

## 人手粒子滤波跟踪器的研究与设计\*

冯志全<sup>1+</sup>, 杨波<sup>1</sup>, 郑艳伟<sup>1</sup>, 唐好魁<sup>1</sup>, 徐涛<sup>1</sup>, 李毅<sup>2</sup>

<sup>1</sup>(济南大学 信息科学与工程学院, 山东 济南 250022)

<sup>2</sup>(济南大学 控制科学与工程学院, 山东 济南 250022)

### Design and Research on Particle Filtering for Human Hand 3D Tracking

FENG Zhi-Quan<sup>1+</sup>, YANG Bo<sup>1</sup>, ZHENG Yan-Wei<sup>1</sup>, TANG Hao-Kui<sup>1</sup>, XU Tao<sup>1</sup>, LI Yi<sup>2</sup>

<sup>1</sup>(School of Information Science and Engineering, University of Ji'nan, Ji'nan 250022, China)

<sup>2</sup>(School of Control Science and Engineering, University of Ji'nan, Ji'nan 250022, China)

+ Corresponding author: E-mail: ise\_fengzq@ujn.edu.cn

Feng ZQ, Yang B, Zheng YW, Tang HK, Xu T, Li Y. Design and research on particle filtering for human hand 3D tracking. *Journal of Software*, 2008,19(Suppl.):87-95. <http://www.jos.org.cn/1000-9825/19/s87.htm>

**Abstract:** Aiming at describing post probability distribution of the state with high dimensionality by using of a small amount of particles, a sampling in particle filtering for human hand 3D tracking is put forward in this paper. First of all, with the specific human-computer conditions, both cognitive psychology features of operators and motion features of the operators' hands are studied, upon which a novel concept, microstructure of state variable, is proposed. Then, a general method to describe the microstructure of state variable is further discussed. At last, a sampling algorithm based on the microstructure of state variable is put forward. The microstructure of state variable provide an unite and efficient mathematic model upon which the sampling algorithm may avoid sampling mass particles with poor quality. In order to validate validity and performance of sampling algorithm, a great deal of experiments are completed, and just fewer particles, compared with conventional particle filtering, may achieve better tracking precision.

**Key words:** human hand 3D tracking; particle filtering; microstructure of state variable

**摘要:** 提出一种粒子滤波方法中粒子样本采样方法,旨在用较少的粒子数描述高维状态变量的后验概率分布.首先,研究特定人机交互条件下操作者的认知心理特点和人手运动特点,在此基础上提出了状态变量微观结构的基本概念,然后,进一步探讨描述微观结构的一般方法,最后,提出了基于状态变量微观结构的粒子采样方法和改进的粒子滤波跟踪算法.状态变量的微观结构为采样算法的设计提供了一种统一、高效的数学模型,以此为

---

\* Supported by the National Natural Science Foundation of China under Grant No.60773109 (国家自然科学基金); the Natural Science Foundation of Shandong Province of China under Grant No.Y2007G39 (山东省自然科学基金); the Key Project of Natural Science Foundation of Shandong Province of China under Grant No.2006G03 (山东省自然科学基金); the Natural Science Foundation of Shandong Province of China for Distinguished Young Scholars under Grant No.JQ200820 (山东省自然科学基金); the Science and Technology Plan of Shandong Province Education Department of China under Grant No.J07YJ18 (山东省教育厅科技计划); the Ph.D. Scientific Research Foundation of University of Ji'nan of China under Grant No.XBS0705 (济南大学博士科研启动基金)

Received 2008-05-01; Accepted 2008-11-25

基础的采样算法可以有效避免对质量比较差的粒子样本进行大量采样.为了验证算法的有效性和性能,进行了大量实验,结果表明,与传统的粒子滤波方法相比,采用少量的粒子样本就可以达到较高的跟踪精度.

**关键词:** 人手三维跟踪;粒子滤波;状态变量微观结构

众所周知,人机交互(HCI)技术正在从以计算机为中心逐步转移到以人为中心,具体表现为:(1) 交互方式逐渐演化为适应人类的行为习惯,而不是计算机程序;(2) 使用多种媒体、多种模式进行交互;(3) 基于多通道和多媒体的自然、高效、智能化、无障碍的 HCI 将是新一代智能 HCI 的主要发展方向.

自然人手的运动跟踪是新型人机交互界面的核心内容之一,实现这种跟踪的主要技术手段之一就是粒子滤波.粒子滤波<sup>[1]</sup>虽然是从上世纪 90 年代中后期才发展起来的一种最优非高斯非线性滤波方法,但目前已被学术界认为是最具发展前景的状态估计方法之一.粒子滤波方法可以有效处理由于遮挡和复杂背景所产生的非高斯分布问题和非线性问题,它在每一时刻输出的是目标位置概率分布的近似,且可以保留多个假设匹配同时随时间传播,使得跟踪丢失时也能重新找到目标并继续跟踪.粒子滤波跟踪方法的另一个重要特点在于,它对于光照变化、复杂背景以及人手自遮挡等因素的影响具有较强的鲁棒性.

自然人手状态采样是基于粒子滤波方法进行自然人手跟踪研究中的关键问题之一,它直接影响到跟踪结果的实时性.但是,粒子滤波方法需要大量样本来表示后验分布,对于以高维为特征的关节式物体来说,即使采用 10 维变量,每个变量取 20 个离散值,也至少需要 320 000 000 个样本才能较好地表示后验分布的典型集<sup>[2]</sup>,这使得实时性跟踪变得遥不可及.因此,很多研究者对如何降低状态空间的维数进行了比较广泛的研究,除了利用人手的动态约束条件和静态约束条件所决定的变量之间的关联性和变量的变化范围,从而在一定程度上降低状态维数之外,人们采用了多种方法研究高维状态的处理方法.比较典型的方法有状态空间分割法<sup>[3-6]</sup>、机器学习法<sup>[7]</sup>、分析-合成法<sup>[1,8]</sup>、约束误差优化法<sup>[9]</sup>以及基于回归-优化方法的关节式物体的姿态估计方法<sup>[10]</sup>等.现有的这些粒子滤波改进算法基本上是在粒子样本的附近进行局部搜索优化,但这类方法过分地依赖于动态模型的预测结果,仍无法很好地解决高维状态空间所带来的困难;把智能计算用于粒子滤波方法后,鲁棒性和精度固然可以得到明显改善,但计算时间开销太大.本文旨在从新的角度探讨高维状态空间采样方法,降低 PF 算法的时间开销.

## 1 状态变量的微观结构

### 1.1 研究思路

抽象地研究 3D 运动人手跟踪,往往容易掉入“维数灾难”的陷阱,难以降低实时性跟踪的难度.事实上,3D 人手模型一般有 26 个自由度,即使在每个自由度上取 2 个值,也有  $2^{26}$  个不同状态.如果进一步考虑模型约束条件,仍然有  $2^{21}$  个不同状态.另一方面,粒子滤波是一种基于随机搜索策略的方法,需要足够多的样本个数才能描述真实运动,所需的样本个数随空间维数的增长呈指数级增长.

一种有效的研究思路就是在特定的 HCI 环境中研究采样方法.本文正是以数字虚拟样机系统的 HCI 过程作为基本背景,以分析操作者在该过程中的心理认知特征和人手运动生理特征为切入点,以减少必需进行随机采样的变量个数为基本手段,对 PF 中的样本采样方法进行研究.

### 1.2 实验方案

当代认知心理学认为,人的认知活动过程是以系列(serially)和序列(sequentially)的方式对信息进行加工处理的,即当人的认知活动发生时,它们是按照一次一个事件或信息的加工方式进行的,强调按顺序依次连续加工的过程.

根据上述基本理论,设计了一整套实验方案,研究操作者在操作过程中的心理认知特征和人手运动的生理特征.首先,通过对数字虚拟样机系统进行分析,提炼出比较典型的 5 个基本操作:(1) 抓取一个物体;(2) 把螺帽套在螺杆上;(3) 拧螺丝;(4) 移动一个物体;(5) 选中一个物体.在操作过程中,不要求、也不暗示操作人员按照某

种特定的运动模式改变手势,只要求其按照自己的方式自然地完成操作目标和操作任务即可.将受试人员分为5组,每组10人,每组人员完成同一个基本操作.在每个基本操作中,受试人员带上数字手套,其手势变化数据由数字手套传感器存入计算机数据库中(如图1所示).图1显示了部分典型操作,每行显示一个操作的部分片断.

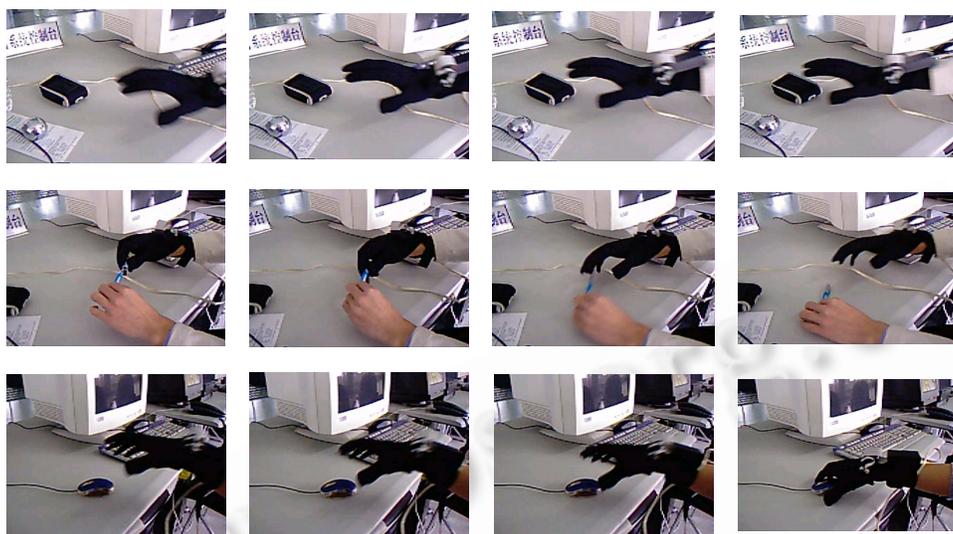


Fig.1 Some of the typical operations devised for exploring both cognitive psychology features of operators and motion features of the operators' hands

图1 本文为研究操作者心理认知特征和人手运动生理特征而设计的几个基本操作

### 1.3 实验结果

对每个操作中连续的5 000帧进行研究,对各组/各人/各帧/各关节的角度变化情况进行分析,重点研究各关节速度和加速度变化情况,图2和图3中分别给出了部分速度和加速度变化曲线图.

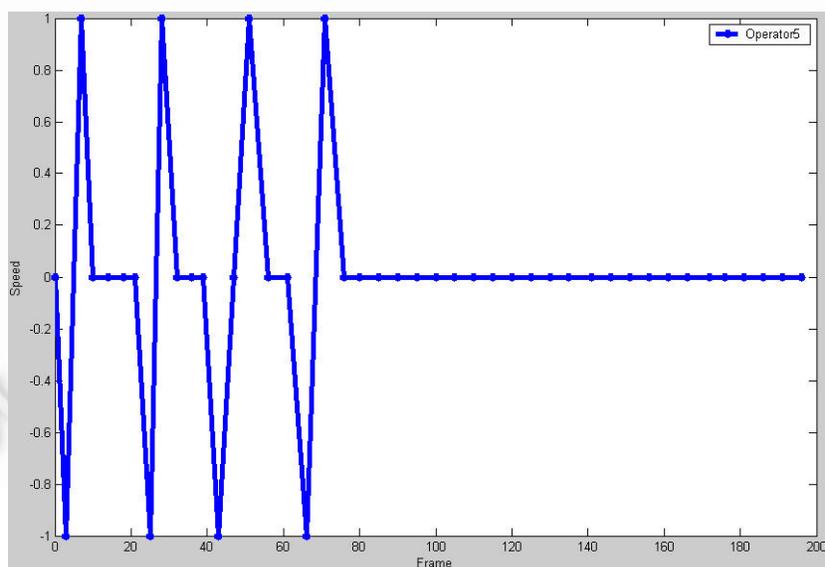


Fig.2 Speed curve of a finger of the operator 5#

图2 5#操作者在实验过程中手指速度变化曲线

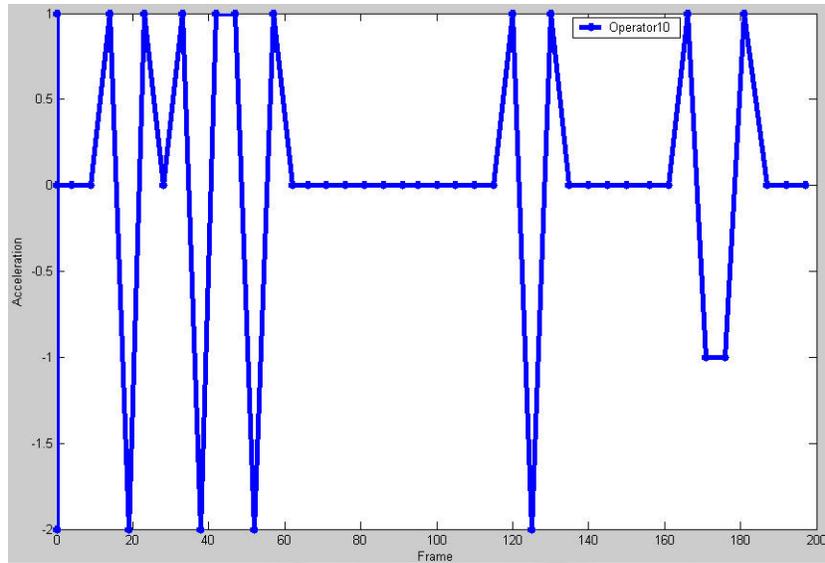


Fig.3 Acceleration curve of a finger of the operator 10#

图3 10#操作者在实验过程中手指加速度变化曲线

#### 1.4 实验结果分析及其主要结论

根据上节的实验结果,得出下面两个基本假说:

假说 1(不变性与无关性假说). 从概率统计意义上说,在绝大部分时间内,操作者手指的速度或加速度基本上保持恒定,且与操作者和基本操作无关.

另一方面,根据认知心理学顺序依次连续加工原理以及人手肌肉兴奋性和收缩性两大基本生理特性,结合对实验者进行调研,操作者在操作过程中往往表现出一般性规律性.

假说 2(阶段性假说). 从概率统计意义上说,任何一个基本操作都由若干基本阶段所构成,在每个阶段可以用速度模型或加速度模型逼近手指的真实运动状态.

#### 1.5 微观结构的基本概念

根据不变性与无关性假说,不同操作者的人手运动过程具有共同的心理认知特征和生理特征,存在着一般性规律;根据阶段性假说,没有必要通过盲目生成大量样本来刻画人手的真实运动状态.当然,我们仍然无法确切判断在某个时刻某状态变量准确的运动形式,因为在多因素的作用下使得状态变量同时呈现出多种强度不同的变化趋势或可能状态,这些变化趋势或可能状态之间往往还存在着联系,本文把与手势状态变量有关的运动形式及其相互关系称为状态变量的微观结构,把不同的运动形式称为通道,把某种通道的概率值称为该通道的通道概率.通道概率越大,说明选择该通道所对应运动形式的可能性也越大.

一个状态变量的微观结构采用 5 个通道进行描述(如图 4 所示),其中,  $S_t$  表示在时刻  $t$  某状态变量的值.匀速通道主要刻画与状态变量匀速运动有关的特性,匀加速运动通道主要刻画与状态变量匀加速运动有关的特性.显然,如果某状态变量既不能进入匀速通道,又不能进入匀加速通道,就要采用其他方法确定该变量的值.遮挡通道描述某手指被其他手指或手掌所遮挡有关的特性.无论是匀速通道、匀加速通道,还是遮挡通道,它们都是建立在随机算法的基础之上,从而有可能使得状态估计结果逐渐远离其真实值.为此,引入误差纠正通道,它主要描述必须采用随机方法进行高密度采样的可能性大小,如果误差纠正概率很大,则意味着要适当增大时间开销以换取精度,采用高密度采样把状态值强制性地修复到正确的范围内.图 4 中的其他通道是指除了上述 4 个通道之外,最后可以进入的通路.

## 2 基于微观结构的粒子采样算法

### 2.1 粒子采样算法的基本思想

一个粒子就是一个手势向量,它可以唯一确定一个三维手势的位置和姿态.状态变量的微观结构为粒子采样算法的设计提供了基本信息,利用这些信息可以建立通道概率与样本数之间的关系,使得通道概率越大的通道被“进入”的可能性也越大.本文提出基于通道概率与样本数之间非线性关系的通道采样方法.当进入某一通道之后,需要在确定性算法和随机性算法之间进行选择,我们用概率方法来实现这种选择.确定性算法的发生概率与对应的通道概率之间成正比例关系.在采用确定性方法计算状态变量值的过程中,近似假设运动系统按照对应通道模型进行运动而忽略随机白噪声的影响;在采用随机性方法计算状态变量值的过程中,近似用上一时刻的高斯分布代替当前时刻的高斯分布进行采样.

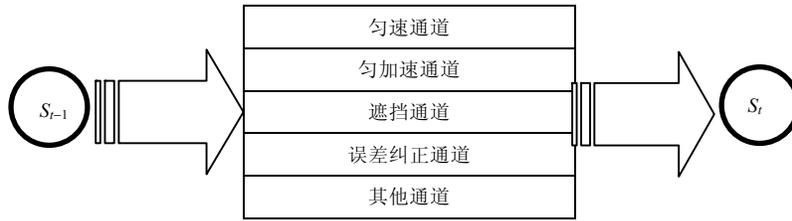


Fig.4 Microstructure of state variable

图4 状态变量的微观结构

本文采用统一的通道概率高斯描述形式:

$$C_i \sim N(\mu_i^R, \sigma_i^R) \quad (1)$$

其中, $C_i$ 表示*i*通道变量( $i$ =匀速通道,匀加速通道,遮挡通道), $\mu_i^R, \sigma_i^R$ 分别表示*i*通道变量相对于当前帧而言前*R*帧对应通道特性的均值和方差.

误差纠正通道概率描述为

$$P=1-e^{-H} \quad (2)$$

其中, $H$ 是预测观测值与实际观测值之间的 Hausdorff 距离.

### 2.2 粒子采样算法描述

输入:(1) 上两帧 3D 手势;(2) 上两帧图像.

输出:当前帧的一个粒子.

算法:

- (1) 刷新状态变量的微观结构以及各通道概率;
- (2) 对每个状态变量
  - 1) 采样通道;
  - 2) 采样确定状态变量值的方法;
  - 3) 采样变量值;
- (3) 返回当前帧的一个粒子.

## 3 改进的 PF 算法

### 3.1 改进的PF算法基本思想

与文献[11]中的方法不同,本文采用精度更高的状态预测算法<sup>[12]</sup>.该方法主要基于双 UT 变换并融入当前帧

观测值;利用本文提出的粒子采样算法产生粒子;根据样本在对应像平面上的投影图像与帧图像之间的 Hausdorff 距离计算粒子权值.

### 3.2 改进的PF算法描述

改进的 PF 算法描述如下:

输入:(1)  $t-1$  时刻的输出状态  $X^{(k-1)}$ ; (2)  $t$  时刻帧图像;

输出: $t$  时刻的输出状态  $X^{(k)}$ .

- (1) 利用 UKFDUT 算法<sup>[12]</sup>对样本  $X^{(k-1)}$  进行预测,得到  $\hat{X}^{(k)}$ ;
- (2) 利用粒子采样算法生成  $N$  个粒子;
- (3) 计算每个粒子的权值;
- (4) 对  $N$  个粒子作加权平均,得到输出状态  $X^{(k)}$ .

## 4 算法分析

### 4.1 算法的时间复杂度分析

假设一个粒子共有  $n$  个变量(本文  $n=26$ ),则计算概率图需要  $n$  重循环,为了进一步得到各个变量的值,又需要进行  $n$  次循环,从而粒子生成算法总共需循环  $2n$  次,时间复杂度为  $O(n)$ .在计算概率图过程中,获取图像特征以得到不变概率和遮挡概率占据主要时间开销.

### 4.2 算法特点分析

(1) 状态变量的多通道概率描述形式可以从不同角度描述状态变化的多层次性特征和多元性特征,从而将采样过程可以建立在更加有效、高效和统一的数据结构的基础之上;(2) 改善了粒子质量,有可能用少量的粒子描述后验状态概率的真实分布,为实现实时性手势跟踪目标创造了基本条件.假设手势只作全局刚性运动,根据本文算法,将产生极少部分非刚性运动粒子;(3) 假设手势发生遮挡问题,根据本文算法,被遮挡部分基本上不参与采样.

## 5 实验结果

对同一帧,取不同的粒子数,图 5~图 11 给出了本文算法和传统 PF 算法<sup>[13]</sup>的实验结果,其中,(a)(b)是传统 PF 算法的实验结果,(c)(d)是本文算法的实验结果,(a)是跟踪得到的 3D 人手模型,(b)是 3D 人手模型在对应帧图像上的投影.

从实验结果可知,在同等条件下,与传统 PF 算法相比,本文算法的跟踪精度得到改善,主要表现在:一是手势姿态更符合录像中操作者手势的实际变化情况,二是跟踪得到的 3D 人手模型在对应帧图像上的投影与帧图像具有更好的吻合度.

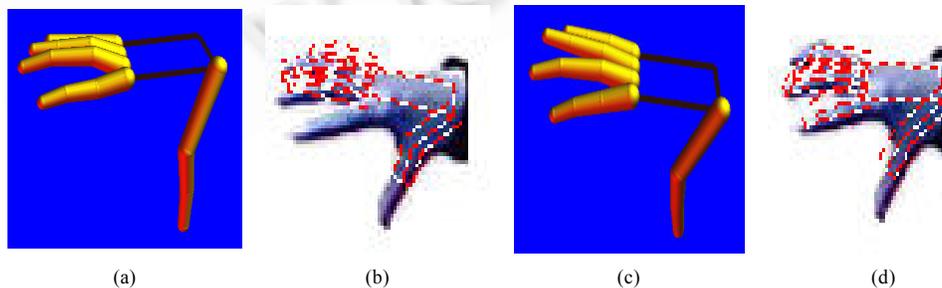


图 5  $N=3$  时的实验结果

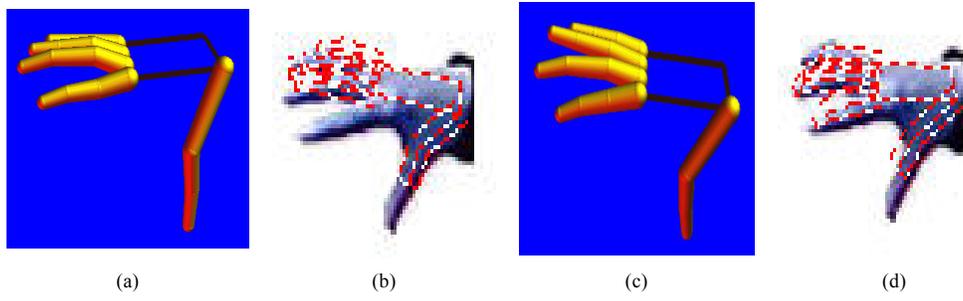


图 6  $N=5$  时的实验结果

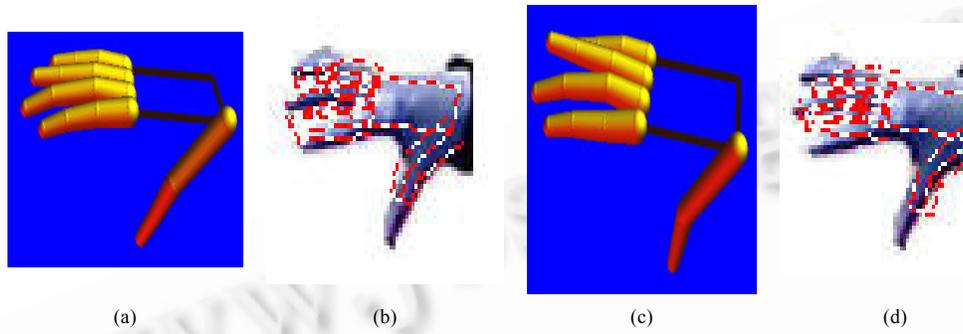


图 7  $N=7$  时的实验结果

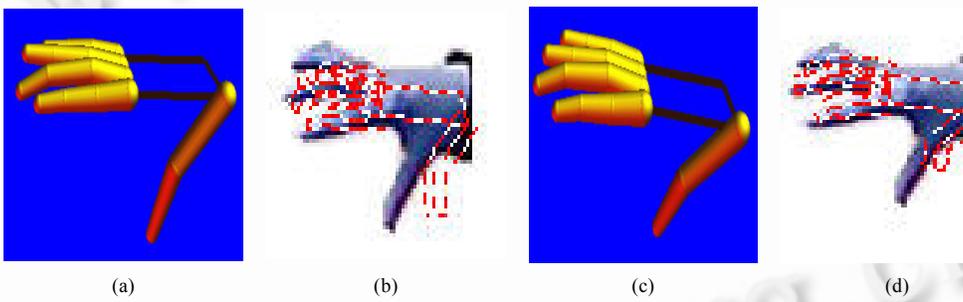


图 8  $N=9$  时的实验结果

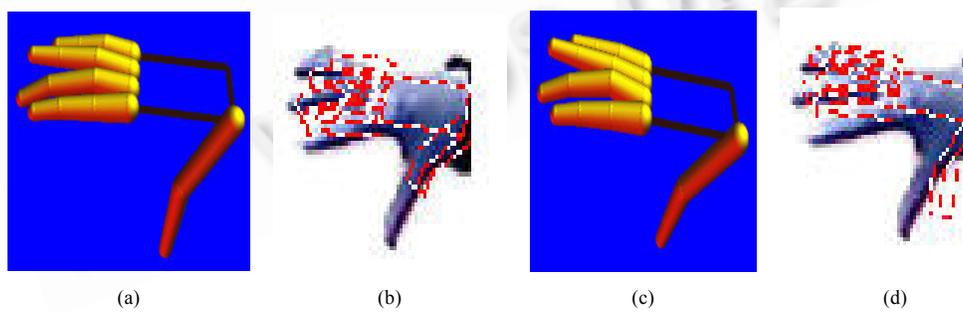


图 9  $N=11$  时的实验结果

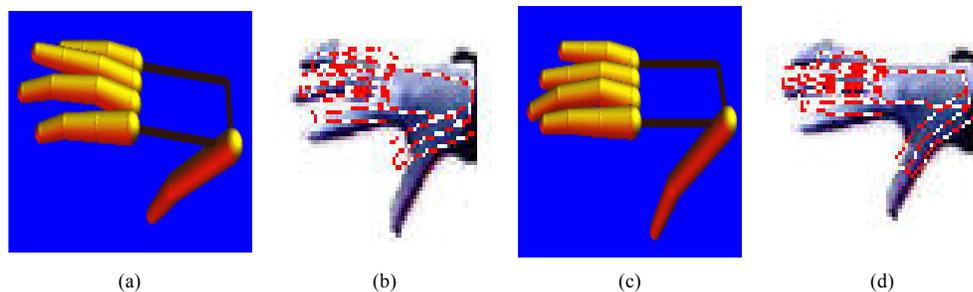
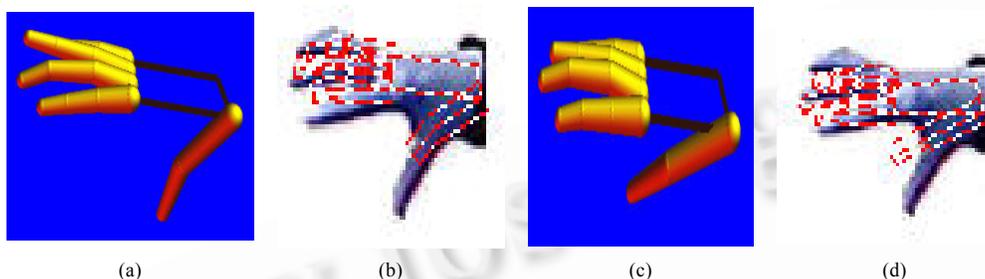
图 10  $N=13$  时的实验结果图 11  $N=15$  时的实验结果

表 1 中给出了在不同粒子数目情况下的计算时间开销(ms/frame).本文算法的时间开销平均降低 14%.

Table 1 Numbers used in our method and the classical PF algorithm

表 1 采用不同跟踪算法时间开销随粒子数目的变化情况

算法	粒子数								
	3	5	6	7	9	10	11	13	15
本文方法(ms)	254	413	320	448	525	434	534	576	512
传统方法(ms)	374	435	435	449	648	545	535	581	635

## 6 结论与展望

微观结构旨在从不同的角度、用不同的方法描述同一个状态变量,将采样过程建立在更加有效、高效和统一的数据结构的基础上.基于人手状态的多通道描述形式融入了状态的先验分布信息,是对状态变化规律的深层次描述,从而可以更快地提取反映系统概率特征的典型粒子,尽最大可能地避免产生不可能或可能性较小的样本,实际上达到降维的目的,用较少的粒子描述状态的后验概率分布,为粒子滤波条件下实时性处理创造了基本条件.

本文的研究仍有许多问题值得进一步探索.在基于自然人手的人机交互过程中,无处不打上操作者心理运动的烙印,显然,本文对操作者心理特征的研究只是局部的和肤浅的,这些心理特征与微观结构的深层次联系形式更需进一步研究.

### References:

- [1] Rehg JM, Kanade T. Visual tracking of high DOF articulated structures: An application to human tracking. In: Eklundh JO, ed. Proc. of the 3rd European Conf. on Computer Vision, Vol.801. Heidelberg: Springer-Verlag, 1994. 35–46.
- [2] Chen R, Liu GY, Zhao GY, Zhang J, Li H. 3D Human motion tracking based on sequential Monte Carlo method. Journal of Computer-Aided Design & Computer Graphics, 2005,17(1):85–92 (in Chinese with English abstract).
- [3] MacCormick J, Isard M. Partitioned sampling, articulated objects, and interface-quality hand tracking. In: Vernon D, ed. Proc. of

- the 6th European Conf. on Computer Vision, Vol.2. LNCS 4247, Heidelberg: Springer-Verlag, 2000. 13–19.
- [4] Deutscher J, Davidson A, Reid I. Automatic articulated partitioning of high dimensional search spaces associated with articulated body motion capture. In: Proc. of the Int'l Conf. on Computer Vision and Pattern Recognition, Vol.2. New York: IEEE Computer Society Press, 2001. 669–676.
- [5] Wu Y, Huang TS. Capturing articulated human hand motion: A divide-and-conquer approach. In: Zucker S, ed. Proc. of the 7th IEEE Int'l Conf. on Computer Vision. New York: IEEE Computer Society Press, 1999. 606–611.
- [6] Zhou HN, Huang TS. Tracking articulated hand motion with eigen dynamics analysis. In: Proc. of the 9th IEEE Int'l Conf. on Computer Vision. New York: IEEE Computer Society Press, 2003. 1102–1109.
- [7] Wachs J, Kartoun U, Edan Y, Stem H. Real-Time hand gesture telerobotic system using the fuzzy c-means clustering algorithm. In: Jamshidi M, ed. World Automation Congress (WAC). Albuquerque: TSI Press, 2002. 403–410.
- [8] Rehg JM, Kanade T. Digiteyes: Vision-Based human hand tracking. Technical Report, CMU-CS-93-220, School of Computer Science, Carnegie Mellon University, 1993.
- [9] Pan CH, Ma SD. Region-Based 3D motion tracking of hand. Journal of Image and Graphics, 2003,8(10):1205–1210 (in Chinese with English abstract).
- [10] Cui JS. Studies on three-dimensional model based posture estimation and tracking of articulated objects [Ph.D. Thesis]. Beijing: Tsinghua University, 2004 (in Chinese with English abstract).
- [11] Feng ZQ, Yang B, Chen YH, Zheng YW, Li Y, Wang ZH. Research on sampling methods in particle filtering based upon microstructure of state variable. Lecture Notes in Computer Science, 2008,5226:251–257.
- [12] Feng ZQ, Yang B, Li Y, Wang ZH, Zheng YW. Research on human hand tracking aiming at improving its accurateness. Journal of Computer Research and Development, 2008,45(7):1239–1248 (in Chinese with English abstract).
- [13] Gordon N, Salmond DJ, Smith AFM. Novel approach to nonlinear and non-Gaussian Bayesian state estimation. IEE Proc.-F, 1993,140(2):107–113.

#### 附中文参考文献:

- [2] 陈睿,刘国翌,赵国英,张俊,李华.基于序列蒙特卡罗方法的3D人体运动跟踪.计算机辅助设计与图形学学报,2005,17(1):85–92.
- [9] 潘春洪,马颂德.基于区域的手指三维运动跟踪.中国图像图形学报,2003,8(10):1205–1210.
- [10] 崔锦实.基于三维模型的关节式物体姿态估计与跟踪方法研究[博士学位论文].北京:清华大学,2004.
- [12] 冯志全,杨波,李毅,王中华,郑艳伟.以改善精度为目标的人手跟踪方法研究.计算机研究与发展,2008,45(7):1239–1248.



冯志全(1964—),男,四川绵阳人,博士,教授,主要研究领域为人手三维跟踪,图像处理,虚拟现实,人机交互.



唐好魁(1972—),男,讲师,主要研究领域为图像处理,图像分析,图像理解,生物特征识别.



杨波(1965—),男,博士,教授,博士生导师,主要研究领域为图像处理,智能计算,计算机网络.



徐涛(1979—),男,讲师,主要研究领域为计算机图形与图像处理.



郑艳伟(1977—),男,讲师,主要研究领域为图像处理.



李毅(1961—),女,副教授,博士生导师,主要研究领域为图像处理,自动化理论,计算机网络.