融合激光测距的低成本摄像机用于环境制图的物体尺寸提取

Sachini Ekanayake[#1], Nilanga Thelasingha[*2], Bavantha Udugama[#3], G.M.R.I. Godaliyadda[#4], M.P.B Ekanayake[#5], B.G.L.T. Samaranayake[#6], J.V. Wijayakulasooriya[#7]电气与电子工程系，佩拉德尼亚大学。斯里兰卡佩拉德尼亚[1]emspe9353@gmail.com，[2] nilanga901@gmail.com，[3] bavantha11@gmail.com，[4]roshangodd@ee.pdn.ac.lk [5]mpb.ekanayake@ee.pdn.ac.lk，[6] tilantha@ee.pdn.ac.lk，[7]jan@ee.pdn.ac.lk，

摘要-It 对于各种应用而言，拥有一种对未知地形进行建图的方法至关重要。对于人类无法到达的区域，应提出一种识别环境的方法。如果能够获得环境地图，探索、灾害救援、交通运输以及许多其他工作都会更加便利。利用立体相机复现人类视觉系统，是一种理想的解决方案。本研究采用基于激光测距的技术，并将其与立体相机融合，以提取物体尺寸用于建图。我们利用摄像机数学模型对畸变进行了标定。通过半全局块匹配[1] 生成视差图，并采用一种新颖的视差图膨胀噪声抑制方法减少噪声。我们使用激光测距仪（LRF）数据以及经过噪声抑制的视觉数据来识别物体参数

关键词- **建图、视差、激光雷达、摄像机标定、立体视觉、尺寸提取**

I. 引言

建图通常是对包含各种物体和特征的环境进行图形化表示。建图可用于提取物体信息，例如高度、尺寸、材料类型、物体类型等。机器视觉通过技术赋予机器人类视觉能力。人类视觉将环境感知为一个色彩丰富的三维空间。

建图发挥着至关重要的作用，因为它能够帮助

环境制图在技术领域有诸多应用，例如军事应用、灾害救援、导航和探索等，也可用于工业中的许多安防场景。对于人类无法进入的场所，这无疑是理想的解决方案。

探索人类无法进入的区域、测量未知地形的尺寸，以及将这一提议的方法应用于无人机，都可用于构建下方地形的三维模型。此外，还可进一步根据具体应用确定材料类型并识别物体。文献中已有多种传感器可用于探测障碍物，包括声呐 [2], 雷达、激光雷达 [3]，视觉系统以及其他接近传感器 [4]。

但其中许多方法需要强大的计算能力，并使用昂贵的设备。

考虑到立体相机成本较低，且已通过立体视觉完成建图，人们早已开发出用于穿越给定点、以尽可能降低行驶成本的路径规划技术[5]。此外，激光测距仪（LRF）是一种精确的测量设备。围绕该主题，人们已开展了各种与激光测距仪相关的应用研究。

因此，通过所提出的方法，我们以较低的计算开销融合激光测距仪数据与立体相机数据，以提取物体的尺寸；其中，激光测距仪用于测量平面内的 X、Y 方向尺寸，立体相机则用于测量 Z 方向尺寸。由此即可完成环境建图。

II. 所提方案

在所提出的方案中，我们使用两台低成本的现成网络摄像头进行立体视觉，并使用一台从 Neato 真空吸尘器中拆下的简单 Neato XV 11 激光测距仪作为传感器。由于摄像机标定是必需的，因此我们对摄像机进行了适当标定，最多使用 14 个畸变系数，不受常规方法的限制。

根据左、右图像获得视差图，并据此找到物体的高度。对于长度和宽度，我们使用激光测距仪，通过它确定物体所在的平面，并按照第[E]节所述构建二维环境。方法方面，已有许多[6] 研究使用立体相机，但进行多次实现会带来很高的计算成本，不利于在合适的平台上实际应用。因此，我们提出从立体相机中提取物体的高度信息，并使用激光测距仪构建二维环境，进而构建三维物体。为此，只需使用左、右两幅图像、物体高度信息以及激光测距仪的扫描信息。这种方法的计算成本并不高，也并非无法实现。

所使用的摄像机各自存在畸变，其中径向畸变和切向畸变是最显著的两种。因此，必须先对摄像机进行标定和去畸变。标定是指确定摄像机的数学模型及其参数，分为两个部分。首先，分别对两台摄像机进行处理和标定，并求得误差参数。此外，为了生成最优视差图，还需要进行校正，以修正畸变

并使图像对齐到同一条水平线上，从而执行立体匹配。

在进行高度提取时，我们考虑了视差图像的最大强度值。因此，为了消除地平面噪声值的影响，拍摄图像时始终保持固定距离。这样，通过像素与厘米的关系，识别最大强度值即可得出物体高度。使用立体相机将三维转换为二维平面时，尺寸会发生变化。因此，我们提出的方法如 F 节所述。

与其进行多次实现，不如采用更高效的方案；考虑到实际应用中的效率，我们提出了一种通过膨胀降低视差图像噪声的新方法。 [6]

获得高度并使用激光测距仪获取物体所在平面后，我们的重点是将其应用于移动平台，并完成地图重建。

获得的尺寸数据可用于在虚拟现实中重建物体，从而准确创建环境的三维模型。识别出的尺寸可用于计算体积和表面积，也可用于深入分析物体，以确定其密度、质量及许多其他材料属性。

如今，大多数领域都采用立体相机结合图像处理的深度提取方法。所提出的方法是使用低成本摄像机获取视差，并采用膨胀方法去除视差中的噪声。

A. 相机模型识别

摄像机本身存在畸变。为了通过摄像机准确地获取深度，必须识别其内参并校正畸变。所使用的简单网络摄像头可以方便地建模为针孔相机。在针孔相机模型中，场景视图通过透视变换将三维点投影到像平面上，表示为：

$$sm' = A(R|t)M'$$

Or,

$$S\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

其中，(X, Y, Z) 是世界坐标空间中三维点的坐标。(u, v) 是投影点在像素中的坐标；主点通常位于像平面中心；A 是相机矩阵；焦距以像素为单位表示。

真实镜头通常存在畸变；若不对其进行建模，就无法利用摄像机进行任何精确测量。针孔摄像机可能存在以下畸变：径向畸变（枕形畸变和桶形畸变）以及切向畸变。

要识别并校正这些误差，需要进行校准过程。上述畸变可以通过以下方式进行数学建模。其中，k_1, k_2, k_3, k_4, k_5, k、 k_{14} 为径向畸变系数，P_1， P_2 为切向畸变系数。

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + t$$

$$x' = \frac{x}{z}$$

$$y' = \frac{y}{z}$$

$$x'' = x'\frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + 2P_1 x' y' + P_2(r^2 + 2x')$$

$$y'' = y'\frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + 2P_2 x' y' + P_1(r^2 + 2y')$$

其中，

$$r^2 = x'^2 + y'^2$$

$$u = f_x x'' + C_x$$

$$v = f_y y'' + C_y$$

B. 摄像机标定

摄像机标定分为两个部分。首先分别对单个摄像机进行标定，并求得误差参数和相机参数。随后将两台摄像机作为一对进行标定，以求得该装置的立体参数。校准过程中使用了尺寸已知的棋盘格，如图 1 所示。



FIGURE1：摄像机标定用棋盘格图像（左和右）

通过摄像机标定，求得了以下参数。

- 径向畸变：径向畸变的校正矩阵 切向畸变：切向畸变的校正矩阵 焦距：摄像机的焦距
- 主点：像平面的主点
-
-

通过立体配置标定获得了以下参数

- 摄像机 2 的旋转 摄像机 2 的像平面相对于摄像机 1 的旋转 摄像机 2 的平移 摄像机 1 的主点到摄像机 2 的主点的距离 基础矩阵 基础矩阵是同一场景中任意两幅图像之间的一种关系，它限定了场景中各点在两幅图像中的投影位置。

摄像机 2 相对于摄像机 1 的旋转矩阵为:

1.0000 0.0027 -0.0036 -
0.0027 1.0000 0.0016
0.0036 -0.0016 1.0000

摄像机 2 相对于摄像机 1 的平移为:

-93.1032 1.2802 0.1104

基础矩阵

-0.0000 -0.0000 0.0016
0.0000 0.0000 0.1269
-0.0018 -0.1274 0.6250

本质矩阵

-0.0048 -0.4414 1.0334
0.1125 0.1493 93.1061
-1.2801 -93.1021 0.1459

表 I 相机参数

参数	左摄像机	右摄像机
径向畸变	[0.0644 -0.2494 -0.6359]	[-0.0101 0.3192 -2.2780]
切向畸变	[6.9078e-04 -0.0011]	[-0.0048 -0.0035]
焦距	[729.9077 729.4782]	[733.1340 732.8580]
主点	[322.5457 226.0965]	[303.3310 224.2269]

使用找到的参数校正图像中的畸变，并通过校正过程将图像水平对齐。

C. 视差图生成。在图2中，我表示

像平面，c 表示像平面的主点。(X, Y, Z) 是

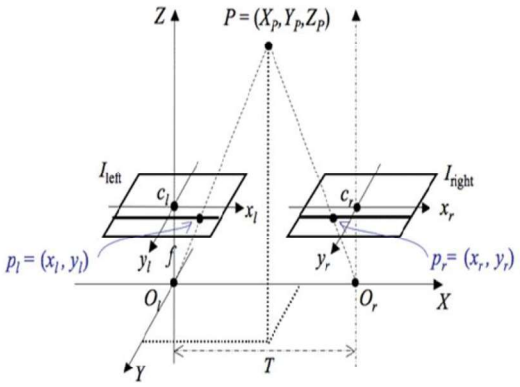


图2：视差的含义

笛卡尔坐标。P 表示真实世界点，p 表示该点在图像中的像。下标 l 和 r 分别表示左和右。根据几何关系，可以推导出真实世界物点与像点之间的关系:

$$Z_p = f\frac{T}{d},$$
$$X_p = x_l\frac{T}{d},$$
$$Y_p = y_l\frac{T}{d},$$

这里，d 表示 y,for 其数学表示如下:

$$d = x_l - x_r,$$

每个像素的视差构成视差图，视差图通常由半全局块匹配（SGBM）[1]生成。图 3 所示的立体图像已被用作生成视差图的基础。



图 3：立方体的立体图像

视差取决于多个参数。为找出这些参数，我们进行了如下实验。

为此，我们通过逐步增大与基线的距离（最高至 1 米）来拍摄图像对。此外，还改变了表面的纹理图案。

我们找出视差并手动调整至最佳状态，同时记录了参数值。

通过以下实验，得出了如下结论:

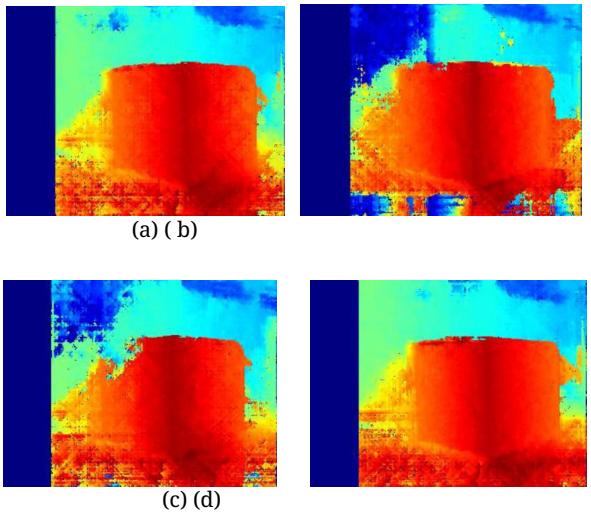
- 显然，视差数量取决于与基线的距离；距离基线越远，视差数量就应越少。由于激光测距仪（LRF）可测量与基线之间的实际距离，只需调整参数，使物体表面的相对深度能够被观测到即可。

- 块大小在很大程度上取决于物体表面的纹理。物体表面的眩光会严重影响视差。地
- 平面上的眩光也会影响视差。此外，对于
- 单色物体，很难计算视差。
-

因此，建议物体和地平面采用具有纹理变化的哑光表面。对于本研究所使用的相机设置，建议与基线的距离约为 0.75米。但该距离可能取决于物体尺度。

根据图4所示的视差结果，对测试环境进行了以下调整：测试环境为一个立方体（30厘米 x 30厘米 x 30厘米），立方体与摄像机之间的距离为70厘米；摄像机高度为8.5，与移动平台的高度相近，因为计划将摄像机安装在移动机器人上。摄像机到物体最近边缘的距离为0.70米。

图4(a)显示了光照条件良好时无纹理表面的视差图，图4(b)显示了光照条件较差时有纹理表面的视差图，图4(c)显示了光照条件良好时无纹理表面的视差图，而图4(c)显示了光照条件良好时有纹理表面的视差图。通过这些观察可以证明，视差受多种因素影响。



(a) (b)

(c) (d)

图4：不同测试条件下的视差图

D. 视差图平滑 视差图生成

采用 SGBM 计算得到的视差图受各种噪声分量影响，准确性不足。因此，使用膨胀来去除噪声。尽管膨胀是一种用于在简单图像处理过程中平滑原始相机图像的技术，但由于膨胀的内部过程会填充与周围像素强度偏差较大的像素，因此将其用于平滑视差图像。视差图像中的噪声正是这种类型，因此膨胀很适用。用核 B 对 A 进行膨胀定义如下：

$$(A \oplus B) = \bigcup_{b \in B} A_b$$

A_b 是平移 ion of A by b.

E. 用于二维地图生成的激光测距仪（LRF）

该激光测距仪（LRF）是由 Xv-11 激光雷达控制器驱动的 Neato 激光传感器。激光传感器向周围发射一串激光脉冲，并测量脉冲反射回来所需的时间 [7]。根据飞行时间，可以准确计算出反射点的距离（精确至 1 毫米）。图 5 显示了激光测距仪的误差特性。随后，传感器将距离数据编码成数据包，并转发至激光雷达控制器。控制器对数据包进行解码后，通过串行端口将数据传送至主机设备（个人计算机或树莓派等单板计算机）。由于百分比误差较小，在距离小于 3 米时，激光测距仪的测量结果可用于准确估算距离。

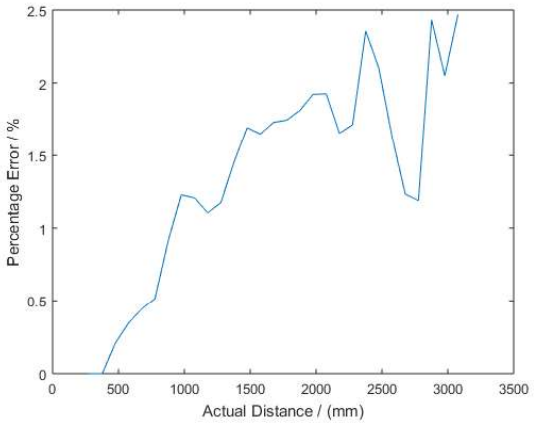


图 5：激光测距仪的误差特性

为了从激光测距仪获取距离数据，我们开发了一种单独的算法。由于激光测距仪在 0–3 米范围内测量距离时的方差较小，因此只考虑 3 米以内的距离。此外，反射率效应也可能导致误差，但其中许多误差可以通过多次采样来校正。考虑到运行速度和功耗，此处采样次数约为五次。如图 6 所示，使用激光测距仪进行多次采样可以生成准确的二维地图。

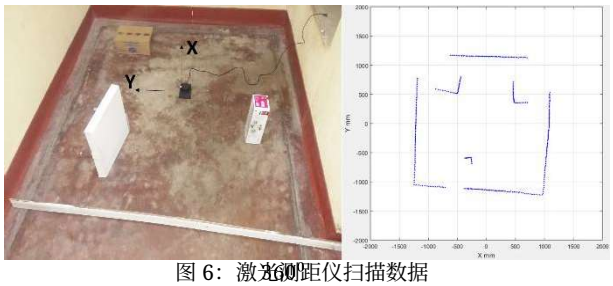


图 6：激光测距仪扫描数据

F. 高度提取

接下来，我们根据获得的视差图，着重找出其中的最大强度，因为如图 7 所示，它对应物体最近的边缘。但地面上存在的相同强度值也会产生影响。因此，图像是在固定距离处拍摄的，并通过像素值确定该值。

采用 Canny边缘检测方法 [8] 绘制包围矩形轮廓，并获得物体的像素高度。由于通过简单计算可得最大强度像素的顶部最小行索引，因此可得到以像素为单位的常数，如下所示。（常数 = 53 像素）。通过 Canny检测器找到的包围矩形会因同一物体的不同图像和场景而变化。因此，提出了另一种方法，通过视差图像提取物体高度

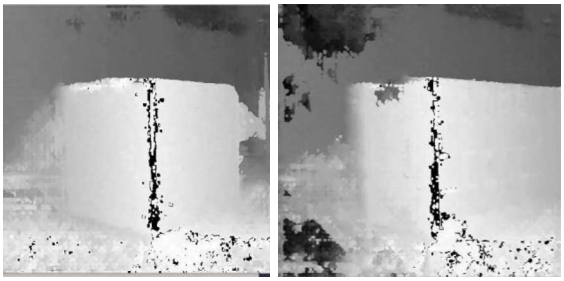


图7：使用视差图最大强度值进行边缘检测的结果

III. 使用最大强度值的高度提取方法。

首先，获取像素强度值，并确定用于筛除最大强度的阈值。因此，使用该阈值筛除视差图中的最大强度值。找到该值后，按行分析整幅视差图，直到找到第一个具有最高强度的像素，并记录其列索引。由于图像的总高度以及从图像底部到物体下端的像素数均已知，因此确定最大强度后，即可轻松计算出物体高度，如图 8 所示。

points’ column index

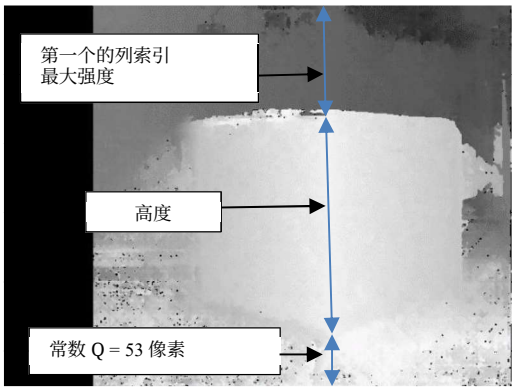


图8：使用视差图的最大强度值进行边缘检测的结果

IV. 像素与实际距离关系

由于始终使用相同的距离和相机高度，这些参数也适用于本实验。使用 CANNY边缘检测器（如图 9 所示），以像素为单位测量立方体（30x30x30）从一个角到另一个角的宽度。

立方体宽度 = 362像素 角之间的实际距离 = 42厘米 1个像素代表的距离 = 0.116 厘米



图9：使用 CANNY边缘检测器矩形确定角到角的距离

V. 结论

本文提出了一种将低成本摄像机与激光测距相结合，用于环境制图中物体尺寸提取的方法。针对摄像机标定，我们摒弃了传统的标定方法，将标定扩展至14个层级。此外，我们引入了膨胀方法来去除视差图中的噪声，而非采用需要多次采样且处理成本高昂的方法。针对物体的x-y细节提取，我们引入了激光测距仪（LRF），并据此模拟和仿真了使用LRF进行二维地图构建的过程，正如第*E*节所述。利用通过所提方法获得的高度，可以在二维地图上生成三维地图，最终实现环境制图。

因此，借助三维地图，该方法可用于识别物体类型和物体距离，还可进一步扩展到识别物体的材料。与使用高成本设备相比，我们提出的方法经济高效且易于实施。

像素尺寸与实际尺寸之比取决于与基线的距离。因此，可以确定两者之间的参数关系。利用这一关系，可以扩展所提出的算法，使其能够在距基线任意可测距离处识别物体的尺寸。

VI. 参考文献

[1] Hirschmuller, H.，《通过半全局匹配和互信息进行立体处理》，模式分析与机器智能，*IEEE*汇刊，第30卷，第2期，第328、341页，2008年2月[2] P. A. Plonski、J. V. Hook、C. Peng、N. Noori 和 V. Isler，使用前向……地面车辆，国际安全、防护与救援机器人研讨会（SSRR），2015年，第1至7页。[3]控制系统、计算与工程国际会议（ICCSCE），2016年，第441至445页。[4] Frueh、Zakhor，《利用结构化红外光捕捉时变场景的2 1/2 D深度和纹理》，三维数字成像与建模，2005年。3DIM2005第五届国际会议，第318、325页，2005年6月13日至16日[5] E. W. Dijkstra，“A Note 论与……相关的两个问题，271，1959年。[6] A. H. A. D. Abeysekara、D. P. Liyanage、W. R. C. B. S. Welikala、G. M. R. I. Godaliyadda、M. P. B. Eakanayake，以及……工业与信息系统国际会议（ICIIS），2015年，第320至325页。[7] K. Konolige、J. Augenbraun、N. Donaldson、C. Fiebig 和 P. ……第8届IEEE机器人与自动化国际会议，2008年，第3002至3008页。[8] John Canny，边缘检测的计算方法，1986年，IEEE 模式分析与机器智能汇刊，1986 年，第 679 至 698 页。

“Navigation around an unknown obstacle for autonomous facing sonar,” in 2015 IEEE

A. N. Cataranga and A. Ramos, “Obstacle detection using a 2D LIDAR system for an Autonomous Vehicle,” in 2016 6th IEEE

Graphs,” Numer. Math., vol. 1, pp. 269–

J. V. Wijayakulasooriya, “Depth map generation for a connaissance robot via sensor fusion,” in 2015 IEEE 10th

Shah, “A low cost laser distance sensor,” in 200

“A detection”, in