Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Nonlinear multifunctional sensor signal reconstruction based on least squares support vector machines and total least squares algorithm*

Xin LIU†, Guo WEI, Jin-wei SUN, Dan LIU

(*Department of Automatic Measurement and Control, Harbin Institute of Technology, Harbin 150001, China*)

†E-mail: xinliu@hit.edu.cn

Received Apr. 14, 2008; Revision accepted Oct. 10, 2008; Crosschecked Mar. 13, 2009

Abstract: Least squares support vector machines (LS-SVMs) are modified support vector machines (SVMs) that involve equality constraints and work with a least squares cost function, which simplifies the optimization procedure. In this paper, a novel training algorithm based on total least squares (TLS) for an LS-SVM is presented and applied to multifunctional sensor signal reconstruction. For three different nonlinearities of a multifunctional sensor model, the reconstruction accuracies of input signals are 0.001 36%, 0.031 84% and 0.504 80%, respectively. The experimental results demonstrate the higher reliability and accuracy of the proposed method for multifunctional sensor signal reconstruction than the original LS-SVM training algorithm, and verify the feasibility and stability of the proposed method.

Key words: Least squares support vector machine, Total least squares, Multifunctional sensor, Signal reconstruction
doi:10.1631/jzus.A0820282 **Document code:** A **CLC number:** TN98

INTRODUCTION

In the last decades, people have paid more attention to the development of multifunctional sensors, which is a new direction in modern sensor technology. Multifunctional sensors have been applied in the field of environmental perception (Sun and Shida, 2002), industry measurement (Yuji and Shida, 2000), aeronautics, astronautics, and micro-mechanical technology. In general, a multifunctional sensor can simultaneously detect several different non-electric signals and greatly reduce the size and consumption of the measurement system. The schematic structure of a general multifunctional sensor is shown in Fig.1, and this multi-variable transfer function model can be described as

$$y_i = f_i(x_1, x_2, \ldots, x_n), \quad i = 1, 2, \ldots, n, \tag{1}$$

where x_1, x_2, …, x_n are the input signals, y_1, y_2, …, y_n are the output signals, and f_i (i=1, 2, ..., n) describes the transfer function of sensitive component i, which in general is nonlinear. The estimation of the input signals can be obtained through the signal reconstruction algorithm, and this process is the so-called sensor signal reconstruction.

By now, many reconstruction algorithms have been proposed to estimate the regression function in multifunctional sensor signal reconstruction (Flammini *et al*., 1999; Sun *et al*., 2004; Liu *et al*., 2007). These methods are based on the empirical risk minimization (ERM) principle, which ensures the actual risk close to the empirical risk for a large sample size. For a small sample size, however, minimizing the empirical risk cannot guarantee a small value for the actual risk, and thus this leads to overfitting and poor generalization capabilities (Vapnik, 1999). A support

* Project supported by the National Natural Science Foundation of China (Nos. 60772007 and 60672008), and China Postdoctoral Science Foundation (No. 20070410258)

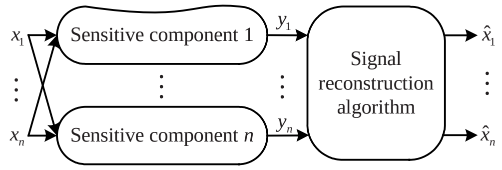


Fig.1 Schematic structure of a multifunctional sensor

vector machine (SVM) and its modified algorithms are new machine learning methods based on statistical learning theory and the structural risk minimization principle, and it can efficiently deal with the small sample size problem, restrain overfitting and improve the generalization capability (Vapnik, 1998).

A least squares support vector machine (LS-SVM), being a recently reported least squares version SVM, involves equality constraints instead of inequality constraints and adopts a least squares cost function. Therefore it expresses the training by solving a set of linear equations instead of a quadratic programming problem, which greatly reduces computational cost (Suykens and Vandewalle, 1999). In this paper, we present a novel solution for LS-SVM based on total least squares (TLS). This method enables us to solve the sensor signal reconstruction problem in the condition where input signals and output signals are both contaminated by noise.

This paper is organized as follows. In Section 2 we briefly review the SVM and LS-SVM algorithms, and then describe the solution procedure for LS-SVM based on TLS. In Section 3, we build up a simulation model of a multifunctional sensor and analyze the nonlinearity degree of this multifunctional sensor transfer function by a simple nonlinearity measurement, and then we discuss the experimental results for different nonlinearity levels obtained by the proposed approach and the original one. We conclude the paper in Section 4.

THEORY AND ALGORITHM

Least squares support vector machine

SVM was originally developed for pattern recognition (Cortes and Vapnik, 1995) and then Vapnik (1998) extended the results to solve the regression problem by promoting a novel so-called ε-insensitive loss function: $|y-f(\boldsymbol{x})|_\varepsilon=\max\{0, |y-f(\boldsymbol{x})|-\varepsilon\}$. Consider a nonlinear regression problem. The dependence of a scalar output y on an input vector $\boldsymbol{x}$ can be described by

$$y = g(\boldsymbol{x}) + v. \quad (2)$$

The function $g(\cdot)$ and the statistics of independent noise v are unknown. All that is available is a training dataset $\{\boldsymbol{x}_i, y_i\}_{i=1}^N$, where $\boldsymbol{x}_i\in\mathbb{R}^n$ is a sample of the input vector $\boldsymbol{x}$ and $y_i\in\mathbb{R}$ is the corresponding value of the equation output. The problem is to provide an estimation of the dependence of y on $\boldsymbol{x}$. And the Vapnik's SVM is intended to estimate the following function:

$$f(\boldsymbol{x}) = \boldsymbol{w}^{\mathrm{T}}\varphi(\boldsymbol{x}) + b, \quad \boldsymbol{x}\in\mathbb{R}^n, \; b\in\mathbb{R} \quad (3)$$

by minimizing the regularized risk functional:

$$\frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} + C\sum_{i=1}^{N}\left|y_i - f(\boldsymbol{x}_i)\right|_\varepsilon, \quad (4)$$

where the nonlinear function $\varphi(\cdot)$ maps the input data into a higher dimensional feature space. However, this function is constructed in an implicit way, and C is the regularization parameter that controls the trade-off between minimizing the model complexity and the empirical risk. Minimizing Eq.(4) is equivalent to the following constrained optimization problem by introducing two sets of non-negative slack variables $\{\xi_i\}_{i=1}^N$ and $\{\xi_i^*\}_{i=1}^N$:

$$\min_{\boldsymbol{w},b,\xi,\xi^*} J(\boldsymbol{w},\xi,\xi^*) = \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \quad (5)$$

s.t.

$$-\varepsilon - \xi_i^* \le \boldsymbol{w}^{\mathrm{T}}\varphi(\boldsymbol{x}_i) + b - y_i \le \varepsilon + \xi_i, \quad \varepsilon, \xi_i, \xi_i^* \ge 0. \quad (6)$$

The slack variables measure the deviation of data outside the ε-insensitive tube, and they are penalized in Eq.(5). It is obvious that training an SVM amounts to solving a complex quadratic programming problem (Smola and Scholkopf, 2004). To avoid the complicated computation, Suykens *et al*. developed a least squares version of SVM for a classification and nonlinear function estimation problem (Suykens and Vandewalle, 1999; Suykens *et al*., 2002). The main difference between LS-SVM and SVM is that the quadratic loss function is taken with equality constraints. Namely, the LS-SVM is formulated as follows:

$$\min_{\boldsymbol{w},b,\nu} J(\boldsymbol{w},\nu) = \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} + \frac{\gamma}{2}\sum_{i=1}^{N}\nu_i^2 \tag{7}$$

s.t.

$$y_i = \boldsymbol{w}^{\mathrm{T}}\varphi(\boldsymbol{x}_i) + b + \nu_i,\ \ i = 1,2,\ldots,N. \tag{8}$$

And the solution to the minimization problem can be obtained through solving the following linear equations by utilizing Lagrange multipliers techniques:

$$\begin{bmatrix} 0 & \mathbf{1}_N^{\mathrm{T}} \\ \mathbf{1}_N & \boldsymbol{\Omega} + \gamma^{-1}\boldsymbol{I} \end{bmatrix}\begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}, \tag{9}$$

where $\boldsymbol{y}=[y_1, y_2, \ldots, y_N]^{\mathrm{T}}$, $\mathbf{1}_N=[1, 1, \ldots, 1]^{\mathrm{T}}$, $\boldsymbol{\alpha}=[\alpha_1, \alpha_2,\ldots, \alpha_N]$ is the Lagrange multipliers vector, and $\Omega_{ij}=(\varphi(\boldsymbol{x}_i))^{\mathrm{T}}\varphi(\boldsymbol{x}_j)=K(\boldsymbol{x}_i, \boldsymbol{x}_j)$. The kernel function $K(\cdot, \cdot)$ should satisfy Mercer's condition, which can reduce the expensive calculations of the inner product in the high-dimensional feature space and has no need for the explicit form of the nonlinear mapping. The typical choices of the kernel function include linear kernel, polynomial kernel, sigmoid kernel, and RBF kernel. Hence the LS-SVM for function estimation is formulated as

$$y(\boldsymbol{x}) = \sum_{i=1}^{N}\alpha_i K(\boldsymbol{x}, \boldsymbol{x}_i) + b, \tag{10}$$

where α_i and b are the solutions to Eq.(9). Therefore LS-SVM solves a set of linear equations instead of a computational quadratic programming problem. Furthermore, in the case of LS-SVM, the additional tuning parameters are less than those for standard SVM, which can significantly reduce the level of complexity control for model selection.

Total least squares algorithm for LS-SVM

Eq.(9) can be represented in the form

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{11}$$

with $\boldsymbol{A} = \begin{bmatrix} 0 & \mathbf{1}_N^{\mathrm{T}} \\ \mathbf{1}_N & \boldsymbol{\Omega} + \gamma^{-1}\boldsymbol{I} \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix}$. As for the square and full rank matrix $\boldsymbol{A}$, the unique solution can be explicitly obtained by matrix inversion. However, the kernel matrix $\boldsymbol{\Omega}$ and observation vector $\boldsymbol{b}$ consist of measurement signals, which are rarely known exactly and inevitably contaminated by observational noise. Thus the matrix $\boldsymbol{A}$ is often nearly ill-conditioned in practice. Though the inverse of a badly conditioned matrix is possible, the solution often exhibits numerical instabilities, which means that small changes in the matrix and observation vector can lead to large changes in the computed solution. Therefore, the TLS method is much more reasonable and stable in the case of noise data by considering the perturbation bias of both matrix and observation vector (Rahman and Yu, 1987).

To solve Eq.(11) by means of the TLS strategy, we suppose $\boldsymbol{E}$ and $\boldsymbol{e}$ are the perturbation effects of $\boldsymbol{A}$ and $\boldsymbol{b}$ respectively, and the TLS algorithm looks for an optimal solution to the following corrected equation that minimizes the perturbation effects on the given $\boldsymbol{A}$ and $\boldsymbol{b}$:

$$(\boldsymbol{A} + \boldsymbol{E})\boldsymbol{x} = \boldsymbol{b} + \boldsymbol{e}. \tag{12}$$

Obviously Eq.(12) can be transformed as follows:

$$(\boldsymbol{B} + \boldsymbol{D})\boldsymbol{z} = \mathbf{0}, \tag{13}$$

where $\boldsymbol{B}=[-\boldsymbol{b}, \boldsymbol{A}]$, $\boldsymbol{D}=[-\boldsymbol{e}, \boldsymbol{E}]$, $\boldsymbol{z} = \begin{bmatrix} 1 \\ \boldsymbol{x} \end{bmatrix}$. Then the TLS solution to Eq.(13) can be formulated as a constrained optimization problem:

$$\min_{\boldsymbol{D},\boldsymbol{x}} \|\boldsymbol{D}\|_{\mathrm{F}}^2 \quad \text{s.t.}\ \ \boldsymbol{b} + \boldsymbol{e} \in \mathrm{range}(\boldsymbol{A} + \boldsymbol{E}), \tag{14}$$

where $\|\boldsymbol{D}\|_{\mathrm{F}}$ denotes the Frobenius norm of matrix $\boldsymbol{D}$, and the solution can be calculated by singular value decomposition (Rahman and Yu, 1987). Actually, it should be noted that the TLS solution to Eq.(11) is the optimal solution to Eq.(12), while the matrix inversion method calculates the solution to Eq.(11) in a straightforward manner.

RESULTS AND ANALYSIS

A simple nonlinearity measurement

It is clear that the performance of the reconstruction algorithm is related to the nonlinearity degree of the function Eq.(1). However, this important parameter is very difficult to discuss or measure explicitly. Box (1971) proposed a formula to estimate the bias between the estimates of the parameters and their true parameter values in the least squares estimators, which can be viewed as an indication of the extent of nonlinear behavior of a given model (Ribeiro *et al*., 2005). Bates and Watts (1980) developed a new nonlinearity measurement method based on the geometric concept of curvature, which showed the intrinsic and parameter effect curvatures associated with a given model (Moreira *et al*., 2006). While useful in many ways, both measures are less useful in the context of measuring nonlinearity in the case of multifunctional sensor signal reconstruction. Since the dependence of the function set Eq.(1) is usually unknown, it is therefore very difficult or of high computational complexity to investigate the first- and second-order derivatives of the function, which are needed for both measures.

Obviously, when unaware of the true functional form of the function set Eq.(1), we can estimate a linear model and judge its performance by some standard statistical criterion. One criterion, root mean square error (RMSE), will deteriorate when the function deviates more and more from a linear model. It is clear that the errors in the linear model will arise entirely as the result of the nonlinearity; the larger errors mean the more nonlinearity. When the function is actually linear, all errors will disappear or near zero. Therefore, the RMSE can be used as a measure of nonlinearity.

Consider a function in which the dependence of a scalar output on a vector input is described by

$$Y = f(\boldsymbol{X}). \tag{15}$$

The function $f(\cdot)$ is unknown. All that is available is a set of dataset $\{\boldsymbol{X}_i, Y_i\}_{i=1}^{N}$. The problem is to find a linear function that optimally approximates Eq.(15). Here, the LS-SVM with linear kernel is used, and then the optimal linear estimation of Eq.(15) can be obtained as follows:

$$\hat{Y} = f^{\text{Linear-LSSVM}}(\boldsymbol{X}). \tag{16}$$

Therefore, the measurement of nonlinearity can be formulated as

$$N_{\text{m}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_i - \hat{Y}_i}{Y_i}\right)^2}, \tag{17}$$

where Y_i and $\hat{Y}_i$ are the ith actual value and estimated value respectively, and the error criterion used here is the relative error (RE) instead of the residual error. Actually, the residual error is not suitable for assessing the accuracy of estimation. The residual error of prediction usually increases with the magnitude of the actual value. However, a small residual error does not guarantee good estimation accuracy, because it does not take into consideration the magnitude of the actual value and thus penalizes the prediction of a large value. Thus it is not adequate for comparing the regression accuracy between a large actual value set and a small one. While the RE reflects the ratio of the estimated value to the actual value, which takes into account the magnitude of the actual value and allows the larger absolute error for a larger actual value, the RE is more suitable for the assessment of the regression accuracy.

Simulation model and its nonlinearity analysis

To verify the feasibility of the proposed method, a physical model of a two-input two-output multifunctional sensor used in the experiment has been built up and shown in Fig.2.

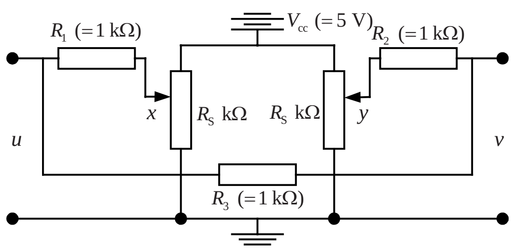


Fig.2 Circuit model of a two-input two-output sensor

The input signals are x and y, which represent the ratio of the slide resistor lower side resistances to the entire resistances. The output signals are u and v (V). According to KCL, the system transfer function can be described as follows:

$$\begin{cases} u/\text{V} = \dfrac{5[2x + y + R_{\text{S}}xy(2 - x - y)]}{3 + R_{\text{S}}[y(1-y) + x(1-x)]}, \\ v/\text{V} = \dfrac{5[2y + x + R_{\text{S}}xy(2 - x - y)]}{3 + R_{\text{S}}[y(1-y) + x(1-x)]}. \end{cases} \tag{18}$$

The resistance of the slide resistor, R_{S} kΩ, can control the nonlinearity of the transfer function. For instance, the nonlinear transfer function becomes a linear function when R_S=0, and it can be detected from experiments that the nonlinearity of the transfer function is growing with increasing R_S. To illustrate the influence of parameter values on model nonlinearity and the usefulness of the proposed nonlinearity measurement, the measurement of nonlinearity and box plots of the RE of the reconstructed input signal x for different parameter values are shown in Fig.3. As for the symmetry properties of Eq.(18), only the reconstructed signal x is discussed. Here, the training input set (x_i, y_i) is a Cartesian product of two input signal sets, which are both composed of 17 equally spaced data points over the interval (0.1, 0.9). Thus we obtained 289 training samples (x_i, y_i, u_i, v_i) generated from the above function for different R_S's. And similar to the training dataset, the test input set is also a Cartesian product that is composed of 81 equally spaced data points in the interval (0.1, 0.9). Therefore we obtained 6561 related test samples.

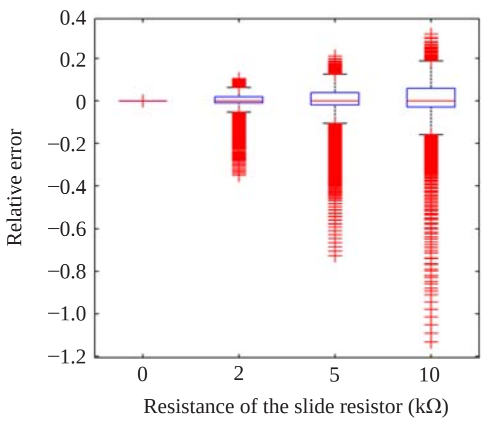


Fig.3 Box plots of the relative error of *x* for different R_S's

As shown in Fig.3, each box plot is based on the results of the test dataset of varying R_S's. For R_S=0, the transfer function Eq.(18) becomes a linear model and we notice that the range of RE is very close to zero and measurement of nonlinearity (1.42E−009) is also approximately zero, which implies that the estimated values from LS-SVM with linear kernel perfectly match the actual values. For other parameter values, the nonlinearity measurement and the range of the RE increase quickly with growing R_S (the measurement of nonlinearity is 0.0349, 0.0675 and 0.0982 for R_S=2, 5 and 10, respectively), which means that a higher parameter value causes a higher nonlinearity. The practical results are in accord with the previous analysis. Furthermore, the proposed measure of nonlinearity is nearly zero for a linear function, non-zero for a nonlinear function, and increases with the deterioration of nonlinearity. Therefore, it could be a useful tool to measure the nonlinearity degree of an unknown function.

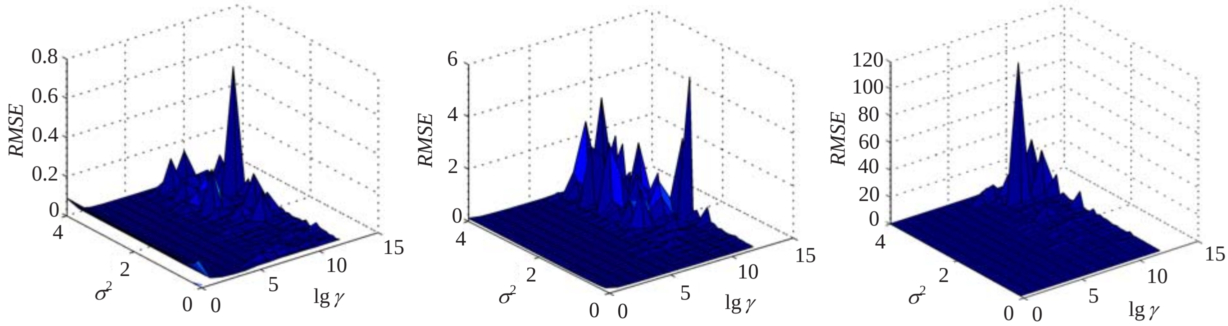
Experiment results and discussion

In the previous section we presented the theoretical results for TLS solution to LS-SVM and analyzed the nonlinearity of the simulation model, which makes the need evident to verify whether the proposed method can lead in practice to a better generalization performance than the original one for different nonlinearity levels. Furthermore, it can be found that the generalization performance of LS-SVM depends on a good setting of the regular parameter and kernel parameter. For a more accurate comparison, we numerically compute the generalization error with an RBF kernel, here the RMSE as a function of regular parameter γ and kernel parameter σ. The RMSE is calculated according to

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \hat{x}_i}{x_i}\right)^2}, \qquad (19)$$

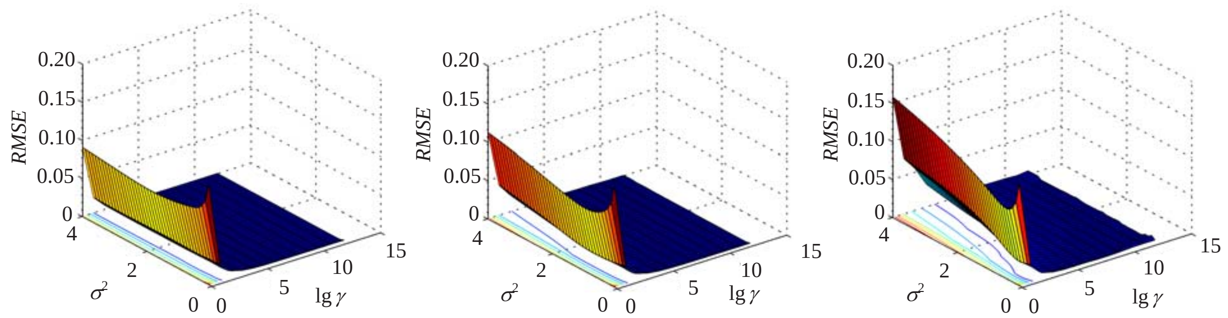
where n is the total number of test sets, x_i and $\hat{x}_i$ are the ith test input signal and related reconstructed signal calculated by LS-SVM with TLS solution or original solution, respectively. In the case of LS-SVM, the generalization performance is influenced by the choice of (γ, σ) when using an RBF kernel.

The experimental setup is the same as in Fig.2. A total of 289 examples were used as the training dataset, and 6561 examples as the test set. Then we built the regression function with the given set of parameters {γ, σ}, training the function for two different solutions to the training set. The performance of the parameter set can be measured by the RMSE on the test set. In particular, we plotted the generalization error of two solutions versus parameters for different nonlinearity levels. The experiment results are shown in figures as surface plots, where Figs.4a~4c are the RMSE of input x using the original solution versus parameters for different nonlinearity degrees, and Figs. 5a~5c are the RMSE of input x using TLS solution versus parameters for different nonlinearity levels.



(a) (b) (c)

Fig.4 RMSE of x using LS-SVM vs parameters σ and γ for different R_S's
(a) R_S=2; (b) R_S=5; (c) R_S=10



(a) (b) (c)

Fig.5 RMSE of x using TLS solution vs parameters σ and γ for different R_S's
(a) R_S=2; (b) R_S=5; (c) R_S=10

As shown in Figs.4a~4c, the risk surfaces are flat and smooth with respect to the whole kernel parameter range and the small regular parameter range, while the risk surfaces show rapid deterioration with respect to the whole kernel parameter range and the large regular parameter range. Furthermore, the shape and range of risk surfaces with respect to the small regular parameter region almost remain unchanged with increasing nonlinearity levels, while the shape of risk surfaces dramatically changes and the range of risk surfaces rapidly increases, such as the range of RMSE from 0 to 0.8 for R_S=2, from 0 to 6 for R_S=5, and from 0 to 120 for R_S=10, with respect to the large regular parameter region.

As shown in Figs.5a~5c, the risk surfaces are flat and smooth and the shape of risk surfaces almost remain unchanged with respect to the whole parameter range for all nonlinearity levels, which means that the neighboring parameter values result in the neighboring risk values and this fact guarantees the stability of the proposed solution for all nonlinearity levels over the whole range of parameters. The range of risk surfaces increases slowly with the increasing nonlinearity levels, such as the range of RMSE from 0 to 0.1 for R_S=2, from 0 to 0.15 for R_S=5, and from 0 to 0.2 for R_S=10.

From the results above we can conclude that the TLS solution is more stable and accurate than the original one. Therefore, it can guarantee a lower generalization risk.

Finally, we can choose the parameter set that performed best. Here two evaluation criteria are used, i.e., the RMSE and the maximum relative error (MRE), and the MRE is defined as

$$MRE = \max_{1\le i\le n}\left|(x_i - \hat{x}_i)/x_i\right|. \qquad (20)$$

The optimal parameters are set and the related results are shown in Table 1, where it can be found that both the MRE and RMSE calculated by the proposed solution are smaller than the original ones for all nonlinearity levels. Moreover, for input signals x and y, the optimal parameter set and two evaluating criteria calculated by the proposed solution are the

Table 1 Reconstruction results comparison of two methods with optimal parameters setting

Nonlinear level	Method	Input signal *x*				Input signal *y*			
		lgγ	σ^2	*MRE* (%)	*RMSE* (%)	lgγ	σ^2	*MRE* (%)	*RMSE* (%)
R_S=2	Original	6	3.5	0.13501	0.02533	6	2.8	0.09989	0.01791
	Proposed	12	1.1	0.00136	0.00012	12	1.1	0.00136	0.00012
R_S=5	Original	4	0.9	0.56499	0.12385	5	0.8	0.36083	0.05463
	Proposed	10	0.4	0.03184	0.00182	10	0.4	0.03184	0.00182
R_S=10	Original	4	0.3	1.11167	0.14249	4	0.3	1.02709	0.18832
	Proposed	6	0.2	0.50480	0.04318	6	0.2	0.50480	0.04318

same for all nonlinearity levels, and this fact is in accord with the symmetry properties of Eq.(18). While for the original solution, MRE and RMSE of the two input signals are totally different for all nonlinearity levels, and the optimal parameter sets of two input signals are the same for R_S=10 only. Therefore, it proves that the proposed method is more robust and accurate than the original one.

CONCLUSION

We proposed a new training algorithm for LS-SVM based on the TLS method. Actually, a wide range of problems in estimation, identification and reconstruction can be transformed into seeking the solution to an overdetermined equation. Though the least squares method is a popular algorithm, it only takes into account the error of the observer vector, while TLS considers the errors of both the transfer matrix and the observer vector. Therefore the estimation accuracy of the TLS solution is better than that of the LS method in the case where input signals and output signals are all contaminated by noise (Zhang *et al*., 1995). The experimental results suggest that the proposed solution is suitable for the multivariable condition and may achieve better generalization performance and stability of signal reconstruction than the original solution under different nonlinearity levels. Hence, the proposed approach can be immediately used by practitioners who are interested in applying LS-SVM to various application domains.

References

Bates, D.M., Watts, D.G., 1980. Relative curvature measures of nonlinearity. *J. Royal Stat. Soc. Ser. B*, **42**:1-25.

Box, M.J., 1971. Bias in nonlinear estimation. *J. Royal Stat. Soc. Ser. B*, **33**:171-201.

Cortes, C., Vapnik, V., 1995. Support-vector networks. *Machine Learning*, **20**(3):273-297. [doi:10.1023/A:1022627411411]

Flammini, A., Marioli, D., Taroni, A., 1999. Application of an optimal look-up table to sensor data processing. *IEEE Trans. Instrum. Meas.*, **48**(4):813-816. [doi:10.1109/19.779179]

Liu, D., Sun, J.W., Wei, G., Liu, X., 2007. Application of moving least squares to multi-sensors data reconstruction. *Acta Autom. Sin.*, **33**(8):823-828 (in Chinese).

Moreira, M.F.P., Ferreira, M.D., Freire, J.T., 2006. Evaluation of pseudo-homogeneous models for heat transfer in packed beds with gas flow and gas-liquid cocurrent downflow and upflow. *Chem. Eng. Sci.*, **61**(6):2056-2068. [doi:10.1016/j.ces.2005.11.003]

Rahman, M.D., Yu, K.B., 1987. Total least squares approach for frequency estimation using linear prediction. *IEEE Trans. Acoust., Speech, Signal Processing*, **35**(10):1440-1454. [doi:10.1109/TASSP.1987.1165059]

Ribeiro, J.A., Oliveira, D.T., Passos, M.L., Barrozo, M.A.S., 2005. The use of nonlinearity measures to discriminate the equilibrium moisture equations for *Bixa orellana* seeds. *J. Food Eng.*, **66**(1):63-68. [doi:10.1016/j.jfoodeng.2004.02.040]

Smola, A.J., Scholkopf, B., 2004. A tutorial on support vector regression. *Stat. Comput.*, **14**(3):199-222. [doi:10.1023/B:STCO.0000035301.49549.88]

Sun, J.W., Shida, K., 2002. Multilayer sensing and aggregation approach to environmental perception with one multifunctional sensor. *IEEE Sensors J.*, **2**(2):62-72. [doi:10.1109/JSEN.2002.1000243]

Sun, J.W., Liu, X., Sun, S.H., 2004. TLS algorithm-based study on multi-functional sensor data reconstruction. *Acta Electron. Sin.*, **32**(3):391-394 (in Chinese).

Suykens, J.A.K., Vandewalle, J., 1999. Least squares support vector machine classifiers. *Neural Processing Lett.*, **9**(3):293-300. [doi:10.1023/A:1018628609742]

Suykens, J.A.K., Brabanter, J., Lukas, L., Vandewalle, J., 2002. Weighted least squares support vector machines: robustness and sparse approximation. *Neurocomputing*, **48**(1-4):85-105. [doi:10.1016/S0925-2312(01)00644-0]

Vapnik, V., 1998. The Nature of Statistical Learning Theory. Springer-Verlag, New York.

Vapnik, V., 1999. An overview of statistical learning theory. *IEEE Trans. Neural Networks*, **10**(5):988-999. [doi:10.1109/72.788640]

Yuji, J., Shida, K., 2000. A new multifunctional tactile sensing technique by selective data processing. *IEEE Trans. Instrum. Meas.*, **49**(5):1091-1094. [doi:10.1109/19.872935]

Zhang, H.Y., Huang, J.D., Fan, W.L., 1995. Total least square method and its application to parameter estimation. *Acta Autom. Sin.*, **21**(1):40-47 (in Chinese).