



DOI:10.12404/j.issn.1671-1815.2404031

引用格式:李晗,葛动元,姚锡凡.面向鱼眼相机标定和畸变处理的深度神经网络[J].科学技术与工程,2025,25(17):7260-7267.

Li Han, Ge Dongyuan, Yao Xifan. Deep neural network for fisheye camera calibration and distortion correction[J]. Science Technology and Engineering, 2025, 25(17): 7260-7267.

面向鱼眼相机标定和畸变处理的深度神经网络

李晗¹, 葛动元^{1*}, 姚锡凡²

(1. 广西科技大学机械与汽车工程学院, 柳州 545006; 2. 华南理工大学机械与汽车工程学院, 广州 510640)

摘要 针对鱼眼相机的传统标定过程烦琐并且不适用于日常场景图像的问题,提出了一种新的基于卷积神经网络的方法,可同时标定鱼眼镜头的内参并进行图像畸变校正。该方法通过预测不同畸变参数下像素点的位移量,从而提高鱼眼相机标定和图像畸变校正的精度;为了进一步提高模型精度和泛化性,在编码部分引入坐标注意力模块,增强对图像位置信息的关注度;最后为了增强图像的细节特征,在跨越连接部分设计了跨尺度融合模块。针对数据集稀缺的问题,还生成了一个新的大规模数据集,标有相应的畸变参数和畸变校正后的图像。实验结果表明:与其他鱼眼相机标定方法相比,重投影误差为0.312 pixel,标定的精度较高;与图像畸变处理方法相比,峰值信噪比(peak signal to noise ratio, PSNR)为38.055 dB,结构相似度(structural similarity, SSIM)为0.874,图像畸变校正的质量较好。

关键词 鱼眼相机标定; 畸变处理; 坐标注意力模块; 跨尺度融合模块

中图分类号 TP391; 文献标志码 A

Deep Neural Network for Fisheye Camera Calibration and Distortion Correction

LI Han¹, GE Dong-yuan^{1*}, YAO Xi-fan²

(1. School of Mechanical and Automotive Engineering, Guangxi University of Science and Technology, Liuzhou 545006, China;
2. School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China)

[Abstract] To address the cumbersome calibration process of fisheye cameras and its inapplicability to everyday scene images, a novel convolutional neural network(CNN)-based method was proposed that simultaneously calibrates the intrinsic parameters of fisheye lenses and corrects image distortion. The accuracy of fisheye camera calibration and image distortion correction was improved by predicting the displacement of pixel points under different distortion parameters. A coordinate attention module was introduced in the encoding part to enhance the model's accuracy and generalization ability to increase attention to image position information. Additionally, a cross-scale fusion module was designed in the skip connections to enhance image detail features. To address the issues of dataset scarcity and incomplete distortion parameter distribution, a new large-scale dataset labeled with corresponding distortion parameters and images after distortion correction was created. Experimental results show that compared to other fisheye camera calibration methods, this method achieves a reprojection error of 0.312 pixel, indicating the highest calibration accuracy. Additionally, compared to other image distortion correction methods, a peak signal to noise ratio(PSNR) of 38.055 dB and an structural similarity(SSIM) of 0.874 are achieved, indicating the best quality of image distortion correction.

[Keywords] fisheye camera calibration; distortion correction; coordinate attention module; cross-scale fusion module

鱼眼镜头通过光的折射来获得大的视场,其视场角能达到180°以上,因此在机器人导航、无人驾驶、深度估计^[1]和虚拟现实等众多领域有着广泛的应用前景。随着视场角的增大,桶形畸变也越来越严重^[2],需要对图像进行畸变处理,传统的鱼眼图像处理需要先求解鱼眼相机的内参和畸变系数^[3],根据标定结果对畸变图像校正,因此鱼眼相机的标定及图像畸变校

正是计算机视觉应用非常重要的预处理步骤。

在此之前,已经有很多关于传统鱼眼相机标定的研究。Geyer等^[4]提出了适用于折反式成像的统一球体模型,该模型最早用于折反式镜片的成像,如椭圆镜、双曲面镜和平面镜。Ying等^[5]扩展了统一球体模型,证明了该模型能够应用到鱼眼相机。Mei等^[6]提出了Mei模型,该模型在统一球体模型的基础

收稿日期:2024-05-30 修订日期:2025-03-10

基金项目:国家自然科学基金(51765007)

第一作者:李晗(2000—),男,汉族,河南安阳人,硕士研究生。研究方向:畸变处理、全景视觉。E-mail:lh1730696606@163.com。

*通信作者:葛动元(1970—),男,汉族,湖南邵阳人,博士,研究员。研究方向:机器视觉、机器学习。E-mail:gordon399@gxust.edu.cn。

上,考虑了径向畸变,并适用于多种镜面。Kannala等^[7]提出的鱼眼相机通用模型也称KB模型,该模型将入射光线和反射光线的关系展开成一系列奇数幂。KB模型根据针孔成像模型显示,不同之处在于KB模型的光线不再直线传播,而是在通过透镜后折射。Scaramuzza等^[8]开发了泰奥尔级数展开式表达图像模型也称Scaramuzz模型,并通过两步最小二乘线性最小化估计系数。上述的模型为鱼眼相机成像的投影失真模型,传统的鱼眼相机标定过程是根据鱼眼相机成像模型的几何关系建立含有相机内参和畸变系数的方程,从而求出最优解,传统相机标定过程是一个优化问题。这些传统标定算法每次标定时需要拍摄不同位置的棋盘格图片,标定过程烦琐,手动交互时需要大量的时间,标定的精度和棋盘格图片的角点检测算法、棋盘格的旋转位移矩阵的估计有关,并且不适用于日常图片。

随着深度学习的发展,为鱼眼相机标定和图像畸变校正提供了很多新的方案,越来越多的人使用端到端的网络进行鱼眼相机标定和畸变处理。Rong等^[9]首次利用卷积神经网络进行相机标定,他们将相机标定作为一个监督分类问题,使用卷积神经网络来学习输入的失真特征并预测畸变参数。Bogdan等^[10]提出了Deepcalib网络,利用Deepcalib网络来预测畸变系数和相机内参矩阵,在监督分类问题的基础上将相机标定扩展了监督回归问题。为了增强网络的扭曲感知,Xue等^[11]在网络中引入了几何特征,通过将鱼眼图片的曲线校正为直线来指导网络预测相机参数和进行畸变纠正,然而该方法并没有考虑到鱼眼图片中的曲线经过校正后仍是曲线的情况。在图像畸变校正方面,Li等^[12]通过编码-解码结构用来预测畸变图像和校正图像之间的位移场,并提出了一种鱼眼图像数据集的制作办法,该网络还考虑了旋转、剪切、透视和波浪失真。Kim等^[13]提出了具有自关注层的全局卷积神经网络对鱼眼图像进行校正。该方法采用扩展卷积神经网络来扩大接收域从输入图像中提取出全局特征,并利用自关注层来提取输入图像的最重要特征,从而完成畸变校正。Liao等^[14]设计了一个局部-全局关联估计网络,该网络通过学习有序畸变来近似真实畸变分布。所提出的有序畸变与图像特征的关系更为明确,增强了神经网络的畸变感知能力。Yang等^[15]提出了一种并行互补结构,该结构有两个并行的编码-解码结构,并利用不同层的外观流对图像特征进行预校正。因此,解码器可以很容易地用剩余的无失真信息重建合理的结果。

针对传统的鱼眼相机标定不适用于日常场景图

像以及鱼眼图像畸变校正不能提供准确参数的问题,现提出一种端到端的卷积神经网络,通过预测畸变图像与校正图像的畸变位移来计算鱼眼相机内参和畸变系数。该网络以U-net为基础,在编码部分引入坐注意力机制(coordinate attention, CA)来增强畸变图像不同的位置畸变位移量,在跨越连接部分引入设计的跨尺度融合模块(cross-scale fusion module, CS-FM)来增强目标特征和抑制背景噪声,并采用空洞空间金字塔池化(atrous spatial pyramid pooling, AS-PP)获取多尺度图片信息,设计重采样损失和结构相似度损失来提高网络的训练精度。

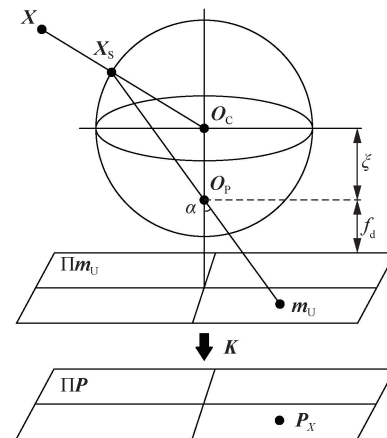
1 数据集构建

1.1 投影失真模型

鱼眼相机有很多投影失真模型,不同的投影失真模型的相机参数和畸变系数各不相同。如Geyer等^[4]的统一球体模型、Mei等^[6]的Mei模型、Kannala等^[7]的KB模型、Scaramuzza等^[8]的Scaramuzza模型。Mei模型的计算复杂度较高,需要在引入畸变系数的情况下再考虑非线性畸变系数,KB模型的精度与多项式的阶数有关,模型模拟针孔成像下光的折射,不适用于视场超过 180° 的鱼眼相机。

研究选择统一球体模型,首先,它是完全可逆的;其次,它可以拟合各种大小的视场和畸变系数,具有很好的泛化性;第三,投影过程和反投影过程都具有封闭型解,计算效率很高。统一球体模型如图1所示。

统一球体模型的投影步骤如下。



X 为三维空间坐标系中的一个空间点; O_C 为相机坐标系的原点; O_P 为等效投影中心; 射线 XO_C 与单位球相交于点 X_S ; ξ 为 O_C 与 O_P 之间的距离,也就是该模型的畸变系数; m_u 为点 X 投影到归一化平面 Πm_u 上的坐标; P_X 为图像平面的像素坐标; K 为相机的内参数矩阵; f_d 为等效焦距; α 为投影点到像素中心的距离与焦距 f_d 的比值

图1 统一球体模型

Fig. 1 Unified spherical model

步骤1 三维坐标系的空间点 X 以射线 XO_C 的形式投影到单位球面上,计算公式为

$$(X_S)_{O_C} = \frac{X_{O_C}}{\|X_{O_C}\|} = [x_s \ y_s \ z_s]^T \quad (1)$$

式(1)中: x_s 、 y_s 和 z_s 为空间点 X_S 在单位球面上的三维坐标值。

步骤2 将单位球面上的点 X_S 从相机坐标系转换为 $O_p = [0 \ 0 \ -\xi]^T$ 的坐标系下,计算公式为

$$(X_S)_{O_p} = [x_s \ y_s \ z_s + \xi]^T \quad (2)$$

步骤3 将单位球面上的点 X_S 以射线 $X_S O_p$ 的形式投影到归一面 Πm_U 上,得

$$m_U = \left[\frac{x_s}{z_s + \xi} \ \frac{y_s}{z_s + \xi} \ 1 \right]^T = [x \ y \ 1] \quad (3)$$

式(3)中: x 和 y 为归一化平面上点的二维坐标值。

步骤4 对归一化平面上的点进行透视变换使其投影到像素平面上,公式为

$$P_x = K m_U = \begin{bmatrix} f_1/d_x & \beta & u_0 \\ 0 & f_2/d_y & v_0 \\ 0 & 0 & 0 \end{bmatrix} m_U \quad (4)$$

式(4)中: f_1 和 f_2 为相机在两个方向上的焦距,默认 $f_1 = f_2$; $(u_0 \ v_0)$ 为图像平面的主点也就是图像得到中心点; β 为偏移系数,理论模型误差较小,因此将其设为零; d_x 为该方向上图像传感器的单位像素的大小; $f_d = f_1/d_x$ 为等效焦距,单位为像素(pixel),因此,可以将式(4)改写为

$$P_x = K m_U = \begin{bmatrix} f_d & 0 & u_0 \\ 0 & f_d & v_0 \\ 0 & 0 & 0 \end{bmatrix} m_U \quad (5)$$

归一化平面上的点到单位球面的反投影函数可以表示为

$$\begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} = \frac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{1 + x^2 + y^2} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \xi \end{bmatrix} \quad (6)$$

1.2 鱼眼图像数据集

针对数据集稀缺的问题,还生成了一个新的大规模数据集,由于每个鱼眼相机的参数和畸变系数是固定值,因此使用不同的鱼眼相机拍摄图像作为训练集并不能满足网络训练的要求,所以数据集一般采用将普通的无畸变图像进行畸变处理得到鱼眼图像的方式进行制作^[16]。传统的鱼眼图像生成方法依靠鱼眼相机的投影失真模型,根据不同的投影模型添加不同的畸变系数来模拟鱼眼成像的效果。然而合成得到的鱼眼图像本身视场并没有增

加,得到的是非现实的数据,另一方面,生成的畸变图像的像素坐标是由原本的图像坐标变化得到的,因此畸变图像的像素位置不具有整齐排列性,需要重新进行排列,这就导致了图像像素的误差,造成了过多的失真。

提出一种基于全景图像的鱼眼图像数据集生成办法,即采用图像平面到三维空间映射的方式生成鱼眼图像。首先建立以相机中心 O_C 为原点的相机坐标系,将全景图像导入相机坐标系中的单位球面上,根据统一球体模型的投影方程,求解出像素坐标点在单位球体上的坐标 X_S ,将 X_S 代入单位球面的全景图像上,采用双线性插值的办法求解出图像各个像素点的红绿蓝(red-green-blue, RGB)值,从而得到畸变图像。考虑到数据集的广泛性,可以对单位球面上的全景图像进行旋转,从而得到多个角度的图片,由于不同分辨率图片的图片中心不一样,当等效焦距 f_d 为定值范围时,对高分辨率图片的泛化性效果不好,因此选取角度 α 为变量来达到控制等效焦距 f_d 的目的。 α 从 0.2 到 0.6 每 0.1 递增,畸变系数 ξ 由 0.5 到 2 每 0.1 递增,由此生成 17 677 张大小为 (512 pixel \times 512 pixel) 的畸变图片数据集,考虑边缘黑边问题,将图像裁剪到 (256 pixel \times 256 pixel) 大小,部分鱼眼图像数据集如图 2 所示。



图2 部分数据集图片

Fig.2 Some dataset images

2 网络结构

设计的网络结构如图 3 所示,由 4 个部分组成,将鱼眼畸变图像[图 4(a)]作为特征图输入深度神经网络[图 4(b)]中,通过网络预测不同畸变参数下像素点位移量[图 4(c)],从而求解出相机内参和畸变系数,最后根据像素点位移量对畸变图像[图 4(a)]进行像素变换得到模型校正后的无畸变图像[图 4(d)]。使用改进后的 U-net 网络作为主干网络,网络在编码部分引入坐标注意力机制来增强畸变图像不同的位置的畸变位移,使模型有更好的拟合性能,在跨越连接部分设计跨尺度融合模块

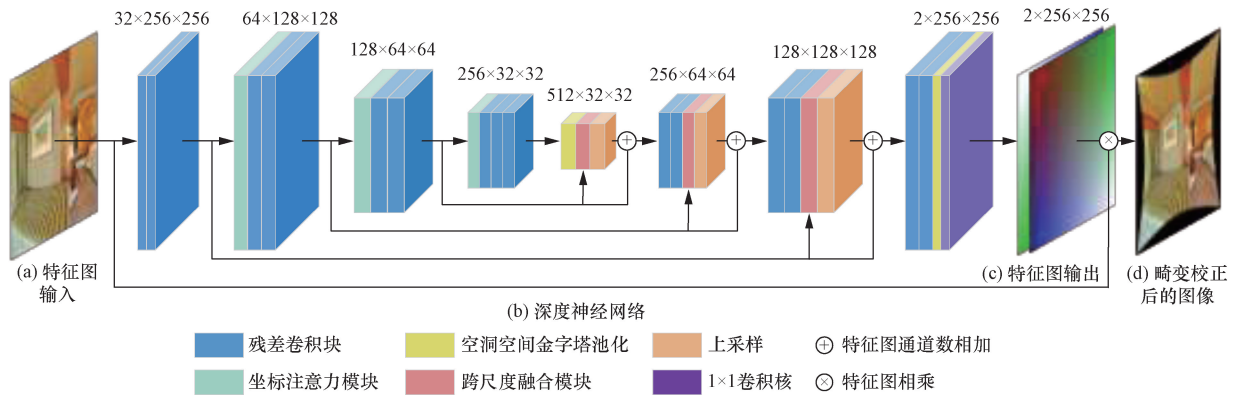
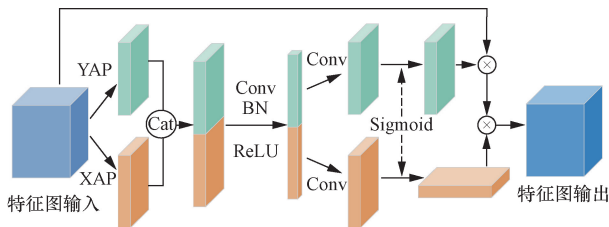


图3 网络结构

Fig. 3 Network architecture



XAP为X方向上最大池化;YAP为Y方向上的最大池化;Cat为特征图拼接;Conv为卷积;BN为批量归一化;ReLU为ReLU激活函数;Sigmoid为Sigmoid激活函数

图4 坐标注意力模块

Fig. 4 Coordinate attention

来增强目标特征和抑制背景噪声,并采用ASPP获取多尺度图片信息,对所采用的模块组合进行了消融实验,以证明网络模块的组合对性能的影响。

2.1 网络主体结构

U-net网络由Ronneberger等^[17]提出,所提网络在U-net网络上进行改进,网络结构如图3所示,主要包括编码器、解码器和跳跃连接3个部分组成。在网络中设计了残差块来防止梯度爆炸,每个残差块由3x3卷积层、BN(batch normalization)层和ReLU激活函数组成。在编码器部分对图像进行4次下采样处理,每次下采样后特征图尺寸减半,通道数加倍;并在下采样后引入坐标注意力机制,增强信息提取能力,针对畸变严重的边缘区域进行提取,并抑制图像中心部分的轻微畸变。解码器部分引入了跨尺度融合模块,利用跳跃连接将编码器的特征进行融合,增强图像局部细节信息,同时抑制噪声;使用3次2x2反卷积进行上采样,逐次减少通道数并扩大特征图尺寸,恢复到与输入图像相同的大小,最后经过ASPP和1x1卷积层生成与输入图像尺寸相同的二维坐标位移量,该坐标位移量可以求解出鱼眼相机内参和畸变系数,并将该二维位移量与输入图像相乘得到畸变校正后的图像。

2.2 坐标注意力模块

在畸变图像中,畸变程度与像素点离图像中心的距离有关,越靠近图像中心畸变程度越弱,畸变程度由图像中心向四周逐渐增加。因此考虑到畸变量与图像像素点的位置关系,在编码层的部分引入了坐标注意力模块,结构如图4所示,利用其强大的特征提取能力,增加畸变程度较大的图像边缘地区的注意力,对畸变程度较小的中心区域进行抑制,从而提高网络的精度。首先,对输入特征图在X和Y方向分别进行平均池化得到两个单通道结果并将其拼接在一起;接着经过一个3x3卷积层、BN层和ReLU激活函数后,拆分成X方向和Y方向的特征图;再通过各自的3x3卷积层和Sigmoid激活函数得到X方向和Y方向上的权重系数矩阵,最后将该权重系数矩阵与分别与特征图相乘得到输出特征图。

2.3 跨尺度融合模块

低水平特征提供了更详细的信息,如边界和空间结构,但对背景噪声很敏感,相比之下,高级特征包含更多的语义信息,有助于定位和抑制噪声。通过跨尺度融合模块可以在增强图像局部细节信息的同时衰减噪声的影响。跨尺度特征融合模块如图5所示,将高水平特征作为Input1,低水平特征作为Input2,将两者通过BN层和一个卷积核为3x3的卷积层,再对Input2的特征进行一个2x2的最大

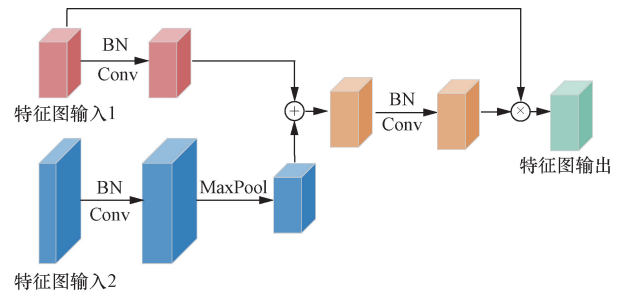


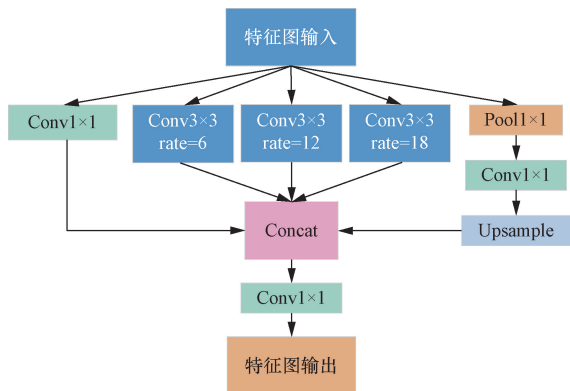
图5 跨尺度融合模块

Fig. 5 Cross-Scale fusion module

池化(Maxpool),使两者的通道数和大小相等,然后再相加,并不改变其通道数和特征图大小,融合后的特征通过BN层和一个卷积核为 3×3 的卷积层再与Input1进行相乘,最后得到大小和通道数与Input1一样的输出特征图。

2.4 空洞空间金字塔池化

空洞空间金字塔池化ASPP的理念源自空间金字塔池化,它能成功地对多个尺度的特征进行重新采样^[18],本文网络模型在网络瓶颈部分和解码器的输出部分引入ASPP,它能够增加网络的感受视野,并在不同尺度上捕获全局信息。ASPP结构如图6所示,输入特征通过一个 1×1 的卷积层来减少通道数,从而降低计算复杂度;然后构建多个平行的空洞卷积层,不同膨胀因子(rate)对应不同的卷积核感受野,因此使用不同膨胀因子的空洞卷积层可以提取不同尺度下的特征,在不增加参数数量的情况下扩展卷积核的感受野,从而捕捉到更大范围的特征;添加全局平均池化层有助于捕捉全局上下文信息,增强网络的全局感知能力。最后,将ASPP各层的输出进行拼接,并通过一个 1×1 的卷积层来整合来自不同尺度的信息,并减少通道数,从而得到最终的输出特征图。



Conv 为卷积,其中 1×1 为卷积的尺寸,rate为膨胀因子;
Concat 为通道数拼接;Upsample 为上采样

图6 空洞空间金字塔池化

Fig. 6 Atrous spatial pyramid pooling

2.5 损失函数

采用的损失函数表达式为

$$L = L_1 + \lambda L_{SSIM} \quad (7)$$

式(7)中: L_1 为平均误差损失; L_{SSIM} 为结构相似度损失; λ 为比例系数。

使用 L_1 损失作为损失函数,其定义为

$$L_1 = \frac{1}{n} \sum \| \mathbf{V}^{gt} - \mathbf{V}^{pred} \|_2 \quad (8)$$

式(8)中: n 为坐标差向量元素总数; \mathbf{V}^{pred} 为网络预测的二维坐标差向量; \mathbf{V}^{gt} 为真实的二维坐标

差向量。然而,当仅使用 L_1 损失时产生的图片有阴影和黑点,因此引入 L_{SSIM} 损失, L_{SSIM} 损失用于生成比较流畅的图片,其中结果相似度SSIM定义为

$$SSIM(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (9)$$

式(9)中: μ_x 为 x 的平均值; μ_y 为 y 的平均值; σ_x 为 x 的方差; σ_y 为 y 的方差; σ_{xy} 为 x 和 y 的协方差; C_1 和 C_2 为用来维持稳定的常数, C_1 为6.5025, C_2 为58.5225。SSIM的范围为 $0 \sim 1$,其值越高,图像畸变处理的效果越好,图像越流畅,因此 L_{SSIM} 定义为

$$L_{SSIM} = 1 - SSIM(x, y) \quad (10)$$

将 λ 设置为0.5,训练目标是最小化这个损失函数。

3 实验与分析

实验在NVIDIA GeForce GTX 3070,CPU为12th Gen Intel(R) Core(TM) i5-12600KF的服务器上进行;软件部分使用Pycharm作为IDE,通过Python编程语言以及深度学习Pytorch框架进行程序设计。模型训练的优化器选择Adam,初始化学率为0.0001,设置训练轮次Epoch为200,Batch size为16。

3.1 评价指标

采用3个评价指标,第一个评价指标为重投影误差(reprojection error, RE),用重投影误差来检验传统鱼眼相机标定算法的精度。根据求解出的相机参数和畸变系数对三维空间点重新进行投影,得到计算的图像点,计算得到的图像点和检测得到的图像点的差值就是重投影误差。重投影误差的值越小表示相机标定的精度越高。另外两个指标为峰值信噪比(peak signal to noise ratio, PSNR)和结构相似度(structural similarity, SSIM),这两个指标为图像校正研究常用的评价指标^[19]。其中PSNR主要是衡量处理后的图像和原始图像中每个像素值的差异,值越大代表处理后的图像质量越好。PSNR仅基于像素间的差异,未考虑人眼对图像的感知特性,不能有效反映图像的结构信息和局部细节,比如模糊、噪声和阴影,因此引入SSIM,SSIM会比较两幅图像的亮度、对比度和结构信息,范围为 $0 \sim 1$,值越大表示处理后的图像更接近原始图像。SSIM能够更好地反映人眼对图像质量的主观感受和图像的视觉效果。

3.2 对比分析

为了验证本文方法在鱼眼相机标定和畸变图

像校正的有效性,将和其他方法进行对比。在畸变图像校正方面,选取了 Bogdon 的 Deepcalib 模型^[10]和两种最新的基于深度学习的 Li 等^[12]和 Yang 等^[15]的畸变图像校正算法,并使用生成的数据集对其进行了训练和测试。Deepcalib 模型的主体是用 Inceptionv3 网络来预测畸变系数和相机内参矩阵,并根据得到的参数和畸变系数来进行图像校正, Li 等^[12]通过编码-解码结构来预测畸变图像和无畸变图像之间的位移场。Yang 等^[15]利用两个并行的编码-解码结构,一个生成不同层的外观流,一个用于畸变图像处理,在跳跃连接中嵌入校正层,并利用不同层的外观流对图像特征进行预校正。这 3 种方法都具有一定的代表性。对测试集中的 1 301 张图片进行几何校正实验,部分实验结果如图 7 所示,根据 2 个评价指标 PSNR 和 SSIM 来评价不同方法的校正效果,结果如表 1 所示。

表 1 不同方法的评价结果

Table 1 Evaluation results of different methods

方法	PSNR/dB	SSIM
Bogdon 等 ^[10]	36.501	0.641
Li 等 ^[12]	36.764	0.806
Yang 等 ^[15]	36.580	0.839
本文方法	38.055	0.874

由表 1 可知,本文方法的 PSNR 为 38.055 dB, SSIM 的值为 0.874。Bogdon 等^[10]和 Li 等^[12]在 PSNR 和 SSIM 方面表现不佳, Yang 等^[15]的 SSIM 值与本文方法相当,但在 PSNR 上稍差。因此选取 Yang 等^[15]的方法对局部区域的细节处理方面进行对比,结果如图 8 所示,由图 8(a)可知, Yang 等^[15]的方法在进行畸变处理后期会有阴影,而本文方法在局部区域的细节处理方面具有更好的质量。因此,本文方法校正后的图像的质量以及校正精度,每个像素值的误差,亮度,对比度和结构高于其他算法。

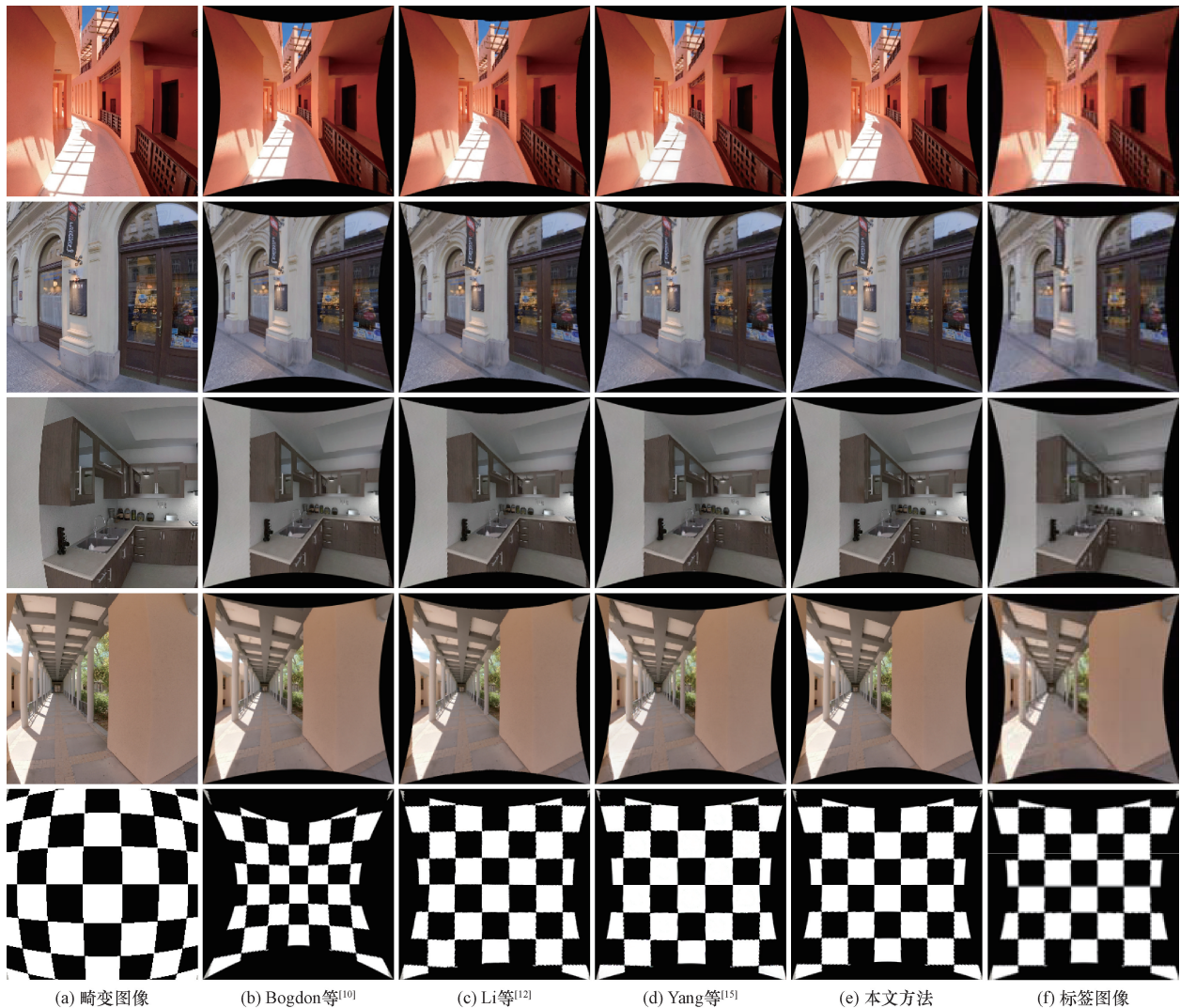


图 7 不同方法的测试集样例图

Fig. 7 Sample images from the test set of different methods

在鱼眼相机标定方面,由于传统方法鱼眼相机标定需要采集不同方位的棋盘格图像,因此首先从数据集中选取一定的棋盘格图像进行鱼眼相机标定,棋盘格图像如图9所示。传统方法的鱼眼相机标定是根据鱼眼图像的成像投影模型建立相关的等式方程,从而求解出相机参数和畸变系数,不同模型的畸变系数不相同。因此,选用 Mei 等^[6]的 Mei 模型、Kannala 等^[7]的 KB 模型、Scaramuzza 等^[8]的 Scaramuzza 模型和基于深度学习的 Bogdan 等^[10]的 Deepcalib 模型进行对比。Deepcalib 模型和本文方法也都使用该图9棋盘格图像进行标定,从而对标定结果进行定量结果分析。Mei 模型和 KB 模型的标定选用 OpenCV 包中的标定算法,Scaramuzza 模型标定选用其提供的 MATLAB 包^[20]进行标定,不同模型的相机参数和畸变系数如表2所示。

根据表2的标定结果,将已知的三维世界坐标点重新投影,得到图像平面上的计算像素点,图像平面上的计算像素点与实际像素点的差值便是重投

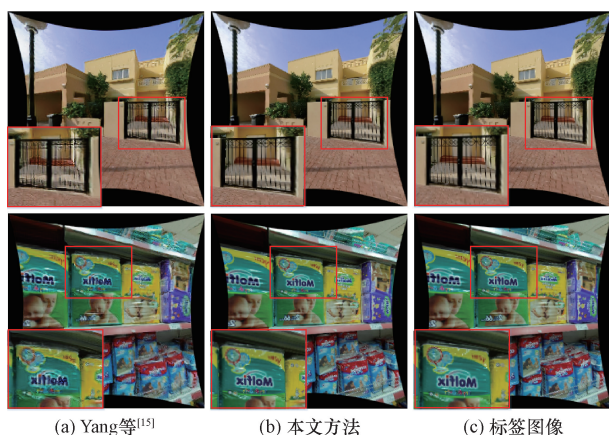


图8 局部细节处理对比图

Fig. 8 Comparison of local detail processing

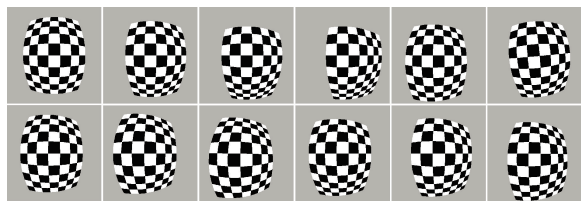


图9 棋盘格图像

Fig. 9 Checkerboard image

表2 标定结果

Table 2 Calibration results

模型	f_d/mm	畸变系数
KB	134.70	0.147, 0.014 8, 0.013 7, 1.44×10^{-5}
Mei 等 ^[6]	222.50	0.655
Bogdon 等 ^[10]	184.16	1.175
Scaramuzza 等 ^[8]	69.74	0, 0.002 1, 1.08×10^{-5} , 0
本文方法	106.66	0.595

影误差 (reprojection error, RE)。不同方法的重投影误差如表3所示,可以看出,本文方法在进行鱼眼相机标定的精度高于其他方法。

表3 不同方法的重投影误差

Table 3 Reprojection errors of different methods

模型	KB	Mei 等 ^[6]	Bogdon 等 ^[10]	Scaramuzza 等 ^[8]	本文方法
RE/pixel	0.371	0.369	16.01	0.693	0.312

3.3 消融实验

通过消融实验研究,以评估所提出的各种模块和损失函数的有效性。保持其他训练条件相同,分别移除坐标注意力机制 (CA)、跨尺度融合模块 (CSFM)、空洞空间金字塔池化 (ASPP) 和 L_{SSIM} 损失,对于 ASPP 使用 2×2 的反卷积来进行替换,使用重投影误差 (RE)、峰值信号比 (PSNR) 和结构相似度 (SSIM) 来验证其对网络预测精度和畸变图像校正的影响。实现结果如表4所示,可以看出,每个模块和 L_{SSIM} 损失函数在提高网络的标定精度和畸变校正效果上都起着非常重要的作用。

表4 消融实验结果

Table 4 Ablation experiments Results

CA	CSFM	ASPP	L_{SSIM}	RE/pixel	PSNR/dB	SSIM
×	√	√	√	0.478	35.847	0.798
√	×	√	√	0.425	36.735	0.818
√	√	×	√	0.346	37.892	0.841
√	√	√	×	0.681	34.724	0.746
√	√	√	√	0.312	38.055	0.874

注:√表示采用该模块,×表示没有采用该模块。

3.4 真实鱼眼图像的畸变校正

在该系列的研究中,根据训练完成的网络,对鱼眼相机拍摄的真实鱼眼图像进行畸变校正实验,以验证本文方法的有效性,即泛化性。并选取 Li 等^[12]、Yang 等^[15]和本文方法进行对比。首先利用 APEXEL-HB195 鱼镜头获取的真实鱼眼图像,并将图像的尺寸压缩到 (256 pixel × 256 pixel),用不同方法对真实鱼眼图像进行畸变校正,校正结果如图10所示。可以看出,本文方法对真实鱼眼图像的校正效果优于其他两种方法,泛化性能较好,可以直接对真实鱼眼图像进行畸变校正处理。

4 结论

针对鱼眼相机标定过程烦琐并且不适用于日常图像以及图像去畸变过程中不能计算相机参数和畸变系数的问题,提出了一种新的基于卷积神经网络的方法,对鱼镜头进行标定的同时进行图像畸变校正,该方法通过卷积神经网络来预测不同畸

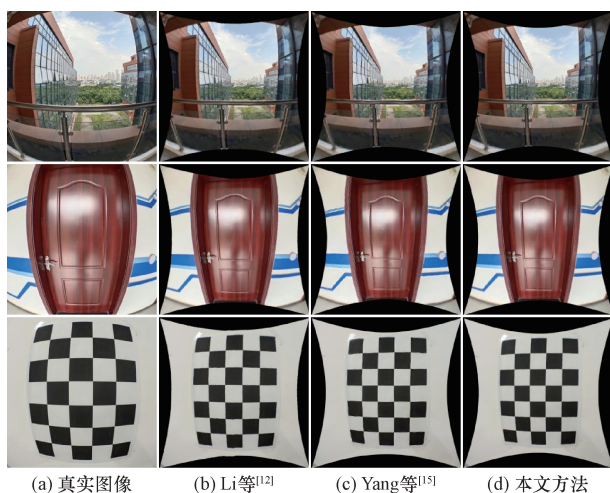


图 10 真实鱼眼图像样例图

Fig. 10 Sample images of real fisheye pictures

变参数引起的像素点的位移量,从而进行鱼眼相机标定和图像畸变校正。在提高了精度的情况下增加了图像校正的泛化性。实验结果表明,与传统相机标定的方法相比,本文方法的精度最高,重投影误差为 0.312 4。与最新基于深度学习的畸变校正方法的对比实验表明,本文方法校正后的图像的质量以及校正精度,每个像素值的误差、亮度、对比度和结构均高于其他算法,其中 PSNR 为 38.055 dB, SSIM 为 0.877。在真实鱼眼图像上的实验进一步证明了本文方法的泛化性。

参 考 文 献

- [1] 张玉亮, 赵智龙, 付炜平, 等. 融合边缘语义信息的单目深度估计[J]. 科学技术与工程, 2022, 22(7): 2761-2769.
Zhang Yuliang, Zhao Zhilong, Fu Weiping, et al. Integrating spatial semantic information for monocular depth estimation[J]. Science Technology and Engineering, 2022, 22(7): 2761-2769.
- [2] 杨宇, 赵成星, 张晓玲. 鱼眼相机的视觉标定及畸变校正[J]. 激光杂志, 2020, 41(9): 20-23.
Yang Yu, Zhao Chengxing, Zhang Xiaoling. Visual calibration and distortion correction of fish eye cameras[J]. Laser Journal, 2020, 41(9): 20-23.
- [3] 张静, 柴兴华, 裴春琴, 等. 一种适用于广角、鱼眼及折反射系统的标定方法[J]. 科学技术与工程, 2018, 18(5): 252-257.
Zhang Jing, Chai Xinghua, Pei Chunqin, et al. A calibration method of wide-angle, fisheye and catadioptric system[J]. Science Technology and Engineering, 2018, 18(5): 252-257.
- [4] Geyer C, Daniilidis K. Catadioptric projective geometry[J]. International Journal of Computer Vision, 2001, 45: 223-243.
- [5] Ying X, Hu Z. Can we consider central catadioptric cameras and fisheye cameras within a unified imaging model[C]//Computer Vision-ECCV 2004. Prague: Springer, 2004: 442-455.
- [6] Mei C, Rives P. Single view point omnidirectional camera calibration from planar grids[C]//Proceedings 2007 IEEE International Conference on Robotics and Automation. Rome: IEEE, 2007: 3945-3950.

- [7] Kannala J, Brandt S. A generic camera calibration method for fish-eye lenses[C]//Proceedings of the 17th International Conference on Pattern Recognition. Cambridge: IEEE, 2004, 1: 10-13.
- [8] Scaramuzza D, Martinelli A, Siegwart R. A flexible technique for accurate omnidirectional camera calibration and structure from motion[C]//Fourth IEEE International Conference on Computer Vision Systems (ICVS'06). New York: IEEE, 2008: 297-304.
- [9] Rong J, Huang S, Shang Z, et al. Radial lens distortion correction using convolutional neural networks trained with synthesized images[C]//Computer Vision-ACC'V 2016. Taipei, China: Springer, 2017: 35-49.
- [10] Bogdan O, Eckstein V, Rameau F, et al. DeepCalib: a deep learning approach for automatic intrinsic calibration of wide field-of-view cameras[C]//Proceedings of the 15th ACM SIGGRAPH European Conference on Visual Media Production. London: ACM, 2018: 1-10.
- [11] Xue Z, Xue N, Xia G S, et al. Learning to calibrate straight lines for fisheye image rectification[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 1643-1651.
- [12] Li X, Zhang B, Sander P V, et al. Blind geometric distortion correction on images through deep learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 4855-4864.
- [13] Kim B, Lee D, Min K, et al. Global convolutional neural networks with self-attention for fisheye image rectification[J]. IEEE Access, 2022, 10: 129580-129587.
- [14] Liao K, Lin C, Zhao Y. A deep ordinal distortion estimation approach for distortion rectification[J]. IEEE Transactions on Image Processing, 2021, 30: 3362-3375.
- [15] Yang S, Lin C, Liao K, et al. Progressively complementary network for fisheye image rectification using appearance flow[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 6348-6357.
- [16] Feng H, Wang W, Deng J, et al. SimFIR: a simple framework for fisheye image rectification with self-supervised representation learning[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2023: 12418-12427.
- [17] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-assisted Intervention-MICCAI 2015. Berlin: Springer, 2015: 234-241.
- [18] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834-848.
- [19] 宋巍, 师丽彪, 耿立佳, 等. 基于改进 U-Net 网络的图像混合畸变校正方法[J]. 液晶与显示, 2023, 38(11): 1580-1589.
Song Wei, Shi Libiao, Geng Lijia, et al. Hybrid distortion image correction method based on improved U-Net networks[J]. Chinese Journal of Liquid Crystals and Displays, 2023, 38(11): 1580-1589.
- [20] Scaramuzza D, Siegwart R. A practical toolbox for calibrating omnidirectional cameras[J]. Vision Systems: Applications, 2007, 17: 297-310.