

一种具有精确位姿的飞机 CAD 模型投影图像库设计方法

付泰 杨力 王斌

(中国卫星海上测控部 江苏 江阴 214431)

摘要 近年来,随着 CAD 技术和 3D 场景理解等方面研究的发展,基于目标 3D CAD 模型进行精确位姿估计成了一种重要方法。然而,直接使用 CAD 模型时往往需要在编程中调用商用 CAD 软件,这不仅需要配置大量兼容性文件而且对计算机图形硬件要求较高,同时软件层次数据频繁交互导致效率低下。因此,可以通过对 CAD 模型进行投影得到其带有位姿真值的图像库,从而将三维问题转化为二维问题,使其大大简化。考虑到程序的集成性,选用 OpenGL 作为 CAD 模型的显示工具。首先标定虚拟相机内外参数,考虑到运动的相对性,保持虚拟相机不动,按设定距离将目标模型以 1° 为步长,俯仰和方位各旋转 $0\sim 90^\circ$ 和 $0\sim 360^\circ$ 进行投影,从而得到具有精确位姿的目标投影图像库。

关键词 复杂三维结构,3D CAD 模型,OpenGL,精确位姿,投影

中图分类号 TP301 **文献标识码** A

Projection Image Library Design Method for Aircraft CAD Model with Accurate Pose

FU Tai YANG Li WANG Bin

(China Satellite Maritime Tracking and Controlling Department, Jiangyin, Jiangsu 214431, China)

Abstract With the development of CAD technology and 3D scene understanding in recent years, accurate pose estimation based on target 3D CAD model has become an important method. However, direct application of the CAD model often requires commercial CAD software to be invoked in programming, in this way, it not only need to configure a large number of compatibility files, but also has higher requirements for computer graphics hardware. Meanwhile, the frequent interaction of software level data leads to low efficiency. Therefore, an image library with the truth values of pose can be obtained by projecting the CAD model, which greatly simplify the problem by means of transforming the three-dimensional problem into a two-dimensional problem. Considering the integration of programs, OpenGL is used as the display tool of CAD model. Firstly, the internal and external parameters of virtual camera are calibrated. Considering the relativity of motion, the virtual image is kept motionless, and the pitch and azimuth are rotated $0\sim 90$ degrees and $0\sim 360$ degrees respectively with the target model projected at 1 degree according to the set distance. The target projection image library with accurate pose is obtained.

Keywords Complex three-dimensional structure, 3D CAD model, OpenGL, Accurate pose, Projection

1 前言

物体识别是计算机视觉应用的一项基本任务,其目的是寻找到一个包含可以区分不同目标物体的有效信息的描述^[1]。目标物体事先已知,其几何特征可以被直接应用于识别任务。随着计算机图像技术的不断发展,3D 模型的制备与现实技术得到了巨大发展,目前可用的 3D CAD 模型的格式有 3ds, obj, max 及 ezm 等,可视化的工具包含 OpenGL, OS-GRender 及 Blender 等。OpenGL 由于具有高度的可移植性及显示的逼真性,被广泛应用于游戏开发、医学显示、气象模拟等领域^[2],考虑到程序的集成性,本文选用 OpenGL 作为 CAD 模型的显示工具。

由于实际目标往往具有十分复杂的三维结构,导致物体和相机发生相对运动时所拍摄图像中的目标外形可能发生较大变化,这使得对具有复杂三维结构的目标的检测及精确位姿估计问题成为图像和视觉领域的重难点之一。针对简单形状基元(如点、线、平面等)3D 结构的检测及姿态估计在计算

机视觉发展的早期就被大量研究,但针对具有复杂 3D 结构的物体的研究仍有欠缺,例如图 1 中的歼 15 飞机。近年来,随着图像处理、模式识别及计算机视觉等领域相关技术的飞速发展,研究人员在目标检测及 3D 场景理解等方面取得了突出的研究成果,这使得针对具有复杂 3D 结构物体的检测及姿态估计存在可能,基于目标 3D 模型的检测及姿态估计的研究受到了越来越多的关注,并取得了一定的研究成果。

基于 3D CAD 模型可以方便地实现采样视点下仿真图像的生成,本文基于待检测飞机的 3D CAD 模型离线生成目标的仿真图像样本库。歼 15 飞机 CAD 仿真模型如图 2 所示。



图 1 歼 15 飞机真实图像



图 2 歼 15 飞机 CAD 仿真模型

考虑到程序的集成性,选用 OpenGL 作为 CAD 模型的显示工具。相比于 max 及 ezm 模型,3ds 及 obj 模型格式简单、易于程序读取。由于 obj 格式的模型相比于 3ds 格式的模型表现的结构更为精细,因此选用 obj 格式的模型。下面介绍基于 obj 模型及 OpenGL 的样本库制备。

2 投影变换

为实现基于 OpenGL 的模型显示,需要理清 OpenGL 模拟成像与真实世界相机成像模型之间的对应关系。此外,相机成像模型采用最为常用的小孔成像模型^[3],其中涉及到的坐标系主要有世界坐标系、相机坐标系、物体坐标系及图像坐标系,OpenGL 中涉及的坐标系主要有世界坐标系、模型坐标系、视点坐标系及设备坐标系。类似于相机拍照的过程,从三维模型到二维图像,OpenGL 的绘制过程需要经历以下阶段^[4]:

- 1) 调整相机位置,对准三维场景(视点变换);
- 2) 调整三维模型在场景中的位置(模型变换);
- 3) 调整相机镜头使三维场景投影在二维成像靶面上(投影变换);
- 4) 对二维成像进行裁剪(视口变换)。

至此即完成了三维场景到二维图像之间的变换。运动是相对的,通过调整相机的位置或通过调整物体的位置均可达到相同的效果,因此,在 OpenGL 处理机制中使用一个模视矩阵表示模型变换及视点变换,OpenGL 的成像过程可表达为^[4]:

$$\mu \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = FPM \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (1)$$

其中, μ 为常数因子, F 为视口矩阵, P 为投影矩阵, M 为模式变换矩阵, $[x_w \ y_w \ z_w \ 1]^T$ 为物体在世界坐标系中的坐标, $[u \ v \ 1]^T$ 为二维图像中的坐标表示。实际相机采集成像的过程可表示为:

$$\mu \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f/dx & 0 & u_0 \\ 0 & f/dy & v_0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (2)$$

其中, f, dx, dy, u_0 及 v_0 均为相机内参数, R 及 T 为旋转矩阵及平移向量,通过对比 OpenGL 的成像过程与实际相机的成像过程,可明确相应参数间的对应关系,为基于 OpenGL 实现模拟图像的生成奠定基础。

3 OpenGL 指令

在采用的 OpenGL 开发库中使用 $glViewport(0,0,scrX,scrY)$ 设置视口矩阵 F ,其中 $scrX$ 及 $scrY$ 为窗口的大小;调用 $glFrustum(x_l, x_r, y_b, y_t, z_{near}, z_{far})$ 设定投影矩阵,其中 (x_l, y_b) 和 (x_r, y_t) 分别为成像平面的左下角及右上角坐标, z_{near} 和 z_{far} 分别为视锥体距视点近、远平面的距离。采用 glm 库^[5]进行相关矩阵的设置与运算,指令语句 $Translatef$ 及 $Rotatef$ 对应模型变换中的平移操作及旋转操作; $LookAt$ 语句用于视点设置,其返回的结果构成模视变换矩阵 M 。下面给出视口矩阵及投影矩阵的表达式:

$$F = \begin{bmatrix} scrX/2 & 0 & 0 & 0 \\ 0 & scrY/2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

令

$$A = \frac{2z_{near}}{x_r - x_l}, B = \frac{x_r + x_l}{x_r - x_l}, C = \frac{2z_{far}}{y_t - y_b}, D = \frac{y_t + y_b}{y_t - y_b}, \quad (4)$$

$$E = -\frac{z_{far} + z_{near}}{z_{far} - z_{near}}, F = -\frac{2z_{far}z_{near}}{z_{far} - z_{near}}$$

则

$$P = \begin{bmatrix} A & 0 & B & 0 \\ 0 & C & D & 0 \\ 0 & 0 & E & F \\ 0 & 0 & -1 & 0 \end{bmatrix} \quad (5)$$

此处,平移和旋转与实际相机成像过程中相同,相机焦距为 f ,水平和垂直方向上每个像元的物理尺寸为 dx 和 dy ,相机靶面的物理尺寸为 lx 和 ly ,则像素数分别为 lx/dx 和 ly/dy 。设相机距离物体的最大可能距离为 s ,则相应的参数关系为:

$$\begin{cases} z_{near} = f \\ z_{far} = s \\ x_l = -lx/2 \\ x_r = lx/2 \\ y_b = -ly/2 \\ y_t = ly/2 \\ scrX = lx/dx \\ scrY = ly/dy \end{cases} \quad (6)$$

通过对比 OpenGL 的成像过程与实际相机的成像过程,可明确相应参数间的对应关系,为基于 OpenGL 实现模拟图像的生成奠定基础。

4 图像库生成

对目标的三维数字模型进行二维投影,本文中保持虚拟相机不动,转动目标得到目标在不同姿态时的二维投影图,也即相当于相机在空间中的不同位置观察目标。本文中使用的 OpenGL 对目标的三维模型进行多角度二维投影,离线生成模型库。

明确参数之间的关系及相关设置之后,进行仿真图像的生成。为方便设置,采用 $LookAt$ 语句对视点进行设置,依据得到的模视变换矩阵并按照既定的旋转顺序(这里采用依次

绕 x, y 及 z 轴的顺序进行)分解出平移向量及 3 个欧拉角,作为对应仿真图像的位姿真值记录。这里在仿真图像的生成过程中只在以飞机模型为中心的上半球面上进行视点的采样,图 3 给出了球面上 5 个采样点下对应的仿真图像。在实际的应用中,可根据飞机的运动轨迹及与相机的相对位置关系等约束进一步缩小采样的范围,如图 3 中深色框选择区域,这样可以进一步减少样本数量,或者采用更加密集的视点采样,以提高后续参数估计的精度。

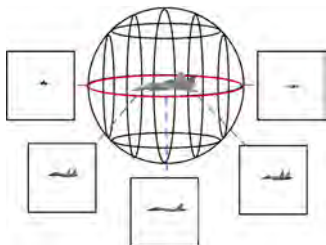


图 3 对三维模型进行二维投影的示意图

在实际仿真过程中,通过控制俯仰角 ψ 、偏航角 θ 及半径 r ,即可实现视点的设置,其中 $\psi \in [0, \pi/2], \theta \in [0, 2\pi)$,如图 4 所示。

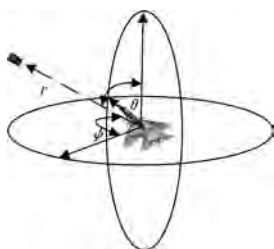


图 4 仿真图像生成中视点设置示意图

采样密度的增加有利于提高参数估计的精度,但样本数量的增加会增加计算负担,因此应综合考虑精度及处理时间的要求,设置合理的采样密度,完成样本数据库的制备工作。

本着算法研究的目的,本文将 ψ 及 θ 的采样步长均设为 1° ,共生成 32760 张仿真图,部分仿真图如图 5 所示。

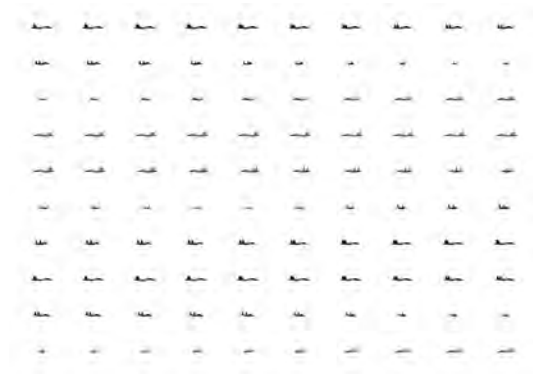


图 5 部分仿真图

结束语 本文以基于 CAD 模型的飞机目标位姿估计问题中的难点为研究目标,通过 OpenGL,保持相机不动,旋转飞机目标来离线建立具有精确位姿的 32760 张飞机图像库,后续工作将把本文算法应用到飞机位姿精确估计中,以解决三维位姿估计问题。

参考文献

(上接第 214 页)

- [11] DOLLAR P, APPEL R, BELONGIE S, et al. Fast Feature Pyramids for Object Detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 36(8): 1532-1545.
- [12] ZHANG S, BAUCKHAGE C, CREMERS A B. Informed Haar-Like Features Improve Pedestrian Detection[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2014: 947-954.
- [13] NAM W, DOLLAR P, HAN J H. Local Decorrelation For Improved Pedestrian Detection[J]. Advances in Neural Information Processing Systems, 2014, 1: 424-432.
- [14] PAISITKRIANGKRAI S, SHEN C, HENGEL A V D. Pedestrian Detection with Spatially Pooled Features and Structured Ensemble Learning[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(6): 1243.
- [15] ZHANG S, BENENSON R, OMRAN M, et al. How Far are We from Solving Pedestrian Detection? [C]//IEEE Conference on Computer Vision & Pattern Recognition. 2016: 1259-1267.
- [16] ZHANG H, XU M, ZHUO L, et al. A novel optimization framework for salient object detection [J]. The Visual Computer, 2016, 32(1): 31-41.
- [17] CHENG M M, ZHANG Z, LIN W Y, et al. BING: Binarized Normed Gradients for Objectness Estimation at 300fps [C]//Computer Vision and Pattern Recognition. IEEE, 2014: 3286-3293.
- [18] 刘涛, 吴泽民, 姜青竹, 等. 基于似物性的快速视觉目标识别算法 [J]. 计算机科学, 2016, 43(7): 73-76.
- [19] BROX T, BREGLER C, MALIK J. Large displacement optical flow [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2009. (CVPR 2009). IEEE, 2009: 41-48.
- [20] DOLLAR P, WOJEK C, SCHIELE B, et al. Pedestrian detection: A benchmark [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2009 (CVPR 2009). IEEE, 2009: 304-311.