(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117994313 A
(43)申请公布日 2024.05.07

(21)申请号 202311455824.3

(22)申请日 2023.11.02

(30)优先权数据
2211433 2022.11.03 FR

(71)申请人 丰达西翁必卡姆公司
地址 法国瑟松塞维涅

(72)发明人 纳比尔·马达利 安东尼·吉勒
帕特里克·焦亚 吕斯·莫林

(74)专利代理机构 北京康信知识产权代理有限责任公司 11240
专利代理师 梁丽超

(51)Int.Cl.
G06T 7/593(2017.01)
G06N 3/0464(2023.01)
G06N 3/08(2023.01)

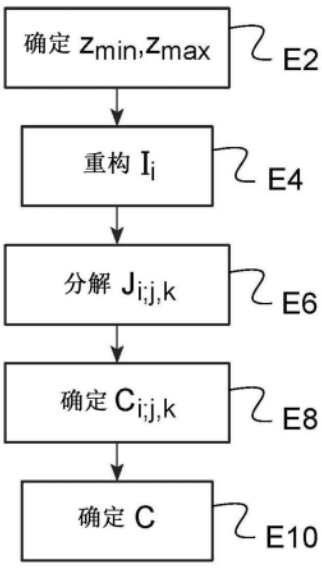
权利要求书1页 说明书8页 附图3页

(54)发明名称

用于估计深度图的方法和装置以及相关联的计算机程序

(57)摘要

本发明涉及用于估计深度图的方法和装置以及相关联的计算机程序。一种用于估计与表示场景的全息图相关联的深度图(C)的方法,包括以下步骤:重构(E4)场景的图像(I_i),每个图像与深度(z_i)相关联;将每个图像分解(E6)成彼此相邻的多个缩略图($J_{i;j,k}$),每个缩略图与所述深度相关联并且包括多个像素;通过在神经元网络的输入端提供与缩略图的所述像素相关联的值,以便在所述网络的输出端获得包括与相关的像素相关联的焦点级别的所述焦点图,来为每个缩略图确定(E8)焦点图($C_{i;j,k}$);以及根据获得的焦点级别,确定(E10)深度图(C)的每个点的深度值。本发明还涉及一种估计装置和一种相关联的计算机程序。



CN 117994313 A

1.一种用于估计与表示场景的数字全息图(H)相关联的深度图(C)的方法,所述方法包括以下步骤:

-从所述数字全息图(H)重构(E4)所述场景的多个图像(I_i),每个重构的图像(I_i)与所述场景的深度(z_i)相关联,

-将每个重构的图像(I_i)分解(E6)成多个缩略图($J_{i;j,k}$),所述多个缩略图中的所述缩略图($J_{i;j,k}$)彼此相邻,每个所述缩略图($J_{i;j,k}$)与对应于相关的重构的图像(I_i)的所述场景的所述深度(z_i)相关联,每个所述缩略图($J_{i;j,k}$)包括多个像素,

-通过在人工神经网络(NN)的输入端提供与相关的所述缩略图($J_{i;j,k}$)的所述像素相关联的值,以便在所述人工神经网络(NN)的输出端获得焦点图($C_{i;j,k}$),为每个缩略图($J_{i;j,k}$)确定(E8)焦点图($C_{i;j,k}$),所述焦点图($C_{i;j,k}$)包括与相关的所述像素相关联的焦点级别,以及

-在对包含与所述深度图(C)的点对应的像素的缩略图($J_{i;j,k}$)分别确定的焦点图($C_{i;j,k}$)中,根据对与所述深度图的所述点相关联的像素分别获得的焦点级别,为所述深度图(C)的每个点确定(E10)深度值($d_{js+q,ks+r}$)。

2.根据权利要求1所述的方法,其中,通过确定在包含所述像素的缩略图中的与深度图(C)的所述点相关联的所有像素中的最高的焦点级别对应的所述深度,对所述深度图(C)的每个点获得相关联的深度值($d_{js+q,ks+r}$)。

3.根据权利要求1或2所述的方法,其中,所述人工神经网络(NN)是卷积神经网络。

4.根据权利要求1至3中任一项所述的方法,其中,重构步骤(E4)以这样的方式实现:重构的图像(I_i)分别与在所述场景的最小深度(z_{min})和最大深度(z_{max})之间的均匀分布的深度相关联。

5.一种用于估计与表示场景的数字全息图(H)相关联的深度图(C)的装置(1),所述装置(1)包括:

-重构模块(5),从所述数字全息图(H)重构所述场景的多个图像(I_i),每个重构的图像(I_i)与所述场景的深度(z_i)相关联,

-分解模块(6),将每个重构的图像(I_i)分解成多个缩略图($J_{i;j,k}$),所述多个缩略图中的所述缩略图($J_{i;j,k}$)彼此相邻,每个所述缩略图($J_{i;j,k}$)与对应于相关的重构的图像(I_i)的所述场景的所述深度(z_i)相关联,每个缩略图($J_{i;j,k}$)包括多个像素,

-确定模块(8),通过在人工神经网络(NN)的输入端提供与相关的所述缩略图($J_{i;j,k}$)的所述像素相关联的值,以便在所述人工神经网络(NN)的输出端获得焦点图($C_{i;j,k}$),为每个缩略图($J_{i;j,k}$)确定焦点图($C_{i;j,k}$),所述焦点图($C_{i;j,k}$)包括与相关的所述像素相关联的焦点级别,以及

-模块(9),用于根据与在包含所述像素的缩略图($J_{i;j,k}$)中的深度图(C)的点相关联的像素分别获得的焦点级别,为所述深度图(C)的每个点确定深度值($d_{js+q,ks+r}$)。

6.一种计算机程序,包括能够由处理器(2)执行的指令,并且所述计算机程序被设计为当这些指令由所述处理器(2)执行时,实现根据权利要求1至4中任一项所述的方法。

用于估计深度图的方法和装置以及相关联的计算机程序

技术领域

[0001] 本发明涉及数字全息的技术领域。

[0002] 尤其涉及一种用于估计与表示场景的数字全息图相关联的深度图的方法和装置。还涉及一种相关联的计算机程序。

背景技术

[0003] 数字全息是一种沉浸式技术,该技术记录由三维场景中存在的物体所衍射的波的特征,以便再现该物体的三维图像。随后,获得的数字全息图包含了能够描述这种三维场景的所有信息。

[0004] 然而,从数字全息图本身提取该信息以便重构三维场景并不简单,因为数字全息图的平面不包括三维场景中物体的空间位置。

[0005] 众所周知,使用关于全息图平面的深度确定方法(或“来自焦点的深度”,根据常用的盎格鲁-撒克逊起源的命名方式),以便从全息图中确定场景的几何信息。这种方法的示例见Grossmann Pavel的文章“来自焦点的深度(Depth from focus)”,Pattern Recognit.Lett.5,63-69,1987。

[0006] 在这种方法中,从在预定间隔内选择的不同焦距处计算的几个全息重构平面中获得重构体积。从这些重构平面中,通过在这些平面的每个平面上应用聚焦算子来估计相关联的深度,以便为每个像素选择焦点最佳的重构平面的深度。

[0007] 然而,这种方法需要很长时间来实现,并且在计算资源方面非常昂贵。

发明内容

[0008] 在这种情况下,本发明提出改进相对于与三维场景相关联的数字全息图的平面确定深度值。

[0009] 更具体地,根据本发明,我们提出了一种用于估计与表示场景的数字全息图相关联的深度图的方法,该方法包括以下步骤:

[0010] -从数字全息图重构场景的多个图像,每个重构的图像与场景的深度相关联,

[0011] -将每个重构的图像分解成多个缩略图,所述多个缩略图中的所述缩略图彼此相邻,每个缩略图与对应于涉及重构的图像的场景的深度相关联,每个缩略图包括多个像素,

[0012] -通过在人工神经网络的输入端提供与涉及缩略图的所述像素相关联的值,以便在人工神经网络的输出端获得包括与涉及像素相关联的焦点级别的所述焦点图,为每个缩略图确定焦点图,以及

[0013] -根据在聚焦图中为包含对应于深度图的所述点的像素的缩略图分别确定的、对与深度图的所述点相关联的像素分别获得的焦点级别,为深度图的每个点确定深度值。

[0014] 因此,经处理的不同的缩略图彼此独立,因为缩略图由不相交的像素的组构成。然后,这种独立性允许更快地实现该方法。此外,由于要分析的区域数量有限(这意味着不同的缩略图),所需的计算资源也更便宜。

[0015] 此外,人工神经网络的使用允许更快地处理所有缩略图,并且更精确地确定与缩略图的像素相关联的焦点级别。

[0016] 根据本发明的方法具有以下其他非限制性的有利特征(单独或以所有技术上可能的组合形式):

[0017] -对深度图的每个点,通过确定包含所述像素的缩略图中的与深度图的所述点相关联的所有像素中的与最高焦点级别对应的深度,获得相关联的深度值;

[0018] -人工神经网络是卷积神经网络;以及

[0019] -重构步骤以这样的方式实现:重构的图像分别与在场景的最小深度和最大深度之间的均匀分布的深度相关联。

[0020] 本发明还涉及一种用于估计与表示场景的数字全息图相关联的深度图的装置,该装置包括:

[0021] -用于从所述数字全息图重构所述场景的多个图像的模块,每个重构的图像与所述场景的深度相关联,

[0022] -用于将每个重构的图像分解成多个缩略图的模块,所述多个缩略图中的所述缩略图彼此相邻,每个缩略图与对应于涉及重构的图像的所述场景的所述深度相关联,每个缩略图包括多个像素,

[0023] -用于为每个缩略图确定焦点图的模块,通过在人工神经网络的输入端提供与涉及所述缩略图的所述像素相关联的值,以便在所述人工神经网络的输出端获得包括与涉及所述像素相关联的焦点级别的所述焦点图,以及

[0024] -用于确定所述深度图的每个点的深度值的模块,根据为包含所述像素的缩略图中的深度图的所述点相关联的像素分别获得的焦点级别。

[0025] 本发明最后涉及一种计算机程序,该计算机程序包括可由处理器执行的指令并且被设计为当这些指令由处理器执行时实现如前所述的方法。

[0026] 当然,本发明的不同特征、变型和实施方式,只要不是不相容的或彼此排斥的,可以以各种组合彼此相关联。

附图说明

[0027] 此外,本发明的各种其他特征从参考附图所做的附加描述中显现出来,附图示出了本发明实施方式的非限制性形式,其中:

[0028] 图1以功能形式表示用于估计深度图的装置,该装置被设计为实现根据本发明的用于估计深度图的方法;

[0029] 图2表示与根据本发明估计的深度图相关联的数字全息图的示例;

[0030] 图3是在根据本发明的估计深度图的方法期间实现的人工神经网络(或人工神经元网络)的架构示例的示意图;以及

[0031] 图4以流程图的形式表示根据本发明的用于估计深度图的方法的示例。

具体实施方式

[0032] 应当注意,在这些附图中,不同变型共有的结构和/或功能元件可以具有相同的附图标记。

[0033] 图1以功能形式表示用于从数字全息图H估计深度图C的装置1(以下也称为装置1)。

[0034] 数字全息图H表示给定的三维场景。这个三维场景包括例如一个或多个物体。三维场景定义在标志(0,x,y,z)中。

[0035] 如图2所示,数字全息图H由(x,y)平面中的像素的矩阵定义。z轴(称作z深度轴)与数字全息图H的(x,y)平面正交。如图2所示,在此,数字全息图H定义在方程平面z＝0中。

[0036] 例如,数字全息图H在此具有1024×1024像素的尺寸。

[0037] 用于估计深度图C的装置1被设计为估计与数字全息图H相关联的深度图C。为此,装置1包括处理器2和存储装置4。存储装置4例如是硬盘或存储器。

[0038] 装置1还包括一组功能模块。该装置包括例如重构模块5、分解模块6、用于确定焦点图$C_{i;j,k}$(或聚焦图)的模块8和用于确定深度值$d_{js+q,ks+r}$的模块9。

[0039] 所描述的不同模块中的每个模块都例如通过计算机程序指令来实现,这些指令被设计为当这些指令由用于估计深度图C的装置1的处理器2执行时实现相关模块。

[0040] 然而,作为变型,前述模块中的至少一个模块可以通过专用电子电路来实现,例如,具有特定应用的集成电路。

[0041] 处理器2还被设计为实现人工神经网络NN,人工神经网络NN参与估计与数字全息图H相关联的深度图C的过程。

[0042] 这种人工神经网络NN的架构的示例如图3中所示。在这个示例中,人工神经网络NN是例如U-Net类型的卷积神经网络。

[0043] 总体而言,这种人工神经网络NN包括根据不同级别分布的多个卷积层,如下所述并如图3所示。有关U-Net类型人工神经网络的更多详情,可见Ronneberger,O.、Fischer,P.和Brox,T.的文章“U-Net:用于生物医学图像分割的卷积网络(U-Net:Convolutional Networks for Biomedical Image Segmentation)”,CoRR,abs/1505.04597,2015年。

[0044] 为了描述人工神经网络NN的架构,我们在此考虑在该人工神经元网络NN的输入端提供图像I_e。实际上,该图像I_e是从数字全息图H得到的图像,如随后将解释的。

[0045] 如图3所示,此处的人工神经网络NN包括第一部分10、连接桥20和第二部分30。

[0046] 第一部分10是所谓的压缩部分。总体而言,该第一部分10具有编码器功能,并且使得可以减小在输入端提供的图像的尺寸,同时保持(保存)其特征。为此,包括四个级别,在此是12、14、16、18。每个级别12、14、16、18包括卷积块Conv和二次采样块D。

[0047] 卷积块Conv包括至少一个卷积层,其核心是大小为n×n的矩阵。优选地,在此,每个卷积块具有两个连续的卷积层。在此,每个卷积层都有内核,内核具有3×3大小的矩阵。

[0048] 然后,卷积层(如果有若干个,则为或多个卷积层)之后是校正线性单元类型的激活功能(或表示“校正线性单元”的ReLu,根据常用的盎格鲁-撒克逊起源的命名方式)。最后,对于应用激活功能后获得的结果,卷积块Conv包括所谓的批归一化。在此,该批由在输入端提供给人工神经网络NN的多个图像组成(或构成)。在如下所述的人工神经网络的学习步骤期间,批尺寸大于或等于1(这说明在输入端提供至少两个图像,以允许训练人工神经元网络)。在随后描述的深度图估计方法的情况下,批尺寸例如在此由重构的图像I_i的数量给出(见下文)。

[0049] 如图3所示,在卷积块Conv的输出端,每个级别12、14、16、18包括二次采样块D。该

二次采样块D使得可以减少在卷积块Conv的输出端获得的结果的维度。这涉及例如将这些维度减少2,例如,通过在尺寸为2×2的像素窗口的四个像素中选择最大像素值(然后称之为“最大池化2×2”,根据常用的盎格鲁-撒克逊起源的命名方式)。

[0050] 因此,以图3所示的作为示例,输入图像I_e在输入端提供给第一部分10的第一级别12。该第一级别12的卷积块Conv和二次采样块D使得可以在输出端获得第一数据$X^{0,0}$。在此,与输入图像I_e相比,该第一数据$X^{0,0}$的维度例如减半。

[0051] 然后,该第一数据$X^{0,0}$作为输入提供给第一部分10的第二级别14,以便在第二级别14的输出端获得第二数据$X^{1,0}$。在此,与第一数据$X^{0,0}$相比,该第二数据$X^{1,0}$的维度例如减半。

[0052] 第二数据$X^{1,0}$作为输入提供给第一部分10的第三级别16,以便在第三级别16的输出端获得第三数据$X^{2,0}$。在此,与第二数据$X^{1,0}$相比,该第三数据$X^{2,0}$的维度例如减半。

[0053] 然后,该第三数据$X^{2,0}$作为输入提供给第一部分10的第四级别18,以便在输出端获得第四数据$X^{3,0}$。在此,与第三数据$X^{2,0}$相比,该第四数据$X^{3,0}$的维度例如减半。

[0054] 因此,人工神经网络NN的第一部分10对输入图像I_e的处理操作可以用以下形式表示:

[0055] $X^{i,j}=D(Conv(X^{i-1,j}))$

[0056] 其中,j=0,i在1和4之间,算子Conv对应于由卷积块Conv实现的处理,并且算子D与二次采样块D相关联。

[0057] 如图3所示,人工神经网络NN在第一部分10的输出端包括连接桥20。该连接桥20使得可以在人工神经网络NN的第一部分10和第二部分30之间建立链接。该连接桥包括如前所述的卷积块Conv。在此,因此接收第四数据$X^{3,0}$作为输入,并提供第五数据$X^{4,0}$作为输出。

[0058] 人工神经网络NN的第二部分30被称为扩展。总体而言,该第二部分30具有解码器功能,并且使得可以形成具有在输入端提供的图像尺寸的图像,并且该图像仅包含处理所必需的特征。

[0059] 为此,第二部分30在此包括四个级别32、34、36、38。与第一部分10类似,我们将第二部分30的第一级别38定义为位于与第一部分10的第一级别12相同的级别上。第二部分30的第二级别36位于与人工神经网络NN的第一部分10的第二级别14相同的级别上。第二部分30的第三级别34位于与人工神经网络NN的第一部分10的第三级别16相同的级别上。最后,第二部分30的第四级别32位于与人工神经网络NN的第一部分10的第四级别18相同的级别上。该定义用于匹配处理相同维度的数据的人工神经网络的级别。

[0060] 每个级别32、34、36、38包括过采样块U、级联块Conc和卷积块Conv(诸如先前在第一部分中介绍的)。

[0061] 每个过采样块U旨在增加在输入端接收的数据的维度。根据常用的盎格鲁-撒克逊起源的命名方式,这是一种“放大”操作。例如,在此,维度乘以2。

[0062] 在过采样块U之后,每个级别32、34、36、38包括级联块Conc。级联块旨在将在相关级别的过采样块U的输出端获得的数据与在人工神经网络NN的第一部分10的级别12、14、16、18中的一个级别的输出端获得的相同尺寸的数据级联。在级联操作中涉及来自人工神经网络NN的第一部分的数据在图3中以虚线示出。

[0063] 该级联块然后允许也在第二部分30中传输在人工神经网络NN的第一部分10中获得的提取的高频的信息。如果没有该级联块Conc,在人工神经网络NN中存在的欠采样和过

采样的多重操作之后,该信息可能会丢失。

[0064] 然后,在级联块Conc的输出端,第二部分30的每个级别32、34、36、38包括卷积块Conv,诸如先前在人工神经网络NN的第一部分10中描述的那样。在此,每个卷积块Conv特别包括至少一个卷积层,其后是校正线性单元类型的激活功能和批归一化操作。

[0065] 基于图3所示的示例,在第二部分30的第四级别32的输入端提供第五数据$X^{4,0}$。然后,过采样块U使得可以在输出端获得第一中间数据X^{int1},该第一中间数据X^{int1}具有与在第一部分10的第四级别18的输出端获得的第四数据$X^{3,0}$相同的维度。该第一中间数据X^{int1}和第四数据$X^{3,0}$然后由级联块Conc级联。然后,在级联块Conc的输出端获得的结果作为输入提供给卷积块Conv,以便在输出端获得第六数据项$X^{3,1}$。

[0066] 然后,在第二部分30的第三级别34的输入端(特别是在过采样块U的输入端)提供第六数据项$X^{3,1}$。在过采样块U的输出端,获得第二中间数据X^{int2},该第二中间数据X^{int2}具有与第三数据$X^{2,0}$相同的维度。第二中间数据X^{int2}和第三数据$X^{2,0}$由级联块Conc级联。在卷积块Conv的输入端提供在级联块Conc的输出端获得的结果,以便在输出端获得第七数据项$X^{2,2}$。

[0067] 然后,如图3所示,在第二部分30的第二级别36的输入端(因此在该第二级别36的过采样块U的输入端)提供第七数据$X^{2,2}$。在该过采样块U的输出端获得第三中间数据X^{int3}。该第三中间数据X^{int3}具有与第二数据$X^{1,0}$相同的维度。然后,第三中间数据X^{int3}和第二数据$X^{1,0}$由级联块Conc级联。在卷积块Conv的输入端提供在级联块Conc的输出端获得的结果,以便在输出端获得第八数据项$X^{1,3}$。

[0068] 然后,在第二部分30的第一级别38的输入端提供该第八数据$X^{1,3}$。然后,过采样块U可获得第四数据X^{int4}。第四数据具有与第一数据$X^{0,0}$相同的维度。然后,第四中间数据X^{int4}和第一数据$X^{0,0}$由级联块Conc级联。在卷积块Conv的输入端提供在级联块Conc的输出端获得的结果,以便在输出端获得最终数据$X^{0.4}$。该最终数据$X^{0.4}$具有与输入图像I_e相同的维度和相同的分辨率。实际上,在此,该最终数据$X^{0.4}$例如与如下所述的焦点图(也称为聚焦图)相关联。

[0069] 因此,人工神经网络NN的第二部分30对第五数据项$X^{4,0}$的处理操作可以用以下形式表示:

[0070] $X^{i,j}=Conv(Conc[X^{i,0};U(X^{i+1,j-1})])$

[0071] 其中,$j\geq1$并且i在0和3之间,算子Conv对应于由卷积块Conv实现的处理,算子Conc对应于由级联块Conc实现的处理,算子U与过采样块U相关联。

[0072] 图4是表示在上述背景下实现的用于估计与数字全息图H相关联的深度图C的方法(或过程)的示例的流程图。该方法例如由处理器2实现。总体而言,该过程是由计算机实现的。

[0073] 如图4所示,该方法开始于步骤E2,在此期间,处理器2确定数字全息图H的三维场景中z坐标的最小深度z_{min}和最大深度z_{max}。这些最小和最大深度例如预先记录在存储装置4中。

[0074] 该方法然后继续步骤E4,用于重构由数字全息图H表示的三维场景的多个二维图像。

[0075] 为此,重构模块5被配置为通过数字全息图H重构场景的n个图像I_i,其中,i是范围

1到n内的整数。

[0076] 在垂直于数字全息图H的深度轴的重构平面中定义每个重构图像I_i。换而言之,每个重构平面垂直于深度轴z。每个重构平面与深度值相关联,使得可以将深度z_i与每个重构图像I_i相关联,索引i是指重构图像I_i的索引。每个深度值定义了数字全息图平面和相关重构平面之间的距离。

[0077] 优选地,在此,重构步骤E4以这样的方式实现:使得与重构图像I_i相关联的深度z_i在最小深度z_{min}和最大深度z_{max}之间均匀分布。换而言之,重构图像I_i沿着深度轴均匀分布在最小深度z_{min}和最大深度z_{max}之间。因此,第一重构图像I_1与数字全息图H的平面间隔最小深度z_{min},而最后重构图像I_n与数字全息图H的平面间隔最大深度z_{max}。

[0078] 与重构图像I_i相关联的重构平面例如以距离z_e两两隔开。每个重构平面之间的距离z_e例如大约为50微米(μm)。

[0079] 优选地,使用由以下公式定义的角度谱的传播来计算在重构步骤E4中获得的n个图像:

[0080] $$I_i(x,y) = F^{-1}\left\{F(H)e^{j2\pi z_i\sqrt{\lambda^{-2}-f_x^2-f_y^2}}\right\}(x,y)$$

[0081] 其中,F和F^{-1}分别对应于傅立叶正变换和傅立叶逆变换,并且f_x和f_y是在傅立叶域中的数字全息图H在数字全息图的第一空间方向x和第二空间方向y上的频率坐标,λ是数字全息图H的采集波长,i是重构图像I的索引,i的范围从1到n,z_i是图像I_i的重构平面中给出的深度。

[0082] 每个重构图像I_i由多个像素定义。优选地,重构图像由与数字全息图H一样多的像素形成。因此,重构图像I_i和数字全息图H具有相同的尺寸。例如,在数字全息图H的尺寸为1024×1024的情况下,每个重构图像I_i也具有1024×1024的尺寸。

[0083] 如图4所示,该方法在步骤E6中继续。在该步骤中,分解模块6被配置为将在步骤E4中获得的每个重构图像I_i分解成多个缩略图$J_{i;j,k}$。换而言之,在该分解步骤E6期间,每个重构图像I_i被分成多个缩略图$J_{i;j,k}$。仍然换而言之,每个缩略图$J_{i;j,k}$对应于相关重构图像I_i的子部分。

[0084] 每个缩略图$J_{i;j,k}$由以下公式定义:

[0085] $J_{i;j,k}=\{|I_i|(j.s:(j+1).s,k.s:(k+1).s)\}$

[0086] 其中,$j = 1 \dots \left\lfloor\frac{s_W}{s}\right\rfloor$并且$k = 1 \dots \left\lfloor\frac{s_H}{s}\right\rfloor$,$s_W$和$s_H$是重构图像$I_i$的维度(分别是高度和宽度),s是缩略图$J_{i;j,k}$的尺寸;$|x_1|$是对应于数据$x_1$和$\lfloor x_2 \rfloor$的模块的符号,该符号对应于数字$x_2$的较低整数部分。符号$y_1:y_2$说明,对于相关变量,缩略图$J_{i;j,k}$在像素$y_1$和像素$y_2$之间定义。换而言之,在此,前面的公式根据维度x在重构图像I_i的像素js和(j+1)s之间以及根据维度y在重构图像I_i的像素ks和(k+1)s之间定义了缩略图$J_{i;j,k}$。

[0087] 每个缩略图$J_{i;j,k}$包括多个像素。该多个像素对应于相关联的重构图像I_i的像素的一部分。

[0088] 在此,缩略图$J_{i;j,k}$彼此相邻。实际上,每个缩略图$J_{i;j,k}$由重构图像I_i的一组连续像素形成。在此,重构图像I_i的像素组(分别形成缩略图$J_{i;j,k}$中的每个缩略图)不相交。换而言

之,这说明与重构图像I_i相关联的缩略图$J_{i;j,k}$彼此不重叠。因此,每个缩略图$J_{i;j,k}$包括不属于与同一重构图像I_i相关联的其他缩略图的像素。换言之,与重构图像I_i相关联的缩略图$J_{i;j,k}$彼此独立。

[0089] 缩略图之间的这种独立性的特性对于根据本发明的方法特别有利,因为这种独立性的特性允许更快的实现。此外,由于要分析的区域数量有限(即不同的缩略图),所需的计算资源也更便宜。

[0090] 因为每个缩略图$J_{i;j,k}$是从与深度z_i相关联的重构图像I_i中导出的,所以每个缩略图$J_{i;j,k}$也与(三维场景的)这个相同的深度z_i相关联。

[0091] 在数字全息图H具有1024×1024的尺寸的情况下,每个缩略图$J_{i;j,k}$可以例如具有32×32的尺寸。

[0092] 在数字全息图H具有$l_H \times l_W$的尺寸的情况下,每个缩略图$J_{i;j,k}$具有$(32 \times s_H) \times (32 \times s_W)$的尺寸,其中,$s_H = l_H/1024$,$s_W = l_W/1024$。

[0093] 缩略图尺寸的这种定义使得可以确保这些缩略图的尺寸适合于数字全息图H的尺寸,从而提高用于估计与数字全息图H相关联的深度图的方法的实现速度。

[0094] 如图4所示,该方法然后在步骤E8中继续。在该步骤期间,处理器2为每个缩略图$J_{i;j,k}$确定焦点图$C_{i;j,k}$(或聚焦图)。该焦点图$C_{i;j,k}$包括多个元素(每个元素由指数js+q、ks+r标识)。聚焦图$C_{i;j,k}$的每个元素与相关的缩略图$J_{i;j,k}$的像素相关联。

[0095] 此处,焦点图$C_{i;j,k}$的每个元素对应于焦点级别(对应于与缩略图$J_{i;j,k}$中的相关像素相关联的焦点级别)。换而言之,与焦点的级别相关的聚焦图$C_{i;j,k}$与缩略图$J_{i;j,k}$中的每个像素相关联。

[0096] 实际上,该步骤E8是经由人工神经网络NN实现的。在输入端,人工神经网络接收每个缩略图$J_{i;j,k}$,并在输出端提供与相关的缩略图$J_{i;j,k}$中的每个像素相关联的焦点级别(也称为聚焦级别)。

[0097] 更具体地,人工神经网络NN在输入端接收缩略图$J_{i;j,k}$的每个像素,并在输出端提供相关联的焦点级别(或聚焦级别)。该聚焦级别例如包括在0和1之间,并且相当于与相关像素相关联的清晰度级别。例如,在像素模糊的情况下,焦点级别接近0,而在像素明显清晰的情况下,焦点级别接近1。

[0098] 有利地,使用人工神经网络允许更快地处理所有缩略图,并且更精确地确定与缩略图的像素相关联的聚焦级别。

[0099] 在实现估计方法之前,学习步骤(图中未示出)允许训练人工神经网络NN。为此,例如,使用计算机计算的全息图。对于这些计算的全息图,场景的精确几何形状(并且因此相关的深度图)是已知的。一组基本图像来自这些计算的全息图。

[0100] 对于该组中的每个基本图像,每个像素都与焦点级别相关联。实际上,对于每个基本图像的每个像素,如果深度图中的对应像素等于相关联的深度,则焦点级别等于1。否则,焦点级别为0。

[0101] 然后,训练步骤包括调整包含在前述不同卷积块中的不同卷积层的节点的权重,以便最小化在人工神经网络NN的输出端获得的聚焦级别(当在该网络的输入端提供基本图像时)和从已知深度图确定的聚焦级别之间的误差。例如,此处可以使用交叉熵损失方法,以便最小化在人工神经网络NN的输出端获得的焦点级别(当在该网络的输入端提供基本图

像时)和从已知深度图确定的焦点级别之间的距离。

[0102] 换而言之,调整不同卷积层的节点的权重,以便将在人工神经网络NN的输出端获得的焦点级别向从已知深度图确定的焦点级别收敛。

[0103] 实际上,人工神经网络NN在输入端接收与每个重构图像I_i相关联的所有缩略图$J_{i;j,k}$,并继续并行处理每个缩略图$J_{i;j,k}$。

[0104] 或者,缩略图$J_{i;j,k}$可以一个接一个地连续处理。

[0105] 因此,在步骤E8结束时,对于每个缩略图$J_{i;j,k}$,处理器2知道相关联的聚焦图$C_{i;j,k}$,该聚焦图$C_{i;j,k}$列出了在与相关缩略图$J_{i;j,k}$的每个像素相关联的人工神经网络NN的输出端获得的聚焦级别。每个聚焦图$C_{i;j,k}$与相应的缩略图$J_{i;j,k}$相关联,从而与(三维场景的)深度z_i相关联。

[0106] 如图4所示,该方法然后包括估计与数字全息图H相关联的深度图C的步骤E10。该深度图C包括多个深度值$d_{js+q,ks+r}$。每个深度值$d_{js+q,ks+r}$与缩略图$J_{i;j,k}$的不同像素中的像素相关联。深度值$d_{js+q,ks+r}$基于在步骤E8中确定的焦点级别来确定。

[0107] 实际上,在步骤E8期间,由于多个重构图像I_i的每个重构图像对应于不同的深度z_i,所以缩略图的像素与不同的聚焦级别相关联(取决于从中导出相关缩略图的重构图像I_i的深度)。换而言之,对于每个像素(与深度图C的深度值$d_{js+q,ks+r}$相关联),若干聚焦级别是已知的。

[0108] 因此,在此,处理器2为与相关深度值$d_{js+q,ks+r}$相关联的每个像素确定聚焦级别最高的深度:

[0109] $d_{js+q,ks+r}=\mathrm{argmax}_{i=1\cdots N}(C_{i;j,k}(q,r))$

[0110] 其中,(js+q,ks+r)是分配给确定的深度值的像素,NN是对应于人工神经网络NN和argmax的实现方式的算子,该算子对焦点级别的最大值的确定进行转换(在人工神经网络的输出端获得)。

[0111] 换而言之,对于每个索引像素(js+q,ks+r),处理器2确定焦点级别最高的深度。然后,该深度对应于深度值$d_{js+q,ks+r}$(深度图C的元素)。

[0112] 或者,可以使用除了确定焦点级别的最大值之外的另一种方法来确定深度值。例如,可以定义由多个相邻像素形成的区域,并且可以通过考虑从定义的像素区域上的焦点级别的平均值观察到最大偏差的深度来确定深度值。

[0113] 因此,在步骤E10结束时,深度图C例如在此从这些确定的深度值中估计。因此,深度图C的每个元素包括与具有索引(js+q,ks+r)的每个像素相关联的深度值$d_{js+q,ks+r}$。

[0114] 这个估计的深度图C最终使得可以具有深度值的矩阵的形式的空间信息,该深度值矩阵表示与数字全息图H相关联的三维场景。

[0115] 当然,上述用于数字全息图的方法以同样的方式应用于多个全息图。对于多个数字全息图,该方法的实现对于每个全息图可以是连续的,或者对于多个数字全息图可以是并行的。

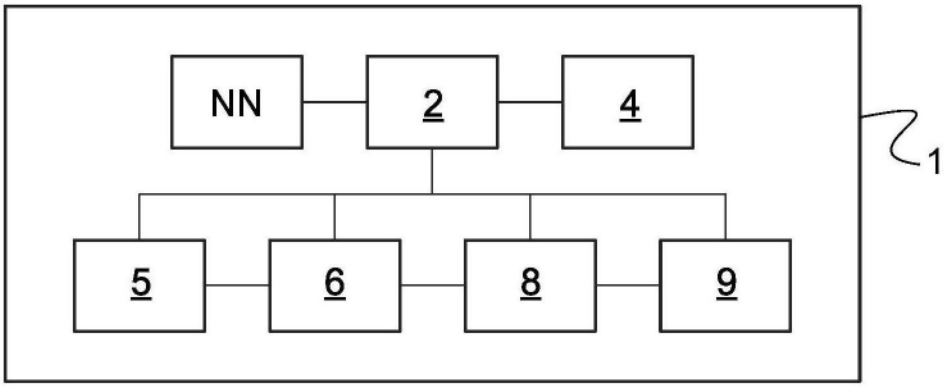


图1

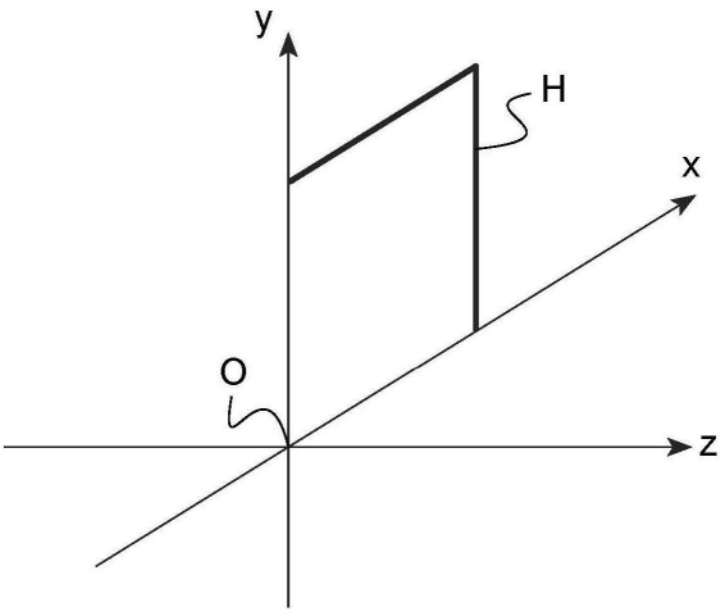


图2

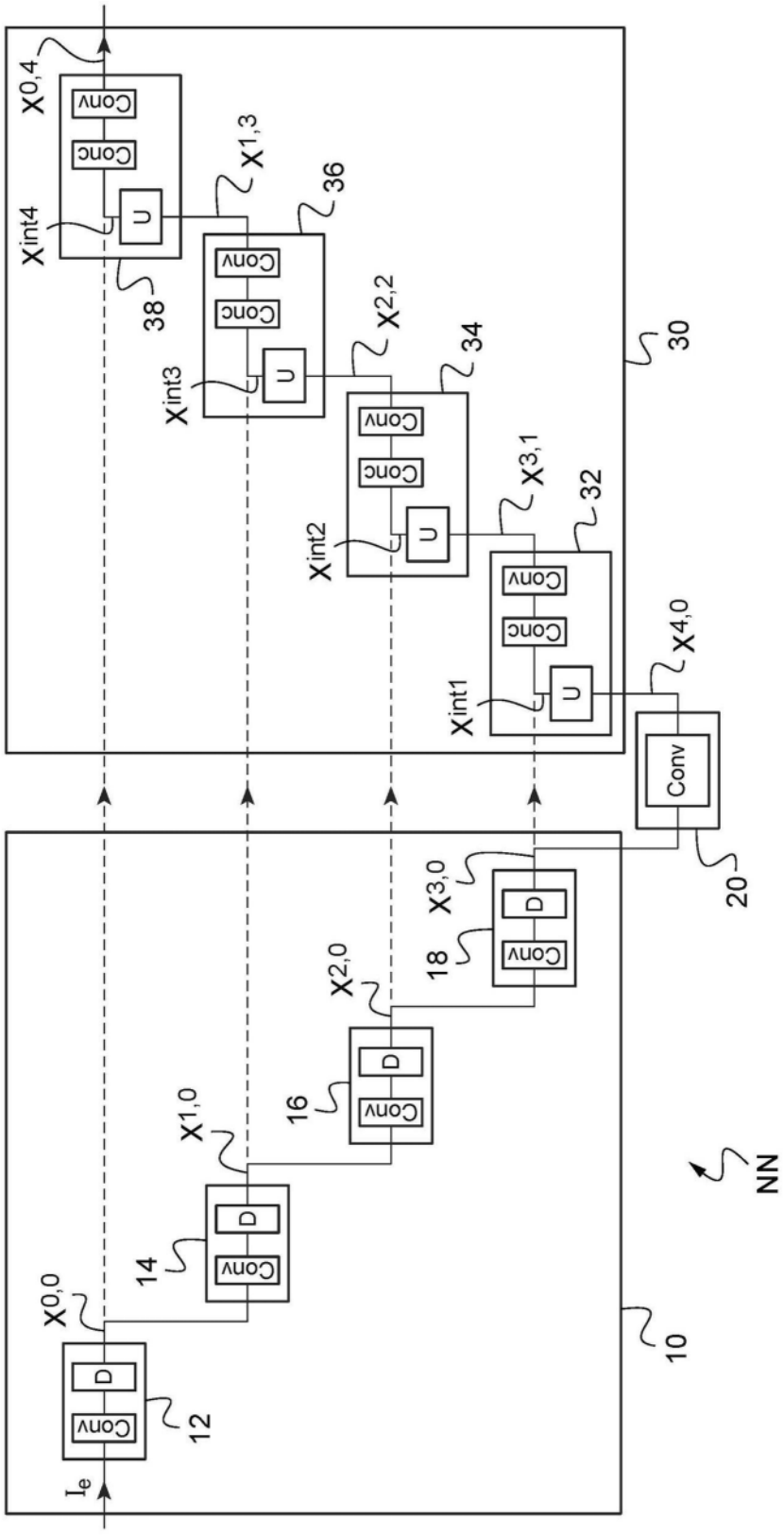

X0,4
Conv
Conc
U
Xint4
38
X1,3
36
Xint3
X2,2
34
Xint2
X3,1
32
Xint1
X4,0
30
Conv
20
X3,0
D
Conv
18
X2,0
16
X1,0
14
X0,0
12
Ie
10
NN

图3

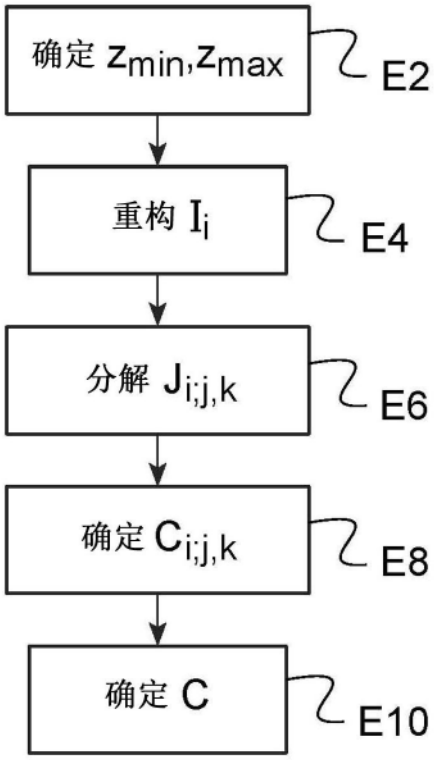


图4