

基于超分辨率角点检测算法的 双目相机高精度标定方法

孙婷^{1,2}, 陈娜¹, 孟然², 龙潜^{3*}, 谢启伟⁴

¹湖北大学数学与统计学学院, 湖北 武汉 430062;

²北京中科慧眼科技有限公司, 北京 100023;

³中国科学院云南天文台, 云南 昆明 650000;

⁴北京工业大学北京现代制造业发展研究基地, 北京 100124

摘要 双目相机标定是研究立体视觉的基础工作, 标定的精度是视觉测量精度的关键。图像角点提取是相机标定的基础, 但在现实应用场景中, 外界影响使获取的图像不清晰, 导致检测到的角点精度低, 从而影响标定的精度。因此, 提出一种基于超分辨率亚像素角点检测的端到端算法, 从特征级解决低质量角点检测问题。首先, 应用盲超分辨率估计低分辨率图像模糊核, 融合低分辨率图像特征重建出高分辨率图; 然后, 在此基础上得到角点亚像素位置; 最后对双目相机进行高精度标定, 并用测距实验对其进行检验。实验结果表明, 所提基于超分辨率的亚像素角点检测方法在真实场景下具有优越性。

关键词 成像系统; 相机标定; 角点检测; 超分辨率; 亚像素; 自监督; 深度学习

中图分类号 TP751.1

文献标志码 A

DOI: 10.3788/LOP223305

High-Precision Binocular Camera Calibration Based on Super-Resolution Corner Point Detection Algorithm

Sun Ting^{1,2}, Chen Na¹, Meng Ran², Long Qian^{3*}, Xie Qiwei⁴

¹School of Mathematics and Statistics, Hubei University, Wuhan 430062, Hubei, China;

²Beijing Smarter Eye Technology Co., Ltd., Beijing 100023, China;

³Yunnan Observatories, Chinese Academy of Sciences, Kunming 650000, Yunnan, China;

⁴Research Base of Beijing Modern Manufacturing Development, Beijing University of Technology, Beijing 100124, China

Abstract Binocular camera calibration is the foundation of stereo vision research, and its accuracy is the key to achieve precision in vision measurement. The basis of camera calibration is image corner extraction, but in real applications, the accuracy of detected corner points is often low due to unclear acquisition of images due to external influences, which affects calibration accuracy. To solve the problem of low-quality corner detection in terms of the feature level, an end-to-end algorithm based on super-resolution subpixel corner detection is proposed. First, the fuzzy kernel of a low-resolution image is estimated using the blind hyperspectral part, and its features are fused to reconstruct its high-resolution version. Subsequently, the subpixel position of the corner points is obtained. Finally, a binocular camera is calibrated with high accuracy and tested via ranging experiments. Experimental results show that the proposed subpixel corner detection method based on super-resolution has advantages in real scenarios.

Key words imaging systems; camera calibration; corner point detection; super resolution; sub-pixel; self-supervision; deep learning

1 引言

近年来, 随着计算机视觉技术在无人驾驶等人工

智能领域的应用, 该技术受到了越来越多的关注。双目立体视觉是计算机视觉研究中非常重要的内容, 类似于人类双目, 基于视差理论可以实现三维测距与定

收稿日期: 2022-12-12; 修回日期: 2023-01-08; 录用日期: 2023-02-06; 网络首发日期: 2023-02-16

基金项目: 国家重点研发计划(2018AAA0103103)、国家自然科学基金(61673381)

通信作者: *longqian@ynao.ac.cn

位^[1-2]。而对双目摄像机进行标定是研究立体视觉的前提,标定的精度直接影响后续三维重建等的效果。基于目标的相机标定方法是指用一个结构已知、精度很高的标定物或控制场作为空间参照物,由三维世界的空间点和二维的图像上的点之间的对应关系形成几何约束关系,然后通过优化算法来求取相机参数^[3]。然而视觉传感器在获取图像时,由于各种光线、噪声、抖动等因素的影响,获得的图像噪声高、质量差、分辨率低,呈现的视觉效果差^[4]。因此要实现高精度双目标定,亟待解决的问题是提高图像分辨率,从而提高标定板图检测精度^[5]。

图像分辨率是指单位长度的像素数,是图像质量的重要指标。在单位长度内的像素数越多,图像质量越高,则图中存储的信息量越大,目标细节就越丰富。最初常用插值方法对图像进行放大,从而增加图像的像素数,即利用像素周围的灰度值来计算像素点的灰度值,计算量小但增加的图像细节信息少。因此,研究者提出超分辨率重建方法。图像的降质过程是由原始高分辨率图像经过模糊、下采样、噪声干扰等退化过程获取到低分辨率(LR)图像。而超分辨率(SR)重建^[6-8]则是该过程的逆过程,即从退化后的低分辨率图像重建出具有更多纹理及细节的高分辨率图像。

近年来,深度学习在超分辨率图像的应用得到广泛提升^[9-12]。super resolution convolutional network (SRCNN)是基于深度学习超分辨率的开山之作。2014年,Dong等^[13]首次根据三层的卷积神经网络模拟低分辨率图像和高分辨率图像之间的非线性映射。2016年,Shi等^[14]提出 efficient sub-pixel convolutional neural network(ESPCN),该网络在低分辨率图像上提取特征,计算得到高分辨率图像,但其存在无法保留图像原有高频信息的问题。generative adversarial network(GAN)的出现很好地克服了高频信息丢失的问题,在此基础上,Ledig等^[15]提出 super-resolution generative adversarial network(SRGAN)。随着网络深度的不断增加,参数不断增加,性能受到影响。且这些网络都是非盲超分辨率的,只能在提前已知正确的退化时产生预期的效果。因此,考虑到现实环境的诸多不确定性和复杂性,盲超分辨率方法——从未知和复杂的退化造成的低分辨率图像中恢复高分辨率的图像的方法是更好的选择。blind super-resolution kernel estimation using an internal-GAN(KernelGAN)^[16]利用GAN的特性,提出用来估计模糊核的分布,对抗模拟生成真实世界的模糊核。然而,它只适用于空间不变的核估计,而不能估计图像单位像素的核。为了解决这个问题,针对图像每个像素点估计一个模糊核,Liang等^[17]提出 mutual affine network(MANet)。

由于角点具有旋转不变性等优点,因此常被用于图像信息的载体。角点也称为“特征点”“兴趣点”等,通常是指图像中梯度变化率非常大的像素点及图像边

缘曲线曲率^[18]是极大值的像素点,反映图像的局部特征^[19-20]。经典的角点提取算法有 Moravec、Harris、smallest univalue segment assimilating nucleus(SUSAN)等。Moravec算法将角点定义为具有低“自相关性”的点。Harris角点检测算法由局部窗口在图像上的移动来判断图像灰度是否发生较大的变化,从而判断其是否存在角点。SUSAN算法用圆形的模板来对图像的像素进行遍历,计算圆形模板区域内像素的灰度值与模板灰度值的差是否小于设定阈值来判断核值相似区(USAN)值的大小,若USAN值小于模板区域像素值的1/2,则认为该点为角点。角点提取算法主要分为基于手工设计和基于深度学习的角点提取算法。基于手工设计的角点提取算法如 scale invariant feature transform(SIFT)^[21]等,通过数学公式对图片进行进化和抽象来提取信息,其鲁棒性和泛化性相比于大规模数据集驱动的深度学习具有天然劣势。因此,基于深度学习提取角点的算法更受大家的欢迎,其鲁棒性以及泛化能力也比前者好。Tian等^[22]提出了利用卷积神经网络在欧几里得空间学习图像块中角点的L2-Net。DeTone^[23]提出的 SuperPoint是端到端的学习特征网络,输入一张图片可以直接输出角点。SuperPoint采用一种自监督方式训练网络提取角点,输出的角点精度较高且网络结构较其他网络更为简单,降低工程的复杂度也减小误差的累计。

本文提出一种基于超分辨率亚像素角点检测的端到端算法,应用盲超部分估计低分辨率图像模糊核,融合低分辨率图像特征重建出高分辨率图,在此基础上得到角点亚像素位置,最后对双目相机进行高精度标定。实验结果表明,所提方法在真实场景下具有优越性。

2 相机标定原理

张正友标定法^[24]相对于传统方法操作更简便,又较自标定方法精度高,因此采用这一标定法。在双目立体视觉系统^[25]中,其成像原理如图1所示。其中,世界坐标系上点 $P(X_w, Y_w, Z_w)$ 在左右相机的投影点分别为 $p_1(u_L, v_L)$ 和 $p_2(u_R, v_R)$,左右相机投影中心的连线为基线 b ,相机焦距为 f ,左右相机的中心点分别是 O_R 和 O_L 。令投影点的齐次坐标为 $\tilde{p}(u, v, 1)^T$,世界坐标点的三维齐次坐标为 $\tilde{P}(X, Y, Z, 1)^T$,由小孔成像模型可知,三维点 \tilde{P} 与其投影点 \tilde{p} 关系为

$$s\tilde{p} = K [R \ t] \tilde{P}, \quad (1)$$

$$K = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

式中: s 为任意的比例因子; $[R \ t]$ 为外参,由世界坐标系到相机坐标系的旋转与平移矩阵构成; K 为相机内参矩阵; (u_0, v_0) 是图像坐标系的原点; α 和 β 是图像坐

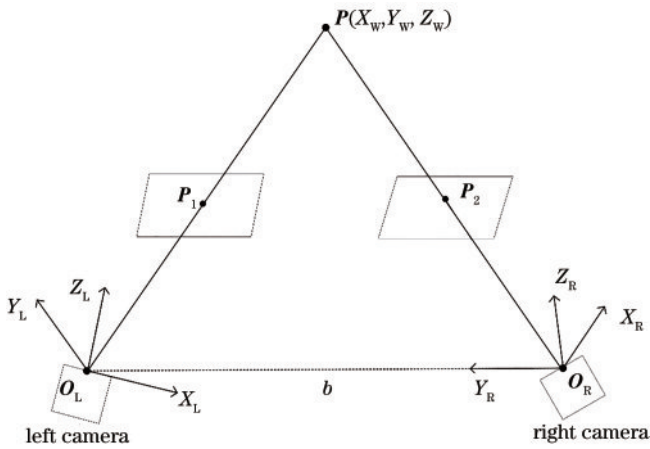


图 1 双目立体视觉成像原理
Fig. 1 Principles of binocular stereo vision imaging

标系下 w 轴的比例因子; γ 是描述图像坐标系两轴的倾斜参数。

在张氏标定法中, 棋盘格标定板图平面假设为 $Z=0$ 平面, 并将旋转矩阵 \mathbf{R} 中的第 i 列元素表示为 \mathbf{r}_i , 则其关系式为

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} [\mathbf{r}_1 \mathbf{r}_2 \mathbf{r}_3 \mathbf{t}] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \mathbf{K} [\mathbf{r}_1 \mathbf{r}_2 \mathbf{t}] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}. \quad (3)$$

由于模型平面位于世界坐标系 $Z=0$ 上, \mathbf{P} 仍表示图像中的一点, 因此 $\mathbf{P} = [X \ Y]^T$, 其齐次坐标 $\tilde{\mathbf{P}} = [X \ Y \ 1]^T$. 因此, 点 \mathbf{P} 和它的在图像上映射点 \mathbf{p} 之间的单应性矩阵 \mathbf{H} 满足:

$$s\tilde{\mathbf{p}} = \mathbf{H}\tilde{\mathbf{P}}, \quad (4)$$

式中: $\mathbf{H} = \mathbf{K} [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{t}] = [\mathbf{h}_1 \ \mathbf{h}_2 \ \mathbf{h}_3]$, 为 8 个自由度的单应性矩阵, 包含内参和外参。由于旋转矩阵的正交性, $\mathbf{r}_1^T \mathbf{r}_2 = 0$ 且 $|\mathbf{r}_1| = |\mathbf{r}_2| = 1$, 有

$$\mathbf{h}_1^T \mathbf{K}^{-T} \mathbf{K}^{-1} \mathbf{h}_2 = 0, \quad (5)$$

$$\mathbf{h}_1^T \mathbf{K}^{-T} \mathbf{K}^{-1} \mathbf{h}_1 = \mathbf{h}_2^T \mathbf{K}^{-T} \mathbf{K}^{-1} \mathbf{h}_2, \quad (6)$$

式中, \mathbf{K} 中的 5 个参数即为待求参数。

综上所述, 所检测到的棋盘格标定板图的角点精度直接影响相机内外参数的精确性。

3 基于超分辨率的角点检测算法

所提基于超分辨率的角点检测 mutual affine based SuperPoint (MASP) 算法利用超分辨率重建出具有更为丰富的角点特征的高分辨率图像, 同时在此基础上检测到其标定板图角点的亚像素位置。

每个待估计的高分辨率像素点根据网络估计出模糊核 k 和提取的图像特征而确定, 并在此基础上再对其通过 SuperPoint 架构提取新的特征, 输出整像素角点概率值的热图 (heatmap), 再由 `soft_argmax` 计算出其亚像素角点位置。将盲超分重建和角点检测过程集成为一个框架下, 实现了端到端的算法。为了更好提升算法的效率, 在图像超分辨部分采用区域相关性判断策略, 只有相关的区域参与模糊核的权重估计。算法框架结构如图 2 所示。所提 MASP 算法优点主要体现在: 1) 可以对真实场景下的未知模糊核进行估计; 2) 可适应多种复杂的标定板图检测, 如标定板图像有噪声和距离过远等; 3) 实现了端到端的模型。

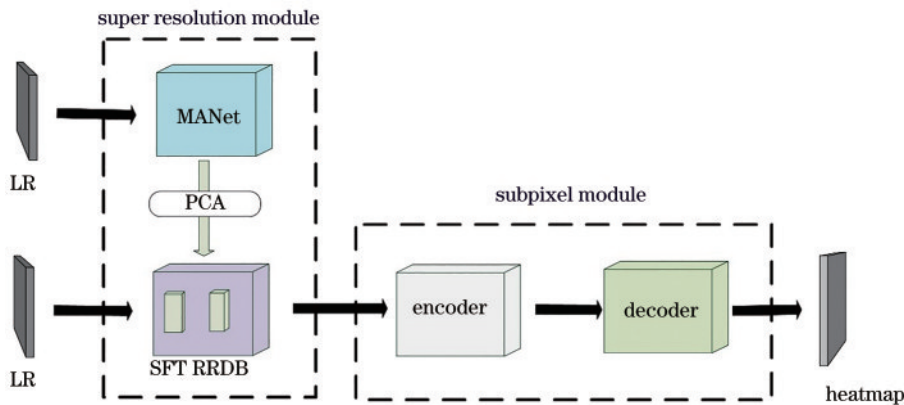


图 2 MASP 算法框架结构图
Fig. 2 MASP algorithm framework structure diagram

3.1 模糊核估计

在大多数盲超分辨率重建方法中, 模糊核 k 普遍被假设为服从高斯分布, 所提出的空间变异核估计也基于这一假设。基于空间变异核对原始高分辨率图像进行降质, 其退化模型可描述为

$$\mathbf{Y} = (\mathbf{K}\mathbf{X})_{\downarrow u} + \mathbf{n}, \quad (7)$$

式中: \mathbf{Y} 为低分辨率图像; \mathbf{X} 为高分辨率图像; \mathbf{K} 为类似

于卷积矩阵的模糊核矩阵, 且 \mathbf{K} 中的每一行对应于 \mathbf{X} 中每个像素的模糊核; $\downarrow u$ 是比例因子为 u 的下采样操作; \mathbf{n} 为各向异性高斯噪声。

MANet 包含特征提取 (feature extraction) 模块和核重构 (kernel reconstruct) 模块, 如图 3 所示。特征提取模块类似 U-Net 结构, 由卷积层、残差块、下采样器和上采样器组成。首先将低分辨率图像输入一个 3×3

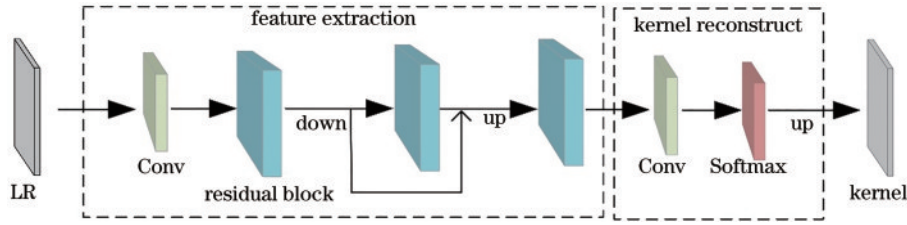


图 3 MANet 整体框架

Fig. 3 MANet overall framework

的卷积层中提取图像特征,然后经过 3 个残差块。其中,每个残差块包含两个互仿射卷积层(MAConv),它们之间有 ReLU 激活的卷积层以学习非线性。在中间残差块的前后,分别使用卷积层和转置卷积层对图像特征进行降采样和上采样。此外,在特征提取模块中添加了两个跳过连接,以利用不同层次的特征,提高表示能力。

特征提取后,核重构模块通过 3×3 卷积层和 Softmax 层来预测每个低分辨率图像像素的核。然后,使用最近邻插值得到高分辨率图像的最终估计核。核估计 $\mathbf{K} \in \mathbf{R}^{h \times w \times H \times W}$,其中, h, w, H 和 W 分别为核高度、核宽度、高分辨率图像的高度和宽度。

MANet 在低分辨率图像输入上有一个 22×22 的中等接收域,这保证了核估计不会受到超过 11 pixel 的其他图像区域块的干扰。同时,相互仿射卷积层不增加额外计算负担,但学习模糊核的表示能力明显提升。相互仿射卷积层框架如图 4 所示。

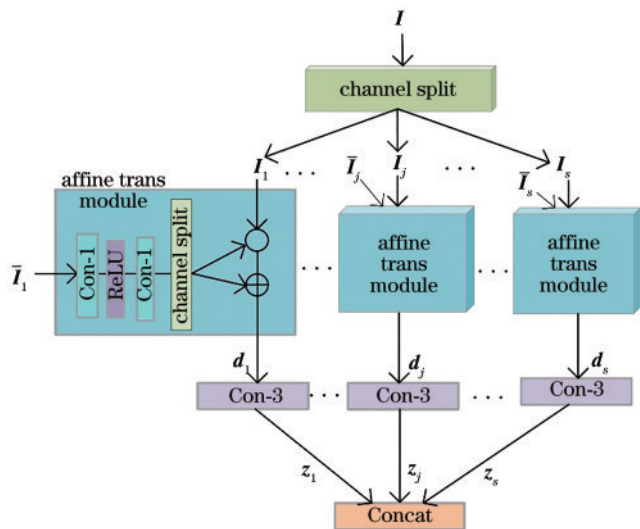


图 4 MAConv 框架

Fig. 4 MAConv framework

设 $\mathbf{I} \in \mathbf{R}^{C_n \times H_i \times W_i}$ 为 MAConv 层的输入特征,经过分割通道(split)将特征 \mathbf{I} 分割成 s 个特征 \mathbf{I}_i ,对于每个 $\mathbf{I}_i \in \mathbf{R}^{C_n \times H_i \times W_i}$,将与 \mathbf{I}_i 互补的分割部分连接表示为 $\widehat{\mathbf{I}}_i$ 。特征 \mathbf{I}_i 和 $\widehat{\mathbf{I}}_i$ 都被传递到仿射变换模块 (affine transformation module) 中,该模块有一个全连接网络 F ,从 $\widehat{\mathbf{I}}_i$ 中学习缩放变换参数 β_i 和位移变换参数 γ_i ,得

到 \mathbf{I}_i 的仿射变换后特征 \mathbf{d}_i :

$$\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_s = \text{split}(\mathbf{I}), \quad (8)$$

$$\beta_i, \gamma_i = \text{split}[F(\widehat{\mathbf{I}}_i)], \quad (9)$$

$$\mathbf{d}_i = \beta_i \otimes \mathbf{I}_i + \gamma_i, \quad (10)$$

式中: \otimes 为哈达玛积。不同分割通道得到的不同特征 \mathbf{d}_i 经过卷积层 Conv_i 后得到新的特征 \mathbf{z}_i ,连接(Concat)不同的特征 \mathbf{z}_i 得到 MAConv 层的输出特征 \mathbf{z} :

$$\mathbf{z}_i = \text{Conv}_i(\mathbf{d}_i), \quad (11)$$

$$\mathbf{z} = \text{Concat}(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_s). \quad (12)$$

3.2 基于空间变异核的高分辨率估计

高分辨率估计方法的性能直接对后续超分辨重建过程产生重要的影响。所提方法利用基于盲超分的核估计得到模糊核矩阵 \mathbf{K} ,通过主成分分析(PCA)将单位像素上模糊核维数从 hw 降为 q ,并将其拉伸为大小为 $q \times \frac{H}{s} \times \frac{W}{s}$ 的 PCA 映射核。将得到的映射核输入不同 spatial feature transform (SFT) + residual in residual dense block (RRDB) 模块的 SFT 层,将映射核与不同卷积层的低分辨率图像特征连接起来,以提高表示能力,有效重建高分辨率图像。其重建图像框架如图 5 所示。

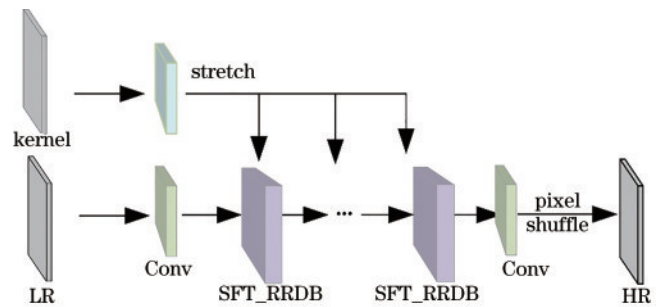


图 5 重建图像框架

Fig. 5 Reconstruct image frame

3.3 基于超分辨的角点检测

在基于超分辨率网络后基于 SuperPoint 架构获取角点。该网络由类似 VGG-Net 的编码器和解码器组成。特征图在编码器中降维并提取特征,解码部分经过 Softmax 层,计算得到图像的每个像素为角点的概率,最后采用子像素卷积的方法将其重组为与输入特征图像尺寸相同大小的热图,其上每个像素点所得到的概率值为该像素点成为角点的概率,如图 6 所示。

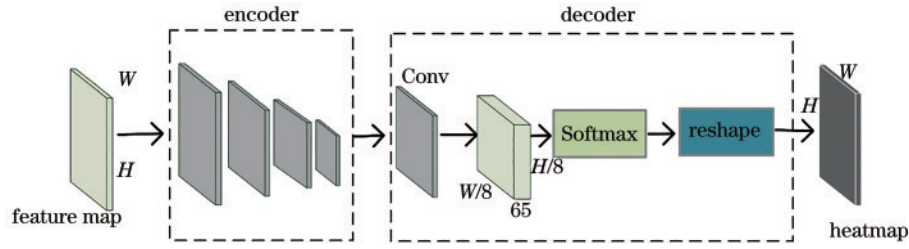


图 6 角点检测框架

Fig. 6 Corner detection frame

根据热图由 `soft_argmax` 算法计算得到角点的亚像素坐标 (x, y) , 其过程如下:

$$\Phi(h_{i,j}) = \frac{\exp(h_{i,j})}{\sum_{k=1}^W \sum_{l=1}^H \exp(h_{k,l})}, \quad (13)$$

式中: $h_{i,j}$ 为热图上像素坐标 (i, j) 的概率值。

$$\psi_x(h) = \sum_{i=1}^M \sum_{j=1}^H M_{i,j,x} \Phi(h_{i,j}), \quad (14)$$

$$\psi_y(h) = \sum_{i=1}^M \sum_{j=1}^H M_{i,j,y} \Phi(h_{i,j}), \quad (15)$$

式中: $M_{i,j,x} = \frac{i}{M}; M_{i,j,y} = \frac{j}{H}$ 。

$$\mathbf{P} = [\psi_y(h), \psi_x(h)]^T, \quad (16)$$

式中: \mathbf{P} 为热图上每个角点对应的亚像素坐标 (x, y) 。

该网络结构简单, 参数组成较少, 网络运算时间短, 可以使用其提取复杂环境下的角点。

4 实验结果与分析

真实场景下的图像由于存在不同噪声(如受传感器材料属性、拍摄环境光线等造成的噪声)、标定板图和相机距离较远时物体成像的像素占比率低等问题, 都会导致所获得的标定图像质量下降, 从而造成相机

的标定误差明显变大。首先, 采用不同角点检测算法对棋盘格虚拟图像(随机生成)进行对比实验; 其次, 在真实低分辨率棋盘格标定图集上用不同的标定方法对双目相机内外参进行标定^[26]; 最后, 采用前一步得到的不同标定参数对测距图像进行校正, 并在此基础上进行测距实验。

4.1 角点检测算法对比实验

对随机生成的 60 张低分辨率棋盘格虚拟图像进行角点检测, 一共包含 573 个角点, 低分辨率棋盘格虚拟图像是通过程序设置绘制生成棋盘格后再对其进行相应的仿射变换和单应性变换得到的, 故可得到其角点的亚像素位置真值。其中, 以棋盘格第一个格子的左上角为坐标系原点, 即 $(0, 0)$ 点, 后文均如此。对于已知的图像真实角点位置 (x, y) , 其角点为 (x', y') , 则检测误差 (E_{error}) 为 $\sqrt{(y - y')^2 + (x - x')^2}$ 。不同分辨率(scale factor)的检测结果如表 1 所示。其中: 检出误差所占比率 (R_{de}) 为不同检测误差区间(单位为 pixel)角点总数 n_i 与模型检测的角点总数 N_i 的比值; 检出率 (R_d) 为 N_i 与真值角点总数 N 的比值。不同分辨率亚像素角点检测的可视化结果如图 7 所示。

表 1 不同分辨率下的角点检测精度比较

Table 1 Comparison of accuracy ratio of corner detection under different resolution

Scale factor	R_{de}					R_d
	$E_{\text{error}}=0-0.2 \text{ pixel}$	$E_{\text{error}}=0.2-0.4 \text{ pixel}$	$E_{\text{error}}=0.4-0.6 \text{ pixel}$	$E_{\text{error}}=0.6-0.8 \text{ pixel}$	$E_{\text{error}}=0.8-0.9 \text{ pixel}$	
$\times 2$	0.4982	0.4175	0.0788	0.0053	0.0000	0.9738
$\times 3$	0.4937	0.4165	0.0789	0.0089	0.0017	0.9720
$\times 4$	0.4046	0.4334	0.1312	0.0251	0.0053	0.9703

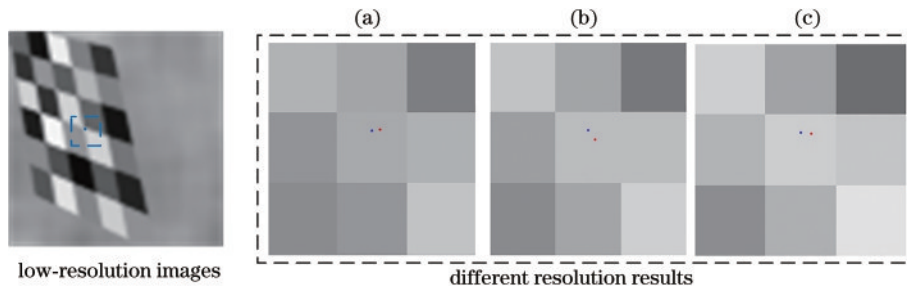


图 7 不同分辨率亚像素角点检测的可视化结果。(a) 2 倍分辨率 (b) 3 倍分辨率 (c) 4 倍分辨率

Fig. 7 Visualization of sub-pixel corner detection at different resolutions. (a) $2\times$ resolution; (b) $3\times$ resolution; (c) $4\times$ resolution

由表 1 可知,不论分辨率的大小为几倍,角点检测的误差都不会超过 0.9 pixel,且误差在 0.2 pixel 以下的检出误差所占比率可以达到 49.82%。

对于图 7 左边的低分辨率图像中的一个角点(蓝色线框内),右边对应于其由不同分辨率检测模型所得到的亚像素角点像素位置。其中,以角点整像素位置为中心取其 3×3 领域的像素块,蓝色点是角点真值像素位置,红色点为不同模型所预测得到的亚像素角点所在像素位置。可见,由不同分辨率角点检测所得到的角点与真值角点在一个像素以内,且与真值角点位

置都十分接近。

不同角点检测算法角点位置检测实验结果如表 2 所示。其中,平均检测误差(E_{md})为所有检测角点误差与模型检测的角点总数的比值。

由表 2 可以看出:MASP 算法的平均检测误差为 0.21 pixel,明显小于其他两个算法;相较于 Harris 算法的 88.13% 和 SuperPoint 算法的 92.67%,MASP 算法的检出率是最高的,达到 97.38%。并且,MASP 算法检测出的角点与真值角点相差 0~0.2 pixel 的检出率达到 49.82%,远高于 Harris 算法的 0.594%。

表 2 不同的角点检测算法比较

Table 2 Comparison of different corner detection algorithms

Algorithm	R_{de}						R_d	E_{md} / pixel
	$E_{error}=0-0.2 \text{ pixel}$	$E_{error}=0.2-0.4 \text{ pixel}$	$E_{error}=0.4-0.6 \text{ pixel}$	$E_{error}=0.6-0.8 \text{ pixel}$	$E_{error}=0.8-0.9 \text{ pixel}$	$E_{error}=0.9-2 \text{ pixel}$		
SuperPoint	0.0960	0.3352	0.3333	0.1920	0.0433	0.0000	0.9267	0.34
Harris	0.0059	0.0633	0.1920	0.3485	0.1485	0.2415	0.8813	0.64
MASP	0.4982	0.4175	0.0788	0.0053	0.0000	0.0000	0.9738	0.21

4.2 双目相机标定实验

采用双目相机拍摄不同视角下的棋盘格标定板

图,对其原始图像进行模糊和下采样处理,制造伪低分辨率图像,如图 8 所示。

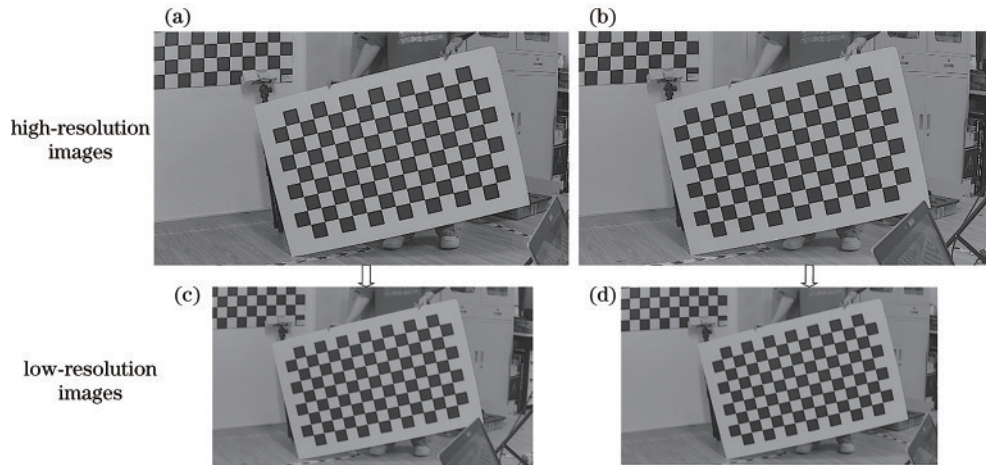


图 8 低分辨率棋盘格标定板左右相机图像。(a)(c)左图;(b)(d)右图

Fig. 8 Low-resolution checkerboard calibration board left and right camera images. (a)(c) Left images; (b)(d) right images

分别由 MASP 和 Harris 进行角点检测后,根据张正友相机标定计算得到左右相机内参以及其双目相机外参。由 Harris 标定得到相机内外参数如表 3 所示,由 2、3 及 4 倍分辨率的 MASP 标定得到的相机内外参数分别如表 4、表 5 及表 6 所示。其中, R_x 、 R_y 和 R_z 为右相机相对于左相机的旋转矩阵通过罗德里格斯公式转换得到的旋转向量参数, T_x 、 T_y 和 T_z 为右相机相对于左相机的平移向量参数, k_1 和 k_2 畸变系数参数, f_x 和 f_y 分别为 x 轴和 y 轴上的归一化焦距, (C_x, C_y) 为图像坐标系的原点参数。Harris 的重投影误差为 0.03 pixel, MASP 的重投影误差均为 0.02 pixel, MASP 算法的重投影误差明显小于 Harris 算法。

表 3 由 Harris 角点检测所得到的标定参数

Table 3 Calibration parameters obtained by Harris corner point detection

Binocular camera parameter	Left camera parameter	Right camera parameter
$R_x=0.0011$	$k_1=-0.51$	$k_1=-0.52$
$R_y=-0.0006$	$k_2=-0.68$	$k_2=-0.22$
$R_z=0.0149$	$f_x=967.70$	$f_x=964.79$
$T_x=-120.5738$	$f_y=965.78$	$f_y=962.77$
$T_y=-0.9858$	$C_x=318.27$	$C_x=317.93$
$T_z=-8.5164$	$C_y=178.92$	$C_y=180.30$

表 4 由 2 倍分辨率的 MASP 角点检测所得到的标定参数
Table 4 Calibration parameters obtained by 2× resolution MASP corner point detection

Binocular camera parameter	Left camera parameter	Right camera parameter
$R_x=0.0010$	$k_1=-0.55$	$k_1=-0.51$
$R_y=0.0002$	$k_2=0.42$	$k_2=-0.40$
$R_z=0.0149$	$f_x=1932.87$	$f_x=1928.53$
$T_x=-120.4570$	$f_y=1929.41$	$f_y=1925.31$
$T_y=-1.1657$	$C_x=637.05$	$C_x=637.41$
$T_z=-4.4702$	$C_y=356.87$	$C_y=359.43$

表 5 由 3 倍分辨率的 MASP 角点检测所得到的标定参数
Table 5 Calibration parameters obtained by 3× resolution MASP corner point detection

Binocular camera parameter	Left camera parameter	Right camera parameter
$R_x=0.0012$	$k_1=-0.56$	$k_1=-0.52$
$R_y=0.0003$	$k_2=0.43$	$k_2=-0.21$
$R_z=0.01486$	$f_x=1934.35$	$f_x=1932.74$
$T_x=-120.6524$	$f_y=1932.60$	$f_y=1925.31$
$T_y=-1.0536$	$C_x=634.11$	$C_x=631.59$
$T_z=-4.3653$	$C_y=358.47$	$C_y=361.42$

从相机标定结果可以看出,相比于Harris角点检测标定方法,由不同分辨率MASP角点检测所得到的相机参数都十分接近,且重投影误差都小于Harris算法,提高了相机标定的精度。

4.3 标定板图测距实验

在真实场景图像实验中,在一个可移动的平台搭载双目摄像机,对同一张黑白棋盘格标定板图进行

表 6 由 4 倍分辨率的 MASP 角点检测所得到的标定参数
Table 6 Calibration parameters obtained by 4× resolution MASP corner point detection

Binocular camera parameter	Left camera parameter	Right camera parameter
$R_x=0.0014$	$k_1=-0.54$	$k_1=-0.51$
$R_y=0.0004$	$k_2=0.00$	$k_2=-0.53$
$R_z=0.01485$	$f_x=1932.99$	$f_x=1922.54$
$T_x=-120.4545$	$f_y=1930.37$	$f_y=1919.69$
$T_y=-1.3821$	$C_x=636.61$	$C_x=633.80$
$T_z=-12.1748$	$C_y=358.22$	$C_y=361.84$

不同距离的拍摄。为了便于操作,采用相对运动的实验方法,即双目相机固定不动,然后由棋盘格标定板图相对其每间隔 1 m 进行拍照,从 3~14 m 共进行 12 组实验。其中,图 9 为 3~14 m 距离下每间隔 1 m 的棋盘格标定板图像。针对真实场景下获取的低分辨率图像对标定及测距的影响问题,对于不同距离下的左右图像,同样对其进行模糊和下采样(2 倍)处理得到左右伪低分辨率图像,由以上得到的两组内外参分别对其进行矫正后,再分别使用 Harris 和 MASP 算法检测出左右图对应的角点并计算出视差值 v_{disp} 。由 $Z = \frac{bf}{v_{disp}}$ ($b = 120 \text{ mm}$) 计算相机标定板图的距离,并与真实测距距离进行对比,误差率 $R_{error} = \frac{|Z - Z_{real}|}{Z_{real}}$ 。实验结果如表 7 所示,其中,Mc 为标定角点检测方法,Mr 为测距角点检测方法,Mc_M 表示用 2 倍分辨率的 MASP 算法标定角点检测,Mc_H 表示用 Harris 算法标定角点检测,Mr_M 表示用 2 倍分辨率的 MASP 法测距角点检测,

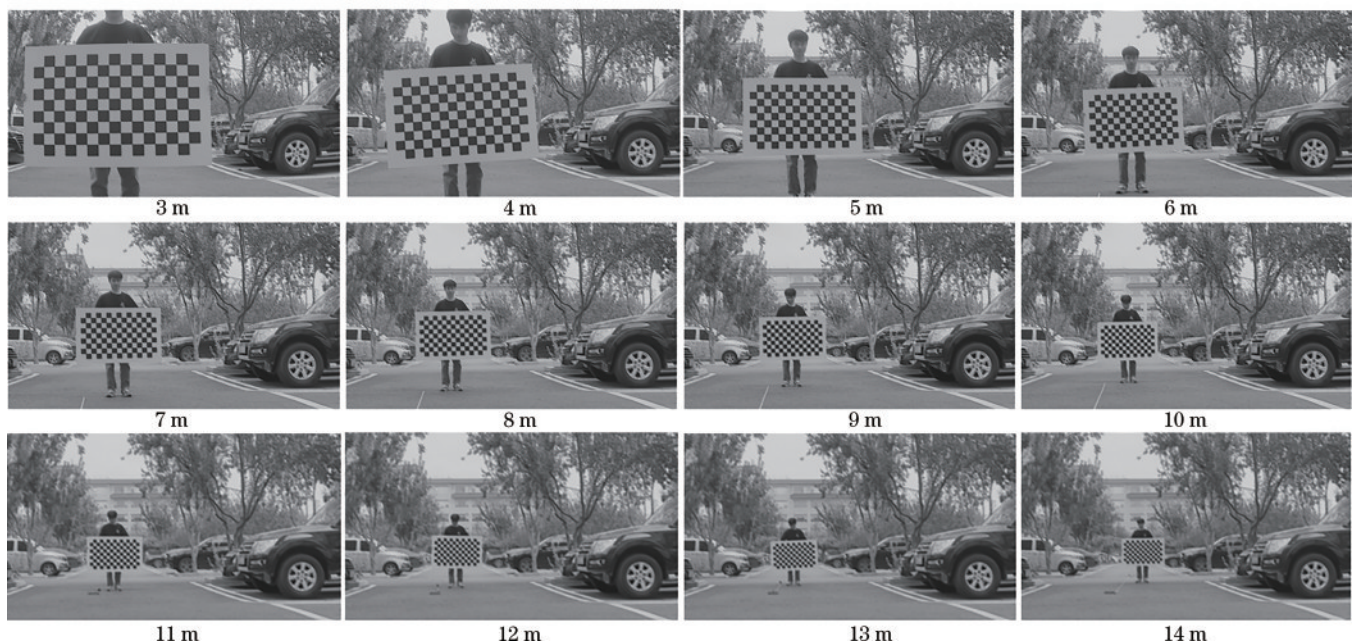


图 9 3~14 m 距离下的标定板图
Fig. 9 Calibration plate diagram at a distance of 3-14 m

表 7 不同内外参数对图像进行校正后的测距误差率

Table 7 Error rate of ranging after correcting the image with different internal and external parameters

Method	R_{error}			
	Mc_MMr_H	Mc_MMr_M	Mc_HMr_H	Mc_HMr_M
D3	0.0076	0.0073	0.0113	0.0112
D4	0.0083	0.0074	0.0100	0.0100
D5	0.0114	0.0099	0.0149	0.0154
D6	0.0190	0.0139	0.0217	0.0211
D7	0.0169	0.0149	0.0248	0.0244
D8	0.0266	0.0153	0.0377	0.0341
D9	0.0273	0.0297	0.0339	0.0312
D10	0.0341	0.0199	0.0390	0.0360
D11	0.0359	0.0146	0.0531	0.0557
D12	0.0384	0.0293	0.0544	0.0396
D13	0.0583	0.0299	0.0384	0.0367
D14	0.1250	0.0399		0.0413
Average error rate	0.0340	0.0193	0.0308	0.0297

Mr_H 表示用 Harris 算法测距角点检测, Mc_MMr_H 即为用 MASP 算法标定角点检测后对测距图像用其参数进行校正, 再用 Harris 算法进行测距实验的角点检测。本实验为保证 b, f 值不受影响, 在这组实验中, 采用最近邻插值法将伪低分辨率图像插值为原来的 2 倍再进行检测。D 表示真实距离, D3 即为真实距离为 3 m。

表 7 数据表明, 用 MASP 角点检测算法标定得到的相机参数对测距图像进行校正后再分别用 MASP 算法和 Harris 算法进行角点检测时, MASP 算法的误差率明显小于 Harris 算法, 且其相对测量误差为 1.93% 也小于 Harris 算法的 3.4%。

用 Harris 角点检测算法标定得到的相机参数对测距图像进行校正后再分别用 MASP 算法和 Harris 算法进行角点检测, 当距离为 14 m 时, Harris 算法检测不到标定板图像上的角点, MASP 算法的误差率也小于 Harris 算法, 且其相对测量误差为 2.97% 也小于 Harris 算法的 3.08%。

同时可看出, 在 3~12 m 时, 用 Harris 算法标定得到的相机参数进行图像校正后两个算法的误差率也同时变大了, 说明 MASP 算法的双目相机标定结果是准确的, 在此基础上再进行测距的效果也更准确。

5 结 论

围绕双目相机标定中的低分辨率棋盘格标定板图像检测问题开展了相关研究, 提出基于 MANet 和 SuperPoint 相结合的超分辨率角点检测算法的双目相机标定方法, 提高了低分辨率标定板图像的检测能力和标定算法的精度。一方面, 通过设计端到端的角点检测算法, 实现了低分辨率棋盘格标定板上亚像素角点的提取。虚拟棋盘格图像检测实验结果显示, 所

提方法检测能力大于 97%。同时, 检测出的角点误差仅为 0.21 pixel, 检测精度提升显著。另一方面, 基于盲超分特性, 在重建出更接近真实世界的超分辨率图像基础上, 角点检测更快速精准。由此可见, 所提方法能够获得精度更高且稳定性更好的双目相机标定结果。

参 考 文 献

- [1] Long Q, Xie Q W, Mita S, et al. Real-time dense disparity estimation based on multi-path Viterbi for intelligent vehicle applications[C]//Proceedings of the British Machine Vision Conference, September 1-5, 2014, Nottingham. London: British Machine Vision Association, 2014.
- [2] Long Q, Xie Q W, Mita S, et al. A real-time dense stereo matching method for critical environment sensing in autonomous driving[C]//17th International IEEE Conference on Intelligent Transportation Systems (ITSC), October 8-11, 2014, Qingdao, China. New York: IEEE Press, 2014: 853-860.
- [3] Luo H G, Zhu L M, Ding H. Camera calibration with coplanar calibration board near parallel to the imaging plane[J]. Sensors and Actuators A: Physical, 2006, 132(2): 480-486.
- [4] Ma C, Pei S S, Sun G L, et al. Disparity estimation based on fusion of vision and LiDAR[J]. International Journal of Wavelets, Multiresolution and Information Processing, 2022, 20(5): 2250014.
- [5] Xie Q W, Hu X Y, Ren L, et al. A binocular vision application in IoT: realtime trustworthy road condition detection system in passable area[J]. IEEE Transactions on Industrial Informatics, 2023, 19(1): 973-983.
- [6] 辛元雪, 朱凤婷, 史朋飞, 等. 基于改进增强型超分辨率生成对抗网络的图像超分辨率重建算法[J]. 激光与光电子学进展, 2022, 59(4): 0420002.
- [7] Xin Y X, Zhu F T, Shi P F, et al. Super-resolution reconstruction algorithm of images based on improved enhanced super-resolution generative adversarial network [J]. Laser & Optoelectronics Progress, 2022, 59(4): 0420002.
- [7] 梁美玉, 杜军平, 刘红刚, 等. 基于时空非局部相似性的视频超分辨率重建[J]. 系统科学与数学, 2016, 36(9): 1397-1409.
- [8] Liang M Y, Du J P, Liu H G, et al. Video super-resolution reconstruction based on spatio-temporal non-local similarity[J]. Journal of Systems Science and Mathematical Sciences, 2016, 36(9): 1397-1409.
- [8] 罗爽, 黄辉, 张凯兵. 基于 Boosting 回归的单帧图像超分辨率重建[J]. 激光与光电子学进展, 2022, 59(8): 0810018.
- [9] Luo S, Huang H, Zhang K B. Boosting regression-based single-image super-resolution reconstruction[J]. Laser & Optoelectronics Progress, 2022, 59(8): 0810018.
- [9] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York:

- IEEE Press, 2016: 770-778.
- [10] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1646-1654.
- [11] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[EB/OL]. (2015-11-19) [2022-09-08]. <https://arxiv.org/abs/1511.06434>.
- [12] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04)[2022-09-08]. <https://arxiv.org/abs/1409.1556>.
- [13] Dong C, Loy C C, He K M, et al. Learning a deep convolutional network for image super-resolution[M]//Fleet D, Pajdla T, Schiele B, et al. Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8692: 184-199.
- [14] Shi W Z, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1874-1883.
- [15] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 105-114.
- [16] Bell-Kligler S, Shocher A, Irani M. Blind super-resolution kernel estimation using an internal-GAN[EB/OL]. (2019-09-14) [2022-09-08]. <https://arxiv.org/abs/1909.06581>.
- [17] Liang J Y, Sun G L, Zhang K, et al. Mutual affine network for spatially variant kernel estimation in blind image super-resolution[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), October 10-17, 2021, Montreal, QC, Canada. New York: IEEE Press, 2022: 4076-4085.
- [18] Xie Q W, Liu R R, Sun Z, et al. A flexible free-space detection system based on stereo vision[J]. Neurocomputing, 2022, 485: 252-262.
- [19] 卢威威, 刘安东, 仇翔, 等. 基于视觉的机器人滚动时域位姿估计[J]. 系统科学与数学, 2021, 41(7): 1772-1787. Lu W W, Liu A D, Qiu X, et al. A moving-horizon-based method for robot pose estimation of vision[J]. Journal of Systems Science and Mathematical Sciences, 2021, 41(7): 1772-1787.
- [20] 魏祥灰, 唐超颖, 王彪, 等. 视觉着陆三维合作目标结构设计及定位算法[J]. 系统工程理论与实践, 2019, 39(11): 2975-2983. Wei X H, Tang C Y, Wang B, et al. Three-dimensional cooperative target structure design and location algorithm for vision landing[J]. Systems Engineering-Theory & Practice, 2019, 39(11): 2975-2983.
- [21] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [22] Tian Y R, Fan B, Wu F C. L2-net: deep learning of discriminative patch descriptor in euclidean space[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6128-6136.
- [23] DeTone D, Malisiewicz T, Rabinovich A. SuperPoint: self-supervised interest point detection and description [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 18-22, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 337-349.
- [24] Zhang Z. A flexible new technique for camera calibration [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [25] Xie Q W, Long Q, Mita S. Integration of optical flow and Multi-Path-Viterbi algorithm for stereo vision[J]. International Journal of Wavelets, Multiresolution and Information Processing, 2017, 15(3): 1750022.
- [26] 刘泽庆, 张玉荣, 赵建新, 等. 靶场静爆实验测速高速相机标定方法[J]. 激光与光电子学进展, 2016, 53(11): 111501. Liu Z Q, Zhang Y R, Zhao J X, et al. High speed camera calibration for velocity measurement in range static explosion experiment[J]. Laser & Optoelectronics Progress, 2016, 53(11): 111501.