



(12) 发明专利

(10) 授权公告号 CN 110084742 B

(45) 授权公告日 2024. 01. 26

(21) 申请号 201910381923.9

(22) 申请日 2019.05.08

(65) 同一申请的已公布的文献号
申请公布号 CN 110084742 A

(43) 申请公布日 2019.08.02

(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

(72) 发明人 吴方印 陈平 杨东

(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
专利代理师 项京 高莺然

(51) Int. Cl.

G06T 3/08 (2024.01)

G06T 7/55 (2017.01)

(56) 对比文件

CN 108921942 A, 2018.11.30

JP H0927969 A, 1997.01.28

CN 105335934 A, 2016.02.17

US 2013147911 A1, 2013.06.13

审查员 孙娟

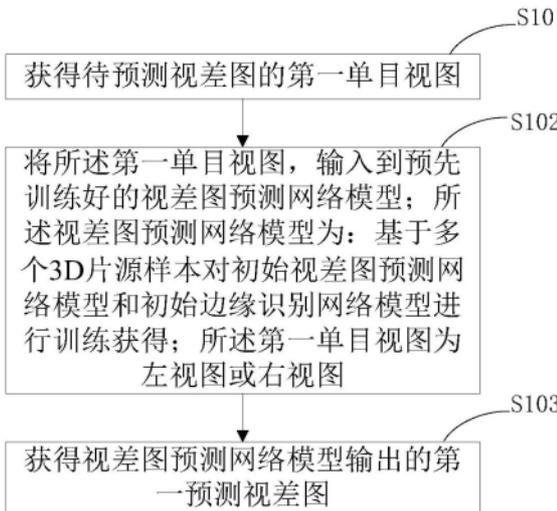
权利要求书10页 说明书26页 附图7页

(54) 发明名称

一种视差图预测方法、装置及电子设备

(57) 摘要

本发明实施例提供了一种视差图预测方法、装置及电子设备,可以获得待预测视差图的第一单目视图;将所述第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为:基于多个3D片源样本对初始视差图预测网络模型和初始边缘识别网络模型进行训练获得;所述第一单目视图为左视图或右视图;获得视差图预测网络模型输出的第一预测视差图。可见,应用本发明实施例,基于多个3D片源样本,对视差图预测网络模型和初始边缘识别网络模型进行训练,获得的视差图预测网络模型在训练过程中引入物体边缘信息,可以使物体边缘的视差过渡明显,物体的边缘更清晰。



1. 一种视差图预测的方法,其特征在于,所述方法包括:

获得待预测视差图的第一单目视图;

将所述第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为:基于多个3D片源样本对初始视差图预测网络模型和初始边缘识别网络模型进行训练获得;所述第一单目视图为左视图或右视图;

获得视差图预测网络模型输出的第一预测视差图;

其中,所述视差图预测网络模型的训练过程,包括:

利用所述多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络;利用所述多个3D片源样本的左视图,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络;利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络;

所述获得待预测视差图的第一单目视图的步骤,包括:

从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得所述当前待预测帧图像的左视图;针对所述当前待预测帧图像,获得上一帧图像与所述当前待预测帧图像之间的第三光流信息图;所述第三光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的;

所述将所述第一单目视图,输入到预先训练好的视差图预测网络模型的步骤,包括:

将所述当前待预测帧图像的左视图和所述第三光流信息图合并,将合并图输入所述视差图预测网络的编码部分,获得下采样图像;将所述视差图预测网络的编码部分获得的下采样图像,输入所述视差图预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第三预测视差图。

2. 根据权利要求1所述的方法,其特征在于,

所述多个3D片源样本为多个3D视频;

所述利用所述多个3D片源样本的左视图,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络的步骤,包括:

按如下预训练步骤,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络:

从所述多个3D片源样本中获取预设数量的连续视频帧,作为第一样本;

将各个第一样本的左视图输入初始视差图预测网络,获得各个第一样本的第一预测视差图;

将各个第一样本的右视图输入所述过渡边缘识别网络,获得各个第一样本的过渡第一边缘图;

基于各个第一样本的第一预测视差图、所述过渡第一边缘图和预设的损失函数,计算第一损失值;

基于所述第一损失值,对所述初始视差图预测网络的网络参数进行调整,在达到第一预设训练次数,获得过渡视差图预测网络。

3. 根据权利要求2所述的方法,其特征在于,

所述视差图预测网络,包括:编码部分和解码部分;

所述将各个第一样本的左视图输入初始视差图预测网络,获得各个第一样本的第一预

测视差图的步骤,包括:

针对各个第一样本中的当前帧图像;

获得上一帧图像与当前帧图像之间的第一光流信息图;所述的第一光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;

将当前帧图像的左视图和所述第一光流信息图合并,将合并图输入所述过渡视差图预测网络的编码部分,获得下采样图像;

将所述下采样图像,输入所述初始视差图预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第一预测视差图。

4. 根据权利要求3所述的方法,其特征在于,

所述第一预测视差图,包括:第一预测左视差图和第一预测右视差图;

所述预设的损失函数为:SSIM+L1损失函数和二阶梯度损失函数;

所述基于各个第一样本的第一预测视差图、所述过渡第一边缘图和预设的损失函数,计算第一损失值的步骤,包括:

基于所述第一预测左视差图和第一预测右视差图,获得第一预测左视图和第一预测右视图;

根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;

根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;

根据预测右视差图和实际右视差,计算出二阶梯度损失函数 C_{ds}^r ;

获得二阶梯度损失函数值图,基于公式一对二阶梯度损失函数值图进行调整,获得调整后的二阶梯度损失函数值;

对二阶梯度损失函数值图进行调整,获得调整后的二阶梯度损失函数值;

按公式二计算最终的第一损失值;

所述公式一为:调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)

所述公式二为: $(C_{ap}^l + C_{ap}^r) + (0.1 * \text{调整后二阶梯度损失函数值图})$ 。

5. 根据权利要求3所述的方法,其特征在于,

所述利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络的步骤,包括:

从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本;

将各个第二样本的左视图输入过渡视差图预测网络,获得各个第二样本的第二预测视差图;

将各个第二样本的右视图输入所述过渡边缘识别网络,获得各个第二样本的过渡第二边缘图;

基于各个第二样本的第二预测视差图、所述过渡第二边缘图和预设的损失函数,计算第二损失值;

基于所述第二损失值,判断所述过渡视差图预测网络和过渡边缘识别网络是否均收敛至平稳,

如果所述过渡视差图预测网络和过渡边缘识别网络均收敛至平稳,则将训练的次数加一,并判断是否达到第二预设训练次数;如果达到第二预设次数,则确定所述过渡视差图预测网络为最终的视差图预测网络;如果未达到第二预设次数,则返回所述从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本的步骤;

如果所述过渡视差图预测网络和过渡边缘识别网络未均收敛至平稳,则将训练的次数加一,并调整所述过渡视差图预测网络的网络参数和过渡边缘识别网络的网络参数,返回所述从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本的步骤。

6. 根据权利要求5所述的方法,其特征在于,

所述将各个第二样本的左视图输入过渡视差图预测网络,获得各个第二样本的第二预测视差图的步骤,包括:

针对各个第二样本中的当前帧图像;

获得上一帧图像与当前帧图像之间的第二光流信息图;所述的第二光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;

将当前帧图像的左视图和所述第二光流信息图合并,将合并图输入所述过渡视差图预测网络的编码部分,获得下采样图像;

将所述下采样图像,输入所述过渡视差图预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第二预测视差图。

7. 根据权利要求6所述的方法,其特征在于,

所述视差图预测网络的编码部分,包括:预设数量的编码卷积层,对输入的左视图和右视图进行合并后,对合并的图像进行预设层数的卷积,获得最后一层卷积层输出的下采样图像;

所述视差图预测网络的解码部分,包括:预设数量的解码卷积层和分别设置在各个解码卷积层之间的时序信息卷积层;

所述将所述下采样图像,输入所述过渡视差图预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第二预测视差图的步骤,包括:

将解码部分输出的下采样图像,输入所述过渡视差图预测网络的解码部分;

所述过渡视差图预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第二预测视差图。

8. 根据权利要求7所述的方法,其特征在于,所述第二预测视差图,包括:第二预测左视差图和第二预测右视差图;

所述预设的损失函数为:SSIM+L1损失函数和二阶梯度损失函数;

所述基于各个第二样本的第二预测视差图、所述过渡第二边缘图和预设的损失函数,计算第二损失值的步骤,包括:

基于所述第二预测左视差图和第二预测右视差图,获得第二预测左视图和第二预测右视图;

根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;

根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;

根据预测右视差图和实际右视差,计算出二阶梯度损失函数 C_{ds}^r ;

获得二阶梯度损失函数值图,基于公式一对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;

对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;

按公式二计算最终的第一损失值;

所述公式一为:调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)

所述公式二为: $(C_{ap}^l + C_{ap}^r) + (0.1 * \text{调整后二阶梯度损失函数值图})$ 。

9. 根据权利要求1所述的方法,其特征在于,

所述视差图预测网络的编码部分,包括:预设数量的编码卷积层,对输入的左视图和右视图进行合并后,对合并的图像进行预设层数的卷积,获得最后一层卷积层输出的下采样图像;

所述视差图预测网络的解码部分,包括:预设数量的解码卷积层和分别设置在各个解码卷积层之间的时序信息卷积层;

所述将所述视差图预测网络的编码部分获得的下采样图像,输入所述视差图预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第三预测视差图的步骤,包括:

所述视差图预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第三预测视差图。

10. 一种视差图预测网络模型的训练方法,其特征在于:包括:

利用多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络;

利用所述多个3D片源样本的左视图,对初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络;

利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络;

所述多个3D片源样本为多个3D视频;所述利用所述多个3D片源样本的左视图,对初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络的步骤,包括:

从所述多个3D片源样本中获取预设数量的连续视频帧,作为第一样本;针对各个第一样本中的当前帧图像;获得上一帧图像与当前帧图像之间的第一光流信息图;所述的第一光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;将当前帧图像的左视图和所述第一光流信息图合并,将合并图输入所述过渡视差图预测网络的编码部分,获得下采样图像;将所述下采样图像,输入所述初始视差图预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第一预测视差图;将各个第一样本的右视图输入所述过渡边缘识别网络,获得各个第一样本的过渡第一边缘图;基于各

个第一样本的第一预测视差图、所述过渡第一边缘图和预设的损失函数,计算第一损失值;基于所述第一损失值,对所述初始视差图预测网络的网络参数进行调整,在达到第一预设训练次数,获得过渡视差图预测网络。

11. 一种二维2D图像转换为三维3D图像的方法,其特征在于,所述方法包括:

获取待转换为3D图像的2D图像;

将所述2D图像作为用于生成3D图像的第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为采用权利要求10的方法训练获得的;所述第一单目视图为左视图或右视图;

获取视差图预测网络模型输出的第四预测视差图;

基于所述第四预测视差图,获得第二单目视图;所述第二单目视图为与第一单目视图对应的右视图或左视图;

基于所述第一单目视图和所述第二单目视图,生成3D图像。

12. 根据权利要求11所述的方法,其特征在于:

所述获取待转换为3D图像的2D图像的步骤,包括:

从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得当前待预测帧图像的左视图;

针对所述当前待预测帧图像,获得上一帧图像与当前待预测帧图像之间的第四光流信息图;所述的第四光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的;

所述视差图预测网络模型,包括:编码部分和解码部分;

所述将所述2D图像作为用于生成3D图像的第一单目视图,输入到预先训练好的视差图预测网络模型的步骤,包括:

将当前待预测帧图像的左视图和所述第四光流信息图合并,将合并图输入所述视差图预测网络模型的编码部分,获得下采样图像;

将所述视差图预测网络模型的编码部分获得的下采样图像,输入所述视差图预测网络模型的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第四预测视差图。

13. 一种视差图预测装置,其特征在于,所述装置包括:

第一单目视图获得单元,用于获得待预测视差图的第一单目视图;

单目视图输入单元,用于将所述第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为:基于多个3D片源样本对初始视差图预测网络模型和初始边缘识别网络模型进行训练获得;所述第一单目视图为左视图或右视图;

第一预测视差图获得单元,用于获得视差图预测网络模型输出的第一预测视差图;

其中,所述装置还包括:视差图预测网络模型训练单元;

所述视差图预测网络模型训练单元,包括:

第一过渡边缘识别网络获得模块,用于利用所述多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络;

第一过渡视差图预测网络获得模块,用于利用所述多个3D片源样本的左视图,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络;

最终视差网络获得模块,用于利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络;

所述第一单目视图获得单元,具体用于从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得所述当前待预测帧图像的左视图;针对所述当前待预测帧图像,获得上一帧图像与所述当前待预测帧图像之间的第三光流信息图;所述第三光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的;

所述单目视图输入单元,具体用于将所述当前待预测帧图像的左视图和所述第三光流信息图合并,将合并图输入所述视差图预测网络的编码部分,获得下采样图像;将所述视差图预测网络的编码部分获得的下采样图像,输入所述视差图预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第三预测视差图。

14. 根据权利要求13所述的装置,其特征在于,

所述多个3D片源样本为多个3D视频;

所述第一过渡视差图预测网络获得模块,具体用于采用如下子模块对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络:

第一样本获取子模块,用于从所述多个3D片源样本中获取预设数量的连续视频帧,作为第一样本;

第一预测视差图获得子模块,用于将各个第一样本的左视图输入初始视差图预测网络,获得各个第一样本的第一预测视差图;

过渡第一边缘图获得子模块,用于将各个第一样本的右视图输入所述过渡边缘识别网络,获得各个第一样本的过渡第一边缘图;

第一损失值计算子模块,用于基于各个第一样本的第一预测视差图、所述过渡第一边缘图和预设的损失函数,计算第一损失值;

第一过渡视差图预测网络获得子模块,用于基于所述第一损失值,对所述初始视差图预测网络的网络参数进行调整,在达到第一预设训练次数,获得过渡视差图预测网络。

15. 根据权利要求14所述的装置,其特征在于,

所述视差图预测网络,包括:编码部分和解码部分;

所述第一预测视差图获得子模块,包括:

第一当前帧图像预处理子模块,用于针对各个第一样本中的当前帧图像;

第一光流信息图获得子模块,用于获得上一帧图像与当前帧图像之间的第一光流信息图;所述的第一光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;

第一下采样图像获得子模块,用于将当前帧图像的左视图和所述第一光流信息图合并,将合并图输入所述过渡视差图预测网络的编码部分,获得下采样图像;

第一预测视差图输出子模块,用于将所述下采样图像,输入所述初始视差图预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第一预测视差图。

16. 根据权利要求15所述的装置,其特征在于,

所述第一预测视差图,包括:第一预测左视差图和第一预测右视差图;

所述预设的损失函数为:SSIM+L1损失函数和二阶梯度损失函数;

所述第一损失值计算子模块,具体用于:

基于所述第一预测左视差图和第一预测右视差图,获得第一预测左视图和第一预测右视图;

根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;

根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;

根据预测右视差图和实际右视差,计算出二阶梯度损失函数 C_{ds}^r ;

获得二阶梯度损失函数值图,基于公式一对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;

对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;

按公式二计算最终的第一损失值;

所述公式一为:调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)

所述公式二为: $(C_{ap}^l + C_{ap}^r) + (0.1 * \text{调整后二阶梯度损失函数值图})$ 。

17.根据权利要求15所述的装置,其特征在于,

所述最终视差网络获得模块,包括:

视频帧获取子模块,用于从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本;

第二预测视差图获得子模块,用于将各个第二样本的左视图输入过渡视差图预测网络,获得各个第二样本的第二预测视差图;

过渡第二边缘图获得子模块,用于将各个第二样本的右视图输入所述过渡边缘识别网络,获得各个第二样本的过渡第二边缘图;

第二损失值计算子模块,用于基于各个第二样本的第二预测视差图、所述过渡第二边缘图和预设的损失函数,计算第二损失值;

第一收敛至平稳判断子模块,用于基于所述第二损失值,判断所述过渡视差图预测网络和过渡边缘识别网络是否均收敛至平稳,

第二预设训练次数判断子模块,用于如果所述过渡视差图预测网络和过渡边缘识别网络均收敛至平稳,则将训练的次数加一,并判断是否达到第二预设训练次数;如果达到第二预设次数,则确定所述过渡视差图预测网络为最终的视差图预测网络;如果未达到第二预设次数,则触发所述视频帧获取子模块执行从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本;

参数调整子模块,用于如果所述过渡视差图预测网络和过渡边缘识别网络未均收敛至平稳,则将训练的次数加一,并调整所述过渡视差图预测网络的网络参数和过渡边缘识别网络的网络参数,触发所述视频帧获取子模块执行从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本。

18.根据权利要求17所述的装置,其特征在于,

所述第二预测视差图获得子模块,包括:

当前图像处理子模块,用于针对各个第二样本中的当前帧图像;

第二光流信息图获得子模块,用于获得上一帧图像与当前帧图像之间的第二光流信息图;所述的第二光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;

第二下采样图像获得子模块,用于将当前帧图像的左视图和所述第二光流信息图合并,将合并图输入所述过渡视差图预测网络的编码部分,获得下采样图像;

第二预测视差图输出子模块,用于将所述下采样图像,输入所述过渡视差图预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第二预测视差图。

19. 根据权利要求18所述的装置,其特征在于,

所述视差图预测网络的编码部分,包括:预设数量的编码卷积层,对输入的左视图和右视图进行合并后,对合并的图像进行预设层数的卷积,获得最后一层卷积层输出的下采样图像;

所述视差图预测网络的解码部分,包括:预设数量的解码卷积层和分别设置在各个解码卷积层之间的时序信息卷积层;

所述第二预测视差图输出子模块,具体用于:

将解码部分输出的下采样图像,输入所述过渡视差图预测网络的解码部分;

所述过渡视差图预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第二预测视差图。

20. 根据权利要求19所述的装置,其特征在于,所述第二预测视差图输出子模块获得的第二预测视差图,包括:第二预测左视差图和第二预测右视差图;

所述预设的损失函数为:SSIM+L1损失函数和二阶梯度损失函数;

所述第二损失值计算子模块,具体用于:

基于所述第二预测左视差图和第二预测右视差图,获得第二预测左视图和第二预测右视图;

根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;

根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;

根据预测右视差图和实际右视差,计算出梯度损失函数 C_{ds}^r ;

获得二阶梯度损失函数值图,基于公式一对二阶梯度损失函数值图进行调整,获得调整后的二阶梯度损失函数值;

对二阶梯度损失函数值图进行调整,获得调整后的二阶梯度损失函数值;

按公式二计算最终的第一损失值;

所述公式一为:调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)

所述公式二为: $(C_{ap}^l + C_{ap}^r) + (0.1 * \text{调整后二阶梯度损失函数值图})$ 。

21. 根据权利要求13所述的装置,其特征在于,

所述视差图预测网络的编码部分,包括:预设数量的编码卷积层,对输入的左视图和右视图进行合并后,对合并的图像进行预设层数的卷积,获得最后一层卷积层输出的下采样图像;

所述视差图预测网络的解码部分,包括:预设数量的解码卷积层和分别设置在各个解码卷积层之间的时序信息卷积层;

所述第三预测视差图输出模块,具体用于:

所述视差图预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第三预测视差图。

22. 一种视差图预测网络模型的训练装置,其特征在于,所述装置包括:

过渡边缘识别网络获得单元,用于利用多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络;

过渡视差图预测网络获得单元,用于利用所述多个3D片源样本的左视图,对初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络;

视差图预测网络获得单元,用于利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络;

所述多个3D片源样本为多个3D视频;所述过渡视差图预测网络获得单元,具体用于从所述多个3D片源样本中获取预设数量的连续视频帧,作为第一样本;针对各个第一样本中的当前帧图像;获得上一帧图像与当前帧图像之间的第一光流信息图;所述的第一光流信息图为先采用光流信息提取算法对各个第一样本进行计算获得的;将当前帧图像的左视图和所述第一光流信息图合并,将合并图输入所述过渡视差图预测网络的编码部分,获得下采样图像;将所述下采样图像,输入所述初始视差图预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第一预测视差图;将各个第一样本的右视图输入所述过渡边缘识别网络,获得各个第一样本的过渡第一边缘图;基于各个第一样本的第一预测视差图、所述过渡第一边缘图和预设的损失函数,计算第一损失值;基于所述第一损失值,对所述初始视差图预测网络的网络参数进行调整,在达到第一预设训练次数,获得过渡视差图预测网络。

23. 一种二维2D图像转换为三维3D图像的装置,其特征在于,所述装置包括:

待转换图像获取单元,用于获取待转换为3D图像的2D图像;

2D图像输入单元,用于将所述2D图像作为用于生成3D图像的第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为采用权利要求10的方法训练获得的;所述第一单目视图为左视图或右视图;

第四预测视差图获取单元,用于获取视差图预测网络模型输出的第四预测视差图;

第二单目视图获得单元,用于基于所述第四预测视差图,获得第二单目视图;所述第二单目视图为与第一单目视图对应的右视图或左视图;

3D图像生成单元,用于基于所述第一单目视图和所述第二单目视图,生成3D图像。

24. 根据权利要求23所述的装置,其特征在于:

所述待转换图像获取单元,包括:

左视图获得模块,用于从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得当前待预测帧图像的左视图;

第四光流信息图获得模块,用于针对所述当前待预测帧图像,获得上一帧图像与当前待预测帧图像之间的第四光流信息图;所述的第四光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的;

所述视差图预测网络模型,包括:编码部分和解码部分;

所述2D图像输入单元,包括:

下采样图像获得模块,用于将当前待预测帧图像的左视图和所述第四光流信息图合并,将合并图输入所述视差图预测网络模型的编码部分,获得下采样图像;

第四视差图输出模块,用于将所述视差图预测网络的编码部分获得的下采样图像,输入所述视差图预测网络模型的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第四预测视差图。

25. 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,处理器,通信接口,存储器通过通信总线完成相互间的通信;

存储器,用于存放计算机程序;

处理器,用于执行存储器上所存放的程序时,实现权利要求1-9任一所述的方法步骤。

26. 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,处理器,通信接口,存储器通过通信总线完成相互间的通信;

存储器,用于存放计算机程序;

处理器,用于执行存储器上所存放的程序时,实现权利要求10所述的方法步骤。

27. 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,处理器,通信接口,存储器通过通信总线完成相互间的通信;

存储器,用于存放计算机程序;

处理器,用于执行存储器上所存放的程序时,实现权利要求11或12所述的方法步骤。

一种视差图预测方法、装置及电子设备

技术领域

[0001] 本发明涉及2D图像转换成3D图像的技术领域,特别是涉及一种视差图预测方法、装置及电子设备。

背景技术

[0002] 目前,将一张2D图像转换成3D图像的方法,通常是:将2D图像做为左视图,预测出左视差图或右视差图,再根据预测的左视差图或右视差图获得预测的右视图,再用左视图和右视图转换成3D图像。

[0003] 其中,根据一张2D图像预测视差图,是2D图像转换为3D图像的重要步骤。现有技术中,进行视差图预测的方法通常是:将左视图输入视差图预测网络模型中,获得预测的视差图;这个网络模型是用大量的2D图像对单一网络模型进行训练获得的。

[0004] 发明人在实现本发明的过程中发现,现有技术至少存在如下问题:

[0005] 在使用单一网络模型进行训练获得的视差图预测网络模型预测出的视差图,对2D图像进行3D图像转换的时候,物体的边缘不够清晰。

发明内容

[0006] 本发明实施例的目的在于提供一种视差图预测方法、装置及电子设备。

[0007] 用以解决在用视差图预测网络模型预测出的视差图,对2D图像进行3D图像转换的时候,物体的边缘不够清晰的问题。

[0008] 具体技术方案如下:

[0009] 第一方面,本发明实施例提供了一种视差图预测的方法,所述方法包括:

[0010] 获得待预测视差图的第一单目视图;

[0011] 将所述第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为:基于多个3D片源样本对初始视差图预测网络模型和初始边缘识别网络模型进行训练获得;所述第一单目视图为左视图或右视图;

[0012] 获得视差图预测网络模型输出的第一预测视差图。

[0013] 可选的,

[0014] 所述视差图预测网络模型的训练过程,包括:

[0015] 利用所述多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络;

[0016] 利用所述多个3D片源样本的左视图,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络;

[0017] 利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络。

[0018] 可选的,

[0019] 所述多个3D片源样本为多个3D视频;

[0020] 所述利用所述多个3D片源样本的左视图,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络的步骤,包括:

[0021] 按如下预训练步骤,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络:

[0022] 从所述多个3D片源样本中获取预设数量的连续视频帧,作为第一样本;

[0023] 将各个第一样本的左视图输入初始视差图预测网络,获得各个第一样本的第一预测视差图;

[0024] 将各个第一样本的右视图输入所述过渡边缘识别网络,获得各个第一样本的过渡第一边缘图;

[0025] 基于各个第一样本的第一预测视差图、所述过渡第一边缘图和预设的损失函数,计算第一损失值;

[0026] 基于所述第一损失值,对所述初始视差图预测网络的网络参数进行调整,在达到第一预设训练次数,获得过渡视差图预测网络。

[0027] 可选的,

[0028] 所述视差图预测网络,包括:编码部分和解码部分;

[0029] 所述将各个第一样本的左视图输入初始视差图预测网络,获得各个第一样本的第一预测视差图的步骤,包括:

[0030] 针对各个第一样本中的当前帧图像;

[0031] 获得上一帧图像与当前帧图像之间的第一光流信息图;所述的第一光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;

[0032] 将当前帧图像的左视图和所述第一光流信息图合并,将合并图输入所述过渡视差预测网络的编码部分,获得下采样图像;

[0033] 将所述下采样图像,输入所述初始视差预测网络的解码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第一预测视差图。

[0034] 可选的,

[0035] 所述基于各个第一样本的第一预测视差图、所述过渡第一边缘图和预设的损失函数,计算第一损失值的步骤,包括:

[0036] 所述第一预测视差图,包括:第一预测左视差图和第一预测右视差图;

[0037] 所述预设的损失函数为:SSIM+L1损失函数和二阶梯度损失函数;

[0038] 所述基于各个第一样本的第一预测视差图、所述过渡第一边缘图和预设的损失函数,计算第一损失值的步骤,包括:

[0039] 基于所述第一预测左视差图和第一预测右视差图,获得第一预测左视图和第一预测右视图;

[0040] 根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;

[0041] 根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;

[0042] 根据预测右视差图和实际右视差,计算出二阶梯度损失函数 C_{ds}^r ;

[0043] 获得二阶梯度损失函数值图,基于公式一对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;

- [0044] 对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;
- [0045] 按公式二计算最终的第一损失值;
- [0046] 所述公式一为:调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)
- [0047] 所述公式二为: $(C_{ap}^l + C_{ap}^r) + (0.1 * \text{调整后二阶梯度损失函数值图})$ 。
- [0048] 可选的,
- [0049] 所述利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络的步骤,包括:
- [0050] 从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本;
- [0051] 将各个第二样本的左视图输入过渡视差图预测网络,获得各个第二样本的第二预测视差图;
- [0052] 将各个第二样本的右视图输入所述过渡边缘识别网络,获得各个第二样本的过渡第二边缘图;
- [0053] 基于各个第二样本的第二预测视差图、所述过渡第二边缘图和预设的损失函数,计算第二损失值;
- [0054] 基于所述第二损失值,判断所述过渡视差图预测网络和过渡边缘识别网络是否均收敛至平稳,
- [0055] 如果所述过渡视差图预测网络和过渡边缘识别网络均收敛至平稳,则将训练的次数加一,并判断是否达到第二预设训练次数;如果达到第二预设次数,则确定所述过渡视差图预测网络为最终的视差图预测网络;如果未达到第二预设次数,则返回所述从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本的步骤;
- [0056] 如果所述过渡视差图预测网络和过渡边缘识别网络未均收敛至平稳,则将训练的次数加一,并调整所述过渡视差图预测网络的网络参数和过渡边缘识别网络的网络参数,返回所述从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本的步骤。
- [0057] 可选的,
- [0058] 所述将各个第二样本的左视图输入过渡视差图预测网络,获得各个第二样本的第二预测视差图的步骤,包括:
- [0059] 针对各个第二样本中的当前帧图像;
- [0060] 获得上一帧图像与当前帧图像之间的第二光流信息图;所述的第二光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;
- [0061] 将当前帧图像的左视图和所述第二光流信息图合并,将合并图输入所述过渡视差预测网络的编码部分,获得下采样图像;
- [0062] 将所述下采样图像,输入所述过渡视差预测网络的解码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第二预测视差图。
- [0063] 可选的,
- [0064] 所述视差图预测网络的编码部分,包括:预设数量的编码卷积层,对输入的左视图和右视图进行合并后,对合并的图像进行预设层数的卷积,获得最后一层卷积层输出的下采样图像;

- [0065] 所述视差图预测网络的解码部分,包括:预设数量的解码卷积层和分别设置在各个解码卷积层之间的时序信息卷积层;
- [0066] 所述将所述下采样图像,输入所述过渡视差预测网络的解码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第二预测视差图的步骤,包括:
- [0067] 将解码部分输出的下采样图像,输入所述过渡视差预测网络的解码部分;
- [0068] 所述过渡视差预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第二预测视差图。
- [0069] 可选的,所述第二预测视差图,包括:第二预测左视差图和第二预测右视差图;
- [0070] 所述预设的损失函数为:SSIM+L1损失函数和二阶梯度损失函数;
- [0071] 所述基于各个第二样本的第二预测视差图、所述过渡第二边缘图和预设的损失函数,计算第二损失值的步骤,包括:
- [0072] 基于所述第二预测左视差图和第二预测右视差图,获得第二预测左视图和第二预测右视图;
- [0073] 根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;
- [0074] 根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;
- [0075] 根据预测右视差图和实际右视差,计算出二阶梯度损失函数 C_{ds}^r ;
- [0076] 获得二阶梯度损失函数值图,基于公式一对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;
- [0077] 对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;
- [0078] 按公式二计算最终的第一损失值;
- [0079] 所述公式一为:调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)
- [0080] 所述公式二为: $(C_{ap}^l + C_{ap}^r) + (0.1 * \text{调整后二阶梯度损失函数值图})$ 。
- [0081] 可选的,
- [0082] 所述获得待预测视差图的第一单目视图的步骤,包括:
- [0083] 从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得当前待预测帧图像的左视图;
- [0084] 针对所述当前待预测视频帧,获得上一帧图像与当前待预测帧图像之间的第三光流信息图;所述的第三光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的;
- [0085] 所述将所述第一单目视图,输入到预先训练好的视差图预测网络模型的步骤,包括:
- [0086] 将当前待预测帧图像的左视图和所述第三光流信息图合并,将合并图输入所述视差预测网络的编码部分,获得下采样图像;
- [0087] 将所述视差预测网络的编码部分获得的下采样图像,输入所述视差预测网络的解

码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第三预测视差图。

[0088] 可选的,

[0089] 所述视差图预测网络的编码部分,包括:预设数量的编码卷积层,对输入的左视图和右视图进行合并后,对合并的图像进行预设层数的卷积,获得最后一层卷积层输出的下采样图像;

[0090] 所述视差图预测网络的解码部分,包括:预设数量的解码卷积层和分别设置在各个解码卷积层之间的时序信息卷积层;

[0091] 所述将所述视差预测网络的编码部分获得的下采样图像,输入所述视差预测网络的解码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第三预测视差图步骤,包括:

[0092] 所述视差预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第三预测视差图。

[0093] 第二方面,本发明实施例提供了一种视差图预测网络模型的训练方法,包括:

[0094] 利用所述多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络;

[0095] 利用所述多个3D片源样本的左视图,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络;

[0096] 利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络。

[0097] 第三方面,本发明实施例提供了一种二维2D图像转换为三维3D图像的方法,所述方法包括:

[0098] 获取待转换为3D图像的2D图像;

[0099] 将所述2D图像作为用于生成3D图像的第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为采用上述任一项的方法训练获得的;所述第一单目视图为左视图或右视图;

[0100] 获取视差图预测网络模型输出的第四预测视差图;

[0101] 基于所述第四预测视差图,获得第二单目视图;所述第二单目视图为与第一单目视图对应的右视图或左视图;

[0102] 基于所述第一单目视图和所述第二单目视图,生成3D图像。

[0103] 可选的,所述获取待转换为3D图像的2D图像的步骤,包括:

[0104] 从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得当前待预测帧图像的左视图;

[0105] 针对所述当前待预测视频帧,获得上一帧图像与当前待预测帧图像之间的第四光流信息图;所述的第四光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的;

[0106] 所述视差图预测网络模型,包括:编码部分和解码部分;

[0107] 所述将所述第一单目视图,输入到预先训练好的视差图预测网络模型的步骤,包

括：

[0108] 将当前待预测帧图像的左视图和所述第四光流信息图合并,将合并图输入所述视差预测网络模型的编码部分,获得下采样图像;

[0109] 将所述视差预测网络模型的编码部分获得的下采样图像,输入所述视差预测网络的解码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第四预测视差图。

[0110] 第四方面,本发明实施例提供了一种视差图预测的装置,所述装置包括:

[0111] 第一单目视图获得单元,用于获得待预测视差图的第一单目视图;

[0112] 单目视图输入单元,用于将所述第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为:基于多个3D片源样本对初始视差图预测网络模型和初始边缘识别网络模型进行训练获得;所述第一单目视图为左视图或右视图;

[0113] 第一预测视差图获得单元,用于获得视差图预测网络模型输出的第一预测视差图。

[0114] 可选的,所述装置还包括:第一视差图预测网络模型训练单元:

[0115] 所述视差图预测网络模型训练单元,包括:

[0116] 第一过渡边缘识别网络获得模块,用于利用所述多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络;

[0117] 第一过渡视差图预测网络获得模块,用于利用所述多个3D片源样本的左视图,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络;

[0118] 最终视差网络获得模块,用于利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络。

[0119] 可选的,

[0120] 所述多个3D片源样本为多个3D视频;

[0121] 所述第一过渡视差图预测网络获得模块,具体用于采用如下子模块对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络:

[0122] 第一样本获取子模块,用于从所述多个3D片源样本中获取预设数量的连续视频帧,作为第一样本;

[0123] 第一预测视差图获得子模块,用于将各个第一样本的左视图输入初始视差图预测网络,获得各个第一样本的第一预测视差图;

[0124] 过渡第一边缘图获得子模块,用于将各个第一样本的右视图输入所述过渡边缘识别网络,获得各个第一样本的过渡第一边缘图;

[0125] 第一损失值计算子模块,用于基于各个第一样本的第一预测视差图、所述过渡第一边缘图和预设的损失函数,计算第一损失值;

[0126] 第一过渡视差图预测网络获得子模块,用于基于所述第一损失值,对所述初始视差图预测网络的网络参数进行调整,在达到第一预设训练次数,获得过渡视差图预测网络。

[0127] 可选的,

[0128] 所述视差图预测网络,包括:编码部分和解码部分;

[0129] 所述第一预测视差图获得子模块,包括:

[0130] 第一当前帧图像预处理子模块,用于针对各个第一样本中的当前帧图像;

[0131] 第一光流信息图获得子模块,用于获得上一帧图像与当前帧图像之间的第一光流信息图;所述的第一光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;

[0132] 第一下采样图像获得子模块,用于将当前帧图像的左视图和所述第一光流信息图合并,将合并图输入所述过渡视差预测网络的编码部分,获得下采样图像;

[0133] 第一预测视差图输出子模块,用于将所述下采样图像,输入所述初始视差预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第一预测视差图。

[0134] 可选的,

[0135] 所述第一预测视差图,包括:第一预测左视差图和第一预测右视差图;

[0136] 所述预设的损失函数为:SSIM+L1损失函数和二阶梯度损失函数;

[0137] 所述第一损失值计算子模块,具体用于:

[0138] 基于所述第一预测左视差图和第一预测右视差图,获得第一预测左视图和第一预测右视图;

[0139] 根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;

[0140] 根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;

[0141] 根据预测右视差图和实际右视差,计算出二阶梯度损失函数 C_{ds}^r ;

[0142] 获得二阶梯度损失函数值图,基于公式一对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;

[0143] 对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;

[0144] 按公式二计算最终的第一损失值;

[0145] 所述公式一为:调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)

[0146] 所述公式二为:($C_{ap}^l + C_{ap}^r$) + (0.1 *调整后二阶梯度损失函数值图)。

[0147] 可选的,

[0148] 所述最终视差网络获得模块,包括:

[0149] 视频帧获取子模块,用于从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本;

[0150] 第二预测视差图获得子模块,用于将各个第二样本的左视图输入过渡视差图预测网络,获得各个第二样本的第二预测视差图;

[0151] 过渡第二边缘图获得子模块,用于将各个第二样本的右视图输入所述过渡边缘识别网络,获得各个第二样本的过渡第二边缘图;

[0152] 第二损失值计算子模块,用于基于各个第二样本的第二预测视差图、所述过渡第二边缘图和预设的损失函数,计算第二损失值;

[0153] 第一收敛至平稳判断子模块,用于基于所述第二损失值,判断所述过渡视差图预测网络和过渡边缘识别网络是否均收敛至平稳,

[0154] 第二预设训练次数判断子模块,用于如果所述过渡视差图预测网络和过渡边缘识

别网络均收敛至平稳,则将训练的次数加一,并判断是否达到第二预设训练次数;如果达到第二预设次数,则确定所述过渡视差图预测网络为最终的视差图预测网络;如果未达到第二预设次数,则触发所述视频帧获取子模块执行从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本;

[0155] 参数调整子模块,用于如果所述过渡视差图预测网络和过渡边缘识别网络未均收敛至平稳,则将训练的次数加一,并调整所述过渡视差图预测网络的网络参数和过渡边缘识别网络的网络参数,触发所述视频帧获取子模块执行从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本。

[0156] 可选的,

[0157] 所述第二预测视差图获得子模块,包括:

[0158] 当前图像处理子模块,用于针对各个第二样本中的当前帧图像;

[0159] 第二光流信息图获得子模块,用于获得上一帧图像与当前帧图像之间的第二光流信息图;所述的第二光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;

[0160] 第二下采样图像获得子模块,用于将当前帧图像的左视图和所述第二光流信息图合并,将合并图输入所述过渡视差预测网络的编码部分,获得下采样图像;

[0161] 第二预测视差图输出子模块,用于将所述下采样图像,输入所述过渡视差预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第二预测视差图。

[0162] 可选的,

[0163] 所述视差图预测网络的编码部分,包括:预设数量的编码卷积层,对输入的左视图和右视图进行合并后,对合并的图像进行预设层数的卷积,获得最后一层卷积层输出的下采样图像;

[0164] 所述视差图预测网络的解码部分,包括:预设数量的解码卷积层和分别设置在各个解码卷积层之间的时序信息卷积层;

[0165] 所述第二预测视差图输出子模块,具体用于:

[0166] 将解码部分输出的下采样图像,输入所述过渡视差预测网络的解码部分;

[0167] 所述过渡视差预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第二预测视差图。

[0168] 可选的,所述所述第二预测视差图输出子模块获得的第二预测视差图,包括:第二预测左视差图和第二预测右视差图;

[0169] 所述预设的损失函数为:SSIM+L1损失函数和二阶梯度损失函数;

[0170] 所述第二损失值计算子模块,具体用于:

[0171] 基于所述第二预测左视差图和第二预测右视差图,获得第二预测左视图和第二预测右视图;

[0172] 根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;

- [0173] 根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;
- [0174] 根据预测右视图和实际右视图,计算出梯度损失函数 C_{ds}^r ;
- [0175] 获得二阶梯度损失函数值图,基于公式一对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;
- [0176] 对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;
- [0177] 按公式二计算最终的第一损失值;
- [0178] 所述公式一为:调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)
- [0179] 所述公式二为:($C_{ap}^l + C_{ap}^r$) + (0.1 *调整后二阶梯度损失函数值图)。
- [0180] 可选的,
- [0181] 所述第一单目视图获得单元,包括:
- [0182] 当前待预测图像左视图获得模块,用于从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得当前待预测帧图像的左视图;
- [0183] 第三光流信息图获得模块,用于针对所述当前待预测视频帧,获得上一帧图像与当前待预测帧图像之间的第三光流信息图;所述的第三光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的;
- [0184] 所述单目视图输入单元,包括:
- [0185] 图像获得模块,用于将当前待预测帧图像的左视图和所述第三光流信息图合并,将合并图输入所述视差预测网络的编码部分,获得下采样图像;
- [0186] 第三预测视差图输出模块,用于将所述视差预测网络的编码部分获得的下采样图像,输入所述视差预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第三预测视差图。
- [0187] 可选的,
- [0188] 所述视差图预测网络的编码部分,包括:预设数量的编码卷积层,对输入的左视图和右视图进行合并后,对合并的图像进行预设层数的卷积,获得最后一层卷积层输出的下采样图像;
- [0189] 所述视差图预测网络的解码部分,包括:预设数量的解码卷积层和分别设置在各个解码卷积层之间的时序信息卷积层;
- [0190] 所述第三预测视差图输出模块,具体用于:
- [0191] 所述视差预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第三预测视差图。
- [0192] 第五方面,本发明实施例提供了一种视差图预测网络模型的训练装置,所述装置,包括:
- [0193] 过渡边缘识别网络获得单元,用于利用所述多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络;
- [0194] 过渡视差图预测网络获得单元,用于利用所述多个3D片源样本的左视图,对所述

初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络;

[0195] 视差图预测网络获得单元,用于利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络。

[0196] 第六方面,本发明实施例提供了一种二维2D图像转换为三维3D图像的装置,所述装置包括:

[0197] 待转换图像获取单元,用于获取待转换为3D图像的2D图像;

[0198] 2D图像输入单元,用于将所述2D图像作为用于生成3D图像的第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为采用上述任一项的训练方法训练获得的;所述第一单目视图为左视图或右视图;

[0199] 第四预测视差图获取单元,用于获取视差图预测网络模型输出的第四预测视差图;

[0200] 第二单目视图获得单元,用于基于所述第四预测视差图,获得第二单目视图;所述第二单目视图为与第一单目视图对应的右视图或左视图;

[0201] 3D图像生成单元,用于基于所述第一单目视图和所述第二单目视图,生成3D图像。

[0202] 可选的,所述待转换图像获取单元,包括:

[0203] 左视图获得模块,用于从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得当前待预测帧图像的左视图;

[0204] 第四光流信息图获得模块,用于针对所述当前待预测视频帧,获得上一帧图像与当前待预测帧图像之间的第四光流信息图;所述的第四光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的;

[0205] 所述视差图预测网络模型,包括:编码部分和解码部分;

[0206] 所述2D图像输入单元,包括:

[0207] 下采样图像获得模块,用于将当前待预测帧图像的左视图和所述第四光流信息图合并,将合并图输入所述视差预测网络模型的编码部分,获得下采样图像;

[0208] 第四视差图输出模块,用于将所述视差预测网络的编码部分获得的下采样图像,输入所述视差预测网络模型的解码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第四预测视差图。

[0209] 第七方面,本发明实施提供了一种电子设备,包括处理器、通信接口、存储器和通信总线,其中,处理器,通信接口,存储器通过通信总线完成相互间的通信;

[0210] 存储器,用于存放计算机程序;

[0211] 处理器,用于执行存储器上所存放的程序时,实现上述任一视差图预测方法步骤。

[0212] 第八方面,本发明实施提供了一种电子设备,包括处理器、通信接口、存储器和通信总线,其中,处理器,通信接口,存储器通过通信总线完成相互间的通信;

[0213] 存储器,用于存放计算机程序;

[0214] 处理器,用于执行存储器上所存放的程序时,实现上述视差图预测模型的训练方法步骤。

[0215] 第九方面,本发明实施提供了一种电子设备,包括处理器、通信接口、存储器和通信总线,其中,处理器,通信接口,存储器通过通信总线完成相互间的通信;

[0216] 存储器,用于存放计算机程序;

[0217] 处理器,用于执行存储器上所存放的程序时,实现上述2D图像转换为3D图像的方法步骤。

[0218] 本发明实施还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任一视差图预测方法的步骤。

[0219] 本发明实施例还提供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任一视差图预测的方法。

[0220] 本发明实施例有益效果:

[0221] 本发明实施例提供的一种视差图预测方法、装置及电子设备,可以获得待预测视差图的第一单目视图;将所述第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为:基于多个3D片源样本对初始视差图预测网络模型和初始边缘识别网络模型进行训练获得;所述第一单目视图为左视图或右视图;获得视差图预测网络模型输出的第一预测视差图。可见,应用本发明实施例,基于多个3D片源样本,对视差图预测网络模型和初始边缘识别网络模型进行训练,获得的视差图预测网络模型在训练过程中引入物体边缘信息,可以使物体边缘的视差过渡明显,物体的边缘更清晰。

[0222] 当然,实施本发明的任一产品或方法并不一定需要同时达到以上所述的所有优点。

附图说明

[0223] 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,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的前提下,还可以根据这些附图获得其他的附图。

[0224] 图1为本发明实施例提供的一种视差图预测的方法的一种流程图;

[0225] 图2为本发明实施例提供的一种视差图预测网络模型的一种具体训练流程图;

[0226] 图3为本发明实施例提供的视差图预测网络模型和边缘识别网络的训练原理图;

[0227] 图4为本发明实施例提供的一种2D图像转换为3D图像的方法的一种流程图;

[0228] 图5为本发明实施例提供的一种2D图像转换为3D图像的方法的另一种流程图;

[0229] 图6为本发明实施例提供的视差图预测的装置的结构图;

[0230] 图7为本发明实施例提供的2D图像转换为3D图像的装置的结构图;

[0231] 图8为本发明实施例提供的一种电子设备的结构示意图;

[0232] 图9为本发明实施例提供的另一种电子设备的结构示意图;

[0233] 图10为本发明实施例提供的另一种电子设备的结构示意图。

具体实施方式

[0234] 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清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0235] 为了解决在用视差图预测网络模型预测出的视差图,对2D图像进行3D图像转换的

时候,物体的边缘不够清晰的问题,本发明实施例提供了视差图预测方法、装置及电子设备。

[0236] 本发明实施例所提供的视差图预测、模型训练和图像转换方法可以应用于任意需要对图像转换、深度图预测和模型训练的电子设备,如:电脑或移动终端等,在此不做具体限定。为了描述方便,以下简称电子设备。

[0237] 本发明实施例提供的视差图预测的方法,如图1所示,该方法的具体处理流程包括:

[0238] 步骤S101,获得待预测视差图的第一单目视图。

[0239] 可实施的,电子设备可以获得待预测视差图的第一单目视图,该第一单目视图可以为左视图,也可以为右视图,在这里不做具体限定。在一种具体的实施方式中,可以从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得当前待预测帧图像的左视图;

[0240] 针对所述当前待预测视频帧,获得上一帧图像与当前待预测帧图像之间的第三光流信息图;所述的第三光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的。在一种具体的实施方式中,可以使用FlowNet2.0算法获得第三光流信息图。可实施的,本发明中提到的光流信息图都可以通过这种算法获得。

[0241] 步骤S102,将所述第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为:基于多个3D片源样本对初始视差图预测网络模型和初始边缘识别网络模型进行训练获得;所述第一单目视图为左视图或右视图。

[0242] 在一种实施方式中,可以将当前待预测帧图像的左视图和所述第三光流信息图合并,将合并图输入所述视差预测网络的编码部分,获得下采样图像;将所述视差预测网络的编码部分获得的下采样图像,输入所述视差预测网络的解码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第三预测视差图。在一种具体的实施方式中,所述视差预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第三预测视差图。

[0243] 所述视差图预测网络,可以包括:编码部分和解码部分。视差图预测网络的编码部分,包括:预设数量的编码卷积层,对输入的左视图和光流信息图进行合并后,对合并的图像进行预设层数的卷积,获得最后一层卷积层输出的下采样图像。所述视差图预测网络的解码部分,包括:预设数量的解码卷积层和分别设置在各个解码卷积层之间的时序信息卷积层。

[0244] 可参见表1所示的本发明实施例的视差图预测网络模型编解码部分的编解码表。

[0245] 表1

[0246]

编码	输入	输出
Conv1	[256, 512, 3]	[256, 512, 32]
Conv2	[256, 512, 32]	[128, 256, 32]
Conv3	[128, 256, 32]	[128, 256, 64]
Conv4	[128, 256, 64]	[64, 128, 64]
Conv5	[64, 128, 64]	[64, 128, 128]
Conv6	[64, 128, 128]	[32, 64, 128]
Conv7	[32, 64, 128]	[32, 64, 256]
Conv8	[32, 64, 256]	[16, 32, 256]
Conv9	[16, 32, 256]	[16, 32, 512]
Conv10	[16, 32, 512]	[8, 16, 512]
Conv11	[8, 16, 512]	[8, 16, 512]
Conv12	[8, 16, 512]	[4, 8, 512]
Conv13	[4, 8, 512]	[4, 8, 512]
Conv14	[4, 8, 512]	[2, 4, 512]
解码		
UpSample1	[2, 4, 512]	[4, 8, 512]
Conv1	[4, 8, 512]	[4, 8, 512]
Conv2	[4, 8, 512]	[4, 8, 512]
UpSample2	[4, 8, 512]	[8, 16, 512]
Conv3	[8, 16, 512]	[8, 16, 512]
Conv4	[8, 16, 512]	[8, 16, 512]
UpSample3	[8, 16, 512]	[16, 32, 512]
Conv5	[16, 32, 512]	[16, 32, 256]

[0247]

Conv6	[16, 32, 256]	[16, 32, 256]
UpSample4	[16, 32, 256]	[32, 64, 256]
Conv7	[32, 64, 256]	[32, 64, 128]
Conv8	[32, 64, 128]	[32, 64, 128]
ConvLSTM1	[32, 64, 128]	[32, 64, 128]
Conv8_out	[32, 64, 128]	[32, 64, 2]
UpSample5	[32, 64, 128]	[64, 128, 128]
Conv9	[64, 128, 128]	[64, 128, 64]
Conv10	[64, 128, 64]	[64, 128, 64]
ConvLSTM2	[64, 128, 64]	[64, 128, 64]
Conv10_out	[64, 128, 64]	[64, 128, 2]
UpSample6	[64, 128, 64]	[128, 256, 64]
Conv11	[128, 256, 64]	[128, 256, 32]
Conv12	[128, 256, 32]	[128, 256, 32]
ConvLSTM3	[128, 256, 32]	[128, 256, 32]
Conv12_out	[128, 256, 32]	[128, 256, 2]
UpSample7	[128, 256, 32]	[256, 512, 32]
Conv13	[256, 512, 32]	[256, 512, 16]
Conv14	[256, 512, 16]	[256, 512, 16]
ConvLSTM4	[256, 512, 16]	[256, 512, 16]
Conv14_out	[256, 512, 16]	[256, 512, 2]

[0248] 如表1所示,编码部分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、第五级联下采样网络、第六级联下采样网络及第七级联下采样网络。每个下采样级联网络分别包含两个卷积层,当然也可以根据实际需要对联下网络的结构进行调整。

[0249] 可实施的,以左视图为例进行说明,从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得当前待预测帧图像的左视图;

[0250] 针对所述当前待预测视频帧,预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得上一帧图像与当前待预测帧图像之间的光流信息图。将当前待预测帧图像的左视图和所述第三光流信息图合并,将合并图输入所述视差预测网络的编码部分,编码部分对样本中的合并图两次卷积分别进行增加通道和减小尺寸处理,得到最后一层卷积层输出的第二编码下采样图像。如表1所示,将尺寸为265*512*3的合并图输入到第一级联下采样网络中,其中,265可表示合并图的宽;512可表示合并图的高;3可表示该合并图的通道数。第一级联下采样网络包含conv1(第一个卷积层)和conv2(第二个卷积层),conv1(第一个卷积层)对265*512*3的合并图进行增加维度的卷积处理得到265*512*32的特征图1,conv2(第二个卷积层)对特征图1进行减小尺寸的卷积处理得到128*265*32的特征图2;再将特征图2经过conv3(第三个卷积层)卷积处理得到128*265*64的特征图3。以此类推,最终经过conv14(第十四个卷积层)卷积处理得到2*4*512的下采样图像。再将下采样

图像解码部分。

[0251] 解码部分包括：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、第五级联上采样网络、第六级联上采样网络及第七级联上采样网络。每个上采样级联网络分别包含上采样及两个卷积层，当然也可以根据实际需要对级联网络的结构进行调整。每个上采样级联网络包含上采样双线性插值增大尺寸的处理和两个卷积层降低维度的处理，其中一个卷积层做降低维度处理，另一个卷积不做降低维度处理。

[0252] 解码部分对经过编码部分获得的下采样图像，对 $2*4*512$ 的图像进行第一上采样双线性插值，增大尺寸处理得到 $4*8*512$ 的上采样中间图像1，conv1(第一个卷积层)对上采样中间图像1卷积处理得到 $4*8*512$ 的上采样特征图1，再将上采样特征图1经过conv2(第二个卷积层)卷积处理得到上采样特征图2。特别说明在此处两次卷积都不做降低通道处理是为了模型需要，可以根据实际情况调整。

[0253] 再将特征图2经过第二上采样双线性插值，增大尺寸处理得到 $8*16*512$ 的上采样中间图像2，conv3(第三个卷积层)对上采样中间图像2卷积处理 $8*16*512$ 的上采样特征图3，再将上采样特征图3经过conv4(第四个卷积层)卷积处理得到上采样特征图4。特别说明在此处两次卷积都不做降低通道处理是为了模型需要，可以根据实际情况调整。

[0254] 再将特征图4经过第三上采样双线性插值，增大尺寸处理得到 $8*16*512$ 的上采样中间图像3，conv5(第五个卷积层)对上采样中间图像3卷积降低通道处理 $16*32*512$ 的上采样特征图5，再将上采样特征图5经过conv6(第六个卷积层)卷积处理得到上采样特征图6。以此类推。

[0255] 特别说明，当经过Conv8后输出 $32*64*128$ 的特征图8，再将特征图8经过加时序信息卷积层(ConvLSTM1)利用上一帧图像的时序信息图对特征图8进行处理得到加入了时序信息的时序特征图1，将时序特征图1传递给下一帧图像，作为时序输入，并且将时序特征图1经Conv8_out输出一次预测的视差图。同时将时序特征图1经过第五上采样双线性插值，继续处理，以此类推。

[0256] 同理在ConvLSTM2、ConvLSTM3及ConvLSTM4之后分别输出一个预测的左视差图和预测的右视差图，如表中Conv8_out、Conv10_out、Conv12_out及Conv14_out所示。相当于一个样本可以输出4次预测的左视差图和预测的右视差图，最终根据这4次预测的左视差图和预测的右视差图求平均损失值。

[0257] 需要说明的是，本发明可选的实施例中设置了七个级联采样网络，在实际实施过程中，可按照实施人员的具体需求设置多于七个或少于七个的级联采样网络。

[0258] 可实施的，所述视差图预测网络模型的训练过程，可以包括：

[0259] 利用所述多个3D片源样本的右视图对初始边缘识别网络进行训练，获得过渡边缘识别网络；利用所述多个3D片源样本的左视图，对所述初始视差图预测网络进行第一预设次数的预训练，获得过渡视差图预测网络；利用所述多个3D片源样本的左视图和右视图，对所述过渡视差图预测网络和过渡边缘识别网络进行训练，获得最终的视差图预测网络。

[0260] 步骤S103，获得视差图预测网络模型输出的第一预测视差图。

[0261] 可见，应用本发明实施例，基于多个3D片源样本，对视差图预测网络模型和初始边缘识别网络模型进行训练，获得的视差图预测网络模型在训练过程中引入物体边缘信息，

可以使物体边缘的视差过渡明显,物体的边缘更清晰。

[0262] 而且,在视差图预测的过程中引入了光流信息和时序信息,使得对同一点的深度变化平缓,解决帧间抖动的问题。

[0263] 本发明实施例还提供了一种视差图预测网络模型的训练方法,包括如下步骤:

[0264] 步骤一,利用所述多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络;

[0265] 步骤二,利用所述多个3D片源样本的左视图,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络;

[0266] 步骤三,利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络。

[0267] 应用本发明实施例,基于多个3D片源样本,对视差图预测网络模型和初始边缘识别网络模型进行训练,获得的视差图预测网络模型在训练过程中引入物体边缘信息,可以使物体边缘的视差过渡明显,物体的边缘更清晰。

[0268] 视差图预测网络模型的训练方法的一种具体实现过程,如图2所示,具体处理流程包括:

[0269] 步骤S201,利用多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络。

[0270] 可实施的,所述多个3D片源样本为多个3D视频,每一个3D视频可以为一个样本。首先可以将初始边缘识别网络在数据集PASCAL VOC Context上进行预训练,该数据集中包含图片样本和对应的边缘图。

[0271] 步骤S202,从所述多个3D片源样本中获取预设数量的连续视频帧,作为第一样本。

[0272] 可实施的,可以固定过渡边缘识别网络的网络参数,然后从多个3D片源样本中获取预设数量的连续视频帧,作为第一样本。

[0273] 步骤S203,将各个第一样本的左视图输入初始视差图预测网络,获得各个第一样本的第一预测视差图。

[0274] 可实施的,可以针对各个第一样本中的当前帧图像;获得上一帧图像与当前帧图像之间的第一光流信息图;所述的第一光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;将当前帧图像的左视图和所述第一光流信息图合并,将合并图输入所述过渡视差预测网络的编码部分,获得下采样图像;将所述下采样图像,输入所述初始视差预测网络的解码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第一预测视差图。

[0275] 步骤S204,将各个第一样本的右视图输入所述过渡边缘识别网络,获得各个第一样本的过渡第一边缘图。

[0276] 步骤S205,基于各个第一样本的第一预测视差图、所述过渡第一边缘图和预设的损失函数,计算第一损失值。

[0277] 可实施的,所述第一预测视差图,包括:第一预测左视差图和第一预测右视差图;

[0278] 所述预设的损失函数为:SSIM+L1损失函数和二阶梯度损失函数;

[0279] 基于所述第一预测左视差图和第一预测右视差图,获得第一预测左视图和第一预测右视图;

- [0280] 根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;
- [0281] 根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;
- [0282] 根据预测右视差图和实际右视差,计算出二阶梯度损失函数;
- [0283] 获得二阶梯度损失函数值图,基于公式
- [0284] 调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)
- [0285] 对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值 C_{ds}^r ;
- [0286] 按公式 $(C_{ap}^l + C_{ap}^r) + (0.1 * \text{调整后二阶梯度损失函数值图})$,计算最终的第一损失值。
- [0287] 步骤S206,基于所述第一损失值,对所述初始视差图预测网络的网络参数进行调整,在达到第一预设训练次数时,获得过渡视差图预测网络。
- [0288] 步骤S207,从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本。
- [0289] 步骤S208,将各个第二样本的左视图输入过渡视差图预测网络,获得各个第二样本的第二预测视差图。
- [0290] 可实施的,针对各个第二样本中的当前帧图像;获得上一帧图像与当前帧图像之间的第二光流信息图;所述的第二光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;将当前帧图像的左视图和所述第二光流信息图合并,将合并图输入所述过渡视差预测网络的编码部分,获得下采样图像。
- [0291] 可以再将所述下采样图像,输入所述过渡视差预测网络的解码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第二预测视差图。
- [0292] 在一种具体的实施方式中,可以将解码部分输出的下采样图像,输入过渡视差预测网络的解码部分;再将所述过渡视差预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第二预测视差图。
- [0293] 第二预测视差图,包括:第二预测左视差图和第二预测右视差图;
- [0294] 预设的损失函数可以为:SSIM+L1损失函数和二阶梯度损失函数;
- [0295] 步骤S209,将各个第二样本的右视图输入所述过渡边缘识别网络,获得各个第二样本的过渡第二边缘图。
- [0296] 步骤S210,基于各个第二样本的第二预测视差图、所述过渡第二边缘图和预设的损失函数,计算第二损失值。
- [0297] 基于所述第二预测左视差图和第二预测右视差图,获得第二预测左视图和第二预测右视图;
- [0298] 根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;
- [0299] 根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;
- [0300] 根据预测右视差图和实际右视差,计算出二阶梯度损失函数;
- [0301] 根据预测右视差图和实际右视差,计算出梯度损失函数 C_{ds}^r ;

[0302] 获得二阶梯度损失函数值图,基于公式一对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;

[0303] 对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;

[0304] 按公式二计算最终的第一损失值;

[0305] 所述公式一为:调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)

[0306] 所述公式二为:($C_{ap}^l + C_{ap}^r$) + (0.1 *调整后二阶梯度损失函数值图)。

[0307] 在一种具体的实施方式中,以根据预测左视图和真实左视图为例,所述SSIM+L1损失函数公式为:

$$[0308] \quad G_{ap}^l = \frac{1}{N} \sum_{i,j} \alpha \frac{1 - SSIM(I_{ij}^l, \tilde{I}_{ij}^l)}{2} + (1 - \alpha) \|I_{ij}^l - \tilde{I}_{ij}^l\|$$

[0309] 其中, G_{ap}^l 表示损失值; N 表示一小批次训练样本的数量; l 表示左视图; α 权重为0.85; I_{ij}^l 表示真实左视图; \tilde{I}_{ij}^l 表示预测左视图; $SSIM(I_{ij}^l, \tilde{I}_{ij}^l)$ 表示预测左视图与真实左视图的结构相似性; $\|I_{ij}^l - \tilde{I}_{ij}^l\|$ 表示预测左视图与真实左视图的绝对值误差L1。

[0310] 根据预测左视图和真实左视图,按照所述SSIM+L1损失函数,获得损失值 C_{ap}^l ;

[0311] 同理,如果是根据预测右视图和真实右视图,按照所述SSIM+L1损失函数,获得损失值 C_{ap}^r ;

[0312] 以根据右视图,预测左视差图为例,二阶梯度损失函数公式为:

$$[0313] \quad G_{ds}^l = \frac{1}{N} \sum_{i,j} |\partial_{xx} d_{ij}^l| + |\partial_{xy} d_{ij}^l| + |\partial_{yy} d_{ij}^l| + |\partial_{yx} d_{ij}^l|$$

[0314] C_{ds}^l 表示损失值, $\partial_{xx} d_{ij}^l$ 代表左视差图在x方向上的一阶导数在x方向上的导数,

$\partial_{xy} d_{ij}^l$ 代表左视差图在x方向上的一阶导数在y方向上的导数, $\partial_{yy} d_{ij}^l$ 代表左视差图在y方向上的一阶导数在y方向上的导数, $\partial_{yx} d_{ij}^l$ 代表左视差图在y方向上的一阶导数在x方向上的导数; i, j 代表像素的坐标。

[0315] 特别说明的是,本发明在视差图预测网络模型训练的过程中,将左视图输入到当前的视差图预测网络模型,根据左视图预测出的右视差图,对该右视差图求二阶梯度损失函数,所以获得的是右视差图的二阶梯度损失值 C_{ds}^r 。

[0316] 基于调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第二边缘图)

[0317] 对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值 C_{ds}^r ;

[0318] 按公式($C_{ap}^l + C_{ap}^r$) + (0.1 *调整后二阶梯度损失函数值图),计算最终

的第二损失值。

[0319] 在一种具体的实施方式中,可以先判断过渡第二边缘图的值小于0.5还是大于0.5;如果小于0.5,则将所述过渡第二边缘图的值改变为0;因为这时候的过渡第二边缘图的值可以认为不准确。如果大于或等于0.5,则将所述过渡第二边缘图的值改变为1;因为这时候的过渡第二边缘图的值可以认为准确,具体的方式在这里不做具体限定。最终得到的边缘识别网络,预测出的边缘部分值为1,非边缘部分值为0。

[0320] 步骤S211,基于所述第二损失值,判断所述过渡视差图预测网络和过渡边缘识别网络是否均收敛至平稳。

[0321] 如果判断的结果为否,即如果过渡视差图预测网络和过渡边缘识别网络不均收敛至平稳,则执行步骤S212;如果判断的结果为是,即如果过渡视差图预测网络和过渡边缘识别网络均收敛至平稳,则执行步骤S213。

[0322] 步骤S212,将训练的次数加一,并调整所述过渡视差图预测网络的网络参数和过渡边缘识别网络的网络参数。返回执行步骤S207。

[0323] 步骤S213,将训练的次数加一,并判断是否达到第二预设训练次数。

[0324] 如果判断的结果为是,即如果达到第二预设训练次数,则执行步骤S214;如果没达到第二预设训练次数,则返回执行步骤S207。

[0325] 步骤S214,确定所述过渡视差图预测网络为最终的视差图预测网络。

[0326] 可见,应用本发明实施例,基于多个3D片源样本,对视差图预测网络模型和初始边缘识别网络模型进行训练,获得的视差图预测网络模型在训练过程中引入物体边缘信息,可以使物体边缘的视差过渡明显,物体的边缘更清晰。

[0327] 而且,在视差图预测的过程中引入了光流信息和时序信息,使得对同一点的深度变化平缓,解决帧间抖动的问题。

[0328] 本发明实施例提供的视差图预测网络模型和边缘识别网络的训练原理图,如图3所示,包括:

[0329] 视差图预测网络和边缘识别网络;其中位置在上的是视差图预测网络,位置在下的是边缘识别网络。

[0330] 可以先将初始边缘识别网络在数据集,比如PASCAL VOC Context上进行训练,该数据集中包含图片样本和对应的边缘图,获得过渡边缘识别网络可实施的,初始边缘识别检测模型可以为HED网络(边缘检测模型)。

[0331] 再固定边缘识别网络的网络参数,对视差图预测网络模型进行训练,从所述多个3D片源样本中获取预设数量的连续视频帧,作为一次训练,可将连续帧的右视图R输入边缘网络,同时将该连续帧与右视图对应的左视图与光流信息图合并输入视差图预测网络模型,在视差图预测网络模型中经过编码解码的操作可以引入所述连续帧中的上一帧的时序信息图,输出引入时序信息和光流信息的预测视差图。根据其中得到的右视差图,和对真实的左视图的采样处理,获得预测的右视图R'与真实的右视图R求SSIM+L1损失函数,再根据其中得到的左视差图,和对真实的右视图的采样处理,获得预测的左视图与真实的左视图求SSIM+L1损失函数,再将预测的右视差图求基于边缘信息的二阶损失函数。根据预设的损失函数计算的损失值调节视差图预测网络,再同时对边缘预测网络和视差预测网络进行训练,直到获得训练好的视差图预测网络。

[0332] 其中,边缘识别网络输出层可以用Sigmoid激活函数,其它层可以用relu激活函数;视差图预测网络输出层可以用tanh激活函数,其它层可以用elu激活函数。

[0333] 可见,应用本发明实施例,基于多个3D片源样本,对视差图预测网络模型和初始边缘识别网络模型进行训练,获得的视差图预测网络模型在训练过程中引入物体边缘信息,可以使物体边缘的视差过渡明显,物体的边缘更清晰。

[0334] 而且,在视差图预测的过程中引入了光流信息和时序信息,使得对同一点的深度变化平缓,解决帧间抖动的问题。

[0335] 参见图4,图4为本发明实施例提供的一种2D图像转换为3D图像的方法的一种流程图;该方法包括:

[0336] 步骤S401,获取待转换为3D图像的2D图像;

[0337] 步骤S402,将所述2D图像作为用于生成3D图像的第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型可以图2所示的训练获得的;所述第一单目视图为左视图或右视图;

[0338] 本实施例中的步骤S401和步骤S402可以与图1中的步骤S101和S102相同,这里不再赘述。

[0339] 步骤S403,获取视差图预测网络模型输出的第四预测视差图;

[0340] 步骤S404,基于所述第四预测视差图,获得第二单目视图;所述第二单目视图为与第一单目视图对应的右视图或左视图;

[0341] 步骤S405,基于所述第一单目视图和所述第二单目视图,生成3D图像。

[0342] 应用本发明实施例,基于多个3D片源样本,对视差图预测网络模型和初始边缘识别网络模型进行训练,获得的视差图预测网络模型在训练过程中引入物体边缘信息,可以使物体边缘的视差过渡明显,物体的边缘更清晰。

[0343] 本发明实施例的视差图预测网络模型,可以包括:编码部分和解码部分。

[0344] 参见图5,图5为本发明实施例提供的一种2D图像转换为3D图像的方法的另一种流程图;该方法包括:

[0345] 步骤S501,从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得当前待预测帧图像的左视图;

[0346] 步骤S502,针对所述当前待预测视频帧,获得上一帧图像与当前待预测帧图像之间的第四光流信息图;所述的第四光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的;

[0347] 步骤S503,将当前待预测帧图像的左视图和所述第四光流信息图合并,将合并图输入所述视差预测网络的编码部分,获得下采样图像;

[0348] 步骤S504,将所述视差预测网络的编码部分获得的下采样图像,输入所述视差预测网络的解码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第四预测视差图。

[0349] 可见,应用本发明实施例,基于多个3D片源样本,对视差图预测网络模型和初始边缘识别网络模型进行训练,获得的视差图预测网络模型在训练过程中引入物体边缘信息,可以使物体边缘的视差过渡明显,物体的边缘更清晰。

[0350] 而且,在视差图预测的过程中引入了光流信息和时序信息,使得对同一点的深度

变化平缓,解决帧间抖动的问题。

[0351] 本发明实施例提供的视差图预测的装置的结构图,如图6所示,包括:

[0352] 第一单目视图获得单元601,用于获得待预测视差图的第一单目视图;

[0353] 单目视图输入单元602,用于将所述第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为:基于多个3D片源样本对初始视差图预测网络模型和初始边缘识别网络模型进行训练获得;所述第一单目视图为左视图或右视图;

[0354] 第一预测视差图获得单元603,用于获得视差图预测网络模型输出的第一预测视差图。

[0355] 可选的,所述装置还包括:第一视差图预测网络模型训练单元:

[0356] 所述视差图预测网络模型训练单元,包括:

[0357] 第一过渡边缘识别网络获得模块,用于利用所述多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络;

[0358] 第一过渡视差图预测网络获得模块,用于利用所述多个3D片源样本的左视图,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络;

[0359] 最终视差网络获得模块,用于利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络。

[0360] 可选的,

[0361] 所述多个3D片源样本为多个3D视频;

[0362] 所述第一过渡视差图预测网络获得模块,具体用于采用如下子模块对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络:

[0363] 第一样本获取子模块,用于从所述多个3D片源样本中获取预设数量的连续视频帧,作为第一样本;

[0364] 第一预测视差图获得子模块,用于将各个第一样本的左视图输入初始视差图预测网络,获得各个第一样本的第一预测视差图;

[0365] 过渡第一边缘图获得子模块,用于将各个第一样本的右视图输入所述过渡边缘识别网络,获得各个第一样本的过渡第一边缘图;

[0366] 第一损失值计算子模块,用于基于各个第一样本的第一预测视差图、所述过渡第一边缘图和预设的损失函数,计算第一损失值;

[0367] 第一过渡视差图预测网络获得子模块,用于基于所述第一损失值,对所述初始视差图预测网络的网络参数进行调整,在达到第一预设训练次数,获得过渡视差图预测网络。

[0368] 可选的,

[0369] 所述视差图预测网络,包括:编码部分和解码部分;

[0370] 所述第一预测视差图获得子模块,包括:

[0371] 第一当前帧图像预处理子模块,用于针对各个第一样本中的当前帧图像;

[0372] 第一光流信息图获得子模块,用于获得上一帧图像与当前帧图像之间的第一光流信息图;所述的第一光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;

[0373] 第一下采样图像获得子模块,用于将当前帧图像的左视图和所述第一光流信息图合并,将合并图输入所述过渡视差预测网络的编码部分,获得下采样图像;

[0374] 第一预测视差图输出子模块,用于将所述下采样图像,输入所述初始视差预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第一预测视差图。

[0375] 可选的,

[0376] 所述第一预测视差图,包括:第一预测左视差图和第一预测右视差图;

[0377] 所述预设的损失函数为:SSIM+L1损失函数和二阶梯度损失函数;

[0378] 所述第一损失值计算子模块,具体用于:

[0379] 基于所述第一预测左视差图和第一预测右视差图,获得第一预测左视图和第一预测右视图;

[0380] 根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;

[0381] 根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;

[0382] 根据预测右视差图和实际右视差,计算出梯度损失函数 C_{ds}^r ;

[0383] 获得二阶梯度损失函数值图,基于公式一对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;

[0384] 对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值;

[0385] 按公式二计算最终的第一损失值;

[0386] 所述公式一为:调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)

[0387] 所述公式二为: $(C_{ap}^l + C_{ap}^r) + (0.1 * \text{调整后二阶梯度损失函数值图})$ 。

[0388] 可选的,

[0389] 所述最终视差网络获得模块,包括:

[0390] 视频帧获取子模块,用于从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本;

[0391] 第二预测视差图获得子模块,用于将各个第二样本的左视图输入过渡视差图预测网络,获得各个第二样本的第二预测视差图;

[0392] 过渡第二边缘图获得子模块,用于将各个第二样本的右视图输入所述过渡边缘识别网络,获得各个第二样本的过渡第二边缘图;

[0393] 第二损失值计算子模块,用于基于各个第二样本的第二预测视差图、所述过渡第二边缘图和预设的损失函数,计算第二损失值;

[0394] 第一收敛至平稳判断子模块,用于基于所述第二损失值,判断所述过渡视差图预测网络和过渡边缘识别网络是否均收敛至平稳,

[0395] 第二预设训练次数判断子模块,用于如果所述过渡视差图预测网络和过渡边缘识别网络均收敛至平稳,则将训练的次数加一,并判断是否达到第二预设训练次数;如果达到第二预设次数,则确定所述过渡视差图预测网络为最终的视差图预测网络;如果未达到第二预设次数,则触发所述视频帧获取子模块执行从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本;

[0396] 参数调整子模块,用于如果所述过渡视差图预测网络和过渡边缘识别网络未均收

敛至平稳,则将训练的次数加一,并调整所述过渡视差图预测网络的网络参数和过渡边缘识别网络的网络参数,触发所述视频帧获取子模块执行从所述多个3D片源样本中获取预设数量的连续视频帧,作为第二样本。

[0397] 可选的,

[0398] 所述第二预测视差图获得子模块,包括:

[0399] 当前图像处理子模块,用于针对各个第二样本中的当前帧图像;

[0400] 第二光流信息图获得子模块,用于获得上一帧图像与当前帧图像之间的第二光流信息图;所述的第二光流信息图为预先采用光流信息提取算法对各个第一样本进行计算获得的;

[0401] 第二下采样图像获得子模块,用于将当前帧图像的左视图和所述第二光流信息图合并,将合并图输入所述过渡视差预测网络的编码部分,获得下采样图像;

[0402] 第二预测视差图输出子模块,用于将所述下采样图像,输入所述过渡视差预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第二预测视差图。

[0403] 可选的,

[0404] 所述视差图预测网络的编码部分,包括:预设数量的编码卷积层,对输入的左视图和右视图进行合并后,对合并的图像进行预设层数的卷积,获得最后一层卷积层输出的下采样图像;

[0405] 所述视差图预测网络的解码部分,包括:预设数量的解码卷积层和分别设置在各个解码卷积层之间的时序信息卷积层;

[0406] 所述第二预测视差图输出子模块,具体用于:

[0407] 将解码部分输出的下采样图像,输入所述过渡视差预测网络的解码部分;

[0408] 所述过渡视差预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第二预测视差图。

[0409] 可选的,所述所述第二预测视差图输出子模块获得的第二预测视差图,包括:第二预测左视差图和第二预测右视差图;

[0410] 所述预设的损失函数为:SSIM+L1损失函数和二阶梯度损失函数;

[0411] 所述第二损失值计算子模块,具体用于:

[0412] 基于所述第二预测左视差图和第二预测右视差图,获得第二预测左视图和第二预测右视图;

[0413] 根据预测右视图和实际右视图计算出SSIM+L1损失函数值 C_{ap}^r ;

[0414] 根据预测左视图和实际左视图计算出SSIM+L1损失函数值 C_{ap}^l ;

[0415] 根据预测右视差图和实际右视差,计算出梯度损失函数 C_{ds}^r ;

[0416] 根据预测左视差图和实际左视差,计算出梯度损失函数 C_{ds}^l ;

[0417] 获得二阶梯度损失函数值图,基于公式一对二阶梯度损失函数值进行调整,获得

调整后的二阶梯度损失函数值；

[0418] 对二阶梯度损失函数值进行调整,获得调整后的二阶梯度损失函数值；

[0419] 按公式二计算最终的第一损失值；

[0420] 所述公式一为:调整后二阶梯度损失函数值图=二阶梯度损失函数值图*(1-过渡第一边缘图)

[0421] 所述公式二为: $(C_{ap}^l + C_{ap}^r) + (0.1 * \text{调整后二阶梯度损失函数值图})$ 。

[0422] 可选的,

[0423] 所述第一单目视图获得单元,包括:

[0424] 当前待预测图像左视图获得模块,用于从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得当前待预测帧图像的左视图;

[0425] 第三光流信息图获得模块,用于针对所述当前待预测视频帧,获得上一帧图像与当前待预测帧图像之间的第三光流信息图;所述的第三光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的;

[0426] 所述单目视图输入单元,包括:

[0427] 图像获得模块,用于将当前待预测帧图像的左视图和所述第三光流信息图合并,将合并图输入所述视差预测网络的编码部分,获得下采样图像;

[0428] 第三预测视差图输出模块,用于将所述视差预测网络的编码部分获得的下采样图像,输入所述视差预测网络的解码部分,利用上一帧图像的时序信息图对所述下采样图像进行处理,输出第三预测视差图。

[0429] 可选的,

[0430] 所述视差图预测网络的编码部分,包括:预设数量的编码卷积层,对输入的左视图和右视图进行合并后,对合并的图像进行预设层数的卷积,获得最后一层卷积层输出的下采样图像;

[0431] 所述视差图预测网络的解码部分,包括:预设数量的解码卷积层和分别设置在各个解码卷积层之间的时序信息卷积层;

[0432] 所述第三预测视差图输出模块,具体用于:

[0433] 所述视差预测网络的解码部分的各个时序信息卷积层,用上一帧图像的时序信息图对各个解码卷积层输出的解码图像进行卷积,卷积的结果图像作为下一解码卷积层的输入图和当前帧图像的时序信息图;各个层时序信息卷积层输出的卷积结果图,作为第三预测视差图。

[0434] 可见,应用本发明实施例,基于多个3D片源样本,对视差图预测网络模型和初始边缘识别网络模型进行训练,获得的视差图预测网络模型在训练过程中引入物体边缘信息,可以使物体边缘的视差过渡明显,物体的边缘更清晰。

[0435] 而且,在视差图预测的过程中引入了光流信息和时序信息,使得对同一点的深度变化平缓,解决帧间抖动的问题。

[0436] 本发明实施例还提供了一种视差图预测网络模型的训练装置,该装置包括:

[0437] 过渡边缘识别网络获得单元,用于利用所述多个3D片源样本的右视图对初始边缘识别网络进行训练,获得过渡边缘识别网络;

[0438] 过渡视差图预测网络获得单元,用于利用所述多个3D片源样本的左视图,对所述初始视差图预测网络进行第一预设次数的预训练,获得过渡视差图预测网络;

[0439] 视差图预测网络获得单元,用于利用所述多个3D片源样本的左视图和右视图,对所述过渡视差图预测网络和过渡边缘识别网络进行训练,获得最终的视差图预测网络。

[0440] 应用本发明实施例,基于多个3D片源样本,对视差图预测网络模型和初始边缘识别网络模型进行训练,获得的视差图预测网络模型在训练过程中引入物体边缘信息,可以使物体边缘的视差过渡明显,物体的边缘更清晰。

[0441] 参见图7,图7为本发明实施例提供的2D图像转换为3D图像的装置的结构图;该装置包括:

[0442] 待转换图像获取单元701,用于获取待转换为3D图像的2D图像;

[0443] 2D图像输入单元702,用于将所述2D图像作为用于生成3D图像的第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为采用权利要求12的方法训练获得的;所述第一单目视图为左视图或右视图;

[0444] 第四预测视差图获取单元703,用于获取视差图预测网络模型输出的第四预测视差图;

[0445] 第二单目视图获得单元704,用于基于所述第四预测视差图,获得第二单目视图;所述第二单目视图为与第一单目视图对应的右视图或左视图;

[0446] 3D图像生成单元705,用于基于所述第一单目视图和所述第二单目视图,生成3D图像。

[0447] 其中,所述待转换图像获取单元,可以包括:

[0448] 左视图获得模块,用于从待转换为3D视频的2D视频中确定一个视频帧作为当前待预测帧图像,获得当前待预测帧图像的左视图;

[0449] 第四光流信息图获得模块,用于针对所述当前待预测视频帧,获得上一帧图像与当前待预测帧图像之间的第四光流信息图;所述的第四光流信息图为预先采用光流信息提取算法,对所述待转换为3D视频的2D视频进行计算获得的;

[0450] 所述视差图预测网络模型,包括:编码部分和解码部分;

[0451] 所述2D图像输入单元,可以包括:

[0452] 下采样图像获得模块,用于将当前待预测帧图像的左视图和所述第四光流信息图合并,将合并图输入所述视差预测网络模型的编码部分,获得下采样图像;

[0453] 第四视差图输出模块,用于将所述视差预测网络的编码部分获得的下采样图像,输入所述视差预测网络模型的解码部分,利用上一帧图像的时序信息图对所述上采样图像进行处理,输出第四预测视差图。

[0454] 可见,应用本发明实施例,基于多个3D片源样本,对视差图预测网络模型和初始边缘识别网络模型进行训练,获得的视差图预测网络模型在训练过程中引入物体边缘信息,可以使物体边缘的视差过渡明显,物体的边缘更清晰。

[0455] 而且,在视差图预测的过程中引入了光流信息和时序信息,使得对同一点的深度变化平缓,解决帧间抖动的问题。

[0456] 本发明实施例还提供了一种电子设备,如图8所示,包括处理器801、通信接口802、存储器803和通信总线804,其中,处理器801,通信接口802,存储器803通过通信总线804完

成相互间的通信,

[0457] 存储器803,用于存放计算机程序;

[0458] 处理器801,用于执行存储器803上所存放的程序时,实现如下步骤:

[0459] 获得待预测视差图的第一单目视图;

[0460] 将所述第一单目视图,输入到预先训练好的视差图预测网络模型;所述视差图预测网络模型为:基于多个3D片源样本对初始视差图预测网络模型和初始边缘识别网络模型进行训练获得;所述第一单目视图为左视图或右视图;

[0461] 获得视差图预测网络模型输出的第一预测视差图。

[0462] 可见,应用本发明实施例,基于多个3D片源样本,对视差图预测网络模型和初始边缘识别网络模型进行训练,获得的视差图预测网络模型在训练过程中引入物体边缘信息,可以使物体边缘的视差过渡明显,物体的边缘更清晰。

[0463] 而且,在视差图预测的过程中引入了光流信息和时序信息,使得对同一点的深度变化平缓,解决帧间抖动的问题。

[0464] 本发明实施例还提供了另一种电子设备,如图9所示,包括处理器901、通信接口902、存储器903和通信总线904,其中,处理器901,通信接口902,存储器903通过通信总线904完成相互间的通信,

[0465] 存储器903,用于存放计算机程序;

[0466] 处理器901,用于执行存储器903上所存放的程序时,实现上述任一项视差图预测网络模型训练方法。

[0467] 本发明实施例还提供了再一种电子设备,如图10所示,包括处理器1001、通信接口1002、存储器1003和通信总线1004,其中,处理器1001,通信接口1002,存储器1003通过通信总线1004完成相互间的通信,

[0468] 存储器1003,用于存放计算机程序;

[0469] 处理器1001,用于执行存储器1003上所存放的程序时,实现上述任一项2D图像转换为3D图像的方法。

[0470] 上述电子设备提到的通信总线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。

[0471] 通信接口用于上述电子设备与其他设备之间的通信。

[0472] 存储器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,存储器还可以是至少一个位于远离前述处理器的存储装置。

[0473] 上述的处理器可以是通用处理器,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。

[0474] 在本发明提供的又一实施例中,还提供了一种计算机可读存储介质,该计算机可

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任一视差图预测方法的步骤。

[0475] 在本发明提供的又一实施例中,还提供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施例中任一视差图预测方法。

[0476] 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。当使用软件实现时,可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时,全部或部分地产生按照本发明实施例所述的流程或功能。所述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。所述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,所述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。所述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。所述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。

[0477] 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,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0478] 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、计算机可读存储介质及计算机程序产品等实施例而言,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。

[0479] 以上所述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。

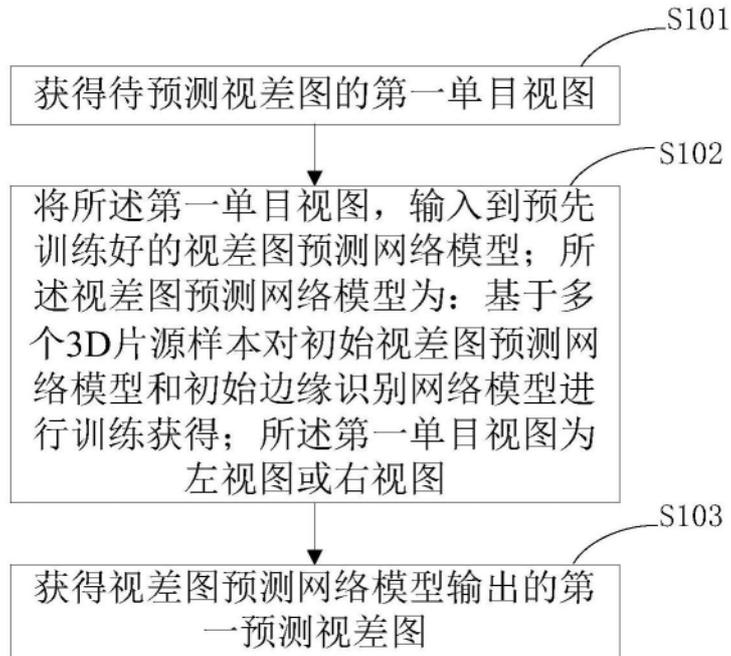


图1

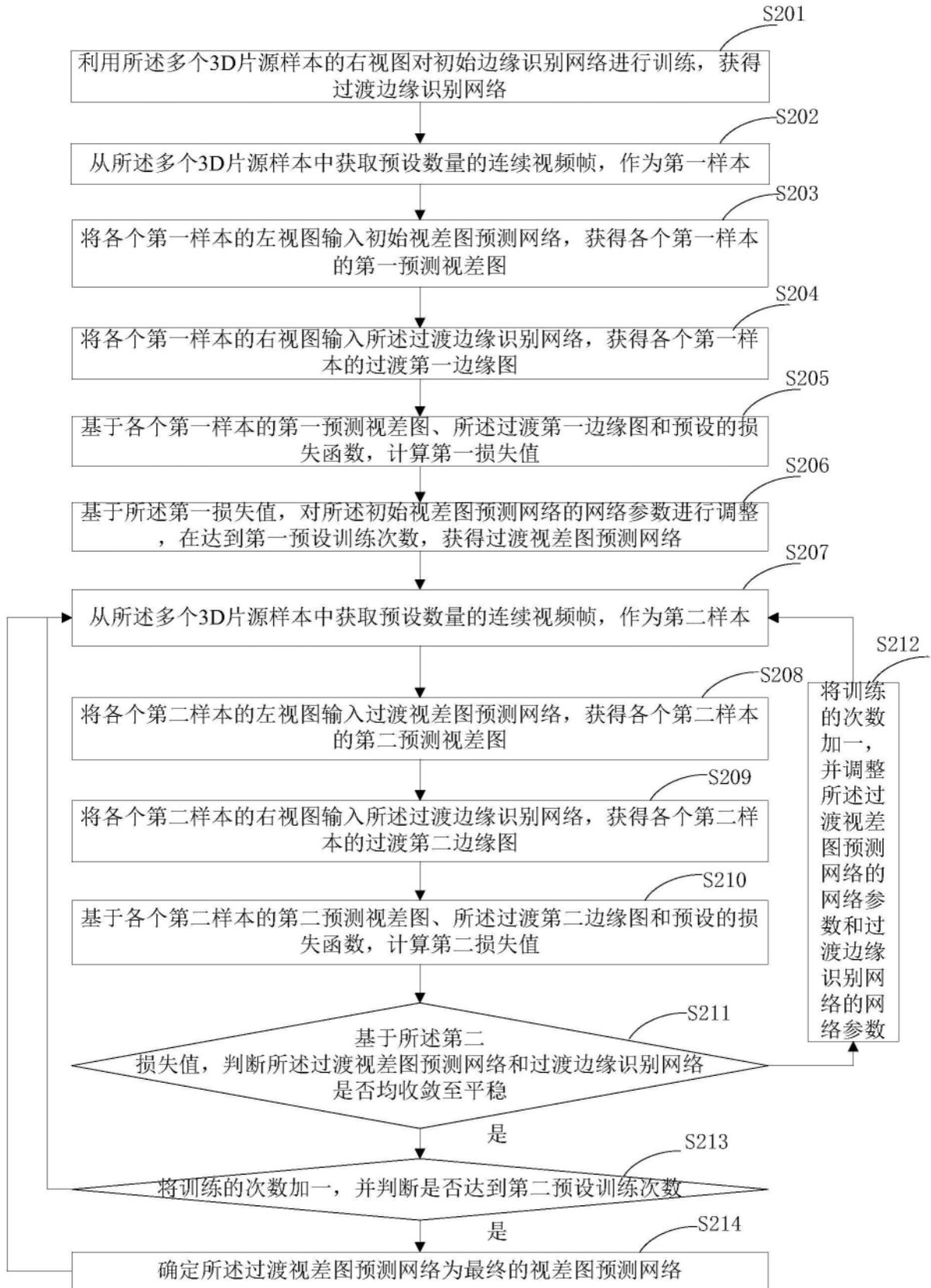


图2

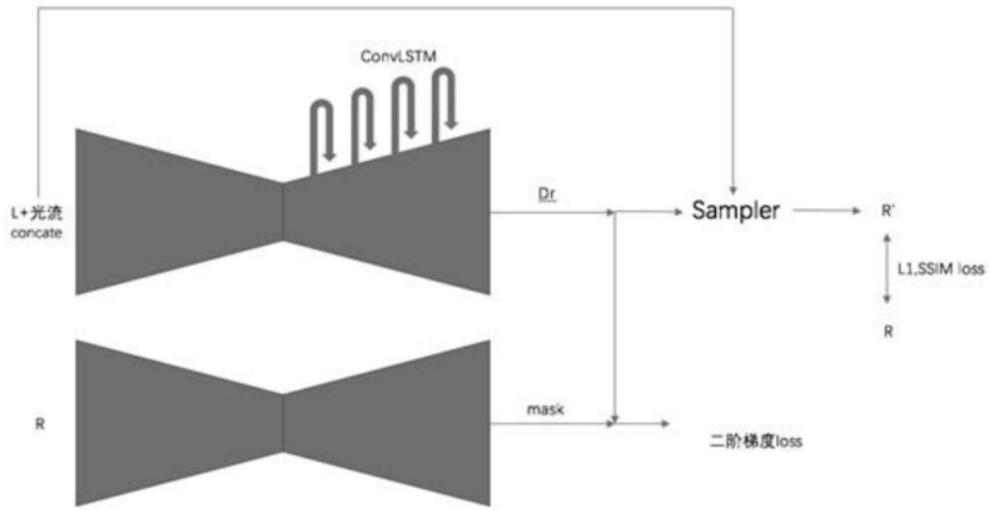


图3

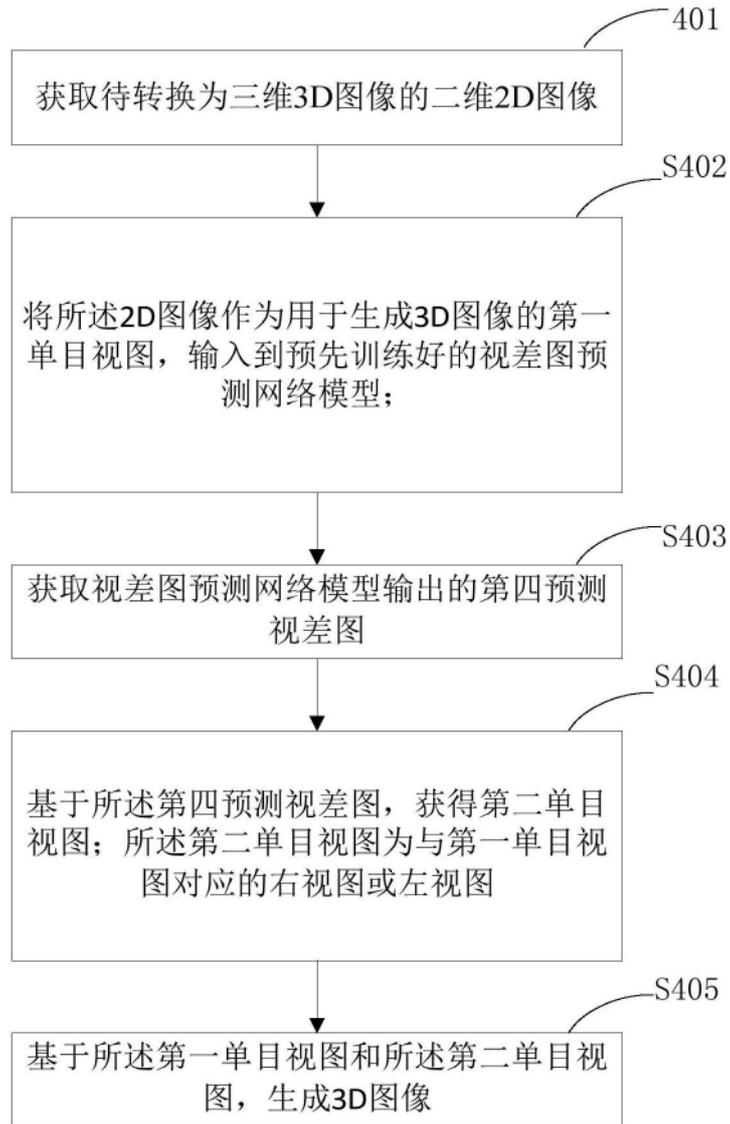


图4

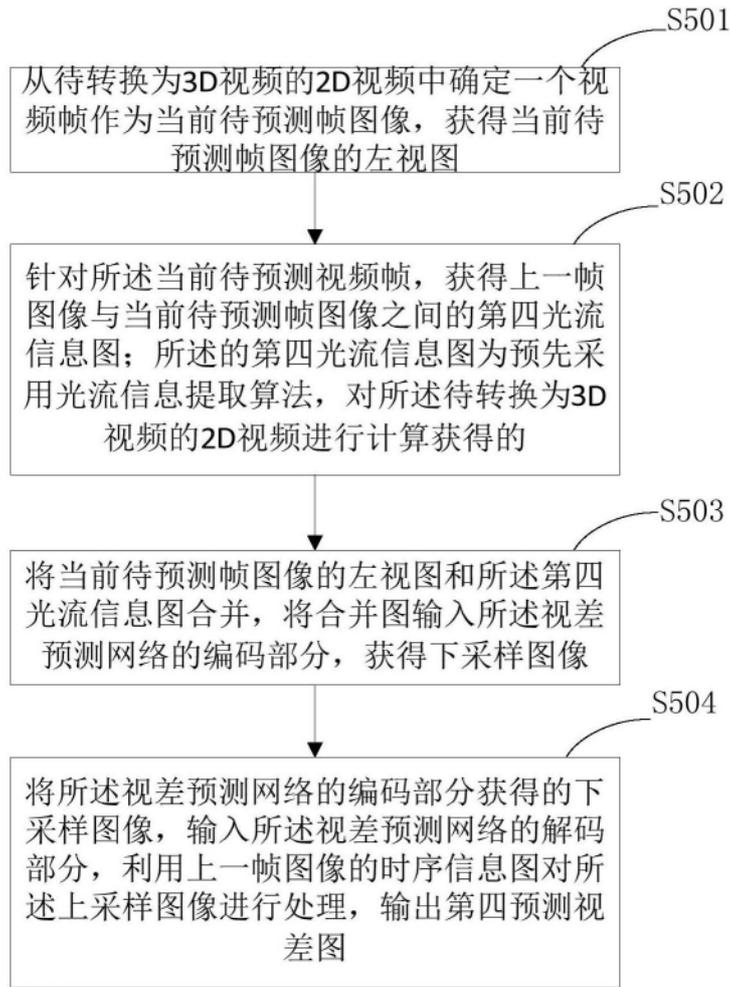


图5

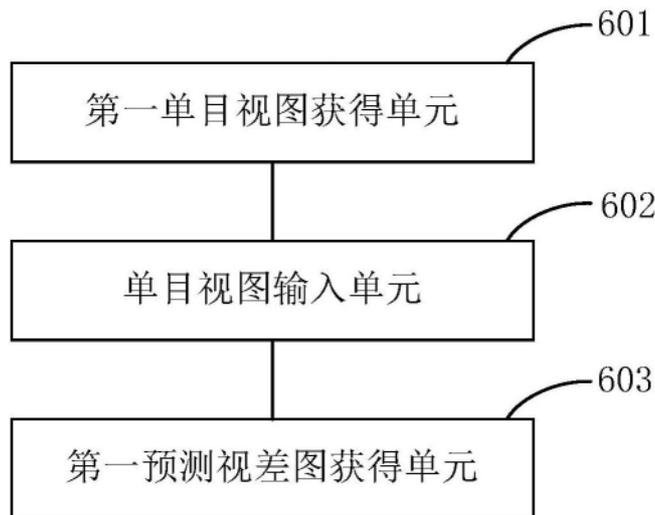


图6

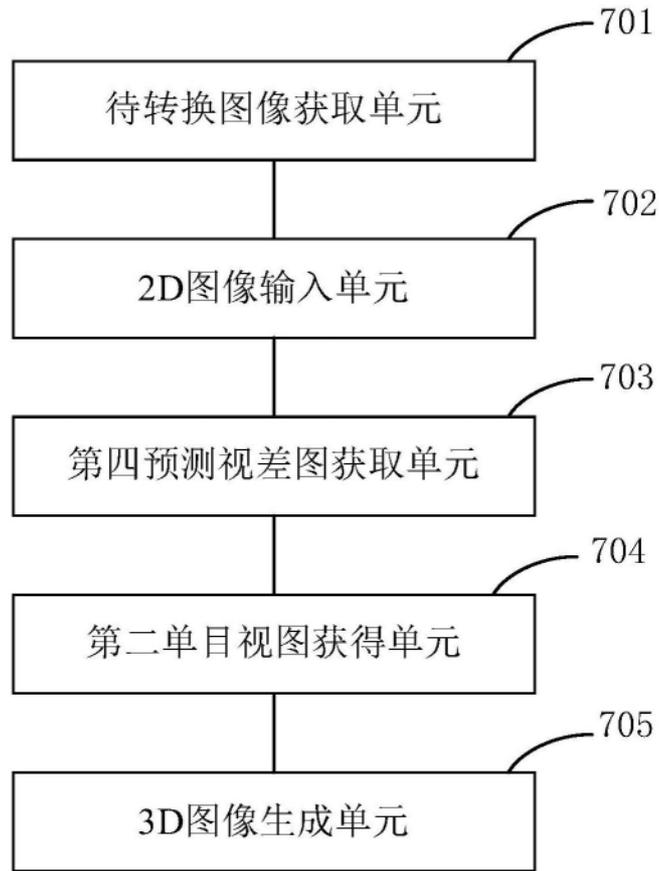


图7

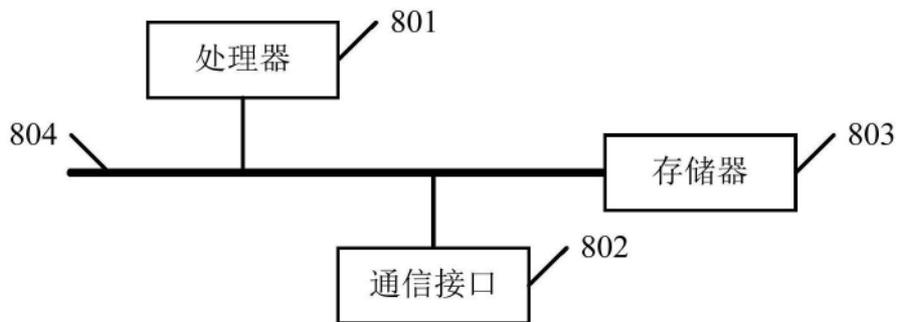


图8

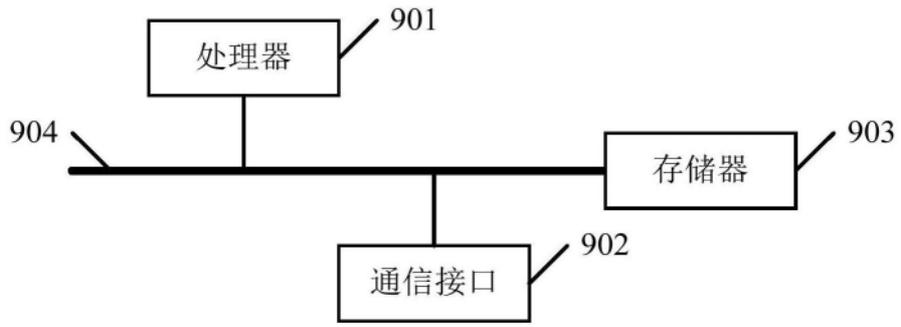


图9

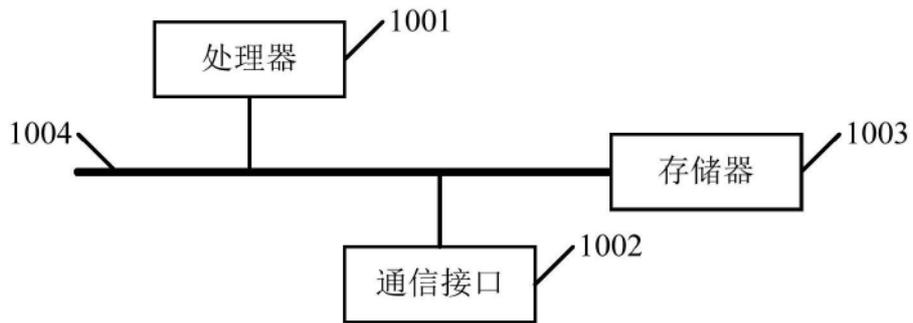


图10