

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号

WO 2023/201783 A1

(43) 国际公布日

2023 年 10 月 26 日 (26.10.2023)

WIPO | PCT

(51) 国际专利分类号:

G06T 7/50 (2017.01)

(21) 国际申请号:

PCT/CN2022/091182

(22) 国际申请日:

2022 年 5 月 6 日 (06.05.2022)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

202210404127.4 2022年4月18日 (18.04.2022) CN

(71) 申请人:清华大学 (TSINGHUA UNIVERSITY) [CN/CN]; 中国北京市海淀区清华园1号, Beijing 100084 (CN)。

(72) 发明人:戴琼海 (DAI, Qionghai); 中国北京市海淀区清华园1号, Beijing 100084 (CN)。岳冬晓 (YUE, Dongxiao); 中国北京市海淀区清华园1

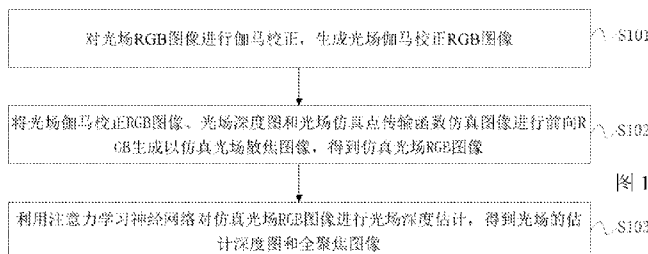
号, Beijing 100084 (CN)。于涛 (YU, Tao); 中国北京市海淀区清华园1号, Beijing 100084 (CN)。吴嘉敏 (WU, Jiamin); 中国北京市海淀区清华园1号, Beijing 100084 (CN)。

(74) 代理人:北京清亦华知识产权代理事务所 (普通合伙) (TSINGYIHUA INTELLECTUAL PROPERTY LLC); 中国北京市海淀区阜成路73号裕惠大厦B座806, Beijing 100142 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE,

(54) Title: LIGHT FIELD DEPTH ESTIMATION METHOD AND APPARATUS, AND ELECTRONIC DEVICE AND STORAGE MEDIUM

(54) 发明名称: 光场深度估计方法、装置、电子设备及存储介质



- S101 Perform gamma correction on a light field RGB image, so as to generate a light field gamma-corrected RGB image
- S102 Perform forward RGB generation on the light field gamma-corrected RGB image, a light field depth map and a light field simulation point transmission function simulation image to simulate a light field defocus image, so as to obtain a simulated light field RGB image
- S103 Perform light field depth estimation on the simulated light field RGB image by using an attention learning neural network, so as to obtain an estimated depth map and a full-focused image of a light field

(57) Abstract: The present application relates to the technical field of image processing, and particularly to a light field depth estimation method and apparatus, and an electronic device and a storage medium. The method comprises: performing gamma correction on a light field RGB image, so as to generate a light field gamma-corrected RGB image; performing forward RGB generation on the light field gamma-corrected RGB image, a light field depth map and a light field simulation point transmission function simulation image to simulate a light field defocus image, so as to obtain a simulated light field RGB image; and performing light field depth estimation on the simulated light field RGB image by using an attention learning neural network, so as to obtain an estimated depth map and a full-focused image of a light field. A light field defocus imaging model is established from a forward direction, so as to improve the depth estimation accuracy; on the basis of an optical image of the light field defocus imaging model, an attention learning neural network is constructed from a reverse direction, so as to quickly estimate a scene depth; and a full-focused image is obtained. Therefore, the accuracy of light field depth estimation is improved, the calculation amount of an algorithm is reduced, the estimation efficiency



PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE,
SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA,
RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO,
RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM,
GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

is improved, and real-time estimation can be performed.

(57) 摘要: 本申请涉及图像处理技术领域, 特别涉及一种光场深度估计方法、装置、电子设备及存储介质, 方法包括: 对光场RGB图像进行伽马校正, 生成光场伽马校正RGB图像; 将光场伽马校正RGB图像、光场深度图和光场仿真点传输函数仿真图像进行前向RGB生成以仿真光场散焦图像, 得到仿真光场RGB图像; 利用注意力学习神经网络对仿真光场RGB图像进行光场深度估计, 得到光场的估计深度图和全聚焦图像。从正向建立光场散焦成像模型提高深度估计准确度, 基于光场散焦成像模型的光学图像, 从逆向构建注意力学习神经网络快速估计场景深度, 同时获得全聚焦图像, 提高了光场深度估计的准确度, 降低了算法计算量, 提高了估计效率, 可以进行实时估计。

光场深度估计方法、装置、电子设备及存储介质

相关申请的交叉引用

- 5 本申请要求清华大学于 2022 年 04 月 18 日提交的、发明名称为“光场深度估计方法、装置、电子设备及存储介质”的、中国专利申请号“202210404127.4”的优先权。

技术领域

- 10 本申请涉及图像处理技术领域，特别涉及一种光场深度估计方法、装置、电子设备及存储介质

背景技术

- 15 光场深度估计是光学三维测量的代表技术之一，在智能制造、机器人视觉、自动驾驶、工业检测、虚拟现实、航空探测及元宇宙等领域具有重要的应用价值。现有光场深度估计方法主要分为两大类：基于优化的估计方法和基于数据驱动的深度学习方法。

 基于优化的传统方法主要基于对场景特性的分析，通过手动设计场景特征，并以此建立匹配代价函数。深度估计被转化为优化代价函数的问题。由于手动设计特征难以充分表达场景结构特点，优化的传统方法在处理遮挡、弱纹理等场景时，难以令人满意。此外，其构建的模型往往存在结构复杂且求解困难等问题，无法满足实时性要求。

- 20 基于数据驱动的深度学习严重依赖训练数据的数量和质量。模型均基于监督学习进行设计和训练，数据集则是通过光线追迹等方式人工合成。一方面，光场成像的特性决定了常规的结构光扫描或深度探测器，难以获得满足光场数据集要求的深度真值。因此，光场视差估计研究尚未能建立包含真实场景准确视差值的大规模数据集。另一方面，合成数据集中的样本数量有限，即使采用数据增强等方法，也难以获得充足的训练数据。这带来的问题是完全在合成数据集上训练的模型，无法在真实场景上获得相近的泛化性能。

 此外，不管是优化的估计方法还是基于数据驱动的深度学习方法，均是从光学系统成像后的图像出发，采用不同的手段估计光学图像的深度信息，即均假设理想的小孔成像模型。然而实际的光学成像系统是存在不同程度的像差的，成像模型带来的误差是影响当前深度估计算法准确度的关键元素之一。

- 30 综上，现有光场深度存在以下问题：

- 1、成像模型的误差以及优化算法误差使得光场深度估计准确度有待提高。
- 2、算法的复杂度导致计算量大，算法效率不高，无法满足实时性要求。

发明内容

本申请提供一种光场深度估计方法、装置、电子设备及存储介质，从正向建立光场散焦成像模型以提高深度估计准确度，基于光场散焦成像模型的光学图像，从逆向构建注意力学习神经网络以快速估计场景深度，同时获得全聚焦图像。

本申请第一方面实施例提供一种光场深度估计方法，包括以下步骤：对光场 RGB 图像进行伽马校正，生成光场伽马校正 RGB 图像；将所述光场伽马校正 RGB 图像、光场深度图和光场仿真点传输函数仿真图像进行前向 RGB 生成以仿真光场散焦图像，得到仿真光场 RGB 图像；利用注意力学习神经网络对所述仿真光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像。

可选地，在本申请的一个实施例中，在得到所述仿真光场 RGB 图像之前，还包括：根据光场相机参数和点源场传播过程，计算物方点经过主透镜的波函数；利用微透镜阵列的相位调制函数对所述波函数进行调制；采集相机像素对特定频率的分量，根据调制后的波函数和所述特定频率的分量计算点传输函数，并对所述点传输函数进行数值模拟及并采用双线性插值、拟合、对称补全和归一化操作，得到仿真点传输函数图；基于所述仿真点传输函数图对物方深度值进行随机采样，模拟仿真随机采样深度下的仿真点传输函数，得到所述光场仿真点传输函数仿真图像。

可选地，在本申请的一个实施例中，所述将所述光场伽马校正 RGB 图像、光场深度图和光场仿真点传输函数仿真图像进行前向 RGB 生成以仿真光场散焦图像，得到仿真光场 RGB 图像，包括：对所述光场深度图进行离散化，得到所述光场深度图对应的二进制深度掩膜；将所述光场伽马校正 RGB 图像与所述二进制深度掩膜相乘，得到对应的光场深度切片 RGB 图；将所述光场深度切片 RGB 图、所述二进制深度掩膜和所述光场仿真点传输函数仿真图像输入光场非线性成像模型，得到所述仿真光场 RGB 图像。

可选地，在本申请的一个实施例中，所述利用注意力学习神经网络对所述仿真光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像，包括：利用所述仿真光场 RGB 图像和 RL 迭代与估计算法获得初始光场聚焦序列，利用光场子孔径图像经所述注意力学习神经网络得到所述光场的估计深度图；将所述光场子孔径图像与所述初始光场聚焦序列级联，通过编解码网络输出估计的光场重聚焦序列，将所述光场重聚焦序列与连续深度体进行相乘，得到所述全聚焦图像。

可选地，在本申请的一个实施例中，在利用注意力学习神经网络对所述仿真光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像之后，还包括：将所述估计深度图与所述全聚焦图像分别与其真值比较，计算损失函数，后向传播误差训练所述注意

力学习神经网络参数。

本申请第二方面实施例提供一种光场深度估计装置，包括：处理模块，用于对光场 RGB 图像进行伽马校正，生成光场伽马校正 RGB 图像；生成模块，用于将所述光场伽马校正 RGB 图像、光场深度图和光场仿真点传输函数仿真图像进行前向 RGB 生成以仿真光场散焦图像，得到仿真光场 RGB 图像；估计模块，用于利用注意力学习神经网络对所述仿真光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像。

可选地，在本申请的一个实施例中，还包括：输出模块，用于在得到所述仿真光场 RGB 图像之前，根据光场相机参数和点源场传播过程，计算物方点经过主透镜的波函数，利用微透镜阵列的相位调制函数对所述波函数进行调制，采集相机像素对特定频率的分量，根据调制后的波函数和所述特定频率的分量计算点传输函数，并对所述点传输函数进行数值模拟及并采用双线性插值、拟合、对称补全和归一化操作，得到仿真点传输函数图，基于所述仿真点传输函数图对物方深度值进行随机采样，模拟仿真随机采样深度下的仿真点传输函数，得到所述光场仿真点传输函数仿真图像。

可选地，在本申请的一个实施例中，所述生成模块，进一步用于，对所述光场深度图进行离散化，得到所述光场深度图对应的二进制深度掩膜；将所述光场伽马校正 RGB 图像与所述二进制深度掩膜相乘，得到对应的光场深度切片 RGB 图；将所述光场深度切片 RGB 图、所述二进制深度掩膜和所述光场仿真点传输函数仿真图像输入光场非线性成像模型，得到所述仿真光场 RGB 图像。

可选地，在本申请的一个实施例中，所述估计模块，进一步用于，利用所述仿真光场 RGB 图像和 RL 迭代与估计算法获得初始光场聚焦序列，利用光场子孔径图像经所述注意力学习神经网络得到所述光场的估计深度图；将所述光场子孔径图像与所述初始光场聚焦序列级联，通过编解码网络输出估计的光场重聚焦序列，将所述光场重聚焦序列与连续深度体进行相乘，得到所述全聚焦图像。

可选地，在本申请的一个实施例中，还包括：训练模块，用于将所述估计深度图与所述全聚焦图像分别与其真值比较，计算损失函数，后向传播误差训练所述注意力学习神经网络参数。

本申请第三方面实施例提供一种电子设备，包括：存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序，所述处理器执行所述程序，以执行如上述实施例所述的光场深度估计方法。

本申请第四方面实施例提供一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行，以执行如上述实施例所述的光场深度估计方法。

本申请的实施例至少具有以下有益效果：

(1) 考虑光场散焦特征的前向成像建模方法提升了深度估计准确度。采用非线性光场散焦成像模型模拟了光场成像过程中不同深度、不同视角的散焦特征，而并非理想的小孔成像模型，能够更加准确对成像过程进行建模，从而提升光场深度估计的准确性。

5 (2) 考虑了遮挡的非线性光场成像模型提升了深度估计准确度。采用 α 合成的光场非线性成像，考虑了遮挡带来的影响，能够提供深度估计准确度。

(3) 深度估计结果更加快速。采用注意力深度神经网络模型，采用对称的注意力图以较少的训练参数确定光场视角的权重系数，相比于传统迭代方法更加快速、高效。

本申请附加的方面和优点将在下面的描述中部分给出，部分将从下面的描述中变得明显，或通过本申请的实践了解到。

10

附图说明

本申请上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解，其中：

图 1 为根据本申请实施例提供的一种光场深度估计方法的流程图；

15 图 2 为根据本申请实施例提供的一种光场深度估计方法的结构框架图；

图 3 为根据本申请实施例提供的一种光场传输模型流程图；

图 4 为根据本申请实施例提供的一种光场前向散焦 RGB 生成过程示意图；

图 5 为根据本申请实施例提供的一种光场深度反演过程示意图；

图 6 为根据本申请实施例提供的一种光场深度反演网络结构图；

20 图 7 为根据本申请实施例的光场深度估计装置的示例图；

图 8 为申请实施例提供的电子设备的结构示意图。

具体实施方式

下面详细描述本申请的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，旨在用于解释本申请，而不能理解为对本申请的限制。

下面参考附图描述本申请实施例的光场深度估计方法、装置、电子设备及存储介质。针对上述背景技术中心提到的现有光场深度成像模型的误差以及优化算法误差使得光场深度估计准确度低，算法的复杂度导致计算量大，算法效率不高，无法满足实时性要求的问题，本申请提供了一种光场深度估计方法，在该方法中，从正向建立光场散焦成像模型以提高深度估计准确度，基于光场散焦成像模型的光学图像，从逆向构建注意力学习神经网络以快速估计场景深度，同时获得全聚焦图像，由此，提高了光场深度估计的准确度，并

30

降低了算法的计算量，提高了估计效率，可以进行实时估计。

具体而言，图 1 为根据本申请实施例提供的一种光场深度估计方法的流程图。

如图 1 所示，该光场深度估计方法包括以下步骤：

在步骤 S101 中，对光场 RGB 图像进行伽马校正，生成光场伽马校正 RGB 图像。

- 5 在本申请的实施例中，基于现有开源光场数据集，对光场 RGB 图像进行伽马校正得到光场伽马校正 RGB 图像，与光场深度图一起送入前向 RGB 生成模型用以仿真散焦光场图像。

在步骤 S102 中，将光场伽马校正 RGB 图像、光场深度图和光场仿真点传输函数仿真图像进行前向 RGB 生成以仿真光场散焦图像，得到仿真光场 RGB 图像。

- 10 可选地，在本申请的一个实施例中，在得到仿真光场 RGB 图像之前，还包括：根据光场相机参数和点源场传播过程，计算物方点经过主透镜的波函数；利用微透镜阵列的相位调制函数对波函数进行调制；采集相机像素对特定频率的分量，根据调制后的波函数和特定频率的分量计算点传输函数，并对点传输函数进行数值模拟及并采用双线性插值、拟合、对称补全和归一化操作，得到仿真点传输函数图；基于仿真点传输函数图对物方深度值进
- 15 行随机采样，模拟仿真随机采样深度下的仿真点传输函数，得到光场仿真点传输函数仿真图像。

- 如图 2 和图 3 所示，计算光场仿真点传输函数仿真图像。具体地，基于光场相机的参数模型，根据光学传输衍射公式计算光场点传输函数（point spread function, PSF），为仿真不同视角光场散焦图像提供散焦线索。相关技术仅考虑了单视角正向传输模型，本申请考
- 20 虑了多视角的光场传输模型。

（1）首先根据光场相机参数及点源场传播过程，物方点经过主透镜的波函数 $U(x, y, z, \lambda)$ 可以表达为：

$$U(x, y, z, \lambda) = \left| \frac{2\pi}{\lambda s} \int_0^\infty r D(r, \lambda, z) J_0 \left(2\pi \sqrt{x^2 + y^2} r \right) dr \right|^2$$

其中， x, y, z 为物点的三维空间距离坐标， z 含义为深度， λ 为波长， r 为孔径平面径向距离， $J_0(\cdot)$ 为第一类零阶贝塞尔函数， $D(r, \lambda, z)$ 表达式如下：

$$D(r, \lambda, z) = \frac{z}{\lambda(r^2 + z^2)} e^{i \frac{2\pi}{\lambda} (\sqrt{r^2 + z^2} - \sqrt{r^2 + d^2})}$$

- 25 其中， s 为透镜与传感器之间距离， d 为物体与透镜距离。

（2）微透镜阵列的相位调制函数为：

$$t(x, y, x_0, y_0) = \sum_{x_0, y_0} \left\{ \text{rect} \left(\frac{x - x_0}{d} \right) \exp \left(\frac{-i\pi n}{\lambda f} \right) \right\} \left\{ \text{rect} \left(\frac{y - y_0}{d} \right) \exp \left(\frac{-i\pi n}{\lambda f} \right) \right\}$$

其中， x_0, y_0 为微透镜阵列的中心坐标， f 为焦距， n 为折射率。

(3) 经过微透镜相位调制后的光场信息为:

$$U'(\omega_x, \omega_y) = F_\omega(U(x, y, z, \lambda) \cdot t(x, y, x_0, y_0))$$

其中, ω_x, ω_y 为对空间 (x, y) 的频域采样, $F_\omega(\cdot)$ 为傅里叶变换操作。

(4) 相机像素对特定频率分量的采集过程为:

$$s(\omega_u, \omega_v) = \text{rect}\left(\frac{\omega_x - \omega_u}{d}\right) \text{rect}\left(\frac{\omega_y - \omega_v}{d}\right)$$

其中, ω_u 为角度 u 对应的空间频率位置, ω_v 为角度 v 对应的空间频率位置, $\text{rect}(\cdot)$ 为矩形窗函数。

(5) 则经微透镜后的点传输函数 PSF 可以表达为:

$$PSF(u, v, \lambda) = \int \|F_\omega(U(x, y, z, \lambda) \cdot t(x, y, x_0, y_0)) \cdot s(\omega_u, \omega_v)\|_2^2 d\omega_x d\omega_y$$

(6) 对上式 PSF 进行数值模拟, 并采用双线性插值、拟合、对称补全和归一化操作, 得到仿真 PSF 图。

(7) 对物方深度值 z 进行随机采样, 模拟仿真随机采样深度下的 PSF, 获得鲁棒的训练样本的 PSF 仿真图像。

可选地, 在本申请的一个实施例中, 将光场伽马校正 RGB 图像、光场深度图和光场仿真点传输函数仿真图像进行前向 RGB 生成以仿真光场散焦图像, 得到仿真光场 RGB 图像, 包括: 对光场深度图进行离散化, 得到光场深度图对应的二进制深度掩膜; 将光场伽马校正 RGB 图像与二进制深度掩膜相乘, 得到对应的光场深度切片 RGB 图; 将光场深度切片 RGB 图、二进制深度掩膜和光场仿真点传输函数仿真图像输入光场非线性成像模型, 得到仿真光场 RGB 图像。

将得到的光场伽马校正 RGB 图像、光场深度图以及仿真获得的 PSF 送入前向 RGB 生成过程, 以仿真光场散焦图像。该前向仿真过程不同于现有的理想小孔成像技术, 通过引用光场点扩散函数对不同深度、不同角度的光学散焦进行建模, 能够更接近真实获得的光场图像。前向生成过程如图 4 所示。

(1) 首先对光场深度图进行离散化, 获得对应深度图的二进制深度掩膜 $\alpha_k(\lambda)$, $k=1, 2, \dots, K$ 。 K 为离散的深度层数。

(2) 将第一步获得的光场伽马校正 RGB 图像与离散的二进制深度掩膜相乘, 获得对应的光场深度切片 RGB 图。

(3) 将光场深度切片 RGB 图、二进制深度掩膜 $\alpha_k(\lambda)$ 以及仿真得到的光场 PSF 送入考虑遮挡基于 α 合成的光场非线性成像模型, 以得到仿真的光场 RGB 图像 $b(\lambda, u, v)$, 如下式所示:

$$b(\lambda, u, v) = \sum_{k=0}^{K-1} \tilde{l}_k(u, v) \prod_{k=k+1}^{K-1} (1 - \tilde{\alpha}_{k'}(u, v)) + \eta$$

其中：

$$\tilde{l}_k := \frac{PSF_k(\lambda, u, v) * l_k(\lambda)}{E_k(\lambda)}, \tilde{\alpha}_k := \frac{PSF_k(\lambda, u, v) * \alpha_k(\lambda)}{E_k(\lambda)}$$

$$E_k(\lambda) := PSF_k(\lambda, u, v) * \sum_{k'=0}^k \alpha_{k'}$$

其中，深度图量化为 K 层， u, v 表示光场视角， λ 为波长， $*$ 表示卷积操作。 $\alpha_k(k = 1, 2, \dots, K)$ 为二进制深度掩膜， $l_k(\lambda)$ 为第 k 层光场深度切片 RGB 图， $PSF_k(\lambda, u, v)$ 表示 u, v 视角下的第 k 层深度处的 PSF， $\tilde{l}_k$ 表征了不考虑遮挡时第 k 层深度处的成像响应， $\tilde{\alpha}_k$ 表征了遮

5 挡对第 k 层深度处成像的影响， η 为加性噪声。

在步骤 S103 中，利用注意力学习神经网络对仿真光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像。

可选地，在本申请的一个实施例中，利用注意力学习神经网络对仿真光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像，包括：利用仿真光场 RGB 图像和
10 RL 迭代与估计算法获得初始光场聚焦序列，利用光场子孔径图像经注意力学习神经网络得到光场的估计深度图；将光场子孔径图像与初始光场聚焦序列级联，通过编解码网络输出估计的光场重聚焦序列，将光场重聚焦序列与连续深度体进行相乘，得到全聚焦图像。

可选地，在本申请的一个实施例中，在利用注意力学习神经网络对仿真光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像之后，还包括：将估计深度图与
15 全聚焦图像分别与其真值比较，计算损失函数，后向传播误差训练注意力学习神经网络参数。

如图 5 所示，基于上述前向模型仿真得到的光场 RGB 图像进行深度估计，首先，采用 RL 迭代与估计算法获得初始光场聚焦序列（focalstack），采用注意力深度学习网络训练、估计光场深度图。

20 具体地，本申请采用的深度反演神经网络如图 6 所示。深度反演网络的输入为第三步仿真得到的多视角光场图像，输出为估计深度图及其对应的全聚焦（All-in-focus）图像。

深度反演网络主要由两部分组成：第一部分采用注意力视角选择神经网络，主要由 2D 卷积、Resblock、SPPModule、CostVolume、AttentionModule、3DCNN 等模块组成，输入为光场子孔径图像，输出为 $H \times W \times K$ 维度（ K 表示深度层数， H, W 分别表示光场图像的长
25 和宽）的深度体 M ，取最大值即可得到光场深度图。

第二部分采用 3DU-Net 编解码网络估计全聚焦图像，输入为光场子孔径图像，首先经过 2DCNN、Resblock、SPPModule 等模块，然后与预估计的光场聚焦序列进行级联，然后

通过编解码网络输出估计的光场重聚焦序列，与连续深度体 M^{AiF} 相乘即可得到全聚焦图像。

深度体 M 与 M^{AiF} 关系定义为：

$$M = \varsigma(M^{AiF})$$

其中：

$$M_{i,j,k} = \frac{\ln(1 + \exp(M_{i,j,k}^{AiF}))}{\sum_{n=1}^K \ln(1 + \exp(M_{i,j,n}^{AiF}))},$$

$$i = 1, \dots, H, j = 1, \dots, W, k = 1, \dots, K$$

其中， $M \in \mathbb{R}^{H \times W \times K}$ ， $M^{AiF} \in \mathbb{R}^{H \times W \times K}$ 为连续深度体， i 和 j 分别表示第 i 行和第 j 列， k 表

示第 k 深度层，为 $M_{i,j,k}$ 和 $M_{i,j,k}^{AiF}$ 分别表示深度体 M 和连续深度体 M^{AiF} 的第 i 行第 j 列第 k 深度层的取值。

将估计的光场深度图与光场全聚焦图像分别与其真值相比较，计算损失函数，后向传播误差训练网络参数。其中，光场深度图的损失函数选用最小均方误差函数，光场全聚焦图像的损失函数选用 SSIM（structural similarity，结构相似性）指标。

根据本申请实施例提出的光场深度估计方法，从正向建立光场散焦成像模型以提高深度估计准确度，基于光场散焦成像模型的光学图像，从逆向构建注意力学习神经网络以快速估计场景深度，同时获得全聚焦图像，由此，提高了光场深度估计的准确度，并降低了算法的计算量，提高了估计效率，可以进行实时估计。

其次参照附图描述根据本申请实施例提出的光场深度估计装置。

图 7 为根据本申请实施例的光场深度估计装置的示例图。

如图 7 所示，该光场深度估计装置 10 包括：处理模块 100、生成模块 200 和估计模块 300。

其中，处理模块 100，用于对光场 RGB 图像进行伽马校正，生成光场伽马校正 RGB 图像。生成模块 200，用于将光场伽马校正 RGB 图像、光场深度图和光场仿真点传输函数仿真图像进行前向 RGB 生成以仿真光场散焦图像，得到仿真光场 RGB 图像。估计模块 300，用于利用注意力学习神经网络对仿真光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像。

可选地，在本申请的一个实施例中，光场深度估计装置 10 还包括：输出模块，用于在得到仿真光场 RGB 图像之前，根据光场相机参数和点源场传播过程，计算物方点经过主透镜的波函数，利用微透镜阵列的相位调制函数对波函数进行调制，采集相机像素对特定频率的分量，根据调制后的波函数和特定频率的分量计算点传输函数，并对点传输函数进行数值模拟及并采用双线性插值、拟合、对称补全和归一化操作，得到仿真点传输函数图，

基于仿真点传输函数图对物方深度值进行随机采样，模拟仿真随机采样深度下的仿真点传输函数，得到光场仿真点传输函数仿真图像。

可选地，在本申请的一个实施例中，生成模块 200，进一步用于，对光场深度图进行离散化，得到光场深度图对应的二进制深度掩膜；将光场伽马校正 RGB 图像与二进制深度掩膜相乘，得到对应的光场深度切片 RGB 图；将光场深度切片 RGB 图、二进制深度掩膜和光场仿真点传输函数仿真图像输入光场非线性成像模型，得到仿真光场 RGB 图像。

可选地，在本申请的一个实施例中，估计模块 300，进一步用于，利用仿真光场 RGB 图像和 RL 迭代与估计算法获得初始光场聚焦序列，利用光场子孔径图像经注意力学习神经网络得到光场的估计深度图；将光场子孔径图像与初始光场聚焦序列级联，通过编解码网络输出估计的光场重聚焦序列，将光场重聚焦序列与连续深度体进行相乘，得到全聚焦图像。

可选地，在本申请的一个实施例中，光场深度估计装置 10 还包括：训练模块，用于将估计深度图与全聚焦图像分别与其真值比较，计算损失函数，后向传播误差训练注意力学习神经网络参数。

需要说明的是，前述对光场深度估计方法实施例的解释说明也适用于该实施例的光场深度估计装置，此处不再赘述。

根据本申请实施例提出的光场深度估计装置，从正向建立光场散焦成像模型以提高深度估计准确度，基于光场散焦成像模型的光学图像，从逆向构建注意力学习神经网络以快速估计场景深度，同时获得全聚焦图像，由此，提高了光场深度估计的准确度，并降低了算法的计算量，提高了估计效率，可以进行实时估计。

图 8 为本申请实施例提供的电子设备的结构示意图。该电子设备可以包括：

存储器 801、处理器 802 及存储在存储器 801 上并可在处理器 802 上运行的计算机程序。处理器 802 执行程序时实现上述实施例中提供的光场深度估计方法。

进一步地，车辆还包括：

通信接口 803，用于存储器 801 和处理器 802 之间的通信。

存储器 801，用于存放可在处理器 802 上运行的计算机程序。

存储器 801 可能包含高速 RAM 存储器，也可能还包括非易失性存储器（non-volatile memory），例如至少一个磁盘存储器。

如果存储器 801、处理器 802 和通信接口 803 独立实现，则通信接口 803、存储器 801 和处理器 802 可以通过总线相互连接并完成相互间的通信。总线可以是工业标准体系结构（Industry Standard Architecture，简称为 ISA）总线、外部设备互连（Peripheral Component，简称为 PCI）总线或扩展工业标准体系结构（Extended Industry Standard Architecture，简称

为 EISA) 总线等。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 8 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。

可选的, 在具体实现上, 如果存储器 801、处理器 802 及通信接口 803, 集成在一块芯片上实现, 则存储器 801、处理器 802 及通信接口 803 可以通过内部接口完成相互间的通信。

处理器 802 可能是一个中央处理器 (Central Processing Unit, 简称为 CPU), 或者是特定集成电路 (Application Specific Integrated Circuit, 简称为 ASIC), 或者是被配置成实施本申请实施例的一个或多个集成电路。

本实施例还提供一种计算机可读存储介质, 其上存储有计算机程序, 其特征在于, 该程序被处理器执行时实现如上的光场深度估计方法。

在本说明书的描述中, 参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本申请的至少一个实施例或示例中。在本说明书中, 对上述术语的示意性表述不必须针对的是相同的实施例或示例。而且, 描述的具体特征、结构、材料或者特点可以在任一个或 N 个实施例或示例中以合适的方式结合。此外, 在不相互矛盾的情况下, 本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

此外, 术语“第一”、“第二”仅用于描述目的, 而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此, 限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本申请的描述中, “N 个”的含义是至少两个, 例如两个, 三个等, 除非另有明确具体的限定。

流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为, 表示包括一个或更 N 个用于实现定制逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分, 并且本申请的优选实施方式的范围包括另外的实现, 其中可以不按所示出或讨论的顺序, 包括根据所涉及的功能按基本同时的方式或按相反的顺序, 来执行功能, 这应被本申请的实施例所属技术领域的技术人员所理解。

应当理解, 本申请的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中, N 个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。如, 如果用硬件来实现和在另一实施方式中一样, 可用本领域公知的下列技术中的任一项或他们的组合来实现: 具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路, 具有合适的组合逻辑门电路的专用集成电路, 可编程门阵列 (PGA), 现场可编程门阵列 (FPGA) 等。

本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，该程序在执行时，包括方法实施例的步骤之一或其组合。

权利要求书

1、一种光场深度估计方法，其特征在于，包括以下步骤：

对光场 RGB 图像进行伽马校正，生成光场伽马校正 RGB 图像；

5 将所述光场伽马校正 RGB 图像、光场深度图和光场仿真点传输函数仿真图像进行前向 RGB 生成以仿真光场散焦图像，得到仿真光场 RGB 图像；以及

利用注意力学习神经网络对所述仿真光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像。

10 2、根据权利要求 1 所述的方法，其特征在于，在得到所述仿真光场 RGB 图像之前，还包括：

根据光场相机参数和点源场传播过程，计算物方点经过主透镜的波函数；

利用微透镜阵列的相位调制函数对所述波函数进行调制；

15 采集相机像素对特定频率的分量，根据调制后的波函数和所述特定频率的分量计算点传输函数，并对所述点传输函数进行数值模拟及并采用双线性插值、拟合、对称补全和归一化操作，得到仿真点传输函数图；

基于所述仿真点传输函数图对物方深度值进行随机采样，模拟仿真随机采样深度下的仿真点传输函数，得到所述光场仿真点传输函数仿真图像。

20 3、根据权利要求 1 或 2 所述的方法，其特征在于，所述将所述光场伽马校正 RGB 图像、光场深度图和光场仿真点传输函数仿真图像进行前向 RGB 生成以仿真光场散焦图像，得到仿真光场 RGB 图像，包括：

对所述光场深度图进行离散化，得到所述光场深度图对应的二进制深度掩膜；

将所述光场伽马校正 RGB 图像与所述二进制深度掩膜相乘，得到对应的光场深度切片 RGB 图；

25 将所述光场深度切片 RGB 图、所述二进制深度掩膜和所述光场仿真点传输函数仿真图像输入光场非线性成像模型，得到所述仿真光场 RGB 图像。

4、根据权利要求 1 所述的方法，其特征在于，所述利用注意力学习神经网络对所述仿真光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像，包括：

利用所述仿真光场 RGB 图像和 RL 迭代与估计算法获得初始光场聚焦序列，利用光场子孔径图像经所述注意力学习神经网络得到所述光场的估计深度图；

30 将所述光场子孔径图像与所述初始光场聚焦序列级联，通过编解码网络输出估计的光场重聚焦序列，将所述光场重聚焦序列与连续深度体进行相乘，得到所述全聚焦图像。

5、根据权利要求 4 所述的方法，其特征在于，在利用注意力学习神经网络对所述仿真

光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像之后，还包括：

将所述估计深度图与所述全聚焦图像分别与其真值比较，计算损失函数，后向传播误差训练所述注意力学习神经网络参数。

6、一种光场深度估计装置，其特征在于，包括：

5 处理模块，用于对光场 RGB 图像进行伽马校正，生成光场伽马校正 RGB 图像；

生成模块，用于将所述光场伽马校正 RGB 图像、光场深度图和光场仿真点传输函数仿真图像进行前向 RGB 生成以仿真光场散焦图像，得到仿真光场 RGB 图像；以及

估计模块，用于利用注意力学习神经网络对所述仿真光场 RGB 图像进行光场深度估计，得到光场的估计深度图和全聚焦图像。

10 7、根据权利要求 6 所述的装置，其特征在于，还包括：

输出模块，用于在得到所述仿真光场 RGB 图像之前，根据光场相机参数和点源场传播过程，计算物方点经过主透镜的波函数，利用微透镜阵列的相位调制函数对所述波函数进行调制，采集相机像素对特定频率的分量，根据调制后的波函数和所述特定频率的分量计算点传输函数，并对所述点传输函数进行数值模拟及并采用双线性插值、拟合、对称补全
15 和归一化操作，得到仿真点传输函数图，基于所述仿真点传输函数图对物方深度值进行随机采样，模拟仿真随机采样深度下的仿真点传输函数，得到所述光场仿真点传输函数仿真图像。

8、根据权利要求 6 或 7 所述的装置，其特征在于，所述生成模块，进一步用于，对所述光场深度图进行离散化，得到所述光场深度图对应的二进制深度掩膜；将所述光场伽马
20 校正 RGB 图像与所述二进制深度掩膜相乘，得到对应的光场深度切片 RGB 图；将所述光场深度切片 RGB 图、所述二进制深度掩膜和所述光场仿真点传输函数仿真图像输入光场非线性成像模型，得到所述仿真光场 RGB 图像。

9、根据权利要求 6 所述的装置，其特征在于，所述估计模块，进一步用于，利用所述仿真光场 RGB 图像和 RL 迭代与估计算法获得初始光场聚焦序列，利用光场子孔径图像经
25 所述注意力学习神经网络得到所述光场的估计深度图；将所述光场子孔径图像与所述初始光场聚焦序列级联，通过编解码网络输出估计的光场重聚焦序列，将所述光场重聚焦序列与连续深度体进行相乘，得到所述全聚焦图像。

10、根据权利要求 9 所述的装置，其特征在于，还包括：

30 训练模块，用于将所述估计深度图与所述全聚焦图像分别与其真值比较，计算损失函数，后向传播误差训练所述注意力学习神经网络参数。

11、一种电子设备，其特征在于，包括：存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序，所述处理器执行所述程序，以实现如权利要求 1-5 任

一项所述的光场深度估计方法。

12、一种计算机可读存储介质，其上存储有计算机程序，其特征在于，该程序被处理器执行，以用于实现如权利要求1-5任一项所述的光场深度估计方法。

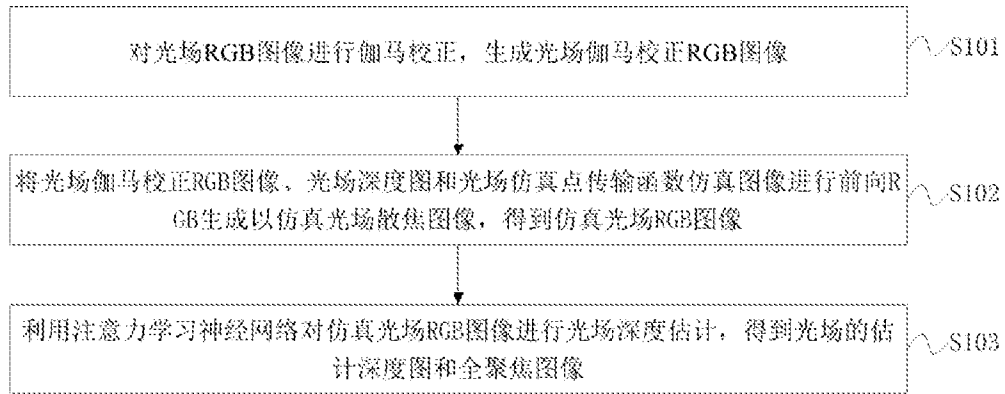


图 1

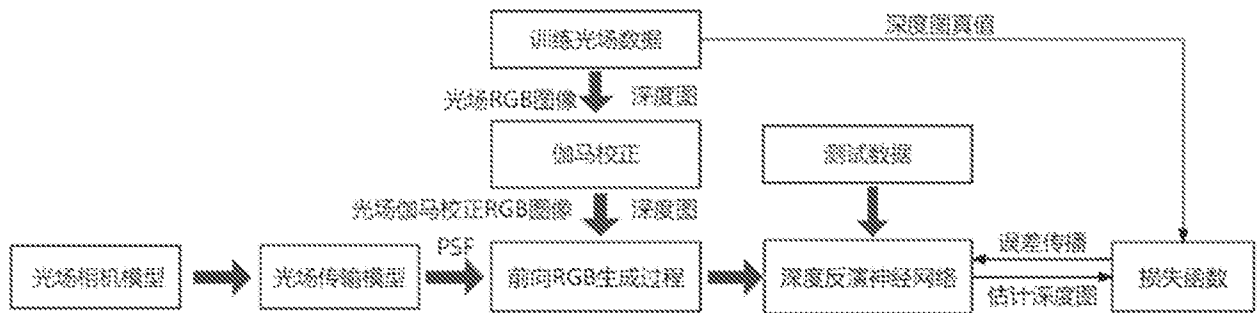


图 2

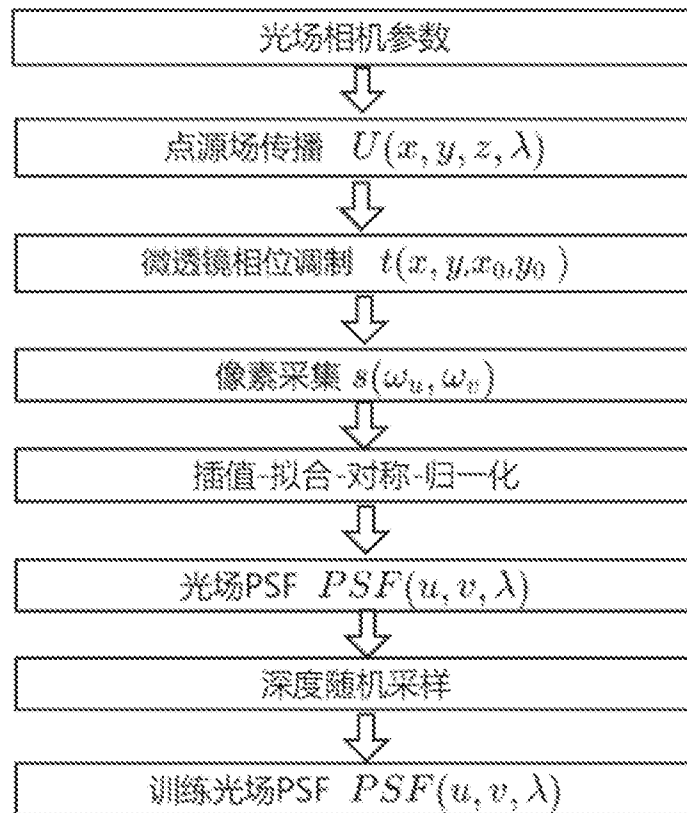


图 3

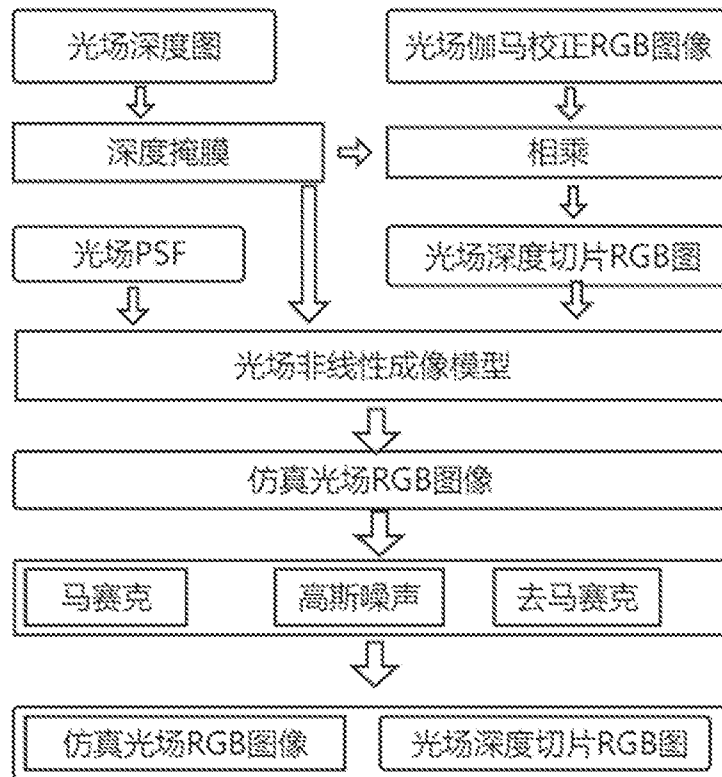


图 4

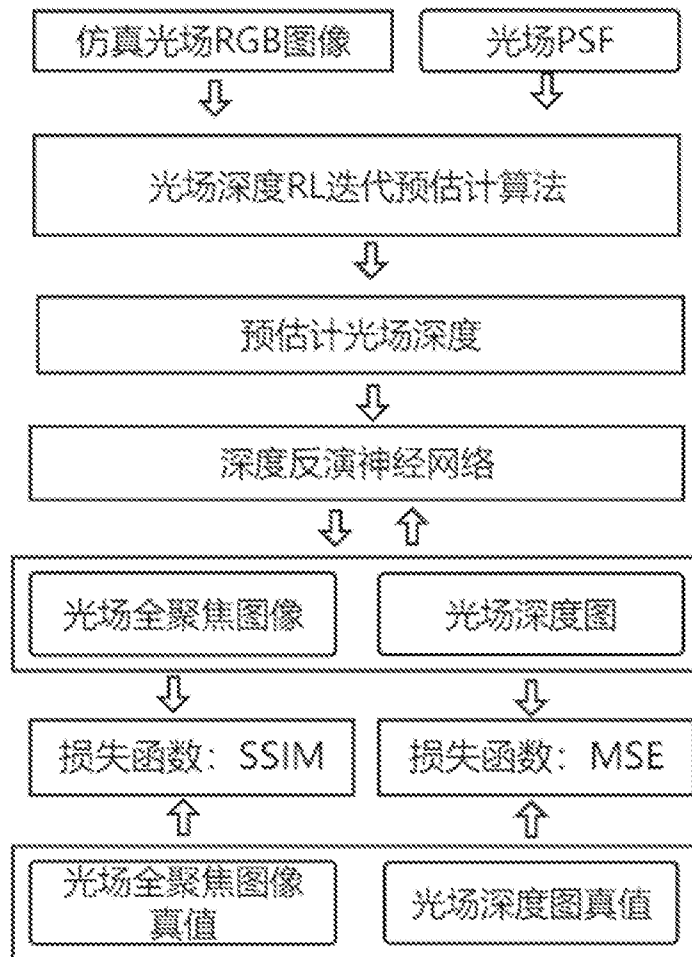


图 5

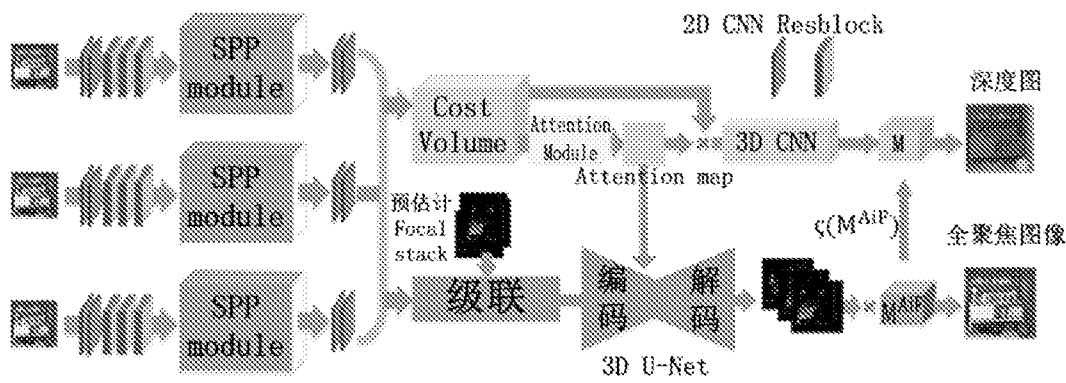


图 6

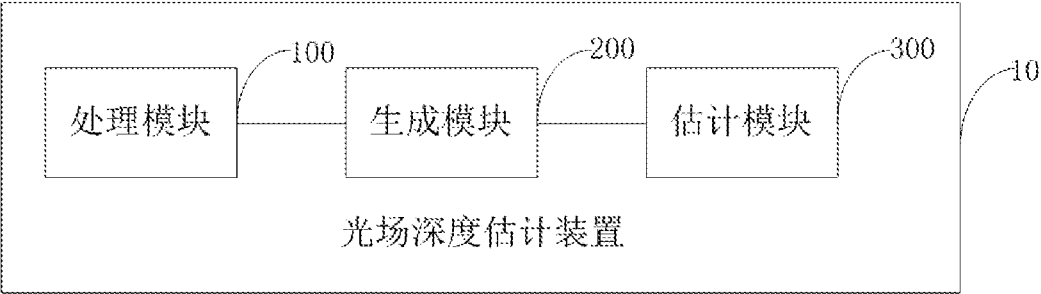


图 7

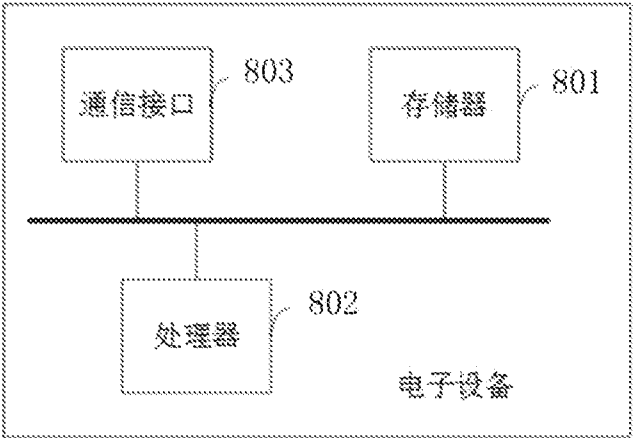


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/091182

A. CLASSIFICATION OF SUBJECT MATTER

G06T 7/50(2017.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

USTXT; VEN; CNABS; CNTXT; EPTXT; WOTXT; CNKI; IEEE: 光场, 深度, 估计, 预测, 伽马, 校正, 红绿蓝, 图像, 深度图, 点传输函数, 点扩散函数, 生成, 仿真, 模拟, 注意力, 神经网络, 散焦, light field, depth, estimate, predict, gamma, revise, correct, RGB, image, depth map, point spread function, PSF, create, emulation, simulate, attention, neural network, defocusing

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 113506336 A (SHANGHAI NORMAL UNIVERSITY) 15 October 2021 (2021-10-15) description, paragraphs 7-36	1-12
A	CN 110443882 A (TSINGHUA UNIVERSITY) 12 November 2019 (2019-11-12) description, paragraphs 5-32	1-12
A	CN 114092540 A (SHANGHAI NORMAL UNIVERSITY) 25 February 2022 (2022-02-25) entire document	1-12
A	CN 113554744 A (TSINGHUA UNIVERSITY) 26 October 2021 (2021-10-26) entire document	1-12
A	CN 111127536 A (TSINGHUA UNIVERSITY) 08 May 2020 (2020-05-08) entire document	1-12
A	CN 112102165 A (BEIHANG UNIVERSITY) 18 December 2020 (2020-12-18) entire document	1-12
A	CN 112150526 A (ZHEJIANG UNIVERSITY) 29 December 2020 (2020-12-29) entire document	1-12



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
 “E” earlier application or patent but published on or after the international filing date
 “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 “O” document referring to an oral disclosure, use, exhibition or other means
 “P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
 “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
 “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
 “&” document member of the same patent family

Date of the actual completion of the international search

09 January 2023

Date of mailing of the international search report

17 January 2023

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/091182**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2020120328 A1 (AVALON HOLOGRAPHICS INC.) 16 April 2020 (2020-04-16) entire document	1-12
A	ZHOU, Xu et al. "Parameter-Free Gaussian PSF Model for Extended Depth of Field in Brightfield Microscopy" <i>IEEE Transactions on Image Processing</i> , Vol. 29, 28 January 2020 (2020-01-28), ISSN: 1057-7149, pages 3227-3238	1-12

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/091182

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	113506336	A	15 October 2021	None			
CN	110443882	A	12 November 2019	CN	110443882	B	11 June 2021
CN	113554744	A	26 October 2021	None			
CN	111127536	A	08 May 2020	None			
CN	112102165	A	18 December 2020	CN	112102165	B	06 December 2022
CN	112150526	A	29 December 2020	None			
US	2020120328	A1	16 April 2020	US	10924727	B2	16 February 2021

国际检索报告

国际申请号

PCT/CN2022/091182

A. 主题的分类

G06T 7/50 (2017.01) i

按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

G06T

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

USTXT;VEN;CNABS;CNTXT;EPTXT;WOTXT;CNKI;IEEE: 光场, 深度, 估计, 预测, 伽马, 校正, 红绿蓝, 图像, 深度图, 点传输函数, 点扩散函数, 生成, 仿真, 模拟, 注意力, 神经网络, 散焦, light field, depth, estimate, predict, gamma, revise, correct, RGB, image, depth map, point spread function, PSF, create, emulation, simulate, attention, neural network, defocusing

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 113506336 A (上海师范大学) 2021年10月15日 (2021 - 10 - 15) 说明书第7-36段	1-12
A	CN 110443882 A (清华大学) 2019年11月12日 (2019 - 11 - 12) 说明书第5-32段	1-12
A	CN 114092540 A (上海师范大学) 2022年2月25日 (2022 - 02 - 25) 全文	1-12
A	CN 113554744 A (清华大学) 2021年10月26日 (2021 - 10 - 26) 全文	1-12
A	CN 111127536 A (清华大学) 2020年5月8日 (2020 - 05 - 08) 全文	1-12
A	CN 112102165 A (北京航空航天大学) 2020年12月18日 (2020 - 12 - 18) 全文	1-12
A	CN 112150526 A (浙江大学) 2020年12月29日 (2020 - 12 - 29) 全文	1-12

☒ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2023年1月9日

国际检索报告邮寄日期

2023年1月17日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

李富贵

电话号码 86-(20)-28958943

C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	US 2020120328 A1 (AVALON HOLOGRAPHICS INC) 2020年4月16日 (2020 - 04 - 16) 全文	1-12
A	ZHOU, Xu 等. "Parameter-Free Gaussian PSF Model for Extended Depth of Field in Brightfield Microscopy" IEEE Transactions on Image Processing, 第29卷, 2020年1月28日 (2020 - 01 - 28), ISSN: 1057-7149, 第3227-3238页	1-12

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/091182

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	113506336	A	2021年10月15日	无	
CN	110443882	A	2019年11月12日	CN 110443882	B 2021年6月11日
CN	113554744	A	2021年10月26日	无	
CN	111127536	A	2020年5月8日	无	
CN	112102165	A	2020年12月18日	CN 112102165	B 2022年12月6日
CN	112150526	A	2020年12月29日	无	
US	2020120328	A1	2020年4月16日	US 10924727	B2 2021年2月16日