

## 基于标志点识别的三维位姿测量方法

阮利锋,王 赓,盛焕烨

(上海交通大学 计算机科学与工程系,上海 200240)

(sharp@sjtu.edu.cn)

**摘 要:**提出一种基于双目视觉原理测量运动目标三维位姿参数的新方法。采用 LED 作为标志点来定义运动目标局部坐标系,通过双目视觉方法计算各 LED 的全局坐标,进而得出运动目标的三维位姿参数。同时根据实验环境的特殊性,提出一种基于优先级的外极线斜率约束匹配算法。实验结果表明该测量方法简单高效,满足精度和实时性的要求。

**关键词:**三维位姿测量;立体匹配;立体视觉

**中图分类号:**TP242.6<sup>+</sup>2;TP391 **文献标志码:**A

## 3D position and attitude measurement based on marking-points recognition

RUAN Li-feng, WANG Geng, SHENG Huan-ye

(Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200240, China)

**Abstract:** A new method based on binocular visual theory to measure 3D position and attitude parameters of a moving object was described. Local coordinate system of the moving object was defined by LED marking-points. By calculating 3D position of each LED, 3D position and attitude parameters of the moving object can be calculated. At the same time, according to the particularity of environment, an algorithm called constraint of epipolar-slope match based on priority was presented. Experimental results show this method is simple and efficient, and satisfies the requirements of precision and real time.

**Key words:** 3D position and attitude measurement; stereo matching; stereo vision

## 0 引言

具有自主飞行能力的小型无人直升机,可用于执行航拍、勘测、自主侦察等任务,具有重要的军事、民用和科学研究价值。无人直升机进行自主飞行控制需要实时感知自身的位姿参数信息,作为反馈参数提供给自主飞行控制算法。目前已有多种不同用途的位姿参数传感设备,包括全球定位系统(Global Positioning System, GPS)、三维陀螺、电子罗盘、动态航姿仪等。但像 GPS 等某些设备由于不适用于室内环境,而无法用于构建室内试验平台。而基于计算机视觉的位姿测量方法能够适用于室内环境,文献[1-3]提出了多种针对不同应用需求的计算机视觉位姿测量方法。

为了构建室内小型无人直升机自主飞行控制试验平台,便于进行直升机动力学模型及自主飞行控制算法相关的研究工作,本文提出了一种基于标志点识别的三维位姿测量方法。该方法以双目视觉原理为基础,同时结合实际应用需求,达到了对运动目标的三维位姿参数进行实时、高精度测量的目的。

## 1 标志点跟踪识别

### 1.1 标志点识别

本文采用白色发光二极管(Light Emitting Diode, LED)作为运动目标的标志点,如图 1(a)所示。为了能够从摄像机捕获的图像中识别标志点,首先需要一种图像二值划分方法,使的标志点成像能够从原始图像中分割出来。目前已有多种进行图像分割的成熟算法<sup>[4]</sup>。但本文中作为标志点的白色 LED 成像的像素点灰度值相对于环境成像的像素点灰度值

来说要大得多,因此可以简单地设定一个灰度阈值,来达到对原始图像进行二值划分的目的,如图 1(b)所示。标志点的识别是整个系统中最耗时的部分,为了达到系统对实时性的要求,并不预先对原始图像进行整体二值划分,而只是在搜索过程中判别某些像素点是否为标志点成像中的像素点。

标志点识别是在以该标志点的预测图像坐标为中心的一个矩形范围内进行的,且该矩形范围可动态调整,如图 1(c)所示。标志点成像是一系列连续像素点所形成的一个近似圆形的小区域范围。因此在定位标志点成像区域时,无须对所有像素点进行连续性搜索,而只须进行间隔性搜索。如图 1(d)所示,只需对交叉点上的像素点进行判别。定位过程是以矩形的中心为起点,向四周进行扩散。应用上述策略,可有效地提高标志点成像区域的定位效率。

当定位到标志点的成像区域时,为了提高标志点的识别精度,需改用连续性搜索的方式来定位标志点成像区域的四个边界(上下、左右)。而标志点成像区域的四个边界的中心则认为该标志点的图像坐标。此外,为了防止由于环境中离散亮点的干扰而造成标志点的错误识别,需要对定位后的四个边界进行判别,判断其所围成的区域是否符合标志点成像区域的形状特征。

### 1.2 标志点跟踪

运动目标是连续运动的,其运动速度较小且不易发生突变。同时实验中摄像机的采样频率较高,达到 20 Hz。因此同一标志点在连续两次识别过程中所处的图像坐标之间的差距较小。由此,根据标志点的历史运动轨迹信息,可以实现标志点图像坐标的预测性跟踪。本文以标志点的当前识别坐标为

收稿日期:2008-05-05;修回日期:2008-07-05。

**作者简介:**阮利锋(1982-),男,浙江慈溪人,硕士研究生,主要研究方向:计算机视觉、智能信息处理;王赓(1970-),男,陕西洋县人,讲师,博士研究生,主要研究方向:嵌入式系统、人工智能;盛焕烨(1943-),男,浙江慈溪人,教授,博士生导师,CCF 会员,主要研究方向:智能信息系统、计算金融、现代物流。

前一次识别坐标和下一次预测坐标的中心为预测方法,预测标志点的图像坐标。实验表明,该方法能够有效地提高标志点的识别效率。

此外,标志点在遇到物理遮挡或超出搜索范围等不确定情况时,可能导致该标志点跟踪丢失。此时应适当地增大搜索范围,以便重新识别该标志点。而较大的搜索范围又会导致识别效率的降低。因此,本文采用了动态的标志点搜索范围,以当前搜索结果来动态调整后续搜索范围,既保证了标志点的有效识别,又提高了标志点的识别效率。

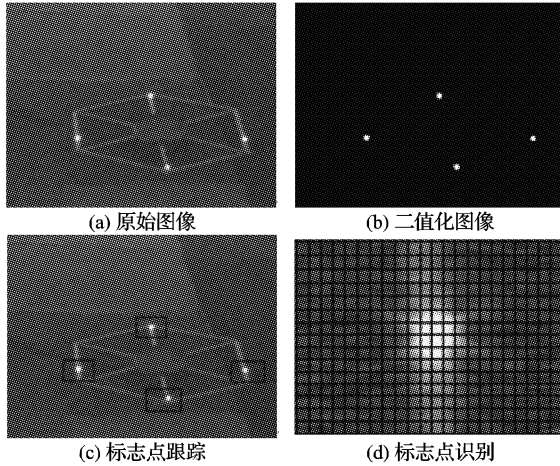


图1 标志点跟踪识别

## 2 标志点三维坐标求解

### 2.1 摄像机模型

本文采用常用的针孔模型来描述摄像机成像模型<sup>[5]</sup>。在该模型中,定义了如图2所示的坐标系,其中包括全局坐标系  $W(X_w, Y_w, Z_w)$ 、摄像机坐标系  $C(X_c, Y_c, Z_c)$ 、像平面坐标系  $P(X_p, Y_p)$  和图像坐标系  $I(U, V)$ 。

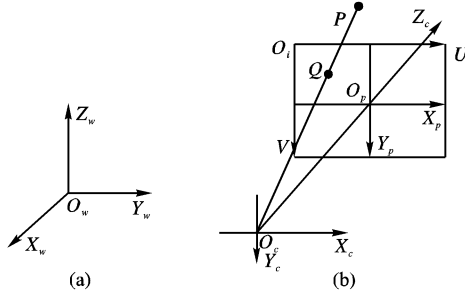


图2 摄像机模型坐标系

由该模型,可得不同坐标系下点坐标之间的转换关系。

1) 点  $P$  在摄像机坐标系下坐标  $P_c$  与其成像点  $Q$  在图像坐标系下坐标  $P_i$  之间的投影关系可由式(1)表示:

$$Z_c P_i = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} P_c = M_1 P_c \quad (1)$$

其中:  $d_x, d_y$  为像素的物理尺寸;  $u_0, v_0$  为点  $O_p$  在图像坐标系下的坐标;  $f$  为焦距;  $Z_c$  为焦点到像平面的距离。  $M_1$  为摄像机内部参数矩阵,只与摄像机内部结构有关。

2) 点  $P$  在全局坐标系下坐标  $P_w$  和摄像机坐标系下坐标  $P_c$  之间的转换关系可由式(2)表示:

$$P_c = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} P_w = M_2 P_w \quad (2)$$

其中:  $R, t$  分别为旋转矩阵和平移矩阵。  $M_2$  为摄像机外部参数矩阵,由摄像机位姿状态决定。

摄像机内外部参数矩阵由摄像机标定获得。目前已有多种摄像机标定方法<sup>[6]</sup>,在此不作讨论。

### 2.2 基于优先级的外极线斜率约束匹配算法

本文所采用的标志点是无差别的白色 LED,因此仅根据识别后的标志点图像坐标,无法对同一 LED 标志点在不同摄像机上的成像进行配对。而特征点匹配则是结合其他信息来实现上述目的,目前已有多种特征点匹配算法<sup>[7-9]</sup>。但在本文中,标志点个数是恒定的,且任意两个标志点之间的距离相对较大。鉴于该特殊环境条件,本文提出了一种基于优先级的外极线斜率约束匹配算法,但该算法并不适用于一般意义上的特征点匹配。

由几何关系可知,点  $P$  在摄像机 2 上的成像点  $P_2$  必在平面  $PO_1O_2$  与摄像机 2 成像面的交线上。该交线称为点  $P_2$  的外极线,而该约束称为外极线约束<sup>[5]</sup>。同理可知,点  $P$  在摄像机 1 上的成像点  $P_1$  在摄像机 2 上的投影点  $Q$  也在该外极线上,如图 3(a) 所示。结合这些几何约束关系,便可实现标志点匹配的目的。

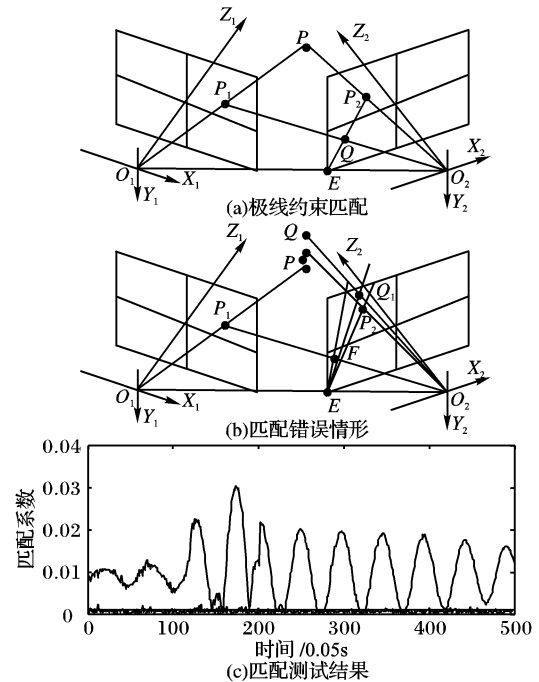


图3 标志点匹配

标志点识别过程中获得的是标志点图像坐标。根据式(1),可得成像点  $P_1, P_2$  在各自摄像机坐标系下的坐标,如式(3)所示:

$$\begin{cases} P_1^1 = (M_1^1)^{-1} Z_c^1 P_i^1 \\ P_2^2 = (M_2^1)^{-1} Z_c^2 P_i^2 \end{cases} \quad (3)$$

根据式(2),可得成像点  $P_1$  在摄像机 2 坐标系下的坐标,如式(4)所示:

$$P_1^2 = M_2^2 (M_1^1)^{-1} P_1^1 \quad (4)$$

继而可得成像点  $P_1$  在摄像机 2 上的投影点  $Q$  在摄像机 2 坐标系下的坐标,如式(5)所示:

$$Q = Z_c^2 [P_1^2(1)/P_1^2(3) \quad P_1^2(2)/P_1^2(3) \quad 1] \quad (5)$$

同理,可得摄像机 1 的光心  $O_1$  在摄像机 2 上的投影点  $E$  在摄像机 2 坐标系下的坐标。

至此,即可得出直线  $EQ, EP_2$  的斜率  $k_1, k_2$ 。由外极线约束可知,若点  $P_1, P_2$  为同一 LED 标志点的成像点,则必有  $k_1 = k_2$ 。但实际上,由于图像噪声、标志点识别误差等因素的影响,  $k_1$  和  $k_2$  并不恰恰相等,即直线  $EQ, EP_2$  之间存在一个夹角。该夹角正切值的绝对值可由式(6)计算如下:

$$K = |(k_1 - k_2)/(1 + k_1 \cdot k_2)| \quad (6)$$

参数  $K$  在本文中称为匹配系数,标志着不同摄像机中的

两个成像点为同一 LED 标志点所成像的可能性的。\$K\$ 值越小,匹配的可能性越大。根据该匹配系数,可以从摄像机 2 中的多个标志点中选出与标识点 \$P\_1\$ 匹配可能性最大的点,作为标识点 \$P\_1\$ 的匹配标志点。

但在运动目标运动过程中,可能存在另一标志点 \$Q\$,其在摄像机 2 上的成像点 \$Q\_1\$ 与点 \$P\_1\$ 之间的匹配系数小于点 \$P\_2\$ 与点 \$P\_1\$ 之间的匹配系数而造成误匹配,如图 3(b) 所示。由此可见,只以匹配系数为判断依据是不充分的,仍需对上述匹配过程进行修正。

标志点是跟踪识别的,即可知点 \$P\_1\$ 在前一次匹配过程中的匹配点是点 \$P\_2\$,而且在误匹配时,点 \$P\_2\$、\$Q\_1\$ 与点 \$P\_1\$ 之间的匹配系数都相当小。鉴于上述特点,本文通过引入优先级的概念,提出了一种基于优先级的外极线约束匹配算法。该算法的基本流程描述如下:1) 预设一个匹配系数阈值。2) 优先计算点 \$P\_1\$ 与前次匹配点 \$P\_2\$ 之间的匹配系数,即点 \$P\_2\$ 具有较高优先级。3) 若该匹配系数小于预设阈值,则认为点 \$P\_2\$ 仍为匹配点,否则取多个标志点中匹配系数最小的点作为匹配点。图 3(c) 为该匹配算法的测试结果。由结果分析可知,该算法能够有效地消除误匹配情况,提高匹配正确率。

### 2.3 标志点三维坐标计算

通过标志点的跟踪识别和匹配过程,可获得同一标志点在不同摄像机中的成像点坐标。根据立体双目视觉的空间三

维点重建方法<sup>[5,10]</sup>,即可得出该标志点的全局坐标。

## 3 运动目标位姿参数求解

### 3.1 位置参数计算

LED 标志点配置及由标志点所定义的运动目标局部坐标系如图 4 所示。4 个标志点构成十字形,且任意两个标志点间的距离可预先确定。根据这些预知信息,结合标志点的三维坐标,可对各 LED 标志点进行区分。由此可知,标志点所定义的局部坐标系是唯一的,且局部坐标系原点 \$O\$ 的全局坐标是该运动目标的位置参数 \$t\$。

### 3.2 姿态参数计算

4 个 LED 标识点的局部坐标是预先确定的,且其全局坐标可通过计算得出。如图 4(b) 所示,由 LED 标志点可确定局部坐标系的 \$X\$ 轴、\$Y\$ 轴,再在 \$Z\$ 轴上取局部坐标为 \$[0 \ 0 \ 1]^T\$ 的点 \$P\_5\$,且其全局坐标可由任意三个标志点全局坐标计算得出。

由点 \$P\_5\$ 及任意两个标志点可定义完整局部坐标系坐标。假设该三点的局部坐标分别为: \$P\_1\$、\$P\_2\$、\$P\_5\$,其相应全局坐标为 \$Q\_1\$、\$Q\_2\$、\$Q\_5\$,根据坐标转换关系:

$$[P_1 \ P_2 \ P_5] = R \cdot [Q_1 - t \ Q_2 - t \ Q_5 - t] \quad (7)$$

可得从全局坐标系到局部坐标系的旋转矩阵 \$R\$。

由坐标系旋转关系<sup>[11]</sup>可知,依次绕 \$Z\$ 轴 \$\gamma\$ 角度、\$X\$ 轴 \$\alpha\$ 角度和 \$Y\$ 轴 \$\beta\$ 角度的旋转矩阵为:

$$R = \begin{bmatrix} \cos\beta \cdot \cos\gamma - \sin\alpha \cdot \sin\beta \cdot \sin\gamma & \cos\beta \cdot \sin\gamma + \sin\alpha \cdot \sin\beta \cdot \cos\gamma & -\cos\alpha \cdot \sin\beta \\ -\cos\alpha \cdot \sin\gamma & \cos\alpha \cdot \cos\gamma & \sin\alpha \\ \sin\beta \cdot \cos\gamma + \sin\alpha \cdot \cos\beta \cdot \sin\gamma & \sin\beta \cdot \sin\gamma - \sin\alpha \cdot \cos\beta \cdot \cos\gamma & \cos\alpha \cdot \cos\beta \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \quad (8)$$

由式(8)可得运动目标的姿态参数,如式(9)所示:

$$\begin{cases} \gamma = -\arctan(R_{21}/R_{22}) \\ \alpha = \arcsin(R_{23}) \\ \beta = -\arctan(R_{13}/R_{33}) \end{cases} \quad (9)$$

至此,便得到了运动目标的三维姿态参数:航向角 \$\gamma\$、俯仰角 \$\alpha\$ 和翻滚角 \$\beta\$。

## 4 实验结果及分析

实验采用利凌 PIH-7330 摄像机、AVENTR SR12575 镜头以及 Matrox Morphis 图像采集卡进行图像采集,采集频率为 20 Hz。运动目标为一个配置了 4 个 LED 标志点的正六边形平面框架,如图 1(a) 所示。其中左右两个 LED 标志点之间的距离为 1 000 mm,上下两个 LED 标志点之间的距离为 1 200 mm。框架用细绳悬于天花板,可自由摆动。

图 5 为测量实验结果,其中图 5(a)~(c) 为位置参数,图 5(d)~(f) 为姿态参数,图 5(g) 为某两个 LED 标志点的测量距离与真实距离之间的误差。

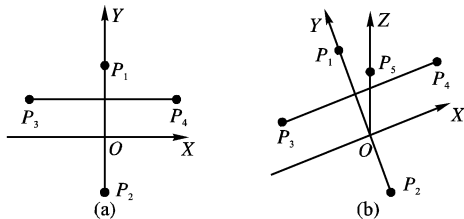


图 4 标志点配置示意图

由静态、动态位姿测量结果及标志点间距离误差分析可知,系统能够达到 \$\leq 3\$ mm 的静态位置参数测量精度, \$\leq 8\$ mm 的动态位置参数测量精度, \$\leq 0.1^\circ\$ 的静态姿态参数测量精度, \$\leq 0.5^\circ\$ 的动态姿态参数测量精度,以及 20 Hz 的测量频率,完全满足小型无人直升机进行室内自主飞行控制实验对位姿参数精度及测量实时性的要求。

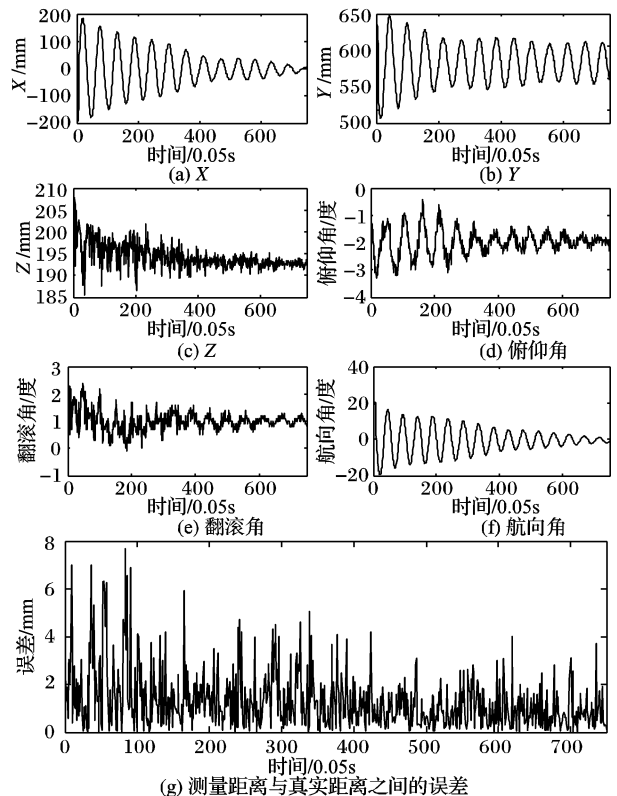


图 5 系统测量实验结果

## 5 结语

本文提出的运动目标三维位姿参数测量方法简单高效,满足测量精度和实时性的要求,能够用于构建室内小型无人直升机自主飞行控制试验平台。但该方法的应用场景并不局限于此。在理论上,该方法可在一定空间范围内用于测量任

(下转第 2862 页)

Toronto 到 Maryland 的 1:1 映射 31 个, 1:n 映射 12 个。

在这两个本体集上, 我们分别用了 GLUE<sup>[6]</sup>、DSSim<sup>[7]</sup> 和 CCOM 三种方法进行了映射实验, 然后对得到的结果进行分析对比。

#### 4.2 评判标准

本文用查全率 (Recall) 和查准率 (Precision) 这两项指标来评估实验结果。查全率指算法识别出的正确的映射规则与人工标注的映射规则的比值, 查准率指算法识别出的正确映射规则与算法识别出的所有映射规则的比值。

$$\text{查全率: } R = \frac{\text{算法识别的正确映射数}}{\text{人工标注的所有映射数}} \quad (13)$$

$$\text{查准率: } P = \frac{\text{算法识别的正确映射数}}{\text{算法识别的所有映射数}} \quad (14)$$

#### 4.3 实验结果与分析

实验结果如图 2 所示。

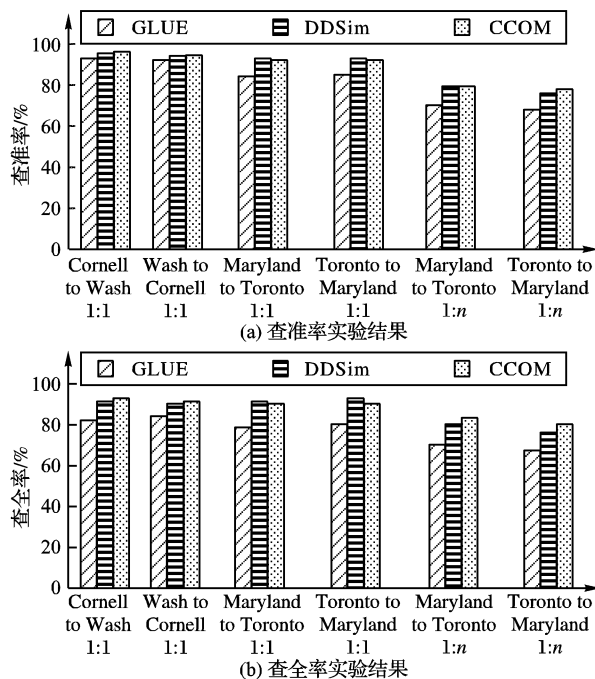


图 2 实验结果比较

从实验结果我们可以看到, CCOM 的映射精度显著高于 GLUE 算法, 而与 DSSim 算法基本相同。与 GLUE 算法相比, 在本体实例数较大时, CCOM 优势较小。但在实例较少时, CCOM 的优势就很明显。这是因为在 CCOM 方法中, 由于概念集群的引入, 可以用概念间的相互关系, 有效抵消实例减少

带来的映射误差。此外, CCOM 相对于 GLUE, 需要的领域专家的协助更少。与 DSSim 相比, 在 1:1 的映射中, 两者相差无几, 但在 1:n 的映射中, CCOM 略胜一筹, 这也得益于概念集群的引入。而在运算的时间上, CCOM 的映射速度更快, 更有效率。在实际操作中, 我们还发现, 如果能够根据实际环境恰当调整算法中的四个权重值, 能够进一步提高映射准确率。

#### 5 结语

本文主要讨论了基于概念集群的本体映射方法 CCOM。为了解决现有的本体映射方法不能兼顾效率与性能的问题, 本文提出了基于概念集群的本体映射方法 CCOM。概念集群是本文提出的一种概念集合, 它由一个核心概念以及本体中与这个核心概念具有语义结构关系的其他概念组成, 这样概念集群就包含了概念间的相互关系信息。CCOM 就是在本体映射过程中, 通过实例联合概率分布来计算, 提升了计算的速度。通过概念集群, 引入概念间的语义结构关系, 提升了映射的效果。通过实验分析, 证明本方法可以得到理想的映射精度与映射速度。

但是, 概念集群的构造对于 CCOM 的效率有很大的影响。目前的方法需要对本体的概念进行多次遍历。对于涉及上万概念的大型映射, 很可能成为新的效率瓶颈。因此, 如何快速有效地创建概念集群, 将是本项工作今后的研究重点。

#### 参考文献:

- [1] 宋炜, 张铭. 语义网简明教程[M]. 北京: 高等教育出版社, 2004.
- [2] CHOI N Y, SONG I Y, HAN H. A survey on ontology mapping[J]. ACM SIGMOD Record, 2006, 35 (3): 34-41.
- [3] 袁洋, 李善平. 基于语义 Web 的本体映射方法综述[J]. 计算机科学, 2004, 31(5): 5-8.
- [4] 唐杰, 梁邦勇, 李娟子, 等. 语义 Web 中的本体自动映射[J]. 计算机学报, 2006, 29(11): 1956-1976.
- [5] SEKINE S, SUDO K, OGINO T. Statistical matching of two ontologies[C]// Proceedings of the SIGLEX99: Standardizing Lexical Resources. Maryland, USA: [s. n.], 1999: 69-73.
- [6] DOAN A, MADHAVAN J, DOMINGOS P, et al. Learning to map between ontologies on the semantic Web[C]// Proceedings of the 11th International Conference on World Wide Web: WWW 2002. New York: ACM Press, 2002: 662-673.
- [7] NAGY M, VARGAS-VERA M, MOTTA E. DSSim-ontology mapping with uncertainty[C]// Proceedings of the 1st International Workshop on Ontology Matching: OM 2006. Athens GA: [s. n.], 2006: 115-123.

(上接第 2858 页)

何能够使用标志点进行标志的运动目标的三维位姿参数, 具有广阔的应用前景。

#### 参考文献:

- [1] 丁辉, 付梦印. 基于标志点识别的自主车视觉导航[J]. 光学技术, 2005, 31(6): 864-867.
- [2] 张志勇, 张靖, 朱大勇. 一种基于视觉成像的快速收敛的位姿测量算法及实验研究[J]. 航空学报, 2007, 28(4): 943-947.
- [3] 肖永利, 张琛. 运动物体位移及姿态参数的一种图像测量方法[J]. 机器人, 2001, 23(3): 266-269.
- [4] GONZALEZ R C, WOODS R E, EDDINS S L. Digital image processing using MATLAB[M]. New Jersey: Prentice Hall, 2003.
- [5] 马颂德, 张正友. 计算机视觉—计算理论和算法基础[M]. 北京: 科学出版社, 1998.
- [6] 李毅, 马颂德, 邱茂林. 计算机视觉中摄像机定标综述[J]. 自动

化学报, 2000, 26(1): 43-55.

- [7] HONG L, CHEN G. Segment-based stereo matching using graph cuts[C]// Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition: CVPR 2004. Washington, DC: IEEE Computer Society, 2004: 74-81.
- [8] WEI YI-CHEN, QUAN LONG. Region-based progressive stereo matching[C]// Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2004: 106-113.
- [9] 陈厚道, 周钢, 王从军, 等. 一种基于极线约束的激光条纹匹配算法[J]. 激光与光电子学进展, 2003, 40(4): 40-43.
- [10] 罗世民, 李茂西. 双目视觉测量中三维坐标的求取方法研究[J]. 计算机工程与设计, 2006, 27(19): 3622-3624.
- [11] DUNN F, PARBERRY I. 3D 数学基础: 图形与游戏开发[M]. 史银雪, 陈洪, 王荣静, 译. 北京: 清华大学出版社, 2005.