Real-time Recognition of Humans By Their Walk

Li Wei
Department of Information Communication, Engineering, TongMyong University, room 515, 1th information building, 179, Sinseonno, Nam-gu, BuSan , Korea
E-mail: 12li-wei@163.com

Anh Viet Nguyen
Department of Information Communication, Engineering, TongMyong University, room 515, 1th information building, 179, Sinseonno, Nam-gu, BuSan , Korea
E-mail: nguyenanhviet82@gmail.com

Eung-Joo Lee*
Department of Information Communication, Engineering, TongMyong University, room 515, 1th information building, 179, Sinseonno, Nam-gu, BuSan , Korea
E-mail: ejlee@tu.ac.kr
*Corresponding author

Abstract: In this paper, we will present a novel method of human recognition using gait information. The position of the moving object is first detected using the difference between two continuous frames. Then, based on the detected moving object, all of the gait information is exploited. Based on the head, shoulders and back contour, the walking direction can be extracted. Depending on this parameter, the lateral or frontal characteristic is focused on in the features detection process. Once all of these characteristics have been exploited, a feature vector is constructed for use in the recognition process. To evaluate the effect of the proposed method and compare it with other methods, we present some simulation results obtained in both indoor and outdoor environments.

Keywords: biometric; moving object detection; gait recognition.

Reference

Biographical notes: Li Wei received his B. S. at Dalian University of Light Industry in China (2002 - 2006), and received his M. S degree at Tongmyong university in Korea. Now he is a Ph.D student of Tongmyong university. His main research tops are image processing, computer vision, Biometrics and face recognition.

Nguyen Anh Viet received his B.S. degree in Information Technology from University of Natural Sciences of Vietnam National University at Ho Chi Minh City (2000 - 2004). From 2005 to 2007 he worked for Dolsoft Inc, an international company in Security and GIS solutions. And now (2007 - 2009) he is a master student.

Eung-Joo Lee received his B. S., M. S. and Ph. D. in Electronic Engineering from Kyungpook National University, Korea, in 1990, 1992, and Aug. 1996, respectively. Since 1997 he has been with the Department of Information & Communications Engineering, Tongmyong University, Korea, where he is currently a professor. From 2000 to July 2002, he was a president of DigitalNetBank Inc.. From 2005 to July 2006, he was a visiting professor in the Department of Computer and Information Engineering, Dalian Polytech University, China. His main research interests includes biometrics, image processing, and computer vision.
1 Introduction

Nowadays, the evolution of society is resulting in many new requirements. For example, automatically monitoring a protected area and generating an alert when an object intrudes into this area. In this respect, building a timekeeping system that does not need any human intervention is necessary. To construct such a system, the problem arises of how to detect and recognize a person walking in this protected zone. In computer vision, there are many approaches that have been proposed to solve this problem. In the history of human recognition, the finger print, iris, and hand characteristics were introduced and have become popular. More recently, the use of the face features has been applied for human recognition. All of the above methods have the potential to provide perfect recognition, but there exists an inconvenience in that they require the cooperation of the person to be recognized. In this context, the gait, which is one of the biometric characteristics, is a particularly useful feature for human recognition at a distance. Gait is defined as the way or manner in which a person walks on foot. Early studies about biometrics suggest that gait is a unique characteristic for every person, so it is possible to identify a person by using the way they walk. (Cutting and Kozlowski, 1977) fitted ellipses to seven rectangular subdivisions of the silhouette and then computed four statistics (first and second-order moments) for each ellipse and, hence, obtained 28 1D signals from the entire silhouette sequence. Finally, they used three different methods of mapping these signals to obtain a single feature vector for classification. M.S. Nixon, J.N. Carter, J.M. Nash, P.S. Huang, D. Cunado, and S.V. Stevenage (Acquah and Nixon, 2003) computed the width of the outer contour of the silhouette and the correlation between the probe silhouette image sequences and those in a data-set. Bobick and Johnson (Bobick and Johnson, 2001) computed the body height, torso length, leg length, and step length for human identification. Using a priori knowledge about the body structure in the double support phase of walking (i.e., when the feet are maximally apart), they estimated these features as the distances between the fiducial points (namely, the midpoint and extreme) of the binary silhouette. Obviously, the accuracy of these measurements is very sensitive to the segmentation noise in the silhouette, even if they are averaged over many frames. Rather than using the entire silhouette, other methods use a signature of the silhouette obtained by collapsing the XYT data into more terses 1D or 2D projection histograms (XT) and horizontal projection histograms (YT).

Gait has many advantages over the other features. The first point that should be mentioned is distance recognition. Not only is the cooperation of the persons being analyzed not required, but also it is difficult for them to conceal themselves. An individual trying to disguise him or herself would probably appear more suspicious. However, as in the case of the other features, gait is also affected by certain factors. Gait describes the way a person walks on foot, so if this person has any physical changes (accident or disease), this feature would be affected. Stimulants such as drugs and alcohols also influence the gait. Another important factor is the person’s clothes; the same person wearing different clothes will create a widely different signature.

In this paper, we propose an algorithm which not only makes use of the advantages of gait, but also overcomes some of its inconveniences. First, in the object detection process, we use the active contours (snake) algorithm to obtain the object boundary. Then, based on the detected object, all of the gait features are extracted and become the input data for the recognition step.

2 PROPOSED METHOD OF RECOGNIZING HUMANS USING THEIR WALK

In general, the proposed algorithm can be separated into 3 main parts: object detection, object feature extraction and recognition vector creation.

The difference between 2 consecutive frames is first used for the purpose of detecting the object position. Then, all of the walking features of the blob are extracted. The formed feature vector is used for the recognition method.

2.1 Object Detection Method

Human tracking and extracting the moving object in the video stream is the pre-processing step in almost all supervision systems. In the field of computer vision, there are a lot of methods that have been proposed to solve this problem. However, the object detection step in this algorithm is very important, because all of the gait features are extracted from the detected object. The method we use herein is a combination of the comparison of 2 continuous frames in the video stream to detect the moving object and background subtraction for the purpose of obtaining the full silhouette. The background subtraction process is the simplest method and is the most effective method of obtaining the moving object. In some other methods, the background image must first be created from several images in the video sequence. This method can give a good result, but it is difficult to apply in real-time systems. Not only does this method require a lot of time to estimate the background image, but also the estimated background image may contain noise.

In the proposed algorithm, the difference between the current frame and the previous frame is computed firstly to detect the position of the moving object. The original image color is converted to the HSL color space for computation.
\[ \text{diff} = f(F_{xy}^i, F_{xy}^{i-1}) \]  

(1)

Where \( \text{diff} \) stands for the difference value of pixel value (HLS) of 2 continuous frames. \( f \) is the difference operation, \( F_{xy}^i, F_{xy}^{i-1} \) stands for pixel value of frame i, i-1 at x and y location.

The first thing to do after the different computations is to connect the object and remove the noise from the mask image that is obtained. To accomplish this, we suggest a method using the Opening Process. The combination of the Erode and Dilate algorithms can not only connect the close parts of the object, but also remove some of the noise.

We denote \( \theta \) as the background image which is created from the first video frame. For every input image, we need to update the background image for the silhouette extraction process in the next step. This is shown in Fig.3

\[ \theta = \theta UF_i \]  

(2)

where \( F_i \) is the current frame.

From the background image and the object boundary in the above steps, we can use the image subtraction method to extract the silhouette. However, when using this method, we must consider the brightness. Therefore, the selection of a suitable threshold for the image subtraction step is very difficult. To solve this problem, we use the following extraction function to indirectly perform differencing (Viet and Lee, 2008).

\[ f(a, b) = 1 - \frac{2\sqrt{(a+1)(b+1)} + 2\sqrt{(256-a)(256-b)}}{(a+1) + (b+1)} \times \frac{2\sqrt{(256-a)(256-b)}}{(256-a)(256-b)} \]  

(3)

Where \( a(x, y) \) and \( b(x, y) \) are the brightness of current image and the background at the pixel position \( (x, y) \) respectively, \( 0 \leq a(x, y), b(x, y) \leq 255, 0 \leq f(a, b) < 1 \)

This function can detect the change in the sensitivity of the difference value according to the brightness level of each pixel in the background image (Viet and Lee, 2008).

### 2.2 Object Boundary Detection Method

From the binary region, it is easy to detect the boundary polygon. Firstly, we use the Dilation and Erosion method to expand and shrink the source binary region. Then, the union of the two resultant images will give us all of the edges of the object, as shown in Fig.5.

From the contour of the silhouette, we use the approximation method to find the boundary polygon. This method is shown in Fig. 6 and the result is shown in Fig. 7. The detail of this method is described below :

1. Select \( P_o \) as the starting vertex \( P_o \rightarrow P_{head} \).
2. Find \( P_i \) which has the maximum distance to \( P_o \), \( P_i \rightarrow P_{head} \).
3. Find \( P_j \) which has the maximum distance to \( P_o P_i \), \( P_j \rightarrow P_{head} \).
4. Select \( P_j \) as the starting vertex
5. Return to 2.

### 3 Gait Features Extraction

Now, we will describe how the gait features are extracted from the object detected in the previous section. First, all of the gait features are listed and the details of how to detect this feature will be described in the next step.

From the above model, we proposed some walking features for human recognition:

- Maximum angle at elbow and knee.
- Walking velocity.
- 2 feet distance.
- Ratio of swing phase and stance phase.
- Ratio of width and height of silhouette over a gait cycle.
- The variant of silhouette perimeter on gait cycle.
- Head direction.

Gait is the way that a person walks, so almost all of the important characteristics belong to his or her legs. However, as mentioned in the discussion of the disadvantages of gait, if the person is wearing a raincoat or any other clothes that cover the leg region, the leg features cannot be extracted. Furthermore, the walking direction also affects the feature selection for recognition. This is the reason why we chose all of the above features. Fig. 9 shows the separated body parts, which are used to extract the features. They are the head, torso, thigh and shank region.

When a person is walking, the most easily recognized features in the leg are the speed and stride length. As in any basic method of speed computation, we must have the distance and time that the person walked. Here, we chose the gait cycle as the standard distance and the time is the number of frames which the person needs to finish a cycle. However, the concept of the gait cycle must first be understood. A gait cycle is defined as the time interval between the point when both feet are in contact with the floor and the point when they recontact the floor in the same leg position. Simply, a gait cycle is composed of 2 consecutive strides. As shown in figure 1, the above array image is the first half of a cycle, which extends from the beginning state to the mid-point, and the bottom array image is from the mid-point to the finished state.

There are two main phases in every gait cycle: the stance and swing. During the stance phase, the foot is
always in contact with the ground, while in the swing phase one foot is on the ground and the other leg is swinging in preparation for the next stride. We can subdivide the stance phase into three separate phases: the first double support (both feet are in contact with the ground), single limb stance (when one foot is swinging and the other foot is on the ground), and the second double support (when both feet are again on the ground). In normal gait, there is a natural symmetry between the left and right sides, but in pathological gait, an asymmetrical pattern very often exists. This is graphically illustrated in Figure 4. Notice the symmetry in the gait of the normal subject between the right and left sides in the stance (62%) and swing (38%) phases; the asymmetry in these phases in the gait of the two patients, who spend less time bearing weight on their involved (painful) sides and the increased cycle time for the two patients compared to that of the normal subject.

In the next phase, the asymmetry of a patient is shown. With the painful leg, the stance phase represents 58% and the swing phase 42%, while in the sound one, these values are 69% and 31%, respectively.

Based on this statistical data, we can choose the ratio of the stance to the swing components as a characteristic for human gait feature classification. In addition, we can also obtain some features from the gait cycle, which will be introduced below, that can give a good recognition result. A simple but effective characteristic is the knee angle when the person is walking. The change of this value of every individual is different. We define this feature as the angle which is made by the thigh and shank leg. In the walking period of some people, their shank swings just a little, so this angle is big, while others that have a strongly swinging shank will have a small angle.

After the silhouette is extracted, the leg region can be estimated by obtaining 35% of the object boundary.

\[
\begin{align*}
\text{LegRT} &= \text{objRectTop} + \text{objRectHeight} * 0.65 \\
\text{LegRL} &= \text{objRectLeft} \\
\text{LegRR} &= \text{objRR} \\
\text{LegRB} &= \text{objRB}
\end{align*}
\]  

In the next step, the skeletal algorithm is applied for the purpose of obtaining the skeleton of the thigh and shank. The first position of each leg is selected at the ankle point. The second point we use is the center point of each leg. The first position of each leg is selected at the ankle point. Where:

\[h_i = D(P_i, P_1 P_3)\]
\[P_{\text{knee}} = P_{\text{hmax}}\]
\[2\text{FootD} = D(LL_A, RL_A)\]

Where: \(h_i\) is the distance from thigh and shank point to \(P_1 P_3\), \(P_{\text{hmax}}\) is the point which has the distance to \(P_1 P_3\) is maximum. \(2\text{FootD}\) is the distance the leg foot and the right foot. \(LL_A\) is the left leg ankle point, \(RL_A\) is the right leg ankle point.

Based on the distance between the two feet, we can calculate the speed of one gait cycle. When this value reaches a maximum, we mark this time as the beginning of the cycle, while after this distance reaches a maximum value 2 more times, the gait cycle is finished. We assign the number of frames from the beginning to the finishing state of this cycle as the walking velocity.

\[WV = \frac{FN}{GC}\]  

To calculate the next feature, the direction of the head, we use the center of gravity of the head region. From the extracted silhouette, it is easy to obtain the binary head region. Then, the head contour is extracted by using the method in II.b. After obtaining the head boundary rectangle, we continue to calculate the coordinates of the centroid point.

\[
\begin{align*}
A &= \frac{1}{2} \sum_{i=0}^{N-1} (x_i y_{i+1} - x_{i+1} y_i) \\
Cx &= \frac{1}{4A} \sum_{i=0}^{N-1} (x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \\
Cy &= \frac{1}{4A} \sum_{i=0}^{N-1} (y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i)
\end{align*}
\]  

In the above equation, \(A\) is the area of the polygon, \(N\) is the number of vertexes, \(x_i, y_i\) are the coordinates of each vertex, and \(Cx, Cy\) are the coordinates of the centroid point. The body centroid is also calculated using the same method. After both of them are extracted, we can calculate the direction of the head when this person is walking. When a person is walking, the silhouette width and height will change from frame to frame. When the distance between the two feet reaches a maximum value or they are overlapped, both the width and height of the person will also be affected. Moreover, because the variation of the ratio of the width to the height is different for everyone, we can use it for human recognition.

Now, we introduce the last important characteristic, the variation of the silhouette perimeter, and the method of detecting it. Whereas most of the features which we analyzed in the previous step belong to separate parts of the human body, this characteristic depends on the whole silhouette. This characteristic is not only affected by the leg region, but also influenced by the swinging of the arm. In the human waking phase, the free arm movement is important to maintain balance and move effectively. For a specific person, the arm swing amplitude will have different features. The arm swing with the maximum amplitude and the leg region width with a maximum value will give a large perimeter value. Otherwise, if the arm swing amplitude is small and the two legs are overlapped, the perimeter will be decreased.

4 Experimental Results

To illustrate the effect of the proposed algorithm, we used the existing USF HumanID Gait Database (Sarkar and Phillips, 2004) and our self-built database which
uses video sequences from different scenes. This includes outdoor surveillance and indoor video. The size of each video is 640*480 and the frame rate is 25 frames/second. Firstly, we briefly describe the USF HumanID gait database and secondly we evaluate the object detection method under different conditions. Then, thirdly, we test the performance of the gait recognition algorithm and finally we compare our proposed algorithm with several other established algorithms for human gait recognition to analyze the advantages and disadvantages of the proposed algorithm.

4.1 A. HumanID Gait Database: Gallery and Probe Data Sets

The USF HumanID outdoor gait (people-walking-sequence) database version 2.1 was built for vision-based gait recognition and is widely used. It consists of 1870 sequences from 122 subjects (people). For each of the subjects, there are the following covariates: change in viewpoint (Left or Right), change in shoe type (A or B), change in walking surface (Grass or Concrete), change in carrying condition (briefcase or No Briefcase), and elapsed time (May or November) between the sequences being compared. There is a set of 12 predesigned experiments for algorithm comparisons. For algorithm training, the database provides a gallery collected in May, with the following covariates: grass, shoe type A, right camera, and no briefcase. The gallery also includes a number of new subjects collected in November. This gallery data set has 122 individuals. For algorithm testing, 12 probe sets are constructed according to the 12 experiments. Detailed information about the probe sets is given in Table 1. More detailed information about the USF HumanID is given in (Sarkar and Phillips, 2004).

Fig 10 shows examples of the gait images. It shows a gait cycle within a sequence. A gait cycle is a series of stances: from full-stride stance and heels-together stance to full-stride stance.

4.2 Object Detection Method Evaluation

The detection method influences the final gait recognition result, because it is the first step of the whole algorithm. Based on the available ground truth data and obtained detection results, the widely used confusion matrix can be used for the evaluation. Fig 11 shows the different relationships between the annotated and computed detection results. Based on this, the values for the true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) can be calculated.

Different measures can be derived from these values, of which the most important are the true positive rate (TPR) which is defined as:

$$TPR = \frac{TP}{TP + FN}$$  \hspace{1cm} (8)

And the false positive rate (FPR) given as:

$$FPR = \frac{FP}{FP + TN}$$  \hspace{1cm} (9)

In the ideal case, a detector would have TPR=1 and FPR=0 which is quite impossible in reality. Thus, the main goal for designing an object detector is to obtain a good tradeoff between these two values. By changing the detection parameter p (the background estimation threshold value), different value pairs (FPR(p), TPR(p)) can be obtained. The receiver operating characteristic (ROC) curve is a well-known way to illustrate this dependency by plotting these value pairs against each other for certain values of p. The larger the area under the ROC curve, the better the performance of the detector. Fig 12 shows the resulting ROC curve for human object detection with different background conditions.

As can be seen from the figure, the detection performance of the object detection method with different conditions varies considerably, since the background buffer updating method is similar for all conditions, and it depends on time. The background can be more stable with a longer time and the object detection result can be more accurate. However, if the object has a similar color to the background, this will confuse the detection algorithm. The detection result will not be tolerated. Table 2 shows some detection and feature extraction results with different background environments.

4.3 Gait Recognition Method Evaluation

The gait recognition method uses the feature vector which is composed of the angles ($\theta_1$, $\theta_2$); walking velocity ($V_{walk}$); distance between the 2 feet ($D_{Feet}$); ratio of the swing phase to stance phase ($R_{swing}$); ratio of width to height of silhouette over a gait cycle ($R_{wh}$); variation of silhouette perimeter over a gait cycle ($P^*$) and head direction ($\theta_{head}$). All of these components affect the recognition results. Table 3 shows the output results for training the NNs by using the feature vector based on the HumanID Gait Database. The input of the NNs has 8 neurons, the data is from the feature vector $V_{feature} = \{\theta_1, \theta_2, V_{walk}, D_{Feet}, R_{swing}, R_{wh}, P^*, \theta_{head}\}^T$, and the output result is composed of 6 neurons which are coded from 0 to 1.

In the recognition stage, we use the testing image to test the NNs, and the recognition results of the pattern (pattern 1012, Registrant 10, image 12) are shown as an example. Table 4 gives the outputs from the NNs for pattern 1012. The output from the NNs is the answer to the recognition result.

We set the threshold to 0.5, so that the output value is set to 1 if it is greater than 0.5, while otherwise it is set to 0. Thus, the output of pattern 1012 is 0001010, which represents 10. Therefore, pattern 1012 was identified as registrant 10.
4.4 Performance Evaluation and Comparison

Table 5 reports all the experiments which compare the proposed algorithm with the existing algorithms. The item "Avg" in the table means the averaged recognition rates of all probes (A-L), that is the ratio of correctly recognized subjects to the total number of subjects in all probes (A-L), that is, the ratio of correctly recognized subjects to the total number of subjects in all probes. The columns labeled A to L are exactly the same tasks as in the baseline algorithm. In the table, the first rows give the performance of the Baseline (Sarkar and Phillips, 2004), hidden Markov model (HMM) (Kale and Sundaresan, 2006), Image Euclidean Distance (IMED) Wang and Zhang (2005), IMED+LDA, LDA (Han and Bhanu, 2004), LDA+Sync (Han and Bhanu, 2004), 2DLDA Wang and Zhang (2005) and 2DLDA+LDA Wang and Zhang (2005). The performance of the proposed algorithm applied to the same gallery set and probe set is fully reported in all of the comparison experiments, namely, GaitFeature+NNs. Finally, the last columns of the table report the average performance of the corresponding algorithm on all probe sets.

From the comparison results in Table 5, it is clear that the average recognition rate of the 12 probes in the proposed method (GaitFeature+NNs) outperforms the previous state-of-the-art algorithms (top part in the table). For example, the HMM algorithm, which is stable in modeling the gait cycles, and the IMED algorithm, which is demonstrated to improve on the conventional LDA. The performances of the different methods have the following relationship: Baseline < HMM < IMED < IMED+LDA < LDA < LDA+Sync < 2DLDA < 2DLDA+LDA < GaitFeature+NNs. The recognition rate of the proposed method reaches 84.639%. Furthermore, the performances for probes D-G and K and L are not satisfactory. Therefore, further studies are required to make them applicable.

5 Conclusion

In summary, in this paper, a new method of human recognition is proposed. First of all, the moving object is detected by using the difference between two continuous frames. In the next step, the base of the silhouette, (viz.) the head shoulder back region, is obtained and the walking direction is estimated. Depending of this parameter, the lateral or frontal features will be focused on in the recognition process. Therefore, the proposed algorithm can recognize human gait features with any walking direction, while the other methods can only be applied in one direction.

Acknowledgment

This research was supported by the MKE (Ministry of Knowledge and Economy), Korea, under the U-Port ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment) (IITA-2009-C1090-0801-0004)

References


Real-time Recognition of Humans By Their Walk

Figure 1  Proposed algorithm overview

Figure 2  Object boundary detection

Figure 3  Updated Background Image frame by frame

Figure 4  Extracted Silhouette

Figure 5  Edge detection from the binary object, a) source image, b) erosion image, c) dilation image, d) result (edge) image.

Figure 6  Polygon detection method

Figure 7  Object boundary detection. In left is the source edge object and the right is the detected boundary with 14 vertexes.
Figure 8  Waking human body analyzing for feature extraction

Figure 9  Body parts separation

Figure 10  Samples in HumanID Gait Database: (a). Video 1; (b) Video 2 and (c) Video 3

Figure 11  Confusion Matrix Illustrating Possible Detection Results

Figure 12  Human Object Detection using ROC Curves
Table 3  Outputs of NNs for Recognizing Pattern 1012

<table>
<thead>
<tr>
<th>NNs Output Neuron</th>
<th>Output of NNs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.025</td>
</tr>
<tr>
<td>2</td>
<td>0.045</td>
</tr>
<tr>
<td>3</td>
<td>0.053</td>
</tr>
<tr>
<td>4</td>
<td>0.987</td>
</tr>
<tr>
<td>5</td>
<td>0.012</td>
</tr>
<tr>
<td>6</td>
<td>0.998</td>
</tr>
<tr>
<td>7</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Note
### Table 1  Twelve Probe Sets for Challenge Experiments

<table>
<thead>
<tr>
<th>Experiment (Probe)</th>
<th>Probe Size</th>
<th>Difference between Gallery and Probe Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>120</td>
<td>View</td>
</tr>
<tr>
<td>B</td>
<td>54</td>
<td>Shoe</td>
</tr>
<tr>
<td>C</td>
<td>54</td>
<td>View and Shoe</td>
</tr>
<tr>
<td>D</td>
<td>121</td>
<td>Surface</td>
</tr>
<tr>
<td>E</td>
<td>60</td>
<td>Surface and shoe</td>
</tr>
<tr>
<td>F</td>
<td>121</td>
<td>Surface and View</td>
</tr>
<tr>
<td>G</td>
<td>60</td>
<td>Surface, Shoe and View</td>
</tr>
<tr>
<td>H</td>
<td>120</td>
<td>Briefcase</td>
</tr>
<tr>
<td>I</td>
<td>60</td>
<td>Briefcase and Shoe</td>
</tr>
<tr>
<td>J</td>
<td>120</td>
<td>Briefcase and View</td>
</tr>
<tr>
<td>K</td>
<td>33</td>
<td>Time, Shoe and Clothing</td>
</tr>
<tr>
<td>L</td>
<td>33</td>
<td>Time, Shoe, Clothing and Surface</td>
</tr>
</tbody>
</table>

### Table 2  Output Results of NNs in Training Approach of Some Data

<table>
<thead>
<tr>
<th>Person NO.</th>
<th>Input Vector</th>
<th>Output of NNs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vfeature1</td>
<td>0.966 0.021 0.010 0.010 0.026 0.010</td>
</tr>
<tr>
<td></td>
<td>Vfeature2</td>
<td>0.970 0.010 0.017 0.021 0.010 0.015</td>
</tr>
<tr>
<td></td>
<td>Vfeature12</td>
<td>0.973 0.016 0.020 0.021 0.012 0.010</td>
</tr>
<tr>
<td>8</td>
<td>Vfeature1</td>
<td>0.014 0.967 0.020 0.010 0.010 0.020</td>
</tr>
<tr>
<td></td>
<td>Vfeature2</td>
<td>0.010 0.935 0.010 0.010 0.010 0.020</td>
</tr>
<tr>
<td></td>
<td>Vfeature12</td>
<td>0.018 0.976 0.010 0.010 0.010 0.010</td>
</tr>
<tr>
<td>9</td>
<td>Vfeature1</td>
<td>0.010 0.010 0.975 0.010 0.012 0.016</td>
</tr>
<tr>
<td></td>
<td>Vfeature2</td>
<td>0.018 0.019 0.976 0.010 0.010 0.010</td>
</tr>
<tr>
<td></td>
<td>Vfeature12</td>
<td>0.012 0.015 0.979 0.010 0.022 0.010</td>
</tr>
<tr>
<td>15</td>
<td>Vfeature1</td>
<td>0.020 0.010 0.010 0.990 0.010 0.010</td>
</tr>
<tr>
<td></td>
<td>Vfeature2</td>
<td>0.014 0.010 0.014 0.985 0.010 0.010</td>
</tr>
<tr>
<td></td>
<td>Vfeature12</td>
<td>0.019 0.010 0.010 0.972 0.010 0.010</td>
</tr>
<tr>
<td>17</td>
<td>Vfeature1</td>
<td>0.012 0.010 0.017 0.010 0.980 0.010</td>
</tr>
<tr>
<td></td>
<td>Vfeature2</td>
<td>0.022 0.010 0.013 0.010 0.971 0.010</td>
</tr>
<tr>
<td></td>
<td>Vfeature12</td>
<td>0.020 0.010 0.010 0.985 0.010 0.010</td>
</tr>
<tr>
<td>22</td>
<td>Vfeature1</td>
<td>0.023 0.027 0.010 0.010 0.010 0.975</td>
</tr>
<tr>
<td></td>
<td>Vfeature2</td>
<td>0.010 0.010 0.011 0.010 0.010 0.977</td>
</tr>
<tr>
<td></td>
<td>Vfeature12</td>
<td>0.012 0.010 0.014 0.010 0.010 0.990</td>
</tr>
<tr>
<td>30</td>
<td>Vfeature1</td>
<td>0.032 0.042 0.098 0.022 0.015 0.095</td>
</tr>
<tr>
<td></td>
<td>Vfeature2</td>
<td>0.020 0.030 0.021 0.035 0.062 0.985</td>
</tr>
<tr>
<td></td>
<td>Vfeature12</td>
<td>0.023 0.034 0.045 0.036 0.032 0.995</td>
</tr>
</tbody>
</table>
Table 4  Object detection and extracted features

<table>
<thead>
<tr>
<th>Source Image</th>
<th>Silhouette</th>
<th>Object Modeling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5  Recognition Rate for Human Gait Recognition

<table>
<thead>
<tr>
<th>Rank</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>Avg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>122</td>
<td>54</td>
<td>54</td>
<td>121</td>
<td>60</td>
<td>60</td>
<td>120</td>
<td>60</td>
<td>120</td>
<td>60</td>
<td>33</td>
<td>33</td>
<td>-</td>
</tr>
<tr>
<td>Baseline</td>
<td>73</td>
<td>44</td>
<td>48</td>
<td>32</td>
<td>22</td>
<td>85</td>
<td>32</td>
<td>84</td>
<td>25</td>
<td>82</td>
<td>12</td>
<td>13</td>
<td>57.680</td>
</tr>
<tr>
<td>HMM</td>
<td>89</td>
<td>46</td>
<td>50</td>
<td>41</td>
<td>33</td>
<td>86</td>
<td>34</td>
<td>87</td>
<td>31</td>
<td>86</td>
<td>15</td>
<td>20</td>
<td>59.352</td>
</tr>
<tr>
<td>IMED</td>
<td>75</td>
<td>45</td>
<td>47</td>
<td>40</td>
<td>23</td>
<td>84</td>
<td>32</td>
<td>79</td>
<td>32</td>
<td>84</td>
<td>14</td>
<td>18</td>
<td>59.874</td>
</tr>
<tr>
<td>IMED+LDA</td>
<td>88</td>
<td>45</td>
<td>46</td>
<td>39</td>
<td>24</td>
<td>78</td>
<td>33</td>
<td>80</td>
<td>35</td>
<td>88</td>
<td>16</td>
<td>17</td>
<td>61.546</td>
</tr>
<tr>
<td>LDA</td>
<td>89</td>
<td>46</td>
<td>47</td>
<td>40</td>
<td>26</td>
<td>79</td>
<td>36</td>
<td>85</td>
<td>36</td>
<td>89</td>
<td>20</td>
<td>19</td>
<td>63.950</td>
</tr>
<tr>
<td>LDA+Sync</td>
<td>83</td>
<td>48</td>
<td>48</td>
<td>41</td>
<td>34</td>
<td>80</td>
<td>42</td>
<td>84</td>
<td>36</td>
<td>90</td>
<td>21</td>
<td>20</td>
<td>65.517</td>
</tr>
<tr>
<td>2DLDA</td>
<td>91</td>
<td>45</td>
<td>50</td>
<td>45</td>
<td>36</td>
<td>81</td>
<td>45</td>
<td>83</td>
<td>37</td>
<td>91</td>
<td>24</td>
<td>25</td>
<td>68.234</td>
</tr>
<tr>
<td>2DLDA+LDA</td>
<td>92</td>
<td>46</td>
<td>52</td>
<td>46</td>
<td>37</td>
<td>85</td>
<td>49</td>
<td>84</td>
<td>40</td>
<td>92</td>
<td>25</td>
<td>28</td>
<td>70.637</td>
</tr>
<tr>
<td>GaitFeature+NNs</td>
<td>102</td>
<td>50</td>
<td>51</td>
<td>115</td>
<td>45</td>
<td>90</td>
<td>50</td>
<td>106</td>
<td>45</td>
<td>102</td>
<td>26</td>
<td>28</td>
<td>84.639</td>
</tr>
</tbody>
</table>