

SVM 在光纤陀螺仪模型辨识中的应用

王宇, 吴峻

(东南大学 仪器科学与工程学院, 南京 210096)

摘要: 针对光纤陀螺仪 (FOG) 的误差特性以及在惯性系统中的运用特点, 提出了利用支持向量机的函数逼近的方法辨识光纤陀螺仪的输入输出模型。从惯性系统的角度, 对光纤陀螺仪的输出信号进行滤波, 从而较好地抑制传感器的噪声扩散, 为惯性系统整体精度的提高提供条件。通过实验表明, SVM 构建的模型能够较好地实现信号滤波功能, 并与传统的标定模型进行对比, 验证了在实时性要求较低的系统中, SVM 模型的输出精度要优于传统的标定模型输出精度。

关键词: 支持向量机; 函数逼近; 模型辨识; 光纤陀螺仪

中图分类号: U666.1

文献标志码: A

Model identification method based on SVM for FOG

WANG Yu, WU Jun

(School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China)

Abstract: In view of the error characteristic of FOG, the function approximation method of Support Vector Machine (SVM) was used to fit the input-output model of FOG. Then, from the perspective of inertial systems, the model was used to filter the output signal of FOG. This can significantly restrain the spread of the sensors' noise, and provide the conditions for improving the accuracy of the whole inertial system. The test result shows that the output accuracy of the model formed by SVM is better than that of the traditional calibration model when the system has low requirements for real-time.

Key words: support vector machine; function approximation; model identification; FOG

目前对模型辨识的研究普遍采用神经网络的方法, 如反馈网络、模糊神经网络、径向基函数网络以及小波神经网络。前面几种网络结构往往没有可靠的理论基础, 通常是基于经验建立的, 从而造成过学习, 过分依赖数据的质量和数量等问题。小波神经网络虽然有小波理论作为支撑, 避免了局部极小值的问题, 但是随着辨识维数的增加, 小波神经元的数量将迅速增加, 从而导致维数“灾难”。

SVM 是一种基于统计学习理论的学习机。它根据结构风险最小化的原理来自动学习目标模型的结构, 是真正意义上的结构可以自动选择的学习机。变结构的神经网络是利用经验启发式算法来修剪结构, 而 SVM 只要通过改变一个控制参数就可以连续调节模型复杂度和结构, 并且这种调节是自动进行的。SVM 的解是全局最优解, 具有很好的泛化性。因此, 选择 SVM 解决模型辨识问题是完全可行的。

1 SVM 的函数逼近原理

在函数逼近问题中, 由于真实函数 f 未知, 只能根据采样所得的样本 $(x_1, y_1), (x_2, y_2), \dots (x_r, y_r)$ (其中 $x_i \in R^d, y_i \in R$) 来求取近似函数 \hat{f} , 使得 f 和 \hat{f} 之间的距离最小。假设 \hat{f} 为线性模型:

$$\hat{f}(x) = \langle w, x \rangle + b \quad (1)$$

其中, w 为待定的常数。根据结构风险最小化准则, 用于函数逼近的支持向量机可以表示为^[1]:

收稿日期: 2007-11-08; **修回日期:** 2008-01-04

基金项目: 国防科技重点预研项目 (51309050402)

作者简介: 王宇 (1978—), 男, 东南大学博士研究生, 从事光纤惯性导航系统研究。E-mail: yu_wang_cn@sohu.com

联系人: 吴峻 (1965—), 男, 副教授, 从事惯性技术及其应用研究。E-mail: wujun@seu.edu.cn

$$\min \frac{1}{2} \cdot \|w\|^2 + c \cdot \sum_{i=1}^r (\xi_i + \xi_i^*) \quad s.t. \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (2)$$

其中, c 是平衡因子; ε 为函数的拟合精度; ξ_i 、 ξ_i^* 为松弛因子; $\|\cdot\|$ 表示向量模。在样本数较少时, 求解上面的支持向量机一般采用对偶理论^[4] (Duality Theory), 把它转化为二次规划问题, 并求得:

$$w = \sum_{i=1}^r (\alpha_i - \alpha_i^*) \cdot x_i \quad (3)$$

根据 Karush-Kuhn-Tucker (KKT) 条件, 在最优解处有^[2]:

$$\begin{cases} \alpha_i \cdot (\varepsilon + \xi_i - y_i + \langle w, x_i \rangle + b) = 0 \\ \alpha_i^* \cdot (\varepsilon + \xi_i^* + y_i - \langle w, x_i \rangle - b) = 0 \end{cases} \quad \begin{cases} (c - \alpha_i) \cdot \xi_i = 0 \\ (c - \alpha_i^*) \cdot \xi_i^* = 0 \end{cases} \quad (4)$$

由此可以得出, 位于不灵敏区内的样本点其对应 α_i 和 α_i^* 都等于零, 外部的点对应有 $\alpha_i = c$ 或 $\alpha_i^* = c$, 而在边界上, ξ_i 和 ξ_i^* 均为零, 因而 α_i 和 $\alpha_i^* \in (0, c)$, 从而有:

$$\begin{cases} b = y_i - \langle w, x_i \rangle - \varepsilon, \alpha_i \in (0, c) \\ b = y_i - \langle w, x_i \rangle + \varepsilon, \alpha_i^* \in (0, c) \end{cases} \quad (5)$$

由式 (5) 可计算出 b 的值。

与 $\alpha_i \neq 0$ 和 $\alpha_i^* \neq 0$ 相对应的样本 x_i (即在不灵敏区边界上或外面的样本) 称为支持向量, 从而有^[3]:

$$w = \sum_{i=1}^r (\alpha_i - \alpha_i^*) \cdot x_i = \sum_{i \in SV} (\alpha_i - \alpha_i^*) \cdot x_i \quad (6)$$

其中, SV 表示支持向量集。

在样本集内删除非支持向量后重新训练所得结果与原结果相同, 也就是说, 在逼近过程中只有那些支持向量才起到作用, 这也就是支持向量定义的来源。

2 基于支持向量机的 FOG 模型辨识

在基于 FOG 构建的惯性系统中, 传统的标定方法是首先进行 FOG 的标定, 得到 FOG 的输入输出关系, 再将其嵌入到惯性系统中。由于标定一般认为 FOG 的输入输出呈线性关系, 因此一般建立线性模型。但是在对 FOG 进行的标定实验中, 选取了不同速率分段进行实验, 标定结果显示标度因数和零漂在不同速率阶段均显示不同特性, 这也验证了输入输出的非线性关系。对陀螺仪的使用者来说, 相对关注的是陀螺仪输入输出之间的关系, 而不是从器件角度去建立分离的模型, 因此, 将光纤陀螺仪作为一个黑箱系统, 挖掘激励和响应之间的数据关系, 利用现代的统计理论解决这种系统的模型辨识, 是解决非线性系统模型较好的途径。

模型辨识就是要在系统输入输出之间建立一种函数关系。根据上述介绍的 SVM 的函数逼近原理可以解决这个问题。非线性函数逼近的基本思想^[5]是先通过非线性变换 $x \rightarrow \varphi(x)$, 将输入空间映射到高维的特征空间 (Hilbert 空间), 然后在特征空间中进行线性逼近, 即 $\hat{f}(x) = \langle w, \varphi(x) \rangle + b$, 从而根据式 (3) 得到:

$$w = \sum_{i=1}^r (\alpha_i - \alpha_i^*) \cdot \varphi(x_i) \quad (7)$$

在支持向量机中, 引入核函数 (Kernel Function) 来简化非线性逼近。核函数 $K(x, x')$ 满足

$$K(x, x') = \langle \varphi(x), \varphi(x') \rangle \quad (8)$$

由式 (1) (7) (8) 可得

$$f(x) = \sum_{i=1}^r (\alpha_i - \alpha_i^*) \cdot K(x_i, x) + b \quad (9)$$

此时 w 不能具体表示; 核函数的引入, 使得非线性函数逼近求解绕过特征空间, 直接在输入空间上求解, 从而避免了计算非线性映射 φ 。

SVM 由训练样本集和核函数完全描述, 因此采用不同核函数 $K(x, x')$ 就可以构造实现输入空间中不同类型的非线性决策面的学习机, 生成不同的支持向量算法。在实际问题中, 通常是直接给出核函数。本文中选取 RBF 径向基核函数 $K(x, x') = \exp(-\|x - x'\|^2 / \sigma^2)$ 作为高维特征空间中的内积函数, 其中 $\sigma = 0.01$ 。核函数确定后, SVM 学习机则只有一个参数可调整 (即平衡因子参数 C)。

将 FOG 视为一个黑箱系统, 运用上面介绍的 SVM 的函数逼近的方法, 求取输入输出之间的关系, 在 SVM 中建立

FOG 的模型。输入量包括了激励和外部环境条件，输出为 FOG 的响应。利用图 1 所示的一维训练关系图，建立多个变量输入的训练关系。求得黑箱系统模型后，再利用图 2 所示的解算关系，进行陀螺仪信号的处理。

训练 FOG 模型过程如下：

- ① 首先确定训练样本窗的宽度 n ，其中， n 越大，训练样本数就越多，包括的信息量就越大，但是系统训练时间就越长。
- ② 将训练样本进行预处理，剔除在数据采集阶段出现的数据奇异。
- ③ 将处理后的训练样本映射到前述的高维空间中。目前常用的映射核函数有多项式核和径向基核，后者的函数拟合能力要强于前者，因此，在模型辨识研究中，一般采用径向基核函数进行高维映射。
- ④ 采用上述训练样本得到 FOG 的学习机模型。

3 实验与结论

在实验室环境下，采用传统的陀螺仪标定模型和 SVM 辨识的模型，对某型 FOG 进行了测试。由于实验室环境下温度变化不大，实验和训练样本中均不考虑温度的影响。FOG 的原始输出为数字量，实验将 FOG 的激励和响应作为训练样本，形成单输入-输出的模型。对该 FOG 进行多组实验，抽取其中的一组作为测试数据，两种模型输出曲线如图 3、4 所示。

图 3 为动态环境下的输出曲线，上图为 SVM 的输出曲线，下图为采用传统标定方法得到的线性模型的输出曲线，

两条曲线的外形差异不大，但是从统计结果来看，SVM 的输出误差的平均值为 -0.0035 ($^{\circ}$)/s，标准差为 $0.007\ 029$ ($^{\circ}$)/s；线性模型输出误差的平均值为 -0.0073 ($^{\circ}$)/s，标准差为 0.02096 ($^{\circ}$)/s，前者较后者的数据离散度有较程度的提高。SVM 输出误差的结果为 0.0079 ($^{\circ}$)/s (RMS)，线性模型的输出误差为 0.0222 ($^{\circ}$)/s (RMS)。从数据离散度和体现综合精度的 RMS 结果来看，SVM 辨识的模型具有较好的抑止噪声的能力，使传感器的输出精度得到提高。图 4 显示的静态环境下的输出曲线与图 3 有相同的结论，其具体统计数据见表 1 所示。

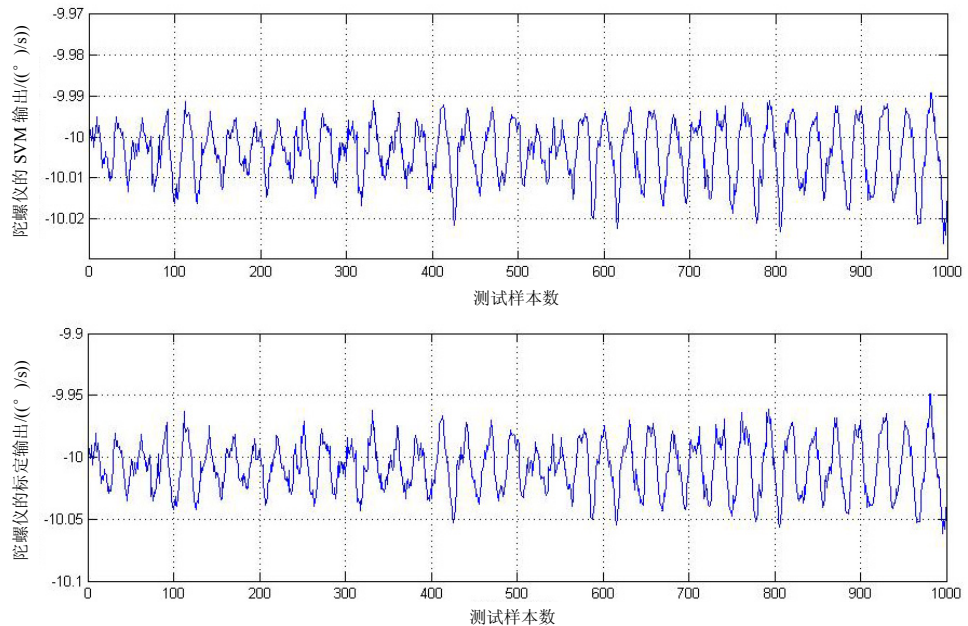


图 3 动态环境下的 FOG 输出曲线
Fig.3 FOG output curve in Dynamic Environment

表 1 静态环境两种模型的输出误差统计

Tab 1 Output error statistics of two models in static environment

静态实验	传统标定模型输出误差值/($^{\circ}$)/s)			SVM 辨识模型输出误差值/($^{\circ}$)/s)		
	平均值	标准差	RMS	平均值	标准差	RMS
1	-0.001 132	0.003 236	0.0034	-0.001 329	0.000 189 1	0.0013
2	-0.001 144	0.003 163	0.0034	-0.001 324	0.000 173 3	0.0013
3	-0.000 993	0.002 968	0.0031	-0.001 310	0.000 149 0	0.0013
4	-0.001 196	0.002 984	0.0032	-0.001 310	0.000 165 1	0.0013

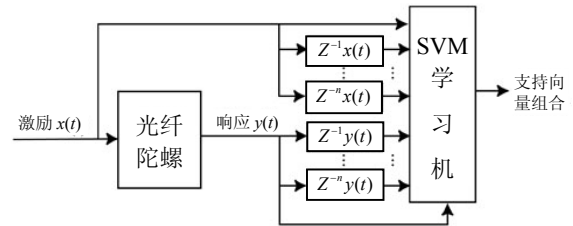


图 1 SVM 训练关系图

Fig.1 Training relationship of SVM

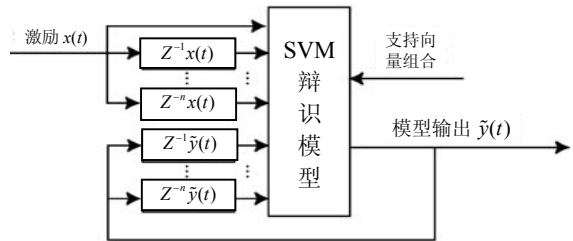


图 2 SVM 测试关系图

Fig.2 Test relationship of SVM

图 5 显示的是另一组实验的静态环境下的输出曲线。SVM 的输出误差的为 $0.025 (^{\circ})/s$ (RMS), 线性模型的输出误差为 $0.0033 (^{\circ})/s$ (RMS)。在这组实验中, 减少 SVM 学习机的训练样本, 力图做到快速完成训练, 快速输出。但是从测试结果来看, 由于训练样本的减少, 造成了学习机的“欠学习”的现象, 使得 SVM 的泛化能力下降, 消噪能力降低。

综上所述, 利用 SVM 的函数逼近的方法, 能够较好地辨识 FOG 的模型, 其输出精度优于传统的标定模型的输出精度。将 SVM 学习机置于惯性系统和传感器之间, 可以减少传感器噪声对惯性系统的干扰, 从而为提高惯性系统的整体精度打下基础。由于 SVM 训练样本需要时间, 因此目前这种方法适用于实时性和动态范围要求不高的环境中, 例如由 FOG 组成的寻北系统。如果将 SVM 应用于在线模型辨识, 那么训练时间是制约其应用的主要因素。缩短训练时间, 就要减小训练样本窗的宽度, 那么数据的信息量就会减少, 就会像图 6 的实验那样, 输出精度将受到影响。这也是基于 SVM 的在线模型辨识方法研究的重要内容。

参考文献 (References):

- [1] Vapnik V N. The nature of statistical learning theory[M]. New York: Springer-Verlag, 1995.
- [2] Suykens J A K. Nonlinear modeling and support vector machines[C]//IEEE Instrumentation and Measurement Technology Conference, Hungary, 2001, 3: 287-294.
- [3] 张学工. 关于统计学习理论与支撑向量机[J]. 自动化学报, 2000, 26 (1): 32-42.
ZHANG Xue-gong. Introduction to statistical learning theory and support vector machines[J]. Acta Automatica Sinica, 2000, 26 (1): 32-42.
- [4] 蔡艳宁, 胡昌华. 基于支持向量回归机的陀螺漂移预测模型[J]. 中国惯性技术学报, 2007, 15 (5): 594-597.
CAI Yan-ning, HU Chang-hua. Gyro drift prediction model based on support vector regression machine[J]. Journal of Chinese Inertial Technology, 2007, 15 (5): 594-597.
- [5] 王晓丹, 王积勤. 支持向量机研究与应用[J]. 空军工程大学学报, 2004, 5 (3): 49-55.
WANG Xiao-dan, WANG Ji-qin. Research and application of support vector machine[J]. Journal of Air Force Engineering University (Natural Science Edition), 2004, 5 (3): 49-55.

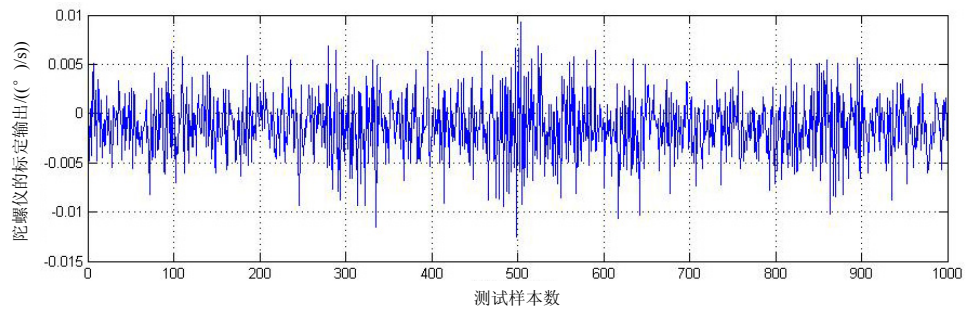
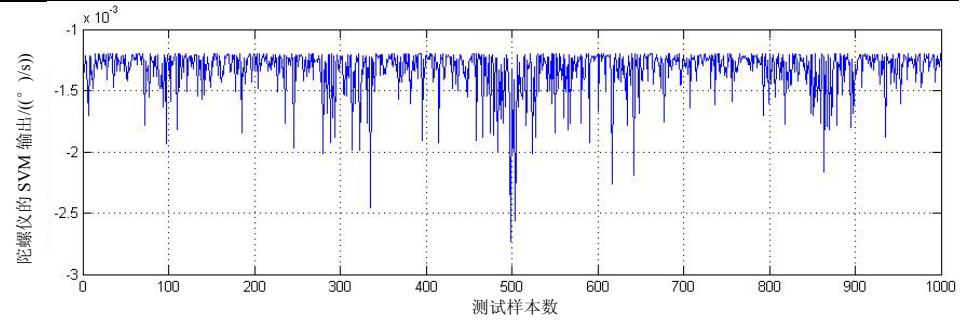


图 4 静态环境下的 FOG 输出曲线
Fig.4 FOG output curve in static Environment

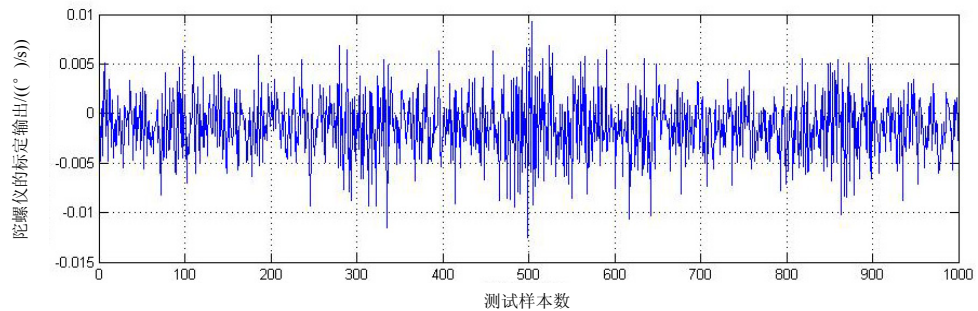
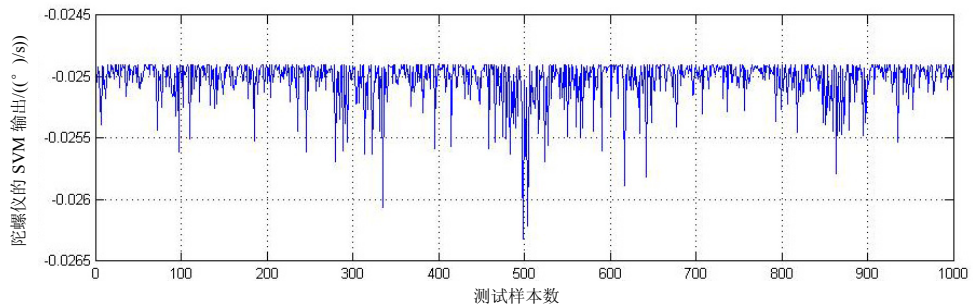


图 5 欠学习的模型输出曲线
Fig.5 Output curve of underfitting model