Sample selection based on kernel-subclustering for the signal reconstruction of multifunctional sensors

2013 Meas. Sci. Technol. 24 025102

(http://iopscience.iop.org/0957-0233/24/2/025102)

Meas. Sci. Technol. **24** (2013) 025102 (7pp) doi:10.1088/0957-0233/24/2/025102

Sample selection based on kernel-subclustering for the signal reconstruction of multifunctional sensors

Xin Wang[1], **Guo Wei**[2,3] **and Jinwei Sun**[2]

[1] Department of Mechanical and Electrical Engineering, Xiamen University, Xiamen 361005, People's Republic of China
[2] Department of Automatic Testing and Control, Harbin Institute of Technology, Harbin 150001, People's Republic of China

E-mail: wg_weiguo@yahoo.com.cn

Received 16 May 2012, in final form 19 November 2012
Published 13 December 2012
Online at stacks.iop.org/MST/24/025102

Abstract
The signal reconstruction methods based on inverse modeling for the signal reconstruction of multifunctional sensors have been widely studied in recent years. To improve the accuracy, the reconstruction methods have become more and more complicated because of the increase in the model parameters and sample points. However, there is another factor that affects the reconstruction accuracy, the position of the sample points, which has not been studied. A reasonable selection of the sample points could improve the signal reconstruction quality in at least two ways: improved accuracy with the same number of sample points or the same accuracy obtained with a smaller number of sample points. Both ways are valuable for improving the accuracy and decreasing the workload, especially for large batches of multifunctional sensors. In this paper, we propose a sample selection method based on kernel-subclustering distill groupings of the sample data and produce the representation of the data set for inverse modeling. The method calculates the distance between two data points based on the kernel-induced distance instead of the conventional distance. The kernel function is a generalization of the distance metric by mapping the data that are non-separable in the original space into homogeneous groups in the high-dimensional space. The method obtained the best results compared with the other three methods in the simulation.

Keywords: sample selection, signal reconstruction, kernel-subclustering, multifunctional sensors

1. Introduction

The multifunctional sensor is a type of sensor that integrates two or more sensitive components, so it can detect several different environment quantities simultaneously with size smaller than multi-sensors. Nowadays, the research work on multifunctional sensors is mainly focused on three aspects: new structures of multifunctional sensors [1–4], new materials for multifunctional sensing [5–8] and signal reconstruction methods [9–11].

Figure 1 shows the principle of the multifunctional sensor and its signal reconstruction process. As a multi-input multi-output (MIMO) function, the input–output function of the multifunctional sensor is complex and nonlinear as shown in (1) because of the crossing sensitivity of sensitive components. Each output signal (y_j, $j = 1, \ldots, m$) is a function of all the inputs (x_i, $i = 1, \ldots n$). The aim of the signal reconstruction methods, such as artificial neural network (ANN), support vector regression (SVR) and B-spline [9–11], is to construct the inverse functions g_i, $i = 1, \ldots n$, and give the estimation of the input measurands ($\hat{x}_i$, $i = 1, \ldots, n$).

In the past, there were two ways to improve the accuracy of signal processing: increase the inverse system parameters

[3] Author to whom any correspondence should be addressed.

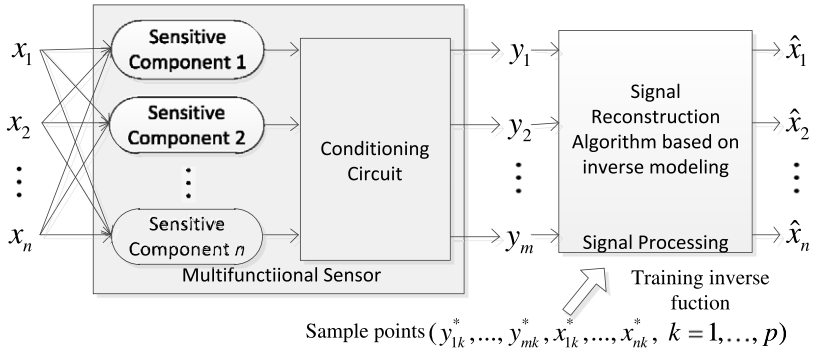

Figure 1. Principle of the multifunctional sensor and signal reconstruction process.

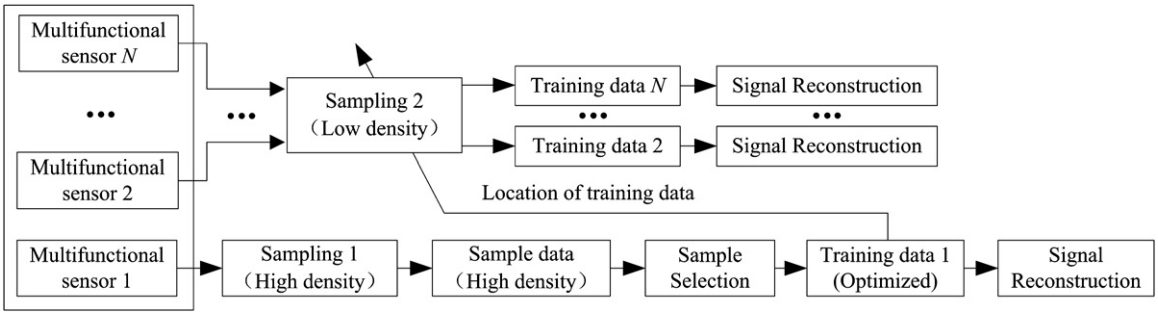

Figure 2. Signal reconstruction optimization for large batches of multifunctional sensors based on sample selection.

or increase the number of sample points. But more parameters meant higher complexity of the algorithms, and more sample points meant more workload for sampling, especially for a large number of sensors. So in this paper, we consider another factor that influences the accuracy, the position of the sample points $(y_{1k}^*, \ldots, y_{mk}^*, x_{1k}^*, \ldots, x_{nk}^*, k = 1, \ldots, p)$, as shown in figure 1. In the past, the sample points for signal processing were given by uniform sampling. It can be improved by a sample-point selection method due to the nonlinearity of the input–output functions of most multifunction sensors. A group of well-selected, representative sample-point positions could help to obtain higher accuracy with fewer sample points effectively;

$$\begin{cases} y_1 = f_1(x_1, \ldots, x_n) \\ \quad \vdots \\ y_m = f_m(x_1, \ldots, x_n) \end{cases} \tag{1}$$

$$\begin{cases} \hat{x}_1 = g_1(y_1, \ldots, y_m) \\ \quad \vdots \\ \hat{x}_n = g_n(y_1, \ldots, y_m) \end{cases}. \tag{2}$$

This sample selection method was very valuable for the signal reconstruction of large batches of multifunctional sensors. It was the key step of the signal reconstruction optimization for large batches of multifunctional sensors (assuming N sensors) we designed, as shown in figure 2. The steps of the optimization were as follows.

(1) Find one multifunctional sensor from the batch of sensors, and give the sample points with a high sampling density by uniform sampling (sampling 1);
(2) use the sample selection method to find the sample points suitable for the signal reconstruction;
(3) give the sample points of all the multifunctional sensors only in the sample position found in step 2 with a low sampling density (sampling 2);
(4) perform the signal reconstruction for all the multifunctional sensors.

Here we assumed that all the multifunctional sensors were qualified and noise was avoided as far as possible during the sampling process. And in step 1, to make sure that the selected multifunctional sensor was qualified, more sensors would be selected as candidates (for example, 4 or 5). Only if there was little difference between their input–output characteristics (no mutation and little random error) could one of them be selected. It was very unlikely that all these sensors had the same problem.

The signal reconstruction optimization provided another way to improve accuracy with a low workload increase for a large number of sensors. In this process, the high sampling density was only performed in one multifunctional sensor. Obviously, in the signal reconstruction optimization, the most important step is the second one. So the sample selection method will be described in detail in the next section.

2. Method

The purpose of the sample selection method is to find the typical points from the sample data; an easy way is to separate the sample data into several groups and use the centroids of each group as the typical points. Therefore, this purpose can be accomplished by a clustering analysis method, because the

purpose of clustering is to identify the natural grouping of data from a large data set to produce a concise representation of a system's behavior [12].

To find an effective clustering method, we used a group of simulation data and tried many popular clustering algorithms (k-means clustering, fuzzy c-means (FCMs) clustering, subtractive clustering, etc) to group them, and then used the cluster centroids to model the inverse system by the signal reconstruction method in [11]. The results showed that the subtractive clustering algorithm can find the most appropriate sample points for inverse modeling.

2.1. Subtractive clustering

Unlike other clustering algorithms, such as FCMs clustering, mountain method, etc, the cluster centroid candidates of subtractive clustering are the sample data themselves. This algorithm can have a large reduction on the number of training samples, based on the density of surrounding data points. Namely, all data points in a small dense zone of one point centroid will be replaced by this typical one. On the other hand, the sparse points in the input space will remain as cluster centroids themselves. So this algorithm is very suitable to select the sample data of multifunctional sensors.

Consider a group of N data points, $Y = \{Y_1, Y_2, \ldots, Y_N\}$, where $Y_i = (y_{1,i}, y_{2,i}, \ldots, y_{n,i})$, $i = 1, \ldots, N$, is an n-dimensional vector. First normalize each point into a unit hyper box range [0 1] to make each dimension identical. Then the subtractive clustering goes as follows.

Step 1: consider each data point as a potential cluster centroid and define a measure of the potential data point Y_i to serve as a cluster centroid:

$$D_i = \sum_{j=1}^{N} \exp\left(-\frac{4}{r_a^2}\|Y_i - Y_j\|^2\right), \qquad i = 1, \ldots, N \tag{3}$$

where $\|.\|$ denotes the Euclidean distance, and $0 < r_a \leqslant 1$ is the clustering radii parameter which defines the neighborhood radius of each point. The constant r_a is a normalized radius defining a neighborhood; the value of D_i is dependent on the number of data points inside this radius. The data outside the radius of Y_i have little influence on its potential. So the point having more data within the radius will get higher potential.

Step 2: select the data point Y_i with the highest D_i, $i = 1, \ldots, N$, as the first cluster centroid Y_1^* and D_i as its potential value D_1^*.

Step 3: reduce the potential value of the remaining data points using (4):

$$D_i = D_i - D_1^* \exp\left(-\frac{4}{r_b^2}\left\|Y_i - Y_1^*\right\|^2\right), \tag{4}$$

where $r_b > 0$ defines the neighborhood of a cluster centroid with which the existence of other cluster centroids is discouraged. The potential value D_i of the points close to the selected cluster centroid will reduce significantly. When $D_i < 0$, the point Y_i is rejected as a cluster centroid forever. So the sample data close to the cluster centroid are replaced by their cluster centroid. Usually, $r_b = 1.25r_a$ or $r_b = 1.5r_a$.

Step 4: select the highest potential value D_k^* from the reduced potential value D_i, and Y_k^* as the next candidate cluster centroid.

Step 5: define the accept ratio ε_{up} above and reject ratio $\varepsilon_{\text{down}}$ below which the candidate centroid will be rejected. If the data were normalized before, the two parameters can use these typical values, $\varepsilon_{\text{up}} = 0.5$ and $\varepsilon_{\text{down}} = 0.15$. Then determine the next cluster centroid by the following criterion.

If $\frac{D_k^*}{D_1^*} > \varepsilon_{\text{up}}$, accept the candidate centroid Y_k^* as the next cluster centroid and go to the next step;

If $\frac{D_k^*}{D_1^*} < \varepsilon_{\text{down}}$, reject Y_k^* and finish the algorithm;

If $\varepsilon_{\text{down}} < \frac{D_k^*}{D_1^*} < \varepsilon_{\text{up}}$, compute the minimum distance $d_{\min}$ between the candidate Y_k^* and all the cluster centroids already selected;

If $\frac{d_{\min}}{r_a} + \frac{D_k^*}{D_1^*} \geqslant 1$, accept the candidate centroid Y_k^* as the next cluster centroid and go to the next step;

else reject Y_k^*, set $D_k^* = 0$ and go to step 4 to find the new candidate cluster centroid.

Step 6: compute the potential value for the remaining data points:

$$D_i = D_i - D_1^* \exp\left(-\frac{4}{r_b^2}\left\|Y_i - Y_k^*\right\|^2\right) \tag{5}$$

then go to step 4.

From the steps of subtractive clustering, we can find that the number of the cluster centroid cannot be determined beforehand. A smaller cluster radius will yield more centroids and a larger one will lead to less centroids. So in practice, it is necessary to test the value of the clustering radii and select an adequate one according to the results obtained.

2.2. Kernel function

The kernel function provides a way of increasing the accuracy of the subtractive method by mapping the data points from input space to a high-dimensional space in which distance is measured using a kernel function. The distances calculated in a high-dimensional space are much more informative than those of the conventional subtractive method calculated in the original space, leading to more accurate selection of the cluster centroids [13, 14].

Considering the data set $Y = \{Y_1, Y_2, \ldots, Y_N\}$ in section 2.1, let ϕ be a nonlinear mapping function from input space to a high-dimensional feature space H:

$$\phi: R^n \to H \quad Y \mapsto \phi(Y). \tag{6}$$

By applying the nonlinear mapping function ϕ, the dot product $(Y_i \cdot Y_j)$ in the input space is mapped to $(\phi(Y_i) \cdot \phi(Y_j))$ in the feature space. The key notion of the kernel function is that the dot product in the feature space can be calculated without the explicit specification of the mapping function ϕ by the kernel function:

$$\phi(Y_i) \cdot \phi(Y_j) = K(Y_i, Y_j). \tag{7}$$

Three commonly used kernel functions are the polynomial kernel function, the Gaussian kernel function and the sigmoidal kernel function, as shown below:

$$K(Y_i, Y_j) = (Y_i \cdot Y_j + c)^d \tag{8}$$

$$K(Y_i, Y_j) = \exp\left(-\frac{\|Y_i - Y_j\|^2}{2\sigma^2}\right) \tag{9}$$

$$K(Y_i, Y_j) = \tanh(\kappa(Y_i \cdot Y_j) + \vartheta), \tag{10}$$

where $c \geqslant 0$, $d \in \mathrm{N}$, $\sigma > 0$, $\kappa > 0$ and $\vartheta < 0$.

2.3. Kernel-subclustering for sample selection

If the data set $Y = \{Y_1, Y_2, \ldots, Y_N\}$ in sections 2.1 and 2.2 is mapped to the feature space by some mapping function ϕ, the Euclidean distance in the feature space can be expressed as follows:

$$\begin{aligned}\|\phi(Y_i) - \phi(Y_j)\|^2 &= (\phi(Y_i) - \phi(Y_j)) \cdot (\phi(Y_i) - \phi(Y_j)) \\ &= \phi(Y_i) \cdot \phi(Y_i) - 2\phi(Y_i) \cdot \phi(Y_j) \\ &\quad + \phi(Y_j) \cdot \phi(Y_j) \\ &= K(Y_i, Y_i) - 2K(Y_i, Y_j) + K(Y_j, Y_j). \end{aligned} \tag{11}$$

The cluster centroid selection procedure of the kernel-subclustering is similar to that of the subtractive method. The Euclidean distance calculation is modified by equation (11), and the other parts of the procedure in section 2.1 are unchanged. So the new equations for the potential value calculation of data points are

$$D_i = \sum_{j=1}^{N} \exp\left(-\frac{4}{r_a^2}(K(Y_i, Y_i) - 2K(Y_i, Y_j) + K(Y_j, Y_j))\right), \quad i = 1, \ldots, N \tag{12}$$

$$D_i = D_i - D_1^* \exp\left(-\frac{4}{r_b^2}\left(K(Y_i, Y_i) - 2K\left(Y_i, Y_1^*\right) + K\left(Y_1^*, Y_1^*\right)\right)\right) \tag{13}$$

$$D_i = D_i - D_1^* \exp\left(-\frac{4}{r_b^2}\left(K(Y_i, Y_i) - 2K\left(Y_i, Y_k^*\right) + K\left(Y_k^*, Y_k^*\right)\right)\right). \tag{14}$$

Therefore, replace equations (3)–(5) by equations (12)–(14); the subtractive clustering method is changed into the kernel-subclustering method. In this paper, the Gaussian function is used as the kernel function in the simulation.

3. Results and discussion

To demonstrate the effectiveness of the proposed method, we applied the kernel-subclustering method, two conventional methods (subtractive clustering and FCM) and uniform sampling to a common circuit model of multifunctional sensors. In the simulation, a fourth-order B-spline method [11] was applied to model the inverse system of the circuit model and reconstruct the measurands using the selected sample points, and then the errors between the signal reconstruction results and the overall samples were compared to evaluate the effectiveness of the four sample selection methods, as shown in figure 3.

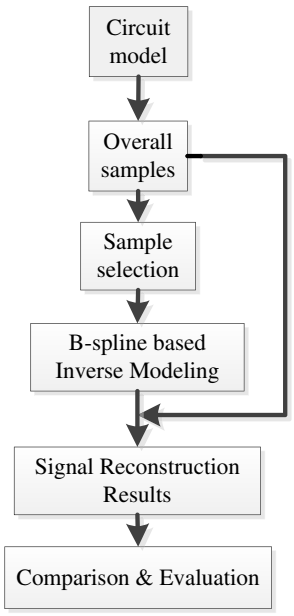

Figure 3. Procedure of the simulation.

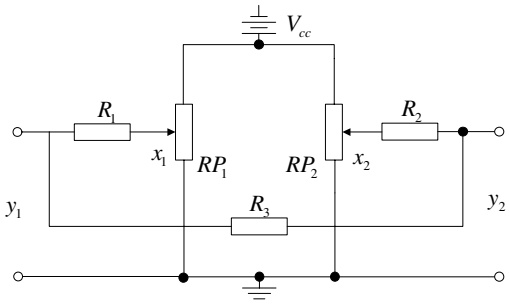

Figure 4. Circuit model of the two-input two-output multifunctional sensor.

3.1. Circuit model of the multifunctional sensor

The circuit model is used as the equivalent of the multifunctional sensor (figure 4). The structure of the circuit model is simple and its input–output characteristic is similar to those of many multifunctional sensors [15]. In the circuit model, $V_{cc} = 5$ V, $R_1 = R_2 = R_3 = 1$ kΩ are the known parameters and $RP_1 = RP_2$. The inputs are the ratios of the lower resistance in the slide rheostats RP_1 and RP_2 to their total resistance. The outputs are the voltages of y_1 and y_2.

Therefore, the transfer function (15) for y_1 of the circuit model can be derived from the Kirchhoff law. The ratio $l = RP_1/R_1 = RP_2/R_2$ indicates the nonlinearity of the circuit model. Then the resistances and voltage are substituted by their values and equations (16) and (17) can be obtained. From these two equations, we can find that the greater the value of l, the higher is the nonlinearity of the model. In this simulation, l is

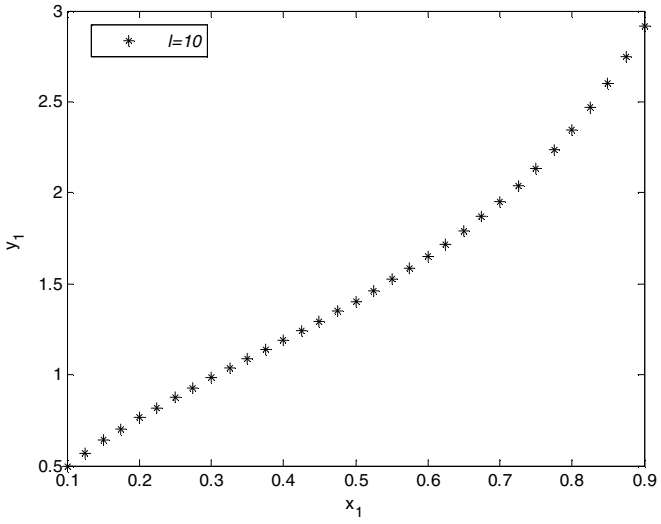

Figure 5. Relationship between input x_1 and output y_1 of the circuit model.

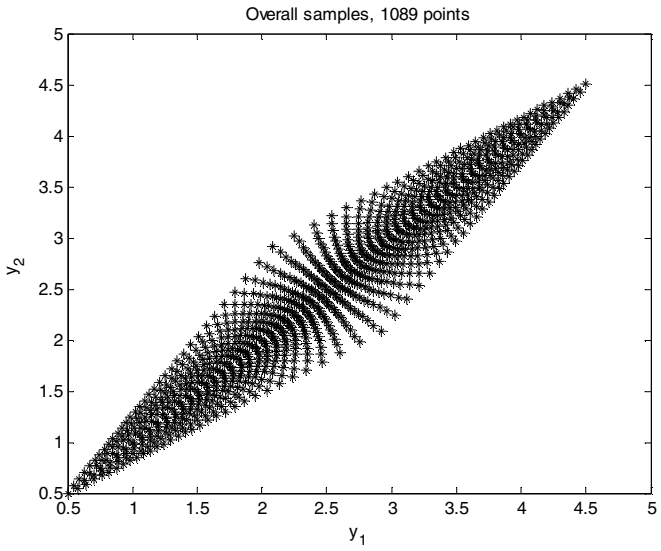

Figure 6. Inputs of the inverse system for the circuit model.

set to 10 which is a relatively large value. Figure 5 shows the relationship between x_1 and y_1 in the case of $l = 10$;

$$y_1 = \frac{V_{cc}\,[2x_1 + x_2 + lx_1x_2(2 - x_1 - x_2)]}{R_1 + R_2 + R_3 + l\,[x_2(1 - x_2) + x_1(1 - x_1)]} \tag{15}$$

$$y_1 = \frac{5\,[2x_1 + x_2 + lx_1x_2(2 - x_1 - x_2)]}{3 + l\,[x_2(1 - x_2) + x_1(1 - x_1)]} \tag{16}$$

$$y_2 = \frac{5\,[2x_2 + x_1 + lx_1x_2(2 - x_1 - x_2)]}{3 + l\,[x_2(1 - x_2) + x_1(1 - x_1)]}. \tag{17}$$

To perform the sample selection methods, the number of overall samples must be sufficient. So the values of x_1 and x_2 were given in the range of [0.1 0.9] at the interval of 0.025 (1089 samples), and the values of y_1 and y_2 can be given using (16) and (17). Therefore, the inputs of the inverse system are y_1 and y_2, and the outputs are x_1 and x_2. Figure 6 shows the inputs of the inverse system for the circuit model, and the sample selection methods are performed on the data set $Y = (y_{1,i}, y_{2,i})$, $i = 1, \ldots, 1089$.

In the simulation, the FCM method ran 500 times with the initial centroids randomly selected from the data set. The parameters of the FCM method were set to a termination criterion $\varepsilon = 1 \times 10^{-5}$ and weight exponent $m = 2.0$. To evaluate the effectiveness of the subtractive clustering and kernel-subclustering methods, we changed the values of the clustering radii parameter r_a to give different numbers of selected samples (120–300 points). Another important parameter r_b was set to $1.25r_a$. Figures 7(*a*)–(*d*) show the select points of the four methods when the number of the samples was 169. From figure 7(*a*), we can see that although x_1 and x_2 were uniformly sampled, the distribution of y_1 and y_2 was non-uniform because of the nonlinearity of the circuit system. So the uniform sampling was not an optimal sample selection method for the inverse modeling.

To further evaluate the performance of samples selected by the four methods, we used the samples to model the inverse systems of the circuit model by a fourth-order B-spline [11] (figure 1) and compute the error between x_1, x_2 and $\hat{x}_1$, $\hat{x}_2$ in the range [0.2 0.8]. The expressions of the inverse systems for $\hat{x}_1$ and $\hat{x}_2$ are

$$\hat{x}_{1,i} = \sum_{j=-K+1}^{L} \sum_{p=-K+1}^{M} [c_{j,p} B_{j,K}(y_{1,i}) B_{p,K}(y_{2,i})], \quad i = 1, \ldots, N \tag{18}$$

$$\hat{x}_{2,i} = \sum_{j=-K+1}^{L'} \sum_{p=-K+1}^{M'} [c'_{j,p} B_{j,K}(y_{1,i}) B_{p,K}(y_{2,i})], \quad i = 1, \ldots, N, \tag{19}$$

where $B_{j,K}()$ and $B_{p,K}()$ are the B-spline basis functions for $y_{1,i}$ and $y_{2,i}$, $\{c_{j,p}\}$ and $\{c'_{j,p}\}$ are the B-spline coefficients, $K = 4$ is the order of the B-spline, and L, M and L', M' are the numbers of the internal B-spline knots for $y_{1,i}$ and $y_{2,i}$ in (18) and (19), respectively. Herein, in the simulation, all the B-spline-based inverse models used the same number of knots ($L = M = 18$, $L' = M' = 18$) and the same coefficient training method (extend Kalman filter, EKF). Obviously, the results for x_1 and x_2 were the same, so we only gave the results for x_1.

The signal reconstruction results are shown in figures 8 and 9 in detail. These two figures show the local and global characteristics, respectively, by maximum relative error (MRE) and mean square error (MSE). Herein, the MSE can be calculated by

$$\mathrm{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{j,i} - \hat{x}_{j,i})^2}, \qquad j = 1, 2. \tag{20}$$

Obviously, no matter the local or global characteristics, the reconstruction results of the uniform sampling were the worst of the four methods. The MRE increased to greater than 1.5% and the MSE increased to 6.72×10^{-4} when the number of selected samples was decreased to 196. Moreover, the MRE of the uniform sampling increased to 2.52% when the number of selected points was 256, whereas the MRE was 0.77% when the number of selected points was 225. This was an obvious

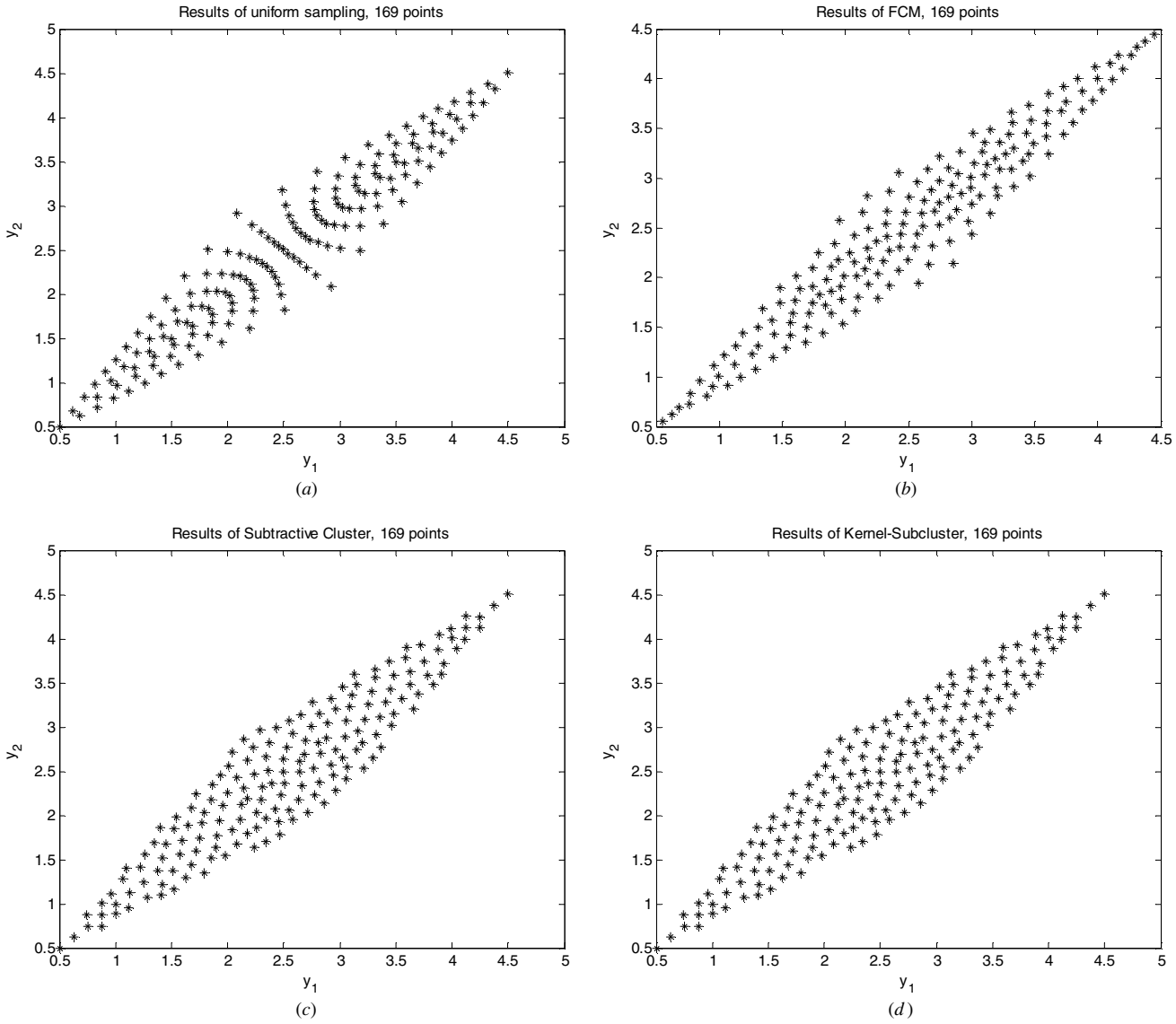

Figure 7. Sample selection results of four methods: (*a*) results of uniform sampling, 169 points; (*b*) results of FCM, 169 points; (*c*) results of subtractive clustering, 169 points; (*d*) results of kernel-subclustering, 169 points.

disadvantage of uniform sampling: key points with important information may be lost by uniform sampling even when the density of the sample rate was increased. Therefore, it was very necessary and meaningful to develop a powerful sample selection method.

In general, the global characteristics of the other three methods were similar, as shown in figure 9, although the kernel-subclustering was a little better than the other two methods when the number of selected points was less than 140. And the local characteristics of the FCM and subtractive clustering were also very close. To ensure that the MRE was less than 1%, the number of selected points must be greater than 160. However, the FCM-based method had a disadvantage that the selected points were different each time we ran the program because of the random initial centroids. So the results of the FCM cannot be guaranteed in a certain run, and figures 8 and 9 only show the best of several runs. Moreover, to improve the signal reconstruction results, we also tried the kernel-based FCM [16], but its results did not improve compared with the FCM.

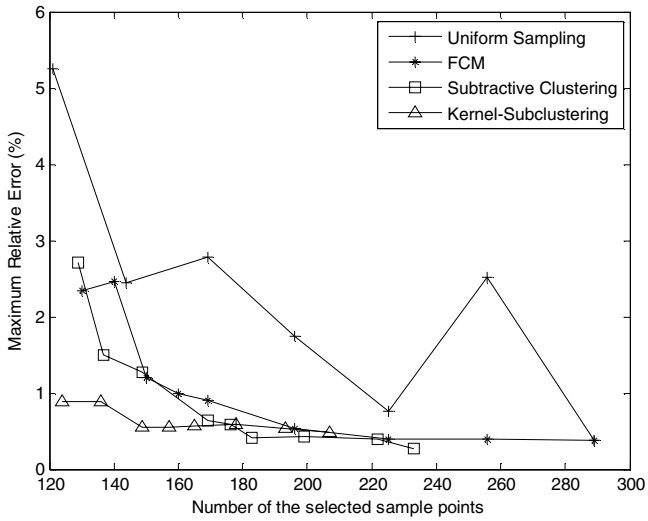

Figure 8. MRE using different numbers of selected samples by the four methods.

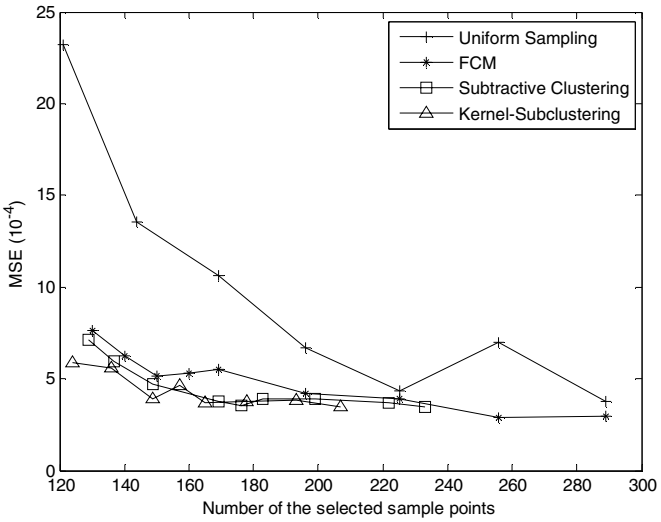

Figure 9. MSE using different numbers of selected samples by the four methods.

The local characteristics of the kernel-subclustering were the best of the four methods. The MRE of the kernel-subclustering was the lowest of the four methods (<1%) when the number of selected points was less than 165 and the MRE was still less than 1% (0.88%) when the number of the selected points was 124. Therefore, considering the global and local characteristics simultaneously, the kernel-subclustering method can be a very good candidate for the sample selection of multifunctional sensors.

4. Conclusion

In this paper, we have presented a signal reconstruction optimization method, as well as its sample selection algorithm kernel-subclustering, to decrease the workload of signal reconstruction for large batches of multifunctional sensors. The simulation was performed to reconstruct the signal of a circuit model whose input–output characteristics were very similar to those of real multifunctional sensors. Four methods (uniform sampling, FCM, subtractive clustering and kernel-subclustering) were applied in the simulation. The simulation results showed that the proposed kernel-subclustering can obtain higher accuracy with fewer selected points than the other three methods. Moreover, the proposed optimization method and its sample selected algorithm can also be a good candidate for the signal reconstruction of traditional sensors. Further work can be focused on applying the proposed method to the signal reconstruction for real multifunctional sensors.

References

[1] Won J, Choa S H and Zhao Y L 2006 An integrated sensor for pressure, temperature, and relative humidity based on MEMS technology *J. Mech. Sci. Technol.* **20** 505–12

[2] Kimoto A and Shida K 2008 A new multifunctional sensor using piezoelectric ceramic transducers for simultaneous measurements of propagation time and electrical conductance *IEEE Trans. Instrum. Meas.* **57** 2542–7

[3] Takao H, Yawata M, Sawada K and Ishida M 2010 A multifunctional integrated silicon tactile imager with arrays of strain and temperature sensors on single crystal silicon diaphragm *Sensors Actuators* A **160** 69–77

[4] Yu H *et al* 2009 A novel design of multifunctional integrated cell-based biosensors for simultaneously detecting cell acidification and extracellular potential *Biosens. Bioelectron.* **24** 1462–8

[5] Cotton D P J, Graz I M and Lacour S P 2009 A multifunctional capacitive sensor for stretchable electronic skins *IEEE Sensors J.* **9** 2008–9

[6] Saleem M *et al* 2009 Cu(II) 5,10,15,20-tetrakis(4′-isopropylphenyl) porphyrin based surface-type resistive-capacitive multifunctional sensor *Sensors Actuators* B **137** 442–6

[7] Kang T J, Lim D-K, Nam J-M and Kim Y H 2010 Multifunctional nanocomposite membrane for chemomechanical transducer *Sensors Actuators* B **147** 691–6

[8] Saleem M *et al* 2009 Surface-type multifunctional sensor based on 5,10,15,20-tetrakis(4′-isopropylphenyl) porphyrin *J. Mater. Sci.* **44** 1192–7

[9] Liu X *et al* 2009 Nonlinear multifunctional sensor signal reconstruction based on least squares support vector machines and total least squares algorithm *J. Zhe Jiang Univ. Sci.* A **10** 497–503

[10] Futane N P *et al* 2010 ANN based CMOS ASIC design for improved temperature-drift compensation of piezoresistive micro-machined high resolution pressure sensor *Microelectron. Reliab.* **50** 282–91

[11] Wang X, Wei G and Sun J 2011 B-spline approximation using an EKF for signal reconstruction of nonlinear multifunctional sensors *IEEE Trans. Instrum. Meas.* **60** 1952–8

[12] Chiu S L 1994 Fuzzy model identification based on cluster estimation *J. Intell. Fuzzy Syst.* **2** 267–78

[13] Muller K R *et al* 2001 An introduction to kernel-based learning algorithm *IEEE Trans. Neural Netw.* **12** 181–201

[14] Girolami M 2003 Mercer kernel-based clustering in feature space *IEEE Trans. Neural Netw.* **13** 780–4

[15] Flammini A, Marioli D and Taroni A 1999 Application of an optimal look-up table to sensor data processing *IEEE Trans. Instrum. Meas.* **48** 813–6

[16] Kannan S R *et al* 2011 Robust kernel FCM in segmentation of breast medical images *Expert Syst. Appl.* **38** 4382–9