

doi: 10.3788/gzxb20164510.1012003

基于单目视觉的静止目标定位方法

邵校, 陶建武

(空军航空大学 飞行器控制系, 长春 130022)

摘 要:以无人机感知与避障为背景,提出了基于运动的单目视觉测距方法.选择摄像机运动过程中在不同位置对同一目标获取的两幅图像,然后利用尺度不变特征变换算法对所选图像进行特征检测和匹配,通过分析同一目标特征点在两幅图像中不同成像位置的变化,结合无人飞行器自身的运动参数,求解出无人飞行器与障碍目标之间的位置信息.采用该方法对不同位置的目标进行定位实验,结果表明该算法的测量精度和时效性可以满足实际避障的要求.

关键词:单目视觉;静态障碍;目标定位

中图分类号:TP242

文献标识码:A

文章编号:1004-4213(2016)10-1012003-8

Location Method of Static Object Based on Monocular Vision

SHAO Xiao, TAO Jian-wu

(Department of Aircraft Control, Aviation University of Air Force, Changchun 130022, China)

Abstract: Aiming at the background of sense and avoiding of unmanned aerial vehicle, a monocular vision method for distance measurement from motion was proposed. Firstly, two different images were chose from a sequence of images for the same target in different position. Then, feature points could be extracted and matched through scale-invariant feature transform algorithm for the two images. Through analysis the different positions for the same feature point on different images and combined with the movement parameters of the aircraft, the information between the aircraft and the obstacle could be got. With the proposed method, the target was located at different locations. The experimental results show that the measurement accuracy and effectiveness of the method can meet the demand of avoidance.

Key words: Monocular vision; Static obstacle; Target localization

OCIS Codes: 120.1680; 150.5670; 150.6910

0 引言

无人机(Unmanned Aerial Vehicle, UAV)具有宽广的搜索和观察范围,可从空间区域获取大量信息,已被广泛应用于军事侦察、城市交通监测、森林火警探测、环境监测与资源勘探等领域.由于大部分无人机的飞行区域是低空,可能存在未知的静止障碍物.为了保证飞行安全,不与障碍物碰撞,无人机感知方法的研究至关重要且具有挑战性^[1-3].

根据传感器的不同,无人机感知方法也有所不同,其中单目视觉感知方法有独特优势.单目视觉感知方法是利用一个摄像机拍摄的图像得到无人机与障碍物

之间的信息,如距离、深度等.该方法具有结构简单、测量实时性好、成本低等优点,因而得到国内外许多学者的关注. Saxena^[4-5]提出了一种改进的马尔科夫随机场(Markov Random Field, MRF)算法用来获取图像中不同点深度之间的关系,但是基于MRF的方法忽略了目标的几何信息,测量结果受目标颜色变化的影响较大. Hoiem^[6]等利用单幅图像重构出了目标的三维信息,然而无法获取目标的绝对深度信息. Hedau^[7]等利用图像中几何结构的相对深度信息恢复出室内空间的场景. Lin^[8]等提出的离焦(Shape From Defocus, SFD)方法利用单张离焦图像估计深度信息,但是该方法无法区分模糊纹理和离焦突变纹理.张法全^[9]等提出了

基金项目:国家自然科学基金(No. 61571462)和吉林省自然科学基金(No. 20140101073JC)资助

第一作者:邵校(1992-),男,硕士研究生,主要研究方向为光电检测定位技术. Email:18844079474@163.com

导师:陶建武(1959-),男,教授,博士,主要研究方向为模式识别. Email:jianwu.tao@gmail.com

收稿日期:2016-04-20;录用日期:2016-06-22

<http://www.photon.ac.cn>

利用图像处理检测图像模糊度,从而判断目标距离,实现测距目的的测量方法.韩延祥^[10]等提出了一种基于特征点的单目视觉测距方法.赵天云^[11]等提出了利用特殊标志点的空间几何信息实现光学相机与标志点相对位姿测量,从而推算出相机的空间位置的方法.但是,上述基于单目视觉的感知方法仅能从单张图像中获取目标或场景的相对信息,并需要已知目标或场景的先验信息,在实际应用中受到一定限制.

本文以无人机感知与规避为背景,提出了基于运动的单目视觉测距方法,基于摄像机的透视成像原理,利用无人机在不同位置获取的同一目标的两幅图像,结合无人机自身的参数信息测量获取无人机飞行过程中遇到障碍物的深度信息,避免发生碰撞.该方法不需要目标的任何先验信息,能够在未知环境下应用.

1 测量原理

在无人飞行器的前端安装一个微型摄像机,其原理为小孔成像,模型如图 1.图中, $O-X_c Y_c Z_c$ 为相机坐标系, O 为投影中心.假设摄像机的光轴与 Z_c 轴重合,与无人飞行器飞行方向平行.长方形表示相机的成像面,它与相机坐标系的 $X_c O Y_c$ 平面平行,其中心为 $O'(0,0,-f)$, f 为相机的焦距. A 表示摄像机前方场景中任意点目标, A' 表示目标 A 在成像面上的投影点(像点),它在相机坐标系中的坐标为 $(x_A, y_A, -f)$.

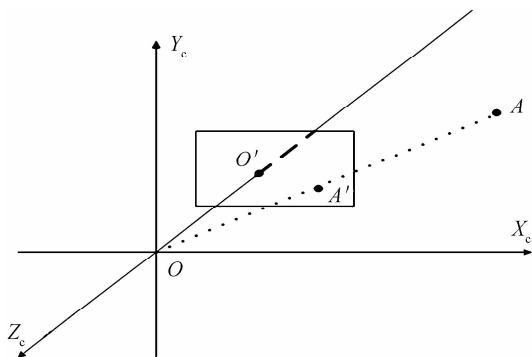


图 1 小孔成像模型
Fig. 1 Pinhole camera model

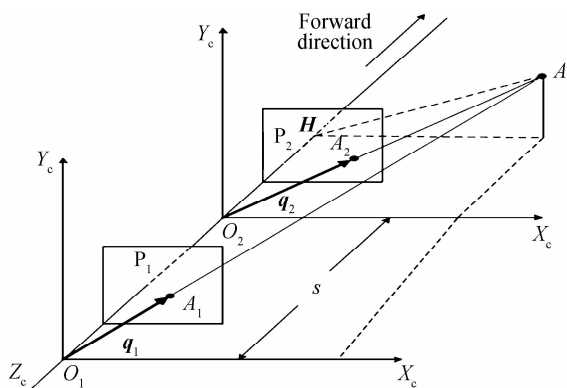


图 2 测量模型
Fig. 2 Measurement model

为测量无人飞行器前端与目标 A 之间的距离,建立测量模型如图 2.拍摄第一张图像时,摄像机的投影中心在 O_1 点, $A_1(x_1, y_1, -f)$ 为目标 A 的像点.当无人飞行器向前飞行时,相机坐标系也随之移动.当向前移动 s 距离时,摄像机拍摄第二张图像.此时摄像机的投影中心在 O_2 点, $A_2(x_2, y_2, -f)$ 为目标 A 此时的像点.在图 2 中,投影面 P_1 表示第一张图像,投影面 P_2 表示第二张图像.设 q_1 表示在相机坐标系 $O_1-X_c Y_c Z_c$ 中,坐标原点 O_1 指向像点 A_1 的向量,即 $q_1 = \overrightarrow{O_1 A_1}$,而向量 Q_1 表示坐标原点 O_1 指向目标 A 的向量,即 $Q_1 = \overrightarrow{O_1 A}$.由摄像机的透视成像原理可知,空间目标 A 点在向量 q_1 的延长线上.因此,向量 Q_1 可写为

$$Q_1 = \lambda_1 q_1 \tag{1}$$

式中, λ_1 为比例因子.同理,在 $O_2-X_c Y_c Z_c$ 相机坐标系下,空间目标 A 可表示为

$$Q_2 = \lambda_2 q_2 \tag{2}$$

式中, $q_2 = \overrightarrow{O_2 A_2}$, $Q_2 = \overrightarrow{O_2 A}$, λ_2 为比例因子.

由式(1)和(2)可知,在 $O_1-X_c Y_c Z_c$ 和 $O_2-X_c Y_c Z_c$ 相机坐标系下,目标 A 的坐标值也与像点 $A_1(x_1, y_1, -f)$ 和 $A_2(x_2, y_2, -f)$ 的坐标值成正比,即

$$\begin{cases} X_1 = \lambda_1 x_1 \\ Y_1 = \lambda_1 y_1 \\ Z_1 = -\lambda_1 f \\ X_2 = \lambda_2 x_2 \\ Y_2 = \lambda_2 y_2 \\ Z_2 = -\lambda_2 f \end{cases} \tag{3}$$

式中,坐标值 (X_1, Y_1, Z_1) 和 (X_2, Y_2, Z_2) 分别表示目标 A 在 $O_1-X_c Y_c Z_c$ 和 $O_2-X_c Y_c Z_c$ 相机坐标系下的坐标值.由图 2 可知,坐标值 (X_1, Y_1, Z_1) 和 (X_2, Y_2, Z_2) 的关系为

$$\begin{cases} X_1 = X_2 \\ Y_1 = Y_2 \\ -Z_1 = -Z_2 + s \end{cases} \tag{4}$$

将式(3)代入式(4),可得

$$\begin{cases} \lambda_1 x_1 = \lambda_2 x_2 \\ \lambda_1 y_1 = \lambda_2 y_2 \\ \lambda_1 f = \lambda_2 f + s \end{cases} \tag{5}$$

整理式(5),可得

$$\begin{cases} \frac{\lambda_1}{\lambda_2} = \frac{x_1}{x_2} = \frac{y_1}{y_2} \\ \lambda_1 - \lambda_2 = \frac{s}{f} \end{cases} \tag{6}$$

从式(6),可以求得比例因子 λ_1 和 λ_2 为

$$\begin{cases} \lambda_1 = \frac{s x_2}{f(x_2 - x_1)} = \frac{s y_2}{f(y_2 - y_1)} \\ \lambda_2 = \frac{s x_1}{f(x_2 - x_1)} = \frac{s y_1}{f(y_2 - y_1)} \end{cases} \tag{7}$$

将式(7)代入式(3)中,可求得目标 A 在 O_2 -XYZ 相机坐标系下的坐标值为

$$A\left(\frac{sx_1x_2}{f(x_2-x_1)}, \frac{sy_1y_2}{f(y_2-y_1)}, \frac{sx_1}{x_2-x_1}\right) \text{ 或}$$

$$A\left(\frac{sy_1x_2}{f(y_2-y_1)}, \frac{sx_1y_2}{f(x_2-x_1)}, \frac{sy_1}{y_2-y_1}\right).$$

由图 2 可得到测量模型在 XOZ 平面的投影示意图,如图 3,这里向量 $\overrightarrow{O_1A_x}$ 是向量 \mathbf{Q}_1 在 XOZ 平面的投影,向量 $\overrightarrow{O_2A_x}$ 是向量 \mathbf{Q}_2 在 XOZ 平面的投影.由图 3 可知,由于向量 $\overrightarrow{O_2A_x}$ 的斜率大于向量 $\overrightarrow{O_1A_x}$ 的斜率,故 $x_2 > x_1$.同理,图 4 为测量模型在 YOZ 平面的投影示意图.由图 4 可以看出 $y_2 > y_1$.因此,可以得到 $x_2 - x_1 \neq 0$ 和 $y_2 - y_1 \neq 0$.

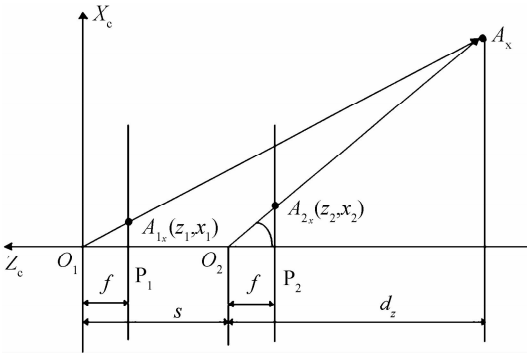


图 3 测量模型在 XOZ 平面的投影示意图

Fig. 3 Projection of measurement model on XOZ plane

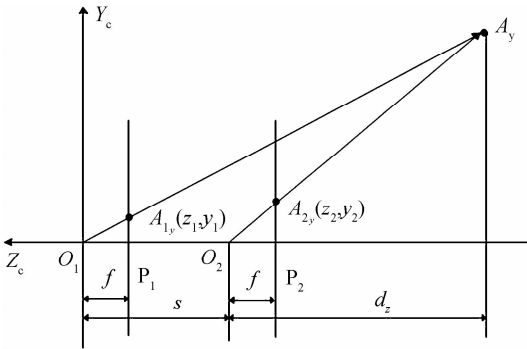


图 4 测量模型在 YOZ 平面的投影示意图

Fig. 4 Projection of measurement model on YOZ plane

由于无人机此时位于 O_2 的位置,无人机与目标 A 之间的距离可表示为 $d = \sqrt{X_2^2 + Y_2^2 + Z_2^2}$.而目标 A 在 O_2 - $X_cY_cZ_c$ 相机坐标系下的 Z 坐标值,则被定义为无人机距目标 A 的深度,可表示为

$$d_z = Z_2 = \frac{sx_1}{x_2 - x_1} = \frac{sy_1}{y_2 - y_1} \quad (8)$$

向量 $\overrightarrow{O_2A_x}$ 与 Z 轴负方向的夹角定义为目标 A 的水平方位角 α ,可表示为 $\tan\alpha = \frac{X}{Z} = \frac{x_2}{f}$.向量 $\overrightarrow{O_2A_y}$ 与 Y 轴正方向的夹角定义为目标 A 的垂直俯仰角 β ,可表示为

$$\tan\beta = \frac{\sqrt{X^2 + Z^2}}{Y} = \frac{\sqrt{f^2 + x_2^2}}{y_2}. \text{ 因此,当像点 } A_1(x_1,$$

$y_1, -f)$ 、 $A_2(x_2, y_2, -f)$ 和无人机移动距离 s 被确定后,就可完全确定无人机与目标 A 的相对位置关系,而无需知道目标或场景的任何先验信息,也无需摄像机标定出旋转矩阵 \mathbf{R} 及平移矩阵 \mathbf{t} .

2 像点的检测与匹配

像点 A_1 和 A_2 是同一目标 A 在前后两张图像上的投影.因此,首先需要从第二张图像中提取出像点 A_2 ,然后在第一张图像中,寻找目标 A 的投影点 A_1 .这个过程通常被称为像点(特征点)检测与匹配.两个像点的匹配精度会影响目标深度的测量精度.目前常用的两种特征点检测与匹配算法是 Harris 算法^[12]和尺度不变特征变换算法(Scale Invariant Feature Transform, SIFT)^[13].Harris 算法的核心思想是拟定模板沿着图像上各个方向移动时,根据图像上灰度值的变化确定该点是否为特征点.对于图像存在平移、旋转时,Harris 算法的性能较好,且具有旋转不变性.但是,Harris 算法对尺度变化很敏感,并不具备尺度不变性,受噪声影响较大,随着图像分辨率的不同,特征点容易产生漂移.由于无人机在运动过程中,所拍摄的目标在两张图片中的尺度、光照、角度等会发生明显变化,并且相机抖动容易导致所拍摄的图像存在大量的噪声,故 Harris 算法不适合这种情况.本文采用 SIFT 算法对所拍摄获取的图片进行特征检测与匹配.该算法对尺度缩放、视角、旋转、光照等变化具有良好的鲁棒性,且抗噪声能力比较强,适于同一目标在不同位置对图像的处理.

SIFT 算法检测与匹配的流程为:

1) 尺度空间极值检测

首先利用高斯卷积核 $G(x, y, \sigma)$ 与原始图像 $I(x, y)$ 进行卷积,可得到在原始图像尺度空间下的高斯图像函数,可表示为

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (9)$$

式中, $*$ 表示图像 x 和 y 两个方向上进行卷积,高斯卷积核 $G(x, y, \sigma)$ 可表示为

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (10)$$

通过构造高斯差分(Difference of Gaussian, DOG)算子,可得到高斯差分图像函数,即

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (11)$$

式中, σ 表示尺度值,其大小决定了图像的平滑程度. σ 比较大时,对应图像的总体特征;而 σ 比较小时,则体现图像的细节特征. k 为常数.

从高斯差分图像函数中寻找尺度空间的极值点时,需要将图像中每一尺度的检测点与同一尺度周围的 8 个相邻点以及纵向相邻尺度对应位置的 9×2 个

点(共 26 个点)进行比较,以确保同时在尺度空间检测局部极值点。

2) 去除不稳定的极值点

由于 DOG 算子边缘响应比较强,可通过拟合三维二次函数去除不稳定的关键点,实现稳定关键点的精确定位,同时增强匹配的稳定性和抗噪性。

首先,采用泰勒级数分解方法,去除对比度低的关键点。函数 $D(x, y, \sigma)$ 在极值点 (x, y, σ) 处的泰勒展开式可表示为

$$D(x) = D + \frac{\partial D^T}{\partial \mathbf{x}} \mathbf{x} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2} \mathbf{x} \quad (12)$$

式中, $\mathbf{x} = (x, y, \sigma)^T$, 对式(12)求导,并令其为 0, 得到极值点 $\hat{\mathbf{x}} = -\frac{\partial^2 D^{-1}}{\partial \mathbf{x}^2} \frac{\partial D}{\partial \mathbf{x}}$ 。将极值点代入式(12)得

$$D(\hat{\mathbf{x}}) = D + \frac{1}{2} \frac{\partial D^T}{\partial \mathbf{x}} \hat{\mathbf{x}} \quad (13)$$

判断 $|D(\hat{\mathbf{x}})|$ 是否小于给定阈值,若小于给定的阈值即为响应较弱的不稳定点,将其丢弃,反之则保留下来。

另外,由于在第一步检测极值点中采用了 DOG 算法,这导致对图像的边缘十分敏感。即许多在边缘附近的像素点被检测为极值点,也包含了一些对目标特征表述能力较差且对噪声敏感的点。通常,这类不明确的极值点在边缘交叉处具有较大的主曲率。根据该性质可以去除边缘响应。函数 $D(x, y, \sigma)$ 的主曲率可以通过 2×2 维的 Hessian 矩阵 \mathbf{H} 计算得到,即

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (14)$$

式中, D_{xx}, D_{xy}, D_{yy} 是函数 $D(x, y, \sigma)$ 的二阶偏导数。矩阵 \mathbf{H} 的特征值与函数 $D(x, y, \sigma)$ 的主曲率是成比例的,设 a 是 \mathbf{H} 的较大特征值, b 是 \mathbf{H} 得较小特征值。则 Hessian 矩阵的迹和行列式的值可分别表示为

$$\begin{aligned} \text{Tr}(\mathbf{H}) &= D_{xx} + D_{yy} = a + b \\ \text{Det}(\mathbf{H}) &= D_{xx} D_{yy} - (D_{xy})^2 = ab \end{aligned} \quad (15)$$

令 r 为两个特征值的比率,即 $a = rb$, 则可得

$$\frac{\text{Tr}(\mathbf{H})^2}{\text{Det}(\mathbf{H})} = \frac{(a+b)^2}{ab} = \frac{(rb+b)^2}{rb^2} = \frac{(r+1)^2}{r} \quad (16)$$

由式(16)可以看出,比值只与两个特征值的比率 r 有关,当两个特征值相等时, $\frac{(r+1)^2}{r}$ 值最小,随着 r 的增大而 $\frac{(r+1)^2}{r}$ 增大。所以,要检测两个特征值的比率小

于某一阈值 R , 只需判断 $\frac{\text{Tr}(\mathbf{H})^2}{\text{Det}(\mathbf{H})} < \frac{(R+1)^2}{R}$, 若小于则保留,反之则剔除。

3) 关键点主方向的分配

对于每一个采样点 $L(x, y)$, 计算其梯度模值 $m(x, y)$ 以及方向 $\theta(x, y)$ 的公式为

$$\begin{cases} m(x, y) = \{ [L(x+1, y) - L(x-1, y)]^2 + \\ \quad [L(x, y+1) - L(x, y-1)]^2 \}^{1/2} \\ \theta(x, y) = \arctan [L(x, y+1) - L(x, y-1)] / \\ \quad [L(x+1, y) - L(x-1, y)] \end{cases} \quad (17)$$

实际计算时,在以关键点为中心的邻域窗口内采样,利用直方图统计邻域像素的梯度方向,梯度直方图的范围是 $0^\circ \sim 360^\circ$, 一个柱代表 10° , 共有 36 个柱。峰值位置则代表了该关键点的主方向,作为该关键点的方向。当存在一个峰值,其能量相当于主峰值的 80% 时,则认为该方向是关键点的辅方向。所以一个关键点可能具有多个方向,这可以增强匹配的鲁棒性。

4) 高维 SIFT 特征向量的生成

首先将坐标轴旋转到关键点的主方向,以确保特征向量的旋转不变性。然后以特征点为中心取 16×16 个窗口,将其分为 4×4 个子区域。每个区域选取 8 个角度方向来计算每个小区域梯度方向的直方图,得到 128 维的 SIFT 特征向量。

5) 特征点匹配

两幅图像生成 SIFT 特征向量后,采用欧氏距离作为两幅图像中关键点的相似性判定度量。求出某一特征点与待匹配图像中所有特征点之间的欧式距离,取出距离最小的那一对特征点。其中欧式距离的计算公式为

$$d(R_i, S_i) = \sqrt{\sum_{j=1}^{128} (r^{ij} - s^{ij})^2} \quad (18)$$

式中, R_i 和 S_i 分别表示两幅待匹配图像的特征点集合, $R_i = (r_{i1}, r_{i2}, \dots, r_{i128})$, $S_i = (s_{i1}, s_{i2}, \dots, s_{i128})$ 。当最近邻点 $d(R_1, S_1)$ 与次近邻点 $d(R_2, S_2)$ 的比值小于某个阈值时,就认为这两个特征点匹配成功。根据应用的实际情况,将该阈值调为 0.1 (Low 实验中采用的值为 0.4)。这样既可以增加匹配的精度,又能够减小匹配点的数量进而提高算法的测距精度和实时性。

3 任意形状目标的测量

在实际应用中,障碍物通常为任意形状目标。由于任意形状目标具有一定大小形状,因此测量时需要选取一些特征点来标记目标的形状。假设前后两张图像 M 和 N 中的任意形状目标 B 经 SIFT 特征检测后的特征点集分别为: $S_M = \{s(m)_1, s(m)_2, \dots, s(m)_i\}$, $S_N = \{s(n)_1, s(n)_2, \dots, s(n)_j\}$, 其中 i 和 j 分别为 B 在图像 M 和 N 中的特征点数目。然后对集合 S_M 和 S_N 进行特征匹配,假设匹配成功的特征点集合为 $S_K = \{s(k)_1, s(k)_2, \dots, s(k)_l\}$, 其中 $s(k)_l$ 表示匹配成功的特征点对, S_K 中包含 l 对特征点。将 S_K 中每对特征点的像点坐标代入目标 A 在 O_2 -XYZ 相机坐标系下的坐标值,即可求得该特征点在相机坐标系下的坐标值,将求得的所有特征点的 X、Y、Z 坐标值分别排序,让其最大值和

最小值的连线作为 X 、 Y 、 Z 坐标轴的边界,从而形成一个矩形立方体.那么, B 就被包涵在这个矩形立方体内,如图 5.最后将这个矩形立方体中具有最小 Z 坐标值的 XY 平面看作等效障碍物平面.因此,这个最小 Z 坐标值就是无人飞行器前端与此平面之间的深度 d_z ,进而可计算此平面四个顶点与飞行器的距离 d 、方位角 α 以及俯仰角 β .此平面(障碍物)是否位于飞行器前进轨迹上,可由此平面四个顶点的坐标值来判定.如果此平面(障碍物)位于飞行器前进轨迹上,飞行器将发生碰撞.

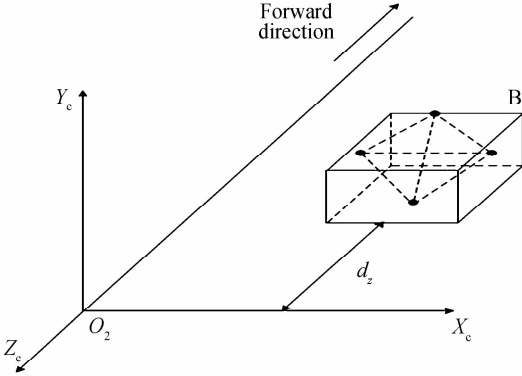


图 5 任意形状目标的测量示意图

Fig. 5 Measurement schematic of arbitrary shape target

测量的计算流程为:

- 1) 从摄像机中选取相隔 Δt 的两张图片;
- 2) 利用 SIFT 算法对选取的两张图片进行特征检测和匹配;
- 3) 获取匹配成功的特征点的坐标 (x_{1i}, y_{1i}) 和 (x_{2i}, y_{2i}) ;
- 4) 将特征点的坐标代入 A 在 O_2 - XYZ 相机坐标系下的坐标值,计算对应相机坐标下的特征点坐标 (X_i, Y_i, Z_i) ;
- 5) 将所有特征点的坐标值分别排序,计算等效障碍物平面的坐标值;
- 6) 利用等效障碍物平面的坐标值,计算飞行器与障碍物的相对距离及角度信息,进而判断飞行器与障碍物是否将发生碰撞.

4 测量误差分析

测量过程是假设飞行方向与光轴方向一致,所计算的距离实际上是障碍物在光轴方向上距离摄像机的距离.但是在实际情况下,由于安装误差等因素,会造成飞行方向与光轴方向产生偏差,如图 6.

假设飞行方向与光轴方向在 X 轴方向上偏离的角度为 φ ,在 Y 轴方向上偏离的角度为 θ .那么此时机体坐标系(飞行方向与光轴方向一致) $[X, Y, Z]^T$ 与相机坐标系(飞行方向与光轴方向有偏差) $[X_c, Y_c, Z_c]^T$ 的关系可表示为

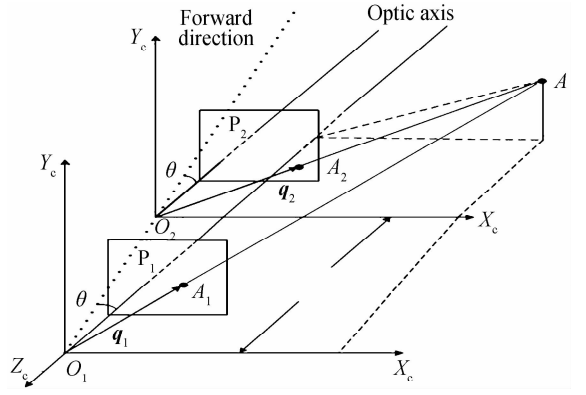


图 6 光轴与前进方向不一致时的测距模型

Fig. 6 Measurement model of the optic axis different from the forward direction

$$[X_c, Y_c, Z_c]^T = \mathbf{M}[X, Y, Z]^T \quad (19)$$

其中旋转矩阵为

$$\mathbf{M} = \begin{bmatrix} \cos\theta & -\sin\theta\sin\varphi & \sin\theta\cos\varphi \\ 0 & \cos\varphi & \sin\varphi \\ -\sin\theta & -\cos\theta\sin\varphi & \cos\theta\cos\varphi \end{bmatrix} \quad (20)$$

当飞行器前进方向与光轴平行时,旋转矩阵 \mathbf{M} 为单位阵.因此,飞行器在 O_1 和 O_2 位置时,目标 A 在相机坐标系下的坐标分别为 $A^1(X_1, Y_1, Z_1)$ 和 $A^2(X_2, Y_2, Z_2)$.当飞行器由 O_1 位置运动到 O_2 位置时,相当于坐标系在 Z 坐标轴方向平移了距离 s .因此,坐标 A^1 与 A^2 满足 $[X_2, Y_2, Z_2]^T = [X_1, Y_1, Z_1]^T + \mathbf{T}$,其中平移矩阵 $\mathbf{T} = [0, 0, s]^T$.

当飞行器前进方向与光轴存在偏差时,假设飞行器在 O_1 和 O_2 位置时刻,目标 A 在相机坐标系下的坐标分别为 $A^{1'}(X_{c1}, Y_{c1}, Z_{c1})$ 和 $A^{2'}(X_{c2}, Y_{c2}, Z_{c2})$,其在投影面 P_1 和 P_2 上像点坐标分别为 $A'_1(x'_1, y'_1, -f)$ 和 $A'_2(x'_2, y'_2, -f)$.那么飞行器在 O_1 位置时,坐标 $A^{1'}$ 可由 A^1 经旋转矩阵 \mathbf{M} 得到,即

$$\begin{cases} X_{c1} = X_1 \cos\theta - Y_1 \sin\theta\sin\varphi + Z_1 \sin\theta\cos\varphi \\ Y_{c1} = Y_1 \cos\varphi + Z_1 \sin\varphi \\ Z_{c1} = -X_1 \sin\theta - Y_1 \cos\theta\sin\varphi + Z_1 \cos\theta\cos\varphi \end{cases} \quad (21)$$

同理,坐标 $A^{2'}$ 也可表示为

$$\begin{cases} X_{c2} = X_2 \cos\theta - Y_2 \sin\theta\sin\varphi + Z_2 \sin\theta\cos\varphi \\ Y_{c2} = Y_2 \cos\varphi + Z_2 \sin\varphi \\ Z_{c2} = -X_2 \sin\theta - Y_2 \cos\theta\sin\varphi + Z_2 \cos\theta\cos\varphi \end{cases} \quad (22)$$

根据 A^1 与 A^2 的关系,可得 $A^{1'}$ 与 $A^{2'}$ 的关系为

$$\begin{cases} X_{c2} = X_{c1} + s\sin\theta\cos\varphi \\ Y_{c2} = Y_{c1} + s\sin\varphi \\ Z_{c2} = Z_{c1} + s\cos\theta\cos\varphi \end{cases} \quad (23)$$

由于目标 A 的坐标值与投影面上像点的坐标值成正比,式(23)可写成

$$\begin{cases} \lambda'_2 x'_2 = \lambda'_1 x'_1 + s\sin\theta\cos\varphi \\ \lambda'_2 y'_2 = \lambda'_1 y'_1 + s\sin\varphi \\ -\lambda'_2 f = -\lambda'_1 f + s\cos\theta\cos\varphi \end{cases} \quad (24)$$

式中, λ'_1 和 λ'_2 为比例因子. 从式(24)可解得

$$\begin{cases} \lambda'_1 = \frac{x'_2 \cdot s \cos \theta \cos \varphi + f s \sin \theta \cos \varphi}{f(x'_2 - x'_1)} = \\ \frac{y'_2 \cdot s \cos \theta \cos \varphi + f s \sin \theta}{f(y'_2 - y'_1)} \\ \lambda'_2 = \frac{x'_1 \cdot s \cos \theta \cos \varphi + f s \sin \theta \cos \varphi}{f(x'_2 - x'_1)} = \\ \frac{y'_1 \cdot s \cos \theta \cos \varphi + f s \sin \theta}{f(y'_2 - y'_1)} \end{cases} \quad (25)$$

进而可求得 $A^z(X_{c2}, Y_{c2}, Z_{c2})$ 为

$$A^z \left(\frac{x'_1 \cdot s \cos \theta \cos \varphi + f s \sin \theta \cos \varphi}{f(x'_2 - x'_1)} x'_2, \right. \\ \left. \frac{x'_1 \cdot s \cos \theta \cos \varphi + f s \sin \theta \cos \varphi}{f(x'_2 - x'_1)} y'_2, \right. \\ \left. - \frac{x'_1 \cdot s \cos \theta \cos \varphi + f s \sin \theta \cos \varphi}{(x'_2 - x'_1)} \right)$$

或

$$A^z \left(\frac{y'_1 \cdot s \cos \theta \cos \varphi + f s \sin \theta}{f(y'_2 - y'_1)} x'_2, \right. \\ \left. \frac{y'_1 \cdot s \cos \theta \cos \varphi + f s \sin \theta}{f(y'_2 - y'_1)} y'_2, \right. \\ \left. - \frac{y'_1 \cdot s \cos \theta \cos \varphi + f s \sin \theta}{(y'_2 - y'_1)} \right).$$

从与第2小节类似的分析可知, $x'_2 - x'_1 \neq 0$ 和 $y'_2 - y'_1 \neq 0$. 此时无人飞行器与目标 A 之间的距离 d' 、深度 d'_z 等信息均可由式 $A^z(X_{c2}, Y_{c2}, Z_{c2})$ 求得.

在实际应用中, 偏离角 φ 和 θ 通常是未知的, 因此, 测量结果会产生一定的误差. 经推导, 距离的相对误差为

$$e_d = \frac{d - d'}{d} = \frac{\Delta \lambda_2}{\lambda_2} = 1 - \cos \theta \cos \varphi - \frac{f}{x_1} \sin \theta \cos \varphi \quad (26)$$

而深度的相对误差为

$$e_{d_z} = \frac{d_z - d'_z}{d_z} = \\ \frac{\lambda_2 f - \lambda'_2 (-x_2 \sin \theta \cos \varphi - y_2 \sin \theta + f \cos \theta \cos \varphi)}{\lambda_2 f} = \\ 1 + \frac{1}{2} \frac{x_2}{f} \sin 2\theta \cos^2 \varphi + \frac{1}{2} \frac{y_2}{f} \cos \theta \sin 2\varphi - \cos^2 \theta \cos^2 \varphi + \\ \frac{x_2}{x_1} \sin^2 \theta \cos^2 \varphi + \frac{1}{2} \frac{y_2}{x_1} \sin \theta \sin 2\varphi - \frac{1}{2} \frac{f}{x_1} \sin 2\theta \cos^2 \varphi \quad (27)$$

从式(26)和(27)可以看出, 相对误差与焦距 f , 偏差角 θ 和 φ 以及特征点的位置有关. 通常采用静态校正方法来减小偏差角 θ 和 φ , 即利用安装后的摄像机对三维位置信息已知的固定目标进行定位测量, 将得到的三维位置信息的测量值与实际值进行比较, 调整摄像机的安装位置, 直至测量值与实际值的误差小于一定的误差.

5 测量实验

5.1 实验系统

实验设备如图7, 六旋翼无人飞行器, 搭载 GoPro

HERO4 的运动摄像机, 结合镜头后焦距为 30 mm, 摄像机的 fps 设为 25, 即每秒拍摄 25 帧图像. 实验环境如图8, 障碍物为前方静止的建筑物, 选取无人飞行器与障碍目标的距离为 250 m、200 m、150 m、100 m、50 m 进行测量.



图7 实验设备

Fig. 7 Experimental equipment



图8 实验环境

Fig. 8 Experimental environment

5.2 实验与分析

实验1参数设置为: 飞行速度 $v=10\text{m/s}$, 采样时间 $\Delta t=0.4\text{s}$, 则移动距离 $s=v\Delta t=4\text{m}$. 测量数据和实际数据对比见表1.

表1 实验1的测量结果

Table 1 Measurement results of experiment 1

Distance to camera/m	Relative error of α	Relative error of β	Relative error of d_z	Relative error of d
50	0.61%	0.69%	1%	1.41%
100	1.34%	1.36%	1.3%	1.83%
150	1.82%	1.8%	1.6%	2.53%
200	2.32%	2.31%	2.05%	3.14%
250	2.78%	2.67%	2.48%	3.68%

从表1数据不难发现, 目标与飞行器之间的距离越小, 方位角、俯仰角、深度以及距离等信息的测量精度越高, 其主要原因是距离越远所获取目标的清晰度越差, 对目标进行特征点检测与匹配时得到的特征点质量越差, 导致测量精度下降. 另一方面, 实际距离目标最近或最远的目标点未必是满足条件的特征点, 导致所测量的特征点的相对位置信息与实际特征点不一致.

实验2参数设置为: 飞行速度 $v=10\text{m/s}$, 改变采

样时间 $\Delta t=1\text{ s}$, 移动距离 $s=v\Delta t=10\text{ m}$. 此时的测量数据和实际数据对比见表 2.

表 2 实验 2 的测量结果

Table 2 Measurement results of experiment 2

Distance to camera/m	Relative error of α	Relative error of β	Relative error of d_z	Relative error of d
50	0.52%	0.57%	0.8%	1.19%
100	0.39%	1.21%	1.1%	1.75%
150	1.35%	1.6%	1.47%	2.31%
200	1.79%	2.05%	1.95%	2.89%
250	2.22%	2.33%	2.32%	3.27%

对比表 1 和表 2, 可以看出当其他参数相同, 而采样时间变大时, 测量精度变高. 这主要由 s 的增大 (即测量基线的增加) 所导致的. 但是, s 增大会使目标在两幅图像中的变化增大, 从而特征点匹配的难度也会增加.

实验 3 参数设置为: 飞行速度 $v=15\text{ m/s}$, 采样时间 $\Delta t=0.4\text{ s}$, 移动距离 $s=v\Delta t=6\text{ m}$. 实验结果见表 3.

表 3 实验 3 的测量结果

Table 3 Measurement results of experiment 3

Distance to camera/m	Relative error of α	Relative error of β	Relative error of d_z	Relative error of d
50	1.12%	1.38%	1.2%	1.63%
100	1.71%	1.82%	1.4%	2.33%
150	2.09%	2.2%	1.8%	2.81%
200	2.59%	2.82%	2.25%	3.24%
250	3.22%	3%	2.64%	3.83%

对比表 1 和表 3, 可以看出当其他参数不变, 而飞行速度变大时, 测量精度降低. 其主要原因是速度快会导致成像模糊, 从而导致特征点匹配误差加大, 使得测量精度下降.

实验 4 参数设置为: 飞行速度 $v=10\text{ m/s}$, 采样时间 $\Delta t=0.4\text{ s}$, 移动距离 $s=v\Delta t=4\text{ m}$, 采用文献[14]的方法对标称值为 30 mm 的镜头标定后得到的精确焦距值为 30.009 mm , 将其带入测量公式中得到实验结果见表 4.

表 4 实验 4 的测量结果

Table 4 Measurement results of experiment 4

Distance to camera/m	Relative error of α	Relative error of β	Relative error of d_z	Relative error of d
50	0.60%	0.67%	0.93%	1.39%
100	1.32%	1.35%	1.29%	1.82%
150	1.81%	1.78%	1.59%	2.51%
200	2.31%	2.29%	2.04%	3.12%
250	2.76%	2.65%	2.47%	3.67%

对比表 1 和表 4 的数据可知, 采用标称值得到的测量误差与采用标定后焦距得到的测量误差相差在 0.05% 以内.

由于算法复杂度主要与特征点的检测及匹配计算

时间有关. 在实验 1 的条件下, 测量过程所花费时间见表 5.

表 5 测量方法的计算时间

Table 5 Computing time of the method

The feature number of A_1 image	The feature number of A_2 image	The feature number of match	Time/s
616	676	25	0.16
676	469	120	0.281
469	654	18	0.129
654	633	87	0.227
633	641	15	0.124

由表 5 可以看出该算法的计算时间与所需匹配的特征点数量有关, 但是总体时间不超过算法的采样时间, 即在一个采样周期 (0.4 s) 内, 该算法可以实现测量定位的计算, 从而满足算法的时效性.

另外, 在实验 2 的参数条件下, 采用文献[10]的方法进行深度测量所得到的实验结果与本文方法的实验结果进行对比, 如表 6.

表 6 不同深度测量方法的测量结果

Table 6 Measurement errors of different methods

Distance to camera/m	Relative error of the method	Relative error of Ref. [10]
10	0.4%	0.3%
20	0.45%	0.35%
50	0.64%	0.62%
100	1.13%	1.23%
150	1.53%	2.13%
200	2.45%	3.45%
250	2.92%	5.16%

从表 6 看出, 在距离小于 50 m 时, 文献[10]方法的测量精度略优于比本文方法; 但是当距离大于 100 m 时, 本文方法的测量精度优于文献[10]方法. 结合无人机视觉感知的应用背景, 若飞行器从发现威胁到做出规避动作需要 10 s 时间, 当飞行速度 $v=10\text{ m/s}$ 时, 则预警距离为 100 m . 在这种背景下, 本文方法将优于文献[10]方法.

6 结论

本文以无人机感知与避障为背景, 提出了一种基于运动的单目视觉测距方法. 该方法利用两幅不同位置图像中目标特征点位置变化与飞行器运动之间的关系, 求解出飞行器与障碍目标之间的位置信息. 并分析了摄像机安装及飞行器抖动对测量结果的影响. 测量实验表明, 在摄像机测量的范围内, 该方法的测量精度随着距离的减小和采样时间变大而增高, 随着飞行速度变大而降低. 当目标距离为 100 m 时, 测量误差在 2% 以内, 可以满足飞行对于避障预警距离的精度要求, 具有实际可行性.

参考文献

- [1] UTT J, MCCALMONT J, DESCHENES M. Test and integration of a detect and avoid system[C]. AIAA 3rd "Unmanned Unlimited" Technical Conference, Workshop and Exhibit, 2004: 6424.
- [2] SUWAL K, CHEN W Z, MOLNAR T. SeFAR integration testbed for see and avoid technologies[C]. Infotech, 2005:91-96.
- [3] SHAKERNIA O, CHEN W Z, GRAHAM S, *et al.* Sense and avoid (SAA) flight test and lessons learned[C]. Proceedings of the AIAA Infotech@Aerospace Conference, 2007:3003.
- [4] BATRA D, SAXENA A, Learning the right model: Efficient max-margin learning in laplacian CRFs[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2012, **6**: 2136-2143.
- [5] SAXENA, A, CHUNG S H. 3-d depth reconstruction from a single still image [J]. *International Journal of Computer Vision*, 2008, **76**(1), 53-69.
- [6] HOIEM D, EFROS A A, HEBERT M. Recovering occlusion boundaries from an image [J]. *International Journal of Computer Vision*, 2011, **91**(3), 328-346.
- [7] HEDAU, V, HOIEM, D, FORSYTH D. Recovering free space of indoor scenes from a single image [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2012, **6**: 2807-2814.
- [8] LIN Jing-yu, JI Xiang-yang, XU Wen-li, *et al.* Absolute depth estimation from a single defocused image[J]. *IEEE Transactions on Image Processing*, 2013, **22**(11), 4545-4550.
- [9] ZHANG Fa-quan, LU Li-ping, SHEN Man-de. Measurement method to object distance by monocular vision [J]. *Acta Photonica Sinica*, 2009, **38**(2): 453-456.
张法全, 路立平, 沈满德, 等. 单目视觉目标距离测量方法研究 [J]. 光子学报, 2009, **38**(2): 453-456.
- [10] HAN Yan-xiang, ZHANG Zhi-sheng, DAI Min. Monocular vision system for distance measurement based on feature points [J]. *Optics and Precision Engineering*, 2011, **19**(5): 1110-1117.
韩延祥, 张志胜, 戴敏. 用于目标测距的单目视觉测量方法 [J]. 光学精密工程, 2011, **19**(5): 1110-1117.
- [11] ZHAO Tian-yun, GUO Lei, ZHANG Li-chuan. A new algorithm of spatial positioning based on mono-vision [J]. *Journal of Northwestern Polytechnical University*, 2009, **27**(1): 47-51.
赵天云, 郭雷, 张立川. 基于单目视觉的空间定位算法 [J]. 西北工业大学学报, 2009, **27**(1): 47-51.
- [12] HARRIS C, STEPHENS M. A combined corner and edge detector [C]. Proceeding of the Fourth Alvey Vision Conference, 1988:147-151.
- [13] LOWE D G. Distinctive image feature from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, **60**(2): 91-110.
- [14] ZHANG Z. A flexible new technique for camera calibration [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2000, **22**(11):1330-1334.