

基于 2 个灭点和局部尺度的三维空间尺度估计

周晓慧¹, 丁晓凤¹, 熊赞晖¹, 彭长歆²

(1. 华南理工大学数学学院, 广东 广州 510641;
2. 华南理工大学建筑学院, 广东 广州 510641)

摘 要: 三维空间尺度估计是三维重建中的一个重要工作, 现实世界中也存在一些基于单幅图像进行三维空间尺度估计的需求。通常情况下, 尺度估计需先对相机进行标定。根据单目图像符合透视原理的特性, 提出了一种基于 2 个灭点和局部尺度信息的方法对相机进行标定, 从而得到单目图像物体中三维空间尺度信息的估计。首先, 从单目图像中选择 2 组互相正交的平行线组, 得到对应 2 个灭点的坐标; 然后, 利用灭点坐标和焦距信息得到世界坐标系和相机坐标系之间的旋转矩阵, 再利用灭点的性质和已知局部尺度信息得到平移向量, 完成单目相机的标定; 最后, 还原二维图像中像素点对应的三维世界坐标值, 计算出图像中 2 个像素点在三维空间的尺度信息。实验结果表明, 该方法能有效地对单幅图像中的建筑物体进行尺度估计。

关 键 词: 单幅图像; 灭点; 局部尺度信息; 相机标定; 尺度估计

中图分类号: TP 391

DOI: 10.11996/JGj.2095-302X.2022030370

文献标识码: A

文章编号: 2095-302X(2022)03-0370-07

3D spatial scale estimation based on two vanishing points and local scale

ZHOU Xiao-hui¹, DING Xiao-feng¹, XIONG Yun-hui¹, PENG Chang-xin²

(1. School of Mathematics, South China University of Technology, Guangzhou Guangdong 510641, China;
2. School of Architecture, South China University of Technology, Guangzhou Guangdong 510641, China)

Abstract: 3D spatial scale estimation is of great significance to 3D reconstruction, and is in demand in the real world in the case of a single image. Generally, the camera must be calibrated before the scale estimation. According to the perspective principle of the monocular image, a method based on two vanishing points and local scale information was proposed to calibrate the camera, thereby obtaining the estimation of the 3D scale information in the monocular image. Firstly, two orthogonal groups of parallel lines were selected from the monocular image to produce the coordinates of two vanishing points. Then the rotation matrix between the world coordinate system and the camera coordinate system was yielded using the vanishing point coordinate and the focal length information. The translation vector was also acquired through the property of vanishing point and the known local scale information. Finally, the 3D world coordinates corresponding to the pixels in the 2D image were restored to calculate the 3D scale information. Experimental results show that this method can effectively estimate the scale of building objects in a single image.

Keywords: single image; vanishing point; local scale information; camera calibration; scale estimation

收稿日期: 2021-09-02; 定稿日期: 2021-09-28

Received: 2 September, 2021; Finalized: 28 September, 2021

基金项目: 国家自然科学基金项目(51978271)

Foundation items: National Natural Science Foundation of China (51978271)

第一作者: 周晓慧(1998-), 女, 硕士研究生。主要研究方向为相机标定和三维重建。E-mail: autoffier@163.com

First author: ZHOU Xiao-hui (1998-), master student. Her main research interests cover camera calibration and 3D reconstruction.

E-mail: autoffier@163.com

通信作者: 熊赞晖(1972-), 男, 副教授, 博士。主要研究方向为图像和视频处理、几何建模和处理、3D重建、3D打印等。E-mail: yhxiong@scut.edu.cn

Corresponding author: XIONG Yun-hui (1972-), professor, Ph.D. His main research interests cover image and video processing, geometry modeling and processing, 3D reconstruction, 3D printing, etc. E-mail: yhxiong@scut.edu.cn

恢复单幅或多幅二维图像中建筑物体的三维几何结构,即三维重建技术^[1-3]是计算机视觉的基础,也是当前人工智能、虚拟现实等众多前沿领域的研究难点和热点。其中,现实世界中存在一些基于单幅图像进行三维重建^[4-6]的需求,例如:从古建筑照片复原照片中的建筑信息。要想从单幅照片或少量照片中对建筑物复原出相同大小的建筑物就需要知道需恢复的建筑物体各部分的尺度信息,这几乎是一个不可能完成的任务。但在已知照片中某些局部尺度的情况下,通过本文的方法,可获取到照片中一些其他物体的尺度信息。通常情况下,尺度估计需先对相机进行标定^[7-9],确定相机的内部参数,再获取三维世界坐标和相机坐标变换中的旋转和平移,之后便可完成一些建筑物体的尺度信息计算工作。由于单目图像符合透视原理的特性及其计算的复杂性,利用灭点的属性进行相机标定^[10-15]是一个比较合适的方法。

本文在文献[10]的基础上提出了利用2个灭点及已知局部尺度进行三维空间尺度估计的方法。该方法只需做出假设:输入的单目图像可以供用户至少选取2组正交的平行线用以计算相应的2个灭点;图像中包含已知真实尺度的参照物体;投影中心是图像中心。其步骤如下:

(1) 从单目图像上选择2组互相正交的平行线组,计算平行线所构成的2个灭点的坐标;

(2) 通过2个灭点坐标和与平行线组平行的一条已知线段的真实尺度,计算单目图像相机标定的旋转矩阵和平移向量;

(3) 将像素坐标转化为世界坐标,根据三维空间物体各点的世界坐标和拓扑关系,估计三维空间物体的尺度。

1 相关工作

1990年,文献[11]首次提出利用灭点进行相机标定,并证明了灭点的3个属性,还根据这些属性找到了相机内外参数与灭点之间的关系。

文献[10]利用2个灭点和一已知线段长度进行相机标定和粗略的三维重建,但该方法并未充分利用灭点的性质使得计算略显繁琐。文献[12]提出一种基于灭点的代数测距方法,通过平行共面空间直线的2个正交集即灭点和参考平面上2个不平行线段的长度,可以唯一地确定参考平面上任意2点之间的距离。但该方法固定的模型参数和参考平面的选择会对尺度估计结果产生不良影响。

文献[13]在总结灭点属性的基础上,提出了一种基于灭点的三维物体重建方法,首先利用灭点计算相机的内、外部参数,再计算出物体上各点的三维空间坐标,然后通过对物体的绘制及纹理映射,实现三维物体的重建。文献[14]结合三维真实场景中的灭点提取、车道检测和车辆检测3个主要步骤,提出了基于摄像机标定的摄像机与正面车辆之间的距离估计方法。文献[15]提出了一种在人工环境下使用灭点检测进行自标定的算法,假设主点为图像的中心,以满足3个灭点的正交性,用于焦距的估计。但文献[13,15]基于灭点的标定方法,则要求图像中存在3组两两正交的平行线组,在实际应用中很难满足该条件,因此应用范围受到制约。

2 尺度估计流程

本文尺度估计的具体方法如图1所示,首先在单目图像上选择2组互相正交的平行线组,进行相应2个灭点坐标的计算;然后利用计算所得的2个灭点坐标和已知局部尺度信息进行相机标定;最后恢复2组平行线组所张成的平面上任意2点间的尺度信息。



图1 尺度估计主要流程

Fig. 1 Main processes of scale estimation

2.1 灭点坐标的计算

如图2所示,已知用户所选择的其中一组平行线组 l_1, l_2 为空间直线 L_1, L_2 在像平面的投影且均经过同一点,即灭点 V 。假设 $l_i(i=1,2)$ 在图片中被选择的端点为 $(x_{i_1}, y_{i_1}), (x_{i_2}, y_{i_2})$,则直线方程为 $E_i x + F_i y + G_i = 0$,其中

$$\begin{cases} E_i = y_{i_2} - y_{i_1} \\ F_i = x_{i_1} - x_{i_2} \\ G_i = y_{i_1} x_{i_2} - y_{i_2} x_{i_1} \end{cases} \quad (1)$$

设灭点在图像上的齐次坐标为 $V = (x_\infty, y_\infty, 1)^T$,投影直线的齐次坐标为 $l_i = [E_i, F_i, G_i]$,则 V 到 l_i 的距离和为

$$K = \sum_{i=1}^2 n_i |E_i x_\infty + F_i y_\infty + G_i| \quad (2)$$

其中, $n_i = \frac{1}{\sqrt{E_i^2 + F_i^2 + G_i^2}}$ 。由于 n_i 只与 E_i, F_i, G_i 有关, 所以求取灭点 V 的图像坐标, 即求解 K 的最小值可转化为线性代数的问题。令 2 条投影直线方程系数构成的对称矩阵为 $M = \sum_{i=1}^2 n_i^2 l_i^T l_i$, 其最小特征值对应的特征向量 $(x', y', N)^T$ 即为所求的归一化的灭点坐标, 则灭点 V 在图像上的非齐次坐标标为 $(x_\infty, y_\infty)^T = \begin{pmatrix} x' \\ y' \\ N \end{pmatrix}^T$ 。

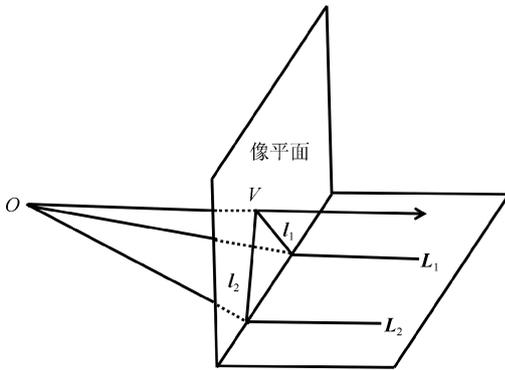


图 2 灭点示意图

Fig. 2 Sketch map of Vanishing Point

2.2 基于 2 个灭点和已知局部尺度的相机标定

在小孔成像的基础上, 相机模型的内外参数决定了像素坐标、图像坐标、相机坐标和世界坐标之间的转换。利用灭点的属性进行相机标定需做如下假设:

- (1) 输入的单目图像可以供用户至少选取 2 组正交的平行线用以计算相应的灭点;
- (2) 图像中包含已知真实尺度的参照物体;
- (3) 投影中心是图像中心。

2.2.1 旋转矩阵的计算

如图 3 所示, 世界坐标系 R_w 以 A 点作为原点, 由 3 个方向的正交基向量 $\alpha, \beta, \gamma (\gamma = \alpha \times \beta)$ 构成: $W = (A, \alpha, \beta, \gamma)$; 相机坐标系 R_c 以 O 点为原点, 由 3 个方向的正交基向量 $\alpha_c, \beta_c, \gamma_c (\gamma_c = \alpha_c \times \beta_c)$ 构成: $C = (O, \alpha_c, \beta_c, \gamma_c)$; 像平面坐标系为 $S = (P, \alpha_c, \beta_c)$, 其中 P 点为相机中心在像平面上的正交投影。世界坐标系 W 到相机坐标系 C 的旋转矩阵为 R , 则有

$$R \cdot \alpha = \alpha_c, R \cdot \beta = \beta_c, R \cdot \gamma = \gamma_c \quad (3)$$

直线 OV_1 和 OV_2 过投影中心, 其分别平行于向量 α' 和 β' , V_1 和 V_2 为灭点, 易知基向量 (α, β, γ) 和基向量 $(\alpha', \beta', \gamma')$ 相似, 因此 $C = (O, \alpha_c, \beta_c, \gamma_c)$ 和

$W = (A, \alpha, \beta, \gamma)$ 之间的旋转变换相同。假设 α', β', γ' 在相机系 C 下的坐标已知, 则有

$$\begin{cases} \alpha'_c = \frac{\overline{OV_{1c}}}{\|OV_{1c}\|} = \begin{pmatrix} V_{1\alpha_c} & V_{1\beta_c} & f \\ \|OV_{1c}\| & \|OV_{1c}\| & \|OV_{1c}\| \end{pmatrix}^T \\ \beta'_c = \frac{\overline{OV_{2c}}}{\|OV_{2c}\|} = \begin{pmatrix} V_{2\alpha_c} & V_{2\beta_c} & f \\ \|OV_{2c}\| & \|OV_{2c}\| & \|OV_{2c}\| \end{pmatrix}^T \\ \gamma'_c = \alpha'_c \times \beta'_c \end{cases} \quad (4)$$

其中, $\overline{OV_{1c}}, \overline{OV_{2c}}$ 为直线 OV_1, OV_2 方向向量在相机坐标系下的坐标; $V_{1\alpha_c}, V_{1\beta_c}, V_{2\alpha_c}, V_{2\beta_c}$ 为灭点 V_1, V_2 在相机坐标系基向量 α_c, β_c 方向上的分量, $\gamma'_c = (\gamma'_{\alpha_c}, \gamma'_{\beta_c}, \gamma'_{\gamma_c})$, $\gamma'_{\alpha_c}, \gamma'_{\beta_c}, \gamma'_{\gamma_c}$ 为 γ'_c 在相机坐标系的基向量 $\alpha_c, \beta_c, \gamma_c$ 方向上的分量, f 为已知相机焦距, 故

$$R \cdot \alpha_w = \alpha'_c, R \cdot \beta_w = \beta'_c, R \cdot \gamma_w = \gamma'_c \quad (5)$$

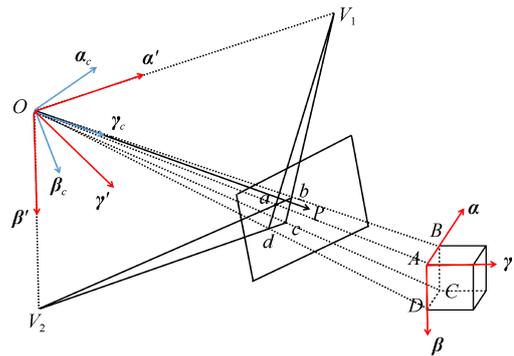


图 3 计算旋转矩阵示意图

Fig. 3 Sketch map for calculating rotation matrix

在世界坐标系 W 中, $\alpha_w = (1, 0, 0)$, $\beta_w = (0, 1, 0)$, $\gamma_w = (0, 0, 1)$, 联合式(3)和式(4), 可得世界坐标系 W 到相机坐标系 C 的旋转矩阵

$$R = \begin{bmatrix} \frac{V_{1\alpha_c}}{\sqrt{V_{1\alpha_c}^2 + V_{1\beta_c}^2 + f^2}} & \frac{V_{2\alpha_c}}{\sqrt{V_{2\alpha_c}^2 + V_{2\beta_c}^2 + f^2}} & \gamma'_{\alpha_c} \\ \frac{V_{1\beta_c}}{\sqrt{V_{1\alpha_c}^2 + V_{1\beta_c}^2 + f^2}} & \frac{V_{2\beta_c}}{\sqrt{V_{2\alpha_c}^2 + V_{2\beta_c}^2 + f^2}} & \gamma'_{\beta_c} \\ \frac{f}{\sqrt{V_{1\alpha_c}^2 + V_{1\beta_c}^2 + f^2}} & \frac{f}{\sqrt{V_{2\alpha_c}^2 + V_{2\beta_c}^2 + f^2}} & \gamma'_{\gamma_c} \end{bmatrix}$$

2.2.2 平移向量的计算

如图 4 所示, 结合图 3, O 为相机中心, AD 为空间中已知真实尺度的线段(如: 已知规格尺寸的砖块), V 为 AD 形成的灭点, ad 为 AD 在像平面上的投影, 过 a 沿 AD 方向作直线 ad' 交 OD 于点 d' 。

假设世界坐标系原点位于点 A , 易知 \overline{OA} 即为相机坐标系 C 变换到世界坐标系 W 的平移向量, 故平移向量 $T = -\overline{OA}$ 。由于 $\overline{ad'} \parallel AD$, 所以 $\Delta Oad'$ 和 ΔOAD 相似, 故有

$$\frac{\|\overline{OA}\|}{\|\overline{Oa}\|} = \frac{\|\overline{AD}\|}{\|\overline{ad'}\|} \Rightarrow \|\overline{OA}\| = \frac{\|\overline{AD}\| \cdot \|\overline{Oa}\|}{\|\overline{ad'}\|} \quad (6)$$

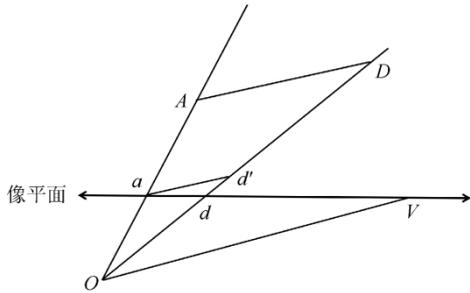


图 4 计算平移向量示意图

Fig. 4 Sketch map of calculating translation vector

由灭点的性质可知 $AD \parallel \overline{ad'} \parallel OV$, 设 a 的相机坐标为 (x_1, y_1, f) , d 的相机坐标为 (x_2, y_2, f) , V 的相机坐标为 (x_3, y_3, f) , 则 \overline{ad} 的方向向量为

$$\overline{d_{od}} = \frac{\overline{od}}{\|\overline{od}\|} = \frac{(x_2, y_2, f)}{\sqrt{x_2^2 + y_2^2 + f^2}} \quad (7)$$

设 $\overline{od'} = t \cdot \overline{d_{od}} = \frac{t \cdot (x_2, y_2, f)}{\sqrt{x_2^2 + y_2^2 + f^2}}$ (t 为实数), 又因

为点 d' 在 OD 上, 且 $AD \parallel \overline{ad'} \parallel OV$, 即 $(t \cdot \overline{d_{od}} - \overline{oa}) \parallel \overline{OV}$, 则有

$$\begin{aligned} \frac{\frac{t}{h}x_2 - x_1}{x_3} &= \frac{\frac{t}{h}y_2 - y_1}{y_3} = \frac{\left(\frac{t}{h} - 1\right)f}{f} \\ \Rightarrow t &= \frac{h(x_1 - x_3)}{x_2 - x_3} = \frac{h(y_1 - y_3)}{y_2 - y_3} \end{aligned} \quad (8)$$

其中, $h = \sqrt{x_2^2 + y_2^2 + f^2}$, 故 $\overline{ad'}$ 的长度可知为 $\|\overline{ad'}\| = \|t \cdot \overline{d_{od}} - \overline{oa}\|$, 因 \overline{OA} 的方向与 \overline{Oa} 的方向相同, 根据式(5)和已知的 AD 长度即可得 \overline{OA} 的长度。因此, 平移向量为

$$T = -\overline{OA} \quad (9)$$

2.3 建筑物估计尺度的计算

如图 5 所示, 像素坐标 $[u, v]$ 向世界坐标 $[X_w, Y_w, Z_w]$ 的转化实质是: 先由像素坐标转化为图像坐标,

再由图像坐标转化为相机坐标, 最后由相机坐标转化至世界坐标。

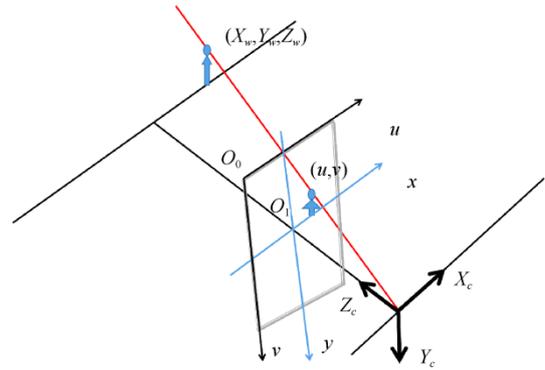


图 5 像素坐标系与世界坐标系示意图

Fig. 5 Sketch map of pixel coordinate system and world coordinate system

以图像左上角为原点 O_0 建立笛卡尔直角坐标系 $u-v$ (单位: 像素), 图像在计算机数组中的列数与行数是像素的横坐标 u 和纵坐标 v 。定义图像坐标系的原点 O_1 为成像平面和相机光轴的交点且在 $u-v$ 坐标系下的坐标设为 (u_0, v_0) ; x 轴与 u 轴平行, y 轴与 v 轴平行; 每个像素在横轴 x 和纵轴 y 上的物理尺寸为 d_x 和 d_y 。则图像中每个像素在 $u-v$ 坐标系下与在 $x-y$ 坐标系下的坐标之间存在如下关系, 即

$$u = \frac{x}{d_x} + u_0, \quad v = \frac{y}{d_y} + v_0 \quad (10)$$

假设图像坐标系以毫米作为物理单位, 则 dx 的单位为毫米/像素, x/dx 的单位为像素, 式(10)可表示为

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (11)$$

至此完成了像素坐标到图像坐标的转化。

如图 6 所示, 根据针孔成像原理, π 为摄像机的像平面, O_c 为摄像机中心, f 为相机焦距, 光轴为以 O_c 为端点垂直于像平面的射线, p 为光轴和成像平面的交点即图像的主点, 图像坐标系为 $O-xy$, 相机坐标系为 $O_c-x_cy_cz_c$ 。记三维空间点 X 在相机坐标系中的齐次坐标为: $X=(X_c, Y_c, Z_c, 1)^T$ 。其像点 m 在图像坐标系中的齐次坐标为: $m=(x, y, 1)^T$ 。根据三角形相似原理可得: $x = \frac{fX_c}{Z_c}, y = \frac{fY_c}{Z_c}$, 其矩阵形式为

$$\begin{aligned} \mathbf{Z}_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} &= \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \end{aligned} \quad (12)$$

可得: $X_c = \frac{\mathbf{Z}_c(u-u_0)dx}{f}$ 和 $Y_c = \frac{\mathbf{Z}_c(v-v_0)dy}{f}$ 。至此完成了像素坐标到相机坐标的转化。

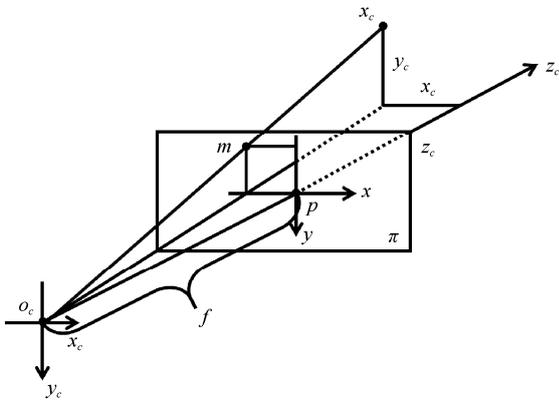


图6 图像坐标系与相机坐标系的关系

Fig. 6 The relationship between image coordinate system and camera coordinate system

而相机坐标到世界坐标的转化实际上是一种刚体变化,对坐标进行旋转和平移操作即可得到转化关系,二者之间的刚体变换数学表示为

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \mathbf{R} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad (13)$$

其中, \mathbf{R} 为 3×3 的旋转矩阵; $\mathbf{T} = (T_x, T_y, T_z)^T$ 为平移向量。其对应的齐次表达为

$$\begin{aligned} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} &= \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_3^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \end{aligned} \quad (14)$$

将上述所示的转换过程合并可得像素坐标和世界坐标之间的数学关系,即

$$\mathbf{Z}_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (15)$$

其中, (u_0, v_0) 为图像中心像素坐标为

$$\mathbf{H} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_3^T & 1 \end{bmatrix}, \quad f_x = \frac{f}{dx}, \quad f_y = \frac{f}{dy}.$$

在计算的过程中,假设生成灭点的2组平行线组所在平面的世界坐标 Z 值是相同的,则推导出二维图像点的三维世界坐标为

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{Z}_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (16)$$

选取已知尺度 d_{true} 参照物的2个像素点 (u_1, v_1) , (u_2, v_2) , 再结合相机参数焦距 f 获取世界坐标 $(X_{w_1}, Y_{w_1}, Z_{w_2})$, $(X_{w_2}, Y_{w_2}, Z_{w_2})$, 则由灭点和已知参照物尺度确定的平面上任意2个图像点之间的真实距离,可通过欧几里得距离公式计算得到

$$d = \sqrt{(X_{w_1} - X_{w_2})^2 + (Y_{w_1} - Y_{w_2})^2 + (Z_{w_1} - Z_{w_2})^2} \quad (17)$$

3 算法实现及实验结果分析

3.1 算法实现

算法实验使用 CPU i5-8500, 8 G 内存 Windows10 系统中 Visual Studio 2015 环境和 Python3.7 环境实现主要代码,还采用了 OpenCV, Numpy, PIL 等第三方代码库。实验主要研究对象是室外场景单目图像,其包含一些已知尺度信息的参照物(如已知尺度信息的墙砖或地砖)。

在单幅单目图像中,基于2个灭点和局部尺度信息的三维空间尺度估计算法(3D spatial scale estimation, SSE)步骤如下:

输入: 单目图像和一条线段的真实尺度值。

输出: 图像建筑尺度估计值。

步骤 1. 用户从图像中选择2组互相正交的平行线组,计算出2个灭点的坐标;

步骤 2. 通过2个灭点坐标和图像中这2组平行线所在平面中的一条已知线段的真实尺度,计算单目图像相机标定的旋转矩阵和平移向量;

步骤 3. 将像素坐标转化为世界坐标,即可得到这2组平行线所在平面上建筑物体的估计尺度。

3.2 误差定义

通过人工实地测量的方式得到图像参照物的真实尺度值, 再利用算法的计算结果得到图像建筑物的真实尺度估算值。用相对误差来对 SSE 算法的有效性进行评估。

相对误差函数定义为: $e_{rel} = \frac{e_{abs}}{d_{true}} \times 100\%$, 其中

d_{true} 为建筑尺度真实值, $e_{abs} = |d - d_{true}|$ 为绝对误差, d 为建筑尺度估算值。

3.3 实验结果及分析

用户对于平行线组的选取和相应的灭点计算如图 7 所示。图中红色和绿色 2 组平行线在真实场景中互相正交, 并确定了尺度估计所需的 2 个灭点。对于每个灭点, 用户(用鼠标)选择一组建筑物上存在的平行线。为了使选择更加容易, 输入图像是超大分辨率原图, 且未进行缩小。具体的灭点坐标计算见 2.1 节。

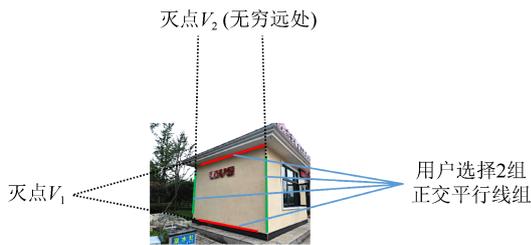


图 7 灭点计算示意图

Fig. 7 Sketch map of extinction point calculation

图 8 中的 #1, #2 和 #3 行表示同一建筑物不同视角下的图像。图 8(a) 为同一个建筑物不同观测角度下所拍摄的图像; 图 8(b) 为用户在相应图像上选择的 2 组正交平行直线(红、绿线组)结果, 用以

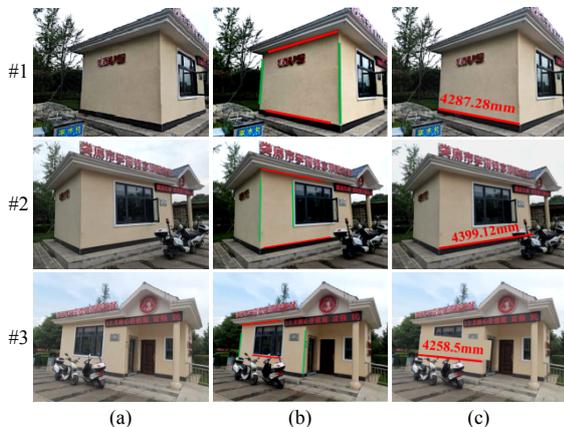


图 8 同一建筑物不同视角下的尺度估计((a)单目图像; (b)平行线选择; (c)尺度估计结果)

Fig. 8 The scale estimation under different viewpoints of the same building (a) Monocular image; (b) Selection of parallel lines; (c) Scale estimation results)

计算平行线组形成的灭点坐标; 图 8(c) 为单目图像建筑物物体尺度估计的结果, 其中红线为建筑物物体尺度估计的范围。

表 1 显示了同一建筑物不同视角下 SSE 算法的实验误差结果。由表 1 可知, SSE 算法在不同视角下对同一建筑物物体进行尺度估计均可获得较好的结果。

表 1 同一建筑物不同视角下尺度估计的误差
Table 1 Errors of scale estimation under different view angles of the same building

数项	尺度估计值 (mm)	尺度真实值 (mm)	相对误差 e_{rel} (%)
#1	4 287.28	4 212.00	1.79
#2	4 399.12	4 322.00	1.78
#3	4 258.50	4 322.00	1.47

图 9 为不同方法对同一建筑物物体进行尺度估计的比较结果。图 9(a) 为建筑物物体室外场景的单目图像; 图 9(b) 为 SSE 算法的尺度估计效果, 其中红线区域即为建筑物物体尺度估计的范围; 图 9(c) 为文献[12]运用灭点和单应性矩阵方法得到的尺度估计效果; 图 9(d) 为文献[13]运用 3 个灭点进行相机标定后的尺度估计效果。其中图 9(d) 前 2 行尺度估计值在图像上没有展示的原因是, 图像中无法准确找到 3 组两两正交的平行线组。

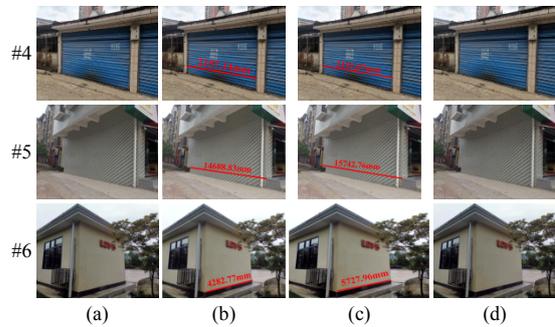


图 9 不同算法间建筑物物体尺度估计的比较((a)单目图像; (b) SSE 算法; (c)文献[12]算法; (d)文献[13]算法)
Fig. 9 Comparison of building scale estimation by different methods ((a) Monocular image; (b) The SSE algorithm; (c) The algorithm from reference [12]; (d) The algorithm from reference [13])

表 2 显示了不同方法对同一建筑物物体进行尺度估计的误差结果, 其中“-”表示未得到相应估计值。从表中可以看出, SSE 算法效果较好, 文献[12-13]所提出的尺度估计方法误差相对较大。其中文献[12]的方法是基于提前设置好的相机模型参数、参考平面和垂直于参考平面的一个参照尺度, 进行与参考平面垂直平面上的相关尺度估计。这种固定的模型参数和参考平面的选择会对估计结果产生一定的影响。而文献[13]的方法则是基于 3 组两两垂直的

灭点对相机进行标定从而得到待测物体的尺度比例关系。理论上来说, 3 个灭点相较于 2 个灭点进行标定应该更加精准, 但是在真实场景的图像中可能因为很难找到相互正交的 3 组平行线形成的 3 个灭点, 反而造成更大的误差。文献[13]算法需要采用 2 张同一物体在同一方向、不同角度下的图像来对其进行优化, 研究更多的是针对建模和仿真, 主要是恢复图像中物体尺度的比例关系, 更多的是满足虚拟现实系统的需求, 因此在对真实场景的单目图像中物体进行估计时会造成较大的误差。

表 2 不同方法进行建筑物尺度估计的误差
Table 2 Errors of building scale estimation by different methods

图像	方法	估计值(mm)	真实值(mm)	$e_{rel}(\%)$
#4	SSE 算法	2 157.14	2 150	0.33
	文献[12]	2 191.07	2 150	1.91
	文献[13]	-	2 150	-
#5	SSE 算法	14 688.83	14 800	0.75
	文献[12]	15 742.76	14 800	6.37
	文献[13]	-	14 800	-
#6	SSE 算法	4 282.77	4 212	1.68
	文献[12]	4 341.73	4 212	3.08
	文献[13]	5 727.96	4 212	36.99

4 结束语

由于单目图像符合透视原理的特性, 用灭点对图像建筑物进行尺度估计是一个比较好的方法。在实际应用中, 有些建筑图像(包括古建筑残存的图像)找不到符合要求的 3 组两两正交的平行线组形成的灭点, 因此无法用基于 3 个灭点的算法来计算。对于这些图像, 如果可以从图像中找到一些已知尺度的物体(如:某个时期古建筑中采用固定尺寸的砖块、地砖和瓦片等), 通过这些已知尺度物体信息, 再结合 2 个灭点就可对图像中建筑物的某些区域进行尺度估计。基于该思想本文提出了利用 2 个灭点和局部尺度信息进行三维空间尺度估计的方法。

实验结果表明, 与文献[12-13]的尺度估计算法相比较, 本文算法有比较好的计算结果。并且该算法对输入的图像限制和所需的几何约束较少, 适用于更多的场景, 泛化性高。但是由于本文算法存在已知尺度线段需平行于某一组所选平行线的限制条件, 实际应用中如果已知尺度线段不平行于任何一组平行线, 可以考虑将其投影到其中一组平行线后再进行平移向量的计算。同时, 因为本文算法是基于传统灭点理论来进行尺度估计, 故和其他灭点算法一样, 缺乏对图像中各像素点的深度估计, 只

能计算 2 组平行线所在平面上的 2 个像素点间的距离, 还无法做到计算图像中任意 2 点间的距离。

参考文献 (References)

- [1] 董鹏辉, 柯良军. 基于图像的三维重建技术综述[J]. 无线电通信技术, 2019, 45(2): 115-119.
DONG P H, KE L J. Overview of 3D reconstruction techniques based on images[J]. Radio Communications Technology, 2019, 45(2): 115-119 (in Chinese).
- [2] LIU C X, KONG D H, WANG S F, et al. Deep 3D reconstruction: methods, data, and challenges[J]. Frontiers of Information Technology & Electronic Engineering, 2021, 22(5): 652-672.
- [3] TAHIR R, SARGANO A B, HABIB Z. Voxel-based 3D object reconstruction from single 2D image using variational autoencoders[J]. Mathematics, 2021, 9(18): 2288.
- [4] LI G Q, WU Z H, LIU Y X, et al. 3D hand reconstruction from a single image based on biomechanical constraints[J]. The Visual Computer, 2021, 37(9-11): 2699-2711.
- [5] YAN C G, SHAO B Y, ZHAO H, et al. 3D room layout estimation from a single RGB image[J]. IEEE Transactions on Multimedia, 2020, 22(11): 3014-3024.
- [6] CHEN Y, TU Z, KANG D, et al. Joint hand-object 3D reconstruction from a single image with cross-branch feature fusion[J]. IEEE Transactions on Image Processing, 2021, 30: 4008-4021.
- [7] MAALEK R, LICHTI D D. Automated calibration of smartphone cameras for 3D reconstruction of mechanical pipes[EB/OL]. [2021-07-10]. <https://arxiv.org/abs/2012.02899>.
- [8] 王沛, 钟凯, 颜福鹏, 等. 基于相位移编码圆的相机离焦快速标定方法[J]. 光学技术, 2021, 47(4): 494-499.
WANG P, ZHONG K, YAN F K, et al. Fast calibration method of camera defocus based on phase shift coding circle[J]. Optical Technique, 2021, 47(4): 494-499 (in Chinese).
- [9] 赫美琳, 高明慧, 李爽, 等. 一种单目相机标定算法研究[J]. 数字通信世界, 2018(5): 11-14.
HE M L, GAO M H, LI S, et al. Research on a monocular camera calibration algorithm[J]. Digital Communication World, 2018(5): 11-14 (in Chinese).
- [10] GUILLOU E, MENEVEAUX D, MAISEL E, et al. Using vanishing points for camera calibration and coarse 3D reconstruction from a single image[J]. The Visual Computer, 2000, 16(7): 396-410.
- [11] CAPRILE B, TORRE V. Using vanishing points for camera calibration[J]. International Journal of Computer Vision, 1990, 4(2): 127-139.
- [12] WANG G H, WU Y H, HU Z Y. A novel approach for single view based plane metrology[C]//2002 International Conference on Pattern Recognition. New York: IEEE Press, 2002: 556-559.
- [13] 余焯, 刘晓平. 基于灭点的三维物体重建[J]. 系统仿真学报, 2008, 20(15): 4069-4072.
YU Y, LIU X P. 3D object reconstruction using vanishing points[J]. Journal of System Simulation, 2008, 20(15): 4069-4072 (in Chinese).
- [14] BOUGHARRIOU S, HAMDAOUI F, MTIBAA A. Vehicles distance estimation using detection of vanishing point[J]. Engineering Computations, 2019, 36(9): 3070-3093.
- [15] LEE S J, HWANG S S. Fast and accurate self-calibration using vanishing point detection in manmade environments[J]. International Journal of Control, Automation and Systems, 2020, 18(10): 2609-2620.