Grey-box modeling of friction

Grey-box modeling of friction: An experimental case-study

R.H.A. Hensen, G.Z. Angelis, M.J.G. van de Molengraft, A.G. de Jager, J.J. Kok

Eindhoven University of Technology, Faculty of Mechanical Engineering
P.O.Box 513, 5600 MB Eindhoven, The Netherlands
Fax : +31 (0)40-2461418 and e-mail : ron@wfw.wtb.tue.nl

Keywords : Friction models, neural networks, polytopic model, identification, extended kalman filtering.

Abstract

Grey-box modeling covers the domain where we want to use a balanced amount of first principles and empiricism. The two generic grey-box models presented, i.e., a Neural Network model and a Polytopic model are capable of identifying friction characteristics that are left unexplained by first principles modeling.

In an experimental case study, both grey-box models are applied to identify a rotating arm subjected to friction. An augmented state extended Kalman filter is used iteratively and off-line for the estimation of unknown parameters. For the studied example and defined black-box topologies, little difference is observed between the two models.

1 Introduction

Friction is to some extent present in all mechanical systems. When this phenomenon is partially neglected, and left unexplained by first principles modeling, it can limit the performance of industrial model-based control systems due to increasing tracking errors and limit cycles. Nevertheless, if detailed prior knowledge about the system is available, and first principles modeling is applicable, it might result in complex friction descriptions not very suitable for the purpose of control. It is often a time consuming job to construct these white-box models. On the other hand, black-box models are easier to construct, but purely rely on the data. If data is sparse in some regions of the operating space one may not expect to identify a reliable model. Furthermore, a black-box model does not extrapolate well and the identified parameters in the chosen model structure do not have a physical meaning. Moreover, engineering knowledge is incompatible with most empirical model representations and is therefore difficult to exploit.

Since both white-box and black-box modeling approaches have their merits as well as their drawbacks, there has in recent years been an increasing interest in combining the best of these two approaches. This approach to modeling has been termed grey-box modeling.

Since we are interested in identifying a rotational mechanical system (that exhibits several distinct friction phenomena) both for control purposes and qualitative friction analysis we want to use a balanced amount of first principles and empiricism.

Two promising grey-box model structures will be compared on this benchmark system, i.e., a Neural Network model and a Polytopic model. It appears that *a priori* unknown friction characteristics can be modeled such as proposed for instance by [11]. In comparison to the presented theoretical friction models in [2] black-box models approximate any nonlinear function without restricting to known system properties such as equilibrium points and odd friction functions. Here, the grey-box model structure is chosen in such a manner that these *prior* known system characteristics, e.g., odd friction function and equilibrium points of the autonomous system, are met.

In literature, the inertia and friction characteristic are often identified seperately, e.g., in the work of Johnson [8] and Held [4]. Here, the well known augmented state Extended Kalman Filter (EKF) [3] is applied for the simultaneous estimation of parameters in both the black- and white-box part of the model. The identification is performed with position sensing and velocity reconstruction, where compared to the work of, for instance, Armstrong-Hélouvry [1] this is done with acceleration sensing. The objective is to identify simulation models, which might give rise to the question whether this is the right objective for modeling for control. Here, the primary goal is to obtain good simulation models.

In Section 2, we will give a description of the rotating arm. The two grey-box models will be discussed in Section 3. An EKF is prosposed to estimate the unknown parameters of the models in Section 4. In Section 5 experimental results are reported to illustrate the grey-box modeling techniques. The paper will be concluded in Section 6.

2 Rotating Arm Characteristics

The system under consideration belongs to the class of nonlinear mechanical systems [10]. The state space equations describing the rotating arm system as shown in Fig. 1 are

$$\frac{d}{dt}\begin{bmatrix} q \\ \dot{q} \end{bmatrix} = \begin{bmatrix} \dot{q} \\ -M^{-1}(\theta)C(\dot{q},\theta) \end{bmatrix} + \begin{bmatrix} 0 \\ M^{-1}(\theta)c_m \end{bmatrix} u \tag{1}$$

where

q	Angular displacement
$\dot{q}$	Angular velocity
$M(q,\theta)$	Effective inertia of the motor-transmission-rotating arm combination
$C(\dot{q},\theta)$	Friction
c_m	Motor gain
u	Motor input current
θ	Parameters

We assume that the friction torque $C(\dot{q},\theta)$ is a nonlinear function of the angular velocity $\dot{q}$ and of the model parameters θ. Here, it is assumed that the friction can be modeled by an odd continuous function

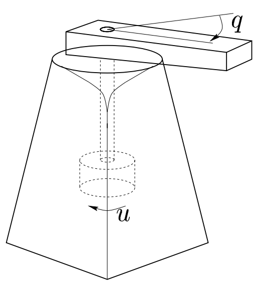

Figure 1: Rotating arm.

$$\hat{C}(\dot{q},\theta) = -\hat{C}(-\dot{q},\theta) \qquad \dot{q} \in \mathbb{R}$$

For angular velocity equal to zero the model friction torque is zero, which results for the model in the same set of equilibrium points as for the system. This friction model does not describe the *not sliding* or *pre-sliding displacement* regime, which means in this case that the friction torque is always equal to zero for zero angular velocity. The choice for this simplified friction model has two reasons

- The *not sliding* regime will be approximated, if the slope of the friction function near $\dot{q} = 0$ is very steep. Then the model may still give acceptable simulation results, i.e., angular displacement near $\dot{q} = 0$ is much smaller than for high angular velocities.
- A continuous friction function will facilitate the numerical solution of the 2^{nd} order differential equation (1).

3 Grey-box Modeling

Grey-box modeling covers the region, where we want to use a balanced amount of first principles and empiricism. In this case, the mechanical model structure is known but the friction component is left unexplained and also the inertia ($M(\theta)$) has to be estimated ($\hat{M}$). Despite of their fixed topology, the grey-box models have to be compatible with prior knowledge and observed data. Two different grey-box models will be demonstrated to model the rotating arm, i.e., (i) the Neural Network model (NN) [9] and (ii) the Polytopic model.

The NN model uses a decomposition of the system in principal functional components. These functional components can be white-box parts or unknown black-box parts. The white-box part of the model consists of the known functional components defined by (1). The NN modeling approach utilizes a neural network to approximate the friction function $C(\dot{q},\theta)$ globally. Since a neural network is a universal approximator [5], it increases the accuracy of the grey-box model.

In the case of the Polytopic modeling, the operating space is decomposed into operating regions. For every operating region a model is defined together with a model validity function. The locally valid models are combined in the operating space to obtain one globally valid nonlinear model. The model structure satisfies the universal approximation property [7], [13]. Since the system is defined by a convex combination of affine models one can associate with this model a polytope in the model space. Therefore, this model type will be called a Polytopic model. The Polytopic model generalizes various model types, e.g., Fuzzy Models [12] and Local Model Networks [6], which all have an equivalent mathematical structure.

3.1 Neural Network

The neural network consists of two layers, i.e., one hidden layer and one output layer. The neural network represents a nonlinear mapping from the network input $\mathbb{R}^r$ into the network output $\mathbb{R}^s$. Here, this mapping is from angular velocity $\dot{q}$ ($\mathbb{R}$) to friction model torque $\hat{C}(\dot{q},\theta)$ ($\mathbb{R}$). Defining the weight matrices for the first and second layers as W_1 and W_2, one can write the neural network output as

$$\hat{C}(\dot{q},\theta) = W_2^T \Sigma(W_1\dot{q} + b_1) + b_2$$

where b_i represents the bias value for the neurons in the i-th layer and $\Sigma(.)$ is a nonlinear operator with $\Sigma(\underline{z}) = [\sigma(z_1), \ldots, \sigma(z_v)]^T$, $\sigma(.)$ a differentiable, nonlinear, monotonic increasing function and v is the number of hidden neurons.

To assure the system properties described in Section 2 to hold, the following restrictions are posed on the neural network topology

- Choose an odd function for $\sigma(.)$ which is equal to zero if its argument is zero.

$$\sigma(z_i) = 1 - \frac{2}{e^{2z_i} + 1}$$

- The first choice together with the set of equilibrium points for the system implies that the bias terms should be zero. Here, the assumption is made that there is no bias in the reconstructed angular velocity $\dot{q}$.

These two restrictions result for the neural network friction approximator in

$$\hat{C}(\dot{q}, \theta) = W_2^T \Sigma(W_1 \dot{q})$$

The linear part of the system dynamics, i.e., the viscous damper characteristic and the input term are described by two other principal functional components. In state space description the grey-box model becomes

$$\begin{aligned} \begin{bmatrix} \dot{q} \\ \ddot{q} \end{bmatrix} &= \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & -\frac{b}{\hat{M}} \end{bmatrix} \begin{bmatrix} q \\ \dot{q} \end{bmatrix}}_{\text{white-box component}} - \underbrace{\begin{bmatrix} 0 \\ \frac{W_2^T \Sigma(W_1 \dot{q})}{\hat{M}} \end{bmatrix}}_{\text{black-box component}} \\ &+ \underbrace{\begin{bmatrix} 0 \\ \frac{c_m}{\hat{M}} \end{bmatrix}}_{\text{white-box component}} u \end{aligned} \qquad (2)$$

where b is the viscous damper constant of the system. For the model parameters this results in

$$\theta = [W_1^T \;\; W_2 \;\; b \;\; \hat{M}]^T$$

3.2 Polytopic Model

The polytopic model is composed of several locally valid models. The structure of each model is chosen equal to the topology of an mechanical system, i.e., $\dot{x} = A_i x + C_i + B_i u$ with $x = [q \; \dot{q}]^T$. With each local model a model validity function $\rho_i : \mathbb{R}^p \to [0,1]$ is associated which, by definition, is close to 1 for those regions in the input and state space where the corresponding local linear model is valid. Here, the partitioning only depends on the angular velocity $\dot{q}$ due to the choice of the nonlinear friction torque as a function of $\dot{q}$. A typical choice for the validity function ρ_i is the Gaussian function

$$\rho_i(\dot{q}, \theta) = e^{-\frac{1}{2} \frac{(\dot{q} - c_i)^2}{\sigma_i}}$$

where c_i is the center and σ_i is the variance of the Gaussian function. Now a set of normalized validation functions $w_i : \mathbb{R} \to [0,1]$ can be defined

$$w_i(\dot{q}, \theta) = \frac{\rho_i(\dot{q}, \theta)}{\sum_{j=1}^N \rho_j(\dot{q}, \theta)}$$

where N is the number of local models used to compose one global model. This definition implies that $\sum_{i=1}^N w_i(\dot{q}, \theta) = 1 \; \forall \dot{q}$. The polytopic friction model becomes

$$\hat{C}(\dot{q}, \theta) = \sum_{i=1}^N w_i(\dot{q}, \theta)(a_i(\dot{q} - c_i) + b_i)$$

where $a_i(\dot{q} - c_i) + b_i$ is the affine model of the friction locally valid around c_i. For the identification of the Polytopic model, the centers c_i, slopes a_i and offsets b_i of the linear models and the variance σ_i of the Gaussian validity functions have to be estimated.

One way to construct an odd function with the polytopic model is to

- Choose an odd number of local models, where one polytopic has no offset $b_1 = 0$ and the corresponding center $c_1 = 0$. Again, the assumption of an unbiased reconstruction of the angular velocity $\dot{q}$ is adopted here.

- The other $N - 1$ models are divided in pairs of two, where the centers are opposite $c_{2i} = -c_{2i+1}$, the variances are equal $\sigma_{2i} = \sigma_{2i+1}$ as well as the slopes $a_{2i} = a_{2i+1}$ and the offsets are again opposite $b_{2i} = -b_{2i+1}$ with $i = 1, \ldots, \frac{N-1}{2}$. An advantage of this construction is the reduction of parameters by a factor 2.

These conditions assure an odd function which gives zero if $\dot{q}$ is zero in order to guarantee the equilibrium property. The state space representation of the polytopic model becomes

$$\begin{aligned} \begin{bmatrix} \dot{q} \\ \ddot{q} \end{bmatrix} &= \sum_{i=1}^N w_i(\dot{q}, \theta) \Bigg\{ \begin{bmatrix} 0 & 1 \\ 0 & -\frac{a_i}{\hat{M}} \end{bmatrix} \begin{bmatrix} q \\ \dot{q} \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ \frac{a_i c_i - b_i}{\hat{M}} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{c_m}{\hat{M}} \end{bmatrix} u \Bigg\} \end{aligned}$$

where the first term on the right-hand side are the normalized validity functions and the second term between the brackets the locally valid linear mechanical models. For the polytopic grey-box model, the model parameters become

$$\theta = [a_1 \; a_{2i} \; b_{2i} \; c_{2i} \; \sigma_1 \; \sigma_{2i} \; \hat{M}]^T \quad i = 1, \ldots, \frac{N-1}{2}$$

So, the modeling problem is reduced to dividing the operating space of the system into a set of operating regimes and identifiying with every operating regime a locally valid mechanical model together with a corresponding validation function.

4 Estimation of the model parameters

The rotating arm will be identified with the objective to obtain simulation models that yield accurate long-term prediction. The model parameters are estimated with an algorithm that minimizes an output error criterion. Due to the smoothness of the proposed nonlinear grey-box models the Extended Kalman Filter (EKF) seems a suitable technique for estimating the model parameters [3]. The filter is able to reconstruct the state of the continuous-time system with discrete-time measurements of the system outputs. This technique, which is based on the assumption that all errors are stochastic, minimizes the variance of the reconstruction error, i.e., the difference between the actual state x_r and the estimated state $\hat{x}$.

The nonlinear parameter estimation procedure for continuous-time mechanical models (1) with discrete measurements will be outlined shortly. First the state $x = [q\ \dot{q}]^T$ is augmented with the unknown parameters, $\theta \in \mathbb{R}^k$ such that the new state $x^* = [x\ \theta]^T$. The parameter estimation problem is converted into a state (x^*) reconstruction problem. Consequently, (1) has to be augmented with k trivial differential equations $\dot{\theta} = 0$, which marks the parameters as constants.

Model errors are w for the state equations and v for the measurement equations. Here, the model errors are considered to be zero mean gaussian noise having intensity matrices $Q(t)$ for the state errors w and $R(t)$ for the measurement errors v. Furthermore, the state errors and measurement errors are assumed to be uncorrelated. The uncertainty in the initial state estimate $\hat{x}^*(0)$ can be expressed by the diagonally choosen initial covariance matrix $P(t_0)$.

To avoid numerical real-time problems due to limited computational time and inaccurate integration schemes, the EKF was implemented off-line. First, an experiment must be performed to obtain measured experimental data. These data should excite all system dynamics we are interested in. Second, the data are passed through the filter several times untill the parameter estimates converge. After each filter pass, the initial system states x and the corresponding covariance matrices are re-initialized with the initial estimates of the first pass. The parameters θ and the corresponding covariance matrix are reset to the final estimates of the previous filter pass. When the parameters have become constant, after applying the iterative EKF, the estimates can be considered as smoothed estimates [3] which implies that effectively an output error criterion is minimized.

5 Experimental Study

In this section an experimental study will be reported, where the parameters of the two proposed models are estimated with experimental data obtained from the rotating arm.

The rotating arm is excited by a motor torque $c_m u$ which together with the measured angle and arm velocity responses of the system are depicted in Fig. 2. Here, we are mainly interested in the friction phenomenon and non-zero angular velocities. The measured angle is differentiated numerically by a high pass filter with a cut-off frequency of 200 [Hz] to reconstruct the angular velocity.

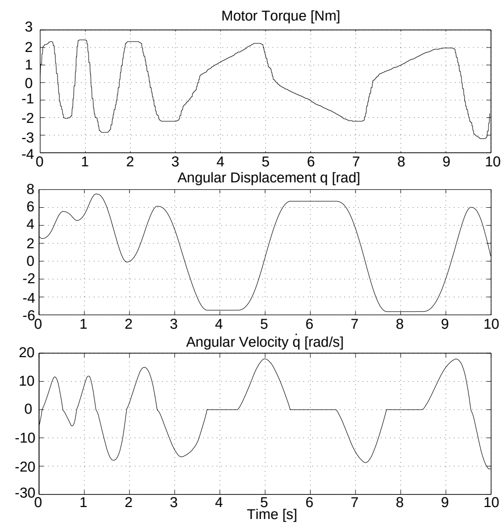

Figure 2: Experimental data.

For the parameter estimation, the two model structures have to be specified, i.e., the number of neurons for the NN and the number of linear models for the polytopic model have to be chosen. Both the number of neurons and the number of linear models are set to 3. Hence, the total number of parameters to be estimated for the NN model becomes 8; 6 parameters for the neural network, one for the viscous damping and one for the inertia. For the polytopic model 7 parameters have to be determined; 6 for the polytopic friction model and one for the inertia. The motor constant c_m is in both cases assumed to be exactly known and set to 16.

The initial state of the system is known, but the model parameters are not known. The initial model parameters are chosen in such a manner that physical known properties, e.g., positive inertia value or positive viscous damper value, are met. Hence, the error variance for the initial state estimates is small while we are not sure of the initial estimates for the model parameters. These considerations lead to the initial variance matrix $P(t_0) = \mathrm{diag}(0, 0, 1, \ldots, 1)$ where the non-zero elements correspond to the variance of initially uncertain parameters. The matrix $Q(t)$ can be seen as the variance of the augmented state model errors. Here the assumption is that the model errors are not cross-correlated. Furthermore the model equations describing constant model parameters and $\frac{d}{dt}q = \dot{q}$ from (1) are regarded as true. Combined this gives a diagonal matrix with $Q(t_0) = Q(t) = \mathrm{diag}(0, Q_{22}, 0, \ldots, 0)$ where $Q_{22} = 0.001$. Due to the finite encoder resolution of 2.10^{-4} [rad] and the differentiation scheme an uncertainty on the angular velocity reconstruction is introduced. To take this into account the variance matrix $R(t)$ is constructed by a diagonal matrix $R(t) = \mathrm{diag}(0.001, 0.01)$ where 0.001 corresponds to the uncertainty in the angle measurement and 0.01 to the uncertainty in the arm velocity reconstruction. The filter tuning is mainly based on experience and trial and error. It is important that the parameters converge to constant values. Different filter tunings will result in different convergence speeds and even parameter divergence can occur.

After 10 filter passes the parameter estimates become constant and the sum of eigen-values of the covariance matrix $P(t)$ is minimal. The identified inertia value is the same for both models, i.e., 0.0292 [kg m^2/rad]. In Fig. 3 both the estimated neural network and polytopic friction model, as a function of the angular velocity $\dot{q}$, are shown. The following different friction phenomena can be distinguished from the estimated friction models: (i) coulomb friction, (ii) static friction, (iii) stribeck effect for low velocities and (iv) viscous friction for high velocities. The obtained simulation models are validated by another

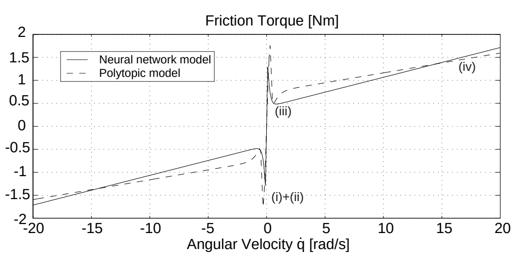

Figure 3: Identified friction models.

experiment. The results of the validation experiment and simulated model responses are shown in Fig. 4. Here the solid lines are the experimental validation data, the dashed lines the Neural Network response and the dash-dotted lines the Polytopic model response. For high velocities, i.e., $|\dot{q}| \geq 20$ [rad/s], which were not present in the training data, the displacement errors become large. Hence, the models exhibit poor extrapolation and good generalization behaviour. The assumption of the continuous friction function, made in Section 2, is justified by the validation responses. For velocities near zero the displacement is indeed much smaller than for high velocities.

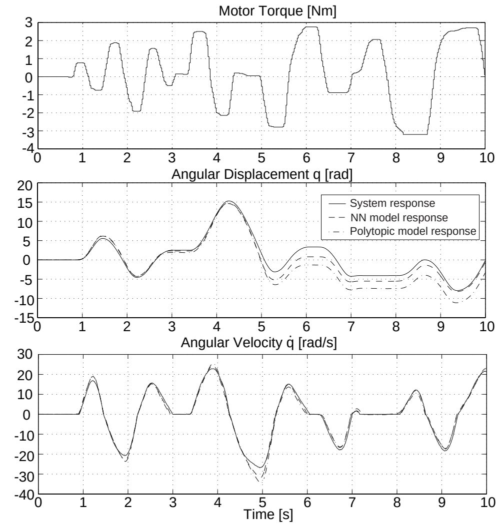

Figure 4: Validation of identified models.

An unexpected change of the friction characteristic of the system was recorded due to maintenance effort. This change was investigated by applying the same input torque as for the validation, as shown in Fig. 5. The dashed lines represent the new system responses,

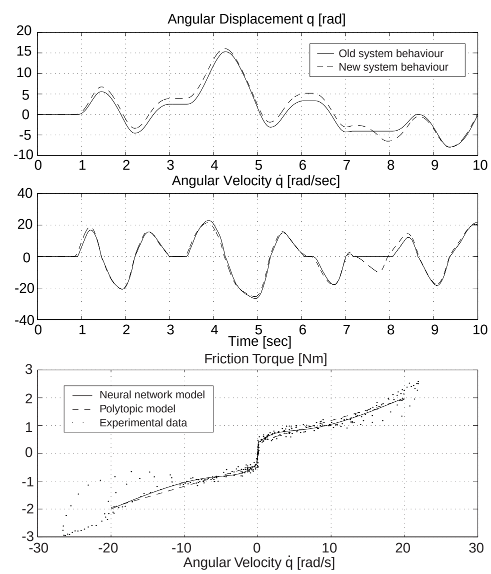

Figure 5: Changed system characteristics and new identified friction models

which indicate a change of the friction characteristic. A new identification experiment was performed to get appropriate measurements of the new system behaviour. The inertia was estimated within 2% difference of the earlier identified inertia. New friction models were identified with the earlier identified friction models as initial estimates and the same identification procedure as above. The static friction is not present any more resulting in a friction characteristic describing: (i) coulomb friction and (ii) viscous friction, as shown in the lower plot of Fig. 5. The black-box models are able to identify different friction characteristics, as shown in Fig. 3 and Fig. 5, without changing the black-box topology, i.e., number of neurons for the NN and number of local models for the Polytopic model. To give insight in the accuracy of the identified friction models the reconstructed friction torque from the identification experiment is also plotted in Fig. 5. The assumption that the friction is an odd function seems to be justified. The neural network friction model gives a slightly better fit of the experimental data than the polytopic friction model. A change in the topology of the polytopic model, e.g., five linear models, might allow us to identify a better friction model due to more freedom in the model. Furthermore, hysteresis curves for high angular velocities are recorded which would prefer dynamic friction models instead of static friction models.

The change in the system characteristics and the poor extrapolation behaviour of the models would prefer an on-line implementation of the EKF to adapt to changes in friction, which will be an important topic in future work.

6 Conclusions

The identification of grey-box models presented in this paper yields good results for an experimental study on a rotating arm which exhibits friction. Although the two proposed grey-box modeling approaches, i.e., Neural Network modeling and Polytopic modeling, are different from a theoretical point of view, both approaches are able to identify a continuous friction function that by construction models *a priori* known system characteristics, e.g., equilibrium points. The friction models can physically be interpreted where friction phenomena such as static friction and the stribeck effect are observed. The black-box elements can represent different friction characteristics without changing the black-box topology. Hence, the proposed grey-box models are favourable for the identification of systems with unkown friction characteristics. This ability to represent the friction characteristic accurately results in simulation models that have good long-term prediction with good interpolation and poor extrapolation properties. In future work the choice of the black-box model topology will be investigated and an extension to dynamic friction models will be made.

The iterative EKF approach appears to be a useful identification tool for the proposed nonlinear continuous-time modeling techniques where discrete-time measurements are available. The filter tuning is an important aspect of the model identification due the possible divergence of the parameter estimates. The EKF is able to identify the friction models based on angular displacement measurements and velocity reconstruction. Due to the time-varying friction characteristics an on-line implementation of the EKF is of future interest.

References

[1] Armstrong-Hélouvry B., "Control of machines with friction", *Kluwer Academic Publishers.*, (1991).

[2] Armstrong-Hélouvry B., Dupont P., and Canudas de Wit C., "A survey of models, analysis tools and compensation methods for the control of machines with friction", *Automatica*, **30**, 1083-1138, (1994).

[3] Gelb A., "Applied optimal estimation", *M.I.T. Press*, (1978).

[4] Held D., Maron C., "Estimation of friction characteristics, inertial and coupling coefficients in robotic joints based on current and speed measurements", *Robot Control 1988, Selected Papers from the 2nd IFAC Symp.*, 207-212, (1988).

[5] Hornik K., Stinchombe M., White H., "Multilayer feedforward networks are universal approximators", *Neural Networks*, **2**, 359-366, (1989).

[6] Johansen T.A., Foss B.A., "Nonlinear local model representation for adaptive systems", *Int. Conf. on Intelligent Control and Instrumentation*, **2**, 677-682, (1992).

[7] Johansen T.A., Foss B.A., "State-space modeling using operating regime decomposition and local models", *Preprints 12th IFAC World Congress*, **1**, 431-434, (1993).

[8] Johnson C.T., Lorenz R.D., "Experimental identification of friction and its compensation in precise, position controlled mechanisms", *Proc. of the Indus. App. Soc. Annual Meeting*, 1400-1406, (1991).

[9] Narendra K.S., Parthasarathy K., "Identification and Control of Dynamical Systems Using Neural Networks", *IEEE Transactions on Neural Networks*, **1**, 4-27, (1990).

[10] Nijmeijer H., Schaft van der A.J., "Nonlinear Dynamical Control Systems", *Springer*, (1990).

[11] Selmic R., Lewis F.L., "Neural network approximation of piecewise continuous functions: application to friction compensation", *Proc. Int. Symp. on Intelligent Control*, 227-232, (1997).

[12] Sugeno M., Kang G.T., "Structure identification of fuzzy model", *Fuzzy Sets and Systems*, **26**, 15-33, (1988).

[13] Wang L.X., Mendel J.M., "Fuzzy Basis Functions, Universal Approximators and Orthogonal Least-Squares Learning", *IEEE Transactions on Neural networks*, **3**, 807-814, (1992).