Hand Recognition Using Thermal Image and Extension Neural Network

Meng-Hui Wang

Department of Electrical Engineering, National Chin-Yi University of Technology, No. 35, Lane 215, Sec. 1, Chung-Shan Rd., Taichung County Taiping City, 411, Taiwan

Correspondence should be addressed to Meng-Hui Wang, wangmh@ncut.edu.tw

Received 18 August 2011; Revised 7 October 2011; Accepted 12 October 2011

Academic Editor: Angelo Luongo

Copyright © 2012 Meng-Hui Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Hand recognition is one of the popular biometry methods for access control systems. In this paper, a new scheme for personal recognition using thermal images of the hand and an extension neural network (ENN) is presented. The features of the recognition system are extracted from gray level hand images, which are taken by an infrared camera. The main advantage of the thermal image is that it can reduce errors and noise in the features extracted stage, which is most important to increase the accuracy of recognition systems. Moreover, a new recognition method based on the ENN is proposed to perform the core functions of the hand recognition system. The proposed ENN-based recognition method also permits rapid adaptive processing for a new pattern, as it only tunes the boundaries of classified features or adds a new neural node. It is feasible to implement the proposed method on a Microcomputer for a portable personal recognition device. From the tested examples, the proposed method has a significantly high degree of recognition accuracy and shows good tolerance to errors added.

1. Introduction

Person recognition and verification is a very important function in many access control systems. Biometric technology is a new method for recognizing the identity of a person based on an already established database of physiological or behavioral characteristics [1, 2]. Usually, these physiological features must be unique, invariant, carry-on, and be permanent, they cannot be imitated or be carried with an individual without remembering. Recently, various biometry techniques have been proposed in literature, such as the fingerprints, hand geometry, palm prints, face recognition, the iris, and speech [3–8]. Comparison among the biometry verification reveals the truth, that each biometry has its own advantages and limitations. Fingerprint-imaging-based systems require good frictional skin, while the iris-based or retina-based identification systems also require a special illumination setup. On the other hand,
worry and uncomfortable feelings of users are the weak points of the above recognition systems. Hand geometry is one of the earliest biometry verification systems [9, 10] and it exists in some commercial systems. The advantages of hand recognition are the ease of acceptance, causing no anxiety, and being easily setup. Earlier hand recognition systems used low-resolution CCD cameras or scanners to capture hand images, where the surrounding environment and lighting affected the quality of the hand images, as well as the following disadvantages: (1) traditional methods have lower distinguishability due to low-resolution hand images; (2) users concerns regarded hygienic issues when touching screens and caused low acceptability.

To improve the problems of the traditional technologies, this paper proposes a new hand-geometry-based recognition system, which uses an infrared camera to acquire the thermal image of user’s hands. There are some advantage of the proposed scheme as follows: (1) without the problem of light interference, photographs can be taken under areas of inadequate lighting; (2) infrared cameras can detect radiation heat emanated from human hands; hence, it is not affected by different lengths of palms, dirt, or wounds; thus, it does not lead to errors and lower discriminability; (3) thermal imaging can be taken by a noncontact and noninvasive image capture devices that can avoid causing any uncomfortable feelings or hygienic concerns of users. Therefore, this paper presents a thermal imaging method to capture hand images, then using the Otsu method extracts the hand features from the gray level images. Moreover, this paper proposes a new recognition method based on ENN to perform the core functions of the hand recognition system. The proposed ENN-based recognition method [11–18] also permits fast adaptive processing for new patterns, as it only tunes the boundaries of classified features or adds a new neural node. It is feasible to implement the proposed method on a Microcomputer for a portable personal recognition device. From the tested examples, the proposed method has a significantly high degree of recognition accuracy and shows good tolerance to errors added.

2. The Operation of the Proposed Recognition System

2.1. The Structure of the Hand Recognition System

The operation of the hand-based recognition system is similar to other biometric authentication devices: sample acquisition, feature extraction, data storage, comparison, and verification. The structure of the proposed hand recognition system is shown in Figure 1. Hand recognition is performed on the thermal image of the palmar surface acquired by an infrared thermal imaging camera. In this paper, an appropriate imaging consists of aluminum alloy, with the features of light weight and greater heat exchange; namely, the temperature of the aluminum alloy does not vary due to hand contact, thus, the aluminum alloy temperature is lower than the temperature of the hand, and the boundary of palmar is easily detected in thermal imaging analysis. A preprocessing module is used to enhance the geometric image of palmar surface, then a feature extraction module is used to compute the geometric parameters based on the processed image. The matching module compares the inputs model with the stored typical models in the database to generate relational degrees through the new recognition method based on the extension neural network as proposed in this paper. The final recognition results can be taken by the decision module according to relational degrees.

2.2. The Detecting Method of the Palmar Boundary

The palmar boundary must be detected for capturing the characteristics of the palmar shape. The boundary can be detected by scanning the pixel points, where the scanning is divided
into two stages. At stage 1, the images are scanned vertically from top to bottom, and the pixel points move from the left to the right. If the left and right adjacent gray values of the pixel points are all greater than 0, the locations of the pixel coordinates are memorized and expressed by white points. However, some places are missed in detection. Hence, at stage 2, the images are scanned from the left to the right in the horizontal direction, and the pixel points move from top to bottom. If the upper and adjacent gray values of the pixel points are all greater than 0, the locations of the pixel coordinates are also memorized. After the scanning action of the two stages, the palm edge coordinates can be thoroughly detected.

The palm edge coordinates contain the valley points and peak points of fingers, as shown in Figure 3. In this paper, the valley and peak points are defined by calculating the Euclidean Distance between a point \( i \) in the coordinate set \( S_b \) of the palm edge and the starting point of the palm contour \( S_1 \), as shown in Figure 4. Equation (2.1) is used to calculate the Euclidean Distance, where \( D_i \) is the distance value between each coordinate and the starting point \( c \). The \( i \) represents the sequence of coordinate contour; \( X_i \) and \( Y_i \) are the coordinate values of each coordinate contour; and \( X_c \) and \( Y_c \) are the coordinate values of the starting point.

\[
D_i = \sqrt{(X_i - X_c)^2 + (Y_i - Y_c)^2} \quad i \in S_b.
\] (2.1)
In the clockwise direction, the palm contour coordinates $S_b$ begin from starting point $c$ and returns to starting point $c$. Equation (2.1) is used to calculate the distribution of distance between $c$ and each coordinate on the palmer contour, as shown in Figure 5.

The distance distribution in Figure 5 illustrates the five filled circles of the peak points and the four slash circles of the valley points, where the indicated coordinates are exactly the peak and valley points in Figure 4. In this paper, another three valley points must be found in order to extract the features of the length and width of the five fingers. In Figure 6, the white solid lines represent the left valley points of the thumb and the index finger, and the right valley point of the little Figure 7. The point-increase method is to utilize the distance between the coordinates of a known valley point and a known peak point to determine the coordinate point of the same distance on the opposite side. As shown in Figure 7, after calculating the distance between the known coordinates of $P$ and $V$ points, the $N$ point coordinate can be found by adding the $P$-$V$ distance from $P$ point on the opposite side of $V$ point, where $N$ point is a new valley point. This method can help to determine all three new valley points.
Figure 5: Distribution condition of the distance between palmar contour coordinates and point c.

Figure 6: Additional valley points.

Figure 7: Feature of palm distance distribution.
2.3. Feature Extraction Method

For the feature extraction of palmar shapes, there are a total of 34 features, which become fixed features after reaching a certain age; hence, they can be used as recognition features. The proposed features in this paper are summarized in Table 1. The palmar contour size of each person is different; therefore, the number of palmar edge pixel points can be regarded as the identification feature after calculation as well as the palmar contour after excluding the fingers. The next step is to determine the gravity point of palm images, where the distance from that point to the palm contour and the distances from that point to the fingers are being used as identification features. The gravity point coordinate can be determined by (2.2) where the \( C_x \) and \( C_y \) are the center coordinates of the image and \( H \) is the Heaviside Step Function and is only equal to 1 when \( \phi > 0 \), otherwise, it is 0. The denominator represents the number of total pixel points within the palm contour, and the numerator represents the summation of \( x \) and \( y \) coordinates within the palm contour. After averaging, the gravity coordinate can be determined and is shown as solid white line circles in Figure 7:

\[
\begin{align*}
C_x &= \frac{\sum_x \sum_y x \cdot H(\phi)}{\sum_x \sum_y H(\phi)}, \\
C_y &= \frac{\sum_x \sum_y y \cdot H(\phi)}{\sum_x \sum_y H(\phi)}.
\end{align*}
\tag{2.2}
\]

After the gravity coordinate is determined, the distribution features of palm gravity distances are extracted, including the distance between the gravity coordinate and valley points of each finger, the distance between the gravity coordinate and the midpoints of valley points, and the horizontal and vertical distances between the gravity coordinate and palm contour, as shown in Figure 7. The finger length extraction method is to calculate the midpoint coordinate between points A and B, and the distance between the midpoint and the H point is the length of the finger, as shown in Figure 8. Normally, people have five fingers; thus, there are five length features.

In addition, three widths of each finger are extracted and regarded as features. As shown in Figure 9, the distance between U and V points is the first width of the index finger; the 2/3 distances between P and U points, and P and V points are coordinately labeled as points X and Y; the distance between them is the second width; the 1/3 distances between P and U points, and P and V points are labeled as points L and R, of which the distance is the third width of the index finger. There are a total of 15 width features. The finger length and width features are not enough to represent the physiological features of a person and cannot determine the uniqueness of each individual, and using only these two features may result in incorrect identifications. The identification features that represent the identity of a person should be increased and enhanced. This study extracted finger sizes and regarded them as identification features. When determining the finger sizes, the contour coordinates of each finger are found, and then the valley points and peak point of each finger can be connected, as shown in Figure 10. At this time, each finger can form its own independent closed curve. Meanwhile, the number of white pixel points is the feature of the finger size.

3. The Proposed Pattern Recognition Method

At the recognition stage, the inputting patterns are compared with the patterns stored in the system database. Learning from a set of training patterns is an important feature of most pattern recognition systems. The neural networks are usually used for pattern recognition;
Figure 8: Finger length features.

Figure 9: Finger width features.

Figure 10: Finger size features.
Table 1: Palm shape features.

<table>
<thead>
<tr>
<th>Item</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hand contour</td>
</tr>
<tr>
<td>2</td>
<td>Palm contour</td>
</tr>
<tr>
<td>3</td>
<td>Horizontal distance of gravity</td>
</tr>
<tr>
<td>4</td>
<td>Vertical distance of gravity</td>
</tr>
<tr>
<td>5</td>
<td>Distance between gravity and thumb</td>
</tr>
<tr>
<td>6</td>
<td>Distance between gravity and index finger</td>
</tr>
<tr>
<td>7</td>
<td>Distance between gravity and middle finger</td>
</tr>
<tr>
<td>8</td>
<td>Distance between gravity and third finger</td>
</tr>
<tr>
<td>9</td>
<td>Distance between gravity and little finger</td>
</tr>
<tr>
<td>10</td>
<td>Contour of thumb</td>
</tr>
<tr>
<td>11</td>
<td>Thumb length</td>
</tr>
<tr>
<td>12</td>
<td>Thumb width</td>
</tr>
<tr>
<td>13</td>
<td>1/3 width of thumb</td>
</tr>
<tr>
<td>14</td>
<td>2/3 width of thumb</td>
</tr>
<tr>
<td>15</td>
<td>Contour of index finger</td>
</tr>
<tr>
<td>16</td>
<td>Length of index finger</td>
</tr>
<tr>
<td>17</td>
<td>Width of index finger</td>
</tr>
<tr>
<td>18</td>
<td>1/3 width of index finger</td>
</tr>
<tr>
<td>19</td>
<td>2/3 width of index finger</td>
</tr>
<tr>
<td>20</td>
<td>Contour of middle finger</td>
</tr>
<tr>
<td>21</td>
<td>Length of middle finger</td>
</tr>
<tr>
<td>22</td>
<td>Width of middle finger</td>
</tr>
<tr>
<td>23</td>
<td>1/3 width of middle finger</td>
</tr>
<tr>
<td>24</td>
<td>2/3 width of middle finger</td>
</tr>
<tr>
<td>25</td>
<td>Contour of third finger</td>
</tr>
<tr>
<td>26</td>
<td>Length of third finger</td>
</tr>
<tr>
<td>27</td>
<td>Width of third finger</td>
</tr>
<tr>
<td>28</td>
<td>1/3 of width of third finger</td>
</tr>
<tr>
<td>29</td>
<td>2/3 of width of third finger</td>
</tr>
<tr>
<td>30</td>
<td>Contour of little finger</td>
</tr>
<tr>
<td>31</td>
<td>Length of little finger</td>
</tr>
<tr>
<td>32</td>
<td>Width of little finger</td>
</tr>
<tr>
<td>33</td>
<td>1/3 of width of little finger</td>
</tr>
<tr>
<td>34</td>
<td>2/3 width of little finger</td>
</tr>
</tbody>
</table>

The advantage of a neural network over other classifiers is that it can acquire experience from the training data, but the training data must be sufficient and compatible to ensure proper training, its convergence of learning is influenced by the network topology and values of learning parameters. To overcome the limitations of the multilayer neural network (MNN) mentioned [12], a new pattern recognition method, based on the ENN is presented for palmer recognition in this paper.

3.1. Structure of the ENN

In this clustering problem of hand recognition, hand’s features and associated person types cover a range of values. Therefore, using the ENN is most appropriate for hand recognition.
The schematic structure of the ENN is depicted in Figure 11. It comprises both the input layer and the output layer. The nodes in the input layer receive the input features and use a set of weighted parameters to generate an image of the input pattern. In this network, there are two connection values (weights) between input nodes and output nodes; one connection represents the lower bound and the other connection represents the upper bound for this classical domain of the features. The connections between the \( j \)th input node and the \( k \)th output node are \( w_{kj}^l \) and \( w_{kj}^u \). This image is further enhanced in the process characterized by the output layer. The output layer is a competitive layer. There is one node in the output layer for each prototype pattern, and only one output node with nonzero output to indicate the prototype pattern that is closest to the input vector.

### 3.2. Learning Algorithm of the ENN

The learning of the ENN is to tune the weights of the ENN to achieve good clustering performance or to minimize the clustering error. Before the learning, several variables have to be defined. Let training set \( \{X_1, T_1\}, \{X_2, T_2\}, \ldots, \{X_Q, T_Q\} \), where \( Q \) is the total number of training patterns, \( X_i \) is an input vector to the neural network, and \( T_i \) is the corresponding target output. The \( i \)th input vector is \( X_i = \{x_{i1}, x_{i2}, \ldots, x_{in}\} \), where \( n \) is the total number of the features. To evaluate the learning performance, the error function is defined below:

\[
E_i = \frac{1}{2} \sum_{i=1}^{Q} \sum_{j=1}^{n_c} (t_{ij} - o_{ij})^2,
\]

where \( t_{ij} \) represents the desired \( j \)th output for the \( i \)th input pattern and \( o_{ij} \) represents the actual \( j \)th output for the \( i \)th input pattern. The detailed supervised learning algorithm can be described as follows.

**Step 1.** Set the connection weights between input nodes and output nodes according to the range of classical domains. The range of classical domains can be directly obtained from previous experience, or determined from training data as follows:

\[
w_{kj}^l = \min_{T_i \in k} \{x_{ij}\}, \quad w_{kj}^u = \min_{T_i \in k} \{x_{ij}\},
\]

for \( i = 1, 2, \ldots, Q; \ j = 1, 2, \ldots, n; \ k = 1, 2, \ldots, n_c \).

**Step 2.** Read the \( i \)th training pattern and its cluster number \( p \)

\[
X_i = \{x_{i1}, x_{i2}, \ldots, x_{in}\}.
\]

**Step 3.** Use the extension distance (ED) to calculate the distance between the input pattern \( X_i \) and the \( k \)th cluster as follows:

\[
ED_{ik} = \sum_{j=1}^{n} \left( \frac{|x_{ij} - \left( w_{kj}^l + w_{kj}^u \right)/2 - \left( w_{kj}^l - w_{kj}^u \right)/2}{\left( w_{kj}^u - w_{kj}^l \right)/2} + 1 \right)
\]

for \( k = 1, 2, \ldots, n_c \).
The proposed extension distance is a new distance measurement; it can be graphically presented as in Figure 12. The proposed ED can describe the distance between the $x$ and a range $\langle W^L, W^U \rangle$, which is different from the traditional Euclidean distance.

Step 4. Find the $m$, such that $ED_{im} = \min\{ED_{ik}\}$. If $m = p$ then go to Step 6; otherwise go to Step 5.

Step 5. Update the weights of the $p$th and the $m$th clusters as follows:

$$
\begin{align*}
\mathbf{w}^{L\text{(new)}}_{pj} &= \mathbf{w}^{L\text{(old)}}_{pj} + \eta \left\{ x_{ij} - \frac{[\mathbf{w}^{L\text{(old)}}_{pj} + \mathbf{w}^{U\text{(old)}}_{pj}]}{2} \right\}, \\
\mathbf{w}^{U\text{(new)}}_{pj} &= \mathbf{w}^{U\text{(old)}}_{pj} + \eta \left\{ x_{ij} - \frac{[\mathbf{w}^{L\text{(old)}}_{pj} + \mathbf{w}^{U\text{(old)}}_{pj}]}{2} \right\}, \\
\mathbf{w}^{L\text{(new)}}_{mj} &= \mathbf{w}^{L\text{(old)}}_{mj} - \eta \left\{ x_{ij} - \frac{[\mathbf{w}^{L\text{(old)}}_{mj} + \mathbf{w}^{U\text{(old)}}_{mj}]}{2} \right\}, \\
\mathbf{w}^{U\text{(new)}}_{mj} &= \mathbf{w}^{U\text{(old)}}_{mj} - \eta \left\{ x_{ij} - \frac{[\mathbf{w}^{L\text{(old)}}_{mj} + \mathbf{w}^{U\text{(old)}}_{mj}]}{2} \right\},
\end{align*}
$$

(3.5)

where $\eta$ is a learning rate, set to 0.2 in this paper. From this step, we can clearly see that the learning process is only to adjust the weights of the $p$th and the $m$th clusters.

Step 6. Repeat Step 2 to Step 5, if all patterns have been classified, then a learning epoch is finished.

Step 7. Stop, if the clustering process has converged, or the total error has arrived at a preset value, otherwise, return to Step 3.
It should be noted that the proposed ENN can take human expertise before the learning, and it can also produce meaningful output after the learning, because the classified boundaries of the features are clearly determined.

3.3. Operation Phase of the ENN

Step 1. Read the weight matrix of the ENN.

Step 2. Read a testing pattern

\[ X_t = \{x_{t1}, x_{t2}, \ldots, x_{tn}\}. \] (3.6)

Step 3. Use the proposed extension distance (ED) to calculate the distance between the tested pattern and every existing cluster by (3.4).

Step 4. Find the \( m \), such that \( ED_{im} = \min \{ED_{ik}\} \), and set the \( O_{im} = 1 \) to indicate the cluster of the tested pattern.

Step 5. Stop, if all the tested patterns have been classified; otherwise go to Step 2.

4. Experimental Results

To demonstrate the proposed method, 600 sets of palmar images with 30 persons were used to test the proposed method. In this case, the structures of the proposed ENN are 30 output nodes and 34 input nodes. If the system randomly chooses 300 instances from the hand image as the training data set, the rest of the instances of the hand image are the testing data sets. The proposed hand recognition system is implemented in a PC with the Visual Basic; the recognition window of the proposed systems is shown in Figure 13. The accuracy of traditional recognition method was compared with the proposed method, and some experimental results are shown as follows.

4.1. Error Analysis of Input Features with Different Cameras

When capturing hand images, the surrounding environment and lighting can affect the quality of the hand images, causing wrong identifications when poor images have been acquired.
Earlier hand identification systems employed CCD cameras and have the following disadvantages: (1) CCD can only produce images in daylight and cannot be used in the dark. A lighting device must be installed to solve this problem; (2) when the traditional CCD identifies hand-shaped characteristics, errors may occur and the identification rate may be lowered due to different lengths of fingernails, dirt, or wounds present during feature extraction. Figure 14 shows the error of input features with different cameras; it should be noticed that, the image of the traditional CCD causes larger average errors, of about 12.7%, whereas the average error rate of the proposed method is only about 5.4%, so it shows that the proposed infrared image method has low light interference and is also not affected by different lengths of palms, dirt, or wounds; thus, it does not lead to errors and lower discriminability.

4.2. Error Analysis of Input Features with Different Finger’s Angles

When traditional biometric systems capture image features, there may be offsets in angles or positions of palm as the user is unfamiliar with the method of application or nervous, causing errors when capturing features; thus, the status identified may be incorrect. Figure 15 shows two thermal images of the hand of a third person; it can be observed that there are slight differences in palm position, finger spacing, and hand angle, in order to prove that the feature capture method, as proposed by this paper, will not result in wrong recognition if the user’s palm angle or position deviates when taking thermal images; moreover, finger spacing may cause large errors in feature capture. This study captures feature values from ten different thermal images of a third person, calculates the error rate, and explains the stability of the thermal image feature capture, as based on the feature error rate. Figure 16 shows the mean error rate of the features of 9 thermal images, where the features of the first image are taken as the standard for the features of 10 thermal images of a third person, the error rate is less than 5.4%, which proves that the feature capture proposed by this paper is good, and the errors caused by the open angle of fingers can be reduced.
4.3. Compare the Recognition Performance with Different Methods

In this paper, the total training samples are 300 sets, and the total testing samples are 300 sets with 30 persons. Table 2 compares the learning performance of the proposed ENN with other existing methods. The results show that the proposed ENN has the shortest training time and highest accuracy of all methods. The proposed ENN-based recognized method in both the training and testing accuracy has a significantly higher diagnosis accuracy of 100% and 99%, respectively, which are higher than the multilayer-neural-network- (MNN-) based methods and K-means-based method. Although the hand recognition system is trained off-line, the training time is not a critical point to be evaluated. It is an index, however, implying in some degree the efficiency of the algorithm developed. As shown in Table 2, it should be noted that

<table>
<thead>
<tr>
<th>Compare items</th>
<th>ENN</th>
<th>MNN-1</th>
<th>MNN-2</th>
<th>MNN-3</th>
<th>K-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>34-30</td>
<td>34-33-30</td>
<td>34-34-30</td>
<td>34-35-30</td>
<td>N/A</td>
</tr>
<tr>
<td>No. of connections</td>
<td>2040</td>
<td>2112</td>
<td>2176</td>
<td>2240</td>
<td>N/A</td>
</tr>
<tr>
<td>Learning times (epochs)</td>
<td>86</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>N/A</td>
</tr>
<tr>
<td>Learning accuracy</td>
<td>100%</td>
<td>91%</td>
<td>93%</td>
<td>95%</td>
<td>70%</td>
</tr>
<tr>
<td>Testing accuracy</td>
<td>99%</td>
<td>88%</td>
<td>90%</td>
<td>86%</td>
<td>68%</td>
</tr>
</tbody>
</table>
the structure of the proposed ENN is simpler than the other neural networks, only 64 nodes and 2040 connections are needed. Moreover, the proposed ENN-based recognized method also permits fast adaptive processes for a large amount of training data or new information, because the learning of ENN is only to tuning low bounds and upper bounds of excited connections, which is rather beneficial when implementing the ENN-based recognized method in a Microcomputer for a real-time hand recognizing device or a portable instrument.

5. Conclusions

This paper presents a novel hand recognized method based on the ENN for biometric authentication. This study applied a thermal imaging camera to capture the palmar images to develop the person’s recognition system; the average errors of the input features with thermal
camera are smaller than average errors of the using traditional CCD camera. According to the experimental results, the proposed recognized features show that the errors caused by the open angle of fingers can be reduced. Compared with other existing methods, the structure of the proposed ENN-based method is simpler and the learning time is faster than other methods. Moreover, the proposed ENN-based hand recognized method also permits fast adaptive processes for the new data, which only tune the boundaries of classified features or add a new neural node. It is feasible to implement the proposed method in a Microcomputer for portable fault detecting devices. We hope this paper will lead to further investigation for industrial applications.

References
