

文章编号: 0253-2239(2010)01-0153-10

# 基于视觉流形正则化的机器人实时定位新方法

吴 华 秦世引

(北京航空航天大学自动化科学与电气工程学院, 北京 100191)

**摘要** 提出了一种基于核主分量分析(PCA)正则化的机器人实时定位算法。此算法以半监督学习完成离线训练, 首先, 以机器人在其预置运动路径上采集到的畸变图像中的稀疏目标面积为观察数据, 将部分标定数据的坐标作为其标签, 然后以核 PCA 所揭示的低维视觉流形为正则化约束条件, 运用最小二乘方法估计无标签数据坐标。在线定位阶段, 利用调和函数估计在线采集到的数据坐标, 从而实现基于无标定单目视觉传感器的机器人在线定位。实验结果表明, 和其他常规的定位方法相比较, 提出的实时定位算法的计算复杂性小、定位精度高、实时性强, 能够满足工业机器人和医疗服务机器人等方面的实时定位要求。

**关键词** 机器视觉; 流形正则化; 核主分量分析; 面积特征

**中图分类号** TP242.6 **文献标识码** A **doi**: 10.3788/AOS20103001.0153

## New Approach to Robot Localization in Real-Time Based on Visual Manifold Regularization

Wu Hua Qin Shiyin

(School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China)

**Abstract** This paper presents a new approach to real-time robot localization using kernel principal component analysis (PCA) regularization. The proposed algorithms are formulated as a semi-supervised learning during offline training. Firstly, sparse area features are extracted from the images captured by the camera mounted on the robot which moves along a predetermined path, and labeled a part of the data with their coordinates. Then, the coordinates of the unlabeled data are estimated by least squares with constraint of regularized low dimensional visual manifold in kernel PCA. In online localization stage, harmonic functions are employed to predict new data coordinates so that the real-time robot localization can be implemented using uncalibrated monocular vision. A series of experiments manifest that the proposed algorithms can outperform other conventional methods with low computational complexity, high localization accuracy and well real-time performance, so as to meet the real-time application requirements of industrial robots and medical service robots.

**Key words** machine vision; manifold regularization; kernel principal component analysis; area feature

## 1 引 言

机器人自主导航中的同步定位与地图构建(SLAM)<sup>[1,2]</sup>使机器人能够借助各类传感器的信息不断估计自身的位置并创建环境地图。智能车辆、工厂、超市中的搬运机器人以及医疗服务等机器人

的大量应用使得 SLAM 系统的需求日益增长。这类机器人由一定的导航策略控制并能够在预置路径上运动并完成各类任务<sup>[3~6]</sup>。如果借助全球定位系统(GPS)或外部视觉监控设备往往会增加工程实施的成本和难度, 而且受到诸如 GPS 的信号强度和外部

**收稿日期**: 2009-04-09; **收到修改稿日期**: 2009-06-15

**基金项目**: 国家 863 计划(2006AA04Z207)、国家自然科学基金(60875072)、教育部博士点基金(20060006018)和中澳国际合作项目(2007DFA11530)资助课题。

**作者简介**: 吴 华(1981—), 男, 博士研究生, 主要从事机器视觉、机器人 SLAM 及机器学习等方面的研究。

E-mail: wishsand@gmail.com

**导师简介**: 秦世引(1955—), 男, 教授, 博士生导师, 主要从事空天飞行器的智能自主控制等方面的研究。

E-mail: qsy@buaa.edu.cn

部遮挡等客观条件的限制,不能为机器人提供快速准确的位置信息。因此,本文提出一种全新的机器人 SLAM 方法,能够解决按照预置路径机器人的自主定位问题,同时易于工程实现。

解决 SLAM 问题的方法有两类。第一类是以滤波方式定位的方法。主要以参考路标为观测量,同时结合机器人的运动模型递推机器人的位置。例如,Se 等<sup>[3]</sup>提出通过立体视觉和路标库的视觉 SLAM 方法。Labrosse<sup>[6]</sup>与 Murillo 等<sup>[7]</sup>利用图像中的特点作为路标估计机器人的位置。此外,通过三维重建的方法也能够求解 SLAM 问题<sup>[4,5]</sup>。Davison 等<sup>[8]</sup>利用扩展卡尔曼滤波递推机器人的位置,但是其算法的初始化和路标的识别决定着定位精度。Eade 等<sup>[9]</sup>利用粒子滤波并借助对很多小边缘的观察推断机器人的位置。但是由于特征点或小边缘的计算都会影响定位系统的处理效率,而且递推过程存在着由累计误差,因此这些方法不适用于在预置路径上往复运动的机器人。第二类 SLAM 方法以解决闭环合差以及全局定位为目标。例如,利用全景图像中旋转不变特征<sup>[10]</sup>和蒙特卡罗方法估计机器人的位置。但是该方法依赖光照条件不变的假设,适应性不强。文献[11]融合激光和摄像机获取的信息来检测闭环合差,但集成在定位系统中需要离线学习以及标定大量数据。文献[12]借助图模型学习采集到的地形图像进而获得机器人位置的后验概率分布,而且其方法能够学习新的地形图像。但是构造表观特征库需消耗时间长,而且不能明确地计算出机器人所在位置的坐标数值。文献[13]利用流形学习的方法估计机器人的位置,揭示了图像中潜在的低维流形,但是由于采用整幅图像作为学习样本,计算复杂度和时间消耗都很大。综上所述,第二类方法存在着特征计算耗时长、无法在线定位或计算位置坐标等问题。

针对在预置路径机器人的定位问题,利用视觉流形作为正则化项的原因;然后利用正则化项作为半监督优化的约束条件,给出机器人 SLAM 的方法,并提出离线学习和在线定位算法;最后完成实验和分析。本文提出的方法不仅能够推断机器人的位置坐标且不会产生闭环合差,同时能够实时在线定位。

## 2 潜在的低维视觉流形

未标定摄像头所采集的图像受到径向畸变的影

响,环境中物体的成像面积会随着机器人的运动而改变,因此多次采集的图像中刚体面积的变化与机器人的二维运动密切相关。提取面积变化需要通过预定颜色阈值分割出 6 个刚性观察目标在畸变图像中的成像面积。用该面积数据向量构造观察矩阵,进而通过核主分量分析(PCA)的方法对面积数据降维求得面积变化的主元。通过分析畸变图像中刚体面积变化与机器人运动的关系,解释了畸变图像中目标的面积变化在低维空间中以低维流形体现,即面积变化的主元,然后据此分析该低维流形存在的原因。

### 2.1 稀疏面积特征中的视觉流形

远小近大,阐述了视觉特征变化与位置变化的关系。数字图像中刚体面积的成像大小在一定条件下由镜头与刚体的距离决定。因此,若以图像中的物体成像面积为特征,然后通过统计方法揭示面积数据中引起变化的主要因素,即数据中方差变化最大的一些主方向。

提取一幅未矫正的畸变图像中刚体的面积数据可通过阈值化其颜色来实现。当然有很多图像分割或多目标检测的方法可以用来实现特定面积的划分和其像素数量的统计,但是为了平衡效率和划分精度,假设观察目标的颜色分布于色彩空间的不同区域以提高色块的分割精度。如图 1(a)所示,1~6号是大小和形状均不相同的单色色块,且每个色块为二维平面以确保面积的计算不会受到其自身遮挡的影响。由于只需统计面积的大小,所以刚体目标的形状并不限于长方形。

面积特征提取其实是多次采集刚体目标的面积并将其排列成矩阵以便后续运算。采集的第  $i$  号面积向量为第  $i$  幅图中刚体面积排成的向量,即

$$\mathbf{O}_i^T = [a_{i1}, a_{i2}, \dots, a_{in}], \quad (1)$$

式中包含  $n$  个预定目标的面积。使用 6 个目标,因此  $n=6$ 。并按照图 1(a)中的顺序依次排入  $\mathbf{O}_i^T$  向量中的每一维。采集  $m$  幅图像所得到的观察矩阵为

$$\mathbf{M}^T = [\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_m], \quad (2)$$

式中每个  $\mathbf{O}_i^T$  对应着每幅图中的面积向量。故  $\mathbf{M}$  中的每行对应着每幅图的面积向量,每列对应同一个目标在不同时刻的面积大小。由此构造的观察矩阵将用于核 PCA 中核矩阵的计算以及后续在线定位中的数据权值的计算。

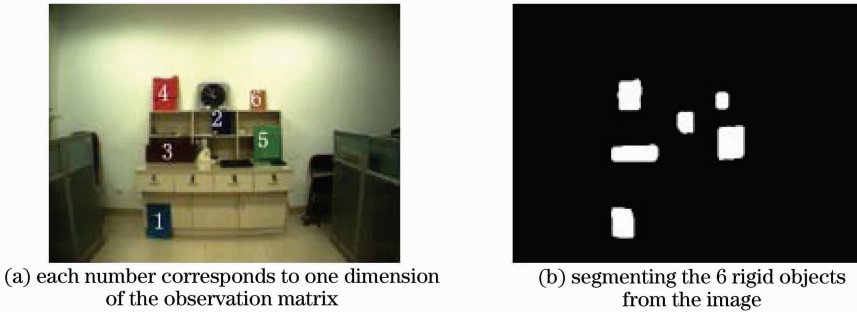


图 1 分割出来的目标被依次排列组成观察向量

Fig. 1 Objects are segmented and arranged as a observation vector

如图 2 所示,利用核 PCA(多用于模式识别中的数据维数简约<sup>[14]</sup>)对  $316 \times 6$ (提取 316 幅图像中的 6 维面积向量)降维。按照特征值降序排列后,取前 6 个特征值比较,如图 2(a)所示。图 2(b)为 Y,其前三个特征值占有所有特征值总和的百分比中深绿与浅绿、扇形表示了前两个特征值反映核矩阵约 99% 的方差。根据以前两个特征值为坐标轴,绘制所有的数据点如图 2(c)所示,从绿色到黄色点的走

向趋势可以看到核 PCA 所得到的前 2 个主元反映了机器人在平面上的二维运动。图 2(d)为原始路径(起始点在左上角),比较图 2(c)与 (d)可见核 PCA 揭示了数据中潜在的低维流形分布,在一定程度上反映了面积变化与真实路径间的非线性关系。两幅图的坐标尺度不一致是因为核 PCA 所得到的主元与真实运动的坐标变化并不完全等价。

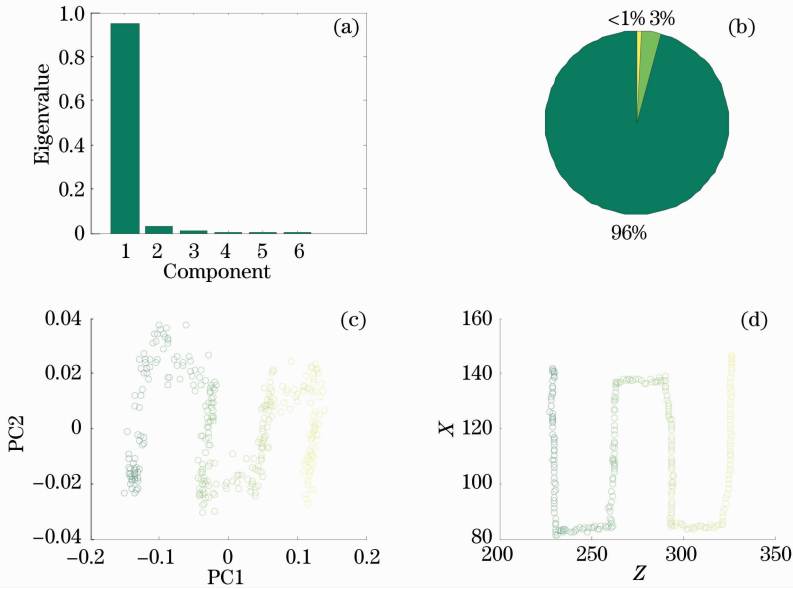


图 2 通过核 PCA 降维观察矩阵( $\sigma=200$ )

Fig. 2 Reducing dimensionalities of the observation matrix using kernel PCA ( $\sigma=200$ )

如图 3(a)和图 3(b)所示,有畸变图像与无畸变图像观察矩阵的降维结果比较可以得出核 PCA 仅能揭示未做畸变矫正的图像中的面积观察矩阵的主元。因此,核 PCA 能够以简约维数的方式提取有畸变图像里目标面积观察矩阵中的低维流形结构,在一定程度上揭示了机器人的位置变化。也就是说通过降维观察矩阵能够估计机器人的面积变化与机器人位置变化的非线性相关。而且这种非线性关系仅存在于畸变图像中的刚体面积数据里,无法在无

畸变图像中找到。

### 2.2 低维变量存在性分析

主元与机器人二维运动的非线性关系可以通过无畸变图像中刚体面积与机器人位姿变化的关系来说明。借助相机几何成像原理和畸变原理可导出畸变图像中刚体面积与机器人位姿变化的非线性关系。由此可推知,机器人在平面上沿 X,Z 轴不断运动所采集的畸变面积反映了机器人平移时的二维坐标变化。

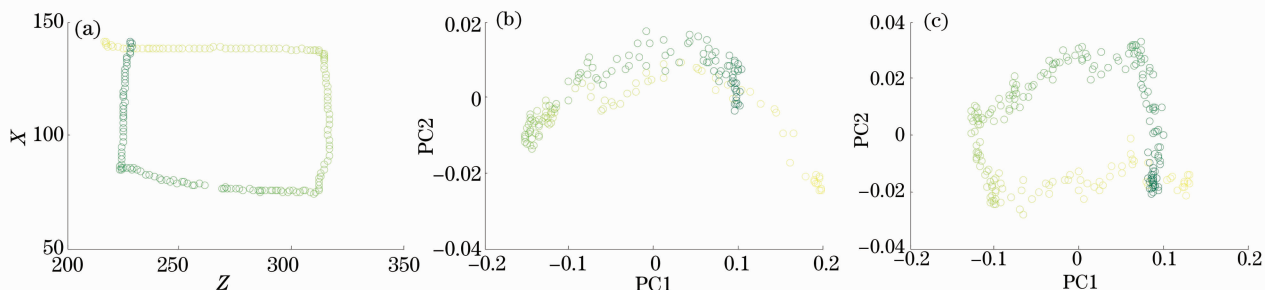


图 3 有畸变和无畸变图像序列观察矩阵的前两个主元示意图。(a) 机器人真实运动路径; (b) 有畸变; (c) 无畸变  
Fig. 3 Top 2 PC of the observation matrices generated from rectified images and distorted images. (a) robot's true motion path; (b) rectified images; (c) distorted images

在畸变系数  $k < 0$  的畸变图像中, 远离图像中心的点偏离其在无畸变图像中的位置越远, 即图像呈枕型失真。文献[15]指出径向畸变是最主要的畸变因素, 且无畸变图像中的点  $(x_u, y_u)$  在畸变图像中的坐标  $(x_d, y_d)$  为

$$\begin{aligned} x_d &= x_u + \delta_x(x_u, y_u) = x_u + kx_u(x_u^2 + y_u^2) \\ y_d &= y_u + \delta_y(x_u, y_u) = y_u + ky_u(x_u^2 + y_u^2) \end{aligned}, \quad (3)$$

式中  $k$  是畸变系数, 在此仅保留到二次方项。因为文献[16]指出保留到二次方项在绝大部分的机器视觉应用中已经能够达到所需精度。不失一般性, 本文假设图像只受到径向畸变影响。

由数字相机成像的几何关系可以计算得到机器人在 3 维空间中的某个位置  $(X + T_x, Y + T_y, Z + T_z)$  观察到的静态刚体目标在图像上的位置  $(x_u, y_u)$  为

$$\begin{aligned} x_u &= \frac{f}{Z + T_z}(X + T_x) \\ y_u &= \frac{f}{Z + T_z}(Y + T_y) = \frac{f}{Z + T_z}Y \end{aligned}, \quad (4)$$

式中  $[T_x, T_y, T_z]$  构成机器人运动过程中的平移矩阵, 描述机器人在不自转的情况下, 沿每个方向平移的量。  $f$  为镜头焦距,  $(X, Y, Z)$  是机器人的初始位置。由于机器人在平坦的地面运动, 此时可认为  $T_y = 0$ 。此外, 对于离散采样的数字图像而言, 其面积为

$$S_{\text{img}} = \iint_D dx_u dy_u \approx \sum_i^n \frac{f^2}{(Z_i + T_z)^2} \Delta X_i \Delta Y_i. \quad (5)$$

因此, 由(5)式可知, 面积的变化仅与  $T_z$  相关。当机器人仅沿  $X$  方向平移运动观察固定的静态刚体目标时, (4) 式说明目标在图像中的位置仅  $x_u$  变化, 此时目标面积没有变化。也就是说, 当机器人平行于  $X$ - $Y$  平面运动时, 刚体的成像面积不会改变。但是就畸变图像中的刚体面而言, 当  $T_x$  改变而  $T_z$  不变时, 面积却会随着  $T_x$  而改变。面积的变化与机器人位置的变化密切相关。这一点可通过(6)式看

出。在畸变成像区域  $D'$  内, 畸变图像的面积是

$$S'_{\text{img}} = \iint_{D'} dx_d dy_d = f(T_x, T_z). \quad (6)$$

由(3)式和(4)式可知, 存在非线性函数  $f$  能将给定的  $T_x, T_z$  映射成刚体成像面积。所以, 畸变图像的面积与机器人的位置坐标变化存在着非线性的关系。只要将核 PCA 所揭示的数据潜在关系作为正则化项来约束机器人的定位模型, 即可有效地利用数据的低维信息来估计机器人所在的位置。

### 3 基于流形正则化的机器人位姿估计模型及算法

如图 2 所示, 仅仅依靠这些低维的流形信息无法明确估计机器人的实际位置。将机器人 SLAM 看作半监督的坐标学习, 即将有标签的数据(即面积向量)拟合到其对应的标签(位置坐标)上, 进而求得其回归模型。运用核 PCA 作为流形正则化项可以有效防止过拟合, 通过低维流形信息使拟合曲线更光滑, 并能很好地估计无标签数据的实际空间坐标。一般而言, 流形正则化采用的是图拉普拉斯(Graph Laplacian)正则化项。由于核 PCA 正则化项与图拉普拉斯正则化项在以完全图为条件的情况下是等价的, 故针对机器人定位建立以最小二乘为目标函数, 以核 PCA 为正则化项的半监督模型。进而提出离线学习和在线定位的两步算法。

#### 3.1 基于核 PCA 正则化的最小二乘法

文献[17~19]指出了多种目标函数形式, 其中最小二乘函数非常适合曲线回归。如果拉普拉斯本征映射<sup>[20]</sup>  $L$  与核 PCA 中的核矩阵  $W$  建立关系, 那么就能够将核 PCA 等价于拉普拉斯本征映射来作为最小二乘的正则化项。核矩阵计算所有数据间的权值, 而拉普拉斯计算  $n$  个近邻间的权值。于是, 一

个完全图的权值矩阵就是核 PCA 中的核矩阵。拉普拉斯本征映射是以拉普拉斯矩阵  $L$  的最小  $m$  个非负特征值为低维嵌入,而核 PCA 是以核矩阵的前  $m$  个最大特征值为低维嵌入。因此核矩阵与拉普拉斯矩阵的特征值对应的特征向量张成相同的特征空间。即  $L = I - D^{-1/2} W D^{-1/2}$ , 其中  $W$  为连接权值矩阵,  $D_i = \sum_j W_{ij}$ 。

核 PCA 正则化形式可写为

$$f^* = \arg \min_{f \in H_K} \sum_{i=1}^l (f - c_i)^2 + \gamma f^T (I - D^{-1/2} W D^{-1/2}) f, \quad (7)$$

式中  $H_K$  为希尔伯特再生空间,  $c_i = (X_i, Z_i)$  是第  $i$  个观察数据的空间坐标,  $f$  是待优化的观察数据坐标,  $\gamma$  为正则化系数。则此凸优化问题的解为

$$f^* = (P + \gamma I - \gamma D^{-1/2} W D^{-1/2})^{-1} C, \quad (8)$$

式中  $P = \text{diag}(p_1, p_2, \dots, p_l, 0, 0, \dots, 0)$  为对角阵, 且前  $l$  个有标签数据对应的每个  $p_i = 1$ 。  $C = [c_1, c_2, \dots, c_l, 0, 0, \dots, 0]$  是标签矩阵, 当模型拟合坐标的  $X$  轴数值时,  $c_i = X_i$ ; 当模型拟合坐标的  $Z$  轴数值时,  $c_i = Z_i$ 。于是, 给定有标签样本的坐标后, 便可通过样本的权值矩阵  $W$  求得无标签数据的坐标。将有标签和无标签的数据坐标用于实时采集数据的坐标估计中, 即可提高计算的效率和定位的精度。

### 3.2 基于调和方程的实时位姿估计

通过估计无标签数据的坐标并结合有标签数据的坐标就能够为基于调和方程的实时位姿估计提供充足的参考数据。由于拉普拉斯本征映射就是调和方程的一种表示, 且核 PCA 与拉普拉斯映射等价, 因此调和方程所得到的估计与核 PCA 的估计一致<sup>[21]</sup>。利用这一性质, 调和方程可以通过实时观察到的面积向量与训练过程中的观察矩阵来快速准确地估计机器人所在的位置。

通过(9)式可实时计算机器人的位置:

$$f_{\text{new}} \approx \frac{1}{d_i} \sum_j W'_{ij} f_j, \quad (9)$$

式中  $d_i = \sum_j W'_{ij}$ 。所以在预置路径上运动的机器人的坐标可以通过加权平均训练数据的坐标求得。其中的权值矩阵是新数据与训练数据间的核矩阵  $W'$ 。利用(9)式即可估计机器人的坐标  $(f_{\text{new}}^X, f_{\text{new}}^Z)$ 。

基于调和方程的机器人在线定位过程将离线训练阶段所估计的无标签数据坐标运用到实时计算机器人位置过程中, 不仅使得计算过程非常简便, 而且从理论上保持了算法的收敛性。根据低维流形存在性条件, 只要数据采集自训练过程所经路径上即可

求得该数据的实际坐标值。

### 3.3 离线地图学习以及在线实时定位算法

根据上述机器人定位方法, 提出了两步定位算法。首先离线学习地形, 通过获得的观察数据作为训练样本, 结合部分已知坐标的数据, 从而确定无标签数据的坐标。通过计算得到的坐标和观察矩阵, 便能在第二阶段利用调和方程实时计算出机器人的位置。前一个阶段离线完成机器人定位过程中最耗时模型学习。故在线阶段的实时定位时间开销大大降低。两部分算法实现过程简单, 具有理论依据和很好的收敛性, 降低了工程应用的风险。

离线学习过程中, 机器人按照固定路径运动, 不断观察目标面积变化记录面积数据并构造观察矩阵。同时置顶相机记录机器人的位置坐标作为机器人每个观察数据的标签。这些有标签的数据结合部分无标签的数据一起构成训练样本集作为第一阶段离线学习的数据。输入为观察矩阵  $M$ , 核参数  $\sigma_x, \sigma_z$ , 正则化参数  $\gamma_x, \gamma_z$  以及真实坐标矩阵  $C$ 。输出为无标签数据的坐标矩阵  $f$ 。计算观察数据的径向基函数(RBF)核矩阵,  $K_x = \exp(-\sigma_x \|\Delta O\|^2)$ ,  $K_z = \exp(-\sigma_z \|\Delta O\|^2)$ , 并令  $W_x = K_x, W_z = K_z$ ; 计算  $L_x = I - D_x^{-1/2}, L_z = I - D_z^{-1/2} W_z D_z^{-1/2}$ ; 计算  $f_x = (P - \gamma L_x)^{-1} C_x, f_z = (P - \gamma L_z)^{-1} C_z$ 。

在线定位过程中, 机器人在预置路径上运动并记录目标成像面积数据, 与离线训练时使用的观察数据相结合计算得到核矩阵, 然后加权平均训练数据的坐标值便求得机器人所在位置的坐标。输入为新观察向量  $O_{\text{new}}$ , 训练数据坐标  $(f_x, f_z)$ , 核参数  $\sigma_x, \sigma_z$ , 训练数据观察矩阵  $M$ 。输出为  $O_{\text{new}}$  的坐标  $f_{\text{new}}$ 。计算  $O_{\text{new}}$  与所有训练数据的 RBF 核矩阵  $K_x = \exp(-\sigma_x \|\Delta O\|^2)$  以及  $K_z = \exp(-\sigma_z \|\Delta O\|^2)$ , 并令  $W_x = K_x, W_z = K_z$ ; 计算  $f_{\text{new}} = \left( \frac{1}{d_x} \sum W_x f_x, \frac{1}{d_z} \sum W_z f_z \right)$ 。

算法通过核 PCA 正则化模型计算无标签数据坐标进而在线估计机器人的位置。其中核参数  $\sigma_x, \sigma_z$  和正则化参数  $\gamma_x, \gamma_z$  决定估计的精度和对流形分布的依赖程度, 经过多次参数评估确定四个参数的取值分别为  $\sigma_x = 267.06, \sigma_z = 454.16, \gamma_x = 0.015, \gamma_z = 0.01$ 。

## 4 实验分析及结果

实验比较不同特征组合并给出最佳的观察向量维数和选用的观察目标组合; 通过改变样本数量分析了最优的有标签数据和无标签数据的数量组合; 最后



以最小均方误差为度量定量评估本文提出的 SLAM 方法的定位精度,并给出算法的时间复杂度。

实验环境如图 4 所示,机器人面向摆有色板的书桌[如图 4(d)]保持平移运动。为了获得较大的视野和显著的径向畸变,实验选用图 4(a)中紫色圈内双目摄像视觉系统的左摄像头为图像获取装置。机器人

在长 420 cm,宽 240 cm 的室内地面以 10 cm/s 的速度运动。相机顶部的双色板是为图 4(c)里置顶摄像机识别计算机器人的位置而放置的。本文所述的机器人控制算法、图像采集和处理以及网络通信均实现在 Grandar 能力风暴移动机器人上。

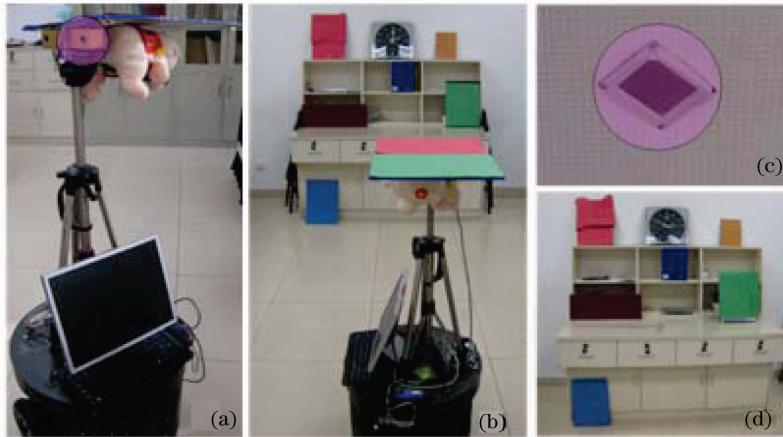


图 4 实验环境及系统设备

Fig. 4 Experimental conditions and system equipments

实验系统由机器人运动控制、位置标定以及数据采集三部分组成。机器人按照预先规划的路径运动。如图 5 所示,机器人自带的摄像采集到的图像如(a1)~(a3)所示,通过分割提取其中的色块

[(b1)~(b3)],与此同时机器人通过无线网络传输信号给连接置顶摄像机的计算机通知其采集图像,并识别出双色板的位置,如(c1)~(c3)所示。利用畸变矫正参数计算出机器人的实际坐标,绘制在二

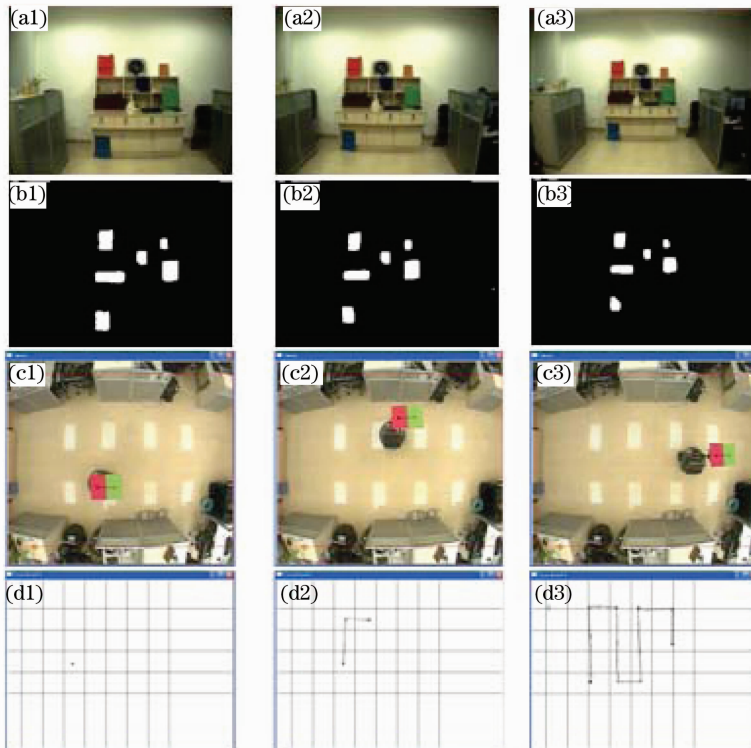


图 5 面积向量的计算以及机器人位置的测量

Fig. 5 Computing the area vectors and measuring the ground truth of the robot location

维平面上如(d1)~(d3)所示。为了同步机器人的数据采集和其位置标定,使用 Visual C++ 编制一套基于无线网络访问接收点(AP)和机器人自带的无线网卡的网络同步数据采集程序。机器人采集每个面积的时间与其位置的计算是完全同步的。为了加快面积采集的速度减小光照强度的影响,采用色彩饱和值(HSV)颜色空间模型将机器人视觉传感设备所采集到的图像映射到 HSV 空间,并通过预定的  $H$  和  $V$  分量阈值分割出 6 个色块,统计像素个数作为其面积值。采用相同的处理方法,与置顶摄像机相连的计算机即可跟踪双色板的中心位置。运用多项式拟合可将置顶相机所采集的畸变图像映射到无畸变图像,然后借助相似三角形的原理即可计算得到双色板中心点的空间位置。以图 4(d)中的目标书桌右侧为原点计算机人的空间位置。系统处理流程如图 5 所示,机器人在 3 个位置采集到三幅图像[图 5(a1)~(a3)],分割后得到每个区域的面积值[图 5(b1)~(b3)]。此时,位置标定模块提

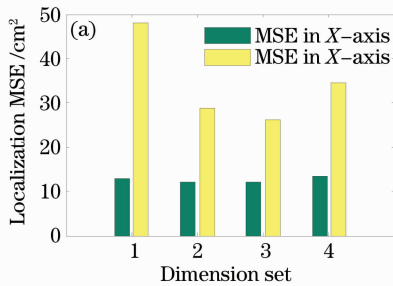


图 6 (a) 4 种组合的定位均方误差;(b) 实际定位结果

Fig. 6 (a) Localization MSE (mean squared error) of 4 different combinations; (b) actual localization results

#### 4.2 有标签和无标签样本数量

选择一定比例的有标签和无标签数据能减轻计算负担,并且减少标定过程的成本。对于所述的定位系统,采用的有标签数据越多越能够提高定位精度,但是在实际应用中标定真实坐标将花费很多的计算资源和人力物力。所以选择合适比例的有标签和无标签数据能够保证计算精度同时使得标定消耗尽可能少。

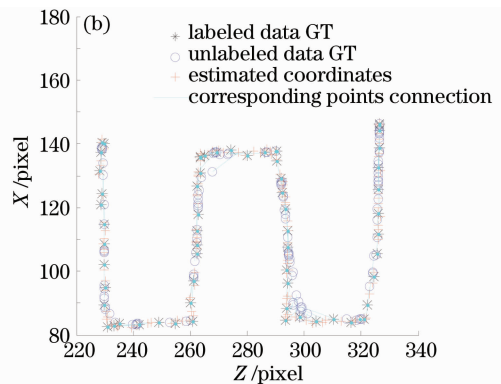
首先固定 64 个均匀抽取的有标签数据,然后分 26 次随机抽取无标签数据。从 300 到 50 个数据,每次抽取时都少抽 10 个数据。将此过程反复进行 5 次,结果如图 7 所示。由图 7(a)中第 22 号数据集可知,当选择 90 个无标签数据,定位的均方误差最小。此时无标签数据量接近有标签数据,无法满足

取置顶相机中双色版的图像坐标并计算其空间位置[图 5(c1)~(c3)],最终将真实路径绘制在窗口中并存储坐标值[图 5(d1)~(d3)]。

#### 4.1 特征选择

6 个观察目标组成的面积数据向量具有部分的冗余特征,选择相关性最小的特征不仅能够提高坐标计算的精度而且还能减少时间和空间的消耗。因此特征选择是数据预处理中最重要的步骤。

为了更好地评估不同特征组合的定位精度,采用不同维的特征构造 4 组特征组合:(a)1,3,4 维,(b)1,4,6 维,(c)1,2,5 维,(d)所有维。分别计算定位误差,可见图 6(a)中 3 号组合的定位误差最小,因此选择 3 号组合作为最佳特征组合。其定位结果与实际路径的比较如图 6(b)所示,其中红色十字为无标签数据坐标的估计位置,青绿色连线越短说明定位精度越高。挑选最优的特征能够减少估计误差,提高数据处理速度,并且在一定程度上排除了面积数据计算过程中由于光照和视角改变所引入的噪声。



在线估计时的数据量需求(一般大于 200 个数据)。因此,本文选择第 17 号数据集,140 个无标签数据,其定位精度仅次于第 22 号数据集。

固定 140 个无标签数据,从有标签数据中随机抽取一定量的数据(数据量从 106 到 37)。重复此过程 5 次,结果如图 8 所示。很明显,选择第 1 组数据能够获得最低的定位误差,然而需要 106 个标定数据,给标定过程带来过大负担。因此选择第 3 号数据集,既能保证定位精度而且 64 个有标签数据不会消耗过多的人力物力标定部分数据的真实位置,且不会给坐标标定系统带来繁重的负担。

所以,选择 64 个有标签数据,140 个无标签数据提供给离线训练模型并使用离线算法求得无标签数据的坐标。选择该数据集时的定位误差小于

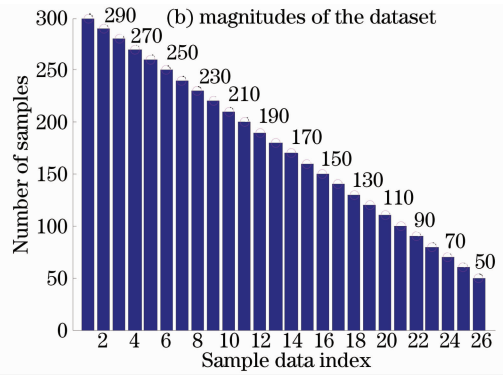
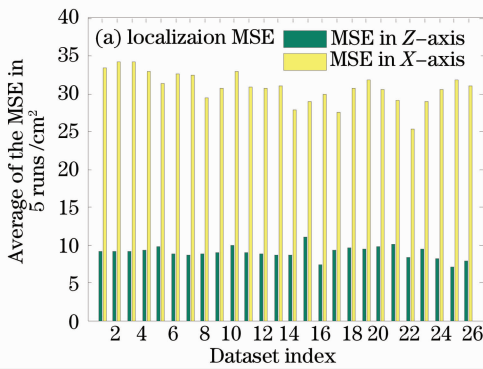


图 7 固定有标签数据量,选择无标签数据量

Fig. 7 Fix labeled data magnitude, select the magnitude of unlabeled data

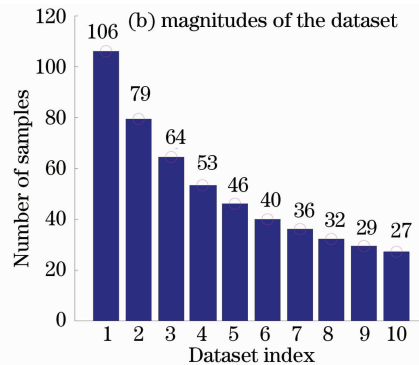
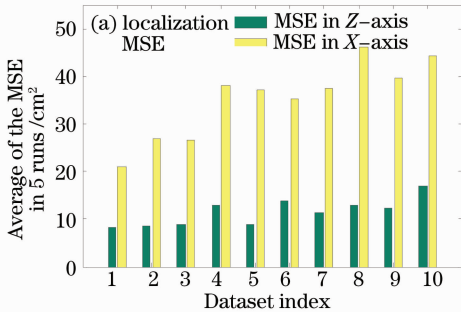


图 8 固定无标签数据量,选择有标签数据量

Fig. 8 Fix unlabeled data magnitude, select magnitude of labeled data

30 cm<sup>2</sup>。综上,当有标签数据与无标签数据的比例在 1:2.2 时可以平衡计算精度和标定消耗。

### 4.3 在线定位精度

评估定位误差是在 4 种不同路径分别采集 500 个测试数据,再将计算得到的无标签坐标带入在线算法。所得的平均在线定位误差如图 9 所示。该定位结果是从 500 个测试数据中随机抽取 100 个测试样本,并重复抽取 5 次计算得到的平均结果。其中 1 号路径的 X 方向误差大于 Z 方向误差,是由于机器人在 X 方向的运动往复两次,精确拟合所需

数据较拟合 Z 方向运动多。

实际定位结果如图 10 所示。图 10(a1)~(a4)为真实路径,图 10(b1)~(b4)是第 3 次抽样的定位结果(省略的图例说明与图 6 相同),图 10(c1)~(c4)是 5 次定位均方误差的比较(蓝色圆点是 X 方向定位均方误差,绿色三角是 Z 方向定位均方误差)。故平均 5 次的统计结果,4 种路径定位误差分别为:X 方向的均方误差小于等于 25 cm<sup>2</sup>,Z 方向的均方误差小于 20 cm<sup>2</sup>。此较图 10(b1)~(b4)可见红色估计点几乎完全位于路径内,而对于复杂路径 2 而言,仅有 3 个路径外的错误估计点。青绿色线都很短,说明该方法能够精确估计机器人的真实坐标。

### 4.4 算法效率

由于采用离线训练完成无标签数据坐标的估计,因此在线定位的算法耗时非常少。离线训练中确定无标签数据坐标的时间为 280 ms,在线定位耗时 15 ms 如表 1 所示。因此,对于 30 Hz 采样频率的视觉传感设备,该方法能够实时处理视频数据并确定机器人的位置。

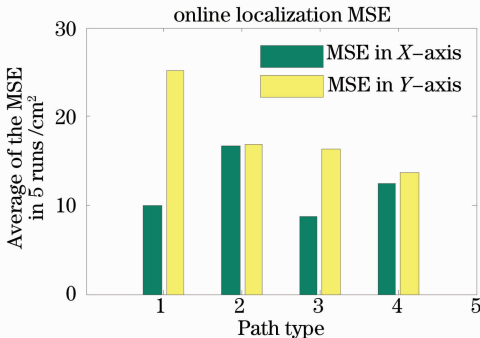


图 9 4 种路径在线定位均方误差

Fig. 9 Online localization MSE of 4 paths



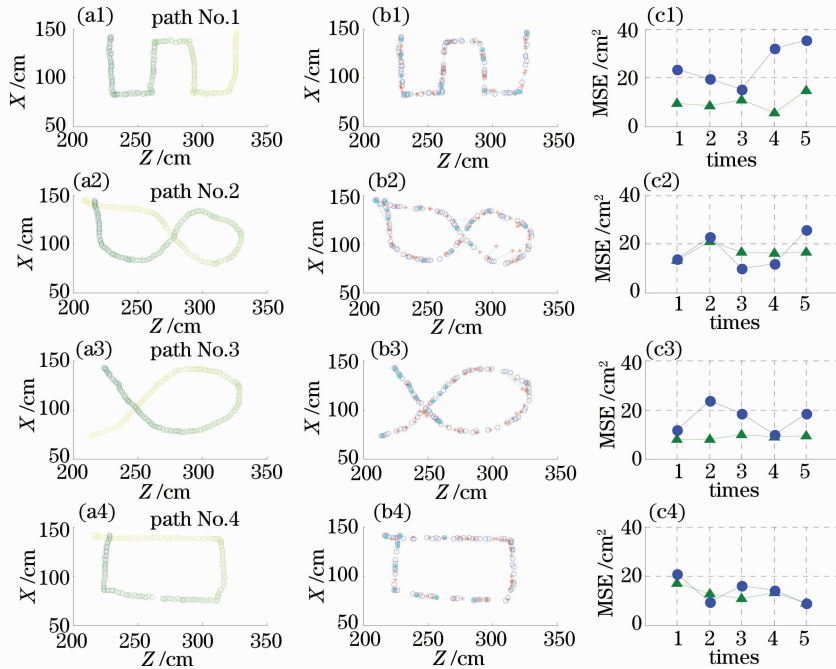


图 10 4 种路径在线定位结果

Fig. 10 Online localization results of 4 paths

表 1 在线定位时间

Table 1 Online localization time

Algorithm steps	Time consumption /ms
Obtaining observation vector (excluding capture time)	2 (C++)
Measuring distances to all the neighbors (5 features, 140 unlabeled data, 64 labeled data)	7 (Matlab 2008a)
Computing kernel matrix	6 (Matlab 2008a)
Computing coordinates	2 (Matlab 2008a)
Total	15

面积分割和在线定位的时间复杂度是在线定位的关键。从 3 通道宽为  $W$  高为  $H$  的彩色图像中分割 6 个区域并计算面积的时间复杂度为已知  $n$  个训练数据, 定位  $d$  维观察向量的时间复杂度为  $O(dn^3)$ 。因此, 训练数据量  $n$  以及图像的大小决定了运算速度。实验从 4 个方面详细阐述了该方法的实施过程、每个阶段的参数确定以及运算效率, 并以定量的方式阐明了本文的定位方法计算速度快, 算法实现简单, 要求标定的有标签数据少等特点。

## 5 结 论

提出的基于核 PCA 正则化的机器人定位方法是一种高效且具有实际应用价值的机器人 SLAM 方法。以低维视觉流形的数据分布为约束条件, 通过最小二乘法估计未标定数据之坐标, 以半监督学习实现机器人定位, 不仅可减轻坐标标定负担, 而且与在线阶段实时位置估计模型相一致。其显著特点是, 有效解决了滤波算法中的闭环合差(loop closure)和地

图匹配算法之运算速度慢、无法实时定位等问题。实验结果表明, 机器人在  $420\text{ cm} \times 240\text{ cm}$  的室内地面按照 4 种不同固定路径运动时, 定位均方误差均小于  $25\text{ cm}^2$ , 且能够以  $30\text{ Hz}$  的速度实时处理视觉传感器所采集的图像, 并在  $15\text{ ms}$  内给出位置估计, 满足工程应用中的实时性和准确性要求。在进一步研究中将考虑改进优化目标函数的形式, 使得算法不仅能够拟合有标签数据, 而且还能通过  $L_1$  范数或者  $L_2$  范数的约束条件, 降低坐标估计的方差, 使估计更准确。

## 参 考 文 献

- 1 T. Bailey, H. Durrant-Whyte. Simultaneous localisation and mapping (SLAM): part II[J]. *IEEE Robotics & Automation Magazine*, 2006, 2(13): 99~110
- 2 H. Durrant-Whyte, T. Bailey. Simultaneous localization and mapping: part I[J]. *IEEE Robotics & Automation Magazine*, 2006, 2(13): 99~110
- 3 S. Se, D. G. Lowe, J. J. Little. Vision-based global localization and mapping for mobile robots[J]. *Robotics, IEEE Transactions on*, 2005, 3(21): 364~375
- 4 Shen Yehu, Liu Jilin, Du Xin. Simultaneous three-dimensional environment reconstruction and localization based on monocular

- vision[J]. *Acta Optica Sinica*, 2008, **28**(5): 907~914
- 沈晔湖, 刘济林, 杜 歆. 单目视觉的同时三维场景构建和定位算法[J]. *光学学报*, 2008, **28**(5): 907~914
- 5 Xu Gang, Zhang Wenming, Lou Fengwei *et al.*. A new 3D reconstruction method based on gray similarity of the projecting grid points[J]. *Acta Optica Sinica*, 2008, **28**(11): 2175~2180
- 徐 刚, 张文明, 楼凤伟 等. 基于网格点投影灰度相似性的三维重建新方法[J]. *光学学报*, 2008, **28**(11): 2175~2180
- 6 F. Labrosse. The visual compass: performance and limitations of an appearance-based method [J]. *J. Field Robotics*, 2006, **10**(23): 913~941
- 7 A. C. Murillo, J. J. Guerrero, C. Sagues. SURF features for efficient robot localization with omnidirectional images[C]. IEEE International Conference on Robotics and Automation, 2007: 3901~3907
- 8 A. J. Davison, I. D. Reid, N. D. Molton *et al.*. MonoSLAM: Real-time single camera SLAM [J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2007, **6**(29): 1052~1067
- 9 E. Eade, T. Drummond. Edge landmarks in monocular SLAM [J]. *Image Vision Comput.*, 2008, In Press, Corrected Proof)
- 10 F. Linäker, M. Ishikawa. Real-time appearance-based Monte Carlo localization[J]. *Robotics and Autonomous Systems*, 2006, **3**(54): 205~220
- 11 K. L. Ho, P. Newman. Loop closure detection in SLAM by combining visual and spatial appearance [J]. *Robotics and Autonomous Systems*, 2006, **9**(54): 740~749
- 12 M. Cummins, P. Newman. FAB-MAP: probabilistic localization and mapping in the space of appearance[J]. *Int. J. Rob. Res.*, 2008, **6**(27): 647~665
- 13 H. Wu, S. Y. Qin. A new method of distance estimation for robot localization in real environment based on manifold learning [C]. Proc. 2007 Internat. Conf. on Wavelet Analysis and Pattern Recognition, Beijing, China: 2007: 589~590
- 14 Yuan Weiqi, Qu Xiaofeng, Ke Li *et al.*. PCA reconstruction error palmprint recognition [J]. *Acta Optica Sinica*, 2008, **28**(10): 1903~1909
- 苑玮琦, 曲晓峰, 柯 丽 等. 主成分分析重建误差掌纹识别方法 [J]. *光学学报*, 2008, **28**(10): 1903~1909
- 15 R. Hartley, A. Zisserman. Multiple View Geometry in Computer Vision[M]. University of Oxford, UK: Cambridge University Press, 2004: 189~193
- 16 R. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses [J]. *Robotics and Automation, IEEE J.* 1987, **4**(3): 323~344
- 17 M. Belkin, I. Matveeva, P. Niyogi. Regression and regularization on large graphs[C]. Proc 17th Annual Conference on Learning Theory. 2004: 824~831
- 18 M. Belkin, P. Niyogi, V. Sindhwani. Manifold regularization: a geometric framework for learning from labeled and unlabeled examples [J]. *The J. Machine Learning Research*, 2006, **7**: 2399~2434
- 19 V. Sindhwani, M. Belkin, P. Niyogi. Geometric Basis of Semi-Supervised Learning[M]. MIT Press, 2006: 209~226
- 20 M. Belkin, P. Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering[J]. *Advances in Neural Information Processing Systems*, 2002, **14**: 585~591
- 21 X. Zhu, Z. Ghahramani, J. Lafferty. Semi-supervised learning using Gaussian fields and harmonic functions[C]. ICML-03, 20th Internat. Conf. on Machine Learning, Washington, DC. 2003: 912~919