Constructing a Takagi-Sugeno Fuzzy Model by a Fuzzy Data Shifter

Horng-Lin Shieh¹ and Ying-Kuei Yang²
¹ Dept. of Electrical Engineering, St. John's University, Taipei, Taiwan
² Dept. of Electrical Engineering, National Taiwan University of Science & Technology, Taipei, Taiwan

Abstract - This paper proposes a fuzzy-based data sifter (FDS) to locate good turning-points to partition a given nonlinear, multi-dimensional data domain into piecewise clusters so that a Takagi and Sugeno fuzzy model can be constructed with fewer rules and less error. One experiment is illustrated and its result has shown the proposed approach has better performance compared with other three related approaches.

Keywords: fuzzy-based data sifter, fuzzy cluster, TS fuzzy model, fuzzy matching degree, turning point.

1 Introduction

Constructing a fuzzy model can be basically categorized as Mamdani's model and Takagi and Sugeno's model (TS model). The Mamdani's model describes the relationship between inputs and outputs of a control system through a set of linguistic control rules and membership functions [1]-[8]. On the other hand, the TS model partitions an input space into several subspaces to describe either a static or a dynamic nonlinear system. One fuzzy rule is then created for each of these clusters. Below is a representation of rules in a TS model:

\[ R^i : \text{IF } x_1 \text{ is } A^i_1 \text{ and } x_2 \text{ is } A^i_2 \text{ and } \ldots \text{ and } x_k \text{ is } A^i_k \text{ THEN } y^i = a^i_0 + a^i_1 x_1 + a^i_2 x_2 + \ldots + a^i_k x_k \]

where \( R^i (i=1,2,\ldots,c) \) represents the \( i \)'th rule, \( A^i_j (j=1,2,\ldots,c, b=0,1,\ldots,k) \) is a constant called consequence parameter, \( y^i \) is the output of the \( i \)'th rule \( (i=1,2,\ldots,c) \), \( A^i_j (j=1,2,\ldots,c, j=0,1,\ldots,k) \) represents a linguistic term characterizing the membership function of the \( j \)'th input variable of the \( i \)'th rule and is called premise parameter.

In this paper, a fuzzy-based data sifter is proposed to better partition a nonlinear system's domain into several piecewise linear subspaces so that a TS model can be constructed with less fuzzy rules and less error for the domain.

2 The Proposed Fuzzy Data Sifter

In Equation (1), the premise parameters \( A^i_j \) are defined in this paper by a Two Sides Gaussian Function (TSGF) [9] that, as shown in Figure 1, indicates a membership function includes 4 parameters \( (\phi_i, \sigma_i, \phi_e, \sigma_e) \) where \( \phi_i, \sigma_i, i=1,2 \), respectively stand for the mean and deviation of the Two Side Gaussian Function.

For TS fuzzy modeling, the data domain is divided into subspaces and each of these subspaces is represented by a linear fuzzy model. The locations where the data domain is divided are named turning-points. Therefore, one of the critical issues on establishing a TS fuzzy model is to locate the turning-points existing in a data domain so that the TS fuzzy model can best represent the system model of the data domain with less error and simplest model structure. This paper proposes a slip-window-based recognition system, namely the fuzzy data sifter (FDS), to search the best turning-points of a nonlinear function. The obtained turning-points are then used to divide the given data domain into clusters. The piecewise linear regression algorithm (PLRA) is then applied to calculate the regression parameters for each of clusters.

The FDS is an iterated sifting system that uses a turning-point sift network (TPSN) to search out the local maximum and minimum data points in the data domain. For simplicity of explanation, the discussion is detailed in a 2-dimension figure as Figure 2. The turning-points can be basically classified into peak pattern in Figure 2(a) and valley pattern in Figure 2(b). The core idea of FDS is to use the 2 patterns as templates to search out all of the turning-points in a data domain. Because FDS is using fuzzy membership degree to indicate the bending level (angle) of a located turning-point, only the peak and valley patterns are used by FDS. The rest of not so
obvious (sharp) patterns can be represented by less membership degrees to indicate their turning angles are not as obvious and critical as peak and valley. The higher membership degree a turning-point has, the higher priority the turning-point has to be used as a dividing point in a data domain. Therefore, the use of only peak and valley patterns as matching templates does not affect the turning-points searching result of FDS. Figure 3 shows a peak and a valley patterns in 3-D.

The TPSN set is to perform the task of searching out all turning-points and generates their fuzzy matching membership degrees against turning-point patterns. A TPSN does the task for a particular turning-point pattern. There are two TPSNs as the peak and valley patterns are used in this paper. A TPSN takes the \( m \) data points in slip window, compares the shape of these data points with the turning-point pattern designated by the TPSN and measures the fuzzy matching degree between the shape and the pattern. The result of a TPSN is a fuzzy turning-point membership degree (FTPD) indicating the matching degree of the turning angle existing in the \( m \) data points covered in the slip window against the designated pattern. Because the length of whole data stream \( n \) is longer than the window width \( m \), values in a FTPD form a matrix. In Figure 4, the FTPD\(^1\) matrix holds the matching results against peak pattern performed by TPSN\(^1\). Similarly, the FTPD\(^2\) matrix holds the matching results against valley pattern performed by TPSN\(^2\). As stated previously, the turning-points with high fuzzy matching degrees will be selected with high priority as points from where to divide a data domain into clusters.

### 2.2 Measuring Fuzzy Matching Degrees Against Turning-Point Patterns

Hypothesize that a multiple-inputs-single-output (MISO) function has input variable \( x_i, i=1,2,…,k \) and output variable \( y \). For the convenience of discussion on how TPSN measures the fuzzy matching degree of a turning-point existing in slip window against a designated turning-point pattern, the multi-dimensional turning-point patterns are projected onto \( x_i-y \) planes. For example, the projection results of peak and valley patterns respectively are shown in Figure 5 and 6 in which the horizontal axis represents a certain input variable \( x_i \) and the vertical axis designates the output \( y \) of system function. The point \( p_s \) is where the center value of pattern bottom \( x_is \) corresponds to the maximum output \( y_{max} \) in Figure 5 and to the minimum output \( y_{min} \) in Figure 6. The \( u \) indicates the bottom width of the referenced pattern.
Assume the \( m \) data points in slip window in Figure 4 are \( \{(x_{i1},x_{j2},\ldots,x_{ij},y_{ij})|1\leq j \leq m\} \). Let \( X_i^{\min} = \min(x_{i1},x_{i2},\ldots,x_{im}) \) and \( X_i^{\max} = \max(x_{i1},x_{i2},\ldots,x_{im}) \) respectively indicate the minimum and maximum values of the \( i \)-th input variable for the \( m \) data points covered by the slip window. That is, the width of pattern bottom \( u = X_i^{\max} - X_i^{\min} \). Let \( y_{\min} = \min(y_1,y_2,\ldots,y_m) \) and \( y_{\max} = \max(y_1,y_2,\ldots,y_m) \) respectively represent the maximum and minimum of the output variable \( y \) for \( m \) data points. For the input value \( x_n \) of a data point in the slip window, the peak and valley patterns in Figures 5 and 6 can be defined as Equations (2) and (3) respectively.

Peak pattern:

\[
T_p(x_n) = \begin{cases} 
   y_{\min} + \frac{(y_{\max} - y_{\min})}{(x_{ij} - x_{i\text{min}})}(x_{ij} - x_{i\text{min}}), & \text{if } x_{ij} < x_n \\
   y_{\max} + \frac{(y_{\max} - y_{\min})}{(x_{ij} - x_{i\text{max}})}(x_{ij} - x_{i\text{max}}), & \text{if } x_{ij} > x_n 
\end{cases}
\]

Valley pattern:

\[
T_v(x_n) = \begin{cases} 
   y_{\min} + \frac{(y_{\max} - y_{\min})}{(x_{ij} - x_{i\text{max}})}(x_{ij} - x_{i\text{max}}), & \text{if } x_{ij} < x_n \\
   y_{\max} + \frac{(y_{\max} - y_{\min})}{(x_{ij} - x_{i\text{min}})}(x_{ij} - x_{i\text{min}}), & \text{if } x_{ij} > x_n 
\end{cases}
\]

That is, in a MISO system, the corresponding output value in a typical peak pattern for the \( h \)-th data point of \( x_i \) input variable is \( T_p(x_{ih}) \). Similarly, the corresponding output value in a typical valley pattern for the \( h \)-th data point of \( x_i \) input variable is \( T_v(x_{ih}) \).

For a data point \( p_h \), \( h=1,2,\ldots,n \), the slip window selects the \( m-1 \) points closest to \( p_h \) making totally \( m \) data points in the slip window. Naturally, the pattern formed by the \( m \) data points often does not conform to the definition of either peak pattern in Equation (2) or valley pattern in Equation (3). For example, for a point \( p_3 \), the slip window selects the closest points of \( p_1 \), \( p_2 \), \( p_4 \) and \( p_5 \) for TPSN to examine the turning degree at point \( p_3 \), as denoted in Figure 7(a).

These 5 points are projected onto \((x_1, y)\) plane, as shown in Figure 7(b) where circles denote the projected points of \( p_1 \sim p_5 \) in the slip window and solid lines denote the projected peak pattern in Figure 5. It is obvious that the pattern formed by \( p_1 \sim p_5 \) does not match against the peak pattern totally, but to some degree. Fuzzy theory is therefore applied here to define the fuzzy matching degree between these two patterns in Figure 7(b). For the value \( x_d(1 \leq i \leq k, 1 \leq j \leq m) \) of the \( i \)-th input variable of the \( j \)-th data point in the slip window, the corresponding \( T_p(x_{ij}) \) value in the peak pattern can be calculated by Equation (2) to form \( Y_{ij}^{\text{fp}} = \{y_{ij1}^{\text{fp}}, y_{ij2}^{\text{fp}}, \ldots, y_{ijm}^{\text{fp}}\} \) where \( y_{ijm}^{\text{fp}} = T_p(x_{ij}) \), \( 1 \leq i \leq k, 1 \leq j \leq m \). Similarly, the value of \( T_v(x_{ij}) \) in the valley pattern can be calculated by Equation (3) to form \( Y_{ij}^{\text{fv}} = \{y_{ij1}^{\text{fv}}, y_{ij2}^{\text{fv}}, \ldots, y_{ijm}^{\text{fv}}\} \), where \( y_{ijm}^{\text{fv}} = T_v(x_{ij}) \), \( 1 \leq i \leq k, 1 \leq j \leq m \).

Figure 7: Sample \( m \) data points and peak pattern comparison

As mentioned previously, the value \( y_j \) of \( j \)-th data point, where \( 1 \leq j \leq m \), is often different from \( y_{ij}^{\text{fp}} \) of a peak pattern.

The matching degree between value \( y_j \) and value \( y_{ij}^{\text{fp}} \) can be measured by a fuzzy membership function defined in Equation (4).

\[
\mu_{j1}(y_j) = \exp\left(-\frac{1}{2}\frac{\sum_{i=1}^{k} (y_j - y_{ij}^{\text{fp}})^2}{y_{\text{range}}}\right)
\]

Similarly, the matching degree between value \( y_j \) and value \( y_{ij}^{\text{fv}} \) in a valley pattern can be measured by a fuzzy membership function defined in Equation (5).

\[
\mu_{j2}(y_j) = \exp\left(-\frac{1}{2}\frac{\sum_{i=1}^{k} (y_j - y_{ij}^{\text{fv}})^2}{y_{\text{range}}}\right)
\]

where \( y_{\text{range}} = y_{\max} - y_{\min} \).
Since there are $m$ data points in the slip window and each data point has its own matching degrees derived from Equations (3) and (3), the overall matching degree between the pattern formed by the $m$ data points in the slip window and a designated pattern can be defined as Equation (6).

$$k_{qs} = \frac{\sum_{j=1}^{N} w_j \mu_{s}(y_j)}{\sum_{j=1}^{N} w_j}$$

where $q=1,2$, $s=1,2\ldots n$ and $\frac{1}{\sum_{j=1}^{N} w_j}$ is normalization factor.

The $s$ indicates the slip window is centered at the $s$-th data point, $q=1$ stands for the overall matching degree against peak pattern and $q=2$ stands for the overall matching degree against valley pattern. The $w_j(j=1,2\ldots m)$ are the weight values on the influence of deciding matching degree for the turning angle in the slip window against peak or valley patterns. The weight $w_1$ is assigned for $p_1$ and $w_2, w_3, \ldots w_m$ are respectively assigned for the rest of $m-1$ points in the slip window. It is obvious that $w_1$ should have the highest weight value since it is where the turning angle takes place, as shown in Figure 5. The second highest weight $w_2$ is assigned for $p_{s-1}$ and the third highest weight $w_3$ is assigned for $p_{s+1}$ and so on. The relationship of these weights is $w_1 > w_2 > w_3 > \ldots$ $> w_m$ indicating the fact that the farther from point $p_s$ the less influence on deciding matching degree of a turning angle at point $p_s$ against a designated pattern.

Equations (4) · (5) and (6), in conjunction with action of slip window, calculate the fuzzy matching degrees against designated patterns for every data point of $\{(x_{i1},x_{i2},\ldots x_{in},y_i)|1 \leq i \leq n\}$. A large value of $k_{qs}$ indicates a certain data point $p_s$ has a turning angle having high matching degree against a designated pattern. Consequently, the task of partitioning a data domain into clusters can be done in accordance of the descending list of $k_{qs}$ value.

Through the work of FDS, the fuzzy matching degrees of each input data point against peak and valley patterns are stored in descending order at FTPD matrix. Since higher matching degree indicates better location to divide the data domain, the values stored at FTPD is selected by the descending order. Therefore, each time a point corresponding to a selected value at FTPD matrix is used to divide the data domain into smaller clusters until the MSE (mean square error) is less than a threshold value.

3. Experiment

Due to page limit, only one 2-inputs-1-out nonlinear data domain is illustrated to show the performance of the proposed method. This example is the nonlinear equation:

$$y = (1-x_1^{-2}+x_2^{-1.5})^2, 1 \leq x_1, x_2 \leq 5$$

cited from [10], as shown in Figure 8. The width of slip window is set as 5 with the weighting value of $w=\{3,2,2,1,1\}$ and threshold value of $\delta=0.08$. After the proposed FDS is applied to the data domain using the peak and valley patterns in Figure 3, the data domain is divided into 3 subspaces resulting in following 3 Takagi-Sugeno fuzzy rules with SME = 0.05384

If $x_1$ is $A_1$ and $x_2$ is $A_2$ then $y = 2.743x_1+0.0435x_2 -0.3562$

If $x_1$ is $A_2^2$ and $x_2$ is $A_2$ then $y = 9.449x_1 -1.194x_2 -2.061$

If $x_1$ is $A_3^1$ and $x_2$ is $A_2^2$ then $y = 5.1x_1 -0.8331x_2 -0.2055$

where the $A_1^1$, $A_2^1$, $A_2^2$, $A_2^3$, $A_3^2$ are defined by a two-sided Gaussian function shown in Figure 1 with mean and standard deviation values indicated in Table 1. Table 2 is the comparisons of the proposed approach against other three methods. The comparison indicates that the proposed approach in this paper results in less rules with less MSE.

![Figure 8: $y = (1-x_1^{-2}+x_2^{-1.5})^2, 1 \leq x_1, x_2 \leq 5$](image)

Table 1: Mean and standard deviation values

<table>
<thead>
<tr>
<th>model</th>
<th>rule number</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugeno and Yasukawa</td>
<td>6</td>
<td>0.079</td>
</tr>
<tr>
<td>[10]$^*$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M.Delgado and Antonio F.</td>
<td>5</td>
<td>0.231</td>
</tr>
<tr>
<td>[11]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. Kim and M. Park and S.</td>
<td>3</td>
<td>0.0551</td>
</tr>
<tr>
<td>Ji[12]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our model</td>
<td>3</td>
<td>0.0538</td>
</tr>
</tbody>
</table>

![Table 1: Mean and standard deviation values](image)
4. Conclusions

In this paper, a Fuzzy Data Sifter is proposed to search out good turning-points to partition an input data domain into piecewise clusters. The TS fuzzy rules and membership functions are then extracted for each of the piecewise clusters to build the Takagi and Sugeno’s fuzzy model. The experiment has shown the good performance of the proposed approach. Furthermore, the FDS is robust in the sense of: (1) the proposed FDS is an open framework that allows the addition of other shapes of standard patterns; and (2) the setup value of threshold has great impact on the number of divided subspaces (clusters).

5. References