺寸相同;

[0074] 第五上采样图像确定单元,用于通过所述第五级联上采样网络对所述第四目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第五上采样图像;其中,所述第五上采样图像的尺寸与所述第一单目视图的尺寸相同;

[0075] 目标右视差图确定单元,用于对所述第五上采样图像进行卷积处理,得到所述第一单目视图对应的目标右视差图。

[0076] 可选地,所述装置还包括:

[0077] 训练样本获取模块,用于获取预设数量不同类型的3D片源作为训练样本;其中,每个训练样本包含第一单目视图及与所述第一单目视图对应的第二单目视图;

[0078] 视图对标定模块,用于分别标定各第一单目视图对应的第二单目视图,得到各视图对;

[0079] 目标网络模型确定模块,用于分别将各所述视图对输入到预设网络模型中进行训

练,得到所述目标网络模型。

[0080] 可选地,所述目标网络模型确定模块,包括:

[0081] 视差图确定单元,用于在各所述视图对中选取一个视图对,将该视图对的第一单目视图输入到预设网络模型中,得到预测的左视差图和右视差图;

[0082] 损失函数值确定单元,用于根据已选取的视图对的第一单目视图、左视差图、第二单目视图及右视差图,计算预设损失函数的当前数值;

[0083] 参数值调整单元,用于按照所述预设损失函数的当前数值,调整所述预设网络模型中用于预测所述左视差图和右视差图的各项参数值;

[0084] 目标网络模型确定单元,用于返回所述视差图确定子模块继续执行,直至所述预设损失函数满足预设输出条件时得到所述目标网络模型。

[0085] 可选地,所述目标网络模型还包括激活函数,所述装置还包括:激活函数确定模块,用于获取拍摄所述3D片源的相机设备参数;根据所述相机设备参数,按照预设函数取值规则,确定所述激活函数的取值;

[0086] 所述视差图确定模块,具体用于利用所述激活函数通过所述目标网络模型进行处理,得到所述第一单目视图对应的视差图;

[0087] 其中,所述激活函数为双曲正切函数 \tanh ,所述激活函数 \tanh 的取值范围为 $[-1, 1]$ 。

[0088] 又一方面,本发明实施例还公开了一种电子设备,包括处理器、通信接口、存储器和通信总线,其中,所述处理器、所述通信接口、所述存储器通过所述通信总线完成相互间的通信;

[0089] 所述存储器,用于存放计算机程序;

[0090] 所述处理器,用于执行所述存储器上所存放的程序时,实现上述2D视频转3D视频的方法中任一所述的方法步骤。

[0091] 又一方面,本发明实施例还公开了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内存储有计算机程序,所述计算机程序被处理器执行时,实现上述2D视频转3D视频的方法中任一所述的方法步骤。

[0092] 又一方面,本发明实施例还公开了一种包含指令的计算机程序产品,当其在计算机上运行时,使得计算机执行上述实现上述2D视频转3D视频的方法中任一所述的方法步骤。

[0093] 本发明实施例公开了一种2D视频转3D视频的方法、装置及电子设备,以实现2D视频转换为3D视频。具体为,将第一单目视图输入到预先构建的目标网络模型中,通过目标网络模型进行处理,得到第一单目视图对应的视差图;根据第一单目视图及视差图得到所述待转换2D视频数据的第二单目视图;通过预设图像处理方式对第一单目视图及所述第二单目视图进行处理,得到待转换2D视频数据对应的3D视频。本发明实施例中通过预先构建的目标网络模型,能够对输入的第一单目视图进行处理,自动输出第一单目视图对应的视差图,使得通过第一单目视图和视差图得到第二单目视图,通过第一单目视图及第二单目视图可得到3D视频。通过本发明实施例,能够对任意场景的第一单目视图得到对应的第二单目视图,实现了方便有效的对不同场景的2D视频转换为3D视频的效果,提高了本发明实施例的实用性。

附图说明

[0094] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍。

[0095] 图1为本发明实施例的一种2D视频转3D视频的方法流程图;

[0096] 图2为本发明实施例的一种2D视频转3D视频的方法中确定视差图的方法流程图;

[0097] 图3为本发明实施例的一种2D视频转3D视频的方法中编码部分流程图;

[0098] 图4为本发明实施例的一种2D视频转3D视频的方法中解码部分流程图;

[0099] 图5为本发明实施例的一种装置2D视频转3D视频的装置结构图;

[0100] 图6为本发明实施例的一种电子设备结构示意图。

具体实施方式

[0101] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行描述。

[0102] 为达到上述发明目的,本发明实施例公开了一种2D视频转3D视频的方法,如图1所示。图1为本发明实施例的一种2D视频转3D视频的方法流程图,方法包括:

[0103] S101,获取待转换2D视频数据的视频图像,将视频图像作为用于转换为3D视频的第一单目视图;其中,第一单目视图为左目视图或右目视图。

[0104] 本发明实施例的2D视频转3D视频的方法,可以通过电子设备实现。具体的,该电子设备可以为服务器、个人电脑或智能手机等设备。

[0105] 本步骤中获取待转换2D视频帧的视频图像,将该视频图像作为该2D视频帧的第一单目视图。该第一单目视图可为使用现有2D视频介质进行单目图像采集的视频图像,具体可为,左眼单目看到的图像或右眼单目看到的图像。本申请中将此时获得的单目图像作为第一单目视图,后面通过目标网络模型得到的视差图确定的图像作为第二单目视图。

[0106] 例如,本步骤中获取的第一单目视图为左目视图。

[0107] S102,将上述第一单目视图输入到预先构建的目标网络模型中,通过上述目标网络模型进行处理,得到上述第一单目视图对应的视差图。

[0108] 本发明实施例中预先构建目标网络模型,该目标网络模型能够将输入的第一单目视图通过特征提取、卷积池化处理,得到待转换2D视频数据的第一单目视图对应的视差图。该处实现的网络结构为:在VGG (Visual Geometry Group,视觉几何组)的基础上实现U型神经网络模型U-Net的编解码网络。本发明实施例中当获取的第一单目视图为左目视图时,可通过目标网络模型得到与该左目视图对应的右目视图;当获取的第一单目视图为右目视图时,可通过目标网络模型得到与该右目视图对应的左目视图。即为训练时,可将左目视图作为模型训练的输入,得到预测右目视图的视差图;也可将右目视图作为模型训练的输入,得到预测左目视图的视差图。以下实施例中详细说明该目标网络模型的构建过程。

[0109] 可选地,通过目标网络模型进行处理包括编码处理和解码处理;S102中将上述第一单目视图输入到预先构建的目标网络模型中,通过上述目标网络模型进行处理,得到上述第一单目视图对应的视差图,可参见图2。图2为本发明实施例的一种2D视频转3D视频的方法中确定视差图的方法流程图,包括:

[0110] S201,通过如下步骤进行编码处理:将第一单目视图输入到预先构建的目标网络模型中,通过目标网络模型对第一单目视图进行多次下采样处理,并对每次下采样处理后

的视图进行第一预设处理,依次得到尺寸由大到小的多个下采样图像。

[0111] 本发明实施例的目标网络模型处理包括第一部分的下采样的编码处理,及第二部分的上采样的解码处理。通过编码及解码处理对第一单目视图进行多层深度学习,得到视差图。

[0112] 本步骤中,将第一单目视图输入到预先构建的目标网络模型中,通过编码处理对第一单目视图进行多次降低尺寸的下采样,并对每次下采样处理后的视图进行第一预设处理。本步骤中,上述第一预设处理可以为多种处理方式,即本步骤的目的为了得到预设尺寸的下采样图像,只要能够实现以上目的图像处理方式均在本实施例的范围内,举例来说:在本发明实施例中,该第一预设处理为增加通道的卷积处理,以及池化处理。

[0113] 对图片进行降低尺寸的下采样处理阶段数可根据实施人员的具体要求设置,保证能够从最小尺寸的下采样图像提取到满足要求的特征即可。下采样次数可以按照实际情况进行设定,例如设定为三次、四次或五次等。

[0114] 在一种可能的实施方式中,目标网络模型的编码处理每次下采样过程可将前一次下采样得到的图片尺寸缩小一半,通过卷积操作增加下采样图片的通道。具体可以包括:

[0115] 步骤a,将第一单目视图进行降低尺寸的下采样处理,得到初始下采样图像,其中,初始下采样图像的尺寸为第一单目视图尺寸的一半(初始下采样图像的宽为第一单目视图宽的一半,且初始下采样图像的高为第一单目视图高的一半)。

[0116] 步骤b,将初始下采样图像进行卷积处理,增加该初始下采样图像的通道数,并对图像进行池化处理,得到目标下采样图像。

[0117] 步骤c,将目标下采样图像作为初始下采样图像,返回上述步骤a继续执行,直至达到设定的次数。

[0118] 下面以三个阶段的下采样为例进行举例说明:将尺寸为 $512*256*3$ 的第一单目视图输入到目标网络模型中,其中,512表示第一单目视图的宽,256表示第一单目视图的高,3表示该第一单目视图的通道数。将 $512*256*3$ 的第一单目视图经过第一次下采样卷积池化处理,得到 $256*128*32$ 的特征图;将 $256*128*32$ 的特征图经过第二次的下采样卷积池化处理,得到 $128*64*64$ 的特征图;将 $128*64*64$ 的特征图经过第三次的下采样卷积池化处理,得到 $64*32*128$ 的特征图。

[0119] S202,通过如下步骤对下采样图像进行解码处理:对于多个下采样图像,按照尺寸由小到大的顺序进行排列;通过目标网络模型依次对排列后的各下采样图像进行上采样处理,并对每次上采样处理后的图像进行第二预设处理,得到第一单目视图对应的视差图;其中,视差图的尺寸与第一单目视图的尺寸相同。本步骤中,上述第二预设处理可以为多种图像处理方式,例如:反卷积处理或者双线性插值处理。

[0120] 本发明实施例中为了让特征图同时具有图像深层信息和图像浅层信息,解码过程中在将图像增加到与编码过程的尺寸相同的图片后,将相同尺寸的编码图像与该解码图像进行合并,得到该尺寸的特征图。具体为,从最小尺寸的下采样图像开始上采样处理,并对将上采样后的特征图进行反卷积或者双线性插值,使得图像的尺寸变为原来图像尺寸的两倍,直至最后输出第一单目视图对应的视差图。对图片进行增加尺寸的上采样处理阶段数与对图片进行降低尺寸的下采样处理阶段数对应。本步骤中,解码处理按照尺寸由小到大的顺序,依次对各下采样图像进行上采样及双线性插值处理,得到第一单目视图对应的视

差图。具体可以包括：

[0121] 步骤1,解码部分按照尺寸由小到大的顺序,选取尺寸最小的下采样图像进行上采样及双线性插值处理,得到图像尺寸增加一半(宽及高均增加一半)、的初始上采样图像。

[0122] 步骤2,对初始上采样图像进行卷积处理得到初始上采样图像的右视差图;

[0123] 步骤3,选取尺寸与初始上采样图像尺寸相同的下采样图像,合并该下采样图像、初始上采样图像及初始上采样图像的右视差图,得到目标图像,其中,目标图像的尺寸与初始上采样图像尺寸相同。

[0124] 步骤4,将目标图像进行上采样及双线性插值处理,得到图像尺寸增加一半的目标上采样图像。

[0125] 步骤5,将目标上采样图像作为初始上采样图像,返回上述步骤2继续执行,直至得到尺寸与第一单目视图尺寸相同的目标上采样图像。

[0126] 步骤6,对尺寸与第一单目视图尺寸相同的目标上采样图像进行卷积处理,得到右视差图,即第一单目视图对应的视差图。

[0127] S103,根据上述第一单目视图及上述视差图得到上述待转换2D视频数据的第二单目视图。

[0128] 通过任意相关的视图转换算法,按照第一单目视图及视差图得到第二单目视图,即待转换2D视频数据的第二单目视图。

[0129] 可选的,在上述根据上述第一单目视图及上述视差图得到上述待转换2D视频数据的第二单目视图之后,上述方法还包括:

[0130] 通过超分辨率卷积神经网络对上述第二单目视图中的所有像素或部分像素进行像素值修正。

[0131] 具体为,将上述第二单目视图输入到超分辨率卷积神经网络中,通过该超分辨率卷积神经网络卷积处理,修正上述第二单目视图中的像素值,经过若干次卷积处理后,得到较高分辨率的第二单目视图。

[0132] 该修正可为将第二单目视图中的所有像素值进行修正,或部分像素值进行修正。

[0133] 具体为,从修正的整体考虑,在训练过程中训练模型对输入图像每个部分都做相同权重比例像素的修正处理,使得最终得到对目标网络模型得到的第二单目视图全部像素值进行修正。

[0134] 另外,从修正的效率考虑,鉴于第一单目视图与第二单目视图的视场范围会存在一些差距,在第二单目视图中一些可见区域在第一单目视图中不可见,因此上述目标网络模型得到的第二单目视图中,对于在第二单目视图中可见在第一区域不可见的区域预测会存在误差。因此,在训练过程中训练模型对输入图像的第二单目视图中可见在第一区域不可见的区域增加权重比例,对该部分重点修正处理,使得最终得到对目标网络模型得到的第二单目视图部分像素值进行修正。

[0135] 本申请的超分辨率卷积神经网络训练过程可为,获取预设数量的3D片源作为训练样本;其中,每个训练样本包含第一单目视图及与第一单目视图对应的第二单目视图;将上述第一单目视图输入到目标网络模型,最终得到第一预测第二单目视图。将该第一预测第二单目视图输入到超分辨率卷积神经网络中,通过超分辨率卷积神经网络进行处理,输出与真实第二单目视图尺寸相同的第二预测第二单目视图,将第二预测第二单目视图与真实

的第二单目视图对比,调整超分辨率卷积神经网络的参数,可设置训练次数,每次输出第二预测第二单目视图时调整网络参数,直至得到训练次数时得到训练好的超分辨率卷积神经网络。进而通过该训练好的超分辨率卷积神经网络对上述目标网络模型进行修正。

[0136] 需要说明的是,本发明对目标网络模型得到的第二单目视图进行修正的模型不做限定,在实施过程中也可使用GAN (Generative Adversarial Network,生成对抗网络)对目标网络模型得到的第二单目视图进行修正。

[0137] 通过超分辨率卷积神经网络可对得到的第二单目视图中低分辨率的像素点进行卷积操作,得到高分辨率的第二单目视图。

[0138] S104,通过预设图像处理方式对上述第一单目视图及上述第二单目视图进行处理,得到上述待转换2D视频数据对应的3D视频。

[0139] 当获取的第一单目视图为左目视图时,得到的第二单目视图即为与该左目视图对应的右目视图;当获取的第一单目视图为右目视图时,得到的第二单目视图即为与该右目视图对应的左目视图。将该第一单目视图和第二单目视图分别作为左、右目视图,通过现有的3D设备观看该第一单目视图及第二单目视图,得到待转换2D视频数据对应的3D视频。或者通过现有对第一单目视图及第二单目视图进行处理得到3D视频的方式,处理该第一单目视图及第二单目视图,得到待转换2D视频数据对应的3D视频。本发明实施例中具体不做限定。

[0140] 在本发明实施例公开的一种2D视频转3D视频的方法中,将第一单目视图输入到预先构建的目标网络模型中,通过目标网络模型进行处理,得到第一单目视图对应的视差图;根据第一单目视图及视差图得到待转换2D视频数据的第二单目视图;通过预设图像处理方式对第一单目视图及第二单目视图进行处理,得到待转换2D视频数据对应的3D视频。本发明实施例中通过预先构建的目标网络模型,能够对输入的第一单目视图进行处理,自动输出第一单目视图对应的视差图,使得通过第一单目视图和视差图得到第二单目视图,通过第一单目视图及第二单目视图可得到3D视频。通过本发明实施例,能够对任意场景的第一单目视图得到对应的第二单目视图,实现了方便有效的对不同场景的2D视频转换为3D视频的效果,提高了本发明实施例的实用性。

[0141] 可选地,在本发明2D视频转3D视频的方法的一种实施例中,目标网络模型中还包括激活函数;方法还包括:获取拍摄3D片源的相机设备参数;根据相机设备参数,按照预设函数取值规则,确定激活函数的取值。

[0142] 上述通过目标网络模型进行处理,得到第一单目视图对应的视差图,包括:

[0143] 利用激活函数通过目标网络模型进行处理,得到第一单目视图对应的视差图;

[0144] 其中,激活函数为双曲正切函数 \tanh ,激活函数 \tanh 的取值范围为 $[-1, 1]$ 。

[0145] 在拍摄3D片源得到第一单目视图时,可采用会聚相机获得单目图像,本实施例中可获取会聚相机的参数,根据会聚相机的参数,选取能够使得视差图中视差小于图片宽度的激活函数,作为调整视差图中视差范围的调整函数。因此本申请中选取双曲正切函数 \tanh 作为激活函数。其中,激活函数 \tanh 的取值范围为 $[-1, 1]$ 。

[0146] 这是因为 \tanh 出来的值在 ± 1 之间,3D电影拍摄使用会聚相机,这样会导致真实的视差图中的会出现有正有负的数值。所以使用 \tanh 对进行合理修正,最终采用 \tanh 激活后的特征图就是本申请的视差图,并且使得视差图中视差范围在 $[-0.5W, 0.5W]$,其中,W为原

始左目视图图片的宽。

[0147] 通过本发明实施例采用的激活函数能够有效对视差图进行修正,使得得到的视差图中的视差能够控制在合理范围内。

[0148] 可选地,在本发明2D视频转3D视频的方法的一种实施例中,上述目标网络模型包括第一级联下采样网络、第二级联下采样网络、第三级联下采样网络、第四级联下采样网络及第五级联下采样网络;且第一预设处理为卷积池化处理时,可通过如下步骤对第一单目视图进行编码处理,可参见图3所示的方法流程图。图3为本发明实施例的一种2D视频转3D视频的方法中编码部分流程图,包括:

[0149] S301,将上述第一单目视图输入上述第一级联下采样网络,通过上述第一级联下采样网络对上述第一单目视图进行下采样及卷积池化处理,得到第一下采样图像。

[0150] 本发明实施例的编码部分包括多个级联下采样网络,每个级联下采样网络包括多个卷积层的卷积处理过程及池化处理过程。前一个级联下采样网络的输出作为下一个级联下采样网络的输入。本发明实施例的编码处理可包括第一级联下采样网络、第二级联下采样网络、第三级联下采样网络、第四级联下采样网络及第五级联下采样网络。

[0151] 本步骤中,将第一单目视图输入第一级联下采样网络,通过第一级联下采样网络包含的多个卷积层及池化层对第一单目视图进行下采样及卷积池化处理。该处实现的网络结构为:在VGG (Visual Geometry Group,视觉几何组)的基础上实现U-Net的编解码网络。该第一级联下采样网络包括多个卷积层和池化层max_pooing,通过多个卷积层对输入的第一单目视图进行卷积处理,通过池化层对第一单目视图进行下采样。

[0152] 例如,将尺寸为 $265*512*3$ 的第一单目视图输入到第一级联下采样网络中,其中,265可表示第一单目视图的宽;512可表示第一单目视图的高;3可表示该第一单目视图的通道数。通过第一级联下采样网络对第一单目视图进行下采样及增加通道的卷积池化处理,得到 $128*256*32$ 的第一下采样图像。

[0153] S302,通过上述第二级联下采样网络对上述第一下采样图像进行下采样及卷积池化处理,得到第二下采样图像。

[0154] 例如,将 $128*256*32$ 的第一下采样图像输入到第二级联下采样网络中,通过第二级联下采样网络对该第一下采样图像进行下采样处理、增加通道的卷积处理及池化处理,得到 $64*128*64$ 的第二下采样图像。

[0155] S303,通过上述第三级联下采样网络对上述第二下采样图像进行下采样及卷积池化处理,得到第三下采样图像。

[0156] 例如,将 $64*128*64$ 的第二下采样图像输入到第三级联下采样网络中,通过第三级联下采样网络对该第二下采样图像进行下采样处理、增加通道的卷积处理及池化处理,得到 $32*64*128$ 的第三下采样图像。

[0157] S304,通过上述第四级联下采样网络对上述第三下采样图像进行下采样及卷积池化处理,得到第四下采样图像。

[0158] 例如,将 $32*64*128$ 的第三下采样图像输入到第四级联下采样网络中,通过第四级联下采样网络对该第三下采样图像进行下采样处理、增加通道的卷积处理及池化处理,得到 $16*32*256$ 的第四下采样图像。

[0159] S305,通过上述第五级联下采样网络对上述第四下采样图像进行下采样及卷积池

化处理,得到第五下采样图像。

[0160] 例如,将16*32*256的第四下采样图像输入到第五级联下采样网络中,通过第五级联下采样网络对该第四下采样图像进行下采样处理、增加通道的卷积处理及池化处理,得到8*16*512的第五下采样图像。

[0161] 可参见表1所示的本发明实施例的一种2D视频转3D视频的方法中编码部分的编码表。

[0162] 表1

Input	Shape	Ops	Output
image	[256,512,3]	convs&max_pooling	pool1
pool1	[128,256,32]	convs&max_pooling	pool2
pool2	[64,128,64]	convs&max_pooling	pool3
pool3	[32,64,128]	convs&max_pooling	pool4
pool4	[16,32,256]	convs&max_pooling	pool5
pool5	[8,16,512]	--	--

[0164] 其中,Input表示输入,Shape表示输入图像的尺寸及通道特征,Ops表示级联下采样网络的操作,Output表示对应该级联下采样网络处理后的输出结果,image表示输入的第一单目视图;pool1表示第一下采样图像;pool2表示第二下采样图像;pool3表示第三下采样图像;pool4表示第四下采样图像;pool5表示第五下采样图像。

[0165] 表2

Name	Encoder Name
pool1	encoder1
pool2	encoder2
pool3	encoder3
pool4	encoder4
pool5	encoder5

[0167] 可将pool1表示的第一下采样图像,作为第一级联下采样网络对应的编码器encoder1的输出特征图;可将pool2表示的第二下采样图像,作为第二级联下采样网络对应的编码器encoder2的输出特征图;可将pool3表示的第三下采样图像,作为第三级联下采样网络对应的编码器encoder3的输出特征图;可将pool4表示的第四下采样图像,作为第四级联下采样网络对应的编码器encoder4的输出特征图;可将pool5表示的第五下采样图像,作为第五级联下采样网络对应的编码器encoder5的输出特征图。

[0168] 需要说明的是,本发明实施例中设置了五个级联上采样网络,在实际实施过程中,可按照实施人员的具体需求设置多于五个或少于五个的级联上采样网络。

[0169] 可见,通过本发明实施例,可实现对第一单目视图进行多极化的下采样处理,得到不同尺寸的特征图,便于后期在解码过程中,对相同尺寸的深层特征图及浅层特征图进行合并处理,使得最后得到的第二单目视图效果更佳。

[0170] 可选地,在本发明2D视频转3D视频的方法的一种实施例中,上述目标网络模型包括第一级联上采样网络、第二级联上采样网络、第三级联上采样网络、第四级联上采样网络

及第五级联上采样网络;且视差图为目标右视差图,第二预设处理为反卷积处理或双线性插值处理时,则通过如下步骤对下采样图像进行解码处理,可见参见图4所示的方法流程图。图4为本发明实施例的一种2D视频转3D视频的方法中解码部分流程图,包括:

[0171] S401,通过上述第一级联上采样网络对上述第五下采样图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第一上采样图像;其中,上述第一上采样图像的尺寸与上述第四下采样图像的尺寸相同。

[0172] 本发明实施例的解码部分对应上述实施例的编码部分,上述实施例的解码部分通过五个级联上采样网络对输入的第一单目视图进行下采样的编码,本实施例通过五个级联上采样网络对编码部分最后的编码特征图进行解码,得到与第一单目视图尺寸对应的第二单目视图。

[0173] 本步骤中,通过第一级联上采样网络对第五下采样图像进行上采样,并对上采样后的图像进行反卷积或双线程差值处理,得到第一上采样图像,该第一上采样图像的尺寸与第四下采样图像的尺寸相同。

[0174] 例如,第一级联上采样网络对 $8*16*512$ 的第五下采样图像进行上采样及双线程差值的卷积池化处理,得到 $16*32*256$ 的第一上采样图像。

[0175] S402,对上述第一上采样图像进行卷积处理,得到第一右视差图。

[0176] 例如,可以通过第一级联上采样网络中的卷积神经网络,对第一上采样图像进行卷积处理,得到第一右视差图。

[0177] S403,将上述第一上采样图像、上述第一右视差图及上述第四下采样图像合并,得到第一目标图像;其中,上述第一目标图像的尺寸与上述第四下采样图像的尺寸相同。

[0178] 可以合并第一上采样图像及第一右视差图,得到第一上采样图像对应的预测第二单目视图,其中,第一上采样图像对应的预测第二单目视图与第一上采样图像的尺寸相同。第四下采样图像与第一上采样图像的尺寸相同,合并第一上采样图像对应的预测第二单目视图与第四下采样图像,得到第一目标图像。

[0179] S404,通过上述第二级联上采样网络对上述第一目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第二上采样图像;其中,上述第二上采样图像的尺寸与上述第三下采样图像的尺寸相同。

[0180] 通过第二级联上采样网络对第一目标图像进行上采样,并对上采样后的图像进行反卷积或双线性插值处理,得到第二上采样图像,该第二上采样图像的尺寸与第三下采样图像的尺寸相同。

[0181] S405,对上述第二上采样图像进行卷积处理,得到第二右视差图。

[0182] 例如,可以通过第二级联上采样网络中的卷积神经网络,对第二上采样图像进行卷积处理,得到第二右视差图。在一种可能的实施方式中,可以通过第二级联上采样网络中的两层卷积神经网络,对第二上采样图像进行卷积处理,得到第二右视差图。

[0183] S406,将上述第二上采样图像、上述第二右视差图及上述第三下采样图像合并,得到第二目标图像;其中,上述第二目标图像的尺寸与上述第三下采样图像的尺寸相同。

[0184] 可以合并第二上采样图像及第二右视差图,得到第二上采样图像对应的预测第二单目视图,其中,第二上采样图像对应的预测第二单目视图与第二上采样图像的尺寸相同。第三下采样图像与第二上采样图像的尺寸相同,合并第二上采样图像对应的预测第二单目

视图与第三下采样图像,得到第二目标图像。

[0185] S407,通过上述第三级联上采样网络对上述第二目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第三上采样图像;其中,上述第三上采样图像的尺寸与上述第二下采样图像的尺寸相同。

[0186] 通过第三级联上采样网络对第二目标图像进行上采样,并对上采样后的图像进行反卷积或双线性插值处理,得到第三上采样图像,该第三上采样图像的尺寸与第二下采样图像的尺寸相同。

[0187] S408,对上述第三上采样图像进行卷积处理,得到第三右视差图。

[0188] 例如,可以通过第三级联上采样网络中的卷积神经网络,对第三上采样图像进行卷积处理,得到第三右视差图。在一种可能的实施方式中,可以通过第三级联上采样网络中的两层卷积神经网络,对第三上采样图像进行卷积处理,得到第三右视差图。

[0189] S409,将上述第三上采样图像、上述第三右视差图及上述第二下采样图像合并,得到第三目标图像;其中,上述第三目标图像的尺寸与上述第二下采样图像的尺寸相同。

[0190] 可以合并第三上采样图像及第三右视差图,得到第三上采样图像对应的预测第二单目视图,其中,第三上采样图像对应的预测第二单目视图与第三上采样图像的尺寸相同。第二下采样图像与第三上采样图像的尺寸相同,合并第三上采样图像对应的预测第二单目视图与第二下采样图像,得到第三目标图像。

[0191] S410,通过上述第四级联上采样网络对上述第三目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第四上采样图像;其中,上述第四上采样图像的尺寸与上述第一下采样图像的尺寸相同。

[0192] 通过第四级联上采样网络对第三目标图像进行上采样,并对上采样后的图像进行反卷积或双线性插值处理,得到第四上采样图像,该第四上采样图像的尺寸与第一下采样图像的尺寸相同。

[0193] S411,对上述第四上采样图像进行卷积处理,得到第四右视差图。

[0194] 例如,可以通过第四级联上采样网络中的卷积神经网络,对第四上采样图像进行卷积处理,得到第四右视差图。在一种可能的实施方式中,可以通过第四级联上采样网络中的两层卷积神经网络,对第四上采样图像进行卷积处理,得到第四右视差图。

[0195] S412,将上述第四上采样图像、上述第四右视差图及上述第一下采样图像合并,得到第四目标图像;其中,上述第四目标图像的尺寸与上述第一下采样图像的尺寸相同。

[0196] 可以合并第四上采样图像及第四右视差图,得到第四上采样图像对应的预测第二单目视图,其中,第四上采样图像对应的预测第二单目视图与第四上采样图像的尺寸相同。第一下采样图像与第四上采样图像的尺寸相同,合并第四上采样图像对应的预测第二单目视图与第一下采样图像,得到第四目标图像。

[0197] S413,通过上述第五级联上采样网络对上述第四目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第五上采样图像;其中,上述第五上采样图像的尺寸与上述第一单目视图的尺寸相同。

[0198] 通过第五级联上采样网络对第四目标图像进行上采样,并对上采样后的图像进行反卷积或双线性插值处理,得到第五上采样图像,该第五上采样图像的尺寸与第一单目视图的尺寸相同。

[0199] S414,对上述第五上采样图像进行卷积处理,得到上述第一单目视图对应的目标右视差图。

[0200] 例如,可以通过第五级联上采样网络中的卷积神经网络,对第五上采样图像进行卷积处理,得到右视差图,即第一单目视图对应的视差图。在一种可能的实施方式中,可以通过第五级联上采样网络中的两层卷积神经网络,对第五上采样图像进行卷积处理,得到右视差图。

[0201] 可参见表3所示的本发明实施例的一种2D视频转3D视频的方法中解码部分的解码表。

[0202] 表3

Input	Ops	Output	
		Name	Shape
encoder5	convs&deconvs	decoder1	[16,32,256]
decoder1	convs & tanh	disps1	[16,32,2]
encoder4& decoder1& disps1	convs&deconvs	decoder2	[32,64,128]
decoder2	convs & tanh	disps2	[32,64,2]
[0203] encoder3& decoder2& disps2	convs&deconvs	decoder3	[64,128,64]
decoder2	convs & tanh	disps3	[64,128,2]
encoder2& decoder3& disps3	convs&deconvs	decoder4	[128,256,32]
decoder4	convs & tanh	disps4	[128,256,2]
encoder1& decoder4& disps4	convs&deconvs	decoder5	[256,512,16]
decoder5	convs & tanh	disps5	[256,512,2]

[0204] 其中,Input表示输入,Ops表示级联上采样网络的操作,其中,convs表示卷积操作,deconvs表示反卷积操作;tanh表示级联上采样网络的激活函数;Output表示对应该级联上采样网络处理后的输出结果;shape表示从该级联上采样网络得到的特征图的尺寸,disps1表示第一级联上采样网络的输出的第一右视差图;disps2表示第二级联上采样网络的输出的第二右视差图;disps3表示第三级联上采样网络的输出的第三右视差图;disps4表示第四级联上采样网络输出的第四右视差图;disps5表示第五级联上采样网络的输出的第一单目视图对应的目标右视差图。每个阶段的解码是先进行上采样及降低通道的convs与增大尺寸的deconvs;在进行两次卷积convs以及采用激活函数激活tanh得到视差图,

encoder&decoder&disps表示将本次级联上采样网络处理得到的上采样图像、与本次级联上采样网络的上采样图像尺寸相同的下采样图像、及本次级联上采样网络得到的视差图进行合并,作为下一级联上采样网络的输入。

[0205] 在本发明实施例中,通过五个级联上采样网络对编码后的图像进行解码,最终得到第一单目视图对应的视差图,能够增加得到的视差图的准确度。

[0206] 在本发明另一实施例中,还具体包括一种预先构建并训练上述目标网络模型的过程,具体为:

[0207] 步骤一,获取预设数量不同类型的3D片源作为训练样本;其中,每个训练样本包含第一单目视图及与第一单目视图对应的第二单目视图。

[0208] 获取预设数量不同类型的3D片源作为训练样本,并使用现有3D视频介质进行单目图像采集第一单目视图及与第一单目视图对应的第二单目视图。其中,第一单目视图可为左眼单目看到的图像;第二单目视图可为右眼单目看到的图像,要求第一单目视图和第二单目视图需要相互对应。本发明实施例将第一单目视图作为采集训练的样本,第二单目视图作为通过第一单目视图得到的预测图像的对比样本。

[0209] 可以获取人工挑选的各类场景下的视频源作为训练数据,并将这些场景分为科幻、陆地战争、海洋、星空、动画风格、室内场景、自然场景、室外场景、其他等9类。在实际训练时可以采用例如OHEM(Online Hard Example Mining,在线难例挖掘算法)、Focal Loss(焦点损失)等算法来解决场景类别不平衡的问题。

[0210] 步骤二,分别标定各第一单目视图对应的第二单目视图,得到各视图对。

[0211] 将同一2D视频图像采集到的第一单目视图及第二单目视图,标记为一个视图对。

[0212] 步骤三,分别将各上述视图对输入到预设网络模型中进行训练,得到上述目标网络模型。

[0213] 在一种可能的实施方式中,上述分别将各上述视图对输入到预设网络模型中进行训练,得到预先构建的目标网络模型,包括:

[0214] 步骤A,在各视图对中选取一个视图对,将该视图对的第一单目视图输入到预设网络模型中,得到预测的左视差图和右视差图;

[0215] 通过各视图对训练预设网络模型,从而得到目标网络模型。具体可以为将第一单目视图作为训练样本,输出预测的左视差图及右视差图。

[0216] 本发明实施例中,对级联上采样网络进行上采样及预设处理后得到的上采样图像,进行第一次卷积操作得到的图像为左视差图,第二次卷积操作得到的视差图为右视差图。即为预设网络模型的最后一层级联上采样网络对上一层级联上采样网络输入的特征图进行处理,得到与输入的视图对的第一单目视图尺寸相同的上采样图像后,在对该上采样图像进行第一次卷积处理,可得到该视图对的左视差图;对该上采样图像进行第二次卷积处理,可得到该视图对的右视差图。

[0217] 通过视差图与单目视图的对应关系,得到预测的单目视图。

[0218] 其中,第一单目视图+右视差图=第二单目视图,第二单目视图+左视差图=第一单目视图。

[0219] 步骤B,根据各已选取的视图对的第一单目视图、左视差图、第二单目视图及右视差图,计算预设损失函数的当前数值。

[0220] 本申请的预设损失函数包括梯度损失函数、感知损失函数及SSIM+L1 (structural similarity, 结构相似性, L1表示预测第二单目视图与真实第二单目视图的绝对值误差)。

[0221] 其中, 梯度损失函数的表达式为:

$$[0222] \quad G_{ds}^l = \frac{1}{N} \sum_{i,j} |\partial_x d_{ij}^l| e^{-\|\partial_x I_{ij}^l\|} + |\partial_y d_{ij}^l| e^{-\|\partial_y I_{ij}^l\|}$$

[0223] 其中, G_{ds}^l 表示梯度函数; N 表示该图片中像素点个数; i 表示任一像素点的横坐标; j 表示该像素点的纵坐标; l 表示第一单目视图; e 表示无理常数; ∂_x 表示图像在x方向的一阶导数; I_{ij}^l 表示预测第二单目视图; d_{ij}^l 表示坐标在 (i, j) 处的视差值; ∂_y 图像在y方向的一阶导数。

[0224] 其中, 感知损失函数的表达式为:

$$[0225] \quad G_j^\phi(x)_{c,c'} = \frac{1}{C_j H_j W_j} \sum_{h=1}^{H_j} \sum_{w=1}^{W_j} \phi_j(x)_{h,w,c} \phi_j(x)_{h,w,c'}$$

$$[0226] \quad G_{style}^{\phi,j}(\hat{y}, y)_{c,c'} = \left\| G_j^\phi(\hat{y}) - G_j^\phi(y) \right\|_F^2$$

[0227] 其中, $G_{style}^{\phi,j}(\hat{y}, y)_{c,c'}$ 表示预测第二单目视图和真实第二单目视图的感知函数; $G_j^\phi(x)_{c,c'}$ 表示矩阵 c 与之转置后的矩阵 c' 的乘积函数; j 表示第 j 张图片; C_j 表示第 j 张图片的通道数; H_j 表示第 j 张图片的高; W_j 表示第 j 张图片的宽; $\phi_j(x)_{h,w,c}$ 表示第 i 张特征图, h 表示图像的高, w 表示图像的宽; c 表示图像的通道; $\phi_j(x)_{h,w,c'}$ 表示第 i 张特征图的转置; \hat{y} 表示预测第二单目视图; y 表示真实第二单目视图; $G_j^\phi(\hat{y})$ 表示预测第二单目视图的感知值; $G_j^\phi(y)$ 表示真实第二单目视图的感知值; $\|\cdot\|_F^2$ 表示2范数, 与点到点的欧氏距离的计算方法一致。

[0228] SSIM+L1的表达式为:

$$[0229] \quad G_{ap}^l = \frac{1}{N} \sum_{i,j} \alpha \frac{1 - SSIM(I_{ij}^l, \tilde{I}_{ij}^l)}{2} + (1 - \alpha) \|I_{ij}^l - \tilde{I}_{ij}^l\|$$

[0230] 其中, G_{ap}^l 表示SSIM+L1整体的损失函数; N 表示图片个数, l 表示第一单目视图; α 表示SSIM损失的权重占比, 取值为0.85; I_{ij}^l 表示真实第二单目视图; \tilde{I}_{ij}^l 表示预测第二单目视图; $SSIM(I_{ij}^l, \tilde{I}_{ij}^l)$ 表示预测第二单目视图与真实第二单目视图的结构相似性; $\|I_{ij}^l - \tilde{I}_{ij}^l\|$ 表示预测第二单目视图与真实第二单目视图的绝对值误差L1。

[0231] 将当前视图对中的第一单目视图、左视差图、第二单目视图及右视差图及已输入

到模型中训练过的样本的第一单目视图、左视差图、第二单目视图及右视差图带入上述公式中,分别计算每种损失函数的当前数值,并求和得到最后的损失值。

[0232] 步骤C,按照预设损失函数的当前数值,调整预设网络模型中用于预测左视差图和右视差图的各项参数值。

[0233] 本发明实施例中,通过设置的损失函数,来反向调节预设网络模型中确定视差图的各项参数值,各参数值包括该预设网络模型各层卷积函数的参数值,池化函数的参数值、双线性插值的参数值或反卷积函数的参数值。

[0234] 具体的,梯度损失函数和生成的视差图的比例设置可以通过多次单一变量的实验得出,设置比例从0.01为起点,10为终点,每次比例设置为上次比例设置的5倍。

[0235] 步骤D,重复执行步骤A到步骤C,直至预设损失函数满足预设输出条件时得到目标网络模型。

[0236] 本发明实施例的预设输出条件可为迭代次数,即为调整完参数后,后一次在前一次调整的基础上训练,使得参数回传调整了预设次数或者训练了预设次数后,可得到目标网络模型。以使通过目标网络模型得到预测的视差图,通过该视差图得到的第二单目视图趋近于真实的第二单目视图。

[0237] 第二方面,本发明实施例公开了一种2D视频转3D视频的装置,如图5所示。图5为本发明实施例的一种装置2D视频转3D视频的装置结构图,包括:

[0238] 第一单目视图确定模块501,用于获取待转换2D视频数据的视频图像,将视频图像作为用于转换为3D视频的第一单目视图;其中,第一单目视图为左目视图或右目视图;

[0239] 视差图确定模块502,用于将第一单目视图输入到预先构建的目标网络模型中,通过目标网络模型进行处理,得到第一单目视图对应的视差图;

[0240] 第二单目视图确定模块503,用于根据第一单目视图及视差图得到待转换2D视频数据的第二单目视图;其中,第二单目视图为与第一单目视图对应的右目视图或左目视图;

[0241] 3D视频确定模块504,用于通过预设图像处理方式对第一单目视图及第二单目视图进行处理,得到待转换2D视频数据对应的3D视频。

[0242] 在本发明实施例公开的一种2D视频转3D视频的装置中,将第一单目视图输入到预先构建的目标网络模型中,通过目标网络模型进行处理,得到第一单目视图对应的视差图;根据第一单目视图及视差图得到待转换2D视频数据的第二单目视图;通过预设图像处理方式对第一单目视图及第二单目视图进行处理,得到待转换2D视频数据对应的3D视频。本发明实施例中通过预先构建的目标网络模型,能够对输入的第一单目视图进行处理,自动输出第一单目视图对应的视差图,使得通过第一单目视图和视差图得到第二单目视图,通过第一单目视图及第二单目视图可得到3D视频。通过本发明实施例,能够对任意场景的第一单目视图得到对应的第二单目视图,实现了方便有效的对不同场景的2D视频转换为3D视频的效果,提高了本发明实施例的实用性。

[0243] 可选地,在本发明2D视频转3D视频的装置的一种实施例中,通过目标网络模型进行处理包括编码处理和解码处理;视差图确定模块502,包括:

[0244] 编码处理子模块,用于将第一单目视图输入到预先构建的目标网络模型中,通过目标网络模型对第一单目视图进行多次下采样处理,并对每次下采样处理后的视图进行第一预设处理,依次得到尺寸由大到小的多个下采样图像;

[0245] 解码处理子模块,用于对于多个下采样图像,按照尺寸由小到大的顺序进行排列;通过目标网络模型依次对排列后的各下采样图像进行上采样处理,并对每次上采样处理后的图像进行第二预设处理,得到第一单目视图对应的视差图;其中,视差图的尺寸与第一单目视图的尺寸相同。

[0246] 可选地,在本发明2D视频转3D视频的装置的一种实施例中,目标网络模型中包括第一级联下采样网络、第二级联下采样网络、第三级联下采样网络、第四级联下采样网络及第五级联下采样网络;且第一预设处理为卷积池化处理时,编码处理子模块,包括:

[0247] 第一下采样图像确定单元,用于将第一单目视图输入第一级联下采样网络,通过第一级联下采样网络对第一单目视图进行下采样及卷积池化处理,得到第一下采样图像;

[0248] 第二下采样图像确定单元,用于通过第二级联下采样网络对第一下采样图像进行下采样及卷积池化处理,得到第二下采样图像;

[0249] 第三下采样图像确定单元,用于通过第三级联下采样网络对第二下采样图像进行下采样及卷积池化处理,得到第三下采样图像;

[0250] 第四下采样图像确定单元,用于通过第四级联下采样网络对第三下采样图像进行下采样及卷积池化处理,得到第四下采样图像;

[0251] 第五下采样图像确定单元,用于通过第五级联下采样网络对第四下采样图像进行下采样及卷积池化处理,得到第五下采样图像。

[0252] 可选地,在本发明2D视频转3D视频的装置的一种实施例中,目标网络模型中包括第一级联上采样网络、第二级联上采样网络、第三级联上采样网络、第四级联上采样网络及第五级联上采样网络,视差图为目标右视差图;第二预设处理为预设处理;第二预设处理为反卷积处理或双线性插值处理时,解码处理子模块,包括:

[0253] 第一上采样图像确定单元,用于通过第一级联上采样网络对第五下采样图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第一上采样图像;其中,第一上采样图像的尺寸与第四下采样图像的尺寸相同;

[0254] 第一右视差图确定单元,用于对第一上采样图像进行卷积处理,得到第一右视差图;

[0255] 第一目标图像确定单元,用于将第一上采样图像、第一右视差图及第四下采样图像合并,得到第一目标图像;其中,第一目标图像的尺寸与第四下采样图像的尺寸相同;

[0256] 第二上采样图像确定单元,用于通过第二级联上采样网络对第一目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第二上采样图像;其中,第二上采样图像的尺寸与第三下采样图像的尺寸相同;

[0257] 第二右视差图确定单元,用于对第二上采样图像进行卷积处理,得到第二右视差图;

[0258] 第二目标图像确定单元,用于将第二上采样图像、第二右视差图及第三下采样图像合并,得到第二目标图像;其中,第二目标图像的尺寸与第三下采样图像的尺寸相同;

[0259] 第三上采样图像确定单元,用于通过第三级联上采样网络对第二目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第三上采样图像;其中,第三上采样图像的尺寸与第二下采样图像的尺寸相同;

[0260] 第三右视差图确定单元,用于对第三上采样图像进行卷积处理,得到第三右视差

图；

[0261] 第三目标图像确定单元,用于将第三上采样图像、第三右视差图及第二下采样图像合并,得到第三目标图像;其中,第三目标图像的尺寸与第三下采样图像的尺寸相同;

[0262] 第四上采样图像确定单元,用于通过第四级联上采样网络对第三目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第四上采样图像;其中,第四上采样图像的尺寸与第一下采样图像的尺寸相同;

[0263] 第四右视差图确定单元,用于对第四上采样图像进行卷积处理,得到第四右视差图;

[0264] 第四目标图像确定单元,用于将第四上采样图像、第四右视差图及第一下采样图像合并,得到第四目标图像;其中,第四目标图像的尺寸与第一下采样图像的尺寸相同;

[0265] 第五上采样图像确定单元,用于通过第五级联上采样网络对第四目标图像进行上采样处理,并对本次上采样处理后的图像进行反卷积处理或双线性插值处理,得到第五上采样图像;其中,第五上采样图像的尺寸与第一单目视图的尺寸相同;

[0266] 目标右视差图确定单元,用于对第五上采样图像进行卷积处理,得到第一单目视图对应的目标右视差图。

[0267] 可选地,在本发明2D视频转3D视频的装置的一种实施例中,装置还包括:

[0268] 训练样本获取模块,用于获取预设数量不同类型的3D片源作为训练样本;其中,每个训练样本包含第一单目视图及与第一单目视图对应的第二单目视图;

[0269] 视图对标定模块,用于分别标定各第一单目视图对应的第二单目视图,得到各视图对;

[0270] 目标网络模型确定模块,用于分别将各视图对输入到预设网络模型中进行训练,得到目标网络模型。

[0271] 可选地,在本发明2D视频转3D视频的装置的一种实施例中,目标网络模型确定模块,包括:

[0272] 视差图确定单元,用于在各视图对中选取一个视图对,将该视图对的第一单目视图输入到预设网络模型中,得到预测的左视差图和右视差图;

[0273] 损失函数值确定单元,用于根据已选取的视图对的第一单目视图、左视差图、第二单目视图及右视差图,计算预设损失函数的当前数值;

[0274] 参数值调整单元,用于按照预设损失函数的当前数值,调整预设网络模型中用于预测左视差图和右视差图的各参数值;

[0275] 目标网络模型确定单元,用于返回视差图确定子模块继续执行,直至预设损失函数满足预设输出条件时得到目标网络模型。

[0276] 可选地,目标网络模型还包括激活函数,装置还包括:激活函数确定模块,用于获取拍摄3D片源的相机设备参数;根据相机设备参数,按照预设函数取值规则,确定激活函数的取值;

[0277] 视差图确定模块,具体用于利用激活函数通过目标网络模型进行处理,得到第一单目视图对应的视差图;

[0278] 其中,激活函数为双曲正切函数 \tanh ,激活函数 \tanh 的取值范围为 $[-1, 1]$ 。

[0279] 为达到上述发明目的,本发明实施例还公开了一种电子设备,如图6所示。图6为本

发明实施例的一种电子设备结构示意图,包括处理器601、通信接口602、存储器603和通信总线604,其中,处理器601、通信接口602、存储器603通过通信总线604完成相互间的通信;

[0280] 存储器603,用于存放计算机程序;

[0281] 处理器601,用于执行存储器603上所存放的程序时,实现上述2D视频转3D视频的方法中任一的方法步骤。

[0282] 上述电子设备提到的通信总线604可以是外设部件互连标准(Peripheral Component Interconnect,PCI)总线或扩展工业标准结构(Extended Industry Standard Architecture,EISA)总线等。该通信总线604可以分为地址总线、数据总线、控制总线等。为便于表示,图中仅用一条粗线表示,但并不表示仅有一根总线或一种类型的总线。

[0283] 通信接口602用于上述电子设备与其他设备之间的通信。

[0284] 存储器603可以包括随机存取存储器(Random Access Memory,RAM),也可以包括非易失性存储器(Non-Volatile Memory,NVM),例如至少一个磁盘存储器。可选的,存储器603还可以是至少一个位于远离前述处理器601的存储装置。

[0285] 上述的处理器601可以是通用处理器,包括中央处理器(Central Processing Unit,CPU)、网络处理器(Network Processor,NP)等;还可以是数字信号处理器(Digital Signal Processing,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现场可编程门阵列(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。

[0286] 在本发明实施例公开了一种电子设备中,将第一单目视图输入到预先构建的目标网络模型中,通过目标网络模型进行处理,得到第一单目视图对应的视差图;根据第一单目视图及视差图得到待转换2D视频数据的第二单目视图;通过预设图像处理方式对第一单目视图及第二单目视图进行处理,得到待转换2D视频数据对应的3D视频。本发明实施例中通过预先构建的目标网络模型,能够对输入的第一单目视图进行处理,自动输出第一单目视图对应的视差图,使得通过第一单目视图和视差图得到第二单目视图,通过第一单目视图及第二单目视图可得到3D视频。通过本发明实施例,能够对任意场景的第一单目视图得到对应的第二单目视图,实现了方便有效的对不同场景的2D视频转换为3D视频的效果,提高了本发明实施例的实用性。

[0287] 又一方面,本发明实施例还公开了一种计算机可读存储介质,所述计算机可读存储介质内存储有计算机程序,所述计算机程序被处理器执行时,实现上述2D视频转3D视频的方法中任一所述的方法步骤。

[0288] 在本发明实施例的一种计算机可读存储介质中,将第一单目视图输入到预先构建的目标网络模型中,通过目标网络模型进行处理,得到第一单目视图对应的视差图;根据第一单目视图及视差图得到所述待转换2D视频数据的第二单目视图;通过预设图像处理方式对第一单目视图及所述第二单目视图进行处理,得到待转换2D视频数据对应的3D视频。本发明实施例中通过预先构建的目标网络模型,能够对输入的第一单目视图进行处理,自动输出第一单目视图对应的视差图,使得通过第一单目视图和视差图得到第二单目视图,通过第一单目视图及第二单目视图可得到3D视频。通过本发明实施例,能够对任意场景的第一单目视图得到对应的第二单目视图,实现了方便有效的对不同场景的2D视频转换为3D视频的效果,提高了本发明实施例的实用性。

[0289] 又一方面,本发明实施例还公开了一种包含指令的计算机程序产品,当其在计算机上运行时,使得计算机执行上述实现上述2D视频转3D视频的方法中任一所述的方法步骤。

[0290] 在本发明实施例的一种包含指令的计算机程序产品中,将第一单目视图输入到预先构建的目标网络模型中,通过目标网络模型进行处理,得到第一单目视图对应的视差图;根据第一单目视图及视差图得到所述待转换2D视频数据的第二单目视图;通过预设图像处理方式对第一单目视图及所述第二单目视图进行处理,得到待转换2D视频数据对应的3D视频。本发明实施例中通过预先构建的目标网络模型,能够对输入的第一单目视图进行处理,自动输出第一单目视图对应的视差图,使得通过第一单目视图和视差图得到第二单目视图,通过第一单目视图及第二单目视图可得到3D视频。通过本发明实施例,能够对任意场景的第一单目视图得到对应的第二单目视图,实现了方便有效的对不同场景的2D视频转换为3D视频的效果,提高了本发明实施例的实用性。

[0291] 在上述实施例中,可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。在计算机上加载和执行计算机程序指令时,全部或部分地产生按照本发明实施例的流程或功能。计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。计算机指令可以存储在计算机可读存储介质中,或者从一个计算机可读存储介质向另一个计算机可读存储介质传输,例如,计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线(DSL))或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、服务器或数据中心进行传输。计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。可用介质可以是磁性介质,(例如,软盘、硬盘、磁带)、光介质(例如,DVD)、或者半导体介质(例如固态硬盘Solid State Disk(SSD))等。

[0292] 需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0293] 本说明书中的各个实施例均采用相关的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于装置、电子设备实施例而言,由于其基本相似于方法实施例,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0294] 以上仅为本发明的较佳实施例而已,并非用于限定本发明的保护范围。凡在本发明的精神和原则之内所作的任何修改、等同替换、改进等,均包含在本发明的保护范围内。

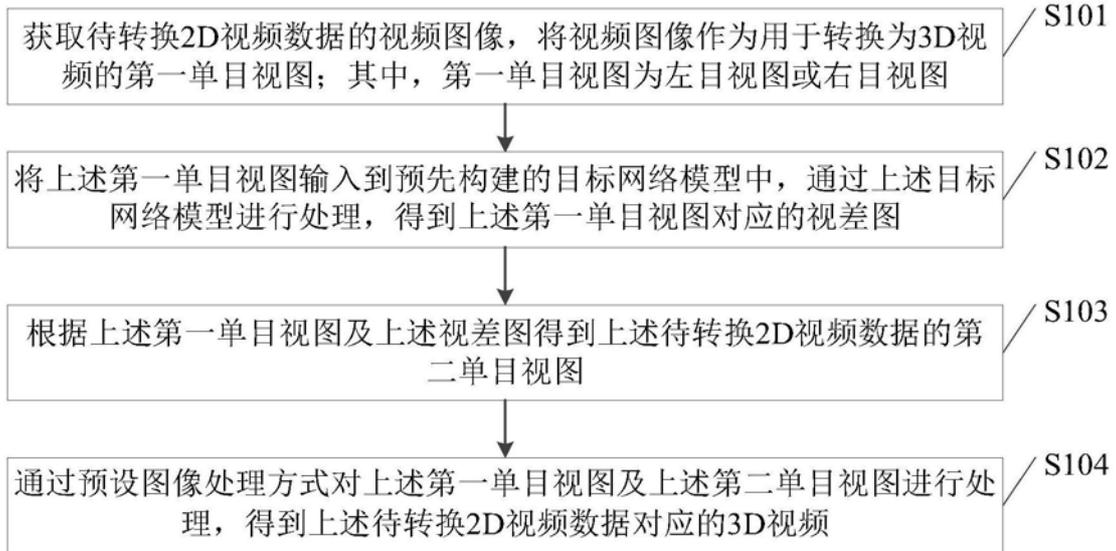


图1

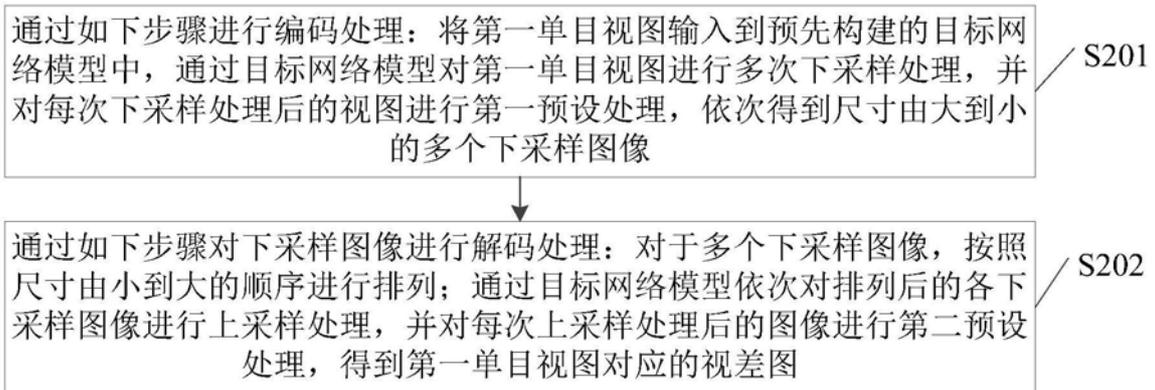


图2

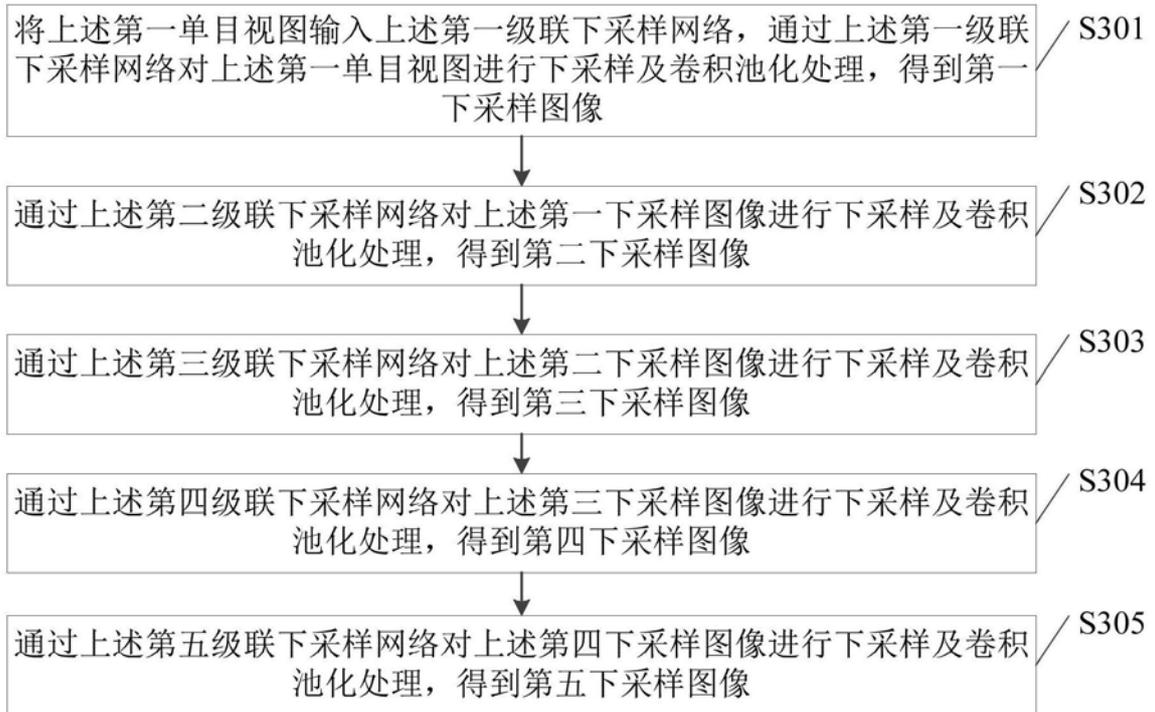


图3

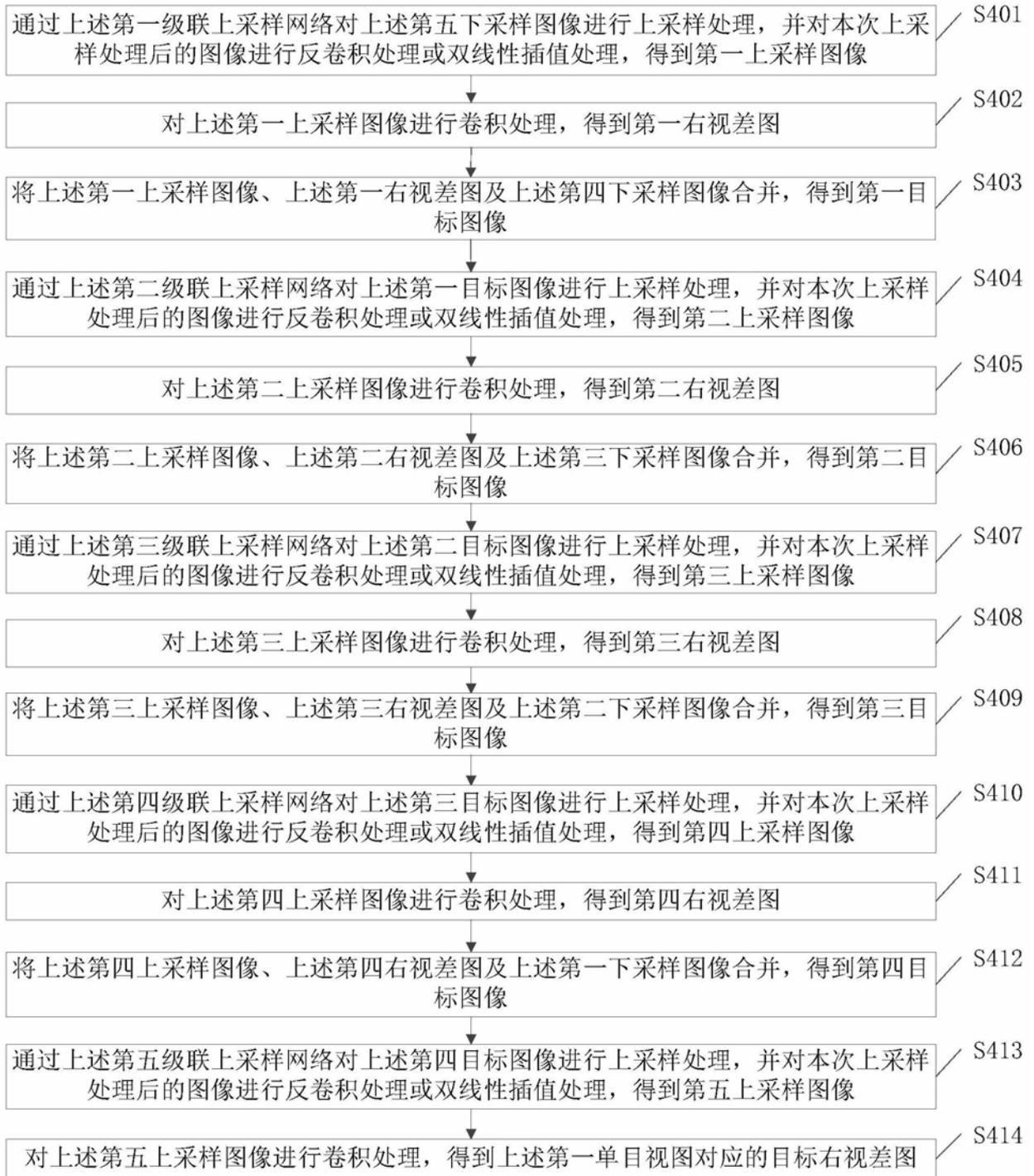


图4

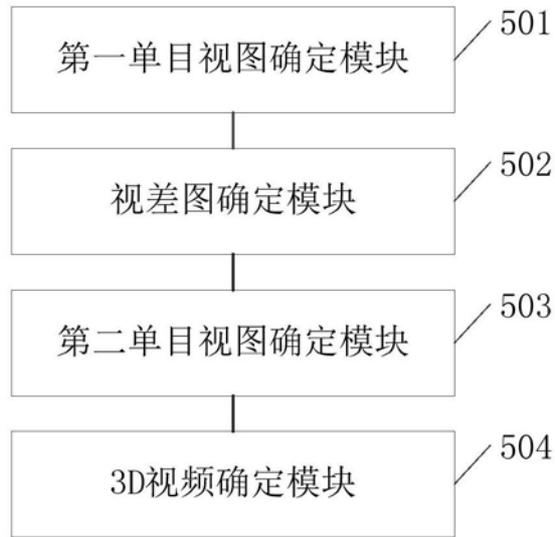


图5

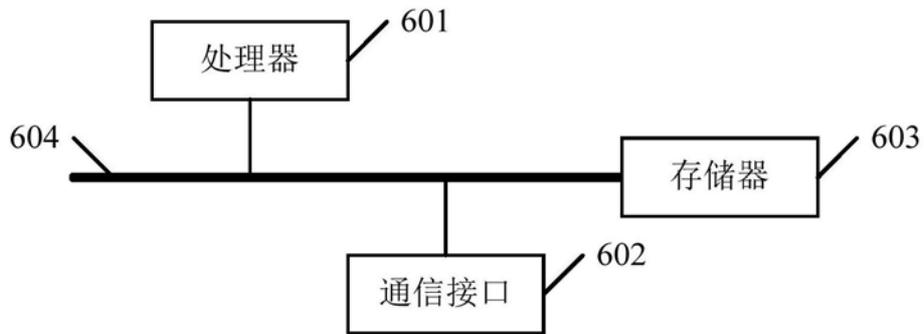


图6