

一种改进 Census 变换的双目匹配测距方法

李大华^{1,2}, 沈洪宇^{1,2*}, 于晓^{1,2}, 高强^{1,2}, 汪宏威^{1,2}

¹天津理工大学电气电子工程学院, 天津 300384;

²天津市复杂系统控制理论及应用重点实验室, 天津 300384

摘要 为了得到一种易实现且精度较高的双目测距方法, 立体校正左右相机的非前向平行结构, 先将改进的 Census 变换算法应用于立体匹配, 得到准确的视差值, 再根据双目视觉特殊的外极线几何结构计算出实际的距离信息。将原始 Census 变换中比较周围像素与中心像素的方案改进为多中心点相互监督融合, 极大地提升了立体匹配精度。在实验室环境下, 利用两个完全相同的互补金属氧化物半导体(CMOS)相机搭建了双目测距实验平台, 详细介绍了测距流程中的硬件、算法以及标定过程。将实验结果应用于实际距离测量, 并与原始的 Census 变换进行对比, 结果表明: 改进的 Census 变换测量误差为 6.4 cm, 2 m 测量误差精度提高了 19.1%, 满足高精度双目测距要求。

关键词 机器视觉; 双目测距; 相机标定; Census 变换; 立体校正; 立体匹配

中图分类号 TP391

文献标识码 A

doi: 10.3788/LOP56.111503

Binocular Ranging Method Using Stereo Matching Based on Improved Census Transform

Li Dahua^{1,2}, Shen Hongyu^{1,2*}, Yu Xiao^{1,2}, Gao Qiang^{1,2}, Wang Hongwei^{1,2}

¹ School of Electrical and Electronic Engineering, Tianjin University of Technology, Tianjin 300384, China;

² Tianjin Key Laboratory for Control Theory & Applications in Complicated Systems, Tianjin 300384, China

Abstract This study proposes a simple and high-precision binocular ranging method. A stereo correction algorithm was first used for the stereo correction of the non-forward parallel structures of the left and right cameras. The improved Census transform algorithm was then applied to obtain accurate disparity values. Finally, the true distance information was calculated based on the special epipolar-line geometry of binocular vision. Further, the multi-center points were used to compare with the surrounding pixels in the original Census transform, which is improved as mutual supervision and fusing of multi-center points. Thus the accuracy of the stereo matching is improved. Two identical complementary metal-oxide-semiconductor (CMOS) cameras were used to build a binocular ranging platform, and the hardware, algorithm and calibration process in the flow chart of ranging were introduced in detail. The experimental results show that the proposed method performs better than the original Census transform. For a 2 m measurement, the accuracy is increased by 19.1% and the measurement error is 6.4 cm, thus meeting the requirements of high-precision binocular ranging.

Key words machine vision; binocular ranging; camera calibration; Census transform; stereo correction; stereo matching

OCIS codes 150.0155; 150.1488; 150.1135

1 引言

随着计算机视觉的发展, 双目立体视觉测量技术研究具有重要的实际意义。利用两个简单的互补

金属氧化物半导体(CMOS)或电荷耦合器件(CCD)相机实现被拍摄物体距离的测量是今后测距的主要方式^[1]。立体视觉测量技术是仿照人类的双眼视觉成像过程。人的两只眼睛从不同的角度观察同一个

收稿日期: 2018-11-22; 修回日期: 2018-12-23; 录用日期: 2019-01-07

基金项目: 国家自然科学基金(61502340)、天津市中青年骨干创新人才培养计划基金(20130830)

* E-mail: 1161508014@qq.com

物体,物体成像在人左右两眼视网膜上的不同位置,视网膜上的距离差称为视差值^[2-3]。利用视差值以及三角测量原理,可以还原出物体的深度信息^[2]。双目立体测距广泛应用于刑侦检测、交通定位测距、机器视觉定位抓取、军事探测以及无人机航拍等研究领域^[4]。因此,针对双目视觉测距精度提升的研究有着深远长久的意义。

2017年, Lee等^[5]提出一种新型的基于代价融合的立体匹配方法的双目测距方法,并给出了详细的推导。目前,大部分学者把提高测距精度的重点放在立体匹配过程中。张丽娟等^[6]提出一种基于模板的灰度相关算法,提出亚像素级别的匹配算法,并利用三次曲面拟合寻点进行立体配准,但是在匹配过程中引入图像压缩的相关方法,极大地降低了立体匹配的准确性。为了保留图像的原始信息,推荐采用 Census 变换算法,此方法的优点是不改变图像的原始特征,操作简单,通过汉明距离确定立体匹配关系,满足工业过程的测距需求^[7]。针对 Census 变换的研究,王琼等^[8]利用变换窗口灰度均值与局部纹理反差值之和代替中心像素灰度值,有效减小了边界不连续区域对匹配的影响,但此方法实现困难,除边界外的点,其他点出现误匹配的概率很高。祝世平等^[9]提出一种区域为十字交叉形状窗口且设有噪声容限的改进 Census 变换,该方法可通过代价计算得到匹配像素,以至得到相对准确的视差值,但是实现过程极其困难,且算法运行时间较长。针对目前 Census 匹配的研究过程,本文提出一种基于 Census 变换的多中心点融合监督的方式,实现过程简单且精度高,适用于一般环境下的测距项目。

2 基本原理

2.1 双目测距原理介绍

双目测距基于视差原理,形同于人类两只眼观察同一物体所呈现的图像的差异,即成像在视网膜上发生了位置平移,根据此原理及数学三角法可以直接测量出两只眼睛与被测物体的距离信息^[10]。双目测距流程中,先将两个 CMOS 相机前向平行摆放,通过双目相机标定工作将两个相机统一在一个坐标系中,接着对被测物体在左右相机中的成像点进行立体匹配,获取被测物体在两个 CMOS 中成像的横坐标的差值,最终利用相似三角形原理测得真实的距离信息。

1) 双目标定工作。双目标定建立在单相机标定的基础上,单相机标定出两个相机各自的内部参数和外部参数,双目标定基于两个相机的内外参数

计算出两个相机间相对位姿关系,并基于这一关系校正两个相机的前向平行结构,使得被拍摄物体在两个相机的成像行对齐。2) 立体匹配工作。立体匹配时 Census 变换的立体匹配算法的误匹配率很高,运用改进的 Census 变换算法,采用多中心点融合监督的方式能大幅提升匹配的准确性,最终结合三角法即可求得距离信息。

2.2 双目标定

此步骤为求解两个相机之间的相对位置关系,即外参,并对左右相机拍摄的图片进行行校正,外极线几何结构如图 1 所示,为对极几何结构,图中 e_l 和 e_r 为两个极点, e_l-e_r 为左极线, e_r-p_r 为右极线, e_l-p_l 所在平面为透视平面,对于此非标准结构应用两个单相机标定内外参数结果得到其校正向量^[11]。两个相机之间的位姿关系可根据两个单相机标定结果进行矩阵运算得到,即根据 $X_l=R_1X_w+t_1$, $X_r=R_2X_w+t_2$, 可得

$$X_l=R_1R_2^{-1}X_r+t_1-R_1R_2^{-1}t_2, \quad (1)$$

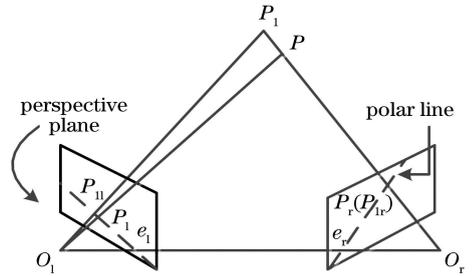


图 1 外极线几何结构

Fig. 1 Geometric structure of epipolar-line

式中: X_l 与 X_r 为左右相机成像平面点坐标; X_w 为被测物体点坐标; R_1 与 t_1 为左相机外参和内参矩阵; R_2 与 t_2 为右相机的外参和内参矩阵。设 $R=R_1R_2^{-1}$, $t=t_1-R_1R_2^{-1}t_2$, 其中 R 和 t 为两相机之间的外参矩阵, 外参矩阵包含了两台相机之间所有的几何信息。双目标定得到右相机相对左相机的旋转和平移向量后, 校正两极线, 创建一个由 α_1 方向开始的旋转矩阵 $R_{rect}=[\alpha_1 \quad \alpha_2 \quad \alpha_3]$, 其中 $\alpha_2=\alpha_1 \times [0 \quad 0 \quad 1]^T$, α_3 为 α_1 与 α_2 的外积, 生成的映射图将两图中行与行之间配准, 提高了后续立体匹配的精度和速度。图 2 为平行校正后的外极线几何结构。其中 l 为距离信息(相机镜头距离物体的深度信息), X_w 为被测物体点坐标, X_l 与 X_r 为左右相机成像平面点, $D=X_l-X_r$ (视差), f 为焦距, O_l 与 O_r 为相机光心。立体匹配后, 将左右相机对应像素点的横坐标相减便可以得到视差, 由相似三角形原理即可得到距离信息

$$l = \frac{bf}{D-b}, \quad (2)$$

式中： b 为相机光心间的距离。

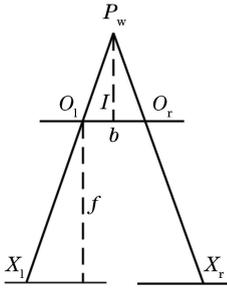


图 2 平行校正的外极线几何结构

Fig. 2 Geometric structure of epipolar-lines for parallel correction

2.3 单相机标定原理

双目标定中关键的一步为单相机标定,求解单相机的内外参数, O_c 为相机光心点,成像平面为图中虚线平面。假设圆形标志区域的左上角的点为世界坐标系原点,虚线平面为相机的成像平面,将其前移构建图像坐标系 $O_{xy}-xy$, 像素坐标系 $O_{uv}-uv$, 世界坐标系 $O_w-X_w Y_w Z_w$, 相机坐标系 $O_c-X_c Y_c Z_c$, 标定过程即可将像素坐标系中点 (u, v) 与其在世界坐标系中对应物体点的坐标 (X_w, Y_w, Z_w) 的对应关系求出^[12-13]。

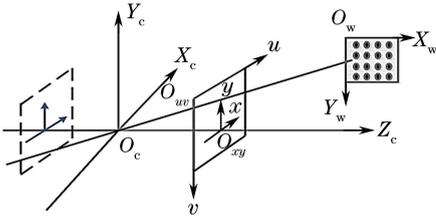


图 3 相机标定原理

Fig. 3 Camera calibration principle

假设图像中一点坐标的矩阵形式为 $m = [u \ v \ 1]$, 对应于世界坐标系中的点坐标为 $M = [X \ Y \ Z \ 1]$, 外参矩阵为 $Q = [r_1 \ r_2 \ r_3 \ t]$, 其中, $r_1 \sim r_3$ 分别为绕 XYZ 轴的旋转向量, t 为平移向量, $t = [T_x \ T_y \ T_z]$, $T_x \sim T_z$ 分别为沿 XYZ 轴的平移向量, 内参矩阵可表示为

$$A = \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (3)$$

式中： f 为焦距； d_x 和 d_y 为单个像素所代表的实际长度尺寸和宽度尺寸； (u_0, v_0) 为相机的光心坐标^[14]。图像中点对应于世界坐标系点的映射关系式为

$$A * Q * M^T = m^T. \quad (4)$$

2.4 改进 Census 变换的立体匹配

原始的 Census 变换算法的步骤为:输入 2 幅图像,1 幅左图,1 幅右图,定义 1 个小窗口为 3×3 或者 5×5 的矩阵,类似掩模,为了方便,后文称其为掩模。掩模是用来遍历左右图像中行上的点,寻找匹配像素。立体校正后可得到行对齐的左右图像。由于校正存在一到两个像素的误差,因此即使校正比较准确,也要在立体匹配时将对应行及其周围行像素扩展一到两行,扩大其匹配范围,提高匹配的准确度。定义一个掩模为

$$\begin{bmatrix} 2 & 4 & 6 \\ 1 & * & 5 \\ 4 & 7 & 8 \end{bmatrix}. \quad (5)$$

掩模中间像素值为空,将掩模遍历左右图像中的对应行,比较掩模周围的灰度值与中间点的灰度值,比中间点小的标记为 1,比中间点大的标记为 0,与其相等的像素也记为 0。假设被遍历的像素值为 3,则矩阵可展开为一条码:10010000,比较左右图像对应点码的相似程度,相似度高则匹配度高。但这种算法存在一定缺点,即存在匹配误差,误差举例如下:

$$\begin{bmatrix} 2 & 4 & 6 \\ 1 & * & 5 \\ 4 & 7 & 8 \end{bmatrix}, \quad (6)$$

$$\begin{bmatrix} 4 & 8 & 12 \\ 2 & * & 10 \\ 8 & 14 & 16 \end{bmatrix}, \quad (7)$$

$$\begin{bmatrix} 1 & 5 & 7 \\ 2 & * & 6 \\ 5 & 8 & 7 \end{bmatrix}. \quad (8)$$

当(5)式经过的中间像素点值为 3,(6)式经过的中间像素点值为 6 时,可以得出(5)式和(6)式的矩阵最终生成的码相同,虽然两式中的中间点的灰度值呈两倍关系,但最终仍会把这两个像素点误匹配为一个点。(7)式所示的矩阵中,当中间像素点值为 4 时,(5)式和(7)式也能匹配成一个像素点,其中 * 代表中间孔像素值。针对此缺点,采用扫描窗口中多中心点融合监督方式进行了改进,改进掩模矩阵示意为

$$\begin{bmatrix} b_1 & b_2 & b_3 & b_4 & b_5 \\ b_6 & b_7 & b_8 & b_9 & b_{10} \\ b_{11} & b_{12} & * & b_{13} & b_{14} \\ b_{15} & b_{16} & b_{17} & b_{18} & b_{19} \\ b_{20} & b_{21} & b_{22} & b_{23} & b_{24} \end{bmatrix}, \quad (9)$$

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ a_6 & * & a_7 & * & a_8 \\ a_9 & a_{10} & a_{11} & a_{12} & a_{13} \\ a_{14} & * & a_{15} & * & a_{16} \\ a_{17} & a_{18} & a_{19} & a_{20} & a_{21} \end{bmatrix} \quad (10)$$

舍弃 3×3 的掩模, 将掩模大小扩充至 5×5 甚至更大。以(8)式为例, 将一个中心点扩充为 4 个新中心点, 得到矩阵如(9)式所示。4 个新的中心点与原始中心点为 8 邻域邻接, 在 4 个中心点下, $a_1 \sim a_{21}$ 为随机的数, 代表像素值, 在改进的算法中, 将 4 个中心点分别在其 8 邻域内比较周围点的汉明距离

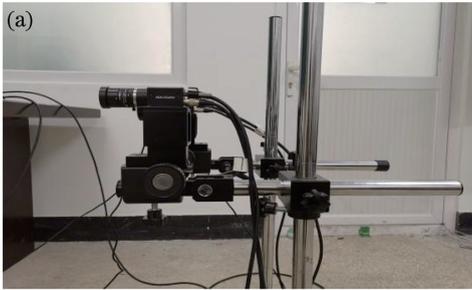


图 4 硬件系统示意图。(a)侧视图;(b)俯视图

Fig. 4 Schematic of hardware system. (a) Side view; (b) top view

3.2 单相机标定

左相机内参矩阵为 $[0.0806856, -5.67737, 3.45139 \times 10^{-5}, 3.45 \times 10^{-5}, 1253.07, 958.157, 2448, 2048]$, 右相机内参矩阵为 $[0.0809745, -6.29119, 3.45483 \times 10^{-5}, 3.45 \times 10^{-5}, 1278.83, 972.503, 2448, 2048]$ 。实验中, 用标定板将相机与世界坐标系建立联系, 设计的标定板是一组 7×7 的由黑色标准圆组成的可提取圆中心点坐标的二维靶标。左相机拍摄的第 1 幅标定板的外参矩阵为

$[0.114368, -0.0898491, 0.468311, 359.189, 1.37263, 91.8985]$, 右相机拍摄的第一幅标定板的外参矩阵为 $[-0.16579, -0.0836154, 0.466578, 358.078, 0.596036, 91.3967]$ 。

3.3 双目标定

双目标定结果为 $[0.272985, 0.00425349, 0.00123705, 1.11599, 0.793002, 0.475306]$, 运行行校正矩阵结果如图 5 所示, 左右相机拍摄的图片被校正为行对齐, 立体校正结果如图 5 所示。

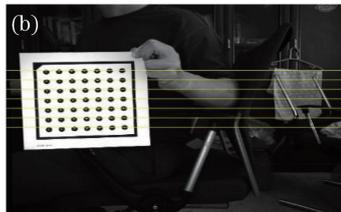
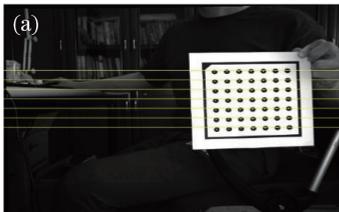


图 5 立体校正结果

Fig. 5 Stereo correction results

3.4 改进 Census 变换的立体匹配

改进 Census 变换的立体匹配实验步骤如下: 1) 输入左右相机原图; 2) 进行图像预处理, 针对复杂的背景图像以及光照因素干扰, 分割提取出感兴趣区域(ROI); 3) 将 ROI 图形作为 Census 立

体匹配程序的输入图像进行计算, 得到匹配结果同时生成视差值; 4) 视差值编码成灰度图像, 反映距离远近。行校正后立体匹配的视差图(针对被测物茶叶罐)如图 6 所示, 距离越近则对应颜色越深。

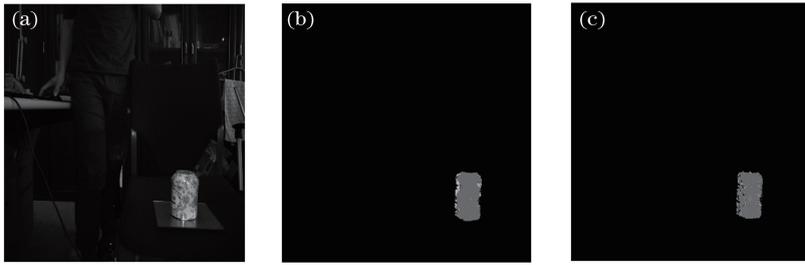


图6 行校正后立体匹配的视差图。(a)左相机原图;(b) Census 变换得到的视差图;(c)改进的 Census 变换得到的视差图
Fig. 6 Disparity map for stereo matching after row correction. (a) Original image of left camera; (b) disparity map obtained by Census transform; (c) disparity map obtained by improved Census transform

图6(a)为左相机原图,需要测量2个相机与茶罐的深度距离信息。图6(b)为运用原始的Census变换得到的视差图像,由图可以看出,茶罐表面出现了许多误匹配点(白色的噪声点)。图6(c)为运用改进的Census变换得到的视差图像,由图可以看出,改进后的算法消除了误匹配,且在茶罐深度图像边缘部分处得到了由于深度距离不同而发生颜色不同的变化特征。

4 分析与讨论

原始的Census变换得到的视差图中,茶叶罐表面出现许多白色的斑点痕迹,这是由误匹配而产生的错误结果。针对这一问题,使用改进的Census变换重新运行程序,得到的视差图效果明显优于原始Census变换的结果,且可以读取视差图中的视差值以进行实际距离的测量。经反复实验,实际距离为2m时的测量误差仅为6.4cm。应用于测距时,

表1 原始Census变换与改进Census变换的误差比较

Table 1 Comparison of errors between original Census transform and improved Census transform

Actual distance /mm	Measured distance by improved Census /mm	Error for improved Census /%	Census measuring distance error /%	Accuracy improvement /%
400	400.84	0.21	2.26	2.05
500	502.65	0.53	3.88	3.35
600	604.98	0.83	4.95	4.12
700	707.14	1.02	6.96	5.94
800	809.52	1.19	8.31	7.12
900	914.04	1.56	10.78	9.22
1000	1019.7	1.97	13.31	11.34
2000	2064.2	3.21	22.31	19.10

5 结论

介绍了一套易实现且精度较高的双目测距系统,给出了其外观构造图,使用2个完全相同的500万像素的黑白相机,定制了带微调结构的铝合金落地支架,固定两相机。通过单相机标定获得其内外

原始Census变换与改进的Census变换的误差比较如表1所示。

改进的Census算法将原始 3×3 的掩模扩展为 5×5 ,采用4个中心点融合监督,极大地降低了误匹配率。由于一些实际项目中需要实时定位测量标记物体的距离,增加掩模大小的确能使结果更加准确,但程序运行时间也会随之增加,无法满足实时性需求。因此,这里在求解视差值的同时,利用相机的特殊结构构建三角形进行求解,如图4所示,当物体正好在两个相机正中心时,校正平行的外极线几何结构为等腰三角形。当被测物体在其他位置时也可以利用三角形的比例关系进行计算,相机成像平面与光心构成平行结构。由于实验室场地的局限性,这里仅对于2m左右的距离进行了百次测量,得到了准确的结论。当被测物距离增加时,左右相机之间的距离应随之提升,后续工作中重新进行了长距离的测距实验并验证了此测距流程的可行性。

参数,利用单相机标定结果进行双目标定获得两相机的相对位置关系。计算出校正向量并对左右相机拍摄的图片进行行对准,并在此基础上运行改进的Census变换的立体匹配算法,得到较为准确的视差值,给出了视差值到实际距离的求解公式。实验中,以茶罐为被测物,运用改进的Census变换得到的视

差图中消除了茶罐表面的误匹配点。为了验证此高精度双目测距系统的准确性,设置多组被测距离进行对比,实际被测距离为 2 m,改进算法的测距误差为 6.4 cm。设置对照组与原始的 Census 算法运行结果进行对比,验证了改进的 Census 算法的优越性,2 m 测量精度提高了 19.1%。所提测距方法得到准确的测量结果有两个前提:在单相机标定过程中需要将标定误差控制在 0.1 pixel 以内;双目标定误差需要控制在 0.2 pixel 以内。Census 变换基于比大小与异或运算,因此所提测距方法的立体匹配速度快,优于其他算法,有望广泛应用于刑侦监控中实时定位跟踪嫌疑人,行驶车辆测速,生产流水线上实时记录多自由度机械臂位姿等实时性测量的项目中。

参 考 文 献

- [1] Zhang E S, Wang S B. Plane-space algorithm based on binocular stereo vision with its estimation of range and measurement boundary[J]. IEEE Access, 2018, 6: 62450-62457.
- [2] Jiang X D, Yu J Y, Zhu Li K, *et al.* Self-calibrated binocular ranging system based on hardware SURF algorithm[J]. Acta Optica Sinica, 2018, 38(10): 1036001.
蒋晓东, 于纪言, 朱立坤, 等. 基于硬件 SURF 算法的自校准双目测距系统[J]. 光学学报, 2018, 38(10): 1036001.
- [3] Zhou F Q, Wang Y X, Chai X H, *et al.* Review on precise measurement technology based on mirror binocular vision[J]. Acta Optica Sinica, 2018, 38(8): 0815003.
周富强, 王晔昕, 柴兴华, 等. 镜像双目视觉精密测量技术综述[J]. 光学学报, 2018, 38(8): 0815003.
- [4] Xiao J S, Tian H, Zou W T, *et al.* Stereo matching based on convolutional neural network [J]. Acta Optica Sinica, 2018, 38(8): 0815017.
肖进胜, 田红, 邹文涛, 等. 基于深度卷积神经网络的双目立体视觉匹配算法[J]. 光学学报, 2018, 38(8): 0815017.
- [5] Lee D H, Chang J Y, Heo Y S. Stereo matching using cost volume fusion for high dynamic range scenes[J]. Electronics Letters, 2017, 53(23): 1522-1524.
- [6] Zhang L J, Yang J H, Jiang Y T, *et al.* Research on target image matching algorithm for binocular CCD ranging [J]. Laser & Optoelectronics Progress, 2014, 51(9): 091002.
张丽娟, 杨进华, 姜雨彤, 等. 基于双目 CCD 测距的目标图像匹配算法研究[J]. 激光与光电子学进展, 2014, 51(9): 091002.
- [7] Zhou L, Xu G L, Li K Y, *et al.* Stereo matching algorithm based on Census transform and modified adaptive windows [J]. Acta Aeronautica et Astronautica Sinica, 2012, 33(5): 886-892.
周龙, 徐贵力, 李开宇, 等. 基于 Census 变换和改进自适应窗口的立体匹配算法[J]. 航空学报, 2012, 33(5): 886-892.
- [8] Wang Q, Jiang X L, Yuan J Y, *et al.* Regional stereo matching algorithm based on improved Census transform[J]. Application Research of Computers, 2015, 32(10): 3164-3167.
王琼, 江晓亮, 袁建英, 等. 基于改进 Census 变换的区域立体匹配算法[J]. 计算机应用研究, 2015, 32(10): 3164-3167.
- [9] Zhu S P, Yan L N, Li Z. Stereo matching algorithm based on improved Census transform and dynamic programming[J]. Acta Optica Sinica, 2016, 36(4): 0415001.
祝世平, 闫利那, 李政. 基于改进 Census 变换和动态规划的立体匹配算法[J]. 光学学报, 2016, 36(4): 0415001.
- [10] Zhang Z G, Tao W B. Pedestrian detection in binocular stereo sequence based on appearance consistency [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(9): 1772-1785.
- [11] McIntire J P, Wright S T, Harrington L K, *et al.* Binocular fusion ranges and stereoacuity predict positional and rotational spatial task performance on a stereoscopic 3D display [J]. Journal of Display Technology, 2015, 11(11): 959-966.
- [12] Han Y, Qian Y J, Sun J, *et al.* A self-calibration technique for on-chip precise clock generator [J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2015, 62(12): 1114-1118.
- [13] Xie Z X, Zhu R X, Zhang A Q. Extrinsic parameters calibration of three-dimensional measurement system for ultra-large scale line-structured light sensor [J]. Chinese Journal of Lasers, 2017, 44(10): 1004003.
解则晓, 朱瑞新, 张安祺. 超大尺度线结构光传感器三维测量系统外参数标定[J]. 中国激光, 2017, 44(10): 1004003.
- [14] Zheng X C, Zhao M, Feng S H. Two-step calibration of probe tip center of planar target [J]. Laser & Optoelectronics Progress, 2018, 55(1): 011201.
郑兴纯, 赵敏, 冯少华. 平面靶标测头中心的两步法标定[J]. 激光与光电子学进展, 2018, 55(1): 011201.