CHAPTER 4
HIDDEN MARKOV MODEL FOR GAIT RECOGNITION

4.1 INTRODUCTION

The walking of a person is considered to be one which is perpendicular to the direction of walk. The walking style (GAIT) feature vector is represented by the following properties; the ability to describe appearance at a level finer than whole body description, without having to segment individual limbs; robustness to noise in video foreground segmentation; and simplicity of representation.

Additional feature that can be considered is the height of an individual which requires calibrating the camera to recover distances, the amount of bounce of the whole body in a full stride, the side-to-side sway of the torso, the maximum distance between the front and the back legs at the peak of the swing phase of a stride, the amount of arm and leg swing.

These features may or may not be used for various reasons such as inaccessibility (the side-to-side sway of torso) or difficulties in obtaining features such as detecting the peaks of swing phase when foreground segmentation is noisy and includes shadows.

Fig.4.1 The silhouette of a foreground walking person is divided into 7 regions, and ellipses are fitted to each region.
The gait appearance feature vector comprises parameters of moment features in image regions, containing the walking person aggregated over time either by averaging or spectral analysis. For each segmented image (Figure 4.1), the centroid is found and the segmented silhouette is divided into 7 parts as in Figure 4.1(b).

The vertical line through the centroid is used to divide the silhouette into front and back sections except for the top portion. The parts above and below the centroid are each equally divided in the horizontal direction resulting in 7 regions that roughly correspond to:

- r1, head / shoulder region;
- r2, front of torso;
- r3, back of torso;
- r4, front thigh;
- r5, back thigh;
- r6, front calf/foot; and
- r7, back calf/foot.

For each of the 7 regions from a silhouette, an ellipse is fit to the portion of foreground object visible in that region (Figure 4.1(b)). The features extracted from each of these regions are the centroid, aspect ratio (l) of major and minor axis of the ellipse and the orientation (α) of major axis of the ellipse, i.e., the region feature vector \( f(r_i) \) is,

\[
f(r_i) = (\text{mean}(x)_i, \text{mean}(y)_i, l_i, \alpha_i), \quad \text{where } i = 1, \ldots, 7. \quad (4.1)
\]

These moment-based features are robust to noise in the silhouettes obtained from background subtraction, as long as the number of noise pixels is small and not systematically biased. The features extracted from each frame of a walking sequences consists of features from each of the 7 regions, i.e. the frame feature vector \( F_j \) of the \( j^{th} \) frame is,

\[
F_j = (f(r_1), \ldots, f(r_7)). \quad (4.2)
\]
In addition to these 28 features, additional feature \( h \), the height (relative to body length) of the centroid of the whole silhouette is used to describe the proportions of the torso and legs. The intuition behind this measure is that an individual with longer torso will have a silhouette centroid that is positioned higher (relative to body length) on the silhouette than someone with a short torso.

Given the region features across a gait sequence, a concise representation across time is required. Two types of features across time are computed: 1) the mean and standard deviation of region features across time and 2) magnitudes and phases of each region feature related to the dominant walking frequency (of 1 step or half of a stride).

The gait average appearance feature vector of a sequence \( s \) is,

\[
s = (\text{mean}_j(h_j), \text{mean}_j(F_j), \text{std}_j(F_j))
\]

where

\( j = 1, \ldots, \text{last frame} \), and \( s \) is 57-dimensional. This feature set is very simple to compute and robust to noisy foreground silhouettes.

The mean features describe the average-looking ellipses for each of the 7 regions of the body; taken together, the 7 ellipses describe the average shape of the body.

The standard deviation features roughly describe the changes in the shape of each region caused by the motion of the body, where the amount of change is affected by factors such as how much one swings his arms and legs.

The gait spectral component feature vector of a sequence is,

\[
t = (\Omega_d, |X_i(\Omega_d)|, \text{phase}(X_i(\Omega_d)))
\]
where

\[ X_i = \text{Fourier Transform } (F_j = 1 \ldots \text{last}(f(r_j))), \]  

(4.5)

\[ \Omega_d \] is the dominant walking frequency of a given sequence.

The magnitude feature components measure the amount of change in each of the 7 regions due to motion of the walking body.

The phase components measure the time delay between the different regions of the silhouette. Because of noise in silhouettes, the time series of region features is also noisy. The power spectra of many of the region features do not show an obvious dominant peak that is the dominant walking frequency. Even when the peak frequencies are found, they often do not agree between different region features. A normalized averaging of power spectra of all region features resulting in a much more dominant peak frequency that are also consistent across all signals are used. The magnitude of each region feature at the dominant frequency \( \Omega_d \) can be directly used, but the phase cannot be directly used because each gait sequence is not predetermined to start at a particular point of a walking cycle. The phases of all region features relative to one particular region feature that is “most stable,” is computed by which the 2\textsuperscript{nd} moment of power spectrum about the peak frequency to measure. The standard phase is that of the x position of the front is the calf/foot region. The gait spectral component feature vector has 57-1(the standard phase) = 56 dimensions.

### 4.2 Principles of Transformation

The transformation of a set of n-dimensional real vectors onto a plane is called a mapping operation. The result of this operation is a planar display. The mapping operation can be linear or non-linear. Fisher, 1936, developed a linear classification algorithm.
Hong and Yang, 1991, implemented a method for constructing a classifier on the optimal discriminant plane by using minimum distance criterion for multiclass classification for less of patterns. Foley, 1972, used the method of considering the number of patterns and feature size.

Siedlecki et al., 1988, gave an overview of mapping techniques. The mapping of the original vector ‘X’ onto a different vector ‘Y’ on a plane is completed by a matrix transformation, which is given by equation (4.6)

\[ Y = AX + b \]  

(4.6)

Where

\[ A = \begin{bmatrix} \varphi_1^T \\ \varphi_2^T \end{bmatrix} \]  

(4.7)

\( b \) is a 2-dimensional vector, \( \varphi_1 \) is a projection vector (also called a discriminant vector) and \( \varphi_2 \) is another projection vector (also called a discriminant vector).

The 2-dimensional vector does not introduce any relevant information, but it is given for the generality of the equation 4.7. The steps involved in the linear mappings are as follows:

**Step 1:** Computation of the Discriminant vectors \( \varphi_1 \) and \( \varphi_2 \). This is exact for a particular linear mapping algorithm.

**Step 2:** Computation of the planar images of the original data points; this is common for all linear mapping algorithms.

Some of the linear mapping algorithms are

i) Principal component mapping, Kittler and Young, 1973.


iii) Least squared error mapping, Mix and Jones, 1982, and
iv) Projection pursuit mapping, Friedman and Turkey, 1974.

Out of the above four methods, generalized declustering optimal Discriminant plane based on a mapping technique of Fisher is used. The vectors \( \varphi_1 \) and \( \varphi_2 \) are discriminant vectors. The plane formed by them is the discriminant plane, which is optimal.

### 4.2.1 Computation of discriminant vectors \( \varphi_1 \) and \( \varphi_2 \)

The Fisher’s criterion is given by

\[
J (\varphi) = \frac{\varphi^T S_b \varphi}{\varphi^T S_w \varphi} \quad \text{and} \quad (4.8)
\]

\[
S_b = \sum_{i=1}^{m} P(\omega_i) (m_i - m_o) (m_i - m_o)^T \quad (4.9)
\]

\[
S_w = \sum_{i=1}^{m} P(\omega_i) E \left[ (x_i - m_i)(x_i - m_i)^T \right] / \omega_i \quad (4.10)
\]

Where

- \( S_b \) is the between class matrix, and
- \( S_w \) is the within class matrix which is non-singular.
- \( P (\omega_i) \) is a prior the probability of the \( i^{th} \) pattern, generally \( P (\omega_i) = 1/m \).
- \( m_i \) is the mean vector of the \( i^{th} \) class patterns, \( i=1, 2, \ldots, m \);
- \( m_o \) is the global mean vector of all the patterns in all the classes.
- \( X = \{x_i, i=1, 2, \ldots, L\} \) is the \( n \)-dimensional patterns of each class.

The discriminant vector that maximizes, \( J \), in equation 4.8 is denoted by \( \varphi_1 \). The vector \( \varphi_1 \) is found as a solution of the Eigen value problem given by equation (4.11):

\[
S_b \varphi_1 = \lambda m_1 S_w \varphi_1 \quad (4.11)
\]

Where

- \( \lambda m_1 \) is the greatest non-zero Eigen value of \( S_b S_w^{-1} \).
The eigenvector corresponding to $\lambda_{m1}$ is $\varphi_1$. Another discriminant vector $\varphi_2$ is obtained by using the same criterion of equation 4.8. The vector $\varphi_2$ should satisfy the equation (4.12).

$$\varphi_2^T \varphi_1 = 0.0$$  \hspace{1cm} (4.12)

The equation 4.12 indicates that the solution obtained is geometrically independent. The discriminant vector $\varphi_2$ is found as a solution of the Eigen value problem, which is given by equation (4.13).

$$Q_p S_b \varphi_2 = \lambda_{m2} S_w \varphi_2$$  \hspace{1cm} (4.13)

Where

$\lambda_{m2}$ is the greatest non-zero eigenvalue of $Q_p S_b S_w^{-1}$ and

$Q_p$ is the projection matrix given by equation (4.14)

$$Q = I - \frac{\varphi_1 \varphi_1^T S_w^{-1}}{\varphi_1^T S_w^{-1} \varphi_1}$$  \hspace{1cm} (4.14)

Where

$I$ is an identity matrix.

In equation 4.11, $S_w$ should be non-singular. It is necessary that, $S_w$ should be non-singular even for a more general discriminating analysis and generating multi-orthonormal (Foley and Sammon, 1975; Liu et al., 1992; and Cheng et al., 1992) vectors, If $S_w$ is singular, $S_w$ should be made non-singular by using singular value decomposition (SVD) method and by perturbing the matrix. Klema and Laub, 1980, discussed the SVD and the method of its computation. Sullivan and Liu, 1984, explained the SVD with application to signal processing. By using equations (4.11 and 4.13), the values of $\varphi_1$ and $\varphi_2$ discriminant vectors are obtained and they are given by:
Fig. 4.2 Plot of discriminant vectors for 5 persons

Figure 4.2 plots 5 different persons based on the FLD output. In this figure, the gait is cluttered for persons 1, 2, 3 and scattered for persons 4 and 5.
4.2.2 Computation of 2-dimensional vector from the original n-dimensional vector

The 2-dimensional vectors set is denoted by $Y_i$. The vector $Y_i$ is given by equation (4.15):

$$Y_i = (u_i, v_i) = \{X_i^T \varphi_1, X_i^T \varphi_2 \} \quad (4.15)$$

The vector set $Y_i$, is obtained by projecting the original vector 'X' of the patterns onto the space spanned by $\varphi_1$ and $\varphi_2$ by using equation 4.15.

Let $X \in \mathbb{X}$ belongs to the $K^{th}$ class. The corresponding 2-dimensional sets of vectors are:

$$y^{(k)} = (y_1^{(k)}, y_2^{(k)}, \ldots, y_L^{(k)}) \quad (4.16)$$

$$\sum_{i=1}^{L_k} I_i - L$$

Where

$$y_i^{(k)} = (u_i^{(k)}, v_i^{(k)}), \quad i = 1, 2, 3, \ldots, L_k \quad (4.17)$$

The 2-dimensional vectors for the training patterns are given in Table 4.1.

<table>
<thead>
<tr>
<th>Pattern No.</th>
<th>U</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.170027</td>
<td>0.26589654</td>
</tr>
<tr>
<td>2</td>
<td>0.186384</td>
<td>0.31478438</td>
</tr>
<tr>
<td>3</td>
<td>0.275119</td>
<td>0.15631626</td>
</tr>
<tr>
<td>4</td>
<td>0.207705</td>
<td>0.093517463</td>
</tr>
<tr>
<td>5</td>
<td>0.18044</td>
<td>0.27104327</td>
</tr>
<tr>
<td>6</td>
<td>0.180694</td>
<td>0.25908034</td>
</tr>
<tr>
<td>7</td>
<td>0.139818</td>
<td>0.2674465</td>
</tr>
<tr>
<td>8</td>
<td>0.148349</td>
<td>0.30747598</td>
</tr>
<tr>
<td>9</td>
<td>0.183</td>
<td>0.27899469</td>
</tr>
<tr>
<td>10</td>
<td>0.163267</td>
<td>0.27702721</td>
</tr>
<tr>
<td>11</td>
<td>0.145098</td>
<td>0.22248487</td>
</tr>
<tr>
<td>12</td>
<td>0.135343</td>
<td>0.26594515</td>
</tr>
<tr>
<td>13</td>
<td>0.175656</td>
<td>0.26430668</td>
</tr>
<tr>
<td>14</td>
<td>0.171221</td>
<td>0.24790403</td>
</tr>
<tr>
<td>15</td>
<td>0.191819</td>
<td>0.24980761</td>
</tr>
</tbody>
</table>
4.3 SUMMARY

This chapter has presented the implementation of hidden Markov model (HMM) for extracting the features representing gait of persons. The transformation of 56 dimensional vector into 2 dimensional vector is presented using FLD. Chapter 5 presents the method of implementing the radial basis function in training and testing the HMM features of gait for different people.