ABSTRACT

In this paper, we present a new technique for the recognition of hand gesture using decision tree method based on information entropy. Some rules are derived from the decision tree using training data, which can classify sixty-five different hand gestures. Normalization for all sensors in a DataGlove are also proposed to model the data variations of each sensor, which result from the same gesture variations. Compared with ANN, the proposed decision tree approach can not only improve the recognition performance by 12.2%, but also overcome the limitation of ANN in tedious training time.

1. INTRODUCTION

Hand gesture recognition is important for the communication between person with hearing ability and the hearing impaired as well as the research on the human-computer interface [1]. There are some approaches used in hand gesture recognition, such as ANN [2][3][4], HMM [5] and recursive induction [6]. The neural network algorithm for classification employs a method of gradient descent. Advantages of neural networks include their high tolerance to noisy data as well as their classify patterns on which they have not been trained. The drawbacks of this model manifest in tedious and time consuming training process and their poor interpretability. Hidden Markov Model (HMM) has been attempted to segment and to recognize continuous gestures. However, it has difficulties in recognition of manual gesture, because manual gesture has few motions. A major problem for regression scheme in hand gesture recognition is that the regression coefficients are hardly estimated by the method of least squares with small change of variables. Classification methods can be compared and evaluated according to the criteria such as predictive accuracy, speed, robustness, scalability and interpretability. In this paper, a new technique is presented for the recognition of hand gesture using decision tree, which always has some advantages such as scalability, training and testing speed, and interpretability. The decision tree method based on information entropy is employed and extended to recognize hand gestures using DataGlove as input devices. Some rules are derived from the decision tree using training data, which can classify sixty-five different hand gestures. Normalization for all sensors in a DataGlove are proposed to model the data variations of each sensor, which result from the same gesture variations. In addition to overcome the decision tree’s limitation on excessively accurate dividing the cases by certain threshold value which possible leads to wrong labels, the normalization can also solve problem of lacking the training data in hand gesture recognition.

The rest part of this paper is organized as follows: in section 2 we discuss the decision tree algorithm based on information entropy. Section 3 present the recognition of hand gesture using decision tree method. Sections 4 discuss the normalization for the noise of the system and the distortion of the same posture. Then we conclude a summary in section 5.

2. DECISION TREE BASED ON INFORMATION ENTROPY

A decision tree can be used to classify a case by starting at the root of the tree and moving through it until a leaf is encountered. At each nonleaf decision node, the case’s outcome for the test at the node is determined and attention shifts to the root of the subtree corresponding to this outcome. When this process finally leads to a leaf, the class of the case is predicted to be that recorded at the leaf.

The most important step in the algorithm is how to select test_attribute [7], in most application, they choose the information gain(T), which measures the average amount of information needed to identify the class of a case in T. Imagine selecting one case at random from a set S of cases and
announcing that it belongs to some class $C_j$. This message has probability

$$p(C_j) = \frac{freq(C_j, S)}{|S|}$$

Where $freq(C_j, S)$ stand for the number of cases in $S$ that belong to class $C_j$

So the information it conveys is

$$-\log_2(p(C_j))$$

To find the expected information from such a message pertaining to class membership, we sum over the classes in proportion to their frequencies in $S$, giving

$$H(C) = -\sum_j p(C_j) \log(p(C_j))$$

Now, consider a similar measurement after $C$ has been partitioned in accordance with the $m$ outcomes of a test $Att_i = \{Att_{ij} | j = 1,2,3,...,m\}$. The expected information requirement can be found as the weighted sum over the subsets, as

$$H(C | Att_i) = \sum_{j=1}^{m} p(Att_{ij}) \sum_{i=1}^{n} p(C_i | Att_{ij}) \log \frac{1}{p(C_i | Att_{ij})}$$

The quantity

$$gain(Att_i) = H(C) - H(C | Att_i)$$

Measure the information that is gained by partitioning $C$ in accordance with the test $Att_i$. The $gain$ criterion, then, selects a test to maximize this information gain.

3. HAND GESTURE RECOGNITION USING DECISION TREE

Two CyberGlove with 18 sensors (see figure 1) and a Pohelmus 3-D tracker with two receivers positioned on the wrist of each CyberGlove are used as input device in our system. In this paper we focus on how to classify the basic hand gesture without considering the position and the orientation.

To deal with the sign data, there are all continuous attributes, we have to select good algorithm for finding appropriate threshold for the continuous attribute to partition the cases. There are some of them that person used. In this paper, we use the follow method:

The training cases $T$ are first sorted on the values of the attribute $Att_i$ being considered. There are only a finite number of these values, for sign data, we normalize them to 0…256,so let us denote them in order as $\{v_1, v_2, ..., v_m\}$. Any threshold value lying between $v_i$ and $v_{i+1}$ will have the same effect of dividing the cases into those whose value of the attribute $Att_i$ lies in $\{v_1, v_2, ..., v_i\}$ and those value in $\{v_i, v_{i+1}, ..., v_m\}$. There are thus only $m-1$ possible splits on $Att_i$, all of which examined. It’s usual to choose the midpoint of each interval as the representative threshold, the $i$th such being

$$\frac{v_i + v_{i+1}}{2}$$

In C4.5 they choose the largest value of $Att_i$ in the entire training set that do not exceed the midpoint above, rather than the midpoint itself, as the threshold [8]. This ensures that all threshold value appearing in trees and/or rules actually occur in the data. For sign data, there may be lots noise in them; we also should catch the noise data even if they are not appearing in the train data. Therefore, the threshold should be a proper number that give a space room for the noise.

Sixty-five basic hand gestures are selected from Chinese sign language. In experience we get five group glove data of the basic units of right hand gesture, and each group have 700 samples. Our four group data are used to train the decision tree, and the remaining was used to test. Sixty-five rules can be derived from the decision tree that can classify the basic hand gesture we defined.

The formulation of the rule is just as follow:
If((6 <=119)&&(14<=139)&&(7<=104)&&(3<=120)&&(2>107)) then basic gesture 21

Testing result are showed in Table 1

<table>
<thead>
<tr>
<th>data</th>
<th>The CRR for Decision tree</th>
<th>The CRR for ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traindata</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Test data</td>
<td>86.2%</td>
<td>96.4%</td>
</tr>
</tbody>
</table>

The result of testing illustrates several important considerations in hand postures recognition based on decision tree. First, the decision tree is sensitive to noisy data caused by system or the signer, and the recognition rate is 86.2%. Second, the recognition rate of the decision tree is higher if the training data is used as testing data. These results indicate certain ways to get around noise problem. The recognition of hand gesture based on decision tree may be improved if the noise for some condition can added to the training data. Then the decision tree can achieve a higher robustness for including the variety range of the testing data.

4. NORMALIZATION

A hand gesture recognizer often encounters two main noises: noise of system and distortion of the same hand gesture. The noise of system such as electrical noise and quantization noises, which of course are present in any electronic recognition system, are in general at a level below the threshold of concern. Nevertheless, noise due to transmission and switching equipment in DataGlove can sometimes be a factor affecting recognizer performance. Figure 3 shows the noise of the fifth sensor in DataGlove. Many factors affect the shape of same gesture of each individual signer. Even for the same signer, there are variety shapes of the same gesture. Figure 4 shows. The variety of basic hand gesture 1. Figure 5 show the curve of the fifth sensor’s data of DataGlove in the variety of hand gesture 1.

![Figure 3: the system noise of the fifth sensor](image)

The noise of system is acuity in local and can be flitted easily if it can be processed separately. However, always the noise of system is combined with the distortion of the hand gesture.

![Figure 4: The variety of basic hand gesture 1](image)

![Figure 5: The noise and distortion of fifth sensor Of basic hand gesture 1](image)

As pointed out previously, if the characteristic of the corrupting noise or distortion are approximately known, our decision tree trained from with the noise-added data, in general, may be perform more robustly than one uses clean data. The problem is how to measure the range of the noise. The noises of the 18 sensors are different with each other because of the different range they are in. Each data of sensor has their max-min range, then their values could be normalized to 0-256. We define seven basic hand gestures to obtain the max-min range of each sensor. Table 2 shows the max value and min value of each sensor (each data was turn to integer)
Table 2: the max-min value of 18 sensors

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>39</td>
<td>83</td>
<td>48</td>
<td>48</td>
<td>56</td>
<td>57</td>
<td>39</td>
<td>63</td>
<td>69</td>
</tr>
<tr>
<td>Max</td>
<td>108</td>
<td>106</td>
<td>161</td>
<td>166</td>
<td>151</td>
<td>168</td>
<td>143</td>
<td>163</td>
<td>217</td>
</tr>
<tr>
<td>Min</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>Max</td>
<td>50</td>
<td>60</td>
<td>40</td>
<td>40</td>
<td>57</td>
<td>44</td>
<td>46</td>
<td>82</td>
<td>45</td>
</tr>
<tr>
<td>Min</td>
<td>111</td>
<td>73</td>
<td>71</td>
<td>129</td>
<td>116</td>
<td>136</td>
<td>145</td>
<td>111</td>
<td>10</td>
</tr>
<tr>
<td>Max</td>
<td>111</td>
<td>73</td>
<td>71</td>
<td>129</td>
<td>116</td>
<td>136</td>
<td>145</td>
<td>111</td>
<td>10</td>
</tr>
</tbody>
</table>

The decision tree reconstructed and retrained using normalized data, the result of testing show in Table 3:

Table 3. The Result of Test decision tree

<table>
<thead>
<tr>
<th>Test data</th>
<th>0% noise</th>
<th>5% noise</th>
<th>10% noise</th>
<th>15% noise</th>
<th>20% noise</th>
<th>30% noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traindata</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>99.8%</td>
<td>100%</td>
</tr>
<tr>
<td>Train data 10% noise</td>
<td>100%</td>
<td>92.3%</td>
<td>98.4%</td>
<td>95.3%</td>
<td>89.2%</td>
<td>78.4%</td>
</tr>
<tr>
<td>Test data</td>
<td>86.2%</td>
<td>92.3%</td>
<td>93.8%</td>
<td>89.2%</td>
<td>92.3%</td>
<td>75.4%</td>
</tr>
<tr>
<td>Test data 10% noise</td>
<td>84.6%</td>
<td>92.3%</td>
<td>93.8%</td>
<td>90.8%</td>
<td>92.3%</td>
<td>76.9%</td>
</tr>
</tbody>
</table>

5 CONCLUSION

In this paper, a new technique is presented for the recognition of hand gesture using decision tree, which always has some advantages such as scalability, training and testing speed, and interpretability. Some rules are derived from the decision tree using training data, which can classify sixty-five different hand gestures. Normalization for all sensors in a DataGlove are proposed to model the data variations of each sensor, which result from the same gesture variations. In addition to overcome the decision tree’s limitation on excessively accurate dividing the cases by certain threshold value which possible leads to wrong labels, the normalization can also solve problem of lacking the training data in hand gesture recognition. Compared with ANN, the proposed decision tree approach can not only perform more robustness and improve the recognition performance but also overcome the limitation of ANN.

6.REFERENCE