

文章编号: 2095-2163(2020)02-0029-05

中图分类号: TP305

文献标志码: A

基于 NVIDIA JetSon TX2 的无人机双目测距算法设计

樊建昌¹, 余粟², 尚钰历³

(1 上海工程技术大学 机械与汽车工程学院, 上海 201620; 2 上海工程技术大学 工程实训中心, 上海 201620;

3 上海星河湾双语学校, 上海 201108)

摘要: 无人机在空中飞行时, 实时获取前方障碍物的距离信息是其实现真正意义上的自主飞行的关键。针对无人机自主飞行过程中需要实时获取前方障碍物距离的需求, 本文以 NVIDIA Jetson TX2 AI 超级计算模块为计算平台, 利用 OpenCV 视觉库的相关函数, 对双目摄像头进行了单目标定、双目标定和立体校正, 并在此基础上设计了以 SGBM 算法作为立体匹配算法的双目测距系统。在图像分辨率为 $752 * 480$ 和 $376 * 240$ 的情况下, 通过 20 次实验取均值对系统的测距算法进行了对比, 实验表明, 在图像分辨率为 $376 * 240$ 的情况下, 测距误差为 4.87%, 双目系统的测量距离达到了 7 m, 帧率达到了 8.9 帧, 基本满足了无人机飞行过程中对于障碍物距离测量的实时性的要求。

关键词: 双目视觉; 双目标定; 立体匹配; 双目测距

Design of UAV binocular ranging algorithm based on NVIDIA JetSon TX2

FAN Jianchang¹, YU Su², SAHNG Zhengli³

(1 School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China;

2 Engineering Training Center, Shanghai University of Engineering Science, Shanghai 201620, China;

3 Shanghai Starriver Bilingual School, Shanghai 201108, China)

[Abstract] When the drone is flying in the air, real-time acquisition of the distance information of the obstacle in front is the key to achieving autonomous flight in the true sense. In view of the need to acquire the distance of obstacles in front of the UAV during the autonomous flight, this paper uses the NVIDIA JetSon TX2 AI supercomputing module as the computing platform, and uses the correlation function of the OpenCV visual library to make a single target and double target fixation and stereo correction for the binocular camera, and based on this, a binocular ranging system with SGBM algorithm as stereo matching algorithm is designed. In the case of image resolutions of $752 * 480$ and $376 * 240$, the ranging algorithm of the system is compared by the average of 20 experiments. The experiment shows that the ranging error is 4.87% in the case of image resolution of $376 * 240$. At 4.78%, the measurement distance of the binocular system reaches 7 meters, and the frame rate reaches 8.9 frames, which basically meets the real-time requirements for obstacle distance measurement during the flight of the drone.

[Key words] binocular vision; dual target fixation; stereo matching; binocular ranging

0 引言

随着无人机技术和计算机科学技术的不断发展, 无人机在商业上的应用场景越来越广泛, 从最初的视频业务如 AR/VR、无人机航拍、物流无人机到后来的智能落地物联网、基础设施巡检、无人安防等领域再到如今 AI 时代正大力发展的互联网产品, 所有的信息都表明了无人机巨大的商用价值。但是随着无人机应用领域的逐渐扩大, 各个应用场景对无人机的性能要求越来越高, 主要包括无人机的续航能力、拍摄、避障等关键技术。其中无人机的避障技术由于直接关系到无人机的安全性, 因此一直以来都是科研人员在无人机研究领域的重点和热点问

题。潘泉等人^[1]提出了一种基于毫米波雷达的多旋翼无人机避障技术的设计方案。此方案通过机载避障传感器毫米波雷达来实现多旋翼无人机避障功能; 刘跃波等人^[2]针对无人机电力巡检的应用场景, 提出了基于 T-S 模糊神经网络, 建立模糊控制规则来对无人机的实时避障进行仿真; 金立军等人^[3]结合定位技术和激光测距技术, 先通过惯性测量系统进行初步定位, 得到无人机的经纬度坐标, 然后利用无人机机载激光测距模块, 测量故障点到无人机的距离来达到无人机避障的目的; 饶俊^[4]利用了单载波频域均衡 (SCFDE) 技术以及优化遥控帧与伪码速率关系及发送位置的方法, 降低了机上测

基金项目: 上海市科学技术委员会染色机器人管理软件系统(17511110204)。

作者简介: 樊建昌(1994-), 男, 硕士研究生, 主要研究方向: 计算机视觉; 余粟(1962-), 女, 教授, 主要研究方向: 机电控制、计算机视觉; 尚钰历(2003-), 男, 学生, 主要研究方向: 计算机视觉。

通讯作者: 余粟 Email: yusu@sues.edu.cn

收稿日期: 2019-11-14

距数据下传误码率,提高了无人机系统的测距精度。与前述研究不同的是,本文利用被动测距技术中的双目视觉测距技术,以 NVIDIA Jetson TX2 智能 AI 计算模块为计算平台,以 SGBM 立体匹配算法为基础,设计了无人机测距系统,在图像分辨率为 $376 * 240$ 、平均误差为 4.87% 的情况下,测距长度上达到了 7 m,帧率达到了 8.9 帧,基本上满足了无人机飞行过程中的实时性要求。

1 摄像机成像模型

常见的理想状态下摄像机成像模型如图 1^[5] 所示。模型中包含了 4 个坐标系,分别是:世界坐标系、相机坐标系、图像坐标系以及图像像素坐标系。其中,物体上的任意一点 P 在世界坐标系中的位置记为 (X_w, Y_w, Z_w) ,在摄像机坐标系 xoy 中的坐标为 (X, Y, Z) ,在图像像素坐标系 $X_fO_fY_f$ 中的坐标为 (u, v) ,结合三角形相似推得点 P 由世界坐标系到图像像素坐标系的转化过程可由式(1)表示^[6]:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} * \begin{bmatrix} R_{3*3} \\ T_{3*1} \\ 0 & 0 \end{bmatrix} * \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (1)$$

$$M * N * P_w. \quad (1)$$

其中, s 为缩放因子; M 为摄像机的内参数矩阵,描述了点 P 由相机坐标系到图像坐标系的转化关系, M 由相机自身的安装精度及透镜特性决定; N 为点 P 由世界坐标系到相机坐标系的转换矩阵,称为摄像机模型的外参数矩阵; $R_{3*3} = [r_1, r_2, r_3]$ 是一个正交矩阵, $T_{3*1} = t_{3*1}$ 。二者分别为点 P 由世界坐标系转换到相机坐标系的旋转矩阵和平移矩阵。内参数矩阵 M 以及外参数矩阵 N 可以由单目相机标定求得。

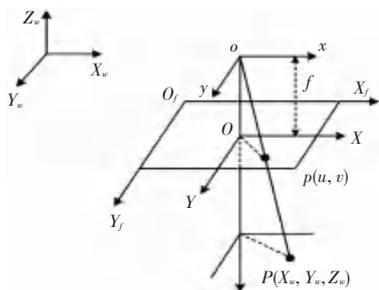


图 1 摄像机成像模型

Fig. 1 Camera imaging model

2 双目测距的基本原理

双目立体视觉测距主要是利用双目相机从不同

视角拍摄同一物体时在图像中得到的不同位置信息来计算图像的视差信息,进一步利用三角形相似的原理求解出物体的三维坐标信息来达到求解物体的距离。其原理图如图 2 所示。

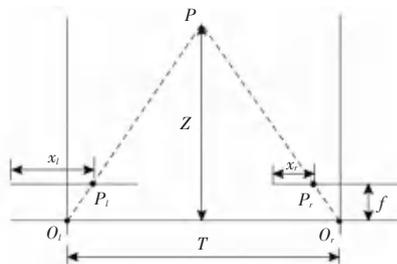


图 2 双目成像原理

Fig. 2 Principle of binocular imaging

在图 2 的平行双目成像系统中, P 表示待测目标上的某一点, O_l, O_r 分别表示左右摄像机坐标系的原点(左右相机的光心), P_l, P_r 分别表示 P 在左右相机 CCD 芯片上的投影点,由图 2 的摄像机成像模型易得, O_l, O_r 与左右相机的 CCD 芯片的中心连线垂直于 P_l, P_r 所形成双目基准平面。 x_l, x_r 分别为物体上的 P 在左右相机成像平面的投影点到成像平面左边界的物理距离, f 为相机焦距, T 为双目相机左右光心之间的物理距离, Z 为物体距离摄像机水平面的真实的物理距离^[7]。

观察图 2,由 $\Delta PP_lP_r \sim \Delta PO_lO_r$ 得:

$$\frac{P_lP_r}{Z-f} = \frac{O_lO_r}{Z}, \quad (2)$$

将 $P_lP_r = T - (x_l - x_r)$ 代入式(2)得:

$$\frac{T + (x_r - x_l)}{Z-f} = \frac{T}{Z}, \quad (3)$$

由相机坐标系和图像坐标系的转换关系可得:

$$x_r - x_l = (u_l - u_r) * d_x, \quad (4)$$

其中, d_x 表示在图像坐标系中,每个像素在 CCD 芯片上所占的实际物理尺寸的大小。将式(4)代入式(3)得:

$$\frac{T + (u_r - u_l) * d_x}{Z-f} = \frac{T}{Z}, \quad (5)$$

$$Z = \frac{f * T}{(u_l - u_r) * d_x} = \frac{T * f_x}{d}, \quad (6)$$

其中, $d = u_l - u_r$ 为物体 P 在左右图像中的视差;

$f_x = \frac{f}{d_x}$ 为相机单目标定时得到的相机的内参数; Z 为要求的物体 P 距离双目相机的距离。

2.1 双目标定原理

双目标定是双目系统中求取左右相机之间的旋

转矩阵 \mathbf{R} 以及平移矩阵 \mathbf{T} 的过程。由图 1~图 2 可得,在不考虑摄像头畸变和相机水平摆放的误差的前提下,物体点 P 由世界坐标系投影到相机坐标系可以用式(7)来表示:

$$\begin{cases} P_l = \mathbf{R}_l * P + \mathbf{T}_l, \\ P_r = \mathbf{R}_r * P + \mathbf{T}_r. \end{cases} \quad (7)$$

进一步可得右相机到左相机的旋转矩阵及平移矩阵如下^[8]:

$$\begin{cases} \mathbf{R} = \mathbf{R}_r(\mathbf{R}_l)\mathbf{R}_r^{-1}, \\ \mathbf{T} = \mathbf{T}_r - \mathbf{R}\mathbf{T}_l. \end{cases} \quad (8)$$

其中, \mathbf{R} 、 \mathbf{T} 即为右相机到左相机的旋转矩阵和平移矩阵。

2.2 立体校正

在实际的双目测距项目中,双目相机由于安装工艺的缺陷常常导致实际的安装情况如图 3 所示^[9],这会使得后期的立体匹配需要在二维方向上进行。所以,在实际的测距过程中,在立体匹配之前常常使用 Bouguet 立体校正使双目相机的左右成像平面平行,并且使得物体在左右图像上的投影点位于同一水平线上。这样做的目的是通过运用极限约束原理使得同一物体在左右两幅图像中的搜索范围从二维降到了一维。

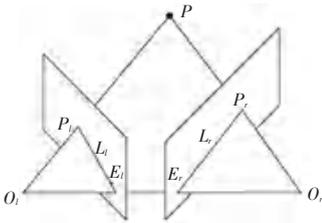


图 3 双目成像系统

Fig. 3 Binocular imaging system

Bouguet 校正矩阵的求解主要分为 2 步,对此可得阐释分述如下。

(1) 求解左右相机共面校正矩阵。由于 \mathbf{R} 是正交实矩阵,则 \mathbf{R} 是正规矩阵^[10],所以有:

$$\mathbf{R} = \mathbf{S} * \mathbf{D} * \mathbf{S}^{-1}, \quad (9)$$

其中, \mathbf{D} 为主对角元素为 \mathbf{R} 的特征值组成的矩阵, \mathbf{S} 为 \mathbf{R} 的特征向量组成的矩阵。进一步拆分式(9)可得:

$$\mathbf{R} = \mathbf{S} * \mathbf{D} * \mathbf{S}^{-1} = (\mathbf{S} * \mathbf{D}^{\frac{1}{2}} * \mathbf{S}^{-1}) * (\mathbf{S} * \mathbf{D}^{\frac{1}{2}} * \mathbf{S}^{-1}), \quad (10)$$

令:

$$\begin{cases} \mathbf{R}_l = \mathbf{S} * \mathbf{D}^{\frac{1}{2}} * \mathbf{S}, \\ \mathbf{R}_r = \mathbf{R}^{-\frac{1}{2}}. \end{cases} \quad (11)$$

则将 2 个相机之间的旋转矩阵 \mathbf{R} 划分为左右相机的合成矩阵 \mathbf{R}_l 、 \mathbf{R}_r , 可以实现左右图像平面共面。但是行没有对齐。

(2) 求解行对校正矩阵。根据式(12)建立行对准矩阵 \mathbf{R}_{rect} 。其中, \mathbf{T} 为由右相机到左相机的平移矩阵。

综合以上 2 步分别计算左右相机的行对准变换矩阵,即:

$$\begin{cases} e_1 = \frac{\mathbf{T}}{\|\mathbf{T}\|}, \\ e_2 = \frac{[-T_y, T_x, 0]}{\sqrt{T_x^2 + T_y^2}}, \\ e_3 = e_1 \times e_2, \\ \mathbf{R}_{rect} = [e_1, e_2, e_3]^T, \end{cases} \quad (12)$$

$$\begin{cases} \mathbf{R}_l = \mathbf{R}_{rect} * r_l, \\ \mathbf{R}_r = \mathbf{R}_{rect} * r_r. \end{cases} \quad (13)$$

2.3 立体匹配

立体匹配是双目测距中最关键的一部分,本文选用立体匹配中最常用的 SGBM 算法作为双目测距中的立体匹配算法,其原理为^[11]:确定左右图像的匹配基元;构建基于多个方向扫描线的代价能量和函数;求取能量代价和函数的最优解。该算法在 OpenCV 中的实现流程具体如下。

(1) 预处理。SGBM 采用水平 Sobel 算子,把图像做处理,再用一个函数将经过水平 Sobel 算子处理后的图像上每个像素点 (P 表示其像素值) 映射成一个新的图像, P_{new} 表示新图像上的像素值。映射函数如下:

$$P_{new} = \begin{cases} 0, & P < -preFilterCap; \\ P + preFilterCap, & -preFilterCap \leq P \leq preFilterCap; \\ 2 * preFilterCap, & P \geq preFilterCap. \end{cases} \quad (14)$$

其中, $preFilterCap$ 是一个常数参数,OpenCV 缺省情况下取 15。

(2) 代价计算。由 2 部分组成:

① 经过预处理后得到的图像的梯度信息利用基于采样^[12]的方法得到的梯度代价。

② 原图像经过基于采样的方法得到的 SAD 代价,其计算公式如下:

$$C(x, y, d) = \sum_{i=-n}^n \sum_{j=-n}^n |L(x+i, y+j) - R(x+i+d, y+j+d)|, \quad (15)$$

(3) 动态规划。SGBM 算法试图通过图像上多个方向上一维路径的约束,来建立一个全局的马尔

科夫能量方程,每个像素最终的匹配代价是所有路径信息的叠加,每个像素的视差选择都只是简单通过 WTA (Winner Takes All) 决定的。多方向能量聚集如图 4 所示。其中,每个方向都按照动态规划的思想进行能量累积,再将所有方向的匹配代价相加得到总的匹配代价,其计算公式可表示为:

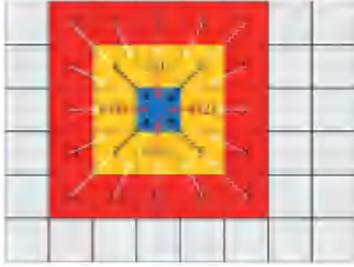


图 4 能量聚集方向

Fig. 4 Energy gathering direction

$$L_r(p, d) = c(p, d) + \min \begin{cases} L_r(p-r, d) \\ L_r(p-r, d \pm 1) + P_1 \\ \min_{i=d_{\min} \dots d_{\max}} L_r(p-r, i) + P_2 \end{cases}$$

$$\min_{i=d_{\min} \dots d_{\max}} L_r(p-r, i),$$

其中, L 为当前路径累积的代价函数; P_1 、 P_2 为像素点与相邻点视差存在较小和较大差异情况下的平滑惩罚, $P_1 < P_2$, 第三项无意义, 仅仅是为了消除各个方向路径长度不同造成的影响。将所有 r 方向的匹配代价相加得到总的匹配代价, 计算公式如下:

$$s(p, d) = \sum_r L_r(p, d). \quad (16)$$

(4) 后处理。主要包括如下内容: 匹配窗口内的视差唯一性检测; 亚像素插值; 左右一致性检测; 连通区域的检测。

3 实验流程及结果

本文以 NVIDIA JetSon TX2 为处理平台和小觅 S1030 (入门版) 双目摄像机搭建实验平台, 详见图 5。这里, 给出研究实验流程及结果的分析表述如下。

(1) 本文以单元格大小为 $5 \text{ m} * 5 \text{ m}$ 的 $12 * 9$ 高精度棋盘格作为标定板, 左右相机分别采集 29 张标定板图像, 经过预处理后利用张氏标定法^[13] 对其进行标定, 得到左右相机的内参数矩阵以及各自的

$$\mathbf{M}_l = \begin{bmatrix} 365.140 & 1 & 0.000 & 0 & 399.323 & 9 \\ 0.000 & 0 & 365.134 & 8 & 234.351 & 6 \\ 0.000 & 0 & 0.000 & 0 & 1.000 & 0 \\ -250.357 & 6 & 505.793 & 9 & -705.366 & 3 & -851.954 & 5 & 0.000 & 0 \end{bmatrix}$$

$$\mathbf{M}_r = \begin{bmatrix} 362.210 & 6 & 0.000 & 0 & 402.323 & 9 \\ 0.000 & 0 & 361.479 & 6 & 234.351 & 9 \\ 0.000 & 0 & 0.000 & 0 & 1.000 & 0 \\ -303.763 & 8 & 799.294 & 3 & 515.475 & 4 & -672.489 & 0 & 0.000 & 0 \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} 9.986 & 790 & 8e-01 & -6.344 & 55e-03 & 5.098 & 845e-02 \\ 5.989 & 03e-03 & 9.999 & 567 & 0e-01 & 7.122 & 42e-03 \\ -5.103 & 144e-02 & -6.807 & 64e-03 & 9.986 & 738 & 4e-01 \end{bmatrix}$$

$$\mathbf{T} = [-1.200 \ 248 \ 976 \ 4e+02, -1.178 \ 263 \ 74e+00, -5.205 \ 820 \ 51e+00]$$

畸变参数, 标定结果界面如图 6 所示, 标定结果数值具体如下:

$$\mathbf{D}_l = \begin{bmatrix} -250.357 & 6 & 505.793 & 9 & -705.366 & 3 & -851.954 & 5 & 0.000 & 0 \end{bmatrix}$$

$$\mathbf{D}_r = \begin{bmatrix} -303.763 & 8 & 799.294 & 3 & 515.475 & 4 & -672.489 & 0 & 0.000 & 0 \end{bmatrix}$$

其中, \mathbf{M}_l 、 \mathbf{M}_r 分别为左右相机的内参数矩阵, \mathbf{D}_l 、 \mathbf{D}_r 分别为左右相机的畸变参数矩阵。



图 5 实验平台

Fig. 5 Experimental platform



图 6 标定结果

Fig. 6 Calibration results

(2) 选取 29 组左右相机的图片, 标定板图片的分辨率为 $752 * 480$ 。利用 OpenCV 视觉库编写双目标定程序, 以左相机坐标系作为双目标定的基准坐标系, 通过 OpenCV 中的 `stereoCalibrate()` 函数计算得到旋转矩阵 \mathbf{R} 和平移矩阵 \mathbf{T} 结果如下:

(3) 利用(1)中得到的左右相机的内参数和畸变参数,以及(2)中的得到的左右相机之间的外参数,再利用通过 OpenCV 视觉库中的 *cvStereoRectify()* 和 *initUndistortRectifyMap()* 函数计算出左右相机的立体校正和畸变校正参数矩阵。

(4) 利用 OpenCV 视觉库中的 *remap()* 函数对左右图像进行立体校正和畸变校正,使得左右相机图像处于行对准的理想状态。

(5) 以 SGBM 立体匹配算法为基础,利用 OpenCV 视觉库编写双目测距的运行算法,计算出左右图像的视差图。在此基础上,进一步利用 OpenCV 视觉库中的 *cvReprojectImageTo3D()* 函数计算得到障碍物距离双目相机的距离,其中 SGBM 算法中的相关参数设置及含义如图 7 所示。

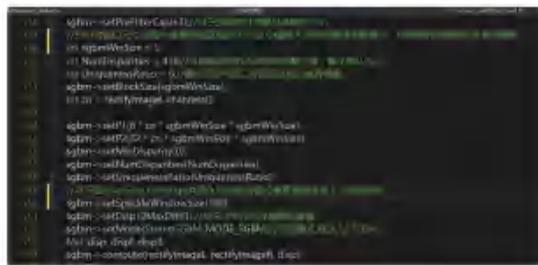


图 7 SGBM 算法及参数

Fig. 7 SGBM algorithm and parameters

本文在以 NVIDIA JetSon TX2 为计算平台上,选取 3 m、5 m、7 m 作为双目测距的基准距离,在近似于无人机空中飞行的简单场景下,采用 20 次测量结果取平均值,得到的实验结果见表 1。

表 1 双目测距实验结果

Tab. 1 Binocular ranging test results

图像分辨率	距离真值/m	测试距离值	误差/%	帧率
752 * 480	3	3.01	0.333	4.5
	5	4.98	0.400	4.4
	7	7.03	0.428	4.5
376 * 240	3	2.98	0.666	8.9
	5	5.06	1.200	8.8
	7	7.34	4.875	8.9

4 结束语

由实验结果可得,对于SGBM算法,在752 *

480 的分辨率下,本文设计的双目系统测距的准确度要高于 376 * 240 的准确率,但是在实时性上,752 * 480 的分辨率在 4.6 帧左右,很难满足无人机的高速飞行,在 376 * 240 的分辨率下,当测量距离在 7 m 时,测距误差在 4.8% 左右,其测距的准确性上不如 752 * 480 的分辨率,但是其帧率达到了 8.9 帧,基本满足了无人机在高速飞行时的实时性要求。

与此同时,本系统也仍存在一定的问題,例如在单目标定时,由于所用的标定板过小,导致标定得到的相机的内参数和外参数的精度还有待提高。测距的准确度也亟需改善,实时性上还可以基于 CUDA 加速做进一步的加速计算,本系统仍有很大的提升空间。

参考文献

[1] 潘泉,王伟,华锡焱,等. 基于毫米波雷达的多旋翼无人机避障技术研究[J]. 中小企业管理与科技(下旬刊),2019(7):170.

[2] 刘跃波,马立新. 电力巡检无人机避障技术研究[J]. 电子测量技术,2019,42(13):1.

[3] 金立军,王恒,王文华,等. 基于机器视觉和激光测距的输电线故障定位[J]. 同济大学学报(自然科学版),2018,46(12):1745.

[4] 饶俊. 一种抗多径效应的非相干无人机测距方案[J]. 电讯技术,2018,58(7):774.

[5] 汪永超,耿丽清. 摄像机标定系统方法研究[J]. 天津职业技术师范大学学报,2019,29(2):59.

[6] 潘雨坤. 基于双目立体视觉的运动目标检测系统设计与实现[D]. 上海:华东师范大学,2019.

[7] 王华锋,李飞. 基于机器学习与双目视觉相结合的测距方法研究[J]. 北方工业大学学报,2019,31(2):23.

[8] 张智,郭文县,苏丽,等. 机器人双目图像实时立体匹配算法[J]. 北京理工大学学报,2017,37(9):975.

[9] 张树立. 双目视觉测距系统的研究[D]. 石家庄:河北科技大学,2019.

[10] ALVES S M, TIAN Zhu, SHOSTAK A, et al. Studying normal tissue radiation effects using extracellular matrix hydrogels[J]. Journal of visualized experiments: JoVE,2019(149):e59304.

[11] 张欢,安利,张强,等. SGBM 算法与 BM 算法分析研究[J]. 测绘与空间地理信息,2016,39(10):214.

[12] BIRCHFIELD S, TOMASI C. Depth discontinuities by pixel-to-pixel stereo[J]. International Journal of Computer Vision,1999,35(3):269.

[13] ZHANG Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000,22(11):1330.

descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2005,27(10):1615.

[15] MIKOLAJCZYK K, SCHMID C. An affine invariant interest point detector[M]// HEYDEN A, SPARR G, NIELSEN M, et al. Computer Vision — ECCV 2002. Lecture Notes in Computer Science. Berlin /Heidelberg:Springer, 2002,2350: 128.

(上接第 28 页)

[13] LEVI G, HASSNER T. LATCH: Learned arrangements of three patch codes[C]//2016 IEEE Winter Conference on Applications of Computer Vision (WACV). Lake Placid, NY, USA :IEEE, 2016:1-9.

[14] MIKOLAJCZYK K, SCHMID C. Performance evaluation of local