

权 利 要 求

1、一种深度估计方法，其特征在于，包括：

以单张图片作为双目图像中的第一图像，经第一神经网络，基于所述第一图像获取所述双目图像中的第二图像；

经第二神经网络，通过对所述第一图像与所述第二图像进行双目立体匹配，获取所述第一图像对应的深度信息。

2、根据权利要求 1 所述的方法，其特征在于，经第一神经网络，基于所述第一图像获取所述双目图像中的第二图像，包括：

经第一神经网络对所述第一图像进行处理，输出 N 个通道的视差概率图；其中，每个通道的视差概率图表示所述第一图像上像素向第一水平方向偏移 i 个视差的概率， $i=0,1,\dots,N-1$ ， N 的取值为大于 1 的整数；

根据所述 N 个通道的视差概率图，将所述第一图像分别向第一水平方向偏移 i 个像素，得到 N 张偏移图；

将所述 N 张偏移图中的各张偏移图分别点乘对应通道的视差概率图，得到 N 个点乘结果；

将所述 N 个点乘结果基于像素进行叠加，得到所述第二图像。

3、根据权利要求 2 所述的方法，其特征在于，所述经第一神经网络对所述第一图像进行处理，输出 N 个通道的视差概率图，包括：

分别通过第一神经网络中至少两个网络深度的网络层对所述第一图像进行特征提取，获得至少两个尺度的特征图；

基于所述至少两个尺度的特征图分别获取至少两个分辨率的、 N 个通道的初步视差概率图；

分别针对每个通道，将所述至少两个分辨率的初步视差概率图放大至所述第一图像的分辨率下进行叠加，得到所述 N 个通道的视差概率图。

4、根据权利要求 2 或 3 所述的方法，其特征在于，所述第一图像中对应前景对象的像素在所述 N 个通道的视差概率图中的对应像素位置，在对应较大视差的通道的视差概率图中具有较大的概率值；所述第一图像中对应背景对象的像素在所述 N 个通道的视差概率图中的对应像素位置，在对应较小视差的通道的视差概率图中具有较大的概率值。

5、根据权利要求 1-4 任一所述的方法，其特征在于，通过对所述第一图像与所述第二图像进行双目立体匹配，获取所述第一图像对应的深度信息，包括：

获取用于表示所述双目图像中一个图像相对于另一个图像中像素的位置关系的相关系数；所述双目图像中一个图像包括所述第一图像或所述第二图像，所述另一个图像对应包括所述第二图像或所述第一图像；

基于所述另一个图像与所述相关系数，生成所述第一图像与所述第二图像的视差图；基于所述视差图获取所述第一图像对应的深度信息。

6、根据权利要求 5 所述的方法，其特征在于，所述获取用于表示所述双目图像中一个图像相对于另一个图像中像素的位置关系的相关系数，包括：

分别对所述一个图像与所述另一个图像进行特征提取；

基于提取的所述一个图像的特征与所述另一个图像的特征，获取所述一个图像与所述另一个图像中像素的位置关系，并输出用于表示所述双目图像中一个图像相对于另一个图像中像素的位置关系的所述相关系数。

7、根据权利要求 6 所述的方法，其特征在于，基于所述另一个图像与所述相关系数，生成所述第一图像与所述第二图像的视差图，包括：

将所述另一个图像的特征与所述相关系数进行叠加，生成所述第一图像与所述第二图像的视差图。

8、根据权利要求 7 所述的方法，其特征在于，将所述另一个图像的特征与所述相关系数进行叠加，生成所述第一图像与所述第二图像的视差图，包括：

将所述另一个图像的特征与所述相关系数进行叠加，得到叠加结果；

提取所述叠加结果的特征，并将提取的所述叠加结果的特征与所述叠加结果进行融合，获得所述第一图像与所述第二图像的视差图。

9、根据权利要求 6-8 任一所述的方法，其特征在于，所述基于所述视差图获取所述第一图像对应的深度信息，包括：

基于所述视差图、拍摄所述第一图像的相机的焦距和所述双目图像对应的双目相机之间的距离，获取所述第一图像对应的深度信息。

10、根据权利要求 1-9 任一所述的方法，其特征在于，还包括：

基于所述第一图像对应的所述深度信息与所述第二图像，获取所述第一图像中场景的三维空间场景。

11、根据权利要求 5-10 任一所述的方法，其特征在于，所述第一神经网络利用第一样本集中的样本双目图像预先训练而得，所述第一样本集包括至少一组第一样本双目图像；和/或，

所述第二神经网络利用第二样本集中的样本双目图像预先训练而得；所述第二样本集包括至少一组第二样本双目图像和视差图标签。

12、根据权利要求 11 所述的方法，其特征在于，所述第一神经网络的训练包括：

经所述第一神经网络，由所述至少一组第一样本双目图像中的第一图像，获取所述至少一组第一样本双目图像中的第二图像并输出；

获取所述第一神经网络输出的第二图像与所述至少一组第一样本双目图像中的第二图像之间的第一差异，并基于所述第一差异对所述第一神经网络进行训练，直至满足第一训练完成条件。

13、根据权利要求 12 所述的方法，其特征在于，获取所述第一神经网络输出的第二图像与所述至少一组第一样本双目图像中的第二图像之间的第一差异，并基于所述第一差异对所述第一神经网络进行训练，直至满足第一训练完成条件，包括：

获取所述第一神经网络输出的第二图像与所述至少一组第一样本双目图像中的第二图像之间在像素上的第一差异；

基于所述第一差异调整所述第一神经网络中网络参数的参数值，直至满足第一训练完成条件；

所述第一训练完成条件包括：所述第一差异小于第一预设阈值，和/或，对所述第一神经网络的训练次数达到第一预设次数。

14、根据权利要求 11-13 任一所述的方法，其特征在于，所述第二神经网络的训练包括：

经所述第二神经网络，获取所述至少一组第二样本双目图像的视差图并输出；

获取所述第二神经网络输出的视差图与所述至少一组第二样本双目图像的视差图标签之间的第二差异，并基于所述第二差异对所述第二神经网络进行训练，直至满足第二训练完成条件。

15、根据权利要求 14 所述的方法，其特征在于，获取所述第二神经网络输出的视差图与所述至少一组第二样本双目图像的视差图标签之间的第二差异，并基于所述第二差异对所述第二神经网络进行训练，直至满足第二训练完成条件，包括：

获取所述第二神经网络输出的视差图与所述至少一组第二样本双目图像的视差图标

签之间在像素上的第二差异；

基于所述第二差异调整所述第二神经网络中网络参数的参数值，直至满足第二训练完成条件；

所述第二训练完成条件包括：所述第二差异小于第二预设阈值，和/或，对所述第二神经网络的训练次数达到第二预设次数。

16、根据权利要求 11-15 任一所述的方法，其特征在于，所述第一神经网络和所述第二神经网络的训练还包括：

利用第三样本集中的样本双目图像和第三样本图像对应的深度图标签对所述第一神经网络和所述第二神经网络进行训练；

其中，所述第三样本集包括至少一组第三样本双目图像和第三样本图像对应的深度图标签。

17、根据权利要求 16 所述的方法，其特征在于，所述利用第三样本集中的样本双目图像和第三样本图像对应的深度图标签对所述第一神经网络和所述第二神经网络进行训练，包括：

经所述第一神经网络，由所述至少一组第三样本双目图像中的第一图像，获取所述至少一组第三样本双目图像中的第二图像；

经所述第二神经网络，获取所述至少一组第三样本双目图像的视差图；

基于所述至少一组第三样本双目图像的视差图获取至少一组深度信息；

获取所述至少一组深度信息与所述至少一组第三样本双目图像的深度图标签之间的第三差异；

基于所述第三差异调整所述第一神经网络和所述第二神经网络中网络参数的参数值，直至满足第三训练完成条件；

所述第三训练完成条件包括：所述第三差异小于第三预设阈值，和/或，对所述第一神经网络和所述第二神经网络的训练次数达到第三预设次数。

18、一种深度估计装置，其特征在于，包括：

图像获取模块，用于以单张图片作为双目图像中的第一图像，经第一神经网络，基于所述第一图像获取所述双目图像中的第二图像；

立体匹配模块，用于经第二神经网络，通过对所述第一图像与所述第二图像进行双目立体匹配，获取所述第一图像对应的深度信息。

19、根据权利要求 18 所述的装置，其特征在于，所述图像获取模块包括：

第一子神经网络单元，用于对所述第一图像进行处理，输出 N 个通道的视差概率图；其中，每个通道的视差概率图表示所述第一图像上像素向第一水平方向偏移 i 个视差的概率， $i=0,1,\dots,N-1$ ， N 的取值为大于 1 的整数；

偏移单元，用于根据所述 N 个通道的视差概率图，将所述第一图像分别向第一水平方向偏移 i 个像素，得到 N 张偏移图；

点乘单元，用于将所述 N 张偏移图中的各张偏移图分别点乘对应通道的视差概率图中，得到 N 个点乘结果；

加法单元，用于将所述 N 个点乘结果基于像素进行叠加，得到所述第二图像。

20、根据权利要求 19 所述的装置，其特征在于，所述第一子神经网络单元，包括至少两个网络深度的网络层，用于：

分别通过至少两个网络深度的网络层对所述第一图像进行特征提取，获得至少两个尺度的特征图；

基于所述至少两个尺度的特征图分别获取至少两个分辨率的、 N 个通道的初步视差概率图；

分别针对每个通道，将所述至少两个分辨率的初步视差概率图放大至所述第一图像的分辨率下进行叠加，得到所述N个通道的视差概率图。

21、根据权利要求19或20所述的装置，其特征在于，所述第一图像中对应前景对象的像素在所述N个通道的视差概率图中的对应像素位置，在对应较大视差的通道的视差概率图中具有较大的概率值；所述第一图像中对应背景对象的像素在所述N个通道的视差概率图中的对应像素位置，在对应较小视差的通道的视差概率图中具有较大的概率值。

22、根据权利要求18-21任一所述的装置，其特征在于，所述立体匹配模块包括：

第一获取单元，用于获取用于表示所述双目图像中一个图像相对于另一个图像中像素的位置关系的相关系数；所述双目图像中一个图像包括所述第一图像或所述第二图像，所述另一个图像对应包括所述第二图像或所述第一图像；

生成单元，用于基于所述另一个图像与所述相关系数，生成所述第一图像与所述第二图像的视差图；

第三获取单元，基于所述视差图获取所述第一图像对应的深度信息。

23、根据权利要求22所述的装置，其特征在于，所述第一获取单元，包括：

第二子神经网络单元，用于分别对所述一个图像与所述另一个图像进行特征提取；

获取子单元，用于基于提取的所述一个图像的特征与所述另一个图像的特征，获取所述一个图像与所述另一个图像中像素的位置关系，并输出用于表示所述双目图像中一个图像相对于另一个图像中像素的位置关系的所述相关系数。

24、根据权利要求23所述的装置，其特征在于，所述生成单元，用于将所述另一个图像的特征与所述相关系数进行叠加，生成所述第一图像与所述第二图像的视差图。

25、根据权利要求24所述的装置，其特征在于，所述生成单元，包括：

叠加子单元，用于将所述另一个图像的特征与所述相关系数进行叠加，得到叠加结果；

融合子单元，用于提取所述叠加结果的特征，并将提取的所述叠加结果的特征与所述叠加结果进行融合，获得所述第一图像与所述第二图像的视差图。

26、根据权利要求23-25任一所述的装置，其特征在于，所述第三获取单元，用于基于所述视差图、拍摄所述第一图像的相机的焦距和所述双目图像对应的双目相机之间的距离，获取所述第一图像对应的深度信息。

27、根据权利要求18-26任一所述的装置，其特征在于，还包括：

获取模块，用于基于所述第一图像对应的所述深度信息与所述第二图像，获取所述第一图像中场景的三维空间场景。

28、根据权利要求22-27任一所述的装置，其特征在于，所述第一神经网络利用第一样本集中的样本双目图像预先训练而得，所述第一样本集包括至少一组第一样本双目图像；和/或，

所述第二神经网络利用第二样本集中的样本双目图像预先训练而得；所述第二样本集包括至少一组第二样本双目图像和视差图标签。

29、根据权利要求28所述的装置，其特征在于，所述第一神经网络，用于由所述至少一组第一样本双目图像中的第一图像，获取所述至少一组第一样本双目图像中的第二图像并输出；

所述装置还包括：

第一训练模块，用于获取所述第一神经网络输出的第二图像与所述至少一组第一样本双目图像中的第二图像之间的第一差异，并基于所述第一差异对所述第一神经网络进行训练，直至满足第一训练完成条件。

30、根据权利要求29所述的装置，其特征在于，所述第一训练模块，用于：

获取所述第一神经网络输出的第二图像与所述至少一组第一样本双目图像中的第二

图像之间在像素上的第一差异；

基于所述第一差异调整所述第一神经网络中网络参数的参数值，直至满足第一训练完成条件；

所述第一训练完成条件包括：所述第一差异小于第一预设阈值，和/或，对所述第一神经网络的训练次数达到第一预设次数。

31、根据权利要求 28-30 任一所述的装置，其特征在于，所述第二神经网络，用于获取所述至少一组第二样本双目图像的视差图并输出；

所述装置还包括：

第二训练模块，用于获取所述第二神经网络输出的视差图与所述至少一组第二样本双目图像的视差图标签之间的第二差异，并基于所述第二差异对所述第二神经网络进行训练，直至满足第二训练完成条件。

32、根据权利要求 31 所述的装置，其特征在于，所述第二训练模块，用于：

获取所述第二神经网络输出的视差图与所述至少一组第二样本双目图像的视差图标签之间在像素上的第二差异；

基于所述第二差异调整所述第二神经网络中网络参数的参数值，直至满足第二训练完成条件；

所述第二训练完成条件包括：所述第二差异小于第二预设阈值，和/或，对所述第二神经网络的训练次数达到第二预设次数。

33、根据权利要求 28-33 任一所述的装置，其特征在于，还包括：

第三训练模块，用于利用第三样本集中的样本双目图像和第三样本图像对应的深度图标签对所述第一神经网络和所述第二神经网络进行训练；

其中，所述第三样本集包括至少一组第三样本双目图像和第三样本图像对应的深度图标签。

34、根据权利要求 33 所述的装置，其特征在于，所述第一神经网络，用于由所述至少一组第三样本双目图像中的第一图像，获取所述至少一组第三样本双目图像中的第二图像；

所述第二神经网络，用于获取所述至少一组第三样本双目图像的视差图；

所述第三训练模块，用于：

基于所述至少一组第三样本双目图像的视差图获取至少一组深度信息；

获取所述至少一组深度信息与所述至少一组第三样本双目图像的深度图标签之间的第三差异；

基于所述第三差异调整所述第一神经网络和所述第二神经网络中网络参数的参数值，直至满足第三训练完成条件；

所述第三训练完成条件包括：所述第三差异小于第三预设阈值，和/或，对所述第一神经网络和所述第二神经网络的训练次数达到第三预设次数。

35、一种电子设备，其特征在于，包括：

存储器，用于存储可执行指令；以及

处理器，用于与所述存储器通信以执行所述可执行指令从而完成权利要求 1-17 任一所述方法的操作。

36、一种计算机程序，包括计算机可读代码，其特征在于，当所述计算机可读代码在设备上运行时，所述设备中的处理器执行用于实现权利要求 1-17 任一所述方法中操作的指令。

37、一种计算机可读存储介质，用于存储计算机可读的指令，其特征在于，所述指令被执行时实现权利要求 1-17 任一所述方法中的操作。