Hand Gesture Recognition Using Multivariate Fuzzy Decision Tree and User Adaptation

Moon-Jin Jeon, Korea Aerospace Research Institute, Korea

Sang Wan Lee, Massachusetts Institute of Technology, USA

Zeungnam Bien, Ulsan National Institute of Science and Technology, Korea

ABSTRACT

As an emerging human-computer interaction (HCI) technology, recognition of human hand gesture is considered a very powerful means for human intention reading. To construct a system with a reliable and robust hand gesture recognition algorithm, it is necessary to resolve several major difficulties of hand gesture recognition, such as inter-person variation, intra-person variation, and false positive error caused by meaningless hand gestures. This paper proposes a learning algorithm and also a classification technique, based on multivariate fuzzy decision tree (MFDT). Efficient control of a fuzzified decision boundary in the MFDT leads to reduction of intra-person variation, while proper selection of a user dependent (UD) recognition model contributes to minimization of inter-person variation. The proposed method is tested first by using two benchmark data sets in UCI Machine Learning Repository and then by a hand gesture data set obtained from 10 people for 15 days. The experimental results show a discernibly enhanced classification performance as well as user adaptation capability of the proposed algorithm.

Keywords: *Hand Gesture Recognition, Learning Algorithm, Model Selection, Multivariate Fuzzy Decision Tree, User Adaptation*

INTRODUCTION

Human computer interaction (HCI) technology has been widely used in various assistive systems for the disabled and the elderly. One of the recent highlighted topics is "understanding a user's intention" from natural human signals such as voice or gesture. Those signals, if successfully recognized, can provide a comfortable and convenient means for the user to interact with an engineering system. For example, a vision-based hand gesture recognition technique can be used to control a multitude of home appliances. Do et al. (2005) developed the Soft Remote Control System which enables the disabled user to control various home appliances using a set of simple hand gestures. Positions of a face and one hand are calculated using images obtained by stereo cameras. A concatenation of those positions constitutes a 3D trajectory of

DOI: 10.4018/ijfsa.2011070102

hands, from which the system recognizes those user's commands.

Critical factors that affect the performance of such systems are known to be false positive errors and inter-person variation / intra-person variation. False positive errors are caused by hand gestures that are meaningless but similar to some hand command gestures. To cope with this problem of false positive error, Yang (2007) proposed a gesture spotting method using the fuzzy garbage model in which an input gesture is classified either as a command gesture or a garbage gesture. The experimental results of the study shows that the command gestures such as "up" or "left" are effectively distinguished from the garbage gestures such as "eating" or "reading". In this paper, we deal with the latter problem of inter-person variation and intra-person variation.

When multiple users access to the system, the user independent (UI) recognition algorithms cannot compete with the user dependent (UD) recognition algorithms in the recognition rate. Furthermore, even for the same user, some characteristics of hand motion vary over time, which results in degradation of performance. The inter-person variation problem can be tackled by properly invoking some of the UD model techniques, model selection methods, or user adaptation strategies. The intra-person variation problem can be tackled by using fuzzy logic owing to its robustness property against uncertainty and ambiguity of human motion.

In particular, fuzzy decision tree learning has been widely used in classification problems due to its two merits: (1) interpretability of the decision tree and (2) capability of fuzzy logic in handling uncertainty and ambiguity (Janikow, 1998). Though the fuzzy decision tree is known to show a good performance in learning and classification tasks, however, it can be vulnerable to an overfitting situation that degrades prediction and adaptation performance. The higher the degree of overlap among membership functions is, the bigger the structure of the fuzzy decision tree becomes.

In this paper, we propose a multivariate fuzzy decision tree (MFDT) structure which effectively prunes the decision tree so as to enhance the classification and adaptation performance. The fuzzy decision tree model can be simplified by using a multivariate concept. Specifically, several recognition models are first built, and the best model that fits a new user is selected by using the maximum likelihood model comparison. Subsequently a user adaptation algorithm is presented, based on the gradient descent method.

To demonstrate the performance of the proposed algorithm, we use IRIS and WINE data set in UCI Machine Learning Repository (Merz et al., 1996) and a hand gesture data set which is collected from 10 people for 15 days. The experimental results show that the classification and user adaptation performances of the proposed method are better than those of other well-known fuzzy decision tree techniques.

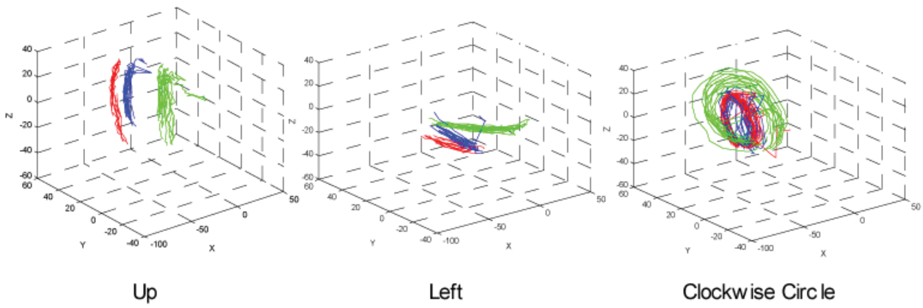
VISION-BASED HAND GESTURE RECOGNITION FOR THE SOFT REMOTE CONTROL SYSTEM

Vision-based hand gesture recognition for the soft remote control system (Do et al., 2005) is carried out in the following steps:

1. Face and hand "region of interests" (ROI) are extracted from camera images.
2. A trajectory of the hand position relative to the face position is calculated.
3. A start position and an end position of the trajectory are segmented.
4. Segmented trajectories are classified.

A set of hidden Markov models (HMM) can be used to recognize the hand gestures that contain temporal and spatial information (Rabiner, 1989). In the learning process, the parameters of the HMMs are optimized to fit the training sequences that correspond to the hand gestures. In the recognition process, the best matching HMM model for the observed sequence is chosen (Yamato et al., 1992). Without any painstaking feature selection process, vector

Figure 1. Hand gesture trajectories of three subjects



quantized (VQ) label and discrete HMM can be used for hand gesture recognition. It can be readily applied to complex gesture recognition problems. In spite of these merits, the HMM-based method has been known to require excessive amount of training data (Bourlard, 1990).

Hong et al. (2000) used finite state machine (FSM) for hand gesture recognition. Gestures are modeled as sequences of states in spatial-temporal space. The trajectories of a hand gesture are a set of points represented by a set of Gaussian spatial regions. In learning hand gestures, the number of states and their spatial parameters are calculated. The temporal information from the segmented hand gesture is then added to the states. The resulting state sequence can be regarded as a FSM recognizer. When a new sample is presented, each gesture recognizer decides whether to remain in the current state or to go to the next state. The sample is classified as the gesture whose FSM recognizer reaches to a final state. What distinguishes FSM from HMM is that FSM aborts a corresponding recognition process when a next point of the trajectory of the sample gesture is located far from the cluster of the FSM model. Hence, FSM is often believed to be simpler and faster than HMM.

Yang (2007) used fuzzy logic for hand gesture recognition due to its capability of easily dealing with ambiguity and uncertainty of human signals such as hand gestures. Using fuzzy sets, the weakness of the system caused by a fixed crisp decision boundary can be overcome. Also multiple membership functions can be utilized to effectively classify multiple classes (Su, 2000).

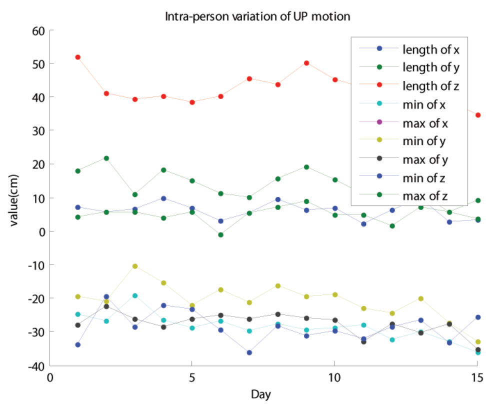
There are two major stumbling blocks that can affect the performance of a hand gesture recognition system. The first is the variability of the characteristics of the hand gestures among different individuals. It is called inter-person variation (Yang et al., 1996). Figure 1 illustrates the trajectories of the hand gestures (including up, left, and clockwise circle) of three subjects. We can easily find that the angle and length of line-type hand gestures and the radius of the clockwise circle gesture of these subjects are different. Figure 2 represents the inter-person variation in a feature space. It can be seen from the Figure 2 that distributions of the feature vectors of user 1 and user 2 are very different.

Secondly, even for the same person, hand gesture characteristics can vary from time to time, which is called "Intra-person variation". Figure 3 illustrates the characteristics of hand gesture which were recorded from one person for 15 days. It can be seen from Figure 3 that these features vary over time. This phenomenon is known to cause misclassification (Su, 2000).

One possible method for reducing the phenomena of inter-person variation and intra-person variation is a technique of loosening the decision boundaries of a hand gesture while not damaging the decision boundaries of other hand gestures (Jung, 2006). However, the false positive error caused by misclassification of unintended motions might be also increased. Figure

Figure 2. Inter-person variation in feature space

Figure 3. Variation of hand gesture characteristic for 15 days



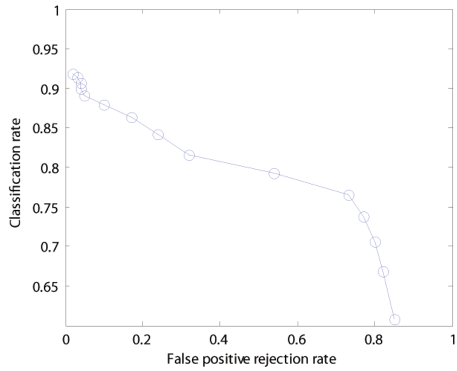
4 shows that classification rate is inversely proportional to false positive rejection rate.

Model construction methods can be stratified according to how training data are organized, such as user independent (UI) model construction and user dependent (UD) model construction.

The UI model is trained with the data of multiple users. On the contrary, each UD model is trained with the data of the corresponding user; the number of users is thus equal to the number of the UD models. Though it is easy to collect training data for the UI model construction, the risk of having lower classification rate is higher than that of the UD model. Conversely, though it is easy to achieve high classification rate with UD models, it is often difficult to gather training data for a new user. If we use a UI model to recognize a new user's hand gestures, the performance could be hampered by inter-person variation. To cope with this problem, various methods of user adaptation and personalized recognition have been studied. Jung (2007) suggested an adaptation method for a UI model and successfully applied to Korean sign language recognition problem. For the facial expression recognition problem, Kim (2004) suggested an example of a user dependent model and a model selection method.

In the next section, we propose a hand gesture recognition system which is capable of handling multiple users and the case of adding a new user. We then construct a personalized hand gesture recognition system, which includes

Figure 4. Relationship between classification rate and false positive rejection rate in hand gesture recognition task



the proposed MFDT learning and classification method. The UD model for each user is trained using the MFDT learning method. A maximum likelihood-based model comparison method is used to select a model that is fitted for a new user's patterns. Adaptation of this model for a new user is conducted by a gradient descent-based adaptation method.

MULTIVARIATE FUZZY DECISION TREE

A fuzzy decision tree provides a powerful means to overcome a major limitation of the decision tree, which stems from a crisp decision boundary. The key characteristic of fuzzy decision tree learning is that a 'single' sample can be designated as a reference point in 'multiple' nodes. Hence, the size and complexity of the trained fuzzy decision tree can be increased excessively, which brings about a poor generalization performance. This weakness of a fuzzy decision tree can be resolved by introducing an attribute vector, as opposed to using a separate attribute that branches each node of a fuzzy decision tree.

In a learning process, each node branches out to its child nodes by the points which maximize information gain of the node. In Figure 5, for example, if each node of the decision tree is branched by only one attribute, the decision tree may need six nodes for a complete separation of the data points. A multivariate decision tree suggested by Yildiz et al. (2001), however, needs just one split. We assert that the same principle can be effectively applied to the fuzzy decision tree. The resulting algorithm we propose in this paper is called a MFDT.

MFDT Learning

While FDT uses a single attribute to split each node, MFDT uses an attribute vector (i.e., multiple attributes) which is obtained by using linear discriminant analysis (LDA). LDA is a dimension reduction method that can be applied to classification problems (Alpaydin, 2004). It finds a projection vector **w** such that separability of the projected data is maximized as follows:

$$\text{Maximize J}(\mathbf{w}) \;=\; \frac{\mathbf{w}^{\mathrm{T}}\mathrm{S}_{\mathrm{B}}\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathrm{S}_{\mathrm{W}}\mathbf{w}}, \tag{1}$$

$$\text{where, } \mathrm{S}_{\mathrm{B}} = \sum_{i=1}^{K}(\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T, \tag{2}$$

$$\mathrm{S}_{\mathrm{W}} = \sum_{i=1}^{K}\mathrm{S}_i, \tag{3}$$

$$\mathrm{S}_i = \sum_{\mathbf{x}\in\ \text{class } i}(\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T, \tag{4}$$

Figure 5. Example of Univariate split and multivariate split for Iris data (Merz et al., 1996)

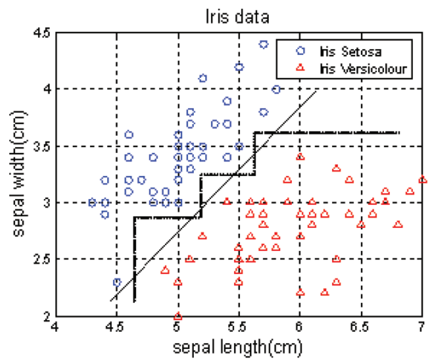
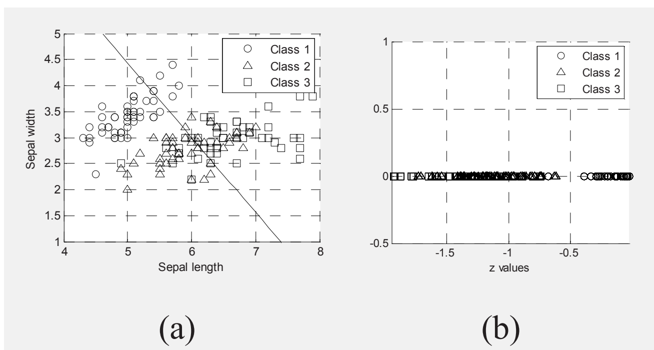


Figure 6. LDA results of Iris data



$$\mathbf{m} = \frac{1}{K}\sum_{i=1}^{K}\mathbf{m}_i \, . \tag{5}$$

Here $\mathbf{x}$ denotes a sample point, K is the total number of classed, and $\mathbf{m}_i$ is the mean of samples of the i_{th} class.

The largest eigenvector of $\mathrm{S}_\mathrm{W}^{-1}\mathrm{S}_\mathrm{B}$ is the solution $\mathbf{w}$ that maximizes $\mathrm{J}(\mathbf{w})$ (Alpaydin, 2004). This vector is used as an attribute vector onto which the given data are projected. Figure 6 represents LDA and projection results of Iris data from UCI Repository (Merz et al., 1996).

The whole process of MFDT learning is as follows.

Step 1. Generate a root node. After generating a root node, assign all training data to the root node.

Step 2. Build a node such that the information gain is maximized.

1. Obtain **w** using LDA, where:

$$\mathbf{w} = \text{the maximum eigenvector of } \mathrm{S}_\mathrm{W}^{-1}\mathrm{S}_\mathrm{B} \, . \tag{6}$$

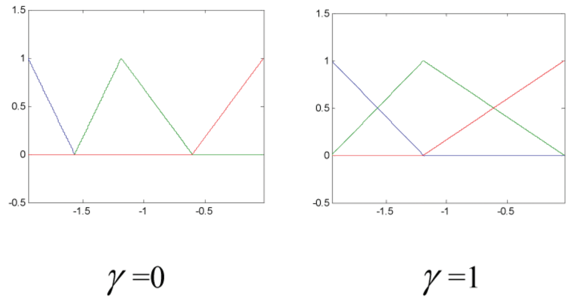
2. Calculate the attribute value using **w**:

$$z = \mathbf{w}^\mathrm{T}\mathbf{x} \, . \tag{7}$$

3. Calculate the entropy of the current node:

$$Entropy(S) = -\sum_i P_i^S \log_2 P_i^S , \tag{8}$$

Figure 7. Degree of overlap γ



where:

$$P_i^S = \frac{N_S^i}{N_S},\ N_S = \sum_i N_S^i \tag{9}$$

with N_S^i = number of data of the i_{th} class and S = a set of the attribute values z assigned in the current node.

4. Find a membership function which maximizes the information gain when the set of the attribute value z in the current node is split by using the membership function.

If the attribute values are arranged in an ascending order, there should be some points by which the class changes when crossing them. If we set N membership functions, we can get all combinations of N-1 such points. Then those combinations are used to make the membership functions. The middle points of the triangular membership functions are calculated as follows:

$$m^{v,c} = \frac{1}{n_v}\sum_{j=1}^{n_v} z_j^v\ ,\ v = 1,\dots,N \tag{10}$$

where n_v denotes the number of the attribute value z in the v_{th} area; z_j^v denotes the j_{th} value of the set of z when z in the v_{th} area are arranged in ascending order; $m^{v,c}$ denotes the middle point of the v_{th} membership function.

We simply define the left and right points of each triangular membership function with their middle points as:

$$m^{v,l} = m^{v,c} - 0.5\left(1+\gamma\right)\left(m^{v,c} - m^{v-1,c}\right),$$
$$m^{v,r} = m^{v,c} + 0.5\left(1+\gamma\right)\left(m^{v+1,c} - m^{v,c}\right), \tag{11}$$

where γ is the degree of overlapping in a membership function (Figure 7).

Using the above parameters of the membership function, we can calculate the fuzzy membership values for the attribute values in the current node.

If we define $\mu_{S_{v|\mathbf{w}}}\left(x\right)$ as the v_{th} membership function of current node, the information gain of the current node is calculated as follows:

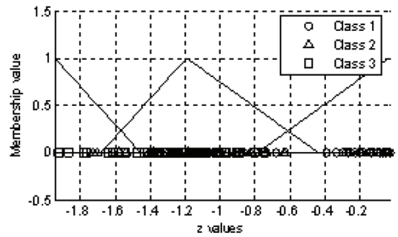
$$Gain(S,\mathbf{w}) = Entropy(S) - \sum_v \frac{N_{S_{v|\mathbf{w}}}}{N_S} Entropy\left(S_{v|\mathbf{w}}\right), \tag{12}$$

where

$$Entropy\left(S_{v|\mathbf{w}}\right) = -\sum_i P_i^{S_{v|\mathbf{w}}} \log_2 P_i^{S_{v|\mathbf{w}}}, \tag{13}$$

$$P_i^{S_{v|\mathbf{w}}} = \frac{C_{S_{v|\mathbf{w}}}^i}{C_{S_{v|\mathbf{w}}}},\ C_{S_{v|\mathbf{w}}} = \sum_i C_{S_{v|\mathbf{w}}}^i, \tag{14}$$

Figure 8. Automatically generated membership function



$$C^{i}_{S_{v|\mathbf{w}}} = \sum_{\substack{class\ of\ z=i \\ z \in Supp\left(S_{v|\mathbf{w}}\right)}} \mu_{S_{v|\mathbf{w}}}\left(z\right), \tag{15}$$

$$S_{v|\mathbf{w}} = \{(z, \mu_{S_{v|\mathbf{w}}}(z)) | \mu_{S_{v|\mathbf{w}}}\left(z\right) : \text{the membership value of the } v_{th} \text{ membership function.}\} \tag{16}$$

5. Calculate the attribute and the membership function in case that the current node is branched by using a single attribute (univariate case); the corresponding attribute vector **w** is as follows:

$$\mathbf{w} = \left[1 \quad 0 \quad . \quad . \quad 0\right]^{T}, \left[0 \quad 1 \quad . \quad . \quad 0\right]^{T}, ..., \left[0 \quad 0 \quad . \quad . \quad 1\right]^{T} \tag{17}$$

6. Select a univariate or multivariate node which has bigger information gain. The branch and child nodes are generated by using the membership function and the **w** that is selected.

Step 3. If the termination conditions are satisfied, make the current node a leaf node which refers to a class label of the majority of training data in the leaf node. If the termination conditions are not satisfied, go back to Step 2 for each child node.

The termination conditions are as follows:

1. The class labels of all data in a current node are the same.
2. The depth of the current node is larger than a predefined maximum value. The depth of a node is the number of nodes from the root node to the current node.

Figure 8 represents a membership function generated by the procedure. Figure 9 represents a MFDT trained by using Iris data set. The trained MFDT achieved 98% classification accuracy.

MFDT Classification

The classification procedure is as follows.

Step 1. Calculate the T-norm of the membership functions and attribute vectors of the nodes from the root node to each leaf node.

Projection of input data into the attribute vector in each node:

$$z_{i} = \mathbf{w}_{i}^{T}\mathbf{x} \tag{18}$$

The T-norm from the root node to the n_{th} leaf node is defined as follows:

$$T_{n} = \prod_{i=\text{root node}}^{n_{th}\ \text{leaf node}} \mu^{i}_{S_{v|\mathbf{w}}}\left(z_{i}\right) \tag{19}$$

The T-norm value is calculated by this method.

Step 2. Calculate the average T-norm of the leaf nodes that belong to the same class. Then

Figure 9. Trained MFDT

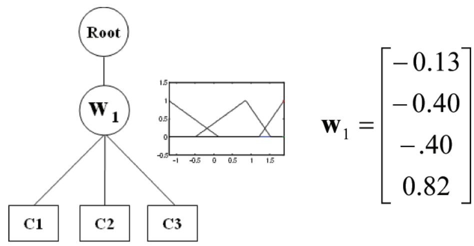
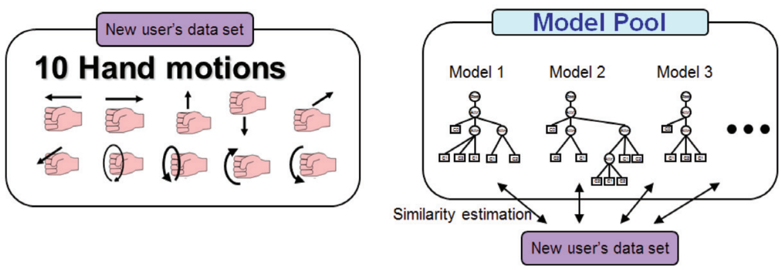


Figure 10. Model Selection process in MFDT-based hand gesture recognition system



classify the input data as the class with the maximum average T-norm value.

A_i : The average value of T-norms of leaf nodes which have the i_{th} class label.

$$class\, C = \arg\max_i A_i \quad (20)$$

MODEL SELECTION AND USER ADAPTATION

Model Selection

A multitude of recognition models that are generated by using multiple users' data sets are kept in a model pool and are used to recognize the input hand gestures.

When a new user starts to use the system, the most appropriate model is selected by using a model selection method (Figure 10). We measure how well the recognition model fits a hand gesture using the maximum likelihood model comparison (Duda et al., 2001).

$$P\left(m_i \mid D\right) = \frac{P\left(D \mid m_i\right) p\left(m_i\right)}{p\left(D\right)} \propto P\left(D \mid m_i\right) p\left(m_i\right) \quad (21)$$

$$\text{Selected Model } \hat{m} = \arg\max_{m_i} P\left(D \mid m_i\right) \quad (22)$$

D : data, m_i : i_{th} model

The model pool is built by using hand gestures of multiple users; A next step is to adapt to new users' patterns or to a change in existing users' patterns.

User Adaptation

The MFDT model selected in the model selection phase is used to recognize the new user's hand gestures. It also can adapt to the patterns of the new user's hand gestures; for instance, it could be adapted by a gradient descent based adaptation method. The MFDT adaptation is a kind of incremental adaptation methods (Fu et al., 2000).

An MFDT model has information about the membership function (refer to Figure 11). The average T-norm value of the leaf nodes which have the same class label as that of input data has to be bigger than any other average T-norm values of the leaf nodes which have other classes. We can select an error function which becomes small if the T-norm of the leaf nodes which have the same class label as that of input data is closed to 1. Then the error function can be adapted to minimize the error.

The adaptation process for the whole leaf nodes is as follows.

1. The input data is projected onto attribute vector of each node which is on the route from the root node to each leaf node. The projected value is given by:

$$z_n = \mathbf{w}_n^T \mathbf{x}, \tag{23}$$

where $\mathbf{w}_n$ is the attribute vector of the n_{th} node from the root node, and:

$$T = \prod_{n=1}^{N} \mu_n^i \left(z_n\right), \tag{24}$$

where N is the depth of the parent node.

2. Update the parameters of the membership function of each node by using a gradient descent method. While the center value of a membership function is calculated by using training data, the left and right values of the membership function are calculated based on the relationship of the center values of successive membership functions:

$$m_n^{i,l} = m_n^{i,c} - 0.5\left(1+\gamma\right)\left(m_n^{i,c} - m_n^{i-1,c}\right), \tag{25}$$

$$m_n^{i,r} = m_n^{i,c} + 0.5\left(1+\gamma\right)\left(m_n^{i+1,c} - m_n^{i,c}\right), \tag{26}$$

where n: node number; i: membership function number; γ: the degree of overlapping in a membership function (Figure 7).

The membership value is calculated as follows:

$$\mu_n^i\left(z_n\right) = \begin{cases} \dfrac{z_n - \left(1-\gamma\right) m_n^{i,c} - \gamma\, m_n^{i-1,c}}{\gamma\left(m_n^{i,c} - m_n^{i-1,c}\right)}, z_n < m_n^{i,c} \\ \dfrac{z_n - \left(1-\gamma\right) m_n^{i,c} - \gamma\, m_n^{i+1,c}}{\gamma\left(m_n^{i,c} - m_n^{i+1,c}\right)}, z_n > m_n^{i,c} \end{cases} \tag{27}$$

$\mu_n^i\left(z_n\right)$: membership value of input data

An error function whose minimization increases T-norm is defined as follows:

$$E = \frac{1}{2}\left(T-1\right)^2 = \frac{1}{2}\left(\prod_{n=1}^{N} \mu_n^i\left(z_n\right) - 1\right)^2. \tag{28}$$

And its adaptation rule is given by a gradient descent method:

$$\frac{\partial E}{\partial m_n^{i,c}} = \begin{cases} \left(\prod_{n=1}^{N} \mu_n^i\left(z_n\right) - 1\right) \cdot \left(\prod_{\substack{k=1 \\ k\neq n}}^{N} \mu_k^i\left(z_n\right)\right) \cdot \dfrac{m_n^{i-1,c} - z_n}{\gamma\left(m_n^{i,c} - m_n^{i-1,c}\right)^2}, z_n < m_n^{i,c} \\ \left(\prod_{n=1}^{N} \mu_n^i\left(z_n\right) - 1\right) \cdot \left(\prod_{\substack{k=1 \\ k\neq n}}^{N} \mu_k^i\left(z_n\right)\right) \cdot \dfrac{m_n^{i+1,c} - z_n}{\gamma\left(m_n^{i,c} - m_n^{i+1,c}\right)^2}, z_n > m_n^{i,c} \end{cases}, \tag{29}$$

$$m_n^{i,c} \leftarrow m_n^{i,c} - \eta \frac{\partial E}{\partial m_n^{i,c}}, \tag{30}$$

$$m_n^{i,l} \leftarrow m_n^{i,c} - 0.5\left(1+\gamma\right)\left(m_n^{i,c} - m_n^{i-1,c}\right), \tag{31}$$

$$m_n^{i,r} \leftarrow m_n^{i,c} + 0.5\left(1+\gamma\right)\left(m_n^{i+1,c} - m_n^{i,c}\right), \tag{32}$$

Figure 11. Parameters of membership function in a node of MFDT

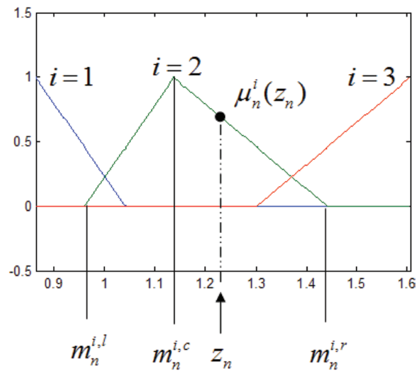


Table 1. Benchmark data sets

Data set	Classes	Instances	Features
Iris	3	150	4
Wine	3	178	13

Table 2. Classification rate of Iris data

Trial	FDT		MFDT	
	Number of nodes	CR (%)	Number of nodes	CR (%)
1	3	92.0	1	96.0
2	16	94.7	10	98.7
3	4	90.7	1	94.7
4	16	93.3	10	98.7
5	4	93.3	3	93.3
6	9	94.7	6	97.3
7	8	96.0	7	96.0
8	7	86.7	3	96.0
9	4	90.7	3	93.3
10	9	96.0	9	100.0
Average	8	92.8	5.3	96.4

Table 3. Classification rate of Wine data

Trial	FDT		MFDT	
	Number of nodes	CR (%)	Number of nodes	CR (%)
1	9	84.1	6	84.1
2	13	85.2	5	92.0
3	10	92.0	5	93.2
4	10	87.5	6	87.5
5	8	94.3	5	89.8
6	13	87.5	7	89.8
7	9	87.5	5	93.2
8	12	81.8	8	89.8
9	12	93.2	6	93.2
10	12	95.5	7	93.2
Average	10.8	88.9	6	90.6

Table 4. Comparison of classification rate of decision trees

Set	C4.5	C5.0	FDT	MFDT
Iris	92.9	92.9	92.8	96.4
Wine	86.6	89.2	88.9	90.6

Table 5. Influence of γ for classification of benchmark data

γ	Iris data		Wine data	
	Number of nodes	CR (%)	Number of nodes	CR (%)
0	2	94.4	2.5	91.6
0.1	2	94.8	3.3	92.6
0.2	2	96.5	3.9	90.9
0.3	3	97.1	4.2	91.0
0.4	4.3	95.9	4.8	89.7
0.5	5.3	96.4	6	90.6
0.6	5.9	96.1	6.4	89.7
0.7	7	96.0	7.6	88.8
0.8	9.2	94.7	9.4	90.9
0.9	11.9	95.5	11	92.7
1	17.9	94.9	15.1	91.9

Table 6. Classes and features of hand gesture data

10 Hand motion classes	16 Attributes
1. Up 2. Down 3. Left 4. Right 5. Forward 6. Backward 7. Clockwise circle 8. Counter clockwise circle 9. Clockwise half circle 10. Counter clockwise half circle	• Length on x, y, z axis • Minimum value on x, y, z axis • Maximum value on x, y, z axis • Time index of minimum value on x, y, z axis • Time index of maximum value on x, y, z axis • Eccentricity

Table 7. Unintended garbage motions

• Drinking water • Reading a newspaper • Fold a blanket	• Stretching • Raising one arm

Table 8. Hand gesture recognition rate

User	FDT			MFDT		
	Number of nodes	RR (%)	FPRR (%)	Number of nodes	RR (%)	FPRR (%)
User 1	28	90.8	26.9	21	88.3	26.7
User 2	24	77.7	63.1	18	90.8	35.0
User 3	16	86.9	58.5	12	98.3	41.7
User 4	17	66.9	73.1	11	90.8	58.3
User 5	18	76.5	61.5	17	93.3	58.3
User 6	23	89.2	51.5	15	86.7	50.0
User 7	16	74.6	62.3	18	86.7	31.7
User 8	22	85.0	53.8	17	85.0	26.7
User 9	21	92.3	75.4	17	89.2	21.7
User 10	20	90.8	53.1	12	96.7	71.7
Average	20.5	83.1	57.9	15.8	90.6	42.2

Figure 12. Average classification rate during user adaptation

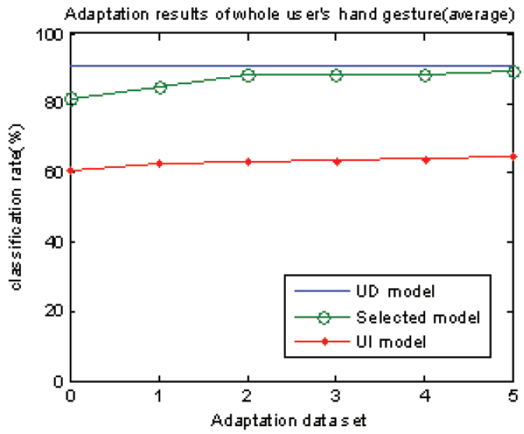


Table 9. Recognition rate using model selection and user adaptation

User	UI model	UI model + Adaptation	Selected model	Selected model + Adaptation
User 1	42.0	55.0	71.7	83.3
User 2	61.0	63.0	80.8	90.8
User 3	69.0	72.5	92.5	95.0
User 4	73.0	68.5	81.7	88.3
User 5	52.5	66.5	67.5	86.7
User 6	72.5	72.5	81.7	90.8
User 7	63.5	59.0	84.2	85.8
User 8	54.0	66.5	75.8	84.2
User 9	52.5	54.0	93.3	94.2
User 10	66.0	71.0	85.0	94.2
Average	60.6	64.9	81.4	89.3

$$m_n^{i+1,l} \leftarrow m_n^{i+1,c} - 0.5\left(1+\gamma\right)\left(m_n^{i+1,c} - m_n^{i,c}\right), \quad (33)$$

$$m_n^{i-1,r} \leftarrow m_n^{i-1,c} + 0.5\left(1+\gamma\right)\left(m_n^{i,c} - m_n^{i-1,c}\right), \quad (34)$$

where η denotes the adaptation size for an adaptation step.

3. Repeat 2. until the T-norm becomes smaller than the predefined value.

EXPERIMENTAL RESULTS

Classification Test using Benchmark Data

We tested the MFDT learning and classification method using Iris and Wine data set of

UCI Machine Learning Repository (Merz et al., 1996). Table 1 provides the specification of the data sets.

We used a 5x2 fold cross validation method to test the proposed method. Test results for Iris data and Wine data are shown in Tables 2 and 3. Note that MFDT achieved higher classification rate with fewer nodes than FDT. It can be seen from Table 4 that MFDT outperforms several conventional decision tree algorithms with a crisp decision boundary such as C4.5, and C5.0. γ is set to be 0.5. Table 5 illustrates how the value of γ influences the number of nodes and the classification rate.

Classification Test Using Hand Gesture Data

The hand gesture data were collected using 3 stereo camera units equipped on the ceiling of an intelligent residential space (Bien et al., 2002). The images were saved at 10 frames per second. The start and end points of hand gestures were manually segmented. We collected 10 kinds of hand gestures from 10 people for 15 days. Table 6 shows the 10 kinds of hand gestures and the 16 kinds of features used to build MFDT models.

The unintended or meaningless motion should be taken into consideration in a learning process of a hand gesture recognition model because a careful consideration of these factors can reduce false positive error. To obtain the data of unintended motions, we have collected 'garbage' data (Table 7) and have assigned them as the 11th class data.

The classification rates and the false positive rejection rates using the UD model of FDT and MFDT are compared in Table 8. Each user's recognition model is trained by using five sets of hand gesture data and two sets of garbage data.

Nine UD models except the selected user's model and a UI model are used for model selection. We have used one data set of a new user for model selection. The selected model then undergoes adaptation. Figure 12 shows user adaptation performances. As the adaptation proceeds, the average classification rate increases.

Table 9 shows an increase in recognition rates of 10 user's hand gestures when using the proposed model selection and user adaptation method. After user adaptation, the recognition rates of the selected model are increased and become close to that of the UD model; however, the recognition rate of UI model does not change much because the adaptation efficiency is decreased due to a huge size of the UI model.

CONCLUDING REMARKS

In this paper, we have proposed the MFDT learning and classification algorithm for robust hand gesture recognition. We have shown that the proposed MFDT method has a better generalization performance than the univariate fuzzy decision tree. The simulation results of classification tests using benchmark data set have shown that MFDT has a better classification performance than a typical general decision tree and fuzzy decision tree. In classification, model selection, and user adaptation tests using hand gesture data, the selected UD models show the best recognition performance.

The adaptation method proposed in this paper adjusts the parameter values of the membership function of each node. For more reliable adaptation of the MFDT, attribute vectors also should be able to adapt to the patterns of the hand gestures.

The proposed MFDT can be applied for various recognition problems. The robust classification and adaptation capability of MFDT can be applied to human signal recognition systems such as a hand posture recognition system or a gait recognition system in the future.

REFERENCES

Alpaydin, E. (2004). *Introduction to machine learning*. Cambridge, MA: MIT Press.

Bien, Z., Park, K.-H., Bang, W.-C., & Stefanov, D. H. (2002). LARES: An intelligent sweet home for assisting the elderly and the handicapped. In *Proceedings of the 1st Cambridge Workshop on Universal Access and Assistive Technology* (pp.43-46).

Bourlard, H., & Wellekens, C. J. (1990). Links between Markov models and multilayer perceptrons. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*(12), 1167–1178. doi:10.1109/34.62605

Do, J.-H., Jang, H.-Y., Jung, S.-H., Jung, J.-W., & Bien, Z. (2005). Soft remote control system in the intelligent sweet home. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3984-3989).

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Fu, H.-C., Chang, H.-Y., Xu, Y. Y., & Pao, H.-T. (2000). User adaptive handwriting recognition by self-growing probabilistic decision-based neural networks. *IEEE Transactions on Neural Networks*, *11*(6), 1373–1384. doi:10.1109/72.883451

Hong, P., Turk, M., & Huang, T. S. (2000). Gesture modeling and recognition using finite state machines. In *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 410-415).

Janikow, C. Z. (1998). Fuzzy decision trees: Issues and methods. *IEEE Transactions on Systems, Man, and Cybernetics*, *28*(1), 1–14. doi:10.1109/3477.658573

Jung, S.-H. (2007). *Incremental user adaptation in Korean sign language recognition using motion similarity and prediction from adaptation history.* Unpublished master's thesis, Korea Advanced Institute of Science and Technology, Daejeon, Korea.

Jung, S.-H., Do, J.-H., & Bien, Z. (2006). Adaptive hand motion recognition in soft remote control system. In *Proceedings of the 7th International Workshop on Human-friendly Welfare Robotic Systems*.

Kim, D.-J. (2004). *Image-based personalized facial expression recognition system using fuzzy neural networks*. Unpublished doctoral dissertation, Korea Advanced Institute of Science and Technology, Daejeon, Korea.

Merz, C. J., & Murphy, P. M. (1996). *UCI repository for machine learning data-bases*. Retrieved from http://archive.ics.uci.edu/ml/

Nam, Y., & Wohn, K. (1996). Recognition of. space-time hand-gesture using hidden Markov model. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (pp. 51-58).

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–285. doi:10.1109/5.18626

Su, M. C. (2000). A fuzzy rule-based approach to spatio-temporal hand gesture recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, *30*(2), 276–281.

Yamato, J., Ohya, J., & Ishii, K. (1992). Recognizing human action in time-sequential images using hidden Markov model. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 375-385).

Yang, S.-E. (2007). *Gesture spotting using fuzzy garbage model and user adaptation.* Unpublished master's thesis, Korea Advanced Institute of Science and Technology, Daejeon, Korea.

Yildiz, O. T., & Alpaydin, E. (2001). Omnivariate decision trees. *IEEE Transactions on Neural Networks*, *12*(6), 1539–1546. doi:10.1109/72.963795

Moon-Jin Jeon received the B.S. degree in electrical engineering and computer science from Hanyang University, Korea, in 2006 and the M.S. degree in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Korea, in 2008. He is currently a researcher of Korea Aerospace Research Institute (KARI). In the company, he participates in development of a low earth orbit satellite. His current research interests include spacecraft mission simulation, soft computing technique for estimating satellite state of heart data.

Sang Wan Lee received his BS degree in electrical engineering from Yonsei University, Korea, in 2003 and the MS and PhD degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Korea, in 2005 and 2009, respectively. He is currently a Postdoctoral Associate in Professor Poggio's lab at Massachusetts Institute of Technology (MIT). His current research interests include machine learning and computational neuroscience with emphasison attention model in the human visual system, hypothetical associative learning, and integrated information in discrete dynamic systems.

Zeungnam Bien (S'72-M'75-SM'91-F'07) received the BS degree in electronics engineering from Seoul National University, Korea, in 1969,and the MS and PhD degrees in electrical engineering from the University of Iowa, in1972 and 1975, respectively. During 1976-1977, he was an assistant professor in the Department of Electrical Engineering, University of Iowa. He, thereafter, joined the Computer Applications and Software Engineering (CASE) Center, Syracuse University, New York. During 1987-1988, he was a visiting professor in the Department of Control Engineering, Tokyo Institute of Technology, Japan. Since 1977, he had been with the Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, until2008, and became professor emeritus. In 2009, he joined the School of Electrical and Computer Engineering, Ulsan National Institute of Science and Technology (UNIST), Korea, where he is currently a chaired professor. His current research interests include intelligent automation and learning control methods, soft computing techniques with emphasis on fuzzy logic systems, and service robotics and rehabilitation engineering systems. He is a fellow of the IEEE and the IFSA (2007).He was a distinguished lecturer of the IEEE Robotics and Automation Society during 2004-2005. He was the president of IFSA during 2003-2005. He has been the director of the Human-Friendly Welfare Robot System Engineering Research Center, KAIST.