Multi-View Gait Recognition Using Enhanced Gait Energy Image and Radon Transform Techniques

Iman Mohammed Burhan and Md. Jan Nordin
Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, UKM, Bangi, Malaysia

Corresponding Author: Iman Mohammed Burhan, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, UKM, Bangi, Malaysia

ABSTRACT

Gait is an overall perceived biometric gimmick that is utilized to distinguish a human at a separation. Gait Recognition (GR) systems encounter several challenges, including viewing angles and translation variations. Hence, GR systems require the development of a robust gait representation model which is invariant in varying conditions. As such, this study presents a gait representation model for Multi-view Gait Recognition Systems (MvGRS) based on Gait Energy Image (GEI) and Radon Transform (RT) on human silhouettes to overcome the challenges in human recognition. In this regard, GEI is utilized for the description of gait features of binary silhouette images which are robust in multi-viewing and varied in appearances. Furthermore, the adoption of Radon Transform (RT) allows for the accommodation of gait representation model with RT features and silhouette alignment. This is to overcome the difficulties in geometrical transformation such as translation, scale and rotation. Consequently robust Principal Component Analysis (PCA) and Partial Least Square (PLS) approaches are accomplished in the reduction of these dimension feature vectors and feature selection. Finally, the recognition of gaits is based on similarity in measurements using Euclidean distance. The experiments were conducted on the public data set of CASIA. The findings from these experiments show that the results are better in comparison with the other methods. Thus, this indicates that the proposed method for gait recognition can outperform the existing methods in gait recognition.

Key words: Biometric identification systems, multi-view gait recognition, gait energy image, view transformation model, radon transform

INTRODUCTION

Biometric Identification Systems (BIS) are increasingly investigating as reliable and efficient technology to person identification and verification (Boulgouris and Chi, 2007). According to Dawson (2002), various biometric measures are used to obtain an individual’s identity. The measures are classified into two distinct categories, namely physiological and behavioral.

The physiological classification includes biometrics measures that bear an immediate estimation of a piece of a human body. To date, the most noticeable and fruitful of these sorts of measures to date are fingerprints, face identification and iris-examines.

The behavioral class extricates qualities focused around an activity performed by a single person (Lu et al., 2007). These sorts of measures are a roundabout estimation of the attributes of the human structure. The fundamental peculiarity of a behavioral biometric is the utilization of
time as a metric. There are several possible measures of behavioral features in the process of human recognition. They include voice dialect, signature and gait with high potential in the application domains, for example at airports, in banks and high-security areas (Foster et al., 2003; Holien, 2008).

The more settled measures incorporate keystroke-sweep and discourse designs. It is alluring that the procedure of biometric distinguishing proof be a robotized one. This is on account of the methodology of manual gimmick extraction which is regarded to be repetitive and also drawn out as a lot of information must be gained and prepared for the generation of a biometric mark. Henceforth the failure to naturally extricate the sought qualities would render the methodology to be infeasible for information sets of practical size in a genuine application (Bobick and Johnson, 2001; Mjaaland, 2009).

Gait recognition is the process of determining persons by the way in which they walk. This is a less unobtrusive biometric which offers the possibility to identify persons at a distance, without any interaction or cooperation from the subject, this is the property which makes is very attractive as a means of identification (Alesse et al., 2012).

In BIS, much consideration has been committed to the utilization of human gait patterns to the examination of human movement as a rule by perceiving the way the individual strolls (Lee et al., 2013). Human concession focused around gait is a generally late approach contrasted with finger impression perception which is a conventional methodology (Bianchi et al., 1998; Jain et al., 2007). As per (Nixon et al., 2010; Wang and Geng, 2010), gait is characterized as the way in which an individual strolls and is one of the few biometric characteristics that can be utilized to recognize people at a separation. Utilizing gait as a biometric inside the domains of machine vision is a noticeably new territory of study. Since its initiation, it has gotten becoming enthusiasm inside the machine vision group. In that capacity, various gait measurements have been produced. The early mental investigations of gait uncovered that gait was a novel individual trademark that incorporates with rhythm and was cyclic in nature.

The methods used for gait recognition are exposed to many different kinds of challenges that have affected the rate of recognition and performance of gait recognition. As stated by Ng et al. (2014) and Zheng et al. (2011), the main challenges include view-invariance and subjects carrying objects or wearing different clothings. Furthermore, translation invariant is another challenging issue that need to be studied further (Boulgouris and Chi, 2007).

MATERIALS AND METHODS

This section presents an overview and details of materials and method. The overall proposed method consists of two individual procedures as gait signature registration and gait recognition as reported by Zheng et al. (2011). The proposed method involves modification in the gait representation model to preserve more effect in translation invariant.

Shortly, as in very nearly all late methodologies for gait recognition, it was expected that the information to our framework is arrangements of double shapes which were concentrated from a gait video sequence utilizing a foundation subtraction process. As likewise exhibited (Boulgouris and Chi, 2007), in all double outlines are at first prepared keeping in mind the end goal to figure the time of strolling. Be that as it may, the extraction of outlines from gait feature successions is normally defective and yields erroneous shapes. Hence, in most useful cases, there will be a requirement for denoising (Boulgouris and Chi, 2007) preceding the application of a gait perception calculation. Hence, it was performed image pre-processing such as noise removal before
gait generate a gait representation model to enhance the gait images. The image enhancement process assists us to improve more accuracy of the GR system accordingly. Subsequently Gait Energy Image (GEI) feature is used for gait representation model in spatial and temporal domain in human silhouettes. The silhouettes are adjusted and Radon transform. Silhouette arrangement is imperative since the Radon transform views the picture focus as the focal point of the change and it requires to align the silhouettes. With extracting of represented feature, it was integrated them and generate a feature vector. The details of the gait representation model also are discussed in following section. The generated feature vector is intended to enhance the Gait Recognition System’s (GRS) accuracy in comparing with previous approaches which they were considered one of features as GEI or Radon. Subsequently, Partial Least Square (PLS) is fulfilled to learn optimal feature representation vector. The PLS is employed as a feature selection approach. Then, a vigorous VTM using vigorous Principal Component Analysis (vigorouPCA) similar to Candes et al. (2011) and Zheng et al. (2011) methods is constructed. In VTM, Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) are done to enhance GEI characteristics and by utilizing SVD, VTM is detailed focused around concentrated created feature vector from multiply viewing angles and various subjects (Kusakunniran et al., 2009). From this time forward, following by phrasing in Phillips et al. (2000) where the reference database is termed as “Gallery”, while testing objects is termed “Probe” subject, the created gait representation of test review point are transformed into that of gallery viewing angles. Finally, it was applied the Euclidean distance to similarity measurement for gait recognition results.

Gait representation model: Human strolling can be considered as intermittent activity. One gait period contains a grouping of strolling postures. In order to represent a gait model, gait period estimation is required (Kusakunniran et al., 2009). Hence, it is required to include the gait sequences (period) integrate with GEI feature to create an accurate gait representation model. Furthermore, Radon Transform (RT) is applied to preserve translation invariants in gait representation model. The detail of representation model is explained in the following sections.

Gait period estimation: Gait period examination has been investigated in past work (Candes et al., 2011; Kusakunniran et al., 2009), serves to focus the recurrence and period of each one watched succession in order to adjust groupings before matching. BenAbdelkader et al. (2002), width time sign of the jumping box of moving silhouette got from a picture grouping is utilized to dissect gait period. This study adopts the similar method as used in Kusakunniran et al. (2009) and Wang et al. (2003) to determine the period of each gait sequence with presenting bounding box similar to BenAbdelkader et al. (2002).

Gait Energy Image (GEI) feature extraction: In this study, gait energy image is created based on gait cycle estimation. According to the results from the period estimation, gait energy image is used as a gait representation for the gait information in spatial and temporal domain. As mentioned in the study of Zheng et al. (2011), gait energy image contains the continuous changes of pose when human walking. Gait energy image holds several key features of the human gait including motion frequency, temporal and spatial changes of human body and global body shape statistic. The silhouettes from background modeling to construct the gait energy image. Given that \( I_{mn} (x, y) \) is a particular pixel located at position \((x, y)\) of \( t \ (t = 1, 2... T) \) from image \( n \ (n = 1, 2... N) \) Gait cycle. All the silhouettes are normalized along the vertical and horizontal directions to a fixed size. Where the width and height of GEI are \( W \) and \( H \). GEI is represented as (Zheng et al., 2011):
Fig. 1(a-c): GEI silhouette samples under different viewing angles, (a) 54°, (b) 90° and (c) 126°

\[ g(x, y) = \frac{1}{T} \sum_{t=0}^{T-1} \sum_{k=1}^{H} I_{t+k} \cdot (x, y) \]  

(1)

where, \( T \) is the number of frames in gait sequence. \( I \) is a silhouette image at frame \( t \), while \( x \) and \( y \) are the image coordinates.

The original GEI feature representation is a 1-D vector, \( g^m \). By concatenating the value of each position in \( I_m \cdot (x; y) \) on all the consecutive rows, where \( m \) represents the mth subject and \( k \) represents the kth viewing angle, the dimension of the \( g^m \) is \( W \times H \). Figure 1 illustrates the GEI of different viewing angles to be examined for this research study.

**Radon Transform (RT):** The Radon transform is exceptionally suitable for gait representation and recognition. The reason is that amid human strolling, there is extensive variety in the edges which are the fundamental leg and arm tomahawks structure regarding, say, a level hub. This implies that the Radon change, brought regarding the focal point of silhouettes, ensures that a great part of the vitality of the first silhouettes will show up in particular coefficients which will differ significantly through time. Hence, imperative decisions about a singular's shape and strolling style can be drawn by recognizing and examining these Radon coefficients (Boulgouris and Chi, 2007; Canton-Ferr er et al., 2010).

This study is enlivened by Boulgouris and Chi (2007) who utilized Radon change for gait identification. Radon transform of silhouettes are utilized for gimmick depiction for this work. Since Radon transform is not exceptionally delicate to clamor, this work can oppose to commotion to a certain degree (Chen et al., 2008). Likewise it was demonstrated that it was interpretation invariant, revolution invariant and scaling invariant (Wang et al., 2007).

In this stage, it was applied Radon Transform to extract features of gait recognition without the effects of translation invariance. Radon Transform is well-known for its extensive variety of imaging applications. It has been characterized in numerous distinctive structures of the most common definition is denoted as below:

\[ R(f)(p, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x, y) \delta(p - x \cos \theta - y \sin \theta) \, dx \, dy \]

(2)

where, \( R(f)(\rho, \theta) \) is the integral line of a 2-D function \( f(x, y) \). The distance of the line is from \( -\infty \) to \( \infty \). The position of the line is determined by two parameters \( \rho \) and \( \theta \) essentially, \( R(f)(\rho, \theta) \) is the integral of \( f \) over the line \( = x \cos \theta + y \sin \theta \). In this study, the 2-D operation is the binary silhouette.
In its discrete structure, a Radon change comprises of a summation of pixel intensities along lines of distinctive bearings. The middle of the silhouette is characterized as the reference point.

A mapping between the space dictated by the direction framework \((x, y)\) and the Radon area controlled by \((\rho, \theta)\) can be built by the application of Radon change to the silhouettes. Figure 2 delineates the procedure for the figuring of Radon coefficients. The silhouette is anticipated onto the \(\rho\) axis that is given a particular bearing. As it were, the pixels along a set of lines that are parallel to the \(s\) axis pivot are summed together. In the meantime a coefficient of \((pi, \theta_i)\) in the Radon area relates to the total of pixels along a line parallel to the pivot in the first silhouette. The area and the introduction of the summation line is focus by \(\rho\) and \(\theta\).

**Feature integration and selection (PLS):** One of the difficulties in gait recognition frameworks is to reduced gait data into model which is suitable for realization. The most effective method to utilize key gait representation and peculiarities to attain proficient gait perception is depicted in this segment. Assuming that there is an activity arrangement which has \(n\) edges meant in the set \(F\) below:

\[
F = \sum_{i=1}^{n} F_i
\]

(3)

where, \(N\) is number of frame and \(F\) is frame, the corresponding Radon transforms are denoted as:

\[
R = \sum_{i=1}^{n} R_i
\]

and the gait feature as:

\[
S = \sum_{i=1}^{n} S_i
\]

the feature vector was generated as:
where, N is number of frame, \( R_i \), \( g_i \) and FS are as gait feature, Radon transform and generated feature vector, respectively.

Each generated feature vector of FS is a matrix with large amount of data. In this regards selection of some features is a solution to overcome the challenges for processing all the feature vectors. As reported in Rosipal and Kramer (2006), Partial Least Square (PLS) choice technique was utilized to concentrate the discriminative piece of the gait feature vector. PLS is a productive regulated measurement of the decremen approach and has the capacity to perform segment of target lessened measurement which is not restricted by the quantity of class of preparing datasets.

To execute PLS characteristic choice methodology, as expressed in Zheng et al. (2011), two sets of gait feature vectors from distinctive subjects are given for the same review plot, i.e., \( FS_n^m \) is gotten from the \( n_{th} \) subject for the \( k_{th} \) viewing angle while \( FS_n^m \) is acquired from the \( n_{th} \) subject for the same viewing angle. PLS figures an ideal projection via hunting the bearing down most extreme after item work between two variables as:

\[
\text{max}_{w_k} [\text{cov}(FS_k^m, FS_n^m)]
\]

where, \( w_k \) is the scholarly projection framework of \( k_{th} \) viewing angle \( \text{cov} \) operation intends to process the covariance between peculiarity representation vectors. Consequently, given another GEI characteristic representation vector \( FS_n^m \), the optimal gait feature vector, \( x_k^m \) for the \( k_{th} \) viewing angle is as follows:

\[
x_k^m = FS_n^m w_k
\]

**Development of view transformation model:** This module explains the designing of View Transformation Model (VTM) implementing a robust Principal Component Analysis technique (PCA) and Truncated Singular Value Decomposition (TSVD) techniques with proposed gait representation model.

**Robust Principal Component Analysis (PCA):** Since, the represented feature vectors from previous stage has the same amount of data of Radon and gait features, it protects a grid with high measurement. Despite the fact that the spoke to gimmick vector has extensive measure of information, the most vital information is not that much and measurement decreasing is needed. In this study, keeping in mind the end goal to lessen the dimensionality of the vector and make identification more productive, vigorous PCA propelled of Candes et al. (2011) and Zheng et al. (2011) was connected to the gait representation model and afterward build View Transformation Model (VTM).

The PCA is a method used to reduce the dimensionality of information. It changes over a set of perceptions of conceivable connected variables into a set of estimations of straight uncorrelated variables called main parts by an orthogonal change. While holding the conceivable variety show in dataset, the PCA will decrease the dimensionality of the information correspondingly (Murukesh and Thanushkodi, 2013). The objective of utilizing PCA is to concentrate a large portion
of the variety of the first variables as peculiarity vectors and diminish the measure of processing required for exact remaking of the information. To figure the PCA, a set of gimmick vectors are made by setting all the first information for a given arrangement in a solitary vector connection grid C. The relationship lattice C is a symmetric grid that serves to diminish the reckoning when figuring the eigenvectors and eigenvalues.

\[ C = \frac{1}{N} \sum_{k=1}^{N} (x_k - \bar{x})(x_k - \bar{x})^T \] \hspace{1cm} (7)

The feature vectors are projected into the new eigenvector space and then finally used for recognition.

**Singular Value Decomposition (SVD):** The goal of using SVD is to process factorization. We factorize the gait representation model based on Singular Value Decomposition (SVD) as described in Zheng *et al.* (2011). A gait framework is made as the left hand side lattice in Eq. 8. The subtle element is as, assuming each one line contains the gait data from diverse subjects under the same review plot while every section incorporates that from the same subjects under distinctive survey edges. All things considered, there are aggregate K review points and M subjects for building VTM. In like manner the factorization procedure is executed based after definition as:

\[ X = \begin{bmatrix} x_1^l & \cdots & x_1^m \\ \vdots & \ddots & \vdots \\ x_N^l & \cdots & x_N^m \end{bmatrix} = p(v^1, \ldots, v^K) = USV^T \] \hspace{1cm} (8)

where, \( U \) is the \( KN_g \times M \) orthogonal matrix. The dimension of \( x_k^m \) that is assumed as \( N_g \times 1 \). \( V \) is the \( M \times N \) orthogonal matrix, while \( S \) is the \( M \times M \) diagonal matrix holds the singular values.

The vector \( v^m \) is a shared gait feature of the \( m_k \) subject from any viewing angle. \( P_k \) is a transforming matrix which can project shared gait feature vector \( v \) to the gait feature vector under a specific viewing angle; \( k \). \( P_k \) is independent of the subject. Based on Eq. 4, given an optimized gait feature vector, \( x_k^m \) from the \( m_k \) subject under the \( j_k \) viewing angle, the learned VTM transform gait feature vectors from the \( j_k \) viewing angle to the \( i_k \) viewing angle and obtain the transformed one is as follows:

\[ x_k^m = P_{k}^i x_k^m \] \hspace{1cm} (9)

where, \( P_k^i \) is the pseudo inverse matrix of \( P_k \). On the other hand, galvanized by the enhancement of Truncated Singular Value Decomposition (TSVD), the shortened rank approximation for SVD attains better achievement in gait recognition. Therefore, the VTM was constructed in term of PCA and TSVD to achieve dimensionality reduction and gait representation factorization and finally gait recognition is remained that following section describes about it.

**Recognition though features similarity measurement:** Once gait information is extracted from gait sequences and projected into a feature space, an appropriate distance metric between feature vectors must be initially defined.

In the recognition stage, utilizing the developed VTM, survey edges of display gait information and test gait data can be changed into the same course. At that point, gait relativity is measured.
without troubles. The basic, yet generally embraced Euclidean distance (Zheng et al., 2011) is utilized as a part of this study for peculiarities likeness estimations. As soon as the two gait features of the same viewing angle, $x_i$ and $x'_i$, are acquired, the affinity of the two traits $(x_i, x'_i)$ is linearly calculated below:

$$d(x_i, x'_i) = \sum_{n=1}^{N} |x(n) - x'(n)|$$

(10)

where, $d(x_i, x'_i)$ is the distance between gait signatures $x_i$ and $x'_i$. $N$ is a dimension of a gait feature. When the value of the smaller value of $d$ records a smaller value, it means that similarity between the representations of the gait feature $x_i$ and $x'_i$ is closer.

RESULTS AND DISCUSSION

One of the most plainly accessible database is the CASIA-B gait database as a multi-view gait dataset (Yu et al., 2006; Zheng et al., 2011). It comprises of 124 subjects under 11 viewing angles (i.e., from 0-180°). Foreach subject there are 10 walking sequences per person per views are recorded, of which 6 sequences under the normal walking circumstances, 2 sequences are for the walking cases whereby the subjects wear, 2 sequences are walking cases whereby the subjects carry knapsacks. This study uses the sequences of normal walking cases. The dataset is prorated into 2 groups the first group comprises of 24 subjects for VTM construction process. The second group comprises of the remaining 100 subjects for the evaluation on multi-views gait recognition. As reported in Zheng et al. (2011), made the GEIs accessible for CASIA database and where it is downloadable at http://www.cbsr.ia.ac.cn/users/szhengl. This study selects normal walking. The experiments carried out evaluate the performance of gait recognition whereby the subject registered and tested in normal walking conditions. Table 1 shows the final results. From the results obtained in Table 1, it can observe that the recognition rate for the viewing angles is a high because it using Radon transform. The findings from these experiments show that the results are better in comparison with the other methods (Zheng et al., 2011; Kusakunniran et al., 2009). The reason for better performance lies in the use of Radon Transform (RT) during the feature extraction stage. Figure 3 and 4 show the comparison of the rate of recognition of the present study and previous research study.

<table>
<thead>
<tr>
<th>Gallery viewing angle (°)</th>
<th>Rate of recognition (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probe gait data under the viewing angle of 126°</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>0.58</td>
</tr>
<tr>
<td>72</td>
<td>0.63</td>
</tr>
<tr>
<td>90</td>
<td>0.75</td>
</tr>
<tr>
<td>144</td>
<td>0.92</td>
</tr>
<tr>
<td>162</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Table 1: Rate of recognition under the condition that the probe gait data is under the viewing angle of 126° and 90° while the gallery gait data is under the viewing angles from 54°-162° and 54°-144°, respectively.
Fig. 3: Comparative analysis of results of the present research with the results of Zheng et al. (2011) and Kusakunniran et al. (2009) methods. The conditions are viewing angle of 128° for probe gait data and viewing angles from 54-162° for gallery gait data.

Fig. 4: Comparative analysis of results of the present research with the results of Zheng et al. (2011) and Kusakunniran et al. (2009) methods. The conditions are viewing angle of 90° for probe gait data and viewing angles from 54-144° for gallery gait data.

CONCLUSION

This study addresses the challenges in extracting features of gait recognition because of the effects of translation invariance. Hence, in order to overcome the difficulties, it is suggested through this study regarding a new multi-views gait recognition based on the integration of the Gait Energy Image (GEI) and Radon Transform (RT) as a gait representation model. The RT features and silhouette alignment preserve scale, orientation and invariance translation. Consequently robust Principal Component Analysis (PCA) and Partial Least Square (PLS) approaches are accomplished to dimension reduction of features vectors and feature selection, respectively. It can show the results of the experiments proved that the objective of using gait energy image based on Radon transform has been is successful in the implementation of gait recognition. Another finding is the feature extraction from the gait is found to be invariant to rotation, translation and scaling. Therefore, when Radon transform was applied to solve the problem of translation invariance, it was found to be effective.
ACKNOWLEDGMENT
The researchers wish to thank Universiti Kebangsaan Malaysia (UKM) and Ministry of Education, Malaysia for supporting this work by research grant FTGS/1/2014/ICT07/UKM/02/2.

REFERENCES


