



(12)发明专利申请

(10)申请公布号 CN 105979244 A

(43)申请公布日 2016.09.28

(21)申请号 201610377597.0

(22)申请日 2016.05.31

(71)申请人 十二维度(北京)科技有限公司

地址 100024 北京市朝阳区五里桥一街1号
院5号楼4层

(72)发明人 赵天奇 渠源 张阳

(74)专利代理机构 北京路浩知识产权代理有限公司 11002

代理人 李相雨

(51)Int.Cl.

H04N 13/02(2006.01)

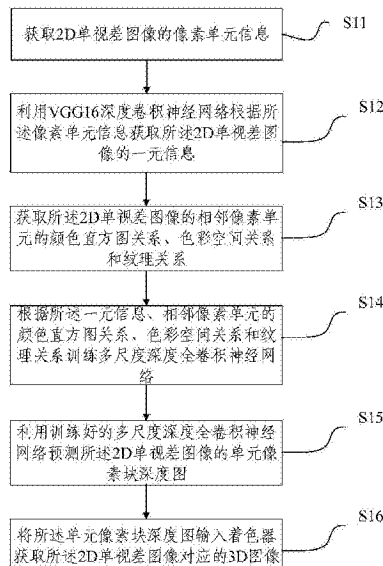
权利要求书2页 说明书8页 附图3页

(54)发明名称

一种基于深度学习的2D图像转3D图像的方法及系统

(57)摘要

本发明提供了一种基于深度学习的2D图像转3D图像的方法及系统。该方法包括：获取2D单视差图像的像素单元信息；利用VGG16深度卷积神经网络根据所述像素单元信息获取所述2D单视差图像的一元信息；获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系；根据所述一元信息、相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络；利用训练好的多尺度深度全卷积神经网络预测所述2D单视差图像的单元像素块深度图；将所述单元像素块深度图输入着色器获取所述2D单视差图像对应的3D图像。本发明避免了现有的人工由2D单视差图像生成深度图像结果不准确、成本高的缺陷，实现了自动将2D图像转换为3D图像。



1. 一种基于深度学习的2D图像转3D图像的方法,其特征在于,包括:
 - 获取2D单视差图像的像素单元信息;
 - 利用VGG16深度卷积神经网络根据所述像素单元信息获取所述2D单视差图像的一元信息;
 - 获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系;
 - 根据所述一元信息、相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络;
 - 利用训练好的多尺度深度全卷积神经网络预测所述2D单视差图像的单元像素块深度图;
 - 将所述单元像素块深度图输入着色器获取所述2D单视差图像对应的3D图像。
2. 根据权利要求1所述的基于深度学习的2D图像转3D图像的方法,其特征在于,所述获取2D单视差图像的像素单元信息包括:
 - 将所述2D单视差图像分割为多个大小相同的像素块;
 - 对各个像素块进行高斯模糊处理获取像素单元信息;
 - 其中,所述像素单元信息包括各个像素块经过高斯模糊后的中心像素的深度值。
3. 根据权利要求1所述的基于深度学习的2D图像转3D图像的方法,其特征在于,所述利用VGG16深度卷积神经网络根据所述像素单元信息获取所述2D单视差图像的一元信息包括:
 - 利用VGG16深度卷积神经网络根据所述像素单元信息获取与所述2D单视差图像对应的多维度的特征图。
4. 根据权利要求1所述的基于深度学习的2D图像转3D图像的方法,其特征在于,所述获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系包括:
 - 利用线性条件随机场获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系。
5. 根据权利要求1所述的基于深度学习的2D图像转3D图像的方法,其特征在于,所述根据所述一元信息、相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络包括:
 - 根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络的第二部分,经过前向传播、反向传播预训练所述多尺度深度全卷积神经网络的第二部分;
 - 根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络,经过前向传播、反向传播微调所述多尺度深度全卷积神经网络。
6. 一种基于深度学习的2D图像转3D图像的系统,其特征在于,包括:
 - 像素单元信息获取单元,用于获取2D单视差图像的像素单元信息;
 - 一元信息获取单元,用于利用VGG16深度卷积神经网络根据所述像素单元信息获取所述2D单视差图像的一元信息;
 - 相邻关系获取单元,用于获取所述2D单视差图像的相邻像素单元的颜色直方图关系、

色彩空间关系和纹理关系；

神经网络训练单元,用于根据所述一元信息、相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络；

单元像素块预测单元,用于利用训练好的多尺度深度全卷积神经网络预测所述2D单视差图像的单元像素块深度图；

3D图像获取单元,用于将所述单元像素块深度图输入着色器获取所述2D单视差图像对应的3D图像。

7.根据权利要求6所述的基于深度学习的2D图像转3D图像的系统,其特征在于,所述像素单元信息获取单元进一步用于:

将所述2D单视差图像分割为多个大小相同的像素块;

对各个像素块进行高斯模糊处理获取像素单元信息;

其中,所述像素单元信息包括各个像素块经过高斯模糊后的中心像素的深度值。

8.根据权利要求6所述的基于深度学习的2D图像转3D图像的系统,其特征在于,所述一元信息获取单元进一步用于:

利用VGG16深度卷积神经网络根据所述像素单元信息获取与所述2D单视差图像对应的多维度的特征图。

9.根据权利要求6所述的基于深度学习的2D图像转3D图像的系统,其特征在于,所述相邻关系获取单元进一步用于:

利用线性条件随机场获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系。

10.根据权利要求6所述的基于深度学习的2D图像转3D图像的系统,其特征在于,所述神经网络训练单元进一步用于:

根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络的第二部分,经过前向传播、反向传播预训练所述多尺度深度全卷积神经网络的第二部分;

根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络,经过前向传播、反向传播微调所述多尺度深度全卷积神经网络。

一种基于深度学习的2D图像转3D图像的方法及系统

技术领域

[0001] 本发明涉及图像处理技术领域,具体涉及一种基于深度学习的2D图像转3D图像的方法及系统。

背景技术

[0002] 近年来,三维显示技术已经成为当前社会中的热点技术,被应用在生活的各种场景中。同时,由于3D技术的兴起,随之而来的2D图像转3D图像的需求也不断增加,但是,在将2D转3D的过程中,通常采用人工处理方式进行,由人工对图像帧进行roto分割、深度信息生成、人工补图等操作,人工将原始2d图像转成深度图像,需要roto后,把图像中的一个roto圈中内容加入深度信息,这个深度信息一般是靠感觉加入,因而深度生成结果倾向于生成深度的人的主观观看感受。对于商业成果的展示,经常要做多次检查及修正。同时,人工处理图像的成本过高。

[0003] 在深度信息生成中,随着深度学习发展迅速,在各个领域都出现了令人欣喜的成果,越来越多的研究采用深度卷积神经网络(Convolutional Neural Networks,CNN)实现深度估计。传统的深度卷积神经网络如AlexNet,vggNet等都是全连接深度卷积神经网络,由卷积层和全连接层构成,通常都是将图像以标准尺寸作为输入,产生非空间排布输出。由于2D转3D目的是通过任意尺寸的单视差2D图像,自动生成3D深度信息图像,因此非空间排布的输出结果无法实现2D转3D的实际应用。

[0004] 传统深度图像信息输入的应用中,通常是以双视差图像作为输入,输出是深度信息估计结果。但是,目前的2D转3D应用中,是通过输入单视差图片和深度图像信息,进行立体显示,所以此类应用无法实现2D转3D的需求。

发明内容

[0005] 本发明所要解决的技术问题是:现有的人工由2D单视差图像生成深度图像结果不准确、成本高的问题。

[0006] 为解决上述技术问题,本发明提出了一种基于深度学习的2D图像转3D图像的方法,该基于深度学习的2D图像转3D图像的方法包括:

[0007] 获取2D单视差图像的像素单元信息;

[0008] 利用VGG16深度卷积神经网络根据所述像素单元信息获取所述2D单视差图像的一元信息;

[0009] 获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系;

[0010] 根据所述一元信息、相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络;

[0011] 利用训练好的多尺度深度全卷积神经网络预测所述2D单视差图像的单元像素块深度图;

- [0012] 将所述单元像素块深度图输入着色器获取所述2D单视差图像对应的3D图像。
- [0013] 可选地,所述获取2D单视差图像的像素单元信息包括:
- [0014] 将所述2D单视差图像分割为多个大小相同的像素块;
- [0015] 对各个像素块进行高斯模糊处理获取像素单元信息;
- [0016] 其中,所述像素单元信息包括各个像素块经过高斯模糊后的中心像素的深度值。
- [0017] 可选地,所述利用VGG16深度卷积神经网络根据所述像素单元信息获取所述2D单视差图像的一元信息包括:
- [0018] 利用VGG16深度卷积神经网络根据所述像素单元信息获取与所述2D单视差图像对应的多维度的特征图。
- [0019] 可选地,所述获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系包括:
- [0020] 利用线性条件随机场获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系。
- [0021] 可选地,所述根据所述一元信息、相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络包括:
- [0022] 根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络的第二部分,经过前向传播、反向传播预训练所述多尺度深度全卷积神经网络的第二部分;
- [0023] 根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络,经过前向传播、反向传播微调所述多尺度深度全卷积神经网络。
- [0024] 本发明还提出了一种基于深度学习的2D图像转3D图像的系统,该系统包括:
- [0025] 像素单元信息获取单元,用于获取2D单视差图像的像素单元信息;
- [0026] 一元信息获取单元,用于利用VGG16深度卷积神经网络根据所述像素单元信息获取所述2D单视差图像的一元信息;
- [0027] 相邻关系获取单元,用于获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系;
- [0028] 神经网络训练单元,用于根据所述一元信息、相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络;
- [0029] 单元像素块预测单元,用于利用训练好的多尺度深度全卷积神经网络预测所述2D单视差图像的单元像素块深度图;
- [0030] 3D图像获取单元,用于将所述单元像素块深度图输入着色器获取所述2D单视差图像对应的3D图像。
- [0031] 可选地,所述像素单元信息获取单元进一步用于:
- [0032] 将所述2D单视差图像分割为多个大小相同的像素块;
- [0033] 对各个像素块进行高斯模糊处理获取像素单元信息;
- [0034] 其中,所述像素单元信息包括各个像素块经过高斯模糊后的中心像素的深度值。
- [0035] 可选地,所述一元信息获取单元进一步用于:
- [0036] 利用VGG16深度卷积神经网络根据所述像素单元信息获取与所述2D单视差图像对

应的多维度的特征图。

[0037] 可选地,所述相邻关系获取单元进一步用于:

[0038] 利用线性条件随机场获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系。

[0039] 可选地,所述神经网络训练单元进一步用于:

[0040] 根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络的第二部分,经过前向传播、反向传播预训练所述多尺度深度全卷积神经网络的第二部分;

[0041] 根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络,经过前向传播、反向传播微调所述多尺度深度全卷积神经网络。

[0042] 本发明提供的基于深度学习的2D图像转3D图像的方法及系统,基于传统深度卷积神经网络,构建出多尺度深度全卷积神经网络,基于单视差图像作为输入,进行深度信息估计,将深度信息输入着色器获取3D图像,避免了现有的人工由2D单视差图像生成深度图像结果不准确、成本高的缺陷,实现了自动将2D图像转换为3D图像。

附图说明

[0043] 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。

[0044] 图1是本发明一个实施例的基于深度学习的2D图像转3D图像的方法的流程示意图;

[0045] 图2是本发明一个实施例的获取像素单元信息的原理图;

[0046] 图3是本发明一个实施例的获取一元信息的原理图;

[0047] 图4a是本发明一个实施例的预训练多尺度深度全卷积神经网络的效果图;

[0048] 图4b是本发明一个实施例的微调整多尺度深度全卷积神经网络的效果图;

[0049] 图5是本发明一个实施例的基于深度学习的2D图像转3D图像的系统的结构示意图。

具体实施方式

[0050] 为使本发明实施例的目的、技术方案和优点更加清楚,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、完整的描述,显然,所描述的实施例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作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。

[0051] 本发明实施例基于使用ImageNet图像数据库预训练好的VGG16模型进行开发,构造出全卷积深度神经网络,网络由七个卷积层构成,前五个卷积层直接采用VGG16模型中的前五层,添加了一些卷积层,以实现更深的全卷积神经网络。

[0052] 图1是本发明一个实施例的基于深度学习的2D图像转3D图像的方法的流程示意

图。如图1所示,本发明实施例方法包括:

[0053] S11:获取2D单视差图像的像素单元信息;

[0054] S12:利用VGG16深度卷积神经网络根据所述像素单元信息获取所述2D单视差图像的一元信息;

[0055] S13:获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系;

[0056] S14:根据所述一元信息、相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络;

[0057] S15:利用训练好的多尺度深度全卷积神经网络预测所述2D单视差图像的单元像素块深度图;

[0058] S16:将所述单元像素块深度图输入着色器获取所述2D单视差图像对应的3D图像。

[0059] 本发明实施例的基于深度学习的2D图像转3D图像的方法,基于传统深度卷积神经网络,构建出多尺度深度全卷积神经网络,基于单视差图像作为输入,进行深度信息估计,将深度信息输入着色器获取3D图像,避免了现有的人工由2D单视差图像生成深度图像结果不准确、成本高的缺陷,实现了自动将2D图像转换为3D图像。

[0060] 在本发明实施例的一种优选的实施方式中,与图1中的方法类似,步骤S11具体包括:

[0061] 将所述2D单视差图像分割为多个大小相同的像素块;

[0062] 对各个像素块进行高斯模糊处理获取像素单元信息;

[0063] 其中,所述像素单元信息包括各个像素块经过高斯模糊后的中心像素的深度值。

[0064] 图2是本发明一个实施例的获取像素单元信息的原理图。如图2所示,本发明实施例对于网络训练的目的在于推断每个像素的深度在单个图像中描绘的一个场景,因此可以假设图像由大小均匀的像素区域构成。原始2D单视差图像大小比如是 640×480 (如图2中第一幅图),按 20×20 的尺寸把原图像划分成相同大小的像素块(如图2中的第二幅图),对各个像素块进行高斯模糊处理,再取各个像素块的中心像素的深度值,组成 32×24 的图像(如图2中的第三幅图),由此得到一个方块深度值矩阵。

[0065] 图1中的步骤S12具体包括:

[0066] 利用VGG16深度卷积神经网络根据所述像素单元信息获取与所述2D单视差图像对应的多维度的特征图。

[0067] 生成一元信息的前部分网络结构(即多尺度深度全卷积神经网络的第一部分)由五个卷积层(conv)和最大池化层(Max-pooling)交替构成,如图2所示,具体流程如下:

[0068] 第一层:卷积层conv_1,卷积核大小为 3×3 ,步长stride为1;

[0069] 第二层:最大池化层,池化区间为 2×2 ,步长stride为2;

[0070] 第三层:卷积层conv_2,卷积核大小为 3×3 ,步长stride为1;

[0071] 第四层:最大池化层,池化区间为 2×2 ,步长stride为2;

[0072] 第五层:卷积层conv_3,卷积核大小为 3×3 ,步长stride为1;

[0073] 第六层:最大池化层,池化区间为 2×2 ,步长stride为2;

[0074] 第七层:卷积层conv_4,卷积核大小为 3×3 ,步长stride为1;

[0075] 第八层:最大池化层,池化区间为 2×2 ,步长stride为2;

[0076] 第九层:卷积层conv_5,卷积核大小为 3×3 ,步长stride为1;

[0077] 第十层:最大池化层,池化区间为 2×2 ,步长stride为2。

[0078] 通过前部分全卷积神经网络,将尺寸为 480×640 ,维度为3的输入图像提取为尺寸为 32×42 ,维度为512的特征图。

[0079] 图1中的步骤S13具体包括:

[0080] 利用线性条件随机场获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系。

[0081] 需要说明的是,本发明实施例将深度卷积神经网络与连续条件随机场(CRF)进行结合,解决了深度连续性问题。在对图像进行全卷积中,虽然每个像素单元都参与计算,但是相邻像素单元间的连续性并没有得到考虑。本发明考虑到马尔可夫独立性假设,其主要观点是,对一个节点,在给定他所连接的所有节点的前提下,他与外接是独立的。也就是说,一个像素单元的连续性只与和它相邻的像素单元有关系,和其他的像素单元的关系是独立的。因此,本发明基于隐马尔可夫模型的线性条件随机场(CRF)提取输入图像相邻像素单元的颜色直方图关系、luv色彩空间关系及纹理关系,作为多尺度的一个输入,解决每个像素单元和其相邻像素单元的特征连续性问题,构建出多尺度深度全卷积神经网络。

[0082] 图1中的步骤S14具体包括:

[0083] 根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络的第二部分,经过前向传播、反向传播预训练所述多尺度深度全卷积神经网络的第二部分;

[0084] 后部分网络(即多尺度深度全卷积神经网络的第二部分)的具体结构如下:

[0085] 第一层:卷积层conv_1,卷积核大小为 3×3 ,步长stride为1;

[0086] 第二层:卷积层conv_2,卷积核大小为 3×3 ,步长stride为1;

[0087] 第三层:单元像素池化层,将尺寸为 32×42 维度为512的输入向量合并成1344;

[0088] 第四层:卷积层conv_3,卷积核大小为 1×1 ,步长stride为1;

[0089] 第五层:卷积层conv_4,卷积核大小为 1×1 ,步长stride为1;

[0090] 第六层:逻辑层logistic,把激活函数作为逻辑函数,如下式:

[0091] $f(x) = (1 + e^{-x})^{-1}$

[0092] 第七层:卷积层conv_5,卷积核大小为 1×1 ,步长stride为1;

[0093] 第八层:CRF损耗层。

[0094] 根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络,经过前向传播、反向传播微调所述多尺度深度全卷积神经网络。

[0095] 本发明构建的多尺度深度全卷积神经网络基于反向传播算法,将训练分为预训练(Pretrain)和微调训练(Finetune)两个部分,在预训练中,本发明实施例首先经过VGG16的卷积层,提取输入图像的一元信息,随后将一元信息、相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系进行结合,传入后面的网络。这样在提取一元信息的网络中,不需要进行反向传递,节省了修改生成一元信息的网络的权重时间。而在微调网络中,将原图作为输入,网络进行到一元特征阶段将其他特征信息与之结合,这样,在进行反向传播中,用来生成一元信息的网络部分权重也会被修正,很好的节省了整体训练时间,提高了训练

效率。

[0096] 本发明实施例采用60291张单视差图像及深度信息图作为训练集(train),10000张单视差图像及深度信息图作为检验集(val),经过预训练20轮(如图4a),微调整训练15轮(如图4b)后,错误率(regression error)降到10以下,并没有发生过拟合现象。

[0097] 本发明实施例网络训练中采用反向传播算法,在前向传播中,网络训练的误差采用平方误差代价函数进行评定,在取得总误差后,对误差进行平均,可以得到每轮的误差率。总误差的平均值计算公式如下:

$$[0098] \quad E^N = \frac{1}{c} \cdot \frac{1}{n} \cdot \sum_{k=1}^n \sum_{i=1}^c (x_k^i - y_k^i)^2$$

[0099] 其中,c代表最终输出的像素单元总数,n代表训练的总轮数,由此 x_k^i 表示第k个像素单元在第n轮的训练后深度值, y_k^i 表示第k个像素单元输入时的深度值, E^N 表示总误差的平均值。

[0100] 因为在全部训练集上的误差只是每个训练样本的误差的总和,所以只讨论对于一个误差样本的反向传播,对于第n个样本的误差计算公式如下:

$$[0101] \quad E^n = \frac{1}{2} \sum_{i=1}^c (x_i^n - y_i^n)^2 = \frac{1}{2} \|x^n - y\|_2^2$$

[0102] 一般,为了在梯度下降过程中增加收敛性,数据集数据先进性归一化处理。根据反向传播规则计算代价损失E关于网络中每一个权值的偏导数,偏导数的计算公式如下:

$$[0103] \quad x^e = f(u^e)$$

$$[0104] \quad u^e = W^e x^{e-1} + b^e$$

[0105] 在反向传播中,反向传播回来的误差可以看作是每个神经元的基的灵敏度,即b变化多少,误差会随着变化多少,也就是误差对基的变化率,即导数,计算公式如下:

$$[0106] \quad \frac{\partial E}{\partial x} = \frac{\partial E}{\partial u} \cdot \frac{\partial u}{\partial x} = \delta$$

[0107] 通过在每一层的误差修正神经元的权重,实现权重更新。将更新后的权重保存为训练模型。由于 $\frac{\partial E}{\partial x} = 1$,所以 $\frac{\partial E}{\partial u} = \frac{\partial E}{\partial x} = \delta$,也就是说基的灵敏度和误差E对一个节点的全部输入u的导数 $\frac{\partial E}{\partial u}$ 是相等的。由于输出层每个神经元的灵敏度不同,所以用点乘来修正神经元灵敏度;反向传播计算公式如下:

$$[0108] \quad \delta^e = (W^{e+1}) \delta^{e+1} \cdot f'(u^e)$$

[0109] 最后,对每个神经元 δ 规则进行权值更新,即对给定的神经元,得到它的输入,然后用这个神经元的 δ 进行缩放。以向量形式表述为,对于第i层,误差对于该层每个权值的导数是该层的输入与该层灵敏度的叉乘。然后得到偏导数乘以一个负学习率就是该层的神经元的权值更新,对于每一个权值W都有一个特定的学习率 η ,计算公式如下:

$$[0110] \quad \frac{\partial E}{\partial W^e} = x^{e-1} (\delta^e)^T$$

$$[0111] \quad \Delta W^e = -\eta \frac{\partial E}{\partial W^e}$$

[0112] 本发明实施例利用训练网络训练出的网络模型,将原始2D图像作为输入,从网络第一层开始逐层前进,直至网络最后一层,输出深度图。由于采用了单元像素块的方式进行深度预测,预测出的深度图像信息由尺寸相等的方块像素信息构成。在实际应用中,需要将单元像素块深度图转换成平滑连续的深度图像才能进行2D转3D应用。因此,在网络预测结

束后,将单元像素块深度图输入着色器(Anat Levin研究的Colorization Using Optimization)进行转换,获得转换后的深度图像。

[0113] 图5是本发明一个实施例的基于深度学习的2D图像转3D图像的系统的结构示意图。如图5所示,本发明实施例的系统包括:像素单元信息获取单元51、一元信息获取单元52、相邻关系获取单元53、神经网络训练单元54、单元像素块预测单元55和3D图像获取单元56;具体地:

[0114] 像素单元信息获取单元51,用于获取2D单视差图像的像素单元信息;

[0115] 一元信息获取单元52,用于利用VGG16深度卷积神经网络根据所述像素单元信息获取所述2D单视差图像的一元信息;

[0116] 相邻关系获取单元53,用于获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系;

[0117] 神经网络训练单元54,用于根据所述一元信息、相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络;

[0118] 单元像素块预测单元55,用于利用训练好的多尺度深度全卷积神经网络预测所述2D单视差图像的单元像素块深度图;

[0119] 3D图像获取单元56,用于将所述单元像素块深度图输入着色器获取所述2D单视差图像对应的3D图像。

[0120] 在一种可选的实施方式中,像素单元信息获取单元51进一步用于:

[0121] 将所述2D单视差图像分割为多个大小相同的像素块;

[0122] 对各个像素块进行高斯模糊处理获取像素单元信息;

[0123] 其中,所述像素单元信息包括各个像素块经过高斯模糊后的中心像素的深度值。

[0124] 进一步地,一元信息获取单元52用于:

[0125] 利用VGG16深度卷积神经网络根据所述像素单元信息获取与所述2D单视差图像对应的多维度的特征图。

[0126] 进一步地,相邻关系获取单元53用于:

[0127] 利用线性条件随机场获取所述2D单视差图像的相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系。

[0128] 进一步地,神经网络训练单元54用于:

[0129] 根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络的第二部分,经过前向传播、反向传播预训练所述多尺度深度全卷积神经网络的第二部分;

[0130] 根据所述一元信息相邻像素单元的颜色直方图关系、色彩空间关系和纹理关系训练多尺度深度全卷积神经网络,经过前向传播、反向传播微调所述多尺度深度全卷积神经网络。

[0131] 本实施例的基于深度学习的2D图像转3D图像的系统可以用于执行上述方法实施例,其原理和技术效果类似,此处不再赘述。

[0132] 本发明提供的基于深度学习的2D图像转3D图像的方法及系统,基于传统深度卷积神经网络,构建出多尺度深度全卷积神经网络,基于单视差图像作为输入,进行深度信息估计,将深度信息输入着色器获取3D图像,避免了现有的人工由2D单视差图像生成深度图像

结果不准确、成本高的缺陷,实现了自动将2D图像转换为3D图像。

[0133] 需要说明的是术语“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。

[0134] 本发明的说明书中,说明了大量具体细节。然而能够理解的是,本发明的实施例可以在没有这些具体细节的情况下实践。在一些实例中,并未详细示出公知的方法、结构和技术,以便不模糊对本说明书的理解。类似地,应当理解,为了精简本发明公开并帮助理解各个发明方面中的一个或多个,在上面对本发明的示例性实施例的描述中,本发明的各个特征有时被一起分组到单个实施例、图、或者对其的描述中。然而,并不应将该公开的方法解释呈反映如下意图:即所要求保护的本发明要求比在每个权利要求中所明确记载的特征更多的特征。更确切地说,如权利要求书所反映的那样,发明方面在于少于前面公开的单个实施例的所有特征。因此,遵循具体实施方式的权利要求书由此明确地并入该具体实施方式,其中每个权利要求本身都作为本发明的单独实施例。

[0135] 以上实施例仅用于说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

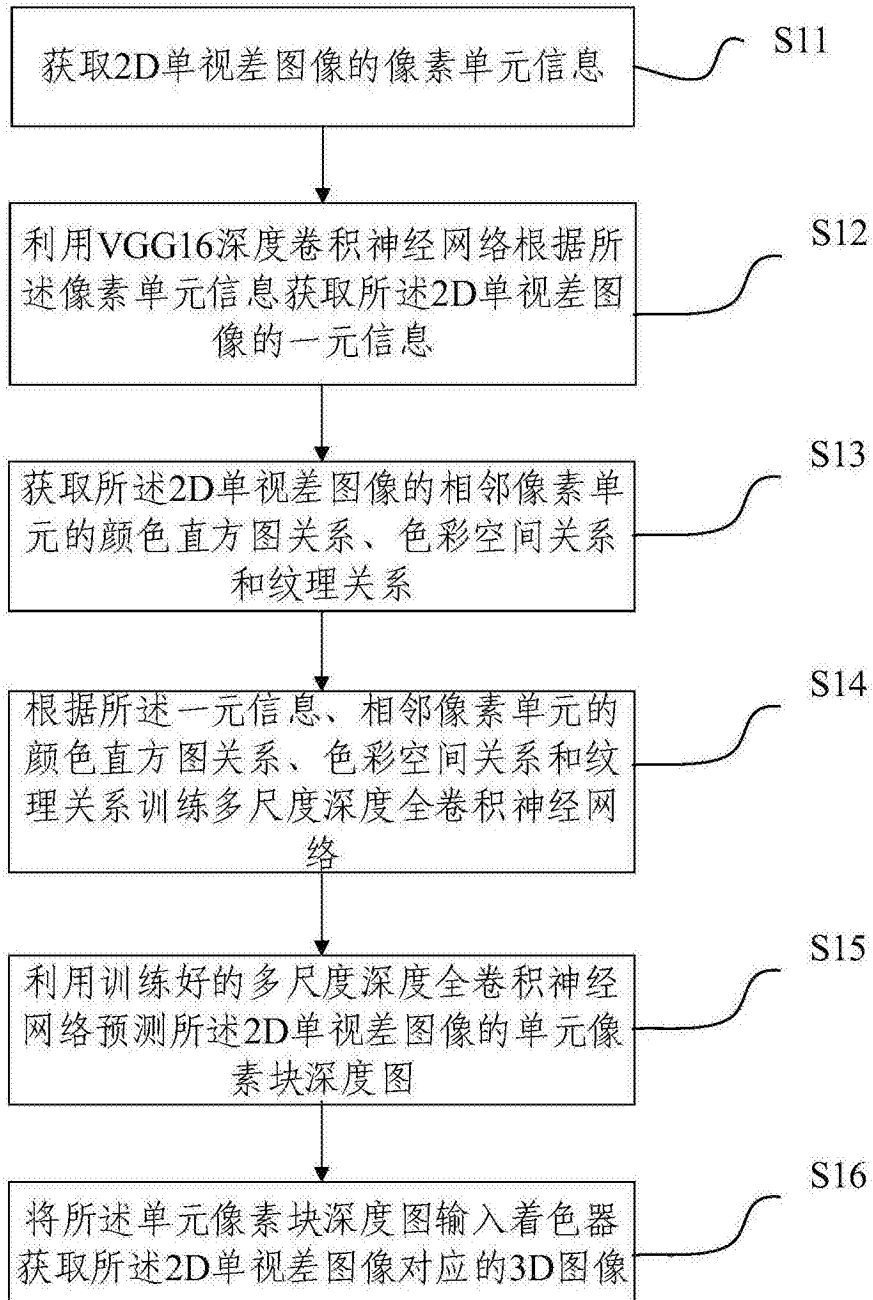


图1

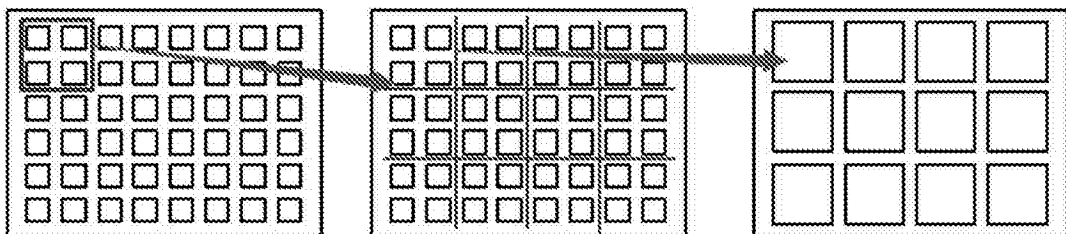


图2

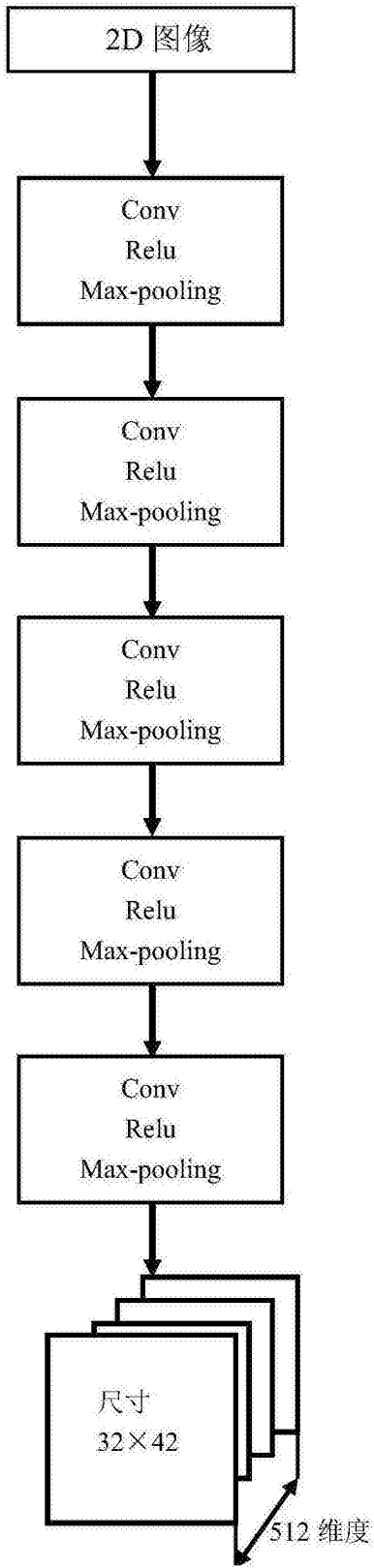


图3

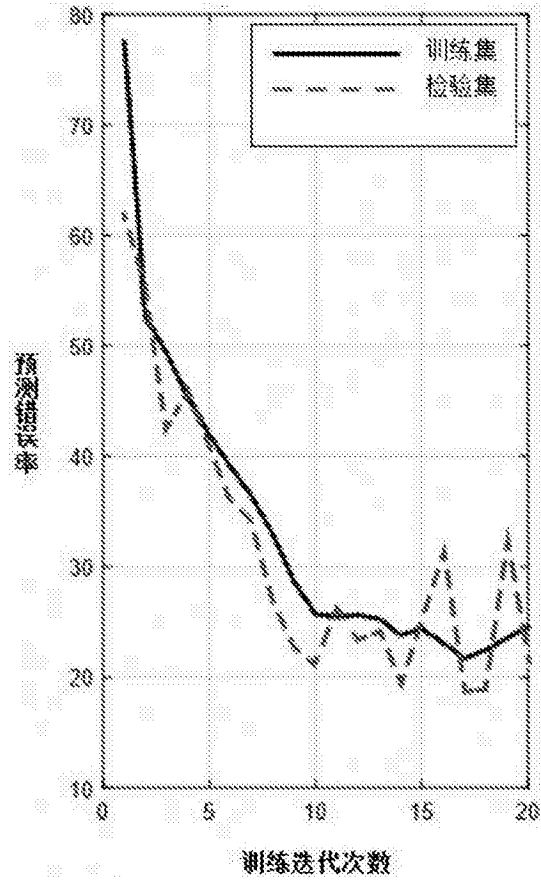


图4a

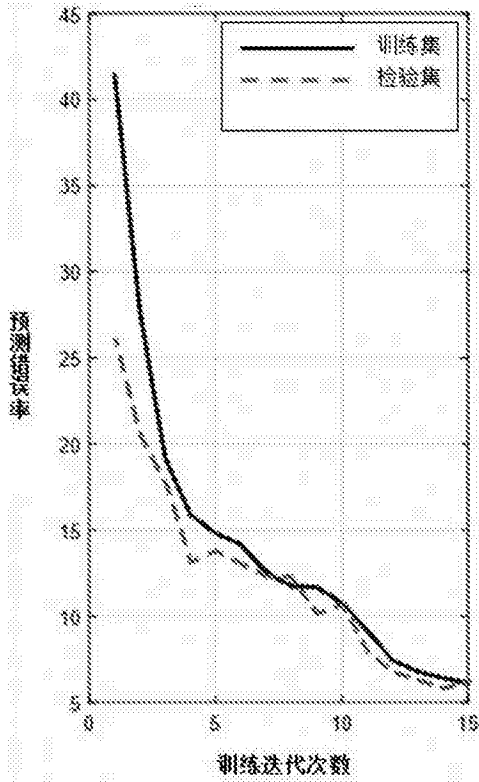


图4b

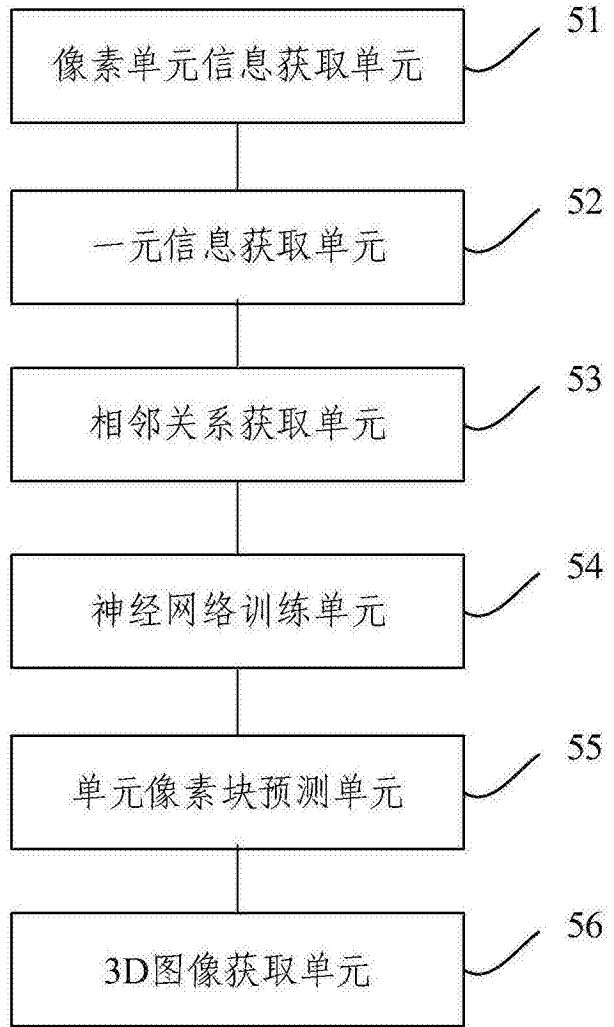


图5