

# 大视场双目视觉定位系统中多目标稀疏匹配

王 鑫<sup>1,2</sup>, 王向军<sup>1</sup>

- (1. 天津大学 微光机电系统技术教育部重点实验室, 天津 300072;
2. 北华航天工业学院 电子与控制工程学院, 河北 廊坊 065000)

**摘 要:** 野外远距离、大视场、双目视觉物体定位监测系统是基于双目立体视觉原理的三维坐标测量系统。存在随机出现的多个无特征目标的正确匹配率低, 导致目标定位精度降低的共性问题。正确判断首个目标出现, 对于同一目标在同侧图像序列连续出现, 形态特征由大到小变化或分裂的情况, 利用质心判断方法排除伪目标, 检测目标正确质心位置; 然后, 两段图像序列时间同步, 同一时刻多个目标基于极线约束进行稀疏匹配, 获得同一目标在两台相机中正确的对应关系。实验证明, 现场连续采集的 9 帧图像中包含 8 个目标, 利用上述方法完成目标检测及匹配, 提高了准确率。

**关键词:** 双目视觉; 稀疏匹配; 目标检测; 极线约束

**中图分类号:** V243.5; TP391.4 **文献标志码:** A **DOI:** 10.3788/IRLA201847.0726001

## Multiple targets sparse matching for binocular vision positioning system with large field of view

Wang Xin<sup>1,2</sup>, Wang Xiangjun<sup>1</sup>

- (1. Key Laboratory of MOEMS of the Ministry of Education, Tianjin University, Tianjin 300072, China;
2. Institute of Electronic and Control Engineering, North China Institute of Aerospace Engineering, Langfang 065000, China)

**Abstract:** A positioning and monitoring system for the objects with long distance, large view field and binocular vision was a three-dimensional coordinate measuring system based on binocular stereo vision principle. There was a common problem that the low correct matching rate of multiple non-feature targets randomly appearing led to the reduced accuracy of target positioning. The key was judging the trigger start time correctly. For the situation of the same target appearing continuously in the same side of the image sequence, or the morphological characteristic of the target changing from large to small or splits, the use of centroid method was to eliminate pseudo targets and detect target centroid position correctly. Then image sequences had time synchronization, and multiple targets realized sparse matching based on polar constraint at the same time. The experiment proves that 9 frame live continuous images contain 8 targets, using the above method to complete the targets detection and automatic positioning, and improve the accuracy rate.

**Key words:** binocular vision; sparse matching; target detection; polar constraint

收稿日期: 2018-02-11; 修订日期: 2018-03-17

基金项目: 国家重点研发计划(017YFC0601900, 2017YFC0601901); 国家教育部支撑项目(625010110);

国家自然科学基金与中国民用航空总局联合研究基金(61179043); 河北省重点研发计划自筹项目(17211708);

河北省高等学校科学技术研究项目(QN2018209)

作者简介: 王鑫(1987-), 女, 讲师, 博士生, 主要从事视觉测量、低空遥感测量及图像处理等方面的研究。Email: wxtju1987@163.com

导师简介: 王向军(1955-), 男, 教授, 博士, 主要从事精密测试与仪器、光电探测与传感技术、影像与视觉测量方面的研究。

Email: xdocuxjw@vip.163.com

## 0 引言

双目立体视觉是计算机视觉中的一个非常重要的分支,是目标跟踪定位<sup>[1-3]</sup>、姿态估计<sup>[4-5]</sup>、三维重建<sup>[6]</sup>、相机标定<sup>[7]</sup>等研究领域的关键技术。利用双目摄像机模拟人的视觉系统,基于视差原理并利用成像设备从不同的位置同时获取被测目标区域的两幅图像,通过计算图像对应点间的位置偏差获得目标三维信息。双目立体视觉测量方法具有系统结构简单、测量效率高、测量精度高等诸多优势,适合于在线非接触测量。

文中研究针对野外大视场、远距离的实际测量环境,搭建了基于双目立体视觉原理的三维坐标测量系统<sup>[8]</sup>。关键解决多个体积近似相等的无特征目标随机稀疏分布至平面,基于双目视觉系统通过两台摄像机获得的图像序列对目标进行立体匹配。解决方法不因测量背景而变化,因此具有普遍适用性。

无特征目标在图像表征上为几十像素的高亮圆斑,多个目标稀疏分布,且无典型特征,很难从物体特征如尺寸、形状、纹理等方面考虑,只能通过空间几何约束搜索角度实现。目前应用较多的立体匹配方法中基于图像几何和场景约束的方法有:极线约束将二维图像中搜索匹配点降为在一条直线上搜索,大大压缩了对应点的搜索区域;几何相似性约束,当视角变化较大,几何特性会发生相应变化;单应性约束<sup>[9]</sup>平面监测区域较理想,起伏区域误差较大;双极线约束<sup>[10]</sup>要求摄像机数目大于两个,对测量系统的复杂度和成本要求更高;ICP 迭代最近点算法<sup>[11]</sup>需要大量点云数据,并且收敛速度慢。

文中提到的双目视觉定位系统用于大视场中物体位置解算,监测范围 500 m×1 000 m,监测区域多为起伏山地,CCD 摄像机距离监测区域中心点约 1~2 km。单一使用上述方法都不能解决文中提出的问题,因此提出了一种基于时间同步和极线约束的立体匹配方法。在充分排除伪目标及重复计算目标,剩余真实目标,一帧图像有多个目标同时出现的情况下,将时间信息与空间信息相结合,确定同一目标在两台摄像机中正确的对应关系解决了超大三维空间下无特征多目标稀疏匹配问题。

## 1 双目视觉定位系统结构

系统由一台中心控制计算机和两侧监测平台通过无线网络构成,如图 1 所示。

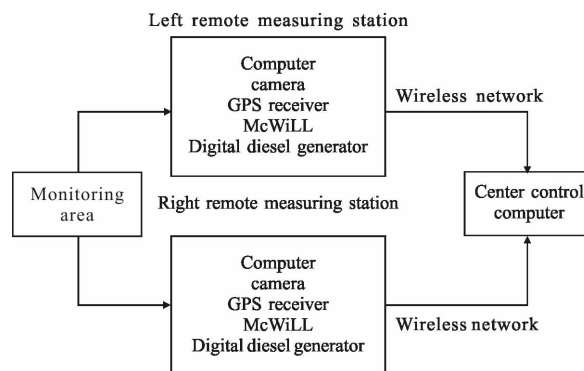


图 1 自动目标定位系统示意图

Fig.1 Block diagram of automatic target positioning system

两侧监测平台受中心控制计算机控制,持续监测目标区域,正确判断首个目标出现,开始连续存储视频图像并进行相应预处理工作,将真实目标质心像面坐标、时间等信息通过无线网络回传给中心控制计算机<sup>[12]</sup>。

中心控制计算机用于远程控制前端测量站,接收前端监测平台上传的目标信息,根据时间同步及极线约束的几何条件得到多个无特征目标正确的对应关系,从而根据双目立体视觉系统模型计算出各个目标的深度定位信息。

## 2 基于摄像机透视投影的定位解算模型

如图 2 所示,利用透视矩阵变换建立大地坐标系、摄像机坐标系及像面坐标系之间的映射关系<sup>[13]</sup>:

$$z_c \begin{bmatrix} u \\ v \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = MX_w \quad (1)$$

式中: $u_0, v_0$ 为像面中点坐标(单位:pixel); $dx, dy$ 为像素在两二维轴向上的物理尺寸; $f$ 为摄像机焦距,可由器件参数得到; $u, v$ 为目标在像面坐标系下的坐标; $X_w, Y_w, Z_w$ 为目标在大地坐标系中的坐标<sup>[14]</sup>; $R$ 为 $3 \times 3$ 正交旋转矩阵,与摄像机三个自由度(自转角 $\alpha$ 、俯仰角 $\beta$ 、倾斜角 $\gamma$ )有关; $T$ 为 $3 \times 1$ 的平移矩阵,反

映了摄像机坐标系相对于大地坐标系在坐标轴上的相对位移。俯仰角  $\beta$ 、倾斜角  $\gamma$  近似为 0, 自转角  $\alpha$  由摄像机标定获得, 平移矩阵由 GPS 获得。

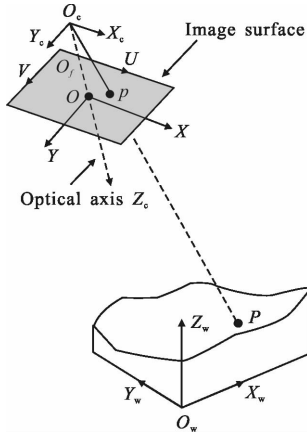


图 2 摄像机透视变换模型

Fig.2 Camera perspective transformation model

正交旋转矩阵  $R$  为:

$$R = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix} \quad (2)$$

公式(1)可写为:

$$z_c \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha & \cos\alpha S_x - \sin\alpha S_z \\ 0 & 1 & 0 & S_y \\ \sin\alpha & 0 & \cos\alpha & \sin\alpha S_x + \cos\alpha S_z \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (3)$$

将公式(3)按矩阵相乘, 结果为下面一方程组:

$$\begin{cases} f\cos\alpha X_w - f\sin\alpha Z_w + f(\cos\alpha S_x - \sin\alpha S_z) = z_c X \\ fY_w + fS_y = z_c Y \\ \sin\alpha X_w + \cos\alpha Z_w + \sin\alpha S_x + \cos\alpha S_z = z_c \end{cases} \quad (4)$$

将方程组中第三个等式代入第一个方程得到:

$$(f - X\tan\alpha)X_w - (f\tan\alpha + X)Z_w + (f - X\tan\alpha)S_x - (f\tan\alpha + X)S_z = 0 \quad (5)$$

上式是单一摄像机某目标点的像面横坐标和三维空间世界坐标系下坐标满足的方程。

设左、右两摄像机的焦距分别为  $f_1, f_2$ , 自转角分别为  $\alpha_1, \alpha_2$ , 平移矢量为  $(S_{x1}, S_{z1}), (S_{x2}, S_{z2})$ , 目标点像面横坐标为  $X_1, X_2$ , 对公式(5)的各项系数设置如下, 其中  $i=1, 2$ :

$$\begin{cases} A_i = f_i - X_i \tan\alpha_i \\ B_i = -(f_i \tan\alpha_i + X_i) \\ C_i = (f_i - X_i \tan\alpha_i)S_{xi} - (f_i \tan\alpha_i + X_i)S_{zi} \end{cases} \quad (6)$$

左、右两摄像机的方程联系为:

$$\begin{cases} A_1 X_w + B_1 Z_w + C_1 = 0 \\ A_2 X_w + B_2 Z_w + C_2 = 0 \end{cases} \quad (7)$$

目标点坐标的表达式为:

$$X_w = \frac{C_2 B_1 - C_1 B_2}{A_1 B_2 - A_2 B_1} \quad (8)$$

$$Z_w = \frac{A_1 C_2 - A_2 C_1}{A_2 B_1 - A_1 B_2} \quad (9)$$

### 3 目标检测

监测平台计算机及相机开启后等待自动触发。触发开始时刻是指目标区域出现首个目标点, 并且相机通过图像处理算法获取到该目标点的信息。从该时刻开始再保存图像序列, 否则存储数据量太大。首个目标出现的原始图像如图 3 所示, 图中白色高亮部分为首个目标。原始图像为加了可见光截止滤镜获取的实测原始图像。



图 3 原始图像

Fig.3 Original image

经过原始图像的预处理, 初步判断确定出所有可能目标点的质心位置, 由于目标点属于炮弹落地爆炸后的影像信息, 火光从无到有, 逐渐变亮, 面积变大, 再到逐渐熄灭, 这个特殊过程使得同一个目标在连续不同帧图像重复出现, 为避免重复计算, 剔除伪目标, 需对此进行处理。对当前目标  $A$  而言, 遍历后一帧所有目标, 存在某一目标  $B$ ,  $A$  的质心位置在  $B$  的像素范围内, 或者  $B$  的质心位置在  $A$  的像素范围内, 则认为  $A, B$  为同一目标。如此进行, 判断目标  $A$  持续帧数, 将其出现的首帧图像的帧号作为该目标的帧号信息。

实际获得的图像中, 目标常常伴有如图 4 所示的变化情况, 即一个目标  $A$  首次出现之后, 在后续连续多帧图像会分裂成两个目标, 如图白色矩形框注的目标, 在第二帧图像上分裂成两个目标  $A', B'$ ,

第三帧图像分裂的两个目标  $A''$ 、 $B''$  像素面积减小，亮度减弱。处理方法是按照图 5 流程判断目标  $A$  与目标  $A'$  质心位置互相包括，目标  $A$  在持续两帧，但不排除第二帧其他目标也与目标  $A$  有同属一个目标的情况，应逐一检查其余所有目标，对第二帧同属于一个目标的多个分裂小目标作出标注。判断第三帧目标是否为目标  $A$  时， $A'$ 、 $B'$  与  $A''$ 、 $B''$  交叉判断质心互相是否包括，即  $A'$  与  $A''$ 、 $A'$  与  $B''$ 、 $B'$  与  $A''$ 、 $B'$  与  $B''$  只要有一个满足，就说明目标  $A$  又持续一帧。

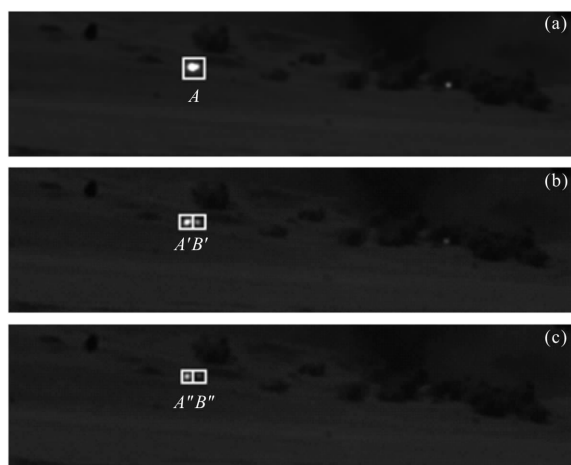


图 4 目标分裂过程图

Fig.4 Diagram of target splitting

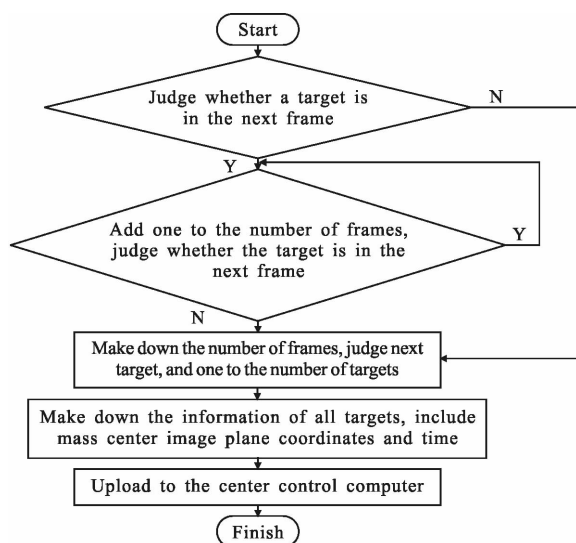
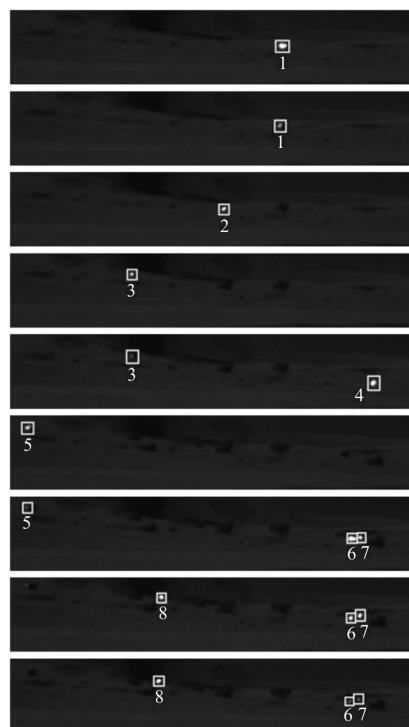


图 5 去除伪目标流程图

Fig.5 Flow chart of removing the false targets

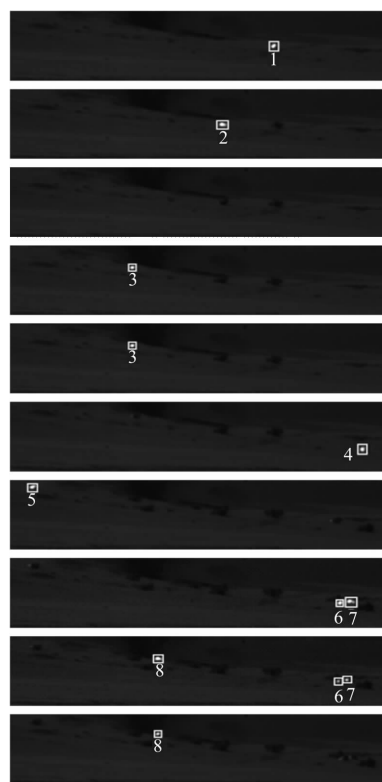
#### 4 无特征多目标稀疏匹配

时间同步后，两侧监测平台触发同时开始，各帧图像对应。图 6 所示为某次现场实验左右监测平台



(a) 左监测平台图像

(a) Images of left monitor platform



(b) 右监测平台图像

(b) Images of right monitor platform

图 6 监测平台图像

Fig.6 Images of monitor platform

获得的原始图片,各九帧,一共八个目标,已在图中依次按序号标注。通过图像显示可知,测试环境为大场景下无特征小目标,对于大视场,目标只有十几或几十像素,目标分布稀疏<sup>[15]</sup>。表 1 为总结的各个目标所在帧号,利于找到它们之间的对应关系。

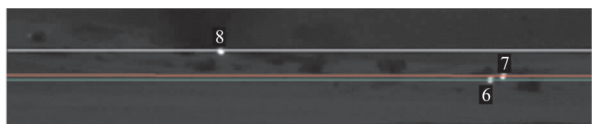
表 1 目标序号信息

Tab.1 Targets number information

Frame number	First round		Second round		Third round	
	Left	Right	Left	Right	Left	Right
1	1	1	-	-	-	-
2	1	2	-	2	-	-
3	2	-	2	-	-	-
4	3	3	-	-	-	-
5	3,4	4	4	4	-	-
6	5	5	-	-	-	-
7	5,6,7	6,7	6,7	6,7	6,7	6,7
8	6,7,8	6,7,8	6,7,8	6,7,8	6,7,8	6,7,8
9	6,7,8	8	6,7,8	8	6,7,8	8

表 1 中第 1、2、3 列已列出左、右八个目标出现的帧号,“-”表示没有目标出现。同一时刻,左、右场景下都只有一个目标的情况认为是同一个目标。由此可知,第一、四、六帧对应为同一目标,找到正确对应关系的目标即可先排除,以免影响其余目标,即新的目标信息为第二轮显示内容。由于相机采集帧率为 25 帧/s,同一目标可能在左右相邻帧被采集到,目标 2 即为该情况。第二轮找到目标 2、4 的对应关系后,剩余第三轮显示信息。前两轮结束时,确定了第 1、3、5 三个目标的对应关系,而不是按照时间先后顺次匹配。第七帧中同时存在两个目标,第八帧中同时存在三个目标,第九帧中三个目标对应一个目标,均无法找到对应关系,因此对于同一时刻的多个目标,再利用极线约束关系找到正确的对应关系。

计算右侧 6、7、8 号目标在左图像上对应极线,依次分别为绿、红、白三色。左侧 6、7、8 号目标在右图像上对应极线,依次分别为绿、红、白三色。如图 7 所示。



(a) 左图像目标与对应极线

(a) Targets in the left image with corresponding polarline



(b) 右图像目标与对应极线

(b) Targets in the right image with corresponding polarline

图 7 目标与对应极线

Fig.7 Targets with corresponding polarline

经过计算可知,表 2 为左侧监测平台采集的第八帧图像中三个目标到右侧三个目标对应的各个极线的距离,表 3 为右侧监测平台采集的第八帧图像中三个目标到各个极线的距离。

表 2 左侧目标到各极线距离(单位:pixel)

Tab.2 Distance from left targets to polarline

(Unit: pixel)

	Left No.6 target	Left No.7 target	Left No.8 target
Right No.6 target corresponding polarline	2.4	6.5	31.3
Right No.7 target corresponding polarline	7.1	2.8	33.5
Right No.8 target corresponding polarline	35.7	38.1	1.8

表 3 右侧目标到各极线距离(单位:pixel)

Tab.3 Distance from right targets to polarline

(Unit: pixel)

	Right No.6 target	Right No.7 target	Right No.8 target
Left No.6 target corresponding polarline	1.2	5.6	35.7
Left No.7 target corresponding polarline	5.1	1.6	38.9
Left No.8 target corresponding polarline	40.4	42.1	2.3

根据表中数据可知,左侧 6 号目标到右侧 6 号目标对应极线距离最短,因此 6 号目标与该极线对应的右侧 6 号目标存在对应关系,反之亦满足。7 号、8 号目标也根据该极线约束关系下得到的点与极线的关系确定其在另一侧的对应点。这一对应关系与图 7 描述一致。

## 5 结论

基于摄像机透视投影模型的双目立体视觉系

统,可以获得大视场下稀疏分布的目标的深度信息。由于目标无背景、纹理、形状等特征,目标总体数量多不唯一,且同一帧图像上同时存在多个目标,在确定其正确对应关系上存在较大困难。目标检测过程中剔除多余伪目标以及重复出现的目标,获得正确目标质心位置,是后续可以确定两侧目标对应关系的重要前提。分析多种几何约束搜索和迭代方法后,利用时间、极线约束几何关系等条件逐步确定多个无特征目标的正确对应匹配关系。当目标分布密集时,对应极线易发生重叠,降低该方法的可靠性,后续将融合其他方法进行进一步研究。

#### 参考文献:

- [1] Wang Tingting, Cai Zhihao, Wang Yingxun. UAV target tracking and location based on binocular vision [J]. *Electronics Optics & Control*, 2016(5): 6–10. (in Chinese)
- [2] Jiang Shan, Zhang Rui, Han Guangliang, et al. Moving object tracking based on multi-feature fusion in the complex background gray image [J]. *Chinese Optics*, 2016, 9(3): 320–328. (in Chinese)
- [3] Wang Xiangjun, Le Bing, Xing Feng, et al. Method of measuring target LOS angular rate with sequence images[J]. *Infrared and Laser Engineering*, 2017, 46(8): 0828001. (in Chinese)
- [4] Li Zhengwei, Wang Jianli, Wu Yuanhao, et al. Method of attitude estimation for space object based on single ground-based telescope[J]. *Chinese Optics*, 2016, 9(3): 371–378. (in Chinese)
- [5] Huang Nannan, Liu Guixi, Zhang Yinzhe, et al. Unmanned aerial vehicle vision navigation algorithm [J]. *Infrared and Laser Engineering*, 2016, 45(7): 0726005. (in Chinese)
- [6] Wang Xin, Yuan Kun, Yu Xiao, et al. Design of binocular vision 3D reconstruction system based on motion recovery [J]. *Optics and Precision Engineering*, 2014, 22(5): 1379–1387. (in Chinese)
- [7] Wan Yilong, Bai Lianfa, Han Jing, et al. Method and realization of significant target recognition and distance ranging in the binocular stereo vision under low illumination [J]. *Infrared and Laser Engineering*, 2015, 44(3): 1053–1060. (in Chinese)
- [8] Wang Xin. Research on objects in the same characteristics matching and positioning based on binocular stereo vision [D]. Tianjin: Tianjin University, 2013. (in Chinese)
- [9] Bentolila J, Francos J M. Homography and fundamental matrix estimation from region matches using an affine error matrix [J]. *Journal of Mathematical Imaging & Vision*, 2014, 49(2): 481–491.
- [10] Guo Lei, Zhu Jigui, Ye Shenghua. Precision registration of corresponding uncode points in large-sized 3D surface measurement using three images [J]. *Infrared and Laser Engineering*, 2008, 37(6): 1079–1082. (in Chinese)
- [11] Wei Shengbin, Wang Shaoqing, Zhou Changhe, et al. An iterative closest point algorithm based on biunique correspondence of point clouds for 3D reconstruction [J]. *Acta Optica Sinica*, 2015, 35(5): 0515003. (in Chinese)
- [12] Lu Yongwei, Wang Xiangjun, Zhang Ninging, et al. Multi-grade target-scoring system based on McWiLL and RS-485 network[J]. *Computer Engineering*, 2013, 39(8): 314–317. (in Chinese)
- [13] Zhang Guangjun. Vision Measurement [M]. Beijing: Science Press, 2008: 34–35. (in Chinese)
- [14] Wang Xiangjun, Wang Jing, Liu Feng, et al. Fast calibration for binocular vision measuring system with a large field[J]. *Optics and Precision Engineering*, 2013, 21(10): 2664–2670. (in Chinese)
- [15] Yan Nian. A study on matching of objects with the same feature for stereo vision [D]. Tianjin: Tianjin University, 2015. (in Chinese)