

文章编号: 1002-0446(2003)03-0259-05

## 机器人灵巧手抓持分类器的设计与实现\*

刘 杰 张玉茹

(北京航空航天大学机器人研究所 北京 100083)

**摘 要:** 机器人灵巧手的抓持分类是抓持规划的一个主要问题. 本文应用模式识别技术设计和实现了一种基于高斯混合模型 GMM 的分类器. 采用 Expectation Maximization(EM) 算法估计 GMM 的参数, 对人手的抓持动作进行识别与分类, 经过人手到机器人手的关节空间运动映射, 实现了机器人灵巧手的抓持动作分类. 该分类器可以应用于在线抓持规划.

**关键词:** 抓持分类; 抓持规划; 高斯混合模型; 模式识别

**中图分类号:** TP24 **文献标识码:** B

## DESIGN AND IMPLEMENT A SYSTEM OF GRASP IDENTIFICATION FOR DEXTEROUS ROBOT HAND

LIU Jie ZHANG Yu-ru

(Robotics Institute of Beijing University of Aeronautics and Astronautics, Beijing 100083)

**Abstract:** The grasp taxonomy is one of the key problems in grasp planning of dexterous robot hands. This paper describes the design and implement of a system of grasp identification based on pattern recognition technology. The feature of human grasp is represented by Gaussian Mixture Model (GMM). The parameters of GMM are estimated by the Expectation Maximization algorithm (EM). The taxonomy of robot hand grasps is achieved by mapping grasps from human hand to robot hand. The effectiveness of the GMM identification method is proved by the experiments. The system can be used in grasp planning.

**Keywords:** taxonomy of grasps, grasp planning, Gaussian mixture model, pattern recognition

### 1 引言(Introduction)

人手抓持分类学的研究方法主要是定性分析, 目前还没有定量的数学描述<sup>[1]</sup>. 本文的目标是给出抓持分类的数学描述, 并应用于机器人灵巧手抓持规划. 我们采用主从策略实现抓持规划, 即用数据手套采集人手运动数据, 并将其映射到机器人灵巧手. 因此, 对灵巧手的抓持分类依赖于对手运动的识别与分类. 为了定量描述人手抓持动作, 本文采用高斯混合模型 GMM (Gaussian Mixed Model) 来有效提取抓取动作的运动特征.

高斯混合模型方法<sup>[2]</sup>是一种基于数值计算的机器学习方法, 广泛应用于模式识别和数据分析等领域<sup>[3]</sup>. Hinton, G. 等<sup>[4]</sup>利用这种方法研究了手写体识

别问题; GMM 也被用于非限定文本的话者识别<sup>[5,6]</sup>, 取得了良好的识别效果. 国内学者在 90 年代中期开始针对 GMM 在语音信号处理、话者识别和手语识别方面的应用展开研究<sup>[7-9]</sup>. 传统的 GMM 方法是无导师的学习方法, 模型中的参数学习算法是用一类训练样本通过最大似然估计方法实现. 理论上, 如果训练样本在样本空间分布得足够稠密, 用这种方法估计的参数在最大似然意义下是最优的.

本文设计实现了机器人灵巧手的抓持分类器. 用数据手套测量操作者的手部关节运动, 用高斯混合模型对手运动进行识别和分类, 通过人手到灵巧手的关节空间映射, 进而实现灵巧手抓取动作分类. 同时还实现了虚拟现实环境中人手和灵巧手运

\* 基金项目: 国家自然科学基金(59985001)和国家教育部博士点基金(2000000605)资助项目.

收稿日期: 2002-10-18

动的可视化, 提供了直观地观察和分析抓持运动的途径.

## 2 模型描述(Model introduction)

对任意多类别样本 GMM 可以很好的近似这些样本空间的概率分布, 传统的 GMM 模型是一种静态的模型, 优点是可以平滑地逼近任意形状的概率密度函数, 而且是一个易于处理的参数模型. 一个完整的 GMM 有 3 个参数. 它们是混合加权值、均值向量和协方差矩阵.

$$\lambda = \{c_i, \mu_i, \Sigma_i\} \quad i = 1, \dots, M \quad (1)$$

其中  $c_i$  表示混合加权值,  $\mu_i$  平均值向量,  $\Sigma_i$  表示协方差矩阵.  $\sum_{i=1}^M c_i = 1$ ,  $M$  是该高斯混合模型的高斯分布的个数, 就是高斯混合模型的阶数. 对每一个抓持动作而言, 用一个  $\lambda$  表示. 采用最大似然估计 (Maximum Likelihood Estimation, MLE) 方法计算参数  $\lambda$ .

设  $\Omega_1, \Omega_2, \dots, \Omega_N$  是  $N$  个不同的类别, 第  $i$  类别  $\Omega_i$  的训练样本为  $x_{i1}, x_{i2}, \dots, x_{iN_i}$ ,  $i = 1, 2, 3, \dots, N$ . 假定每个类别出现的先验概率相同. 对于给定的一个样本  $x$ , 它隶属  $\Omega_i$  的先验概率  $p(x|\Omega_i)$  如下

$$p(x|\Omega_i) = \sum_{k=1}^M c_{ik} p_{ik}(x) \quad (2)$$

其中,  $x$  是  $p$  维向量, 且  $p_{ik}(x) = N(\mu_{ik}, \Sigma_{ik})$ , 即为正态分布.  $\mu_{ik}, \Sigma_{ik}$  分别是均值向量和协方差矩阵,  $c_{ik} \geq 0$ ,  $\sum_{k=1}^M c_{ik} = 1$ , 这些是要被估计的参数, 高斯概率密度函数  $p_{ik}(x) = N(\mu_{ik}, \Sigma_{ik})$ .

$$p_{ik}(x) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma_{ik}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x - \mu_{ik})^T \Sigma_{ik}^{-1} (x - \mu_{ik})\right\} \quad (3)$$

对于给定的测试样本向量  $x$ , 它隶属的类别可由下述关系确定

$$I = \underset{1 \leq i \leq N}{\operatorname{argmax}} p(x|\Omega_i) \quad (4)$$

## 3 模型训练(Model training)

模型的学习方法采用最大似然估计的一种常用方法期望值最大化演算法 (Expectation Maximization Algorithm, EM) 计算高斯混合模型的参数  $\lambda$ . 先从给定的初始模型  $\lambda$  开始, 计算训练样本

中每一个特征向量的在每一个高斯分布的统计概率, 再用这些统计概率来计算每一个高斯分布的期望值, 然后以这些期望值反过来最大化高斯混合模型的参数值, 得到相应的参数  $\lambda$ , 重复上述步骤, 直到收敛为止, 下面是详细的计算公式.

第  $i$  个混合的事件概率值

$$P(i|x_t, \lambda) = \frac{c_i P_i(x_t)}{\sum_{j=1}^M c_j P_j(x_t)} \quad (5)$$

第  $i$  个混合的加权估计值

$$\bar{c}_i = \frac{1}{T} \sum_{t=1}^T P(i|x_t, \lambda) \quad (6)$$

第  $i$  个混合的平均值估计值

$$\bar{\mu}_i = \frac{\sum_{t=1}^T P(i|x_t, \lambda) x_t}{\sum_{t=1}^T P(i|x_t, \lambda)} \quad (7)$$

第  $i$  个混合的协方差估计值

$$\bar{\sigma}^2 = \frac{\sum_{t=1}^T P(i|x_t, \lambda) x_t^2}{\sum_{t=1}^T P(i|x_t, \lambda)} - \bar{\mu}_i^2 \quad (8)$$

其中  $x_t$ ,  $t = 1, 2, \dots, T$  表示训练样本个数. 当  $i = 1, 2, \dots, N$ , 根据上述公式迭代计算模型  $c_{ik}, \mu_{ik}, \Sigma_{ik}$  的参数.

### 3.1 特征提取方法

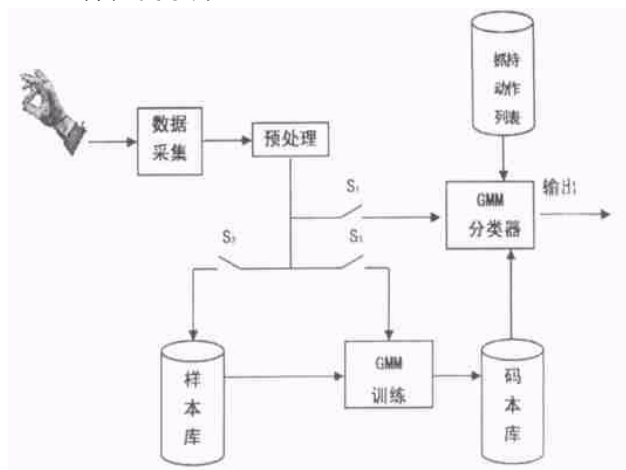


图1 基于 GMM 的人手的抓持动作识别系统

Fig. 1 A GMM-based human grasp recognition system

在基于数据手套的灵巧手在线抓持规划系统中, 通过操作者佩戴数据手套采集人手的运动, 人手到灵巧手的关节空间的运动映射方法让灵巧手实现了类似人手的动作. 根据抓持动作自身的特点——单手手势, 选取单只具有 18 个传感器的 CyberGlove

型号数据手套作为抓持动作的输入设备. 我们研究的是手的预抓持动作, 即手与被抓持物体接触前的静止手势. 因此, 进行抓持分类所需要的数据是由数据手套输出的手形信息选择手的关节转角作为特征描述数据, 不包含运动轨迹和手运动过程中的时间信息.

### 3.2 训练和识别过程

首先为抓持动作列表中的每个抓持动作采集若干组样本, 组成样本数据库. 然后以抓持动作为单位, 为每个抓持动作训练一个 GMM 模型, 即利用公式(5)~(8)进行参数估计, 并将结果存入标准抓持动作码本库中.

从图 1 中不难看出训练过程可以有两种不同的方式: 离线(off-line)学习和在线(on-line)学习. 离线学习独立于识别过程, 它是利用事先建好的样本库

中的样本进行训练, 通常用于建立码本库; 而在线学习伴随识别过程, 它直接利用设备输出的实时样本进行训练, 一般用于对已有抓持动作码本库进行补充或者修改.

分类过程就是模式识别过程, 识别过程实质上是在候选码本集中选取一个能够最好地描述观察信号的模型. 根据识别系统结构图 2 所示的过程, 首先对由设备输入的手的完成抓持动的手势信号  $G$  的每一帧  $G(t) (t= 1 \dots T)$  相应的位置信息进行规整校准. 然后进行特征提取和运动映射, 得到对应于手势信号  $G$  的一个关键帧  $O$ , 最后计算其对应于候选抓持动作码本集中的每一个码本  $\lambda_v (1 \leq v \leq V)$  的概率,  $G$  所属类别由下式最大对数似然概率确定.

$$v^* = \underset{1 \leq v \leq V}{\operatorname{arg\,max}} \log P(O | \lambda_v) \quad (9)$$

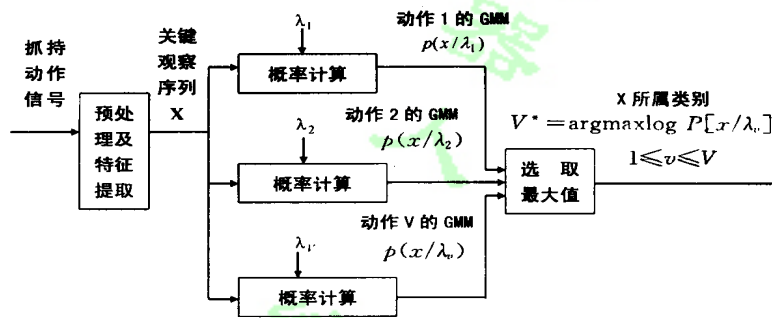


图 2 GMM 分类器方框图

Fig. 2 The diagram of GMM classifier

## 4 试验验证(Experimentation and result)

根据上述原理, 在 Win2000 平台上采用 Visual C++ 6.0、OpenGL 和 Cyberglove 软件开发包 VirtualHand Suite2.5 实现了一个 GMM 分类器系统. 基于虚拟现实环境的主要交互界面如图 3 所示.

### 4.1 抓取类型的确定

人手的抓取动作大致分为三类: 抓 wrap/grasp; 握 grip; 捏 pinch. 详细划分为如下几种抓取模式. 强力抓取 power grasp 手拿水杯、圆柱抓取 cylindric grasp 拿象棋子、夹握 chuck grip 拿铅笔、钩握 hook grip 提旅行箱、跨握 span grip 拿大瓶子、侧捏 lateral pinch 用手拿一页纸、精密捏 precision pinch 用手操作针<sup>[10]</sup>; 不过由于目前的驱动方式和机械结构的加工制造水平, 多指灵巧手在尺寸、结构和运动学上还无法做到和人手完全一样, 无法完成许多人手的动作. 因此选择了钩握、强力抓取、圆柱抓取、侧捏、夹握五种灵巧手可以完成的动作作为待识别的

抓持动作.

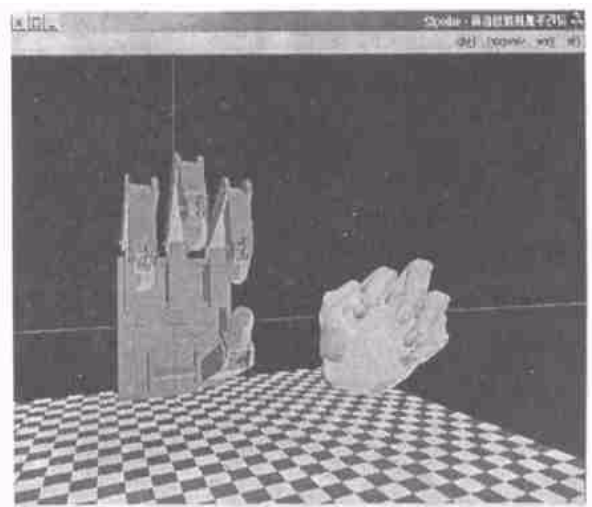


图 3 抓持分类软件主界面

Fig. 3 Main interface of the grasping simulator

### 4.2 数据采集和特征提取

基于数据手套软件开发包 VirtualHand Suite 开

发了数据手套的数据采集和处理模块, 在手套采样速度 38400 波特率条件下测量得到人手的原始数据如下: 手的各关节的原始传感数据, 即小指 MPJ (pinkie MPJ)、小指 PIJ (pinkie PIJ)、无名指 MPJ (ring MPJ)、无名指 PIJ (ring PIJ)、中指 MPJ (middle MPJ)、中指 PIJ (middle PIJ)、食指 MPJ (index MPJ)、食指 PIJ (index PIJ)、拇指 MPJ (thumb MPJ)、拇指 IJ (thumb IJ)、小指无名指间外展 (pinkie-ring abduction)、无名中指间外展 (ring-middle abduction)、中食指间外展 (middle-index abduction)、拇指外展 (thumb abduction)、拇指旋转

度 (thumb rotation/TMJ)、腕的外展 (wrist yaw)、腕的曲度 (wrist pitch)、掌的弧度 (palm arch) 这些角度体现了手指的弯曲与伸展, 手掌的曲率以及手腕的旋转, 输出大小与相应传感器所在位置的弯曲程度成线性关系, 范围为 0~255. 详细的人手的角度定义参见图 3. 考虑我们所的机器人灵巧手结构上和人手不同主要体现在手掌上, 因此我们选择如表 1 所示的特征. 一共 15 个测量数据构成特征向量. 人手到灵巧手的运动映射目前采用关节空间的直接映射.

表 1 Cyberglove 测量数据和特征提取

Table 1 The raw data of Cyberglove and the feature extraction

测量数据	小指 MPJ	小指 PIJ	无名指 MPJ	无名指 PIJ	中指 MPJ	中指 PIJ	食指 MPJ	食指 PIJ	拇指 MPJ
特征向量	是	是	是	是	是	是	是	是	是
测量数据	拇指 IJ	小指无名指间外展	无名指中指间外展	中食指间外展	拇指外展	拇指旋转度	腕的外展	腕的曲度	掌的弧度
特征向量	是	是	是	是	是	是			



图 4 人手解剖图

Fig. 4 Anatomical definitions of the human hand

从数据手套的原始数据中提取手形特征, 构成训练样本和测试样本的特征向量. 针对每类抓持动作, 测试者带上数据手套做 40 次相同的抓持动作, 通过数据采集、数据预处理和特征提取模块得到对应与每一类抓持动作的 40 组数据, 其中 20 组作为训练样本, 5 组作为测试样本, 采用 GMM 模型量化抓持动作的数据. 首先利用训练样本根据公式(5)~(8)进行模型的参数估计, 然后对每一个测试样本利用系统识别模块实现分类.

### 4.3 实验参数确定

GMM 模型中参数的初值采用向量量化 (Vector Quantization, VQ) 中的 K 平均值修正法 (Modified K-means). 根据实际情况确定 GMM 分

类器对抓持动作的分类试验参数如下:

- I. 训练样本特征向量的维度 15;
- II. 高斯分布个数高斯混合模型 GMM 的阶数 M 为 5;
- III. 样本种类(待识别的模式, 抓持动作类别) 5;
- IV. 每一类抓持动作训练样本个数 30;
- V. 每一类抓持动作测试样本个数 10.

表 2 采用 GMM 分类器对抓持动作分类结果

Table 2 The experiment result of GMM classifier for grasp

抓持动作	强力抓取	钩握	圆柱抓取	侧捏	夹握
训练样本	20	20	20	20	20
测试样本	10	10	10	10	10
未正确识别样本	0	1	1	2	2
识别率 $v$	0.88				

### 4.4 实验结果

识别率定义为正确分类的测试样本占测试总样本的个数.  $v = 1 - 6/50 = 0.88$ , 识别率达到 88%.

## 5 结论 (Conclusion)

为了量化人手的抓持动作, 本文建立了抓持分类的数学描述, 应用模式识别技术设计实现了基于高斯混合模型 GMM 的分类器, 采用期望值最大化

演算法(EM)计算高斯混合模型的参数,建立了 5 类抓持动作的 GMM 模型,为研究人手和灵巧手的运动提供了有力的分析工具.基于该方法开发的分类系统可以识别人手的抓持动作,并将其映射到机器人灵巧手.还可以建立典型抓持动作数据库,依据抓持策略从中选择满足任务要求的抓持动作,通过调用抓持数据库来完成局部自主的抓持.因此,该分类系统可以用于基于数据手套的多指灵巧手离线和在线抓持规划.

### 参考文献 (References)

- [1] Cutkosky M. On Grasp Choice, Grasp Models, and the design of hands for manufacturing tasks[J]. IEEE Transactions on Robotics and Automation, 1989, 5(3)
- [2] Aitkin M, Wilson G. Mixture models, outliers, and the EM algorithm[J]. Technometrics, 1980, 22: 325- 331
- [3] Nandakishore Kamhatla. Local models and Gaussian mixture models for statistical data processing[D]. Dept of Computer Science and Engineering, Oregon Graduate Institute of Science & Technology, 1996
- [4] Hinton G, Revon M, Dayan P. Recognizing handwritten digits using mixtures of linear models[J]. In Tesauro, Touretzky & Leen, eds, Advances in Neural Information Processing Systems, 1995, 7: 1015- 1022
- [5] Reynolds D A. Speaker identification and verification using Gaussian mixture speaker models[J]. Speech Communication, 1995, 17: 91- 108
- [6] Reynolds D A. Speaker identification and verification using gaussian mixtures speaker models[A]. ESCA Workshop on Automatic Speaker Recognition[C]. Identification and Verification, 1994. 27- 30
- [7] Jiyong Ma, Wen Gao. The supervised Gaussian mixture model[J] of Comput Sci & Technol[J]. 1998, 13(5): 471- 471
- [8] 牟晓隆, 胡起秀, 吴文虎. 与文本无关的复合策略说话人辨别系统[J]. 清华大学学报, 1997, 37(3): 16- 19
- [9] 吴江琴, 高文, 陈熙霖, 马继涌. 多层 DGMM 识别器在中国手语识别中的应用[J]. 软件学报, 2000, 11(11): 1430- 1439
- [10] Iberall T. Human prehension and dexterous robot hands[J]. The Int J Robotics Research, 1997, 16(3): 285- 299

### 作者简介:

刘 杰 (1975-), 男, 博士研究生. 研究领域: 模式识别, 虚拟现实, 机器人多指手运动规划.

张玉茹 (1959-), 女, 博士, 教授, 博导, 研究领域: 机器人机构学、医用机器人、人机交互技术.

(上接第 234 页)

## 5 结论(Conclusion)

本文给出了完整的基于 USBL/Doppler 和扩展卡尔曼滤波的水下机器人实时位置估计算法,为具有反馈延时的水下机器人位置控制提供了一个可行的解决思路,仿真结果表明,该算法具有很好的位置估计精度.

### 参考文献 (References)

- [1] Myung-Hyun Kim. Nonlinear control and robust observer design for marine vehicle[D]. Virginia State: Virginia State University, 2000
- [2] Louis Whitecomb etc. Advances in doppler-based navigation of underwater robotic vehicle[A]. In Proceedings of the IEEE International Conferences on Robotics & Automation[C], 1999. 399- 406
- [3] Bennamou M, etc. The Development of an Integrated GPS/INS/SONAR Navigation System for Autonomous Underwater Vehicle Navigation[A]. Proceedings of IEEE OCEAN' s96[C]. 1996. 256- 261
- [4] Mohamed A H, etc. Adaptive kalman filter for INS/GPS[J]. Journal of Geodesy, 1999. 193- 203
- [5] Schmiegel A U. A Stabilized model-based kalman filter for underwater navigation[A]. In Proceedings Conference on BAE SYSTEMS Signal & Data Processing[C]. 2002
- [6] Andreas Huster etc. Relative position estimation for intervention-capable AUVs by fusing vision and inertial measurements[A]. In Proceedings of the 12th International Symposium on Unmanned Untethered Submersible Technology[C]. Durham NH: August 2001. Autonomous Undersea Systems Institute.
- [7] White NA. MMAE Detection of Interference/jamming and Spoofing in a DGPS-aided Inertial System[D]. MS Thesis: Air Force Institute of Technology, Ohio, 1996
- [8] Wang Jinling. Kinematics GPS Positioning with Adaptive Kalman Filtering Technology[A]. Proc IAG' 97[C]. 1997

### 作者简介:

邢志伟 (1970-), 男, 博士研究生, 研究领域: 水下机器人控制.

封锡盛 (1942-), 男, 中国工程院院士, 博士生导师, 研究领域: 水下机器人.