Human Verification System Based on Gait Recognition using Active Horizontal Levels (AHL) Feature

By
Manhal Saleh Al-Mohammad [*], Gouda I. Salama [**],
Tarek A. Mahmoud [**], and Osama F. Saleh [***]

Abstract:

Nowadays, Gait is a crucial field for many Pattern Recognition researchers. It is considered as a good way for biometric authentication in many surveillance systems. The most important issue in gait recognition is the features extraction from the silhouettes which are converted from walkers' images or videos. In this paper a new method has been introduced to identify night walker images, which are captured by infrared cameras. The new method depends on the silhouette's presence percentage on different horizontal levels. Minimum distance classifier has been used to choose the active horizontal levels that lead to the highest identification rate. A proposed algorithm which is used to select the Active Horizontal Levels (AHLs) has been presented. The proposed method was evaluated against CASIA silhouette database, to recognize walker into one, or multi silhouettes. The experimental results reveal the effectiveness of our new method against other Gait recognition methods to achieve a better recognition rate.

Keywords:

Feature Extraction, Gait Recognition, Biometric Authentication, Active Horizontal Level

*   Syrian Armed Forces
**   Egyptian Armed Forces
***   Cairo Higher Institute for Engineering, Computer Science and Management

1. Introduction:

Gait has played an important role in biometric authentication due to its unique characteristics compared with other biometrics. Gait can be captured at a distance and without requiring the prior consent of the observed subject. Most other biometrics such as fingerprints [1], face [2], hand geometry [3], iris [4], voice [5], and signature [6] can be captured only by physical contact or at a close distance from the recording probe. Gait also has the advantage of being difficult to hide, steal, or fake.

The techniques used for gait recognition can be divided into two categories: model-based methods and motion-based methods.

Model-based methods aim to explicitly model human body or motion, and they usually perform model matching in each frame of a walking sequence so that the parameters such as trajectories are measured on the model. Tao *et al.* [7] introduced a set of Gabor-based human-gait appearance models and propose a general tensor discriminant analysis (GTDA) to solve the carrying status in gait recognition. Rong Zhang *et al*. [8] proposed a 2-step, model-based approach, in which reliable gait features we re extracted by fitting a five-link biped human locomotion model for each image to avoid shape information. Sundaresan *et al.* [9] proposed a hidden-Markov models-based framework for individual recognition by gait. Tan *et al.* [10] used Orthogonal diagonal projection for gait recognition. Tan *et al.* [11] recognized night walkers based on one pseudo shape representation of Gait. The effectiveness of model-based methods, especially in body structure/motion modeling and parameter recovery from a walking video, is still limited allowing for current imperfect vision techniques (e.g., tracking and localizing human body accurately in 2D or 3D space has been a long-term challenging and unsolved problem). Further, the computational cost of model-based methods is relatively high.

However, motion-based approaches can be further divided into two main classes. The first class called the state space methods. These methods considered gait motion to be composed of a sequence of static body poses and recognized it by considering temporal variations of observations with respect to those static poses. Liang Wang *et al* [12] performed eigenspace transformation to time-varying silhouette shapes in features extraction. However, Kale *et al* [13] used Hidden Markov Models in gait recognition.

The second class called the spatiotemporal methods. These methods generally characterized the spatiotemporal distribution generated by gait motion in its continuum. BenAbdelkader *et al.* [14] used image self-similarity plots of a moving person to recognize gait. More recently, Hayfron-Acquah *et al.* [15] described an automatic gait recognition method using the spatiotemporal symmetry. Collins *et al* [16] established a method based on template matching of body silhouettes in key frames for human identification.

This paper is organized as follows: the proposed architecture is presented in Section 2. Section 3 describes the preprocessing phase. In Section 4 Feature Extraction phase and AHL algorithm are introduced. Matching phase is presented in Section 5. Two Experiments are performed in Section 6 to illustrate the performance of the proposed architecture. The final section concludes the paper.

2. Architecture of the Proposed Model:

The architecture of the proposed model for human verification system based on Gait Recognition passes through two processes: the Learning process (Enrolment) and the Authentication process (Identification) as shown in Figure 1. The major functional units for each process will be introduced in the subsequent sections. The proposed model has been implemented using Visual Basic 6 as a programming language.

The first Phase in the learning process is the preprocessing. The preprocessing Phase includes resizing the silhouette images in CASIA (Chinese Academy of Sciences) Gait Database [17] to 100 x 75 pixels each. Then the silhouette resizing is followed by defining the person's region of interest. The second Phase is the feature extraction. In this Phase each captured gait image is represented by the Active Horizontal Levels (AHLs) feature vector that will be explained later. The third Phase is building the training database that contains the AHLs feature vector of each sample for each person.

Similarly, the Identification process is partitioned into three Phases which include preprocessing, feature extraction, and matching. The preprocessing and feature extraction Phases are essentially the same like the learning process. The matching process is based on the distance between the test sample feature vector and each sample in the training database.
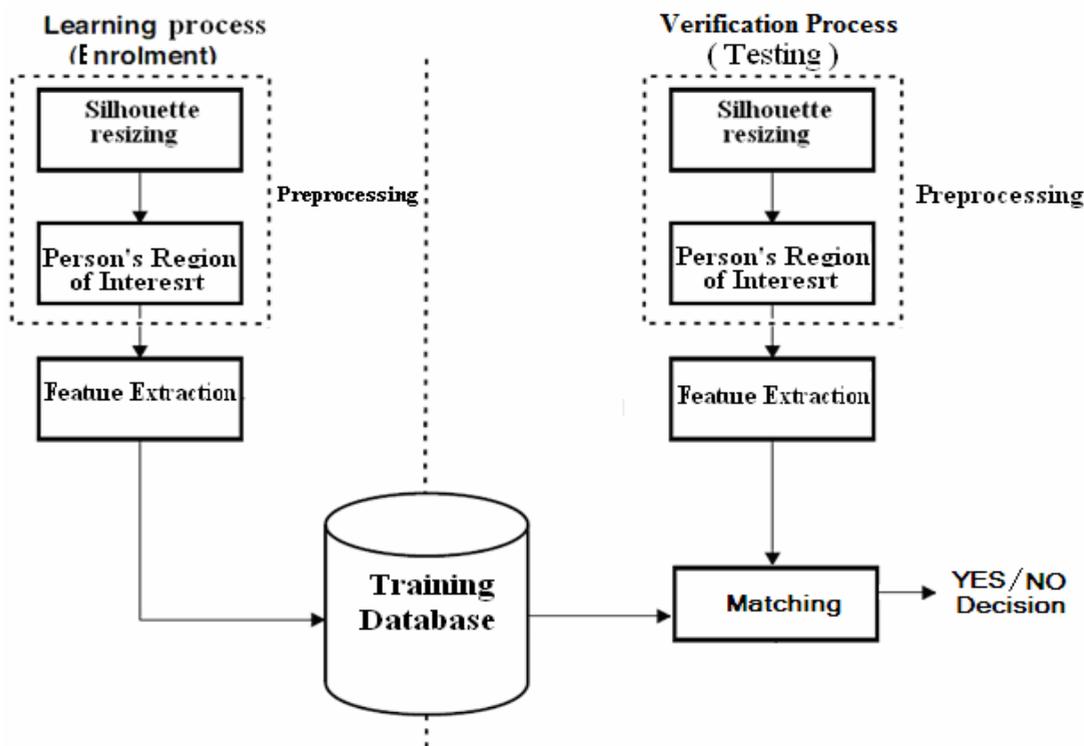
Figure (1): The proposed architecture of the proposed model

3. Preprocessing Phase:

This section describes the preprocessing phase. Two procedures are done to reformulate the images: resizing the images and determination the image region of interest in CASIA datasets. So, the preprocessing of the images was performed, first by resizing the images and then finding the region of interest in all images. The main reason for the preprocessing is two folded. The first one is the attempt to reduce the size of pattern vector. The second reason is to isolate only information that distinguishes individuals for images. The CASIA gait database [17] is one of the largest databases that is used in human verification using gait. The CASIA gait database has three data sets (A, B, and C), the Data set C was used in our experiments. It was captured by an infrared (thermal) camera at night. It contains 153 different walking persons, and takes into account four walking conditions: normal walking, slow walking, fast walking, and normal walking with a bag. Figure 2 shows a sample night gait images captured for a person in four different walking conditions (Normal, Fast, Slow, and Normal with a bag representation). Figure 3 shows sample silhouettes for a person walking Normal, Fast, Slow, and Normal with a bag.

| (a) | (b) | (c) | (d) |
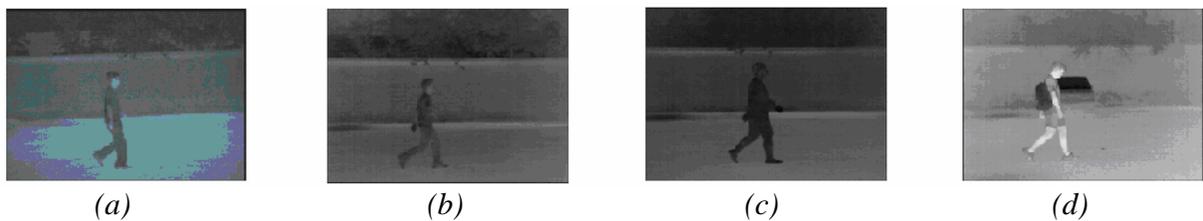
Figure (2): *Night gait images Samples in the thermal infrared spectrum.*
(a) Normal walking. (b) Fast walking. (c) Slow walking. (d)Normal walking with a bag.
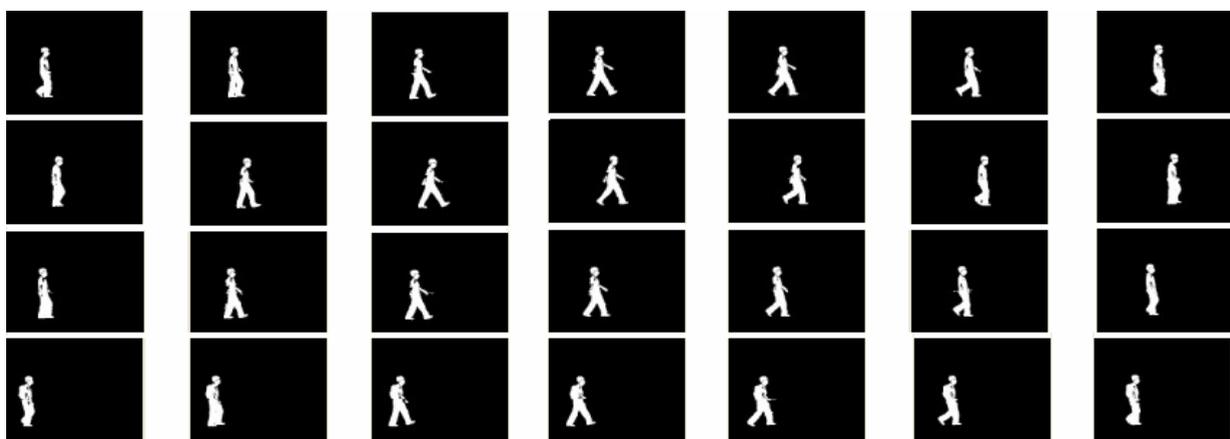
Figure (3): *Silhouettes Samples from the CASIA infrared Night Gait Dataset.*
1st row: normal walking; 2nd row: fast walking; 3rd row: slow walking;
4th row: normal walking with a bag.

The walking person in the gait silhouette is detected by scanning the image firstly from top, left to right until the first pixel belongs to the object is reached. Then, from bottom, left to right until the first pixel belongs to the object is reached. Figure 4 shows a sample silhouettes for a person walking Normal, Fast, Slow, and Normal with a bag after defining the person's region of interest. For instance, the CASIA Database provides images that are 300×240 pixels. If all rows of the image are concatenated into only one vector, the dimension of the problem would be to classify a vector with 72,600 elements. This dimension is too high for computing, and even though having such capacity, satisfactory results is not guaranteed. Therefore, reducing the dimensionality of the problem while focusing only on parts of the scene that effectively identify the individual.

| (a) | (b) | (c) | (d) |
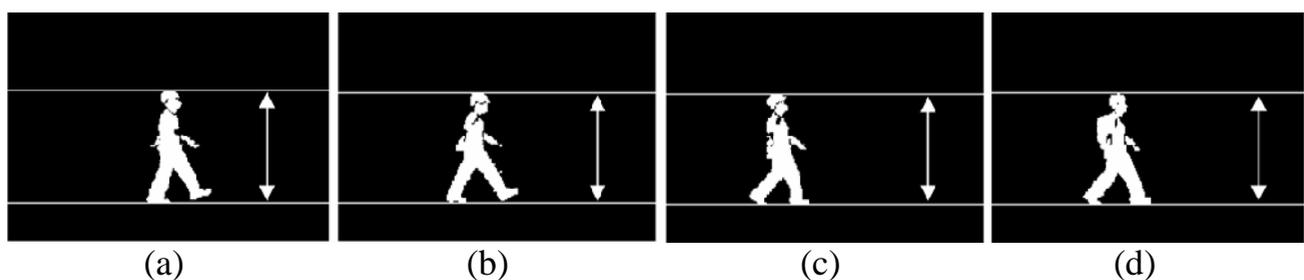
Figure (4): *Person's region of interest for night gait silhouette Sample.*
(a) Normal walking. (b) Fast walking. (c) Slow walking. (d)Normal walking with a bag.

4. Feature Extraction Phase:

In this phase, two algorithms are introduced, Feature Extraction algorithm and Active Horizontal algorithm.

4.1 Feature Extraction algorithm:

The key to the success of any gait recognition system is the gait features extraction. Though, this paper resorts to the use of appearance features to characterize human gait. More precisely, each preprocessed silhouette is projected on many horizontal levels (for example 5%, 10%, …, 99% of the silhouette's height as shown in Figure 5). At each horizontal level the valid number of human pixels is recorded. Thus, each normalized silhouette records a vector of the valid number of human pixels in these horizontal levels to obtain a vector of human pixels counter along all horizontal levels. We denote the presence counter for the L projection level by X_L, where L varies from 1 to N and thus $X = \{X_1, X_2,\ldots\ldots, X_N\}$. These levels fulfill the description of the gait pattern, then the extracted feature vector x is divided by the maximum projection level value for normalization. Figure 5 shows the horizontal levels of a sample silhouette for a person walking Normal. List 1 introduces feature extraction algorithm.
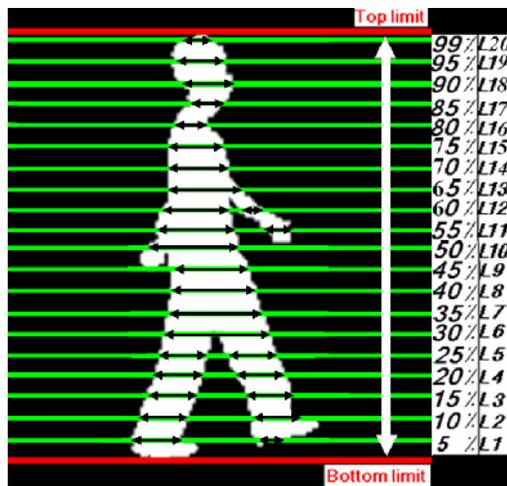
Figure (5): *Horizontal levels for a person walking Normal*

List 1: *Feature Extraction Algorithm*

Initialization:

        $l = 1$, $l = 1$, …, N (where l is the horizontal level and N is maximum number of levels).

Step-1: Compute X[l], where X[l] is the number of pixels for person's appearance on horizontal level l.

Step-2: Increment l.

Step-3: Repeat Step-1 and Step-2 until l = N.

Step-4: Find maximum value X_{Max} for X[l] elements.

Step-5: Divide each element in X[l] by X_{Max} to generate Normalized Image Feature Vectors (NIFV).

4.2 Active Horizontal Levels algorithm:

The second algorithm towards feature extraction phase is finding the active horizontal levels that can be used for authentication a person that is AHL algorithm.

First, select a sequence NIFV's and consider them as a Reference NIFV's (RNIFV's). Second, find the indices of all persons that identifies the minimum distance between each different sequences types groups (4 normal walking, 2 fast walking, 2 slow walking and 2 normal walking with bag) NIFV's and RNIFV's for first horizontal level, then check if these indices refers to the true matched persons to get the True Matched Rate (TMR). Third, find the highest value between the different TMR's, which refers to the first AHL. Next, use the obtained AHL with other horizontal levels to get the next AHL. Finally, repeat the previous steps whatever the TMR not decreased to get all possible AHL's. AHL algorithm is presented in List-2.

List 2: *Active Horizontal Levels Algorithm*

```
main()
  {
    Reset Lb=[]; // AHL's in case of normal walking with bag
    Reset Ln=[]; // AHL's in case of normal walking
    Reset Lf=[]; // AHL's in case of fast walking
    Reset Ls=[]; // AHL's in case of slow walking
    Do {
          l =1
          Do {
                  B[l] = TMR(Bag, 1, Lb & l); // TMRs in case of normal walking with bag
                  N[l] = TMR(Normal, 4, Ln & l); // TMRs in case of normal walking
                  F[l] = TMR(Fast, 2, Lf & l); // TMRs in case of fast walking
                  S[l] = TMR(Slow, 2, Ls & l); // TMRs in case of slow walking
                   l ++;
              }While (l    N)
          Get index of maximum value for: B[l], N[l], F[l] and S[l], and it to AHL's arrays:
          Lb, Ln, Lf and Ls
          Use the obtained AHL with other horizontal levels to get the next AHL
        } While (value of TMR increases)
  }


TMR (Si, j, L)
      {
        TMI = 0, // is number of True Matching Image
        for t = 1 to j  // t …. is sequence number for sequence type
          {
             for p = 1 to Np // p …. is number of persons in case of training
                {
                   for k = 1 to Ni // k … is number of images / person in case of training
                      {
                       for p' = 1 to Np // p' … is number of persons in case of test
                          {
                           for k' = 1 to Ni // k'… is number of images /person in case of test
                            {
                            P-index = min distance(Sbag,2,p,k , SSi,t,p',k', L) // index of person
                            that identifies min distance between Sbag,2,p,k and  SSi,t,p',k' for
                            active levels L
                             }
                          }
                          if  P-index = p then TMI++;
                      }
                }
          }
          return (TMI / (Np * j * Ni));
      }
```

TMR is calculated as follows:

$$TMR_{bag} = \frac{Number\ of\ true\ matching\ with\ bag\ images}{Total\ of\ with\ bag\ images}$$

$$TMR_{normal} = \frac{Number\ of\ true\ matching\ normal\ images}{Total\ of\ normal\ images}$$

$$TMR_{fast} = \frac{Number\ of\ true\ matching\ fast\ images}{Total\ of\ fast\ images}$$

$$TMR_{slow} = \frac{Number\ of\ true\ matching\ slow\ images}{Total\ of\ slow\ images}$$

5. Matching Phase:

Gait recognition is a traditional pattern classification problem which can be solved by measuring similarities between reference patterns and test samples. Classification can be based upon various distance measures like Euclidean distance, cross correlation, and Mahaaloblis distance. Euclidean distance is adopted here upon:

For a certain level l, let T represent a test sequence and R_i represent the ith reference sequence. We may classify this test sequence into class c that can minimize the similarity distance between the test sequence and all reference patterns by:

$$c = \arg\min_i d_i(T, R_i), \quad ,i = 1,2,3,...,n$$

Where n is the number of reference patterns and d is the Euclidean distance measure between the test sequence and all reference patterns as:

$$d(T, R_i) = \sqrt{\sum_{j=1}^{m} (T_j - R_{ij})^2}$$

Where:

$$T = [T_1 \quad T_2 \quad ... \quad ... \quad T_m] \ ,$$

$$R_i = [R_{i1} \quad R_{i2} \quad ... \quad ... \quad R_{im}] \ ,i = 1,2,3,...,n \ , \text{and}$$

m is the number of extracted horizontal active level features.

6. Discussion and Experimental results

To evaluate the proposed model, two experiments were performed using 60 persons from the CASIA-C database. In CASIA-C database, each walking person has four normal, two slow, two fast and two normal carrying a bag sequences.

Experiment 1:

In this experiment, we use the collection of the two normal walking sequences of each person carrying a bag as the enrolment data. However for testing data, four normal, two

slow, two fast and two normal carrying a bag walking sequences for each person are used. Thus, we build our enrolment data set by a total of $60 \times 2 = 120$ sequences and each sequence is composed of 55 images. On the other hand, the testing data set consists of $60 \times 10 = 600$ sequences composing a total of 33000 images.

Table (1): TMR results based on training with two normal walking sequences carrying a bag and testing with four normal, two slow, two fast and two normal carrying a bag walking sequences

Experiments	Normal walking		Fast walking		Slow walking		Normal walking with bag	
True Matching Rate (%)	One image	Sequence	One image	Sequence	One image	Sequence	One image	Sequence
	35.45%	90%	33.12%	83.33%	29.97%	76.67%	68.94%	100%

Table (1) shows the True Matching rate (TMR) results based on the collection of the two normal walking sequences of each person carrying a bag as the enrolment data and four normal, two slow, two fast and two normal carrying a bag walking sequences for each person for testing data using the Euclidean distance classifier for both one image and video snapshot. It could be noticed that the resultant true matching rate for normal, fast, slow, and normal carrying bag using one image are 35.45%, 33.12%, 29.97%, 68.94% respectively. On the other hand, the resultant true matching rate for normal, fast, slow, and normal carrying bag using video sequence are improved to 90%, 83.33%, 76.67%, and 100% respectively.

Table (2) shows the active horizontal levels based on training with two normal walking sequences carrying a bag and testing with four normal, two slow, two fast and two normal carrying a bag walking sequences, and shows the active horizontal levels based on union Boolean operation.

Table (2): Active Horizontal Levels

Horizontal level	Cases of testing			union
	normal	fast	slow	
99%			✓	✓
95%	✓	✓	✓	✓
90%	✓	✓	✓	✓
85%	✓	✓	✓	✓
80%				
75%				
70%				
65%				
60%				
55%				
50%	✓			✓
45%			✓	✓
40%				
35%	✓	✓	✓	✓
30%	✓	✓	✓	✓
25%	✓	✓	✓	✓
20%	✓	✓	✓	✓
15%	✓	✓	✓	✓
10%	✓	✓	✓	✓
5%		✓		✓

Figure (6) shows the active horizontal levels based on union Boolean operation on the active horizontal levels based on training with two normal walking sequences carrying a bag and testing with four normal, two slow, two fast and two normal carrying a bag walking sequences. We use these levels in Experiment 2 to decrease the time of computing matching.
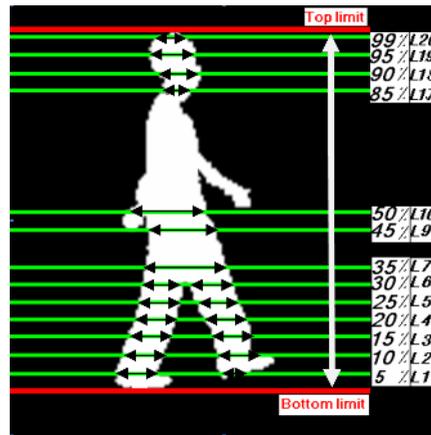
Figure (6): Active Horizontal Levels in case of normal enrolment

Experiment 2:

Since there is not a general gait database for evaluating gait recognition performance, and the total persons of each database vary greatly from 4 to 124. So it is difficult to compare the performance of different algorithms. In this experiment, we use the collection of the four normal walking sequences of each person as the enrolment data. However for testing data, four normal, two fast, two slow and two normal carrying a bag walking sequences for each person are used. Thus, we build our enrolment data set by a total of $60 \times 4 = 240$ sequences and each sequence is composed of 55 images. On the other hand, the testing data set consists of $60 \times 10 = 600$ sequences composing a total of 33000 images.

Figure (7) shows the TMR results at different number of persons based on the collection of the four normal walking sequences of each person as the enrolment data and four normal, two slow, two fast and two normal carrying a bag walking sequences for each person for testing data using the Euclidean distance classifier for both one image and video snapshot. It could be noticed that the resultant true matching rate for normal, fast, slow, and normal carrying bag using one image is 61%, 40%, 41%, and 34% respectively. On the other hand, the resultant true matching rate for normal, fast, slow, and normal carrying bag using video sequence are improved to 96.67%, 91.67%, 85%, and 80% respectively.

Obviously, it could be seen that the results by using video snapshots are much better than using only one image for different number of persons walking with different styles.
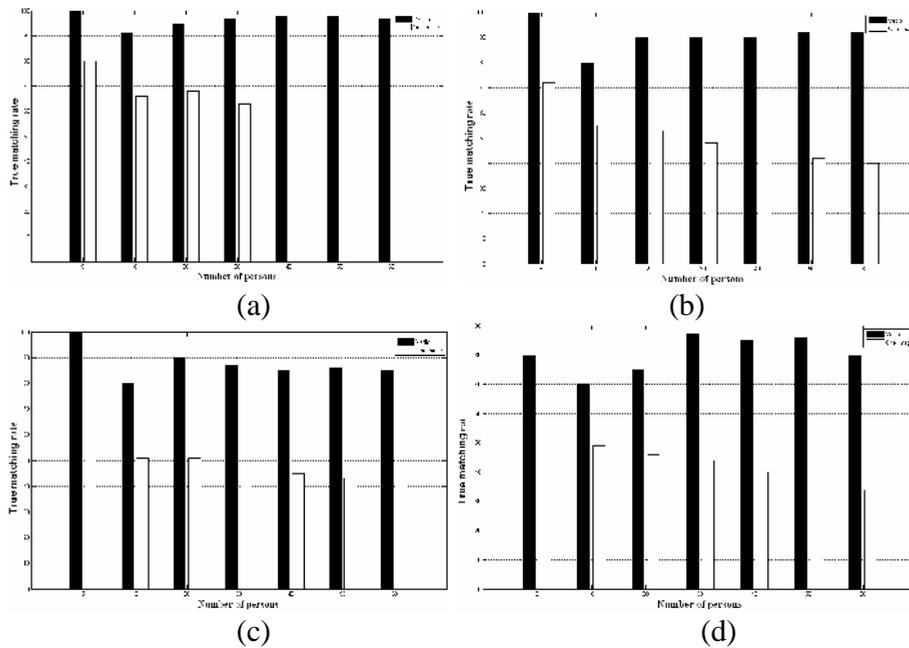
Figure (7): *TMR results for different number of persons based on training with four normal walking and testing with (a) Four normal, (b) Two fast,(c) Two slow, (d) Two normal walking with a bag respectively*

It is worth pointing out the difference between Tan *et al* in [14-15] and ours. We use focus on the cues in horizontal projections, but Tan *et al* in [14] use a focus on the cues in two diagonal directions and Tan *et al* in [15] use a focus on the cues in pseudo shape representation of Gait. Figure (8) shows the comparison of gait recognition performance of our proposed method with the orthogonal diagonal projections method in [14]. Both methods are evaluated by using one image at different number of persons in test database of CASIA dataset-C. It is seen that the TMR of our proposed algorithm is superior to that algorithm proposed in [14]. Figure (9) shows the comparison of gait recognition performance of our proposed method with the PCA, PCA+LDA, and PCA+LDA+RCA feature extraction methods in [15].
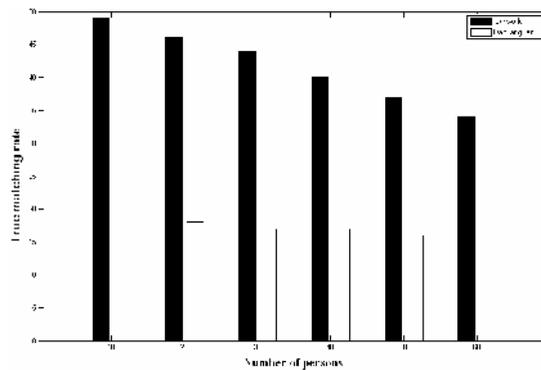
Figure (8): *The TMR comparison of our proposed method with the orthogonal diagonal projections method in [14] by using one image at different number of persons*

All the mentioned methods are evaluated by using video snapshots consisting of 55 sequences each and applied at 60 persons from test database of CASIA dataset-C. Figure 9 (a) shows the TMR of our proposed algorithm is 97%, less than PCA (98.4%), less than PCA+LDA (98%) and less than PCA+LDA+RCA (98.4%). Figure 9 (b) shows the TMR of our proposed algorithm is 90%, as good as PCA (91.8%), less than PCA+LDA (93%) and less than PCA+LDA+RCA (93.7%). Figure 9 (c) shows the TMR of our proposed algorithm is 90%, better than PCA (82.4%), as good as PCA+LDA (90.6%) and as good as PCA+LDA+RCA (91.3%). Figure 9 (d) shows the TMR of our proposed algorithm is 76%, superior than PCA (24.4%), (24.2%) and PCA+LDA+RCA (24.7%).
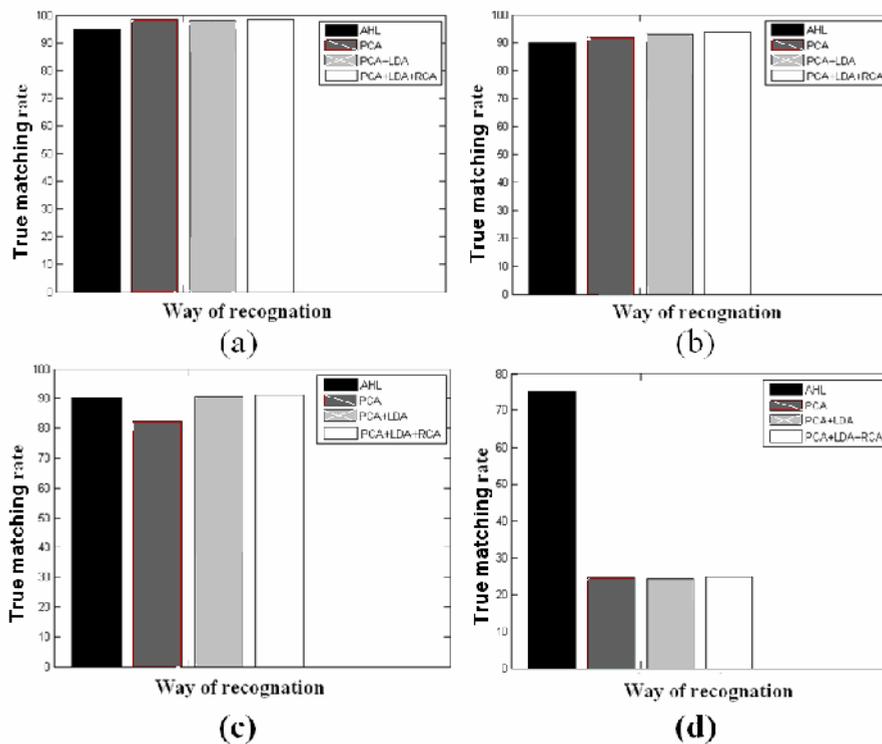
Figure (9): The comparison of gait recognition performance of our proposed method with the PCA, PCA+LDA, and PCA+LDA+RCA methods in [15] applied on 60 persons based on training with four normal walking and testing with (a) Four normal, (b) Two fast, (c) Two slow, (d) Two normal walking with a bag respectively.

7. Conclusions:

This paper has addressed the problem of gait recognition based on appearance features in human silhouettes, with considering the issues of distance metrics and scales. Our major contribution lies in the offering of a promising method to extract gait features (AHL). These features are invariant under Scale, Transform, Speed, Illumination, and Type of clothes. Experimental results on the CASIA Infrared Night Gait Dataset [17] indicate that the proposed algorithm is superior to that algorithm proposed in [14-15] in case normal walking with bag.

References:

[1] K. Ito, H. Nakajima, K. Kobayashi, K., T. Aoki, and T. Higuchi: "A Fingerprint Matching Agorithm Using Phase-only Correlation". IEICE Trans. Fundamentals, Vol.E87-A, No.3 (2004) 682-691.

[2] V. Jain, E. Learned-Miller, and A. McCallum: "People-LDA Anchoring Topics to people using face recognition". In ICCV, 2007.

[3] Alexandra L.N. Wong and Pengcheng Shi: "Peg-Free Hand Geometry Recognition Using Hierarchical Geometry and Shape Matching". IAPR Workshop on Machine Vision Applications, Nara, Japan, pp. 281-284, Dec. 2002.

[4] Miyazawa, K., Ito, K., Aoki, T., Kobayashi, K.: "A Design of an Iris Matching Algorithm based on Phase-only Correlation". Int. Conf. Image Processing (2005) (in press).

[5] Nakasone, H., Beck, S. (2001): "Forensic Automatic Speaker Recognition". Proc. 2001: A Speaker Odyssey, The Speaker Recognition Workshop, ISCA, Chania, Crete, Greece, 18-22 June 2001.

[6] Y. Qi and B.R. Hunt, "A multi resolution approach to computer verification of handwritten signatures: "IEEE Transactions Image Processing". vol. 4, no. 6, pp. 870–874, June 1995.

[7] D. Tao, X. Li, X. Wu, and S. Maybank: "Human carrying status in visual surveillance". in Proc. IEEE Conf. Comput. Vis. and Pattern Recog., 2006, vol. 2, pp. 1670–1677.

[8] Rong Zhang, Christian Vogler and Dimitris Metaxas: "Human Gait Recognition". In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04), 2004.

[9] A. Sundaresan, A. Roy-Chowdhury, and R. Chellappa: "A hidden Markov model based framework for recognition of humans from gait sequences". in Proc. Int. Conf. Image Process. 2003, vol. 2, pp. 93–96.

[10] D. Tan, K. Huang, S. Yu, and T. Tan. "Orthogonal diagonal projection for gait   Recognition: ". in Proceedings of the International Conference on Image Processing (ICIP'07), vol. 1, pp. 337–340, September-October 2007.

[11] D. Tan, K. Huang, S. Yu, T. Tan: "Recognizing Night Walkers Based on One Pseudo shape". Representation of Gait. CVPR 2007.

[12] Liang Wang, Weiming Hu, Tieniu Tan: "A New Attempt to Gait-based Human Identification". icpr, vol. 1, pp.10115, 16th International Conference on Pattern Recognition (ICPR'02) - Volume 1, 2002

[13] A. Kale, A. Rajagopalan, N. Cuntoor, and V. Kru¨ger: "Gait-Based Recognition of Humans Using Continuous HMMs". Proc. Int'l Conf. Automatic Face and Gesture Recognition, 2002.

[14] C. BenAbdelkader, R. Culter, H. Nanda, and L. Davis: "EigenGait: Motion-Based Recognition of People Using Image Self-Similarity". Proc. Int'l Conf. Audio- and Video-Based Biometric Person Authentication, pp. 284-294, 2001.

[15] J. Hayfron-Acquah, M. Nixon, and J. Carter: "Human Identification by Spatio-Temporal Symmetry". Proc. Int'l Conf. Pattern Recognition, 2002.

[16] R. Collins, R. Gross, and J. Shi: "Silhouette-Based Human Identification from Body Shape and Gait". Proc. Int'l Conf. Automatic Face and Gesture Recognition, 2002.

[17] CASIA Gait Database: "http://www.sinobiometrics.comi".