Depth video-based gait recognition for smart home using local directional pattern features and hidden Markov model

Md. Zia Uddin¹, Jeong Tai Kim² and Tae-Seong Kim³

Abstract
Gait recognition at smart home is considered as a primary function of the smart system nowadays. The significance of gait recognition is high especially for the elderly as gait is one of the basic activities to promote and preserve their health. In this work, a novel method was proposed for human gait recognition by processing depth videos from a depth camera. The gait recognition method utilizes local directional patterns (LDPs) for local feature extraction from depth silhouettes and hidden Markov models (HMMs) for recognition. The LDP features were first extracted from the depth silhouettes of a human body from each frame of a video containing human gait. The dimension of the LDP features was reduced by principal component analysis. Then, each HMM was trained using the LDP features. Finally, the recognition was done with a maximum likelihood calculation of the trained HMMs of different gaits. We focused on training and recognizing two kinds of gaits here, namely, normal and abnormal. The proposed approach shows superior recognition performance over other traditional methods of gait recognition.

Keywords
Gait recognition, Depth information, Local directional pattern (LDP), Hidden Markov model (HMM)

Accepted: 14 January 2014

Introduction
Recently, smart home is attracting a lot of attentions by many researchers to promote peoples' health.¹⁻⁵ It can play a key role to support the elderly or disabled people and to support their life and health requirements. For instance, it can support the disabled people with the help of a smart wheel chair, rehabilitation robots, synthetic voice generation for control and command, etc. It can support the visually impaired people with the help of smart remote controllers or audible beacons. It can also support the hearing-impaired people using smart alarms or display screens. Smart home can monitor peoples' lifestyle by means of collecting data via sensors such as video sensors. Smart home can also provide comfort via environmental controls using intelligent environmental control devices for comfortable lighting, air conditioning, ventilation, etc.

Among viable recognition of various human activities at smart homes, gait recognition is considered to be a primary function nowadays.⁶ In this regard, video sensors or cameras are commonly being used over various motion sensors which require body attachment.⁷ Figure 1 shows a conceptual setup of a smart room that uses a video-based gait recognition system where a video camera is installed to obtain continuous time-series video frames of gaits which are analysed for gait recognition.
recognition. When the gait begins to deviate from the normal, the system can generate an alarm to warn the user or a healthcare provider.

To represent the human body in video, binary body silhouettes are the most commonly employed from which useful features are derived. For instance, Agarwal and Triggs used binary silhouettes for human body pose recovery. Howe proposed a binary silhouette-based approach to analyse various human movements. However, binary silhouettes are not efficient enough to represent human activities properly in the videos because of their limited white or black flat pixel intensity distribution. Furthermore, it is not possible to obtain the difference between the far and near parts of human body from the binary silhouettes. To overcome such limitations of the binary representation of human body silhouettes, a depth information-based representation was proposed. In the depth-based representation, the pixel values reflect distance information from an object to the camera: hence, it can provide better activity information than the binary silhouettes. Therefore, depth silhouettes should allow one to come up with more efficient activity recognition than the binary silhouette-based recognition. For instance, depth information was successfully used in gesture recognition.

Regarding the silhouette features, the most common feature extraction technique applied in video-based human activity analysis is principal component analysis (PCA). However, as it extracts global features of human body silhouettes of gait activity, one can obtain more robust human activity features if one uses local feature extraction approaches such as independent component analysis (ICA) and local binary pattern (LBP). ICA is a higher order statistical approach than PCA and it produces local features that basically resulted in an improved recognition performance over PCA for human activity recognition. LBP was introduced originally for the purpose of texture analysis. The LBP features represent more detailed local features than ICA as they focus on extracting pixel features independent of illumination. The LBP features seem to be more robust due to their tolerance against illumination changes and their computational simplicity. Later, in the study by Jabid et al., the improved LBP was renamed as local directional pattern (LDP) and it was used to represent local facial features in the analysis of facial expression images. The LDP features also have the tolerance against illumination changes like LBP, but they become more robust features than LBP as they include the gradient information for each pixel. Thus, in this study, we have adopted the LDP features to depth silhouettes of human gait for more robust gait recognition. For gait modelling and recognition, we have used hidden Markov models (HMMs) that have been used effectively in many time-sequential modelling and recognition works.

In this work, we propose a depth silhouette-based gait recognition system and show superiority over the traditional recognition system using the LDP features with HMM. To extract the depth silhouette features from the gait video images, LDP was used which was further extended by PCA. Then, HMMs were used to model time-sequential information and recognize normal and abnormal gaits. Our proposed system was compared against the traditional binary and depth silhouette-based gait recognition systems.

**Proposed gait recognition methodology**

We start our gait recognition method with the processing of depth silhouettes from the time-sequential gait video images. First, gait activity features were extracted via LDP and PCA from the depth silhouettes. To model or train different activities, discrete symbol sequences were generated from the features using vector quantization and applied on the corresponding gait HMM. To recognize gait in an image sequence, the testing symbol sequence was applied to all trained HMMs and one was chosen with the highest likelihood. Figure 2 shows the basic steps of training and testing of gaits through HMMs.

**Depth silhouette acquisition**

The depth images of different gaits were acquired by a depth camera that reflect the range of each pixel in the
scene to the camera as a greyscale value such as the near ranged pixels have bright and the far ones dark intensity values (i.e. the higher pixel intensity indicates the near distance and the lower the far). Figure 3(a) and (b) shows 10 generalized depth silhouettes from normal and abnormal gaits, respectively.

**LDP features**

The LDP assigns an eight-bit binary code to each pixel of an input depth image. This pattern is calculated by comparing the relative edge response values of a pixel in eight different directions. Kirsch, Prewitt and Sobel edge detectors are some representative edge detectors that can be used for edge responses. Amongst which, the Kirsch edge detector detects the edges more accurately than the others as it considers all eight neighbours. Given a central pixel in the image, the eight directional edge response values \( \{m_k\}, k = 0, 1, \ldots, 7 \) are computed by the Kirsch masks \( M_k \) in eight different orientations centered on its position. Figure 4 shows these masks.

The presence of a corner or an edge represents high response values in some particular directions, and therefore, it indicates the \( p \) most prominent directions in order to generate LDP. Here, the top-\( p \) directional bit responses \( b_k \) were set to 1. The remaining bits of eight-bit LDP pattern were set to 0. Finally, the LDP code was derived by equation (1). Figure 5 shows the mask response as well as the LDP bit positions and Figure 6 an exemplary LDP code with \( d = 4 \). In our work, we adopted the same value of \( d \) as it was noticed during the gait recognition experiments that the greater value of \( d \) than 4 could not improve the recognition performance

\[
LDP_p = \sum_{k=0}^{7} B_k(m_k - m_p) \times 2^k, \quad B_k(a) = \begin{cases} 1 & a \geq 0 \\ 0 & a < 0 \end{cases}
\]

(1)
where \( m_p \) is the \( p \)-th most significant directional response.

Thus, a depth silhouette image was transformed to the LDP map using the LDP code. The image textual feature is presented by the histogram of the LDP map of which the \( q \)-th bin can be defined by equation (2) as

\[
T_q = \sum_{x,y} I(LDP(x,y) = q), \quad q = 0, 1, \ldots, 1
\]  

Figure 4. Kirsch edge masks in eight directions.

where \( n \) is the number of the LDP histogram bins (normally \( n = 256 \)) for an image \( I \). Then, the histogram of the LDP map is presented by equation (3) as

\[
H = (T_0, T_1, \ldots, T_{n-1})
\]

To describe the LDP features, a depth silhouette image is divided into non-overlapping rectangle regions and the histogram is computed for each region. Furthermore, the whole LDP feature \( F \) is expressed as a concatenated sequence of histograms, represented by equation (4)

\[
F = (H^3, H^2, \ldots, H^s)
\]

where \( s \) represents the number of non-overlapped regions in the image. Figure 7 shows a gait depth image is divided into 64 small regions from which LDP histograms are extracted and concatenated into the LDP descriptor.
After obtaining the LDP silhouette descriptors, we proceeded to the dimension reduction process as the LDP features for a silhouette is still with a high dimension. The most widely used technique for dimension reduction is PCA that identifies the directions of maximum variation in data. After extracting the top \( O \) principal components (i.e. \( E_O \)), the depth silhouette representation in the PCA feature space becomes, as equation (5)

\[
P_i = F_i E_O
\]  

where \( P_i \) is the PCA projection of the LDP features of the \( i \)th depth silhouette. Figure 8 shows the top 20 eigenvalues corresponding to the eigenvectors (PCs) after applying PCA on the covariance matrix of all the gait depth silhouette LDP feature vectors. Basically, the eigenvector (PC) corresponding to the largest eigenvalue indicates the axis of the largest variance and the next one indicates the second largest variance and so on. Typically, the eigenvalues close to zero values carry negligible variance and hence can be excluded in this regard. For our experiments, we considered 50 PCs after applying PCA over the training database of gait activities. More details regarding PCA are available in earlier works.\(^{10,11,13}\)

Figure 6. LDP code. LDP: local directional pattern.

Figure 7. A gait image is divided into small regions from which LDP histograms are extracted and concatenated into LDP descriptor.

Figure 8. Top 20 eigenvalues corresponding to the eigenvectors (PCs).

**Gait modelling, training, and recognitions**

Before applying HMM, it is necessary to symbolize the gait features. Hence, a codebook was generated using vector quantization from the training feature vectors. With the extracted features via LDP on each depth silhouette, the next step was to symbolize the features. This step was used to group the features so that a discrete set of clustered features was used to represent each gait. Here, the vector quantization algorithm of Linde et al.\(^ {21} \) was used to generate a codebook that helps to generate the discrete symbols from the gait depth silhouettes. The codebook size was empirically determined as 32, and hence, the possible symbol for a gait depth silhouette was in the range from 1 to 32.

Figure 9 shows the basic steps for the codebook generation and symbol selection, respectively. Figure 10 represents the symbols for a sample testing feature vector sequence of different gaits where the sequences follow separate patterns though the same symbols can be shared by different gait feature vectors in some cases. A HMM can be denoted as \( M = (\Xi, \pi, A, B) \) where \( \Xi \) represents the states, \( \pi \) initial probability of the states, \( A \) state transition probability matrix, and \( B \) observation symbols’ probability matrix. HMMs have been applied in various applications successfully such as activity\(^ {10–13} \) and speech recognition.\(^ {19} \) More details
regarding HMMs can be obtained from Lawrence and Rabiner.\textsuperscript{19}

In training HMM, each gait was represented by a distinctly trained HMM. For \( N \) activities, there would be a set of \( N \) trained HMMs. Figure 11 represents the HMM structure used in this work and the transition probabilities of a trained abnormal gait HMM using the proposed LDP features. To recognize a gait in a video, a symbol sequence obtained from the testing gait image sequence was applied on all the trained HMMs to calculate the likelihoods and one HMM is chosen with the highest likelihood.

**Experimental results of gait recognition**

Our depth gait database was built for two gait activities (i.e. normal and abnormal gaits) where each clip contained variable number of frames. To train and test each gait model, 15 and 40 image sequences with the length of 50 were applied, respectively. Here, the traditional as well as the proposed LDP-based silhouette

---

**Figure 9.** Steps for (a) codebook generation and (b) symbol selection. LDP: local directional pattern; PCA: principal component analysis.

**Figure 10.** The symbols from codebook for a test feature vector sequence of normal and abnormal gaits.

**Figure 11.** The HMM structure and the transition probabilities of an abnormal gait HMM using LDP features (a) before and (b) after training.
feature extraction approaches were examined successively. The traditional silhouette-based recognition was applied first. Hence, PCA, ICA, and LBP were utilized to evaluate their performances on the depth silhouette-based gait recognition.

We started with the binary silhouette-based gait recognition experiments. In this regard, LBP and LDP could not be applied as the binary silhouettes only represent a flat binary pixel valued distribution. First of all, PCA with HMM were applied first and obtained the mean recognition rate of 78.75% which was the lowest recognition performance among all experiments. Then, ICA with HMM were applied and obtained a little improved recognition rate of 81.25%. Table 1 represents the binary silhouette-based gait recognition experiments.

After the binary silhouette-based gait recognition experiments, we continued to the depth silhouette-based experiments. In this regard, PCA with HMM were applied and a mean recognition rate of 81.25% was achieved. Then, to improve the recognition performance, ICA was employed and a slightly improved recognition rate of 83.75% was obtained. Since LBP is a powerful tool to obtain the robust local features, LBP was applied then and a better recognition rate of 87.50% over others was achieved. As PCA focuses on projecting data based on the covariance than LBP alone, PCA on LBP features should generate improved results for gait recognition. Thus, we applied LBP-PCA features and achieved an improved recognition rate of 90%. Then, we continued to apply the proposed LDP-PCA features for gait recognition. Although ICA and LBP also focus on local features, the LDP-based features focus on the local features considering the gradient information of the depth silhouettes, providing the possibility of better discrimination over normal and abnormal gait patterns. By default, a histogram of 256 bins was applied for LDP. After obtaining the features by concatenating the histograms for each of the 64-segmented regions of every depth image of the training gait database, it was noticed that some positions were never used and hence they were ignored during the experiments. As aforementioned, the PCA features basically find the direction of the high covariance lies in the input data and hence PCA on LDP should produce more robust features than the other approaches for gait recognition. Thus, LDP-PCA was applied with HMM and it achieved the highest recognition rate of 96.25% that indicates the superiority over the traditional approaches.

Table 1. Gait recognition results based on the different features from the binary silhouettes.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Gait</th>
<th>Recognition rate (%)</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA</td>
<td>Normal</td>
<td>80</td>
<td>78.75</td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>77.50</td>
<td></td>
</tr>
<tr>
<td>ICA</td>
<td>Normal</td>
<td>82.50</td>
<td>81.25</td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>80</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Gait recognition results based on the different features from the depth silhouettes.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Gait</th>
<th>Recognition rate (%)</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA Normal</td>
<td>80</td>
<td>81.25</td>
<td></td>
</tr>
<tr>
<td>PCA Abnormal</td>
<td>82.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICA Normal</td>
<td>82.50</td>
<td>83.75</td>
<td></td>
</tr>
<tr>
<td>ICA Abnormal</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBP Normal</td>
<td>90</td>
<td>87.50</td>
<td></td>
</tr>
<tr>
<td>LBP Abnormal</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBP-PCA Normal</td>
<td>90</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>LBP-PCA Abnormal</td>
<td>97.50</td>
<td>96.25</td>
<td></td>
</tr>
</tbody>
</table>

PCA: principal component analysis; ICA: independent component analysis; LBP: local binary pattern; LDP: local directional pattern.

Conclusion
In this paper, a novel approach has been proposed for human gait recognition for smart home by using the
depth silhouettes and their LDP features in combination with HMM. The LDP-based features of the depth silhouettes were investigated and produced superior recognition results over the traditional silhouette features from PCA, ICA, and LBP. The proposed system can be employed in smart homes for better human gait recognition and services.

**Authors’ contribution**

Md. Zia Uddin provided original ideas of using local directional pattern features and their use in Hidden Markov Model to recognize gaits of residents at smart home. Jeong-Tai Kim provided helpful comments and valuable discussions on the use of the proposed methodology at smart home. Tae-Seong Kim provided technical comments and discussions on the implementation of the proposed methodology and its application to smart home.

**Funding**

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIP) (No. 2008-0061908).

**Conflict of Interest**

None declared.

**References**