



引用格式:陈建明,时铭慧. 结合改进的 SIFT 算法的双目视觉测距[J]. 轻工学报,2018,33(5):90-96.

中图分类号:TP751 文献标识码:A

DOI:10.3969/j.issn.2096-1553.2018.05.012

文章编号:2096-1553(2018)05-0090-07

结合改进的 SIFT 算法的双目视觉测距

Binocular vision ranging combined with improved SIFT algorithm

陈建明,时铭慧

CHEN Jianming, SHI Minghui

华北水利水电大学 电力学院,河南 郑州 450045

College of Electric Power, North China University of Water Resources and Electric Power, Zhengzhou 450045, China

关键词:

双目视觉测距;摄像机标定;SIFT 算法

Key words:

binocular vision ranging;
camera calibration;
SIFT algorithm

摘要:针对采用双目视觉测距进行无人机实时巡线测距时计算复杂、实时性不高的问题,提出了结合改进的 SIFT 算法的双目视觉测距方案.该方案采用类单目的双目视觉模型,以减少计算变量;采用缩小尺度空间个数、降低特征向量维度、街区距离代替欧式距离的方法对传统的 SIFT 算法进行改进,以提高两摄像机同时拍摄的两幅图片同一特征点的匹配效率.结果表明:该方法匹配时间缩短了 32%,测量最大误差率降至 3.75%,可满足无人机快速实时巡线测距的精度要求.

收稿日期:2018-06-21

作者简介:陈建明(1962—),男,河南省洛阳市人,华北水利水电大学教授,主要研究方向为检测与自动化装置、模式识别与人工智能.

Abstract: In order to solve the problem of complex calculation and low real-time performance when using binocular vision ranging for UAV real-time line patrol, a binocular vision ranging scheme combining improved SIFT algorithm was proposed. A binocular vision model similar to monocular vision was adopted to reduce computation variables, the traditional SIFT algorithm was improved by reducing the number of scale spaces, reducing the dimension of feature vectors, and replacing the Euclidean distance with the block distance, so as to improve the matching efficiency of the uniform feature points of two images taken simultaneously by two cameras. The results showed that the matching time was reduced by 32% and the maximum error rate was reduced to 3.75%. The method could meet the accuracy requirements of UAV fast real-time line patrol ranging.

0 引言

视觉测距是指通过图像采集装置将被采集目标转换成图片信息并传给专门的处理系统的检测方式,相比于超声、红外、激光等传统检测方式,视觉测距以其携带信息丰富、测量精准等诸多优势在人工智能方面的应用越来越广泛^[1-2]。在测量无人机与输电线路的距离时,上述传统的测距方式不能准确检测到目标,故通常选择视觉测距的方式。视觉测距根据图像采集装置的个数可以分为单目视觉、双目视觉和多目视觉^[3-5],其中双目视觉是通过比较某一目标点在两台摄像机上成像点的差距,利用相对运动学原理,还原出距离的方法。采用双目视觉测距的核心^[4]就是匹配出两台摄像机拍摄两幅图片中的同一特征点。常用的特征匹配算法有 SIFT 算法、SURF 算法、KAZE 算法等^[6-9],其中 SIFT 算法具有旋转不变性,对噪声、视角变化比较稳定,鲁棒性较好;但其计算复杂,实时性不够。针对 SIFT 算法的不足,宋佳乾等^[10]利用 Canny 边缘检测算法获得图像的边缘点坐标,然后通过 K-L 变换降低算法复杂度;刘佳等^[11]采用“回”字形双层方邻窗将特征点邻域区域划分成四部分,建立 32 维特征点描述符向量。上述两种做法虽然都在一定程度上降低了计算的复杂程度,但同时也大大降低了点对的正确匹配率。李奇敏等^[12]以曼哈顿距离作为最邻近特征点查询中的相似性度量简化计算,但实时性仍然不高。

鉴于此,本文拟采用改进的 SIFT 算法,在实现图像中同名点匹配的同时减少计算量,结合双目视觉模型^[13-14],测量无人机在巡线过程中与目标输电线路的距离,以满足无人机快速实时巡线测距的精度要求,从而为无人机巡线提供参考。

1 双目视觉测距

1.1 测距模型

常见的双目视觉测距的模型有两种:一种是两台摄像机左右平行摆放,两台摄像机连线与光轴在同一水平平面内,且垂直于光轴,如图 1a)所示;另外一种为两台摄像机上下平行摆放,两台摄像机连线与光轴在同一竖直平面内,且垂直于光轴,如图 1b)所示。以上两种方法在应用中同时存在一个问题:同一点在两摄像机中的成像会发生偏移,与光心连线的角度会超过 180° ,所以在计算过程中就不能单单用距离来表示无人机与输电线路的直线距离,还必须加上角度。为减少变量,在计算时只考虑长度而不考虑角度的影响,本文创新性地采用如图 1c)所示的类单目的双目视觉测距模型,其中,两条竖实线表示两个成像平面,两台摄像机相同,焦距均为 f , o_1 和 o_2 分别是两台摄像头的焦点,两台摄像头距离为 d , d_1 和 d_2 分别为两台摄像头距离目标的直线距离, S'_1 和 S'_2 分别为目标点在两成像平面上所成的点。两台摄像机前后摆放,连线近似看作与光轴重合,相当于摄像机在远处拍摄一张照片之后沿光轴向目标物

体靠近,因此,目标点在两台摄像机成像平面中的点与光心连线角度相等。

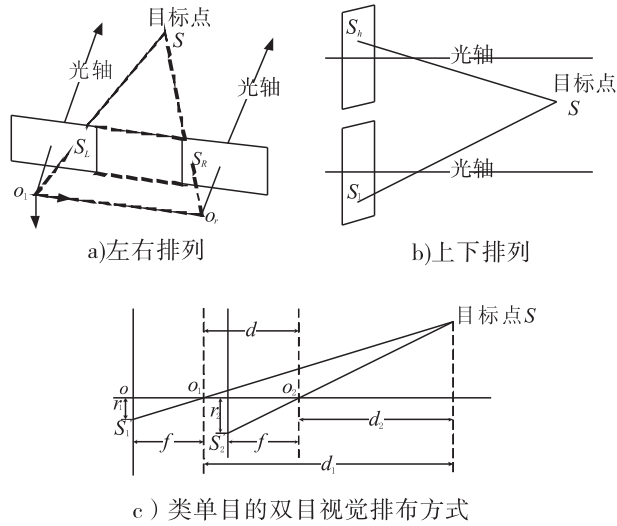


图1 双目视觉测距摄像机排布方式
Fig.1 The arrangement of binocular vision ranging camera

1.2 测距原理

本文所用的双目视觉测距的两光屏成像如图2所示,图2a)是后摄像机即距离目标物较远的摄像机成像,图2b)是前摄像机即距离目标物较近的摄像机成像,在像素坐标系中, S'_1 点的坐标为 (i,j) , S'_2 点的坐标为 (m,n) ,在世界坐标系中,目标点 S 的坐标为 (x,y,z) ,设目标点与光轴所在直线的距离为 R 。

由图1c)和图2可得

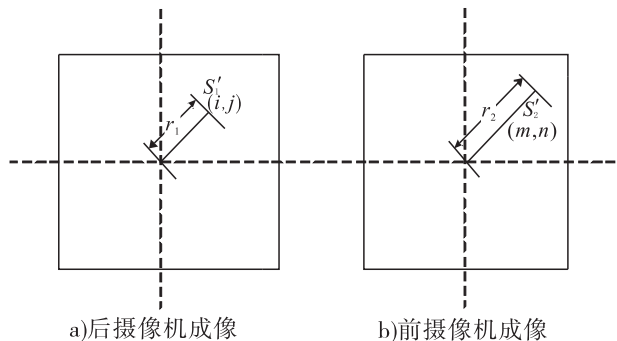


图2 双目视觉测距两光屏成像图
Fig.2 Two-screen imaging of binocular vision ranging

$$r_1 = \sqrt{i^2 + j^2}$$

$$r_2 = \sqrt{m^2 + n^2}$$

$$\frac{r_1}{f} = \frac{R}{d_1} \Rightarrow R = \frac{r_1 d_1}{f}$$

$$\frac{r_2}{f} = \frac{R}{d_2} \Rightarrow R = \frac{r_2 d_2}{f}$$

$$d_1 - d_2 = d$$

将上述公式联立,可得近端摄像头焦距距离目标点的距离为

$$d_2 = \frac{r_1 d}{r_2 - r_1} = \frac{d \sqrt{i^2 + j^2}}{\sqrt{m^2 + n^2} - \sqrt{i^2 + j^2}}$$

2 传统的 SIFT 算法及其改进

采用类单目视觉的双目视觉测距的核心就是匹配出两台摄像机拍摄两幅图片中的同一特征点,本方案中采用改进的 SIFT 算法匹配同一特征点。

2.1 传统的 SIFT 算法

SIFT 算法即尺度不变特征转换算法,用于两幅图片中特征点的匹配,主要由特征提取和特征匹配两个过程组成.特征提取主要由4个步骤组成,SIFT 算法流程图如图3所示。

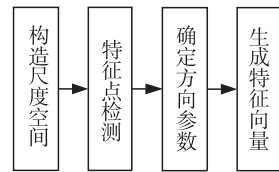


图3 SIFT 算法流程图

Fig.3 SIFT algorithm flow chart

SIFT 算法具有相对较高的准确性和良好的尺度、旋转、光照等不变特性,鲁棒性较高,但是计算复杂,实时性相对较差,直接应用于距离测量效果不佳.通过实验测量30幅图片 SIFT 算法各个步骤的处理时间,整理统计结果见表1。

由表1可知,SIFT 算法的实现时间主要消耗在构建尺度空间和生成特征向量这两个步骤,耗时占总时间的近80%,所以从这两个步

表 1 SIFT 算法各步骤耗时统计表

Table 1 Time-consuming statistics of SIFT algorithm steps

算法步骤	所用时间 /s	占比 /%
构建尺度空间	1.84 ~ 3.92	30 ~ 50
特征点检测	0.85 ~ 1.12	12
确定方向参数	1.91 ~ 3.13	10
生成特征向量	0.18 ~ 0.62	30 ~ 50

骤对 SIFT 算法进行改进以缩短处理时间有较大的空间.

2.2 改进的 SIFT 算法

分析传统 SIFT 算法耗时长长的原因后,总结其主要存在的问题有两个方面:一是构建尺度空间具有一定的复杂性,构建的尺度空间层数越多,所消耗的时间就越多;二是 SIFT 描述子具有高维度(4 × 4 × 8 = 128 维),加大了算法空间复杂性与时间复杂度,并且数据量大,存储不易,严重影响了算法的实时性.

针对以上问题,对 SIFT 算法提出改进措施如下.

改进 1 本文用 2 × 2, 4 × 4, 6 × 6, 8 × 8, 10 × 10 的矩形积分表示 5 个不同的尺度空间,降低尺度空间的维度,然后将原图像降维采样,再用同样的方法构造第二组、第三组……构建积分金字塔完成后,同一组中两两图像作差,最终构建积分差分金字塔.

改进 2 采用 32 维特征向量代替 128 维特征向量.以关键点为中心,取 16 × 16 的窗口,在每个 2 × 2 的小块上,计算 8 个方向的梯度方向直方图,绘制每个梯度方向的累加值,即可形成一个种子点.这样对于一个种子点会产生 2 × 2 × 8 = 32 维的特征向量,即特征描述子.

改进 3 以街区距离代替欧式距离,作为最邻近特征点查询中的相似性度量,降低算法的运算难度,从而缩短匹配时间,提高算法的匹配效率.

欧式距离为

$$l = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

街区距离为

$$l_j = \sum_{i=1}^n |x_i - y_i|$$

式中, n 为特征向量维数.

由定义可知,在 128 维的情况下,计算欧式距离需要 128 次平方,128 次减法,127 次加法,1 次开方.降维之后采用街区距离计算,可大大降低运算量,提高数据处理速度,增强系统匹配的实时性.

3 实验结果与分析

3.1 传统的与改进的 SIFT 算法对比实验

为了验证本文改进 SIFT 算法的有效性,通过实验测定传统的 SIFT 算法与改进的 SIFT 算法的时间对比,结果见表 2,正确匹配率对比结果见表 3.

由表 2 可知,改进的 SIFT 算法的运算时间明显缩短,相较于传统算法时间缩短了 32%.

表 2 传统的与改进的 SIFT 算法运算时间对比表

Table 2 Comparison of operation time between the traditional and improved SIFT algorithm

算法步骤	构建尺度空间	特征点检测	确定方向参数	生成特征向量
传统 SIFT 算法	1.93	1.04	2.23	0.48
改进 SIFT 算法	1.12	0.63	1.76	0.35

表 3 传统的与改进的 SIFT 算法正确匹配率对比表

Table 3 Comparison table of correct matching rate between the traditional algorithm and improved SIFT algorithm

算法	实验组别	特征点数 / 个	正确匹配数 / 个	正确匹配率 / %
传统 SIFT 算法	1	355	275	77.5
	2	375	280	74.7
改进 SIFT 算法	1	300	274	91.3
	2	298	280	94.0

由表3可知,改进的SIFT算法在特征点个数降低不明显的情况下,缩短了匹配时间,提高了匹配效率,增强了实时性,可用于下一步双目视觉测距实验.

3.2 结合改进的SIFT算法的双目视觉测距实验

为了验证结合改进的SIFT算法的双目视觉测距方案的可行性,本文搭建了前后两台摄像机相距0.5 m的实验平台,拍摄多组照片,对不同距离的特征点进行检测匹配、计算距离,实验流程如图4所示.

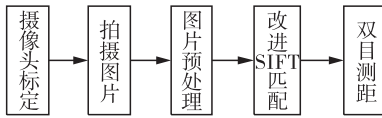


图4 结合改进的SIFT算法的双目视觉测距实验流程图

Fig.4 Experimental flow chart of binocular vision ranging combined with improved SIFT algorithm

摄像机标定作为信息采集处理的前端,标定效果直接影响后续信息的处理,摄像机模型是光学成像几何关系的简化,通常采用小孔成像模型研究摄像机标定问题.摄像机标定实验基于Matlab R2014a,采用camera Calibrator标定工具箱进行.搭建基线长度约为500 mm的沿光轴的双目视觉系统.通过推导像素坐标系、图像物理坐标系、相机坐标系和世界坐标系这四大坐标系之间的联系,得到像素坐标系与世界坐标系的关系:

$$Z_c \begin{bmatrix} u \\ v \\ l \end{bmatrix} =$$

$$\begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} =$$

$$\begin{bmatrix} a_x & 0 & u_0 & 0 \\ 0 & a_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} =$$

$$M_1 M_2 X_w = M X_w$$

据此,分别对前、后摄像头进行单目标定实验,求出摄像机的内参数和外参数,然后利用两台摄像机标定的.mat文件,导入双目标定程序,求出平移矩阵和旋转矩阵,完成标定.单目标定结果如下(M_1 为前内参矩阵, M_2 为后内参矩阵):

$$M_1 =$$

$$\begin{bmatrix} 1 & 046.325 & 7 & 0 & 489.461 & 9 & 0 \\ 0 & 1 & 055.164 & 1 & 640.886 & 2 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

$$M_2 =$$

$$\begin{bmatrix} 1 & 031.969 & 1 & 0 & 515.856 & 5 & 0 \\ 0 & 1 & 050.785 & 9 & 697.026 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

分别对双目视觉的两台摄像机完成单目标定后,利用不同的图片对应的外参矩阵完成对双目的视觉标定,可以获得两台摄像机对应的旋转和平移关系,即旋转矩阵 R 和平移变量 T :

$$R = \begin{bmatrix} 0.916 & 7 & 0.051 & 5 & 0.352 & 0 \\ 0.009 & 7 & 1.000 & 0 & 0.000 & 3 \\ 0.314 & 5 & 0.009 & 3 & 0.976 & 4 \end{bmatrix}$$

$$T = [-509.907 & 2 \quad -3.414 & 1 \quad 2.062 & 11]^T$$

通过标定结果可以看出,旋转矩阵 R 近似于单位矩阵,基本无旋转,符合所建立的双目立体视觉模型,从平移变量可以看出,两台摄像机的距离为509.907 2 mm,与实际距离500 mm相近,标定实验的成功可为后续摄像头拍摄图像的使用奠定基础.

由于无人机的工作环境比较特殊,光照、运动等容易在图片中产生噪声,在图片采集完毕后,需要对图片进行预处理,其原则是在尽可能

消除噪声的基础上保留更多的细节特征. 通过实验对比,本文采用小波法和独立分量分析法进行图像去噪. 根据输电线路贯穿图片、线径固定、平行直线的典型特征,可通过设定长度阈值、宽度阈值和平行线判定的方法将输电线路目标提取出来作为改进的 SIFT 算法匹配的基础.

将距离目标物近的摄像机与目标物的距离定义为所求距离,在 4 ~ 20 m 内以 1 m 为梯度间隔选取 17 组数据,测量结果见表 4.

表 4 实验结果统计表

Table 4 Statistical table of experiment results

实际距离/m	测量距离/m	误差/m	误差率/%
20	20.75	0.75	3.75
19	18.58	-0.42	2.21
18	18.59	0.59	3.28
17	17.55	0.55	3.25
16	16.30	0.30	1.90
15	14.67	-0.33	2.20
14	14.28	0.28	2.00
13	13.35	0.35	2.70
12	12.42	0.42	3.50
11	10.78	-0.22	2.00
10	10.29	0.29	2.95
9	9.21	0.21	2.32
8	7.87	-0.13	1.63
7	7.21	0.21	3.00
6	6.20	0.20	3.33
5	5.12	0.12	2.40
4	3.90	-0.10	2.50

由表 4 可知,测量结果最大的误差率为 3.75%,同等实验条件下的单目视觉测距的误差率为 7.25%,降低了 48.3%,可满足实际设备的精度要求. 表 4 中误差产生的原因可能是: 1) 摄像机的位置有偏差,摄像机的安装位置没有完全按照几何模型设置,造成结果不准确; 2) 成像设备分辨率低,导致空间分辨率低; 3) 无人机运动速度较快,运算的延时造成误差.

4 结语

针对传统测距方法应用于无人机自动巡线效果不理想的情况,本文提出了一种结合改进的 SIFT 算法的类单目的双目视觉测距方案,该方案创新性地采用两台摄像设备前后排列的方式,减少计算过程中的变量;采用缩小尺度空间个数、降低特征向量维度、街区距离代替欧式距离的方法改进 SIFT 算法,提高运算效率. 通过计算推导,证明了该方案的科学性;通过实验,验证了该方案的有效性,匹配时间缩短了 32%,测量最大误差率降至 3.75%,较单目视觉测距降低了 48.3%. 当然,目前将图像测距技术运用于无人机的巡线还处于探索阶段,在后续的工作中需要建立更加精确的模型,换用更高分辨率的摄像设备,以及继续改进算法以进一步缩短运算时间.

参考文献:

- [1] 陈奕君. 基于多传感器信息融合的机器人避障研究[D]. 大庆:东北石油大学,2004.
- [2] 张天,杨晨曦,朱颖,等. 多传感器信息融合在四足机器人避障中的应用[J]. 传感器与微系统,2015,34(5):150.
- [3] 汪启跃,王中字. 基于单目视觉的航天器位姿测量[J]. 应用光学,2017,38(2):250.
- [4] 陈爱鸾. 基于 SIFT 算子的双目视觉三维测量技术研究[D]. 广州:广东工业大学,2015.
- [5] 余彩云,胡贤洪,唐坤,等. 一种多目视觉三维测距系统设计[J]. 制导与引信,2015,36(9):20.
- [6] LOWE D G. Distinctive image features from scale-invariant key points [J]. International Journal of Computer Vision, 2004,60(2):91.
- [7] BAY H, ESS A, TUYTELAARS T, et al. Speeded-up robust features (SURF) [J]. Compute VisImage Underst,2008,110:346.

- [8] ALCANTARILLA P F, BARTOLI A, DAVISON A J. KAZE Features[C]//European Conference on Computer Vision, Berlin Heidelberg: Springer,2012:214.
- [9] ENDRES F,HESS J.3D mapping with an RGB-D camera [J]. IEEE Transactions on Robotics, 2014,30(1):177.
- [10] 宋佳乾,江西原.基于改进 SIFT 特征点匹配的图像拼接算法[J].计算机测量与控制,2015,23(2):512.
- [11] 刘佳,傅卫平,王雯,等.基于改进 SIFT 算法的图像匹配[J].仪器仪表学报,2013,34(5):1107.
- [12] 李奇敏,李扬宇.基于改进 SIFT 算法的双目视觉距离测量[J].传感器与微系统,2017,36(11):119.
- [13] 高宏伟,王辉,刘玉影,等.基于图像测距的汽车防撞系统关键技术研究[J].电子测量与仪器学报,2017,31(6):820.
- [14] 栾长一.基于定焦平移序列图像的测距技术研究[D].沈阳:沈阳理工大学,2013.

本刊数字网络传播声明

本刊已许可中国知网、万方数据资源系统、维普网、博看网、超星、中国科技论文在线、中教数据库等在其系列数据库产品中以数字化方式复制、汇编、发行、信息网络传播本刊全文。其相关著作权使用费与本刊稿酬一并支付。作者向本刊提交文章发表的行为即视为同意本刊上述声明。