

# 基于激光和视觉信息的机器人 目标跟踪方法

闵华清, 黄欣欣, 罗荣华

(华南理工大学 计算机科学与工程学院, 广东 广州 510006)

**摘 要:**文中介绍了一种利用移动机器人的激光信息和摄像头信息实时跟踪目标的方法。实现了对人的准确迅速的跟踪。通过大量提取照到人的双脚的激光特征作为样本集,描述了一种基于条件随机场(CRF)模型的 Rao-Blackwellized particle filter(RBPF)算法,CRF 的 observation potential 可以直接从样本数据中获得。采用类似栅格滤波方法计算样本的后验概率。RBPF 算法根据后验概率进行权值的更新和采样实现对激光特征的实时跟踪,从而实现人的跟踪。根据人的位置信息可以确定人在摄像头图像窗口中的大概位置,提取该位置的 SURF 特征,从而获得人在图像中的精确位置。

**关键词:**条件随机场;栅格滤波;Rao-blackwellized particle filter;SURF

中图分类号:TP31

文献标识码:A

文章编号:1673-629X(2010)04-0113-04

## Robot Target Tracking Approach Based on Laser and Vision Information

MIN Hua-qing, HUANG Xin-xin, LUO Rong-hua

(Department of Computer Science and Engineering, South China University of Technology,  
Guangzhou 510006, China)

**Abstract:** Introduce a real-time target tracking method for a mobile robot using laser range data and camera images. It can track a person quickly and accurately. Obtained plenty of laser features which hit the person's legs as exemplars set, describe a Rao-Blackwellized particle filter(RBPF) algorithm based on CRF model, where the observation potentials are learned from data. In order to compute the posterior the grid filtering is used. The RBPF algorithm updated the weights and sample based on the posterior to realize the real-time tracking. According to the position of the person can estimate the person's position in the image window, get the precision position of the person by computing the SURF features in this area.

**Key words:** conditional random field; grid filter; RBPF; SURF

## 0 引言

现在已经有多种机器人应用在办公室<sup>[1-3]</sup>、超市<sup>[4]</sup>、医院<sup>[5]</sup>和博物馆<sup>[6,7]</sup>。跟踪是机器人特别是移动机器人要具备的一项非常重要的能力,这里面包含了目标物的探测、识别和跟踪等具有挑战性的任务。

在文献[8~10]提出了基于激光的目标跟踪方法。这种传感器的优点是能够得到准确的距离信息,但是它缺少视觉信息不能将你要跟踪的目标与其他目标区分开来,摄像头可以提供这样的信息,基于图像信息的人脸识别、皮肤分割、轮廓匹配已经得到很好的发展。

基于 Kalman 滤波和粒子滤波的运动物体探测方法已经使用了。在非线性和非高斯的情况下,粒子滤波比 Kalman 滤波的表现要好,但是单靠视觉信息要机器人进行跟踪很难,因为机器人的很小的动作可能导致图像不稳定,此外还有光照条件的变化也会对图像造成非常大的影响。

传感器和图像信息的融合已渐渐成为研究的热点,在文献[11]提出一种用粒子滤波将图像数据和声纳数据融合的机器人目标跟踪方法。将粒子滤波应用于图像帧需要非常大的计算量和时间,减少粒子的数目可以相应减少计算量的问题,但是有可能导致跟踪目标的丢失。文献[12]提出了一种基于概率模型的融合激光和视觉信息的跟踪算法。这种算法的主要缺点就是需要大量的训练样本,而且训练样本不太容易得到。而且,图像信息采用的是轮廓信息过于简单,不具

收稿日期:2009-07-18;修回日期:2009-10-15

基金项目:广州市科技攻关项目(2007Z32D3151)

作者简介:闵华清(1956-),男,教授,博士生导师,研究方向为机器人相关领域。

备尺度不变和旋转不变性。

文中基于条件随机场实现对人的双脚状态的跟踪。人的图像跟踪采用的是 SURF 特征,这种特征比 SIFT 特征快<sup>[13]</sup>,而且状态非常稳定具有尺度不变性和旋转不变性。

## 1 条件随机场状态跟踪

### 1.1 条件随机场

图 1(a)描述了在跟踪中广泛应用的 HMM,  $\theta$  表示在  $t$  时刻的状态,  $I^t$  表示  $t$  时刻的观察信息。该模型描述的状态和观察的联合分布如下:

$$p(\theta^{0 \cdots T}, I^{1 \cdots T}) = p(\theta^0) \prod_{t=1}^T [p(\theta^t | \theta^{t-1}) p(I^t | \theta^t)]$$

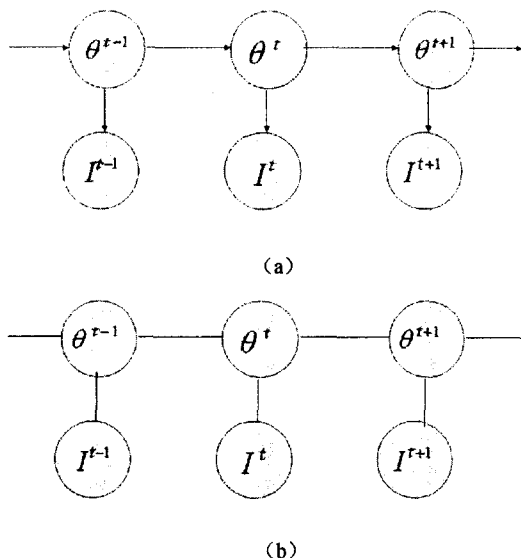


图 1 HMM 和 CRF 模型图

图 1(b)描述了一个链式 CRF。HMM 是在状态和观察的联合分布上进行建模, CRF 直接在状态的条件分布上面进行建模。

$$p(\theta^{0 \cdots T}, I^{1 \cdots T}) = \frac{1}{Z} p(\theta^0) \prod_{t=1}^T [\phi(\theta^t | \theta^{t-1}) \phi_0(\theta^t)]$$

其中  $Z$  是一个归一化常量,  $\phi(\theta^t, \theta^{t-1})$  是 motion potential,  $\phi_0(\theta^t) = \phi(I^t, \theta^t)$  是 observation potential, 只要 observation potential 确定, CRF 可以被用来执行在线的跟踪<sup>[14]</sup>。

$$p(\theta^t, I^{1 \cdots T}) \propto \phi_0(\theta^t) \int \phi(\theta^t, \theta^{t-1}) p(\theta^{t-1} | I^{1 \cdots t-1}) d\theta^{t-1} \quad (1)$$

跟 HMM 比较该模型的最大优势在于  $\phi_0(\theta^t)$  比  $p(I^t | \theta^t)$  更容易得到。

定义 observation potential 如下:

$$\phi(I^t, \theta) = N(S(I^t, \theta^t); 0, \sigma^2)$$

其中  $S$  为相似度衡量。

### 1.2 后验概率的计算

令  $\Theta$  是包含所有有效状态的状态空间。被分割成  $N$  个互不相交的细胞  $C_i, \Theta = \bigcup_{i=1}^N C_i, \Theta = \bigcup_{i=1}^N C_i \cap C_j = \emptyset, i \neq j$ 。样本  $\theta_i \in C_i$ , 称  $\theta_i$  为栅格。暂时先验的表示如下<sup>[14]</sup>:

$$p(\theta^t \in C_i | \theta^{t-1} \in C_j) = \int_{C_i} \int_{C_j} \phi(\theta^t, \theta^{t-1}) d\theta^{t-1} d\theta^t \approx \phi(\theta_i, \theta_j) | C_i | | C_j | \quad (2)$$

$| C_i |$  是第  $i$  个细胞的容积。第  $j$  个到第  $i$  个细胞的转移概率  $T_{ij}$ 。

$$T_{ij} = \frac{\phi(\theta_i, \theta_j) | C_i |}{\sum_{k=1}^N \phi(\theta_k, \theta_j) | C_k |} \quad (3)$$

观察值和状态属于某个细胞的 compatibility 写成如下的形式:

$$\phi_0(C_i) = \int_{C_i} \phi_0(\theta) d\theta \approx \phi_0(\theta_i) | C_i | \quad (4)$$

从而状态属于第  $i$  个细胞的后验概率由公式(1), (2), (4) 可得<sup>[14]</sup>:

$$p(\theta^t \in C_i | I^{1 \cdots T}) \approx \frac{1}{Z} \phi_0(\theta_i) | C_i | \sum_{j=1}^N T_{ij} p(\theta^{t-1} \in C_j | I^{1 \cdots t-1}) = \frac{1}{Z} \phi_0(\theta_i) \sum_{j=1}^N S_{ij} p(\theta^{t-1} \in C_j | I^{1 \cdots t-1}) \quad (5)$$

其中  $S_{ij} = | C_i | T_{ij}$  是时间独立的可以通过离线算得。公式(5)非常重要, 下面的 RBPF 算法就是根据。

### 1.3 激光样本的获得与运用

类似于文献[12]中激光特征的获得方法。将机器人固定, 而人在机器人前面各个方向行走, 提取出碰到人的双脚返回的激光束。然后按照文献[12]中的方法进行归一化, 计算两个激光特征的 chamfer 距离。从而得到一个相似度矩阵。根据相似度矩阵采用聚类算法把训练数据聚类成  $K$  个  $SF_i (i = 0 \cdots k)$  类<sup>[15]</sup>, 然后在每个类中选取一个代表, 该代表能最小化该点到类中所有其他点的距离。这个代表就对应上述 CRF 中的  $\theta$ 。  $SF_i$  对应栅格滤波中的细胞  $C_i$ 。

### 1.4 跟踪算法描述(RBPF)

文中采用 RBPF 算法实现跟踪。RBPF 维持一个赋有权值的样本集  $S_k$ 。

$$S_k = \{ \langle s_k^{(l)}, \omega_k^{(l)} \rangle | 1 \leq l \leq N \}$$

下标  $k$  表示  $k$  时刻,  $l$  表示第  $l$  个粒子。  $s_k^{(l)}, \omega_k^{(l)}$  分别表示  $k$  时刻第  $l$  个粒子和它的权值。  $s_k^{(l)} = \langle \vartheta_k^{(l)}, \theta_k^{(l)} \rangle$ , 其中  $\vartheta_k^{(l)}$  是一个卡尔曼滤波器用来估计人的位置跟速度  $\vartheta_k^{(l)} = \langle \mu_k^{(l)}, \sum_{k=1}^l \rangle$ 。  $\theta_k^{(l)}$  就是激光样本。RBPF 的每个粒子都维持一个卡尔曼滤波器。

算法的描述如下:

- 1) 输入:  $S_{k-1} = (\langle s_{k-1}^{(l)}, \omega_{k-1}^{(l)} \rangle | 1 \leq l \leq N)$ , 激光数据  $I_k$ 。
- 2) 初始化  $S_k$  为空。
- 3) // 为每个粒子进行卡尔曼预测  
For  $l: = 1 \cdots N$  do  
 $\vartheta_{k|k-1}^{(l)} = \langle \mu_{k|k-1}^{(l)}, \sum_{k|k-1}^{(l)} \rangle$
- 4) // 更新重要性权值  
 $\hat{\omega}_{k-1}^{(l)} \propto \omega_{k-1}^{(l)} \sum_{C_i} p(\theta^* \in C_i | I^{1:k})$
- 5) // 根据新的权值进行采样, 生成新的粒子集  
For  $l: = 1 \cdots N$  do  
从  $S_{k-1}$  中根据更新的权值  $\hat{\omega}_{k-1}^{(l)}$  采样  $s_{k-1}^{(l)} = \langle \vartheta_{k-1}^{(l)}, \theta_{k-1}^{(l)} \rangle$   
根据  $p(\theta^* \in C_i | I^{1:K})$  按比例生成新的  $\theta_k^{(l)}$   
卡尔曼滤波器利用  $\theta_k^{(l)}, \vartheta_{k|k-1}^{(l)}$  信息更新位置估计  $\vartheta_k^{(l)}$   
 $s_k^{(l)} = \langle \vartheta_k^{(l)}, \theta_k^{(l)} \rangle$   
 $S_k = S_k \cup \{s_k^{(l)}, \frac{1}{N}\}$
- 6) return  $S_k$

## 2 SURF 特征

SURF 是一种尺度不变和旋转不变的特征描述算子。与 SIFT 算法相比, 它的最大优势是它的速度快。它对特征点的探测采用一种快速海塞矩阵探测器 (Fast-Hessian Detector)。文中采用的是  $(4 \times 4) \times 4 = 64$  位的描述器。而 SIFT 采用  $(4 \times 4) \times 8 = 128$  位的描述器。在对在线快速的探测、描述、匹配, SIFT 描述器的高维性, 成为它的最大缺点。而这也是 SURF 的优势所在<sup>[13]</sup>。由于 SURF 的尺度不变性和旋转不变性, 以及快速性, 适合机器人实时的目标跟踪。

### 2.1 SURF 模板样本的获得

虽然 SURF 具有旋转不变性, 但是旋转超过一定得角度也会导致目标无法识别。用书做实验发现, 学习时书平面跟摄像头平面平行, 识别时书平面跟摄像头平面的夹角超过 30 度时会导致识别失败。在文中, 将机器人固定, 人在机器人前面初始位置与机器人平行, 然后在原地每次旋转一定得角度 (大约 25 度), 直至旋转 (360 度)。这样得到 14 副图像只选取其中人的部分计算 SURF 特征。得到了图像的模版数据库。这些都可以实现离线计算, 所以并不影响实时跟踪的速度。

### 2.2 匹配度的计算

根据文中第二部分的计算结果可以大致确定人在

图像中的候选区域, 从而不用计算整副图像。计算候选区域的 SURF 特征点, 假如有  $N$  个。

模版图像的特征点有  $M$  个, 匹配到的点有  $V$  个, 则定义匹配度  $P$ :

$$P = \frac{V}{N} \times \frac{V}{M} \quad (6)$$

$P$  值最大的模版作为匹配模版。

## 3 实验分析

文中的所有实验都在实验室的家庭服务机器人上完成。该机器人装有一个双目摄像头, 和一个 SICK 激光传感器。其中摄像头大概距离地面 165cm, 分辨率为  $320 \times 240$ , 可以对人的全身进行摄像。激光传感器距离地面大概 45cm 可以探测到人的双脚。

为了测试该算法在移动机器人上面的性能, 必须获得人图像的样本数据和激光样本数据。首先与摄像头平行得到人的正面图像, 然后每次旋转大约 30 度, 得到 11 副图像 (0 度和 360 度重合) 作为样本库中的样本。

在图 2~图 3 中是集中较特殊位置的样本图像。正面 (大约 0 度), 正侧 (大约 45 度), 侧面 (大约 90 度), 背侧 (大约 135 度), 背面 (大约 180 度)。原始图



正面 正侧 侧面 背侧 背面

图 2 原始图像



正面 正侧 侧面 背侧 背面

图 3 SURF 图像

像,依次为(正面,正侧,侧面,背侧,背面)。

图 4 跟各个模板的匹配情况如下:和正面图像有 4 个点匹配,正侧图像有 1 个点匹配,和其他图像没有点匹配。

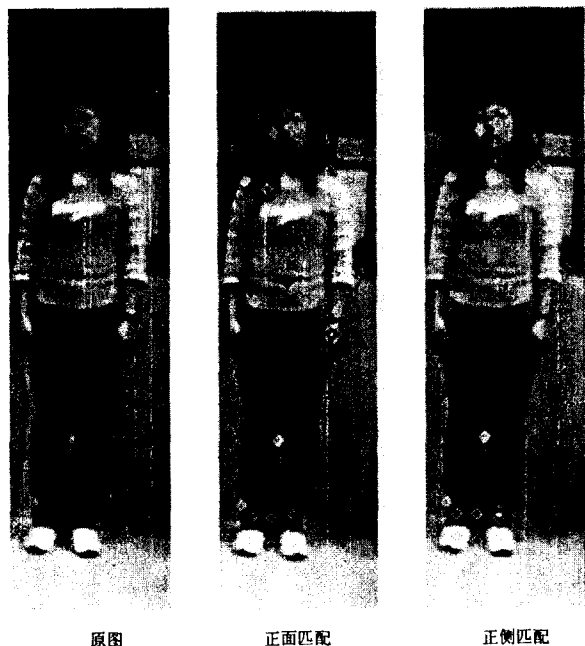


图 4 图像匹配

激光特征的获得相对比较复杂。固定机器人,人在机器人前面以不同的朝向行走,提取 3000 帧激光数据。每帧 180 束,提取从人的双脚返回的激光束。然后用 K 均值算法对样本进行聚类。实验中  $K = 30$ , RBPF 的粒子数为 100。利用该算法可以实现对单个人的准确迅速的跟踪。

#### 4 结束语

文中介绍一种基于条件随机场模型的 RBPF 机器人目标跟踪算法。RBPF 对激光特征进行采样实现人的跟踪。RBPF 根据栅格滤波方法计算得到的后验概率进行权值的更新和采样。每个粒子维持一个卡尔曼滤波器来估计人的位置和速度。SURF 特征的尺度不变性和旋转不变性大幅度减少了图像样本的数目,提高了准确性。

#### 参考文献:

- [1] Arras K, Vestli S. Hybrid, high-precision localization for the mail distributing mobile robot system MOPS[C]//Proc. of the IEEE International Conference on Robotics & Automation (ICRA). Leuven, Belgium: [s. n.], 1998.
- [2] Horswill I. Polly: A vision-based artificial agent[C]//Proceeding of the Eleventh National Conference on Artificial Intelligence. [s. l.]: AAAI, MIT Press, 1993: 824-829.
- [3] Simmons R, Goodwin R, Haigh K, et al. A layered architecture for office delivery robots[C]//Proc. of the First International Conference on Autonomous Agents. Marina del Rey, CA: [s. n.], 1997.
- [4] Endres H, Feiten W, Lawitzky G. Field test of a navigation system: Autonomous cleaning in supermarkets[C]//Proc. of the IEEE International Conference on Robotics & Automation (ICRA). Leuven, Belgium: [s. n.], 1998.
- [5] Engelberger J F. Health-care robotics goes commercial: The helpmate experience[J]. Robotica, 1993, 11: 517-523.
- [6] Burgard W, Cremers A, Fox D, et al. Experiences with an interactive museum tourguide robot[J]. Artificial Intelligence, 1999, 114(1-2): 3-55.
- [7] Thrun S, Bennewitz M, Burgard W, et al. MINERVA: A second generation mobile tour-guide robot[C]//Proc. of the IEEE International Conference on Robotics & Automation (ICRA). Detroit, Michigan: [s. n.], 1999.
- [8] Fod A, Howard A, Mataric M. Laser-based people tracking[C]//Proc. of the IEEE International Conference on Robotics & Automation (ICRA). Washington DC: [s. n.], 2002: 3024-3039.
- [9] Bennewitz M, Burgard W, Thrun S. Using EM to learn motion behaviors of persons with mobile robots[C]//Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Lausanne, Switzerland: [s. n.], 2002.
- [10] Schulz D, Burgard W, Fox D. People tracking with a mobile robot using sample-based joint probabilistic data association filters[J]. International Journal of Robotics Research (IJRR), 2003, 22(2): 99-116.
- [11] Vadakkepat P, Liu Jing. Improved Particle Filter in Sensor Fusion for Tracking Randomly Moving Object[J]. IEEE Transactions on Instrumentation and Measurement, 2006, 55(5): 1823-1832.
- [12] Schulz D. A Probabilistic Exemplar Approach to Combine Laser and Vision for Person Tracking[C]//Proceedings of Robotics: Science and Systems II (RSS 2006). Philadelphia, PA, USA: [s. n.], 2006.
- [13] Tuytelaars H B T, Gool L V. SURF: Speeded Up Robust Features[M]//Lecture Notes in Computer Science. In: ECCV. [s. l.]: [s. n.], 2006: 404-417.
- [14] Taycher L, Shakhnarovich G, Demirdjian D, et al. Conditional Random People: Tracking Humans with CRFs and Grid Filters in[M]//Computer Vision and Pattern Recognition. New York, USA: [s. n.], 2006.
- [15] 曹付元, 梁吉业, 姜广. 基于邻域模型的 K-means 初始聚类中心选择算法[J]. 计算机科学, 2008(11): 181-183.