Change Energy Image for Gait Recognition: An Approach Based on Symbolic Representation

Mohan Kumar H P
Dept of MCA, PES College of Engineering, Mandya, Karnataka, India-571401
email: mohanhallegere@yahoo.com

Nagendra Swamy H S
Dept of studies in Computer Science, Manasagangothri, Mysore, Karnataka, India
email: hsnswamy@compsci.uni-mysore.ac.in

Abstract — Gait can be identified by observing static and dynamic parts of human body. In this paper a variant of gait energy image called change energy images (CEI) are generated to capture detailed static and dynamic information of human gait. Radon transform is applied to CEI in four different directions (vertical, horizontal and two opposite cross sections) considering four different angles to compute discriminative feature values. The extracted features are represented in the form of interval-valued type symbolic data. The proposed method is capable of recognizing an individual when he/she have variations in their gait due to different clothes they wear, in different normal conditions and carrying a bag. A similarity measure suitable for the purpose of establishing similarity match for gait recognition. Experiments are conducted on CASIA database B and the results have shown better recognition performance compared to some of the existing methods.

Index Terms — Change energy images, interval valued features, subject (individual person), Radon transform, representation, similarity measure.

I. INTRODUCTION

Gait recognition is a method of identifying an individual based on his/her style of walking, which is considered to be unique among individuals. Gait image sequences are captured at a distance without individual cooperation to make alert or to give early warnings when the person is far away [1, 2]. Gait recognition has gained much attention in the field of biometric research due to its ability in identifying an individual at a distance despite insufficient resolution in acquired images. Mainly gait recognition approaches are of two types, namely model-based and model-free [3]. In model-based approach modeling human body is achieved considering information such as ankle elevation, angle between joints, position of hip etc [4]. The method is found to be simple and straightforward but it is computationally expensive [5]. Model-free approach directly uses the silhouette sequences for extraction of features without any gait model. In model-free, binary gait silhouettes could be of low quality without any texture or colour information and both static (torso) and dynamic (hand and leg) features are used for gait recognition. Most of the investigations on gait recognition are based on model-free approach. Gait recognition finds its applications in the field of surveillance and other applications like robotics, clinical analysis and computer animation [1].

Several attempts have been made by many researchers to provide an efficient and effective gait recognition system. Following are the few interesting works found in the literature in this direction. Mridul Ghosh et al [6] proposed gait recognition technique, which uses corner points from edge of the image. Control points are selected from the determined corner points as features and Euclidean distance between the control points in a cyclic order are computed and stored in a database. sobel operator is used to find the edges in a gait image. Experimental results show the improvement is found in terms of recognition. Ju Han et al [7] proposed spatio-temporal representation of gait called Gait energy image for recognition of individual by gait. Change with time (temporal) information is acquired in a single image through Gait energy image. Recognition is achieved by combining statistical gait features obtained from real and synthetic poses (templates). The method is less sensitive to silhouette distortions like spurious pixels and scale changes. Experimental study shows the approach has given acceptable results. Xiaoxiang Li et al [8] have proposed a structural gait energy image (SGEI) by combining foot energy image (FEI) and Head energy image (HEI). The results show that the SGEI is better than GEI. Yumi Iwashita et al [9] extracted shadow area of a subject. Gait stripe and gait contour are the two gait features extracted by spherical harmonics and the same has been used in the reference database. A simple NN classifier using Euclidean distance measure is used for identification. Author claims that the proposed method outperforms the methods based on Fourier transform, gait energy image and active energy image. Toby et al [10] proposed a method of capturing motion part and static part such as hand, leg and torso respectively thru templates called motion silhouette contour templates (MSCT) and static silhouette template (SST). A simple
Thus the symbolic representation of a subject in the consolidated to form an interval type symbolic data. Obtained for various instances of a subject is to these images. The radon transformation values particularly for gait recognition techniques. There is a scope for much more analysis in this direction data analysis approach for gait recognition. However, reported in [25, 26] explored the suitability of symbolic [24]. It is also observed from the survey that the work explored in online signature verification and recognition and the concept of symbolic data analysis has also been representation and recognition has been explored in [23]

Copyright © 2014 MECS

II. PROPOSED METHODOLOGY

The motivation for the proposed approach is to capture energy variations in static and dynamic part in a gait sequence of a gait cycle by generating change energy images (CIE) and to extract discriminative feature values for gait recognition. Input to our proposed methodology is a sequence of silhouettes of a gait cycle. A gait cycle is a basic unit of gait and it refers to the time interval between successive instances of initial foot-to-floor contact for the same foot. We have assumed that the speed is constant within any specific subject gait sequence cycle (length of different instances of a subject like change in viewing angle, change in cloth type, carrying a bag is constant). However, speed could vary among reference (training) and probe (test) sequences. The proposed approach consists of the following steps. In the first step, sequence of change energy images (CIE) is generated from a sequence of silhouettes of a gait cycle. In the second step, Radon transform is applied to a sequence of change energy images. Standard deviation and maximum value from the radon transformed values is computed for four different angles, which covers the vertical, horizontal and two opposite cross sections (directions). In the third step, similarity between probe gait sequence and all reference gait sequences is established and the recognition task is accomplished.

A. Change Energy Images (CIE)

The CASIA dataset B, used in our experiments, contain some silhouettes with spurious pixels and small holes. Morphological operations such as erosion and dilation are applied to remove spurious pixels and holes in the silhouettes. The silhouettes are extracted by applying bounding box and then silhouette images are normalized. Change energy images from pre-processed silhouettes for each gait cycle are generated as follows. Let

\[ I = \{I_1, I_2, \cdots, I_n\} \]  

be the \( n \) silhouette images in a given gait cycle. For a gait cycle consisting of \( n \) silhouettes, a set of \( n-1 \) change energy images are obtained. Let

\[ CEI = \{CEI_1, CEI_2, \cdots, CEI_{i}, \cdots, CEI_{n-1}\} \]  

be the set of change energy images.

First change energy image \(CEI_1\) is computed as
\[
CEI_1 = \frac{1}{2} \sum (I_1(x,y) + I_2(x,y))
\]  
(3)

Where \(I_1(x,y)\) and \(I_2(x,y)\) are the intensity values of first and second silhouette images in a given gait cycle at \((x,y)\) coordinates. The Subsequent CEI is computed as
\[
CEI_i = \frac{1}{2} \sum (CEI_{i-1}(x,y) + I_{i+1}(x,y))
\]  
(4)

Where \(I_{i+1}\) is the \(i+1\)th image in \(I\) and \(CEI_{i-1}\) is the \(i-1\)th image in \(CEI\).

\[
CEI_{n-1} = \frac{1}{2} \sum (CEI_{n-2}(x,y) + I_n(x,y))
\]  
(5)

The set of change energy images \(CEI\) generated for an instance of normal walking is shown in Fig. 1.

**Figure 1. Sequence of change energy images**

**B. Feature Extraction**

The proposed feature extraction technique applies Radon transform on the CEIs obtained as discussed in the previous section. The Radon transform maps the 2-D image in Cartesian coordinate system \((x,y)\) into polar coordinate system \((\rho, \theta)\). The Radon transform applied on CEI captures the intensity variations across various angles. The more theoretical description about Radon transform is found in [27]. The radon transform is applied on the change energy images as follows
\[
R = \text{radon}(CEI, \theta)
\]  
(6)

The features chosen are maximum value \(r\) and standard deviation \(\sigma\) of Radon transformed vector \(R\) for a given angle \(\theta\). The \(\theta\) chosen here are 0, 90, 15, and 150 degree, since these angles cover the vertical, horizontal and two opposite cross sections summation line respectively in radon transform of change energy images.
\[
r = \max(R)
\]  
(7)

\[
\sigma = \text{std}(R)
\]  
(8)

Thus the features extracted for every CEI in a gait sequence are \((r_1, \sigma_1), (r_2, \sigma_2), (r_3, \sigma_3), (r_4, \sigma_4)\), which represents the maximum value \(r\) and the standard deviation \(\sigma\) of Radon transformed vector \(R\) respectively for \(\theta = 0, 15, 90, 150\).

**C. Representation**

Since the gait of a person varies slightly due to change in carrying condition, change in clothes and different normal conditions. The Radon features obtained for these instances also contain variations. These variations are handled by consolidating the features in the form of an interval type data as explained below.

Let
\[
S = \left[S_1, S_2, \ldots, S_i, \ldots, S_N\right]
\]  
(9)

be the \(N\) number of subjects.

Let
be the \( n \) instances of the subject \( S_I \).

Let
\[ s_{hi} = [\text{CEI}_1, \text{CEI}_2, \ldots, \text{CEI}_i, \ldots, \text{CEI}_T] \]  

be the change energy images for \( s_{hi} \).

Where \( T \) is the length of \( s_{hi} \) (total number of change energy images). \( T \) is same for a particular subject \( s_{hi} \) \( (i = 1,2, \ldots, n) \) but \( T \) may vary across subjects, i.e., for different subjects \( S_I \) \( (I = 1,2, \ldots, N) \).

The two feature values \( r \) and \( \sigma \) computed from \( CEI_t \) of an instance \( s_{hi} \) of subject \( S_I \) at \( m \) different angles is given by
\[
s_{hi}(t) = [r_{hi1}, \sigma_{hi1}, r_{hi2}, \sigma_{hi2}, \ldots, r_{him}, \sigma_{him}] \]  

The vector representing the features of \( s_{hi} \) is given by
\[
s_h = [T, RF_{h1}, RF_{h2}, \ldots, RF_{hn}, \ldots, RF_{ht}] \]  

Where \( RF_{hit} = s_{hi}(t) \) and \( T \) is the number of change energy images of \( s_{hi} \).

The minimum and maximum value of a feature \( r \) for \( kth \) angle of \( n \) instances of a subject \( S_I \) is given by
\[
r_{nk}^{-} = \text{Min}(r_{n1}(t), r_{n2}(t), \ldots, r_{nk}(t), \ldots, r_{nn}(t)) \]  

\[
r_{nk}^{+} = \text{Max}(r_{n1}(t), r_{n2}(t), \ldots, r_{nk}(t), \ldots, r_{nn}(t)) \]  

The minimum and maximum value of feature \( \sigma \) for
\[
\sigma_{nk}^{-} = \text{Min}(\sigma_{n1}(t), \sigma_{n2}(t), \ldots, \sigma_{nk}(t), \ldots, \sigma_{nn}(t)) \]  

\[
\sigma_{nk}^{+} = \text{Max}(\sigma_{n1}(t), \sigma_{n2}(t), \ldots, \sigma_{nk}(t), \ldots, \sigma_{nn}(t)) \]  

Thus, the reference sequence of a subject \( S_I \) \((i = 1,2, \ldots, N)\) in the knowledge base is represented in the form of interval-valued type symbolic feature vector as
\[
S_I = [T_I, RF_{I1}, RF_{I2}, \ldots, RF_{In}, \ldots, RF_{IT}] \]  

Where \( T_I \) is the number of change energy images of a subject \( S_I \) and the value of \( T_I \) could be different for different \( S_I \).

\( RF_{it} \) is the feature values of \( tth \) change energy images of all \( n \) instances consolidated in the form of interval for \( m \) different angles and is given by
\[
RF_{it} = \left\{ [r_{i1}(t), r_{i2}(t), \sigma_{i1}(t), \sigma_{i2}(t), [r_{i1}(t), r_{i2}(t), \sigma_{i1}(t), \sigma_{i2}(t), \ldots, \ldots, \ldots, [r_{in}(t), r_{in}(t), \sigma_{in}(t), \sigma_{in}(t), \ldots, \ldots, \ldots, [r_{Im}(t), r_{Im}(t), \sigma_{Im}(t), \sigma_{Im}(t)] \ldots \ldots, \ldots, [r_{IT}(t), r_{IT}(t), \sigma_{IT}(t), \sigma_{IT}(t)] \right\} \]  

Where for each \( kth \) angle \( r_{nk}^{-} \) and \( r_{nk}^{+} \) and \( \sigma_{nk}^{-} \) and \( \sigma_{nk}^{+} \) is obtained as shown in equation (14), (15), (16) and (17) respectively.

The probe (test) sequence of an instance of subject \( S_I \) to be tested is represented as crisp feature vector and is given by
\[
s_p = [T_p, PF_{I1}, PF_{I2}, \ldots, PF_{Ip}, \ldots, PF_{IT}] \]  

Where \( T_p \) is the length of the probe sequence and \( PF_{it} \) is the crisp feature values of maximum radon value and standard deviation of radon values for \( m \) different angles and is given by
\[
PF_{it} = [r_{i1}(t), \sigma_{i1}(t), r_{i2}(t), \sigma_{i2}(t), \ldots, r_{Im}(t), \sigma_{Im}(t)] \]  

Table I shows the feature values representing the \( CEI_t \) of different instances of a subject \( S_I \) and the interval-valued representation after consolidating the corresponding feature values.

### Table I. Example of Crisp Feature Values Representing CEI_t of All the Instances of a Subject and the Corresponding Consolidated Interval-Valued Features in Symbolic Representation

<table>
<thead>
<tr>
<th>Instances</th>
<th>( r )</th>
<th>( \sigma )</th>
<th>( r )</th>
<th>( \sigma )</th>
<th>( r )</th>
<th>( \sigma )</th>
<th>( r )</th>
<th>( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0(^\circ)</td>
<td>0(^\circ)</td>
<td>15(^\circ)</td>
<td>15(^\circ)</td>
<td>90(^\circ)</td>
<td>90(^\circ)</td>
<td>150(^\circ)</td>
<td>150(^\circ)</td>
</tr>
<tr>
<td>Bag1</td>
<td>30293</td>
<td>13141</td>
<td>23226</td>
<td>8190</td>
<td>8421</td>
<td>2064</td>
<td>19009</td>
<td>5415</td>
</tr>
<tr>
<td>Cloth1</td>
<td>32370</td>
<td>10972</td>
<td>24291</td>
<td>8325</td>
<td>9347</td>
<td>1981</td>
<td>19115</td>
<td>6105</td>
</tr>
<tr>
<td>Normal1</td>
<td>31002</td>
<td>10562</td>
<td>23453</td>
<td>8468</td>
<td>9003</td>
<td>1562</td>
<td>18437</td>
<td>5489</td>
</tr>
<tr>
<td>Normal2</td>
<td>31460</td>
<td>11159</td>
<td>23981</td>
<td>8392</td>
<td>9217</td>
<td>1681</td>
<td>18463</td>
<td>5568</td>
</tr>
<tr>
<td>Normal3</td>
<td>32117</td>
<td>10998</td>
<td>23775</td>
<td>8992</td>
<td>1821</td>
<td>18081</td>
<td>5871</td>
<td></td>
</tr>
<tr>
<td>Interval</td>
<td>[31460, 32370]</td>
<td>[10562, 11341]</td>
<td>[23453, 24291]</td>
<td>[8047, 8648]</td>
<td>[8421, 9347]</td>
<td>[1562, 2064]</td>
<td>[18081, 19115]</td>
<td>[5415, 6105]</td>
</tr>
</tbody>
</table>
A. Similarity Computation

In order to recognize a probe sequence $S_p$, features are extracted from the probe gait sequence as discussed in sub section B of section II and represented as shown in equation (20). An example of feature vector values (of type crisp) of $PF_t$ of probe sequence $S_p$ is shown in Table II for 4 different angles. The obtained crisp feature vector is compared with the symbolic feature vector of the reference gait sequences (equation 18) in the gait knowledge-base.

| TABLE II. EXAMPLE OF CRISP VALUES OF PROBE SEQUENCE PF T FOR 4 ANGLES |
|-----------------|-----------------|
| Features $r$ and $\sigma$ | values          |
| $n^{(0\circ)}$   | 32090           |
| $\sigma^{(0\circ)}$ | 11186           |
| $n^{(15\circ)}$  | 24002           |
| $\sigma^{(15\circ)}$ | 8412            |
| $n^{(90\circ)}$  | 9340            |
| $\sigma^{(90\circ)}$ | 1998            |
| $n^{(150\circ)}$ | 18863           |
| $\sigma^{(150\circ)}$ | 5897            |

To compute similarity between probe sequence and each reference sequence, the length of $T_p$ in the probe sequence and reference sequence $T_l$ are checked. Since the value $T_l$ may vary from subject to subject, there is a possibility that the value $T_l$ (number of change energy images of reference gait sequence of a subject) may be different. Suppose if $T_l < T_p$ then, the first $T_l$ number of CEIs of a probe sequence of a subject is compared with the reference sequence of a subject and the corresponding similarity value will be computed. If $T_l = T_p$ then, the $T_l$ number of CEIs of a reference sequence of a subject is compared with the $T_p$ number of CEIs of a probe sequence of a subject and the corresponding similarity value will be computed.

In all the above three cases similarity between probe and reference sequence of all the subjects $S_l (l = 1, 2 \ldots N)$ in the knowledge-base is computed using equation (22).

$$TotalSimilarity (s_p, s_l) = (sim_p + sim_l) \quad (22)$$

Where

$$\text{sim}_{l} = \sum_{i=1}^{n} \sum_{t=1}^{m} \left( \text{sim}(r_{l}(t), r_{l}(t)) \right) - |T_p - T_l| \quad (23)$$

for $l = 1$ to $N$

$$T = \text{Min} \left( T_p, T_l \right) \quad (24)$$

$$\text{sim}_{l} = \begin{cases} 1 & \text{if } \left( r_{l}(t) > r_{l}(t) \text{ and } r_{l}(t) \leq r_{l} \right) \\ \frac{r_{l} - r_{l}(t)}{r_{l} - r_{l}} & \text{if } \left( r_{l}(t) > r_{l}(t) \text{ and } r_{l}(t) > r_{l} \right) \\ 0 & \text{if } \left( r_{l}(t) < r_{l}(t) \text{ and } r_{l}(t) > r_{l} \right) \end{cases} \quad (25)$$

$$r_i = r_{l} - d \quad (26)$$

$$r_i = r_{l} + d \quad (27)$$

$$d = r_{l} - r_{l} \quad (28)$$

And

$$\text{sim}_{l} = \sum_{i=1}^{n} \sum_{t=1}^{m} \left( \text{sim}(r_{l}(t), r_{l}(t)) \right) - |T_p - T_l| \quad (29)$$

for $l = 1$ to $N$

$$T = \text{Min} \left( T_p, T_l \right) \quad (30)$$

$$\text{sim}_{l} = \begin{cases} 1 & \text{if } \left( r_{l}(t) > r_{l}(t) \text{ and } r_{l}(t) \leq r_{l} \right) \\ \frac{r_{l} - r_{l}(t)}{r_{l} - r_{l}} & \text{if } \left( r_{l}(t) > r_{l}(t) \text{ and } r_{l}(t) > r_{l} \right) \\ 0 & \text{if } \left( r_{l}(t) < r_{l}(t) \text{ and } r_{l}(t) > r_{l} \right) \end{cases} \quad (31)$$

$$\sigma_i = \sigma_{l} - d \quad (32)$$

$$\sigma_i = \sigma_{l} + d \quad (33)$$

$$d = \sigma_{l} - \sigma_{l} \quad (34)$$

III. EXPERIMENTS

In order to study the performance of the proposed method of gait recognition, we have conducted an experiment on the standard CASIA Dataset B. The dataset consists of 124 individuals (subjects). Each subject consists of 10 sequences, out of which 2 sequences is with carrying a bag, 2 sequences is walking wearing different clothes and 6 sequences is in normal
conditions. The gait silhouettes used are in 90 degree (side view) viewing angle as this view provides more gait information than the silhouettes taken from other view angles. In the Dataset, some subjects do not have (complete) all the sequences and some of the sequence do not have a complete gait cycle. Therefore, only 120 subjects out of 124 subjects are considered in our experiment. We have measured the performance of the proposed approach for identification using cumulative match scores (CMS) suggested in [28].

The task of identification is an attempt to find given probe sequence to be one of the reference sequences, since closed-set identification is considered, which guarantees the availability of the subject in the reference database. For identification, first sequence of carrying a bag named as B1 (bag1), first sequence of coat named as C1 (cloth1) and first three different normal walking sequences named as N1 (normal1), N2 (normal2) and N3 (normal3) are used for training and second sequence of carrying a bag named as B2 (bag2), second sequence of coat named as C2 (cloth2) and rest of the sequences of normal walking named as N4 (normal4), N5 (normal5) and N6 (normal6) are used for testing. To measure the performance of identification, reference sequences \( S_I (I = 1, 2, \ldots N) \) are sorted, based on the similarity values computed with the given probe sequence \( s_p \). If the correct reference sequence has highest similarity value with its corresponding probe sequence, then it is ranked as 1, otherwise its rank will be decided based on its position in the sorted order. We report results in terms of cumulative match scores (CMS). Table III shows the identification rate of the proposed methodology at rank 1, 5 and 10. The Cumulative Match curve for the proposed system shown in Fig. 2 represents that the performance at rank 1 is the correct classification rate (CCR). Table IV shows the average correct classification rate (CCR) at rank 1 for other approaches and the proposed approach.

Also, the performance of the proposed system is compared with NLPR database (similar covariates i.e., normal sequences without coat and bag), which consist of only 20 subjects for lateral view with 124 subjects of normal sequences of CASIA dataset B. It is found that our method for 124 subjects achieved 95% CCR for normal sequences compared with the other methods shown in Table V for only 20 subjects.

### TABLE III. IDENTIFICATION RATES AT DIFFERENT RANKS

<table>
<thead>
<tr>
<th>Probe</th>
<th>Identification rate/Rank (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>N4</td>
<td>95.83</td>
</tr>
<tr>
<td>N5</td>
<td>94.16</td>
</tr>
<tr>
<td>N6</td>
<td>95.00</td>
</tr>
<tr>
<td>C2</td>
<td>84.16</td>
</tr>
<tr>
<td>B2</td>
<td>88.33</td>
</tr>
</tbody>
</table>

### IV. CONCLUSION

In this paper, features are extracted from change energy images (CIE) a variant of gait energy image. Change energy images provide static as well as dynamic information in detail. Radon transform is applied on change energy images with only few different angles (different directions) to extract discriminative features. The feature values are represented in the form of interval values, since it contains variations of gait feature due to change in cloth, carrying a bag and different normal conditions. Gait recognition is achieved by comparing crisp values of test (probe) sequence with interval values of training sequence (reference) in the gait knowledge-base. The results obtained from our experimentations have shown considerable improvement in terms of correct recognition rate (CCR), when compared to some of the existing methods.
REFERENCES


Mohan Kumar H P obtained MCA and M.Sc Tech by research from University of Mysore, India in 1998 and 2009 respectively. He is working as Assistant professor in department of MCA, PES College of Engineering, Mandya, Karnataka, India. He is currently working towards his Ph.D degree in the area of computer vision at University of Mysore.

Nagendraswamy H S obtained his M.Sc and Ph.D degrees from University of Mysore, India in 1994 and 2007 respectively. He is currently working as Associate Professor in the Department of Studies in Computer Science, University of Mysore, Manasagangothri, Mysore, Karnataka, India. His focused areas of research include Shape analysis, Texture analysis, Precision agriculture, Symbolic data analysis, Fuzzy theory, Biometric, and Video analysis.