

一种改进的粒子滤波算法及其在说话人跟踪中的应用

曹洁¹ 李伟^{1,2}

¹(兰州理工大学 计算机与通信学院, 兰州 730050)

²(中国人民解放军 91666 部队, 浙江 舟山 316000)

E-mail: lwyz815@163.com

摘要: 针对非线性、非高斯系统状态的在线估计问题, 提出一种改进的粒子滤波算法, 该算法综合考虑"优选建议分布函数"和"重采样"两种并行改进滤波性能的方法. 首先通过 Unscented 卡尔曼滤波器产生系统的状态估计, 并在协方差预测阶段引入衰减记忆因子, 削弱滤波器对历史信息的依赖, 增强当前量测信息对滤波器的修正作用, 从而产生一个优选的建议分布函数, 有效抑制了粒子退化现象; 接着在重采样阶段引入 MCMC (Markov Chain Monte Carlo) 方法来构造马尔科夫链产生服从目标分布的粒子, 使样本更加多样化, 有效避免了粒子枯竭问题. 最后, 通过系统仿真及说话人跟踪实验, 证明了该算法的有效性.

关键词: 粒子滤波; 衰减记忆因子; 重采样; 说话人跟踪

中图分类号: TP391

文献标识码: A

文章编号: 1000-4220(2012)03-0664-05

Improved Particle Filter Algorithm with Application to Speaker Tracking

CAO Jie¹, LI Wei^{1,2}

¹(College of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)

²(PLA Troops of 91666, Zhoushan 316000, China)

Abstract: An improved particle filter based on "optimum proposed distribution function" and "resampling" two parallel improving filtering methods is proposed for the on-line estimation problem of non-Gauss nonlinear systems. In order to weaken the effect of historical information and enhance the effect of up-to-date measurement, we introduce attenuation memory factor for generating the important density function based on the Unscented Kalman Filter (UKF) for a better performance in inhibiting the particle degradation problems in the new algorithm. The new particle filter can solve the problem of sample impoverishment by introducing Markov Chain Monte Carlo into resampling stage to construct a Markov Chain and produce particles with target distribution. Finally, system simulation and speaker tracking experimental results show that the proposed new particle filter is superior to the standard particle filter on filtering performance.

Key words: particle filter; attenuation memory factor; resampling; speaker tracking

1 引言

建立智能化的室内人机交互环境是人工智能研究领域的一个重点, 有效促进了人与机器之间的自然交互, 也为后续的各项研究奠定了基础, 比如: 说话人跟踪、说话人识别、语音识别以及情感分析等, 其中, 说话人跟踪是后续各项研究的基础^[1-3].

说话人跟踪问题是一个典型的非线性、非高斯问题^[2], 传统的卡尔曼滤波方法已无法满足跟踪精度的需求. 文献[4-7]采用粒子滤波的方法对说话人进行跟踪, 取得了很好的效果. 粒子滤波是一种针对非线性、非高斯噪声系统的统计滤波方法, 其主要思想是通过一组加权的采样粒子来表示状态的后验概率密度. 随着计算机技术的高速发展, 受到越来越广泛的重视, 成为当前的研究热点. 但是, "粒子退化"问题一直是该算法走向实际应用的一大障碍, 为解决"粒子退化"问题,

研究人员提出了很多改进算法, 主要依赖思想就是"优选建议分布函数"和"重采样". 目前, 多数研究均选择状态转移函数 $p(x_k | x_{k-1})$ 作为建议分布函数, 但是该建议分布函数没有考虑最新的量测信息, 滤波精度较差; 而"重采样"方法虽然有效克服了粒子退化问题, 却引入了粒子枯竭 (Particle Impoverishment) 现象.

本文综合考虑"优选建议分布函数"和"重采样"两种并行改进滤波性能的方法, 提出了一种改进的粒子滤波算法, 即: 基于 MCMC (Markov Chain Monte Carlo) 的衰减记忆无味粒子滤波 (MCMC-MAUPF). 该方法首先在无偏粒子滤波 (Unscented Particle Filter: UPF) 的基础上引入衰减记忆因子 S , 削弱滤波器对历史信息的依赖, 增强当前量测对滤波器的修正作用, 从而产生一个优选的建议分布函数, 较好地克服了权值退化问题; 接着, 在重采样阶段采用 MCMC (Markov Chain Monte Carlo) 方法构造马尔科夫链产生服从目标分布

收稿日期: 2010-10-20 基金项目: 甘肃省自然科学基金项目 (1010RJZA046) 资助; 甘肃省教育厅研究生导师基金项目 (0914ZTB003) 资助; 甘肃省财政厅项目 (0914ZTB148) 资助. 作者简介: 曹洁, 女, 1966年生, 教授, 主要研究方向为智能交通系统、多传感器信息融合; 李伟, 男, 1982年生, 硕士研究生, 工程师, 主要研究方向为多传感器信息融合、说话人跟踪与识别.

的样本,有效保持了采样粒子的多样性,使其能够更加有效地逼近后验概率密度函数,以较少的采样粒子数获取较高的估计精度.仿真实验表明,基于MCMC的衰减记忆无味粒子滤波器的性能明显优于标准粒子滤波器及无味粒子滤波器.最后,将该方法应用到说话人跟踪系统中.

2 粒子滤波算法

2.1 标准粒子滤波算法

粒子滤波算法(Particle Filter: PF)为系统的状态估计问题提供了一个通用的框架,其基本思想主要是利用一系列随机采样的样本,来代替状态的后验概率分布.当粒子个数变得足够大的时候,通过这种随机抽样的方法就可以得到状态后验概率分布的有效近似.用 $\{x_{0:k}^i, \omega_k^i\}_{i=1}^N$ 表示系统后验概率密度函数 $p(x_{0:k}|z_{1:k})$ 的粒子集合,则有:

$$p(x_{0:k}|z_{1:k}) \approx \sum_{i=1}^N \omega_k^i \delta(x_{0:k} - x_{0:k}^i) \quad (1)$$

其中 $\{x_{0:k}^i, i=1, \dots, N\}$ 是支持样本集,相应的权值为 $\{\omega_k^i, i=1, \dots, N\}$,且满足 $\sum_{i=1}^N \omega_k^i = 1$,而 $x_{0:k} = \{x_j, j=0, \dots, k\}$ 表示到时刻 k 系统所有状态的集合.其相应的权值递推公式为:

$$\omega_k^i \propto \omega_{k-1}^i p(z_k|x_k^i) p(x_k^i|x_{k-1}^i) / q(x_k^i|x_{k-1}^i, z_k) \quad (2)$$

其中 $p(z_k|x_k^i)$ 为似然概率分布, $p(x_k^i|x_{k-1}^i)$ 为转移概率分布, $q(x_k^i|x_{k-1}^i, z_k)$ 为重要性概率密度分布.

利用粒子滤波器进行目标跟踪,需要建立相应的状态模型和观测模型.把目标跟踪问题抽象为对相应状态向量的估计.针对说话人跟踪问题,具体模型描述如下.

2.1.1 状态模型

假设说话人的状态为 $\chi_k = [x_k, \dot{x}_k, y_k, \dot{y}_k]^T$,其中 (x_k, y_k) 和 (\dot{x}_k, \dot{y}_k) 分别为说话人的位置和速度.考虑到说话人跟踪系统中观测方程的强非线性,以及说话人运动的随意性,本文采用适应性较强的布朗运动建立说话人运动模型,其相应的运动方程称为朗之万(Langevin)过程^[8].假设说话人在 x 与 y 两个方向上的运动是相互独立的,在 x 轴方向上,说话人运动模型的状态方程可以单独描述为:

$$\begin{bmatrix} \dot{x}_k \\ \dot{\dot{x}}_k \end{bmatrix} = \begin{bmatrix} 1 & \Delta T \\ 0 & a_x \end{bmatrix} \begin{bmatrix} x_{k-1} \\ \dot{x}_{k-1} \end{bmatrix} + \begin{bmatrix} 1 \\ b_x \end{bmatrix} u_x \quad (3)$$

其中 $a_x = e^{-\beta_x \Delta T}$, β_x 为常数; $b_x = v_x \sqrt{1 - a_x^2}$, v_x 为稳态均方根速度; u_x 为单位方差的高斯白噪声.

2.1.2 观测模型

在说话人跟踪系统中,观测量一般包括到达时间延迟、输出波束能量等特征,由于联合波束形成-相位变换(Steered Response Phase Transform: SRP-PHAT)声源定位算法将波束形成方法的鲁棒性、短时分析特性与相位变换方法对环境的不敏感性相结合,使声源定位系统对噪声和混响的灵敏度降低,提高了系统的鲁棒性和定位的精确性^[7,8].同时,由于该算法只需采用较短的数据进行分析,特别适用于对运动的说话人进行实时跟踪.因此,本文采用波束输出能量作为观测量.

假设声源信号为 $s(t)$,经多径传播后,第 i 个麦克风接收

到的信号 $m_i(t)$ 可以表示为:

$$m_i(t) = s(t) * h_i(t) + v_i(t) \quad (4)$$

其中 $v_i(t)$ 为噪声, $h_i(t)$ 是声源与第 i 个麦克风之间的冲激响应,它是麦克风位置和声源位置的函数,"*"为卷积运算符.

设声源的空间位置向量为 χ ,语音信号 $m_i(t)$ 的傅里叶变换为 $M(\omega)$,麦克风数目为 M ,直达路径中声源信号至第 i 个麦克风的传播时延为 τ_i ,SRP-PHAT声源定位算法的波束输出能量 $P(\chi)$ 定义为:

$$P(\chi) = \sum_{i=1}^M \sum_{j=1}^M \int_{-\infty}^{+\infty} \frac{M_i(\omega) M_j^*(\omega)}{|M_i(\omega) M_j^*(\omega)|} e^{j\omega(\tau_i - \tau_j)} d\omega \quad (5)$$

2.2 无味粒子滤波算法

标准粒子滤波算法在很大程度上依赖于模型,如果模型不准确或者量测噪声突然增大,标准采样粒子的权值方差会随着时间递增,从而产生权值退化现象,导致算法的滤波精度降低.目前,选择合理的建议分布函数是一种缓解退化现象的有效方法,研究人员利用UKF对非线性问题的处理能力来产生建议分布函数,相应的滤波器被称作无味粒子滤波器(UPF),其目的就是为了在建议分布函数中融入最新的量测信息,改善滤波效果.UPF算法的具体描述如下^[9].

2.2.1 初始化:

从先验概率 $p(x_0)$ 中采样 N 个粒子:

$$\hat{x}_0^i = E(x_0^i)$$

$$P_0^i = E[(x_0^i - \hat{x}_0^i)(x_0^i - \hat{x}_0^i)^T]$$

2.2.2 循环迭代:

① 重要性采样

在UKF算法的基础上对粒子进行更新,创新Sigma点:

$$\chi_{k-1}^i = [\hat{\chi}_{k-1}^i, \hat{\chi}_{k-1}^i \pm \sqrt{(n_k + \lambda) P_{k-1}^i}]$$

时间更新:

$$\chi_{k|k-1}^i = f(\chi_{k-1}^i), \hat{x}_{k|k-1}^i = \sum_{j=1}^{2L} \omega_j^m \chi_{k|k-1}^j$$

$$P_{k|k-1}^i = \sum_{j=0}^{2L} \omega_j^c [\chi_{j,k|k-1}^i - \hat{\chi}_{k|k-1}^i][\chi_{j,k|k-1}^i - \hat{\chi}_{k|k-1}^i]^T$$

$$z_{k|k-1}^i = h(\chi_{k-1}^i), \hat{z}_{k|k-1}^i = \sum_{j=1}^{2L} \omega_j^m z_{j,k|k-1}^i$$

量测更新(融入最新观测值)

$$P_{\hat{z}_k \hat{z}_k} = \sum_{j=1}^{2L} \omega_j^c [z_{j,k|k-1}^i - \hat{z}_{k|k-1}^i][z_{j,k|k-1}^i - \hat{z}_{k|k-1}^i]^T$$

$$P_{\hat{x}_k \hat{x}_k} = \sum_{j=1}^{2L} \omega_j^c [\chi_{j,k|k-1}^i - \hat{\chi}_{k|k-1}^i][z_{j,k|k-1}^i - \hat{z}_{k|k-1}^i]^T$$

$$K_k = P_{\hat{x}_k \hat{x}_k} P_{\hat{z}_k \hat{z}_k}^{-1}, \hat{x}_k^i = \hat{x}_{k|k-1}^i + K_k(z_k - \hat{z}_{k|k-1}^i)$$

$$P_k^i = P_{k|k-1}^i - K_k P_{\hat{z}_k \hat{z}_k}^{-1} K_k^T$$

$$\text{采样: } \hat{x}_k^i \sim q(x_k^i | x_{0:k-1}^i, z_{1:k}) = N(\hat{x}_k^i, P_k^i)$$

重要性权值更新为:

$$\omega_k^i = \omega_{k-1}^i \frac{p(z_k | x_k^i) p(x_k^i | x_{k-1}^i)}{q(\hat{x}_k^i | x_{0:k-1}^i, z_{1:k})}$$
 然后归一化权值

② 重采样: 根据有效粒子数情况判断粒子退化情况,加入粒子的重采样过程.

③ 状态估计: 根据最终得到的粒子集和相应的权值 $\{x_{1:k}^i, \omega_{1:k}^i\}$,则状态估值 $\hat{x}_k = \sum_{i=1}^N \omega_k^i x_k^i$

3 改进的粒子滤波算法

3.1 衰减记忆因子

由于UKF是增长记忆性滤波^[10],在系统的动态模型变化规律精确已知,且无需考虑计算误差的情况下,UKF能够给出一个随着量测信息的积累,精度不断提高的估计结果.但是在实际的动态系统中,模型的变化信息是很难完全掌握的.因此,随着系统的变化,必然导致相应参数的变化,而UKF在估计中还是按照没有变化的参数进行计算,直接导致了估计误差协方差的范数随着时间的增长不断减小,削弱了最新量测信息对状态估计的修正作用.而动态系统的最新量测包含了较为丰富的系统变化信息,因此,要增强滤波的稳定性和鲁棒性,必须增强最新量测信息对状态估计误差的修正作用,降低历史信息对估计的影响程度.因此,本文在UPF算法中引入记忆衰减因子S,生成更加合理的建议分布函数.

例如,考虑采用 t_k 时刻的观测噪声协方差矩阵来表示 t_n 时刻的观测噪声协方差矩阵:

$$R_{n|k} = E[V_k V_k^T | t = t_n] = S^{n-k} R(k) \quad n > k \quad (6)$$

这里 V_k 是 t_k 时刻的观测噪声, $R(k)$ 是 V_k 的协方差阵, S 是一个经验值,且 $S < 1$.上式表示,处于现在时刻的 t_n ,把过去时刻 t_k 的观测噪声协方差阵 $R(k)$ 中的每个元素都放大了 S^{n-k} 倍.也就是说,使过去时刻的观测噪声协方差随着背离时间的增加,以数 S^{n-k} 倍增大,这就相当于随着观测时间的增加,逐渐放弃过去的量测信息,降低了滤波器对历史信息的依赖,更加准确地逼近系统状态的后验概率分布.该方法反映在具体的算法上就是在UPF算法的量测更新过程中引入衰减记忆因子S,即:

$$P_{z_k \hat{x}_k} = S^{n-k} \sum_{i=0}^{2L} W_i^{(c)} [z_{j,k|k-1}^i - \hat{z}_{j,k|k-1}^i] [z_{j,k|k-1}^i - \hat{z}_{j,k|k-1}^i]^T \quad (7)$$

3.2 改进的粒子滤波算法

在标准粒子滤波算法中,采样粒子的退化程度可以通过有效粒子数 N_{eff} 来衡量:

$$N_{eff}(k) = \frac{N}{1 + \text{var}(\omega_k^* i)} = \frac{N}{E(\omega_k^* i)^2} \quad (8)$$

其中 $\omega_k^* i = \frac{p(x_k^i | y_{1:k})}{q(x_k^i | x_{k-1}^i, y_{1:k})}$.有效粒子数越小,表明权值退化现象越严重.克服粒子退化的另一种方法就是重采样^[11],

通过重采样将粒子采样计算仅限于权值较大的粒子,避免将计算量浪费在对后验概率贡献很小的粒子计算上.但是,该方法最终导致粒子枯竭现象的产生^[12],使得粒子集不能有效的逼近后验概率密度函数,从而影响滤波性能.在重采样阶段引入MCMC方法,不仅可以有效克服粒子退化现象,还可以提高粒子的采样效率,避免粒子枯竭现象.

MCMC算法是一种利用马氏链无记忆性的随机模拟方法.它的基本思想是构造一条非周期、不可约的马氏链,当迭代次数 $n \rightarrow \infty$ 时,得到的马氏链的分布可以任意接近状态的真实分布.目前产生这样马氏链的方法有很多,常用的主要有Gibbs采样和Metropolis Hastings(M-H)采样,其中Gibbs采样是M-H采样的一个特例,因此,本文考虑采用Metropolis

Hastings采样算法来实现MCMC转移,具体步骤如下:

- (1) 任选初始状态 $x_0 = x$;
- (2) 由转移核 $q(\cdot | x)$ 产生一个尝试移动 x' ;
- (3) 生成 $U \sim (0, 1)$ 随机数 u ,如果 $u \leq \alpha(x, x')$,则令 $X_1 = x'$,否则保持当前状态不变,即: $X_1 = X_0 = x$;
- (4) 重复上述步骤,依次生成 X_2, X_3, \dots, X_n .

下面考虑在衰减记忆无味粒子滤波(MAUPF)的重采样阶段融入MCMC中常用的MH算法,融合后的新算法称为MCMC-MAUPF,其具体的实现步骤如下:

3.2.1 初始化

在 $k = 0$ 时刻,从先验分布 $p(x_0)$ 中抽取粒子 $x_0^{(i)}$,并设置:

$$\hat{x}_0^i = E(x_0^i) \\ P_0^i = E[(x_0^i - \hat{x}_0^i)(x_0^i - \hat{x}_0^i)^T]$$

初始化的权值设为:

$$\omega_k^{(i)} = 1/N$$

3.2.2 预测和更新

在UKF算法的基础上按照式(7)引入衰减记忆因子S,并对粒子进行更新,创新Sigma点:

$$\chi_{k-1}^i = [\hat{x}_{k-1}^i \quad \hat{x}_{k-1}^i \pm \sqrt{(n_k + \lambda) P_{k-1}^i}]$$

时间更新:

$$\chi_{k|k-1}^i = f(\chi_{k-1}^i) \quad \hat{x}_{k|k-1}^i = \sum_{j=1}^{2L} \omega_j^m \chi_{j,k|k-1}^i$$

$$P_{k|k-1}^i = \sum_{j=0}^{2L} \omega_j^c [\chi_{j,k|k-1}^i - \hat{x}_{k|k-1}^i] [\chi_{j,k|k-1}^i - \hat{x}_{k|k-1}^i]^T$$

$$z_{k|k-1}^i = h(\chi_{k-1}^i) \quad \hat{z}_{k|k-1}^i = \sum_{j=1}^{2L} \omega_j^m z_{j,k|k-1}^i$$

量测更新(融入最新观测值)

$$P_{z_k \hat{x}_k} = S^{n-k} \sum_{j=1}^{2L} \omega_j^c [z_{j,k|k-1}^i - \hat{z}_{j,k|k-1}^i] [z_{j,k|k-1}^i - \hat{z}_{j,k|k-1}^i]^T$$

$$P_{x_k \hat{x}_k} = \sum_{j=1}^{2L} \omega_j^c [x_{j,k|k-1}^i - \hat{x}_{k|k-1}^i] [x_{j,k|k-1}^i - \hat{x}_{k|k-1}^i]^T$$

$$K_k = P_{x_k \hat{x}_k} P_{z_k \hat{x}_k}^{-1} \quad \hat{x}_k^i = \hat{x}_{k|k-1}^i + K_k(z_k - \hat{z}_{k|k-1}^i)$$

$$P_k^i = P_{k|k-1}^i - K_k P_{z_k \hat{x}_k}^{-1} K_k^T$$

采样: $\hat{x}_k^i \sim q(x_k^i | x_{0:k-1}^i, z_{1:k}^i) = N(\bar{x}_k^i, P_k^i)$

重要性权值更新为:

$$\omega_k^i = \omega_{k-1}^i \frac{p(z_k | \hat{x}_k^i) p(\hat{x}_k^i | x_{k-1}^i)}{q(\hat{x}_k^i | x_{0:k-1}^i, z_{1:k}^i)}$$

归一化的权重为:

$$\bar{\omega}_k^i = \omega_k^i / \sum_{i=1}^N \omega_k^i$$

3.2.3 重采样

给定重采样的初始阈值 N_{th} ,如果有 $N_{eff} < N_{th}$,则令 $\bar{\omega}_k^{(i)} = 1/N$,从重要性概率密度 $q(\hat{x}_k^i | x_{0:k-1}^i, z_{1:k}^i)$ 中采样得到粒子

x_k^{*i} ,如果 $u < \min\{1, \frac{p(z_k | x_k^{*i})}{p(z_k | \hat{x}_k^i)}\}$,就接受 x_k^{*i} ,即: $x_k^i = x_k^{*i}$,否

则丢弃 x_k^{*i} ,保留重采样的粒子 \hat{x}_k^i ,即: $x_k^i = \hat{x}_k^i$.

3.2.4 输出

经过上面算法的迭代,则滤波分布的经验概率分布,系统状态估计以及误差协方差的输出值分别为:

$$\hat{p}(x_k | z_{1:k}) = \frac{1}{N} \sum_{i=1}^N \delta(x_k - x_k^i)$$

$$\hat{x}_k = \frac{1}{N} \sum_{i=1}^N x_k^i$$

$$P_k = \frac{1}{N} \sum_{i=1}^N (\hat{x}_k - x_k^i) (\hat{x}_k - x_k^i)^T$$

4 实验结果与分析

4.1 非线性系统状态估计仿真实验

考虑到系统的非线性、非高斯性, 实验中首先采用了如下非线性系统模型^[13]对 PF、UPF 和 MCMC-MAUPF 三种算法的滤波性能进行了比较分析.

$$\begin{cases} x_t = 1 + \sin(0.4\pi t) + 0.5x_{t-1} + v_{t-1} \\ y_t = \begin{cases} 0.2x_t^2 + n_t, & t \leq 30 \\ 0.5x_t - 2 + n_t, & t > 30 \end{cases} \end{cases} \quad (9)$$

其中 过程噪声 $v_{t-1} \sim \text{Gamma}(3, 2)$, 量测噪声 $n_t \sim N(0, 0.00001)$

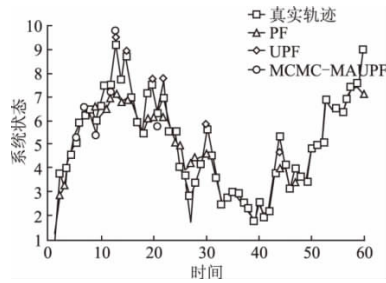


图 1 N = 40 时系统的状态估计
Fig. 1 The state estimation of N = 40

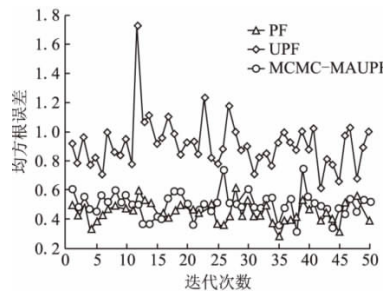


图 2 均方根误差曲线
Fig. 2 The RMSE curve

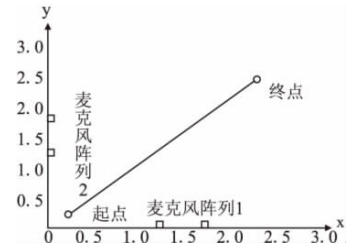


图 3 麦克风阵列的摆放位置
Fig. 3 The placement of microphone array

性能明显优于 UPF, 保持了较高的跟踪精度.

4.2 MCMC-MAUPF 在说话人跟踪中的应用

为进一步验证本文方法的有效性, 将其应用到基于麦克风阵列的说话人跟踪试验中. 试验中模拟普通会议室情况下的声学环境, 如图 3 所示, 在 $3\text{m} \times 3\text{m}$ 的房间内 x, y 两个方向上, 分别放置两组麦克风阵列 (每个阵列包含两个麦克风). 说话人在房间内沿对角线方向运动, 并保持持续发声状态. 麦克风阵列利用接收到的说话人语音信息, 在整个空间范围内搜索波束输出能量的最大值, 从而确定声源位置. 用如图 4 所示的语音信号作为说话人语音信号, 模拟说话人的跟踪过程. 麦克风之间的距离为 50cm . 说话人沿着如图 3 所示的房间对角线运动, 说话人运动的起点为 $(0.2, 0.2)$, 持续发音 6s , 运动的终点如图 3 所示. 麦克风收到的信号用 $\text{image}^{[5-7]}$ 模型仿真, 然后加上 20dB 的高斯白噪声获得. 以 16kHz 的频率对语音信号进行采样, 每 32ms 的语音作为一帧估计的延迟.

实验过程中, 各滤波器使用相同的输入观测值和初始状态, 以保证比较效果的客观性. 说话人运动模型中的各参数设定为: $\alpha_x = \beta_x = 10$, $v_x = 0.5\text{ms}^{-1}$, $\Delta T = 32\text{ms}$, 房间混响时间 $T_{60} = 200\text{ms}$, 初始状态值为 $\hat{X}_0 = (0.2, 0.2, 0.2, 0.1)^T$, 初始状态的估计协方差为 $\hat{P}_0 = \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix} \times \beta^2$, 其中 $C = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$,

采用两个不同阶次观测模型对系统状态进行观测, 观测时间为 $T = 60$. 在实验中, 将衰减记忆因子固定为经验常数: $S = 2$, 采样粒子数固定为 $N = 40$. 系统状态的估计采用均值估计器, 即:

$$\hat{x}_t = \frac{1}{N} \sum_{j=1}^N x_t^j \quad (10)$$

一次独立实验的均方误差定义为:

$$\text{MSE} = \left(\frac{1}{T} \sum_{t=1}^T (\hat{x}_t - x_t)^2 \right)^{1/2} \quad (11)$$

从图中可以看出, 在粒子数很低的情况下, 粒子滤波的跟踪性能明显下降, 有时候出现了明显的错误, 但是 UPF、MCMC-MAUPF 因为引入了最新的量测信息, 仍然能够保持很好的跟踪效果. 由于 MCMC-MAUPF 引入了衰减记忆因子, 同时进行了 Metropolis Hastings 采样算法, 在减少滤波器对历史信息的依赖性, 增强最新量测信息修正作用的同时, 有效增强了采样粒子数的多样性, 较好地克服了粒子枯竭现象, 其跟踪

FFT 变换长度为 $L = 1024$, 窗函数为汉明窗, 声源高度固定为 1.5m . 同时, 在文中引入相对跟踪误差作为衡量精度的标准, 相对跟踪误差越大, 则定位精度越低.

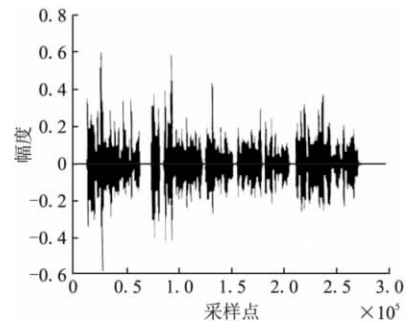


图 4 试验中采用的语音信号
Fig. 4 The speech signal used in the experiment

算法中采用系统状态转移函数 $p(x_t | x_{t-1})$ 作为重要性概率密度函数, 根据系统的状态方程生成随机采样粒子 $\{x_k^i\}_{i=1}^N$, 将波束输出能量 $p(x_k^i)$ 作为似然函数. 粒子权值的递推公式为:

$$\omega_t^i = \omega_{t-1}^i p(x_t^i) \quad (12)$$

说话人位置的最小误差均方差估计为:

$$\hat{\chi}_k = \sum_{i=1}^N \chi_i^k \omega_i^k \quad (13)$$

试验中,分别采用 PF、UPF 和 MCMC-MAUPF 三种滤波算法估计说话人的位置,为验证 MCMC-MAUPF 滤波性能的优越性,试验中仍然取采样粒子数 $N=40$,衰减记忆因子 $S=2$,三种方法在 x 轴方向的跟踪结果如表 1 所示,从表中可以

表 1 三种算法在 x 方向的跟踪误差比较

Table 1 The error comparison of three algorithms in x direction

时间(S)	0	1	2	3	4	5	
位置(m)	0.20	0.70	1.20	1.70	2.20	2.70	
PF 跟踪结果	0.15	0.75	1.26	1.77	2.32	2.88	
UPF 跟踪结果	0.24	0.68	1.25	1.74	2.27	2.79	
本文算法跟踪结果	0.22	0.71	1.22	1.67	2.26	2.77	
相对误差 (%)	PF	25.0	7.1	5.0	4.1	5.5	6.7
	UPF	20.0	2.8	4.2	2.4	3.2	3.3
	本文算法	10.0	1.4	1.7	1.8	2.7	3.0

看出:跟踪初始阶段,由于滤波器的初始化,三种滤波算法的跟踪精度都不是很好,随后在 1S 到 3S 的时间内,三种滤波器都能保持较好的跟踪效果,其中 UPF 和 MCMC-MAUPF 因为融入了最新的观测信息,其跟踪性能明显优于标准粒子滤波算法,而本文改进算法的跟踪精度明显优于 PF 和 UPF 两种算法;随着时间的推移,房间的混响越来越强,多径效应的影响增强,跟踪精度均有所下降,但是,本文算法仍然能够保持较高的精度。

5 结论

为了解决粒子滤波算法的粒子退化问题以及重采样方法导致的样本衰竭现象,本文综合考虑"优选建议分布函数"和"重采样"两种并行改进粒子滤波性能的方法,提出了一种改进的粒子滤波方法:MCMC-MAUPF 算法。该算法通过引入衰减记忆因子产生更加优秀的建议分布函数,有效抑制了粒子退化问题;同时,在重采样阶段引入 MCMC 采样算法,通过构造马尔科夫链有效提升了粒子的多样性和采样效率,较好地避免了因重采样引起的粒子枯竭现象。计算机仿真和说话人跟踪试验证明了新算法的有效性。下一步的研究中,将考虑引入采样粒子的在线自适应调整,以增强算法的实时性,同时,考虑将该方法应用到音视频融合的说话人跟踪系统中。

References:

[1] Shivappa S T, Rao B D, Trivedi M M. Audio visual fusion and tracking with multilevel iterative decoding: framework and experimental evaluation [C]. IEEE Journal of Selected Topics in Signal

Processing, 2010.

- [2] Shivappa S T, Trivedi M M, Rao B D. Person tracking with audio-visual cues using the iterative decoding framework [C]. In 5th IEEE International Conference on Advanced Video and Signal Based Surveillance 2008.
- [3] Shivappa S T, Trivedi M M, Rao B D. Hierarchical audio-visual cue integration framework for activity analysis in intelligent meeting rooms [C]. In IEEE CVPR Workshop: VISU'09, 2009.
- [4] Hou Dai-wei, Yin Fu-liang, Chen Zhe. Quasi-monte carlo filter for speaker tracking [J]. Acta Automatic Sinica, 2009, 35(7): 1016-1021.
- [5] Hou Dai-wen, Yin Fu-liang. An IMM particle filter method for speaker tracking [J]. Acta Electronica Sinica, 2010, 38(4): 835-841.
- [6] Jin Nai-gao, Yin Fu-liang. Quantum evolutionary particle filter algorithm with application to speaker tracking [J]. Signal Processing, 2008, 24(6): 982-987.
- [7] Jin Nai-gao, Yin Fu-liang, Chen Zhe. Microphone array speaker tracking using layered sampling particle filtering [J]. Acta Electronica Sinica, 2008, 36(1): 194-198.
- [8] Vermaak J, Blake A. Nonlinear filtering for speaker tracking in noisy and reverberant environments [C]. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, USA: IEEE, 2001: 3021-3024.
- [9] Zhang Miao-hui, Liu Xian-xing. Target tracking algorithm based on MCMC unscented particle filter [J]. System Engineering and Electronics, 2009, 31(7): 1810-1813.
- [10] Tarn T, Zarbosky J. A practical non-driking filter [J]. AIAA Journal, 1970, 8(6): 1127-1133.
- [11] Doucet A. On sequential simulation-based methods for bayesian filtering [R]. University of Cambridge, CUED/F-INFENG/TR. 310, 1998.
- [12] Andrieu C, Djuric P, Doucet A. Model selection by MCMC computation [J]. Signal Processing, 2001, 17(81): 19-37.
- [13] R Van der Merwe. A doucet the unscented particle filter advance in neural information processing systems [M]. MIT, 2000.

附中文参考文献:

- [4] 侯代文, 殷福亮, 陈喆. 基于拟蒙特卡洛滤波的说话人跟踪方法研究 [J]. 自动化学报, 2009, 35(7): 1016-1021.
- [5] 侯代文, 殷福亮. 基于粒子滤波的交互式多模型说话人跟踪方法 [J]. 电子学报, 2010, 38(4): 835-841.
- [6] 金乃高, 殷福亮. 量子进化粒子滤波算法及其在说话人跟踪中的应用 [J]. 信号处理, 2008, 24(6): 982-987.
- [7] 金乃高, 殷福亮, 陈喆. 基于分层采样粒子滤波的麦克风阵列说话人跟踪方法 [J]. 电子学报, 2008, 36(1): 194-198.
- [9] 张苗辉, 刘先省. 基于 MCMC 无味粒子滤波的目标跟踪算法 [J]. 系统工程与电子技术, 2009, (31) 7: 1810-1813.