Fault detection based on novel fuzzy modelling

Zhen Luo [1,2], Huajing Fang [2], Xiaoyong Liu [2] and Xuquan Chen [1,2]
[1]Guangxi Special Equipment Supervision and Inspection Institute, Guangxi 530219, P. R. China

[2]Department of Control Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, P. R. China

ABSTRACT

The Fault detection which is based on fuzzy modeling is investigated. Takagi-Sugeno (TS) fuzzy model can be derived by structure and parameter identification, where only the input-output data of the identified system are available. In the structure identification step, Gustafson-Kessel clustering algorithm (GKCA) is used to detect clusters of different geometrical shapes in the data set and to obtain the point-wise membership function of the premise. In the parameter identification step, Unscented Kalman filter (UKF) is used to estimate the parameters of the premise's membership function. In the consequence part, Kalman filter (KF) algorithm is applied as a linear regression to estimate parameters of the TS model using the input-output data set. Then, the obtained fuzzy model is used to detect the fault. Simulations are provided to demonstrate the effectiveness of the theoretical results.

KEYWORDS

Unscented Kalman filter, Kalman filter, data driven, Gustafson-Kessel clustering algorithm, fuzzy modeling, fault detection

1. INTRODUCTION

Since fault detection (FD) &diagnosis technique is essential to improve the safety and reliability of dynamic systems, recently more and more attention has been paid to FD. Many FD methods have been developed to detect and identify sensor and actuator faults, such as analytical redundancy [1-3], a neural network [4], parameter identification method based on Fourier Transform [5], testing the covariance matrix of the innovation sequence [5], testing the eigen values of the sample covariance matrix [6]. All of these methods mentioned above are based upon model. Models with good accuracy are necessary to improve the correct diagnostic of faults. However, with the rapid development of industrial technology, the modern industrial processes have become more and more complicated and large-scale. It is thus difficult to construct an effective and explicit physical model to characterize these dynamical systems. Sometimes, it is even impossible to model nonlinear systems by analytical equations [7]. Here, the fuzzy modeling has been extensively used to model complex nonlinear system through a set of measured input-output data [8]. The Takagi-Sugeno (TS) model, which uses the fuzzy modeling technique, approximate the nonlinear system by smoothly interpolating affine local models. Each local model contributes to the global model in a fuzzy subset of the space characterized by a membership function [9]. Then, based on the established system model, FD can be carried out.
A TS fuzzy model is usually constructed in three steps:

Step 1: fuzzy clustering

Among various clustering algorithm, the Gustafson-Kessel clustering algorithm (GKCA) [10] has been widely studied and applied by many researchers. The Gustafson-Kessel (GK) algorithm, which is the fuzzy generalization of the Adaptive Distance Dynamic Clusters algorithm, searches for ellipsoidal clusters. An advantage of the GK algorithm over Fuzzy C-means (FCM) [11] is that GK can detect clusters of different shape and orientation in one data set.

Multidimensional fuzzy sets can be derived by clustering in the product space, but they are generally difficult to be dealt with. In order to solve this problem, projected one-dimensional fuzzy sets are usually preferred.

Step 2: estimate the parameters of the premise's membership function

In order to obtain a fuzzy model, the premise's membership functions must be expressed in a form that allows computation of the membership degrees for input data not contained in the data set. Each update estimation of the parameter vector corresponding to a nonlinear equation is computed from the previous estimate and the new input data (here the input data are the point-wise values of the membership function). To achieve this step, Unscented Kalman filter (UKF) is used to approximate the point-wise defined membership functions by some suitable nonlinear functions, for example normal function. The UKF as an improvement to the extended Kalman filter (EKF) is one of the most widely used approach to analyze the stochastic nonlinear systems [12]. This method is based on the unscented transform (UT) technique, a mechanism for propagating mean and covariance through a nonlinear transformation [13, 14]. The state vector is represented by a minimal set of carefully chosen sample points, called sigma points, which approximate the posterior mean and covariance of the Gaussian random variable with a second order accuracy[15, 16]. In contrast, the linearization technique used in the EKF can only achieve first order accuracy. Generally, the prediction accuracy of UKF is better than EKF when the model is highly nonlinear. The UKF has the same level of computational complexity as that of EKF, both of which are within the order $O(L3)$. Furthermore, Since the nonlinear models are used without linearization, the UKF is not necessary to compute the Jacobian or Hessians matrices [17, 18].

Step 3: estimation of the TS parameters

For linear dynamic systems with white process and measurement noise, the Kalman filter is an optimal estimator [19]. The KF is used as a linear regression to efficiently choose the parameter values of the consequent part (TS parameters) of the fuzzy model from the input-output data of the identified system.

[20] proposes a fuzzy-modeling scheme combining GKCA and KF. GKCA is used to detect clusters of different geometrical shapes in the data set and to obtain the point-wise membership functions of the premise. After that, a KF is first used to estimate the parameters of the premise's membership function. Then, the KF is also used as a linear regression to efficiently choose the parameter values of the consequent part (TS parameters) of the fuzzy model from the input-output data of the identified system. KF process is a linear recursive minimum mean-square estimation procedure. However, it can only deal with linear equation. In this sense, in order to obtain the linear part, α-cut of the considered membership function is taken. Obviously, there are some drawbacks with this algorithm

In order to obtain membership function for the premise's fuzzy sets, the multidimensional fuzzy set defined point-wise in the row of the partition matrix is projected onto the regressor. Division of each membership function into two sets (sets are obtained for each regressor and sets are obtained for all regressors). It needs large amount of computation

The selection of α relies entirely on experience. Too much or too small will influence on membership function approximation. Moreover, the cut data causes the loss of useful information.

The straight line which is used to approximate each set is not very good fit of the data.

In this paper, we point-wise define the premise membership function for the regressor, each membership function as a sets (c sets are obtained for each regressor and cn sets are obtained for all regressors). This will halve the amount of computation. Obviously, it is more accurate to approximate the point-wise defined membership function by some suitable nonlinear function. UKF is utilized to approximate the nonlinear function. A common FD approach is to keep tracking residuals of measurement and compare them against a set threshold value. The residual can be generated by the comparison of actual sample points and estimated sample points .

The structure of this paper is as follow. In section 2, the TS fuzzy models principle is summarized. Section 3 gives new fuzzy modeling algorithm. Fault detection is presented in section 4. Section 5 presents some simulation results. Conclusions are given in section 6.

2. TS FUZZY MODELS

In this section, the TS fuzzy models principle is summarized. The TS fuzzy model can represent or model any unknown nonlinear system $y = f(x)$, on the basis of some available input-output data $x_t = [x_{1t},\ x_{2t} ..., x_{nt}]^T$ and y_t .

In the TS fuzzy model, the rule consequents are crisp function of the model inputs

$$R_i: \text{ if } x \text{ is } A_i(x)\text{, then } y_i = a_i^T x + b_i \quad i = 1, 2, ..., c \tag{1}$$

where is the dimensional input invariable， is the output invariable. and are the dimensional TS parameters. denotes the th rule and is the number of rules in the rule base. is the premise multivariable membership function of the th rule. is the parameter vector of the th rule
Multi dimensional fuzzy sets can be derived by clustering in the product space. But it is generally difficult to be interpreted, so projected one-dimensional fuzzy sets are usually preferred.

The TS fuzzy system is described as a set of c fuzzy rules where the i th rule is as follows:

$$R_i: \text{ if } x_1 \text{ is } A_{i1}(x_1) \text{ and } \ldots \text{and } x_n \text{ is } A_{in}(x_n)\text{ , then } y_i = a_i^T x + b_i \quad i = 1, 2, ..., c \tag{2}$$

The final output of the TS fuzzy model for an arbitrary input sample can be calculated using the following expression:

$$\widehat{y} = \frac{\sum_{i=1}^{c} \beta_i(x)(a_i^T x + b_i)}{\sum_{i=1}^{c} \beta_i(x)} \tag{3}$$

where $\beta_i(x)$ represent the firing strength of the i th rule, and it has calculated by the following equation:

$$\beta_i(x) = \prod_{j=1}^{n} \mu_{A_{ij}} \tag{4}$$

$\mu_{A_{ij}}$ is the membership function of the fuzzy set A_{ij}.

3. NEW FUZZY MODELING ALGORITHM

The loop of establishing and training T-S fuzzy model by Gustafson-Kessel clustering algorithm (GKCA) and UKF is as follows:

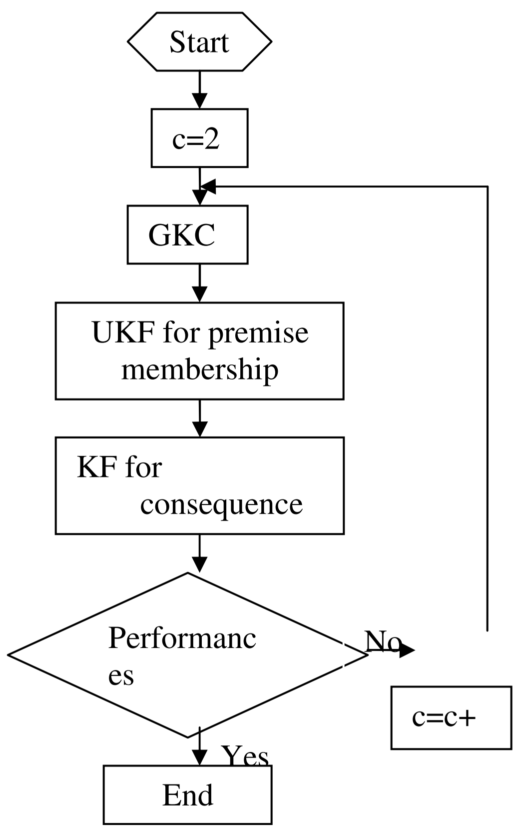

Fig.1 Flow chat of the new fuzzy modelling

A detailed description of the algorithm is provided next.

1. From the input–output sequences $\{(x_k, y_k)\}_{t=1}^{N}$, partition the data into a set of local linear submodels by using GKCA in the product space $X \times Y$.

2. Obtain the membership functions for the premise variables by using cluster projections and UKF.

Estimate the consequent parameters by KF algorithm.

3.1 FUZZY CLUSTERING

Fuzzy clustering is an important tool to identify the structure in data. The Gustafson-Kessel(GK) algorithm, which is the fuzzy generalization of the Adaptive Distance Dynamic Clusters algorithm, searches for ellipsoidal clusters. An advantage of the GK algorithm over Fuzzy C-means (FCM) is that GK can detect clusters of different shape and orientation in one data set.

The original objective functional for the GK algorithm is the following:

$$J = \sum_{i=1}^{c}\sum_{t=1}^{N} \mu_{it}^{m} (z_t - v_i)^T M_i (z_t - v_i) \tag{5}$$

The following restrictions hold

$$\sum_{t=1}^{N} \mu_{it}^{m} = 1 \quad 1 \le t \le N \tag{6}$$

where is a positive-definite symmetric matrix related to the covariance matrix of the th prototype, is a weighting exponent that determines the fuzziness of the resulting cluster(for a crisp , fuzzy model , but typically). is the number of data points, is the number clusters, is the th data point, is the th cluster centre, is the degree of the membership of the th data point in the th cluster centre.

3.2 PREMISE MEMBERSHIP FUNCTION

UKF as an improvement to extended Kalman filter (EKF) is the most widely used approach to analyze the stochastic nonlinear system and show good performance in many cases[12]. In this sense, we propose to use UKF as a nonlinear regression as follow: consider cn sets, each set represents the nonlinear part of the point-wise set of a certain premise's membership function (for example Gauss type membership function). So we obtain cn parameter vector. In each set, we will have N_j data (samples), where j denotes the j th set.

Then, each set can be modeled by the following measurement equation

$$A_t^j = h(x_t^j, \theta^j) + v_t^j \qquad j = 1,2...cn$$
$$= \exp(-\frac{(x_t^j - d^j)^2}{2\sigma^{j2}}) + v_t^j \qquad t = 1,2...N_j \tag{7}$$

where $\theta^j = [\sigma^j, d^j]$ is the parameters vector, v_t^j is the measurement noise, N_j is the number of data(samples) in the j th set and the j denotes the j th nonlinear regression.

θ^j will be considered as a state variable, so the state equation will be

$$\theta_t^j = F^j \theta_{t-1}^j + w_{t-1}^j \tag{8}$$

where is the value of the state variable at the moment , is the state noise, is the state transition matrix we assume that and are uncorrelated zero mean Gaussian random vectors and their covariance matrices are and

The procedure for implementing the UKF can be summarized as follows:

Step 1: sigma points calculation

$$\chi_{t-1}^j = [\hat{\theta}_{t-1}^j,\ \hat{\theta}_{t-1}^j \pm \sqrt{(n+\lambda)\hat{P}_{t-1}}] \tag{9}$$

Step 2: prediction

$$\chi_{t|t-1}^j = F^j \chi_{t-1}^j$$

$$\hat{\theta}_{t|t-1}^j = F^j \hat{\theta}_{t-1|t-1}^j$$

$$\hat{P}_{t|t-1}^j = F^j P_{t-1|t-1}^j F^{jT} + Q \tag{10}$$

Step 3: update.

$$\hat{A}_{t|t-1}^j = h(\hat{\theta}_{t|t-1}^j)$$

$$\gamma_{l\ t|t-1}^j = h(\chi_{l\ t|t-1}^j)$$

$$\hat{P}_{A_t^j A_t^j}^j = \sum_{l=0}^{2n} \omega_l^c (\gamma_{l\ t|t-1}^j - \hat{A}_{t|t-1}^j)(\gamma_{l\ t|t-1}^j - \hat{A}_{t|t-1}^j)^T + R$$

$$\hat{P}_{\theta_t^j A_t^j}^j = \sum_{l=0}^{2n} \omega_l^c (\chi_{t|t-1}^j - \hat{\theta}_t^j)(\gamma_{l\ t|t-1}^j - \hat{A}_{t|t-1}^j)^T + R$$

$$K_t = \hat{P}^j_{\theta_t^j A_t^j} \hat{P}^{j\ -1}_{A_t^j A_t^j}$$

$$\hat{\theta}_t^j = \hat{\theta}^j_{t|t-1} + K_t (A_t^j - \hat{A}^j_{t|t-1})$$

$$\hat{P}_t = \hat{P}_{t|t-1} - K_t \hat{P}^{j\ -1}_{A_t^j A_t^j} K_t^T \tag{11}$$

where, $\omega_0^c = \lambda / (n+\lambda) + (1-\alpha^2+\beta)$, $\omega_l^m = \omega_l^c = \lambda / 2(n+\lambda)$ $\quad l = 1...2n$, $\lambda = n(\alpha^2-1)$. ω_l^c is a set of scalar weights, and n is the state dimension; the parameter α determines the spread of the sigma points around $\hat{x}$ and is usually set to $1e-4 \le \alpha \le 1$, The constant β is used to incorporate part of the prior knowledge of the distribution of x , and for Gaussian distributions, $\beta = 2$ is optimal.

Often it is expected that the system parameters do not vary or, if they do, the variation is much slower than that of the system state. So, we will take $F^j = I$.

3.3 ESTIMATION CONSEQUENT PARAMETERS

In this section, KF algorithm is used to compute the consequent parameters from the data set and the estimated premise's membership functions.
From (3), we have

$$\hat{y} = \sum_{i=1}^{c} \psi_i(x)(a_i^T x + b_i) \tag{12}$$

where $\psi_i(x) = \dfrac{\beta_i(x)}{\sum_{i=1}^{c} \beta_i(x)}$ is the normalized activation value of the i th rule.

$$\hat{y} = \left[\psi_1(x)\left[x\ 1\right] \ \ \psi_2(x)\left[x\ 1\right] ... \psi_c(x)\left[x\ 1\right]\right]\left[a_1\ b_1 ... a_c\ b_c\right]^T$$

$$\Phi = \left[a_1\ b_1 ... a_c\ b_c\right]^T,\ \Omega = \left[\psi_1(x)\left[x\ 1\right] \ \ \psi_2(x)\left[x\ 1\right] ... \psi_c(x)\left[x\ 1\right]\right] \tag{13}$$

The measurement equation can be taken as the following form

$$\hat{y} = \Omega_t \Phi_t + v_t \tag{14}$$

Then, θ^j will be considered as a state variable, so the state equation will be

$$\Phi_t = F\Phi_{t-1} + w_{t-1} \tag{15}$$

Now, we can use KF to estimate the TS parameter vector Φ_t as follows:

$$\hat{\Phi}_{t|t-1} = F\hat{\Phi}_{t-1|t-1}$$

$$P_{t|t-1} = FP_{t-1|t-1}F^T + Q$$

$$K_t = P_{t|t-1}\Omega_t^T(\Omega_t P_{t|t-1}\Omega_t^T + R)^{-1}$$

$$\hat{\Phi}_{t|t} = \hat{\Phi}_{t|t-1} + K_t(y_t - \Omega_t\hat{\Phi}_{t|t-1})$$

$$P_{t|t} = P_{t|t-1} - K_t\Omega_t P_{t|t-1} \tag{16}$$

where $\hat{\Phi}_t$ is the estimated value of Φ_t. Also, we will take $F = I$.

4. PROCESS FAULT DETECTION

The fuzzy FD system is based on fuzzy models identified directly from data. A model is used to estimate the nominal output signals.

$$\Delta = y - \hat{y} \tag{17}$$

where y is the output of the system and $\hat{y}$ is the output of the model in normal operation. When any component of Δ is bigger than a certain threshold, the system detects a fault.

5.SIMULATION

The sampling period of training data and testing data are all 0.05h. We choose 800 data points without fault and 600 data points with fault 1 after collecting 200 data points without fault. The number of clusters $c = 12$.

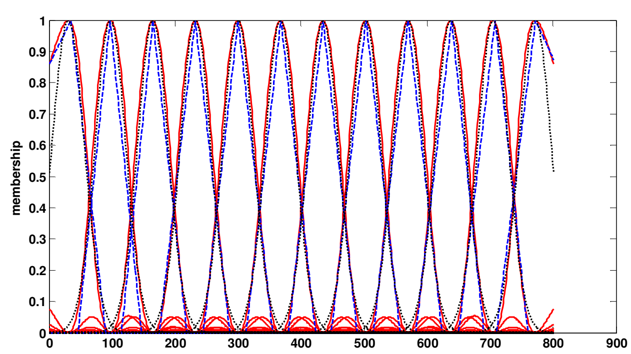

Fig.2 the red solid line indicates the actual membership function, the black dotted line indicates the premise membership function based on UKF, the blue dashed line indicates the premise membership function based on KF

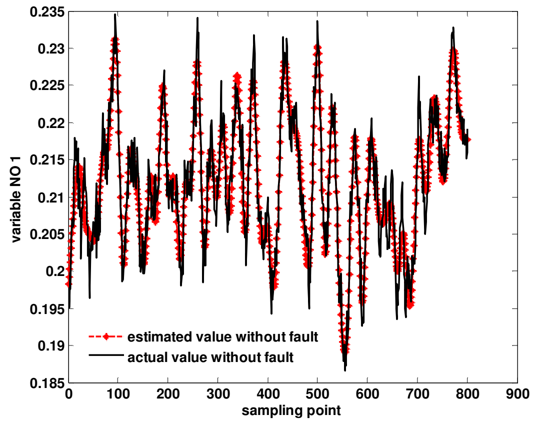

Fig.3 the actual sample point without fault and estimated sample point without fault

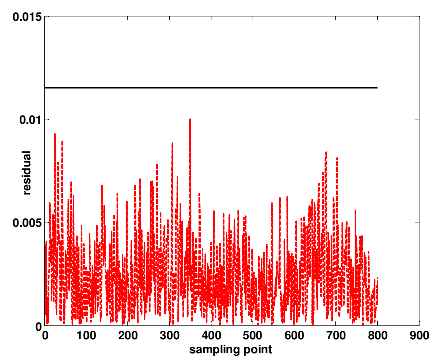

Fig.4 the residual with fault

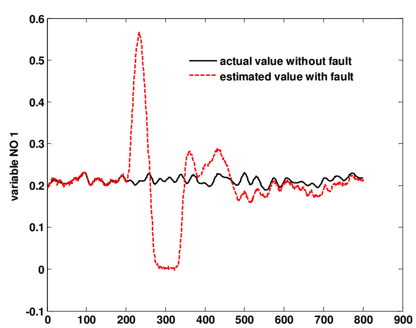

Fig.5 the actual sample point without fault and the estimated sample point with fault

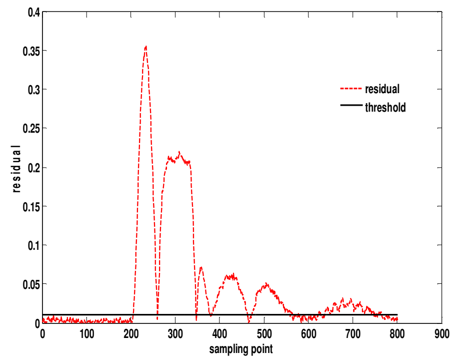

Fig.6 the residual with fault

The membership functions is given in Fig.2. The red solid line indicates the actual membership function, the black dotted line indicates the premise membership function based on UKF, the blue dashed line indicates the premise membership function based on KF. It can be seen from the figure that both approaches track the true state. However, the obtained premise membership based on UKF is more close to the true value than based on KF. It means that estimating the premise membership based on UKF is superior to based on KF.

In Fig.3, we can see that the actual sample point can be tracked well by the estimated sample point. Fig.4 shows that the absolute value of the residuals which is derived by subtracting the estimated value of the actual value is lower than the threshold. The simulation results consistent with the actual case.

When the fault 1 occurs at the step 200, the simulation results are given in Fig.5-6. The estimated sample points have deviated from actual sample points after 200 steps in Fig5 due to the fault, but the actual sample point can be tracked well by the estimated sample point after the step 500 because of feedback effect. The fault makes absolute value of the residuals exceed the threshold at step 206 which is given in Fig.6. I.e. the fault can be detected after occurrence of 6 steps.

6.CONCLUSION

In this paper, a fault detection algorithm based on data driven is proposed. The proposed algorithm is composed of two steps: (1) fuzzy modeling contain fuzzy clustering, determination of premise membership functions and TS parameters. (2) Fault detection based on the established model. In the first step, a fuzzy modeling scheme on the basis of the GKCA and UKF to estimate the premise membership, KF is used to estimate the TS parameter. The performances of the proposed algorithm are demonstrated on the TE challenge problem.

REFERENCES

[1] Zhang YM, Li XR, Ieee(1997)"Detection and diagnosis of sensor and actuator failures using interacting multiple-model estimator", In: 36th IEEE Conference on Decision and Control: Dec 10-12 1997; San Diego, Ca; pp4475-4480.

[2] Maybeck PS(1999)" Multiple model adaptive algorithms for detecting and compensating sensor and actuator/surface failures in aircraft flight control systems", International Journal of Robust and Nonlinear Control, 9(14):pp1051-1070.

[3] Rago C, Prasanth R, Mehra RK, Fortenbaugh R(1998), Ieee:"Failure Detection and dentification and Fault Tolerant control using the IMM-KF with applications to the Eagle-Eye UAV". In: 37th IEEE Conference on Decision and Control: Dec 16-18 ; Tampa, Fl; pp4208-4213.

[4] Alessandri A(2003)"Fault diagnosis for nonlinear systems using a bank of neural estimators". Computers in Industry , 52(3)pp271-289.

[5] Perhinschi MG, Lando M, Massotti L, Campa G, Napolitano MR, Fravolini ML(2002), Acc"On-line parameter estimation issues for the NASA IFCSF-15 fault tolerant systems". In: 20th Annual American Control Conference (ACC): May 08-10 2002; Anchorage, Ak; pp191-196.

[6] Hajiyev C, Caliskan F(2005)" Sensor and control surface/actuator failure detection and isolation applied to F-16 flight dynamic". Aircraft Engineering and Aerospace Technology , 77(2)pp152-160.

[7] Mendonca LF, Sousa JMC, da Costa J(2009)" An architecture for fault detection and isolation based on fuzzy methods" Expert Systems with Applications 36(2)pp1092-1104.

[8] Amann P, Perronne JM, Gissinger GL(2001)"Frank PM: Identification of fuzzy relational models for fault detection". Control Engineering Practice, 9(5)pp555-562.

[9] Skrjanc I, Blazic S(2005)" Agamennoni O: Identification of dynamical systems with a robust interval fuzzy model". Automatica , 41(2)pp327-332.

[10] Gustafson D, Kessel W(1979)"Fuzzy clustering with a fuzzy covariance matrix." Proceedings of IEEE CDC, San Diego, CA,pp 761-766.

[11] Nascimento S, Mirkin B, Moura-Pires F(2003)" Modeling proportional membership in fuzzy clustering". Ieee Transactions on Fuzzy Systems , 11(2)pp173-186.

[12] Xiong K, Zhang HY, Chan CW(2006)" Performance evaluation of UKF-based nonlinear filtering". Automatica , 42(2)pp261-270.

[13] Julier SJ, Uhlmann JK(1997)" A new extension of the Kalman filter to nonlinear systems", vol. 3068.

[14] Julier SJ, LaViola JJ(2007)" On Kalman filtering with nonlinear equality constraints". Ieee Transactions on Signal Processing, 55(6)pp2774-2784.

[15] Kandepu R, Foss B, Imsland L(2008)"Applying the unscented Kalman filter for nonlinear state estimation". Journal of Process Control, 18(7-8)pp753-768.

[16] Julier SJ, Uhlmann JK(2004)"Unscented filtering and nonlinear estimation". Proceedings of the Ieee , 92(3)pp401-422.

[17] Li PH, Zhang TW, Ma B(2004)" Unscented Kalman filter for visual curve tracking". Image and Vision Computing, 22(2)pp157-164.

[18] Dini DH, Mandic DP, Julier SJ(2011)"A Widely Linear Complex Unscented Kalman Filter". Ieee Signal Processing Letters, 18(11)pp623-626.

[19] Simon D, Simon DL(2010)" Constrained Kalman filtering via density function truncation for turbofan engine health estimation". International Journal of Systems Science, 41(2)pp159-171.

[20] Chafaa K, Ghanai M, Benmahammed K(2007)" Fuzzy modelling using Kalman filter". Iet Control Theory and Applications, 1(1)pp58-64.

[21] Downs J J, Vogel E F(1993)"A plant-wide industrial process control problem". Computers and Chemical Engineering, 17(3)pp245-255.

Authors

Zhen Luo She was born in 1983. She is a Ph.D. candidate in Huazhong University of Science and Technology. Her research interests include Kalman filtering state estimation and fault diagnosis of networked control sys tems.

Huajing Fang was born in 1955. He is a Professor in Huazhong University of Science and Technology. His research interests include complex networked control systems, dynamics and control of autonomous vehicle's swarm, fault diagnosis and fault-tolerant control.

Xuquan Chen was born in 1984. he is a Ph.D. candidate in Huazhong University of Science and Technology. His research interest is signal processing.