Identification of Key Variables Using Fuzzy Average With Fuzzy Cluster Distribution

Yanfeng Hou, *Student Member, IEEE*, Jacek M. Zurada, *Fellow, IEEE*, Waldemar Karwowski, William S. Marras, and Kermit Davis

***Abstract*—Identification of the significance of input variables is very important for complex systems with high-dimensional input space. In this paper, a method using fuzzy average with fuzzy cluster distribution is proposed. To avoid the interference of different distributions of the sampling data, the distribution of fuzzy clusters in the sampling data is considered, instead of the original data set. To discover the input–output relationship, the methods of fuzzy rules and fuzzy C-means are first used to partition the original sampling data set into fuzzy clusters. A new data set with the same distribution of the fuzzy clusters is produced. The fuzzy average method is then applied to the new data set. By doing so, the interference of distribution of the original sampling data is removed. This method is straightforward and computationally easy. The performance is tested on both benchmark data and real-world data.**

***Index Terms*—Fuzzy cluster, variable identification.**

I. Introduction

In modeling of complex systems whose input–output relationship is not well understood, it is very helpful to find out the significance of each variable to the output of the system. If inputs that have little or no influence on the output can be removed and emphasis is put on the important variables, a more parsimonious and more effective model can be built. Moreover, a better understanding to those systems can be achieved if how individual inputs affect the EMG/forces is discovered.

A. Problem Statement

The problem being investigated can be stated as follows.

For a system with one output variable and n associated input variables, m sampling data points are obtained. Each data point represents a joint measurement of all variables involved. The input data vectors are of such form

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \cdots \\ \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} x_1^1 & x_1^2 & \cdots & x_1^m \\ x_2^1 & x_2^2 & \cdots & x_2^m \\ \cdots & & & \\ x_n^1 & x_n^2 & \cdots & x_n^m \end{bmatrix} \tag{1}$$

Manuscript received February 21, 2005; revised February 18, 2006 and May 2, 2006. This work was supported by the National Institute for Occupational Safety and Health under a research grant on the "Development of a Neuro-Fuzzy System to Predict Spinal Loading as a Function of Multiple Dimensions of Risk."

Y. Hou and J. M. Zurada are with the Department of Electrical and Computer Engineering, University of Louisville, Louisville, KY 40292 USA (e-mail: y0hou002@louisville.edu; jacek.zurada@louisville.edu).

W. Karwowski is with the Department of Industrial Engineering, University of Louisville, Louisville, KY 40292 USA (e-mail: karwowski@louisville.edu).

W. S. Marras is with the Institute for Ergonomics, The Ohio State University, Columbus, OH, 43210 USA (e-mail: marras.1@osu.edu).

K. Davis is with the Department of Environmental Health, University of Cincinnati, Cincinnati, OH, 45267 USA (e-mail: kermit.davis@uc.edu).

Digital Object Identifier 10.1109/TFUZZ.2006.889897

where x_i^j denotes the jth measurement of input variable $x_i (i = 1, 2, \ldots n; j = 1, 2, \ldots, m)$.

The output data vectors are of the following form:

$$\mathbf{y} = [y^1 \quad y^2 \quad \cdots \quad y^m] \tag{2}$$

where y^j denotes the jth $(j = 1, 2, \ldots, m)$ measurement of the output variable y.

From the above sampling data of input–output pairs, the relationship between each input variable and the output variable (the $x_i - y$ relationship, $i = 1, 2, \ldots n$) should be found.

The difficulty of finding the $x_i - y$ relationship resides in that the change of y is caused by the joint influence of all the input variables, instead of only the influence of x_i.

B. Related Work

This problem has a broad applicability and some methods have been proposed. In [1], Bartlett used a neural network method. During training, the algorithm automatically constructs an appropriate neural network architecture. The importance of each input variable is provided as a by-product. This method cannot determine in what way each input variable affects the outputs. Also, different neural networks need to be developed, which is time consuming.

Yuan and Klir used the full class of Mahalanobis distances to search by an evolutionary algorithm for the optimal distance—one under which the fuzzy c-means algorithm produces a fuzzy partition of the given data set that is as close as possible to the given crisp partition [2]. The contribution of each variable to this partition is then inferred from parameter values of the optimal Mahalanobis distance. This method employs evolutionary computation and is computationally intensive. The result obtained is not straightforward. Also, it brings no information on how input variables affect the outputs.

Sugeno and Takahiro proposed an iterative algorithm for the input identification [3]. Different models are generated to search for the optimal combination of variables. The total number of combinations is $2^n - 1$, where n is the number of input variables. For a high-dimensional system with many input variables, the number of models needed will be very large.

Chiu used a backward selection procedure that starts with all possible variables and reduces one variable at each stage [4]. Premises in the fuzzy rules of an initial model are systematically removed to search for the best simplified model. However, after the iteration process of searching, different results may be obtained when three different criteria for model selection are used. Then a Takagi–Sugeno type model needs to be generated for each solution. The final conclusion is made based on the

comparison of the model errors, which makes the method complicated and time consuming.

In [5] and [6], Lin *et al.* proposed their "fuzzy curves" method. For each input variable x_i, m data points are plotted in the $x_i - y$ space. A fuzzy rule is defined according to each sampling data point $(x_i^j, y^j)(i = 1, 2, \ldots, n, j = 1, 2, \ldots, m)$ in the following form:

$$R^j: \text{ IF } x_i \text{ is } \mu_{ij}(x_i) \text{ THEN } y \text{ is } y_j$$

where $\mu_{ij}(x_i)$ is a Gaussian membership function of x_i^j. A "fuzzy curve" can be produced using the defuzzification method, which stands for the $x_i - y$ relationship. The significance of the input variables is ranked according to the ranges covered by the fuzzy curves.

This method is easy to understand and to calculate. The result obtained is straightforward. Lin *et al.* used this method in fuzzy-neural system modeling to determine model structure and set the initial weights in the model [5]. This method was also used in many other papers such as [6]–[13]. In [6], the method is used to eliminate spurious inputs and dependent inputs. In [7], this method was used to set the initial parameters of the neurofuzzy model. In [8], it was used to rank the input variables and determine the optimal rules describing the behavior of the system. In [9], it was used to generate fuzzy models for short-term load forecasting.

When trying to apply this method to the electromyographic (EMG) signal estimation system for manual lifting tasks, it is found that it did not always work well. The distribution of the sampling data set will affect the result. In other words, significance of the input variables obtained from this method may vary from sampling to sampling. Obviously, this should not happen because influence of each input variable is an inherent property of a system, regardless of the distribution of the sampling data.

In Section II, limitation of the fuzzy curve method is pointed out and is improved with fuzzy average with fuzzy cluster distribution (FAFCD). In Section III, the proposed method is tested on benchmark data and the EMG estimation system.

II. Methods

As mentioned in the prior section, in the method of fuzzy curves, a fuzzy rule is defined according to each sampling data point (x_i^j, y^j). From m data points, m fuzzy rules can be obtained. The fuzzy membership functions for input variable x_i are Gaussian membership functions centered at x_i^j

$$\mu_{ij}(x_i) = \exp\left(-\left(\frac{x_i - \overline{x}_i^j}{\sigma}\right)^2\right) \tag{3}$$

where $\overline{x}_i^j$ and σ are center and width of the membership function, respectively.

Then the fuzzy curve is produced from defuzzification

$$C_i(x_i) = \frac{\sum_{j=1}^{m} y^j \mu_{ij}(x_i)}{\sum_{j=1}^{m} \mu_{ij}(x_i)}. \tag{4}$$

The authors of [5] demonstrated and validated this method using a nonlinear system defined as

$$y = \left(2 + x_1^{-1.5} - 1.5\sin(3x_2)\right)^2, \quad 0 \le x_1, \quad x_2 \le 3 \tag{5}$$

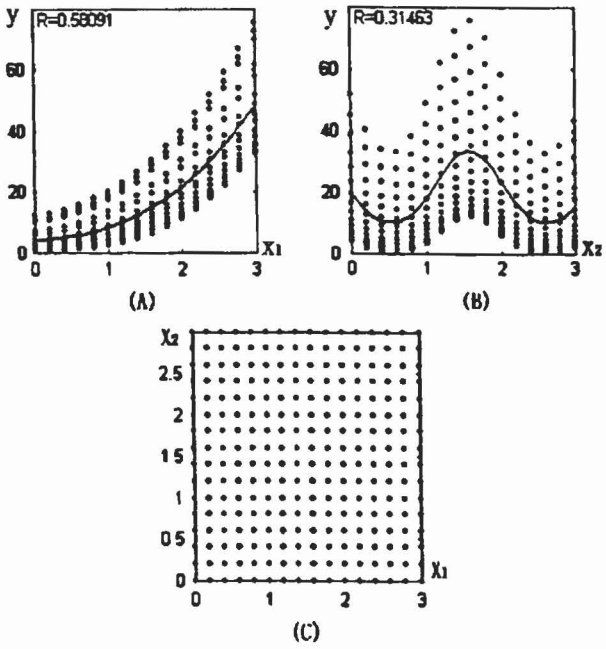

Fig. 1. (A) The $x_1 - y$ relationship from defuzzification. (B) The $x_2 - y$ relationship from defuzzification. (C) The distribution of inputs x_1 and x_2.

where x_1 and x_2 are two input variables and y is the output variable.

Here the defuzzified curves of y in $x_1 - y$ and $x_2 - y$ space are plotted in Fig. 1(A) and (B), respectively. The sample data were generated using formula (5) with x_1 and x_2 uniformly distributed in the interval [0,3] in Fig. 1(C). R in the figure is the ratio of the range of y covered by the curve to the whole range of y. It is used to represent significance of the corresponding input variable instead of using the ranges covered by the fuzzy curves as in [5].

It can be seen that the curves can correctly reflect the $x_1 - y$ relationship and $x_2 - y$ relationship. However, can this still work if the distribution of the sample data changes? Fig. 2(A) gives us a different result of the $x_2 - y$ relationship when the data were generated with x_1 and x_2 shown in Fig. 2(B). It seems that if the sampling data are not uniformly distributed, the curves will be distorted. This example shows that the fuzzy curve method puts restrictions on the distribution of sampling data. Below the reason will be pointed out and a new approach will be derived.

A. The Method Depends on Distribution of Sampling Data

As stated before, the importance of the input variables is ranked according to the ratio of the range of y covered by the curve produced from defuzzification to the whole range of y. Let us define the ratio as influence rate R; then for input variable x_i

$$R_{x_i} = \frac{C_i\left(x_i^u\right) - C_i\left(x_i^l\right)}{a} \tag{6}$$

where $C_i(x_i^u)$ is the highest point on the curve and $C_i(x_i^l)$ is the lowest point on the curve. a is the whole range of y.

$C_i(x_i^u)$ and $C_i(x_i^l)$ are calculated from (4). The membership of each value of x_i to all the m membership functions is cal-

culated for (4). In (3), the width of the Gaussian membership function is often taken as about 20% of the length of the input interval of x_i. If the width σ is very small, only those membership functions with a center (mean of the Gaussian function) close to current value of x_i will have a significant value, while the membership functions far from it will have a value close to zero. When $\sigma \to 0$

$$R_{x_i}(\sigma \to 0) = \lim_{\sigma \to 0} \left(\frac{C_i\left(x_i^u\right) - C_i\left(x_i^l\right)}{a} \right). \tag{7}$$

Substituting (3) and (4) into (7) yields

$$\begin{aligned} & R_{x_i}(\sigma \to 0) \\ & = \frac{1}{a} \lim_{\sigma \to 0} \left[\frac{\sum_{j=1}^m y^j \mu_{ij}\left(x_i^u\right)}{\sum_{j=1}^m \mu_{ij}\left(x_i^u\right)} - \frac{\sum_{j=1}^m y^j \mu_{ij}\left(x_i^l\right)}{\sum_{j=1}^m \mu_{ij}\left(x_i^l\right)} \right] \\ & = \frac{1}{a} \lim_{\sigma \to 0} \left[\frac{\sum_{j=1}^m y^j \exp\left(- \left(\frac{x_i^u - \overline{x}_i^j}{\sigma} \right)^2 \right)}{\sum_{j=1}^m \exp\left(- \left(\frac{x_i^u - \overline{x}_i^j}{\sigma} \right)^2 \right)} \right. \\ & \qquad \left. - \frac{\sum_{j=1}^m y^j \exp\left(- \left(\frac{x_i^l - \overline{x}_i^j}{\sigma} \right)^2 \right)}{\sum_{j=1}^m \exp\left(- \left(\frac{x_i^l - \overline{x}_i^j}{\sigma} \right)^2 \right)} \right] \end{aligned} \tag{8}$$

when $\sigma \to 0$, only those membership functions with their centers equal to x_i^u and x_i^l need to be taken into account. The memberships of x_i^u and x_i^l to the other membership functions are zero. Suppose there are s membership functions with a center equal to x_i^u and h membership functions with a center equal to x_i^l; then (8) becomes

$$R_{x_i}(\sigma \to 0) = \frac{1}{a} \left(\frac{\sum_{k=1}^s y_k^u}{s} - \frac{\sum_{k=1}^h y_k^l}{h} \right) \tag{9}$$

where $y_k^u (k = 1, 2, \ldots, s)$ are the values of y when $u_{ij}(x_i^u) = 1$; and $y_k^l (k = 1, 2, \ldots, h)$ are the values of y when $u_{ij}(x_i^l) = 1$.

Equation (9) indicates that the range of the curve (when $\sigma \to 0$) is the difference between the average value of y at $x_i = x_i^u$ and the average value of y at $x_i = x_i^l$.

If σ is not approaching zero, the value of C_{x_i} at $x_i = x_i^u$ takes into account those data points around $x_i = x_i^u$. But it is still a weighted average. Since it has a meaning of average carried in a fuzzy sense, we call it fuzzy average. After expressing the value of the defuzzified curve for x_i at $x_i = x_i^u$ as a weighted average that takes the points around $x_i = x_i^u$ into account, the next thing needed is to find out what determines the value of this average. For simplicity, the condition when $\sigma \to 0$ is considered. In this condition, the value of the fuzzy average at $x_i = x_i^u$ is the arithmetical average of y at $x_i = x_i^u$

$$C_i(x_i = x_i^u) = \frac{\sum_{k=1}^s y_k^u}{s}. \tag{10}$$

$C_i(x_i = x_i^u)$ is the fuzzy average value of y at $x_i = x_i^u$ in the $x_i - y$ space. The value of $C_i(x_i = x_i^u)$ depends on the values of y at $x_i = x_i^u$ (the values of y_k^u). The values of y_k^u are decided by both the system function and the values of the input variables.

Let us define the system function as

$$\mathbf{y} = f_s(x_1, x_2, \ldots, x_n) = f_s(\mathbf{x}) \tag{11}$$

where f_s is the system function and $\mathbf{y}$ is the output vector.

From (10), the average value of y at $x_i = x_i^{(u)}$ in the $x_i - y$ space is

$$C_i\left(x_i = x_i^u\right) = \frac{\sum_{k=1}^s f_s\left(x_1^k, x_2^k, \ldots, x_n^k\right)}{s} \tag{12}$$

where $k = 1, 2, \ldots, s$, assuming s data points at $x_i = x_i^u$. For easier notation, fuzzy average of y in the $x_1 - y$ space is considered. Using $\mathbf{x}_{(2-n)}$ to represent $[x_2, x_3, \ldots, x_n]$, the fuzzy average of y in the $x_1 - y$ space becomes

$$C_1(x_1) = \frac{\sum_{k=1}^t f_s\left(x_1, \mathbf{x}_{(2-n)}^k\right)}{t} \tag{13}$$

where t is the number of data points at each value of x_1 (they are not the same for different values of x_1).

From (13), it can be seen that only when vectors $\mathbf{x}_{(2-n)}^k (k = 1, 2, \ldots, t)$ are the same for all x_1, $C_1(x_1)$ is determined only by x_1, so that function (13) can reflect the $x_1 - y$ relationship. Otherwise $C_1(x_1)$ is determined by all the input variables, and therefore cannot reflect the relationship between x_1 and y.

Fig. 1 is the two-dimensional example with two input variables x_1 and x_2, in which x_1 has the same values at each value of x_2 [Fig. 1 (C)]. The fuzzy average of y in the $x_2 - y$ space reflects the $x_2 - y$ relationship correctly. While in Fig. 2, x_1 does not have the same values at each value of x_2. Therefore, the fuzzy average of y cannot correctly reflect the $x_2 - y$ relationship.

Thus it can be stated that to find out the $x_i - y$ relationship using fuzzy average method, each of the input variables should have the same values along the axis of x_i, respectively. For the real-world data, this requirement is normally hard to meet. However, if many sampling data points spread all the range of x_1, the fuzzy average of y in $x_2 - y$ space can reflect the $x_2 - y$ relationship as long as x_1 has roughly the same distribution at any value of x_2. Fig. 3 shows the result of this condition. In Fig. 3(B), x_1 was generated randomly using Matlab command "rand" with uniform distribution. The fuzzy average method works in this situation [Fig. 3(A)].

For many practical applications, it cannot be assumed that all input variables have the same distribution along x_i axis. This means for a certain system, if different sampling data sets are used, the fuzzy average may be different. Then the conclusion for the importance of input variables may be different. Therefore, a more representative data set is needed to determine the influence of input variables.

B. Change the Distribution of Sampling Data Using Fuzzy Clustering

Using fuzzy average method, the significance of variable x_i can be correctly evaluated without the interference of other input variables only when all other input variables have the same distribution along x_i axis. To transform the sampling data set into this form, fuzzy clustering is used to change the distribution of

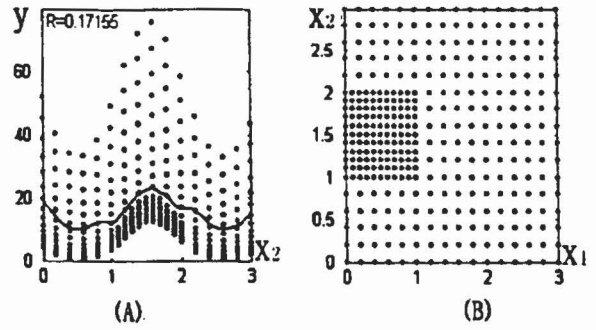

Fig. 2. (A) The $x_2 - y$ relationship from defuzzification. (B) The altered distribution of inputs.

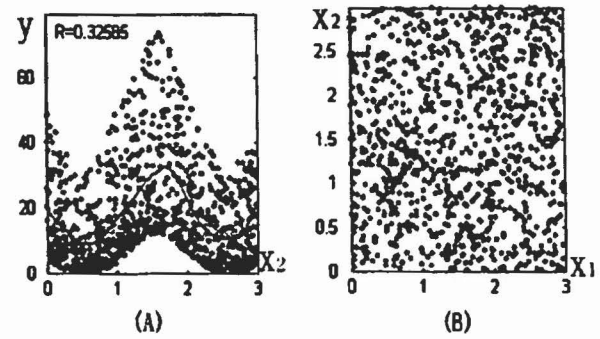

Fig. 3. (A) The $x_2 - y$ relationship from defuzzification. (B) The random distribution of inputs.

the data set. Again the example of the system with two input variables x_1 and x_2 shown in Fig. 1 is used.

Suppose more data points fall into a small region than into other regions in the $x_1 - x_2$ space [as in Fig. 2(B)]. Fuzzy clustering methods are considered to divide the data points into groups. The number of data points in each group (fuzzy cluster) will be different since the distribution of the data is uneven. If one data point (for instance, the fuzzy cluster center) is used to represent each group, a new data set with the distribution of fuzzy clusters can be obtained. Since a different number of sampling data in small regions will be replaced by the same number of cluster center, a new data set with better distribution may be obtained.

Below, the fuzzy C-means method is used to cluster the data. Then, after discussing the drawback in this application, an improved method is proposed.

1) Fuzzy C-Means Method: Each of the data points represents a point in the n-dimensional Euclidean space (n is the input dimension). The purpose of clustering is to partition the data set into clusters in such a way that data points in each cluster are highly similar to each other, while data points assigned to different clusters have low degrees of similarity.

Fuzzy C-means (FCM) allows one data point to belong to two or more clusters [14], [17]. It provides a method that groups data points in multidimensional space into a specific number of clusters. It is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} \mu_{ij}^m \|x_i - b_j\| \tag{14}$$

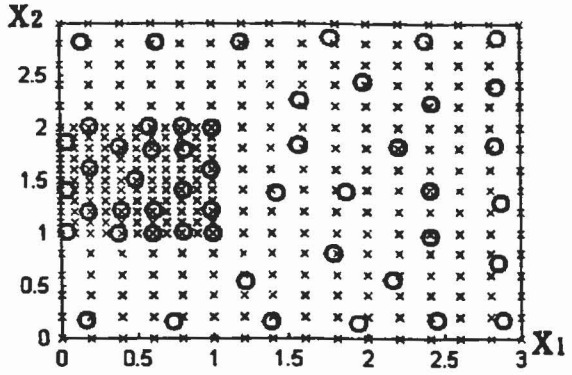

Fig. 4. Clusters generated using FCM. "X" is the original data point; "O" is the fuzzy cluster center.

where m is the number of clusters, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the ith of n-dimensional measured data, b_j is the n-dimensional center of the cluster, and $\|*\|$ is a norm expressing the distance between measured data and cluster center.

FCM is used to cluster the data shown in Fig. 2(B). The number of clusters is predefined as 50. Then the produced 50 centers of the clusters are used to form a new data set. After this process, the distribution of the obtained new data set is shown in Fig. 4. The crosses are the original data points and the circles are the centers of the clusters.

Apparently, the distribution of the data did not change. Why can FCM clustering not change the distribution? It is known that during the iteration process of FCM, the membership is updated though

$$\mu_{ij} = \left[\sum_{k=1}^{C}\left(\frac{\|x_i - c_j)\|}{\|x_i - c_k)\|}\right)^{\frac{2}{m-1}}\right]^{-1} \tag{15}$$

and the cluster centers are updated though

$$b_j = \frac{\sum_{i=1}^{N} x_i \cdot \mu_{ij}^m}{\sum_{i=1}^{N} \mu_{ij}^m}. \tag{16}$$

The membership and the cluster centers are updated to minimize the total weighted distance between data points and the cluster centers of the fuzzy partition. Thus it is reasonable that in areas where more data points exist, there must be more cluster centers in order to make the total weighted distance between all data points and the cluster centers smaller. Therefore the data distribution cannot be changed only using FCM.

2) Generate Even Cluster Distribution: To generate better cluster distribution, the input space is partitioned using fuzzy rules before applying FCM. A fuzzy rule base is built for the nonlinear system. Those data points that can excite a particular fuzzy rule with high firing strength are grouped to the same partition. The fuzzy rule base is in the following form.

IF x_1 is A_{11} and x_2 is A_{21} and ... and x_n is A_{n1} THEN y is y^1.

IF x_1 is A_{12} and x_2 is A_{22} and ... and x_n is A_{n2} THEN y is y^2.

...

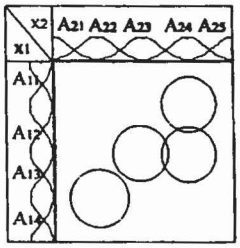

Fig. 5. Partitions generated by fuzzy rules (with fixed width σ').

IF x_1 is A_{1m} and x_2 is A_{2m} and ... and x_n is A_{nm} THEN y is y^m

where $A_{ij}(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m)$ and y^j are fuzzy sets of x_i and y, respectively.

The fuzzy partitions generated by the fuzzy rules are shown in Fig. 5. If the width σ' of Gaussian membership function is the same for all the fuzzy sets, the partition is an even partition.

The method is implemented as follows: the first sampling data point is taken as the center of a cluster, and a corresponding fuzzy rule is built. The parameters are chosen as follows:

$$\bar{x}_i^j = x_i^j \tag{17}$$

$$\sigma' = \frac{1}{30} \times d \tag{18}$$

where d is the range of the input variables (normalized to the same range).

For every sampling data point, the firing strength (degree of fulfillment) of each existing rule is calculated

$$G_j = \prod_{i=1}^{n} (\mu_{ij}(x_i)) = \prod_{i=1}^{n} \exp\left(- \left(\frac{x_i - \bar{x}_i^j}{\sigma'} \right)^2 \right). \tag{19}$$

AND operation is used in (19).

If the firing strength

$$G_j \geq \beta \tag{20}$$

then the sampling data point is close to the data points in the partition. Thus it belongs to this partition. β is a predefined threshold as the least acceptable degree, and it determines the extent of the similarity to be classified into the partition. If the firing strength is less than the threshold β, then a new fuzzy rule (a new partition) should be created.

After all the data are partitioned, FCM algorithm is used to cluster data points in each small partition. The same number of clusters is set for each small partition. Or, if the partition is small enough, only one cluster is set for each partition, and its center is found by FCM. It would be nice to use the centers of the clusters to represent the clusters. But for real-world systems, the corresponding output of the system to the centers are not available, if the centers are not coincident to the existing data points. So the closest sampling data point to the center of a cluster is used to represent the cluster. The closest data point is decided by its Euclidean distance to the center

$$\mathbf{x}_{\text{closest}} = \|\mathbf{x}_i - \mathbf{x}_{\text{center}}\|_{\min}. \tag{21}$$

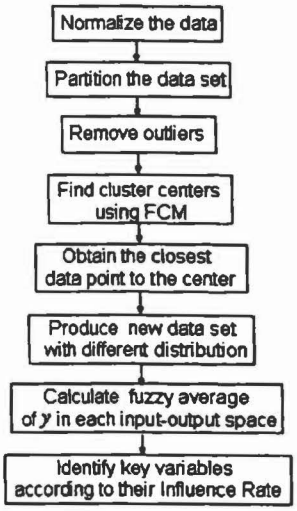

Fig. 6. Procedure of FAFCD.

There is a loss of information during this process, but the number of partitions can be controlled to make sure only redundant data points are removed while keeping enough data points to represent all the sampling data in the input space. This is done by adjusting σ'. If $\sigma' \to 0$, then each sampling data point is a partition; if $\sigma' \to \infty$, only one partition obtained.

When only one cluster is set for each partition, the number of clusters is the same as the number of partitions. So if σ' is too small, the distribution may not change and the number of clusters will be large; if σ' is too large, some clusters may be combined into other clusters and their representation is lost when only the cluster center is kept. σ' is taken as 1/30 of the range of the normalized input variables. This can partition the input space into many small partitions, which can represent the input space adequately, and at the same time the redundant sampling data points in each partition are removed.

The procedure of FAFCD is shown in Fig. 6.

Distribution of the data in Fig. 2(B) becomes better after being processed by the above method [see Fig. 7(B)]. Hence the fuzzy average of y in $x_2 - y$ space can reflect the $x_2 - y$ relationship correctly now [Fig. 7(A)]. These results are very similar to those generated in Fig. 1, which is a uniform distribution.

C. Other Considerations

Some considerations and improvements should be implemented.

1) Data Normalization: Before applying FAFCD, the data set should be preprocessed. Since different variables are measured in different units and with different numerical ranges, a bias may be introduced to the process. Thus the data need to be normalized

$$\tilde{x}_i^j = \frac{x_i^j - \left(x_i^j\right)_{\min}}{\left(x_i^j\right)_{\max} - \left(x_i^j\right)_{\min}} \tag{22}$$

where $\tilde{x}_i^j$ is the value after normalization; x_i^j is the data to be normalized; $(x_i^j)_{\min}$ is the minimum value of vector x_i; and $(x_i^j)_{\max}$ is the maximum value of vector x_i.

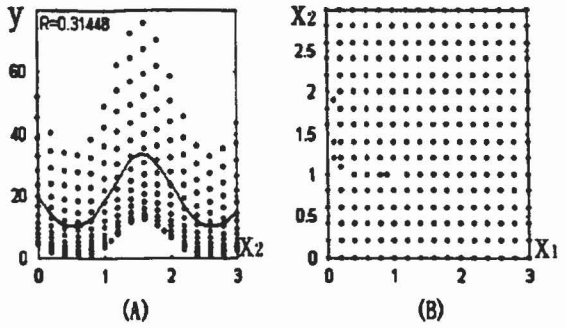

Fig. 7. Results using FAFCD. (A) The $x_2 - y$ relationship from defuzzification. (B) The distribution of inputs.

2) Calculate Less Membership Functions: When using the fuzzy average method, all existing membership functions μ_{ij} for each value of x_i in (4) are calculated. If the center of a membership function ($\overline{x}_i^j$) is far away from the current value of x_i, which means $x_i - \overline{x}_i^j$ is big, the value of the membership function becomes very small and has not much influence on the fuzzy average value. For instance, when $x_i - \overline{x}_i^j = 3\sigma$, the value of the membership function is very small

$$\mu_{ij}(x_i) = \exp\left(-\left(\frac{3\sigma}{\sigma}\right)^2\right) = \exp(-9) = 1.23e-4 \tag{23}$$

Therefore only the membership functions whose centers are in the range of 3σ need to be calculated. The values of the other membership functions are very small and can be neglected. This can reduce the number of membership functions to be calculated. For the example used in the simulation section, there are 29 880 data points, which means for each value of x_i, 29 880 membership functions need to be calculated. If only the membership functions whose centers are in the range of 3σ are calculated, this number can be reduced to 1/10 of the total number of membership functions. So the fuzzy average can be obtained much faster with almost no quality decrease.

3) Remove Outliers: During our clustering process, outliers often become individual clusters. When the cluster center is used to form a new data set, the outliers are kept. If the sampling data set has a large number of data points, those clusters containing very few data points (predefine a threshold) can be removed. Outliers normally can be removed by doing so. The threshold should be determined according to the total number of data points.

III. Simulations and Results

The performance of FAFCD is evaluated using two example data sets. In the first experiment, the algorithm is evaluated using the well-known example of gas furnace data set given by Box and Jenkins. In the second experiment, it is evaluated using the automobile fuel consumption prediction problem. Finally the method is applied to the EMG signal estimation system.

A. Evaluate FAFCD on the Gas Furnace Data Set

The method is applied to the well-known Box and Jenkins gas furnace data from [15]. This data set was also used in [3], [4], and [16] to identify the key variables. This time series data set has the gas flow rate $u(t)$ as input and CO_2 concentration $y(t)$ as output. There are 296 such input–output data pairs. As in [3] and [4], a dynamic process model is extracted from the data by taking ten variables $y(t-1), y(t-2), \ldots, y(t-4), u(t-1), u(t-2), \ldots, u(t-6)$ as input and $y(t)$ as output. Using $\mathbf{x}(t)$ to denote the ten-variable input vector, 290 $\mathbf{x}(t) - y(t)$ data pairs obtained. FAFCD is applied to this data set to find significance of input variables.

Fig. 8 shows the relationship between some input variables $(y(t-1), y(t-4), u(t-1), u(t-4))$ and the output. The inputs are normalized to (0,1). Table I shows the influence rate (defined earlier) of all the input variables.

According to Table I, the top three input variables are determined as $y(t-1), u(t-4)$, and $u(t-3)$. This result is similar to [3]. In [3], different models were generated while searching for the optimal combination of variables. There were 24 models generated before the top three input variables were identified. If the importance of all the ten input variables needs to be ranked, the number of models needed will be even larger. The method used in [4] is also complicated and time consuming. To determine the important variables, premises in the fuzzy rules of an initial model are systematically removed to search for the best simplified model. After the searching process, different importance of top three input variables was obtained using three different criteria for model selection. Then for each solution, a Takagi–Sugeno type model was generated. The final conclusion was made based on the comparison of the model errors. Compared to these methods, the FAFCD approach is more straightforward and computationally efficient.

B. Evaluate FAFCD on Automobile Fuel Consumption Prediction Problem

The automobile fuel consumption prediction problem uses variables such as model year and horsepower to predict miles per gallon (MPG) of the automobiles. FAFCD is used to find the significance of those input variables to the MPG of the automobile.

The data set used here is from the UCI Machine Learning Repository. It consists of 392 complete instances without missing values. Each instance is composed of six input variables (number of cylinders, displacement, horsepower, weight, acceleration, and model year) and one output variable (MPG).

The input variables and output variable are denoted as $x_i (i = 1, 2, \ldots, 6)$ and y, respectively. Using FAFCD, the $x_i - y$ relationship can be obtained for each input variable. Fig. 9 shows the relationship between two input variables (x_4, x_5) and MPG. From the result, it can be found how each input variable affects the output.

The influence rate of each input variable is obtained using FAFCD. Table II shows the results for the six input variables. From the influence rate of each input variable, the conclusion is obtained that the variables weight and horsepower have the most significant influence on the MPG of an automobile.

If the distribution/density of the original data set is changed by repeating some of the data points (Table III), the influence rate of acceleration obtained without clustering (the method used in [5]) becomes 0.6076 [Fig. 10(A)], while the influence

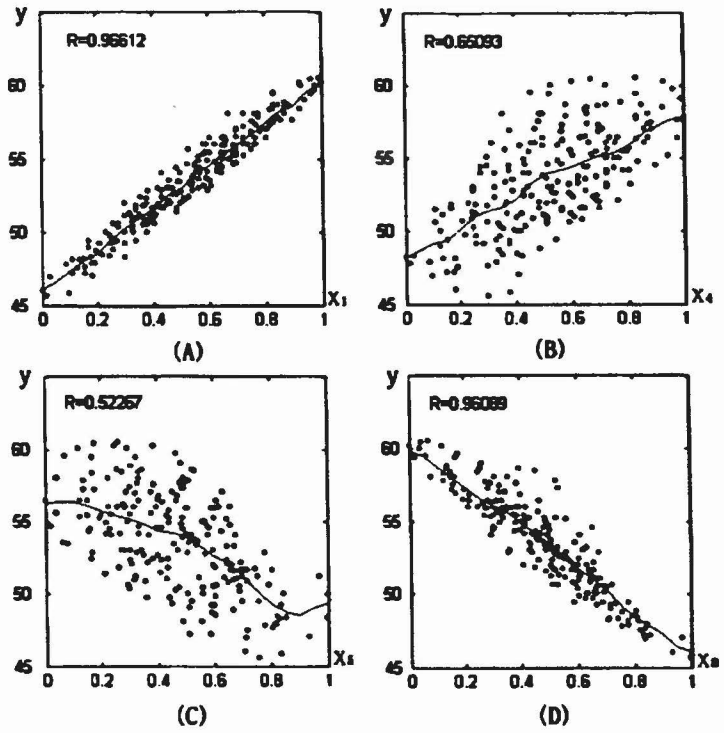

Fig. 8. Relationship between some input variables and the output obtained using FAFCD. x_1, x_4, x_5, and x_8 stand for $y(t-1), y(t-4), u(t-1)$, and $u(t-4)$, respectively. (A) $y(t-1) - y(t)$ relationship. (B) $y(t-4) - y(t)$ relationship. (C) $u(t-1) - y(t)$ relationship. (D) $u(t-4) - y(t)$ relationship.

TABLE I
INFLUENCE RATE OF INPUT VARIABLES FOR GAS FURNACE DATA

Input	Variable Name	Influence Rate
x_1	$y(t-1)$	0.9661
x_2	$y(t-2)$	0.8965
x_3	$y(t-3)$	0.7445
x_4	$y(t-4)$	0.6509
x_5	$u(t-1)$	0.5227
x_6	$u(t-2)$	0.6745
x_7	$u(t-3)$	0.9491
x_8	$u(t-4)$	0.9609
x_9	$u(t-5)$	0.9317
x_{10}	$u(t-6)$	0.9117

rates of other variables remain almost unchanged. Then the conclusion becomes that acceleration is the most important variable, which is incorrect. If the FAFCD method is used, it can give an influence rate of acceleration of 0.5612 [Fig. 10(B)]. Thus the conclusion of significance of variables will not be affected. That is because in FAFCD, the distribution of fuzzy clusters is used instead of using the distribution of the original data set.

C. Evaluate FAFCD on EMG Signal Estimation System

EMG signal estimation system is a model in which kinematic parameters during a motion (kinematics variables) and anthropometric characteristics of subjects (subject variables) are used to estimate the corresponding EMG signals in ten trunk muscles generated during the manual lifting motion of the subjects. In this system, which variables affect the EMG signals is unknown because the muscle activities are not completely understood and are still under study. Our objective is to find out those variables that have significant influence on EMG signals.

1) Input Variables: Not knowing which variables affect the EMG signals, all the associated kinematic variables and subject variables are recorded. Tables IV and V show all the kinematic variables and subject variables, respectively. The 12 kinematic variables are dynamic variables which change their values during the motion. The 15 subject variables are static variables which are the anthropometric characteristics of the subjects, and they are the same during a motion for a particular subject. The sampling data set contains six trials of motions conducted by 249 subjects. Each trial has 20 sampling data points. Every subject conducted all the trials. Therefore the total number of data points N is

$$N = n_{\text{subjects}} \times n_{\text{trials}} \times 20 = 29\,880 \tag{24}$$

where n_{subjects} is the number of subjects and n_{trials} is the number of trials. Each data point consist of 27 input variables $(x_i, i = 1, 2, \ldots, 27)$ and ten output variables $(y_j, j = 1, 2, \ldots, 10)$. The output variables are EMG signals of ten trunk muscles.

2) Results: To calculate the $x_i - y_j$ relationships and identify key variables of the system, the distribution of the original data set without clustering is first used. The results obtained are not satisfactory. For some of the input–output relationships, there is a drop in the middle of the range of variable x_i. An example is shown in Fig. 11(A). This relationship is surely incorrect for

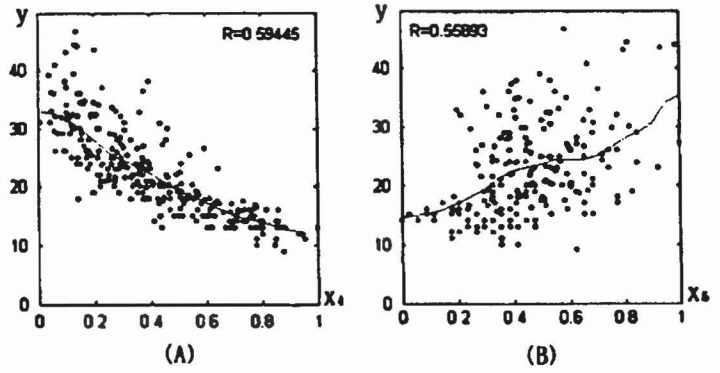

Fig. 9. Relationship between two input variables and the output obtained using FAFCD. (A) Weight (x_4)-MPG (y) relationship. (B) Acceleration (x_5)-MPG (y) relationship.

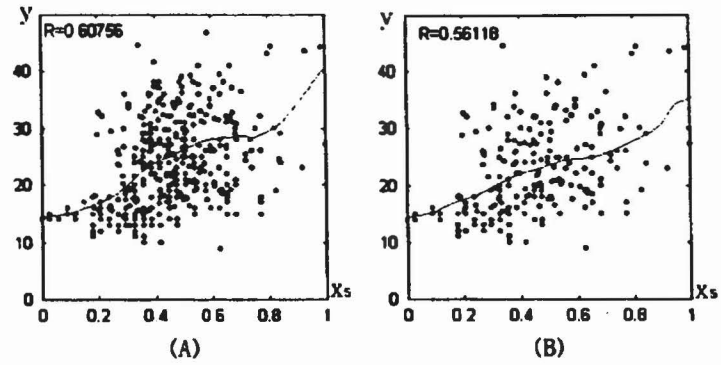

Fig. 10. Results on the modified MPG data using different methods. (A) Without clustering (method used in [5]). (B) Using FAFCD.

TABLE II
INFLUENCE RATE OF INPUT VARIABLES FOR THE FUEL CONSUMPTION

Input	Variable Name	Influence Rate
x_1	Number of cylinders	0.3852
x_2	Displacement	0.4488
x_3	Horsepower	0.5851
x_4	Weight	0.5945
x_5	Acceleration	0.5589
x_6	Model year	0.4239

TABLE III
DATA POINTS REPEATED TWICE

MPG	Cylind.	Displ.	Horsepower	Weight	Accel.	year
35	4	72	69	1613	18	71
33	4	91	53	1795	17.5	75
33	4	91	53	1795	17.4	76
36	4	79	58	1825	18.6	77
39.4	4	85	70	2070	18.6	78
36.1	4	91	60	1800	16.4	78
31.8	4	85	65	2020	19.2	79
40.8	4	85	65	2110	19.2	80
44.3	4	90	48	2085	21.7	80
43.4	4	90	48	2335	23.7	80
36.4	5	121	67	2950	19.9	80
40.9	4	85	77	1835	17.3	80
33.8	4	97	67	2145	18	80
37	4	85	65	1975	19.4	81
38	6	262	85	3015	17	82
37	4	91	68	2025	18.2	82

some variables and is uninterpretable from the ergonomics point of view. Sometimes when different portions of the data set were used, different relationships were obtained. Therefore the distribution of the sampling data affected the result. Certain conditions may have appeared more frequently during the motion than other conditions and thus have distorted the fuzzy average curve.

TABLE IV
KINEMATIC VARIABLES (DYNAMIC)

Sagittal Trunk Moment	Sagittal Trunk Angle
Sagittal Trunk Velocity	Sagittal Trunk Acceleration
Lateral Trunk Moment	Lateral Trunk Angle
Lateral Trunk Velocity	Lateral Trunk Acceleration
Axis Trunk Moment	Axis Trunk Angle
Axis Trunk Velocity	Axis Trunk Acceleration

TABLE V
SUBJECT VARIABLES (STATIC)

Age	Upper Leg Length	Trunk Depth (pelvis)
Body Weight	Upper Arm Length	Trunk Breadth (pelvis)
Standing Height	Lower Leg Length	Trunk Depth (xyphoid)
Shoulder Height	Lower Lrm Length	Trunk Breadth (xyphoid)
Elbow Height	Spine Length	Trunk Circumference

Then FAFCD is used to obtain the input–output relationships on the new data set produced using fuzzy clustering. Procedures in the flowchart as shown in Fig. 6 were followed. All input variables in the original data set are normalized to the range of [0,1]. The output variables (the normalized EMG signals) are

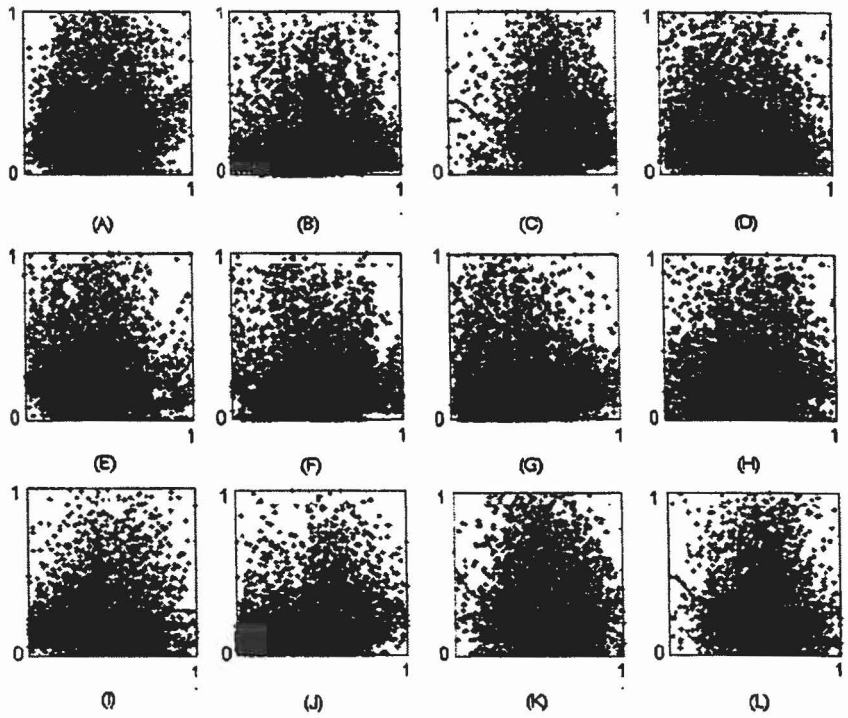

Fig. 12. Relationship between the kinematic variables (x axis) and EMG signals (y axis) of the muscle right latissimus dorsi (RLD). Variables from (A) to (L) are in the same sequence as in Table VII.

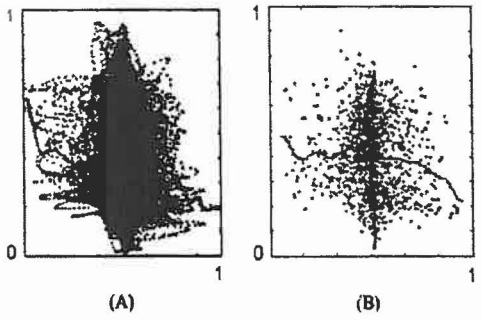

Fig. 11. The relationship between axis trunk velocity (x axis) and EMG signals (y axis). (A) Without clustering and (B) after clustering.

TABLE VI
SUMMARY ABOUT THE CLUSTERS

Total number of data points	29880
Number of clusters	4816
Range of number of data points in the clusters	1 - 1456
Number of clusters with less than 10 data points	66

also in the range of [0,1]. Then the data points were clustered as described in Section II. A summary of cluster properties is shown in Table VI. Those clusters containing fewer than ten data points (about 0.033% of the total) are considered as outliers and are removed from the data set.

Then FCM algorithm was used to find centers of the rest clusters. By comparing the Euclidean distance of each data point in a cluster to the center of this cluster, the closest data point to the center is found. Using this data point to represent the corresponding cluster, a new data set with a different distribution to the original data set was obtained. On this new data set, the fuzzy average of y_j in each $x_i - y_j$ space was calculated. Fig. 11(B) shows the result of the same example as in Fig. 11(A), using FAFCD. As expected, the drop in Fig. 11(A) disappeared and the result has a clear physical explanation now. Fig. 12 shows the relationship between all kinematic variables and EMG signals of muscle right latissimus dorsi. Fig. 13 shows the relationship between all subject variables and EMG signals of this muscle. The relationships of inputs to the other muscles can be obtained similarly. With these relationships, a better understanding to the muscle activities can be gained. At the same time, the importance of the input variables is indicated by their influence rate R.

Table VII shows the influence rate of each kinematic variable to each output. The first row is the name of the muscle; the first column is the name of the kinematic variable. Table VIII shows the influence rate of each subject variable to each output. The first row is the name of the muscle; the first column is the name of the subject variable. Based on the influence rate, key variables can be identified.

According to Tables VII and VIII, significance of the kinematic variables and subject variables are ranked as shown in Figs. 14 and 15, respectively. It is clear that kinematic variables are more significant than subject variables. Thus, these 12 kinematic variables should all be selected as inputs in modelling. This is a reasonable conclusion and agrees with our hypothesis. As for subject variables, four variables (standing height, shoulder height, lower arm length, and spine length) are more significant than the others. These variables should also be taken as inputs in modelling. However, by examining the two variables "standing height" and "shoulder height" in Table VIII, it

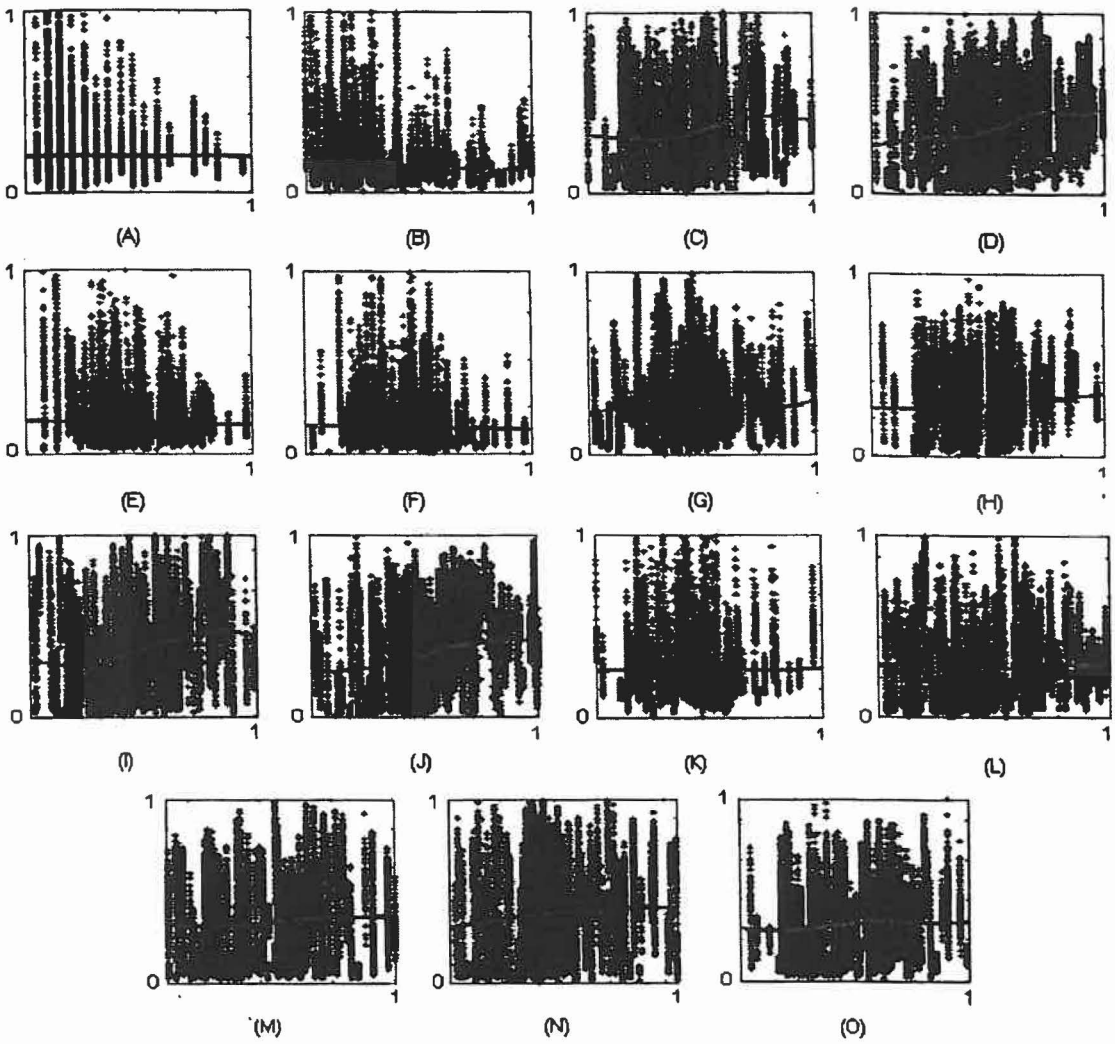

Fig. 13. Relationship between the subject variables (x axis) and EMG signals (y axis) of the muscle right latissimus dorsi. Variables from (A) to (O) are in the same sequence as in Table VIII.

is found that the influence rates of these two variables are very similar, for every muscle. In other words, these two variables are correlated. They are dependent variables to each other. Therefore, one of them can be removed.

3) Validation: As stated before, the result of FAFCD shows that kinematic variables are more significant than subject variables. This is a reasonable conclusion from the ergonomics point of view since kinematic variables are directly related to the muscular activities. For kinematic variables, it is certain that the variable "sagittal trunk moment" is the most significant variable in the EMG estimation system for manual lifting tasks. Result in Fig. 14 shows that FAFCD has correctly identified significance of this variable. For subject variables, four of them (standing height, shoulder height, lower arm length, and spine length) have been identified by FAFCD as the most significant subject variables. Coincidentally, these four variables apparently affect the moment arm during manual lifting tasks that involve a bending motion.

To further validate the results given by FAFCD, two feedforward neural network models for EMG estimation were constructed. The "basic model" with one hidden layer described in [18] was used. In this traditional neural network structure, every input variable has an equivalent position. In Model I, all 12 kinematic variables and 15 subject variables are used as input; in Model II, only variables identified as significant variables are used, including all kinematic variables and three subject variables (shoulder height, lower arm length, and spine length). As stated before, standing height is correlated with shoulder height, so it is removed from Model II. The input dimension is decreased from 27 to 15. EMG signals of ten trunk muscles are output variables of the neural networks.

Model performance was tested based on 1494 lifting trials (the same data set used for FAFCD). The conservative cross-validation method was employed. Every time 1/4 of the data is taken out for test and the other 3/4 used to train the model. A different 1/4 is taken for test and the rest used for training the next time. Continue this process until all the data are evaluated for test. Simulation results showed that removing 12 less significant subject variables did not notably decrease the performance of the EMG estimation model. Mean absolute error of all mus-

TABLE VII
INFLUENCE RATE OF KINEMATIC VARIABLES (SAG. = SAGITTAL, LAT. = LATERAL, MOM. = MOMENT, ANG. = ANGLE, VEL. = VELOCITY, AND ACC. = ACCELERATION)

	RLD	LLD	RES	LES	RRA	LRA	REO	LEO	RIO	LIO	Average
Sag. trunk mom.	0.208	0.220	0.182	0.149	0.049	0.087	0.177	0.223	0.260	0.133	0.1688
Lat. trunk mom.	0.046	0.113	0.332	0.293	0.129	0.110	0.054	0.101	0.093	0.207	0.1478
Axis trunk mom.	0.136	0.209	0.169	0.220	0.081	0.089	0.037	0.062	0.015	0.089	0.1107
Sag. trunk ang.	0.116	0.100	0.149	0.121	0.057	0.058	0.073	0.100	0.187	0.143	0.1104
Lat. trunk ang.	0.135	0.094	0.106	0.109	0.042	0.061	0.065	0.047	0.088	0.139	0.0886
Axis trunk ang.	0.124	0.079	0.172	0.310	0.135	0.141	0.066	0.068	0.122	0.165	0.1382
Sag. trunk vel.	0.055	0.069	0.199	0.196	0.031	0.114	0.106	0.090	0.124	0.225	0.1209
Lat. trunk vel.	0.045	0.042	0.172	0.094	0.050	0.162	0.079	0.061	0.131	0.145	0.0981
Axis trunk vel.	0.072	0.074	0.090	0.087	0.050	0.207	0.029	0.090	0.181	0.081	0.0961
Sag. trunk acc.	0.139	0.095	0.118	0.167	0.071	0.046	0.095	0.053	0.142	0.104	0.1030
Lat. trunk acc.	0.160	0.088	0.100	0.079	0.077	0.087	0.028	0.070	0.104	0.077	0.0870
Axis trunk acc.	0.168	0.080	0.179	0.139	0.059	0.078	0.079	0.126	0.116	0.070	0.1094

TABLE VIII
INFLUENCE RATE OF SUBJECT VARIABLES (LEN. = LENGTH, DEP. = DEPTH, PEL. = PELVIS, BR. = BREADTH, XY. = XYPHOID, AND CIR. = CIRCUMFERENCE)

	RLD	LLD	RES	LES	RRA	LRA	REO	LEO	RIO	LIO	Average
Age	0.002	0.001	0.005	0.005	0.004	0.001	0.002	0.002	0.001	0.003	0.0026
Body weight	0.088	0.066	0.071	0.103	0.013	0.026	0.030	0.029	0.015	0.052	0.0493
Standing height	0.118	0.074	0.120	0.158	0.048	0.097	0.045	0.056	0.018	0.087	0.0821
Shoulder height	0.120	0.080	0.118	0.155	0.038	0.099	0.049	0.059	0.021	0.081	0.0820
Elbow height	0.043	0.013	0.031	0.057	0.016	0.016	0.025	0.005	0.022	0.027	0.0255
Upper leg len.	0.032	0.021	0.013	0.003	0.005	0.009	0.013	0.023	0.003	0.008	0.0130
Lower leg len.	0.099	0.060	0.086	0.123	0.036	0.055	0.043	0.036	0.016	0.086	0.0640
Upper arm len.	0.078	0.020	0.030	0.053	0.008	0.025	0.024	0.019	0.006	0.029	0.0292
Lower arm len.	0.090	0.081	0.151	0.164	0.019	0.062	0.039	0.034	0.033	0.112	0.0785
Spine len.	0.082	0.055	0.117	0.125	0.067	0.087	0.035	0.048	0.066	0.077	0.0759
Trunk dep. pel.	0.009	0.007	0.022	0.009	0.007	0.016	0.018	0.012	0.010	0.008	0.0118
Trunk br. pel.	0.055	0.051	0.022	0.030	0.045	0.018	0.033	0.004	0.036	0.008	0.0302
Trunk dep. xy.	0.038	0.044	0.027	0.057	0.040	0.020	0.020	0.017	0.013	0.035	0.0311
Trunk br.xy.	0.080	0.056	0.066	0.093	0.032	0.032	0.018	0.021	0.027	0.063	0.0488
Trunk cir.	0.047	0.040	0.032	0.052	0.014	0.058	0.055	0.027	0.037	0.031	0.0393

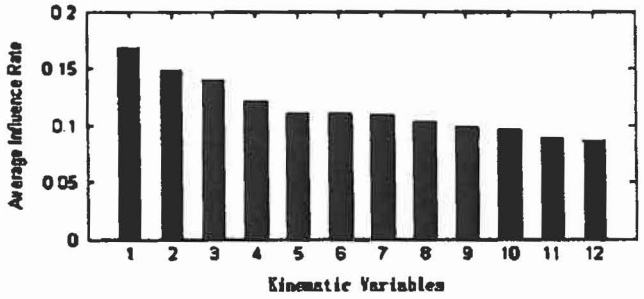

Fig. 14. Rank kinematic variables by their average influence rate (1: sagittal trunk moment, 2: lateral trunk moment 3: axis trunk angle 4: sagittal trunk velocity, 5: axis trunk moment, 6: sagittal trunk angle, 7: axis trunk acceleration, 8: sagittal trunk acceleration, 9: lateral trunk velocity, 10: axis trunk velocity, 11: lateral trunk angle, 12: lateral trunk acceleration).

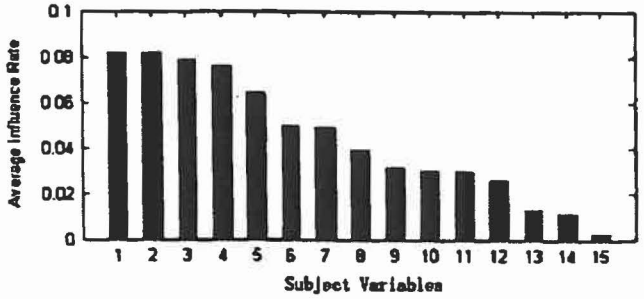

Fig. 15. Rank subject variables by their average influence rate (1: standing height, 2: shoulder height, 3: lower arm length, 4:spine length, 5: lower leg length, 6: body weight, 7: trunk breadth (xyphoid), 8: trunk circumference, 9: trunk depth (xyphoid), 10: trunk breadth (pelvis), 11: upper arm length, 12: elbow height, 13: upper leg length, 14: trunk depth (pelvis), 15: age).

cles for Model I and Model II are 7.47% and 7.52%, respectively.

IV. CONCLUSION

FAFCD can find out the significance of specific input variables and how they influence the output, without the interference of the distribution of sampling data set. The method is straightforward and easy to implement. Knowing the significance of candidate input variables, complexity of the model and time of modelling may be greatly reduced.

REFERENCES

[1] E. B. Bartlett, "Self determination of input variable importance by neural networks," *Neural, Parallel Sci. Comp.*, vol. 2, no. 1, pp. 103–114, 1994.

[2] B. Yuan and G. Klir, "Data-driven identification of key variables," in *Intelligent Hybrid Systems—Fuzzy Logic, Neural Networks, and Genetic Algorithms*, D. Ruan, Ed. Boston, MA: Kluwer Academic, 1997, pp. 161–186.
[3] M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modeling," *IEEE Trans. Fuzzy Syst.*, vol. 1, pp. 7–31, Feb. 1993.
[4] S. L. Chiu, "Selecting input variables for fuzzy models," *J. Intell. Fuzzy Syst.*, vol. 4, pp. 243–256, 1996.
[5] Y. Lin and G. A. Cunningham, "A new approach to fuzzy-neural system modeling," *IEEE Trans. Fuzzy Syst.*, vol. 3, pp. 190–198, May 1995.
[6] Y. Lin, G. A. Cunningham, S. V. Coggeshall, and R. D. Jones, "Nonlinear system input structure identification: Two stage fuzzy curves and surfaces," *IEEE Trans. Syst., Man, Cybern. A*, vol. 28, pp. 678–684, Sep. 1998.
[7] M. F. Azeem, M. Hanmandlu, and N. Ahmad, "Structure identification of generalized adaptive neuro-fuzzy inference systems," *IEEE Trans. Fuzzy Syst.*, vol. 11, pp. 666–681, Oct. 2003.
[8] E. C. Morabito and M. Versaci, "A fuzzy neural approach to localizing holes in conducting plates," *IEEE Trans. Magn.*, vol. 37, no. 5, pp. 3534–3537, Sep. 2001.
[9] S. E. Papadakis, J. B. Theocharis, J. Kiartzis, and A. G. Bakirtzis, "A novel approach to short-term load forecasting using fuzzy neural networks," *IEEE Trans. Power Syst.*, vol. 13, pp. 480–492, May 1998.
[10] A. H. Sung, "Ranking input importance in neural network modeling of engineering problems," in *Proc. IEEE World Congr. Comp. Intell.*, May 1998, vol. 1, pp. 316–321.
[11] B. Bouqata, A. Bensaid, R. Palliam, and A. F. Gomez Skarmeta, "Time series prediction using crisp and fuzzy neural networks: A comparative study," in *Proc. IEEE/IAFE/INFORMS 2000 Conf. Comp. Intell. Financial Eng.*, Mar. 26–28, 2000, pp. 170–173.
[12] M. F. Azeem, M. Hanmandlu, and N. Ahmad, "A new criteria for input variable identification of dynamical systems," in *Proc. IEEE Region 10 Int. Conf. Global Connectivity Energy, Comput, Commun. Contr.*, Dec. 17–19, 1998, vol. 1, pp. 230–233.
[13] S. Papadakis and J. Theocharis, "An efficient fuzzy neural modeling approach using the fuzzy curve concept," in *Proc. 3rd IEEE Int. Conf. Electron., Circuits, Syst.*, Oct. 13–16, 1996, vol. 1, pp. 279–282.
[14] E. S. Schrich, J. Ke, L. O. Hall, and D. B. Goldgof, "Fast accurate fuzzy clustering through data reduction," *IEEE Trans. Fuzzy Syst.*, vol. 11, pp. 262–270, Apr. 2003.
[15] G. E. P. Box and G. M. Jenkins, *Time Series Analysis, Forecasting and Control*, 2nd ed. San Francisco, CA: Holden Day, 1976.
[16] A. E. Gaweda, J. M. Zurada, and R. Setiono, "Input selection in data-driven fuzzy modeling," in *Proc. 10th IEEE Int. Conf. Fuzzy Syst.*, Dec. 2–5, 2001, vol. 3, pp. 1251–1254.
[17] S. Auephanwiriyakul and J. M. Keller, "Analysis and efficient implementation of a linguistic fuzzy c-means," *IEEE Trans. Fuzzy Syst.*, vol. 10, pp. 563–582, Oct. 2002.
[18] Y. Hou, J. M. Zurada, and W. Karwowski, "Prediction of EMG signals of trunk muscles in manual lifting using a neural network model," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN 2004)*, Budapest, Hungary, Jul. 25–29, 2004, pp. 1935–1940.
[19] J. Zurada, W. Karwowski, and W. S. Marras, "A neural network-based system for classification of industrial jobs with respect to the risk of low back disorders," *Appl. Ergon.*, vol. 28, no. 1, pp. 49–58, 1997.
[20] Y. Hou, J. M. Zurada, W. Karwowski, and W. S. Maras, "A fuzzy approach for key variables identification of EMG evaluation system," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN 2005)*, Montreal, PQ, Canada, Jul. 31–Aug. 4 2005, pp. 2520–2525.
[21] W. Lee, W. Karwowski, W. S. Marras, and D. Rodrick, "A neuro-fuzzy model for estimating electromyographical activity of trunk muscles due to manual lifting," *Ergonomics*, vol. 46, no. 1–3, pp. 285–309, Jan. 15, 2003.

Yanfeng Hou (S'05) received the B.S. degree in electrical engineering from Sichuan University, Sichuan, Chinam and the M.S. degree in electrical engineering from Wuhan University, Hubei, China. He is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Louisville, KY.

His research interests are in the areas of artificial neural networks, fuzzy logic, data mining, and other artificial intelligence and machine learning-based methodologies with focus on ergonomics and biomedical applications.

Mr. Hou is a member of Tau Beta Pi and the IEEE CIS Standards Committee.

Jacek M. Zurada (M'82–SM'83–F'96) is the Samuel T. Fife Alumni Professor and Chair of Electrical and Computer Engineering Department at the University of Louisville, Louisville, KY. He was a Visiting Professor at Princeton University, National University of Singapore, Nanyang Technological University, and the University of Karlsruhe, Germany. He is the author or coauthor of more than 250 journal and conference papers in the area of neural networks, computational intelligence, data mining, image processing, and VLSI circuits. He was coeditor of *Knowledge-Based Neurocomputing* (Cambridge, MA: MIT Press, 2000), author of the introduction to *Artificial Neural Systems* (PWS, 1992), contributor to *Progress in Neural Networks* (Ablex, 1994/1995), and coeditor of *Computational Intelligence: Imitating Life* (New York: IEEE Press, 1994). He is an Associate Editor of *Neurocomputing*. He has delivered numerous invited plenary conference presentations and seminars throughout the world.

Dr. Zurada was an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART I: FUNDAMENTAL THEORY AND APPLICATIONS and PART II: ANALOG AND DIGITAL SIGNAL PROCESSING. In 2001–2003, he was a member of the Editorial Board of the PROCEEDINGS OF IEEE. In 1998–2003, he was Editor-in-Chief of IEEE TRANSACTIONS ON NEURAL NETWORKS. He has received a number of awards for distinction in research and teaching, including the 1993 Presidential Award for Research, Scholarship and Creative Activity. In 2001, he received the University of Louisville President's Distinguished Service Award for Service to the profession. In 2003, he received the title of Professor by the President of Poland, Aleksander Kwasniewski and the Honorary Professorship of Hebei University in China. Since 2005, he has served as Foreign Member of the Polish Academy of Sciences. He is Past President of the IEEE Computational Intelligence Society. He is a Distinguished Speaker of IEEE CIS.

Waldemar Karwowski received the M.S. degree in production engineering and management from the Technical University of Wroclaw, Poland, in 1978, the Ph.D. degree in industrial engineering from Texas Tech University, Lubbock, in 1982, and the D.Sc. (Dr.Hab.) degree in management science from the Institute for Organization and Management in Industry, Warsaw, Poland, in 2004.

He is the J. B. Speed School of Engineering Alumni Scholar for Research and a Professor of Industrial Engineering and Director of the Center for Industrial Ergonomics, University of Louisville, Louisville, KY. His research, teaching, and consulting activities focus on work system design, organizational and management ergonomics, human-system integration and safety of advanced manufacturing, and neurofuzzy modeling and fuzzy systems applications.

Dr. Karwowski is a Fellow of the International Ergonomics Association, Human Factors and Ergonomics Society, Institute of Industrial Engineers, and The Ergonomics Society. He received the Doctor of Science Honoris Causa degree from the South Ukrainian State K. D. Ushynsky Pedagogical University of Odessa, Ukraine, in 2004 and the Technical Unviersity of Kosice, Slovakia, in 2006.

William S. Marras holds the Honda Endowed Chair in the Department of Industrial and Systems Engineering, The Ohio State University, Columbus. He is Director of the Biodynamics Laboratory and has joint appointments in the Departments of Orthopaedic Surgery, Physical Medicine, and Biomedical Engineering. His research is centered on occupational biomechanics issues. He is a pioneer in how motion influences the risk of musculoskeletal disorders in the workplace. His research includes workplace biomechanical epidemiologic studies, laboratory biomechanics studies, mathematical modeling, and clinical studies of the low back and upper extremity.

Kermit Davis received the doctorate degree from The Ohio State University, Columbus, in 2001.

He was trained in industrial engineering with specialization in occupational ergonomics and low back biomechanics with special interest in multiple exposures, both physical and psychosocial stressors. He is with the Department of Environmental Health, University of Cincinnati, Cincinnati, OH. His current research is focusing on the effect of physical workplace demands as well as mental workload on the responses within the lower back. He has published numerous articles about the impact of workplace stressors on the lower back including studies evaluating warehousing, patient handling, alternative modes of handling (e.g., team lifting, one hand lifting, pushing/pulling), injured populations, and ergonomic interventions (e.g., back belts, lifting hoists, adjustable fork lifts).

Dr. Davis has received several major research awards including the Alphonse Chapanis Student Paper Award from the human Factors and Ergonomics Society, Volvo Award For Low Back Pain Research in Biomechanical Studies from the International Society for the Study of the Lumbar Spine, Alice Hamilton Award from the National Institute for Occupational Safety and Health, and Promising Young Scientist Award from the International Society of Biomechanics.