Gait Recognition with Shifted Energy Image and Structural Feature Extraction

Xiaxi Huang and Nikolaos V. Boulgouris, *Senior Member*, *IEEE*

Abstract

In this paper, we present a novel and efficient gait recognition system. The proposed system uses two novel gait representations, the Shifted Energy Image and the Gait Structural Profile, that have increased robustness to some classes of structural variations. Furthermore, we introduce a novel method for the simulation of walking conditions and the generation of artificial subjects that are used for the application of Linear Discriminant Analysis. In the decision stage, the two representations are fused. Thorough experimental evaluation, conducted using one traditional and two new databases, demonstrates the advantages of the proposed system in comparison to current state-of-the-art systems.

Keywords: Gait recognition, surveillance, biometrics

I. INTRODUCTION

The human walking style, termed gait, has been known to have the capacity to be used for recognition purposes [3]. Biometric systems based on gait recognition aim to identify individuals based on their idiosyncratic walking style. Gait is a very appealing biometric trait due to its unobtrusiveness and the ease with which it can be captured. It has been shown that gait has promising discriminatory power as a biometric trait, and it can be combined with other biometrics [4] in order to achieve reliable authentication results.

A gait recognition system normally consists of three modules. Firstly, a pre-processing module, i.e., background subtraction and silhouette extraction from the original frames. Secondly, a gait

feature extraction module. Finally, a data storage/retrieval and recognition decision module. Several methods for gait recognition have been proposed with most research focusing on the second module.

The averaged silhouette [5], also known as Gait Energy Image (GEI) [6], is a simple gait representation with high discriminative power. It is calculated as the averaged frame in one or all of the gait cycles in a gait sequence. In order to extract the most relevant and informative features from a GEI, a supervised and an unsupervised feature selection method were introduced in [7]. However, this selection process did not compensate for GEI's sensitivity to variations of observed body structure that could originate from heavy clothing, fatigue, or carried objects. This aspect of gait recognition algorithms will be explored in more detail in the ensuing parts of the present paper.

Instead of condensing the holistic information from all gait frames into one template, some other approaches explore holistic information extracted on a frame by frame basis. In [8], a Hidden Markov Model (HMM) is used to capture the Frame to Exemplar Distance, which represents the gait information. In [9], a generic gait walking model was constructed using a population Hidden Markov Model (pHMM), the states of which represented gait stances over one gait cycle. The corresponding gait stances in each gait cycle were averaged over one gait sequence in order to form a dynamically-normalized gait cycle of fixed length that can be used for recognition. In [10], three gait features were proposed based on the Radon transform and the weighted Krawtchouk moments. Other approaches construct gait features based on human body models [11], [12].

The common characteristic of the feature extraction approaches outlined above is that they rely on side-view gait sequences. Other approaches, however, are based on sequences taken from other view angles. Frontal-view gait recognition methods have been proposed in [13] and [14]. In [15], three view angles were considered by using two gait representations: Motion Silhouette Image (MSI) and Gait Energy Image (GEI). In [16], the impact of view-angle variations on recognition performance was investigated. Each of the reference and test sets have 11 view-angles. Two typical gait features - GEI and key Fourier descriptors (KFDs) - were extracted from each reference and test gait sequence. Subsequently, the reference and test features that were extracted from every view-angle were used to separate recognition results. It was shown that superior results were obtained when the gallery and probe sets have the same view-angles.

A more systematic approach for mitigating view-dependency and moderate occlusion problems of fixed cameras was presented in [17], which aimed at view-invariant representation of human appearance.

From the above discussion, the conclusion that can be reached is that the side-view normally yields better results than other view angles. Furthermore, model-based features are usually extracted from side-view sequences. At the same time, there are limitations for the existing approaches: 1) no specific provision is made for the very common situation in which the walking subjects wear heavy clothing or carry an object, 2) it is routinely assumed that more than one gait cycles are available in each gait sequence. However, in practical scenarios, the above two assumptions will not always hold. Instead, in typical everyday situations the subjects will walk casually, and the way they walk may be affected by heavy clothing, carried objects, fatigue, or other factors. Conditions encountered in practical scenarios will have an impact on the shape of the extracted silhouettes and will pose difficulties to the extraction of appropriate discriminative features. Furthermore, in some cases, e.g., in sequences captured in smaller areas, no more than one gait cycle may be captured and recognition will have to rely on limited information.

In this paper, we propose a gait recognition system, which is aimed to overcome such obstacles and achieve superior performance. The proposed system has the following features:

- A novel gait representation, termed Shifted Energy Image (SEI), that is resistant to walking changes that originate from heavy clothing, carried objects, or other factors. A further gait representation, named Gait Structural Profile (GSP) that is combined with the SEI and yields superior performance.

- A novel method for dimensionality reduction based on artificial subjects and Linear Discriminant Analysis (LDA). The artificial subjects are generated based on rotation and the introduction of artificial clothing.

- A fusion method for the combination of recognition results obtained using these two representations.

The proposed algorithm is experimentally assessed using the CMU (Carnegie Mellon University) database [18], the CASIA (Chinese Academy of Sciences) database [19], and the ACTIBIO database, which was compiled under the European Commission project ACTIBIO[1]. Experimental
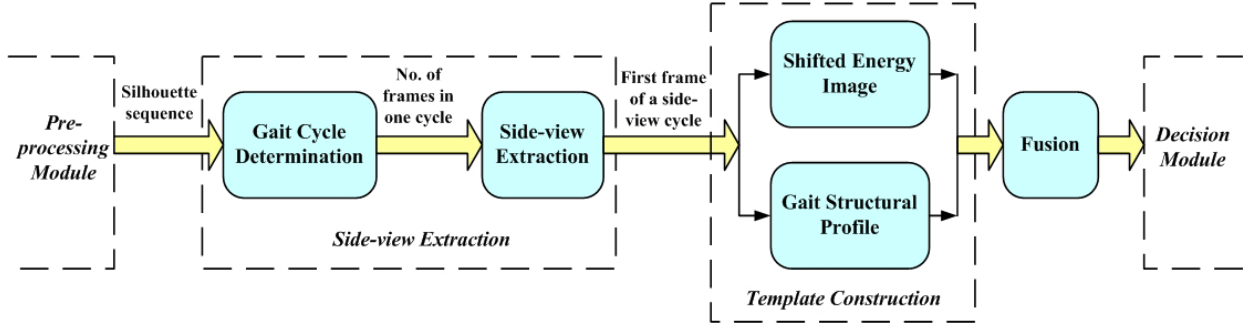
---

[1]http://www.actibio.eu/

Fig. 1.   Block diagram of the proposed gait recognition system.

evaluation shows the benefits of the proposed algorithm over existing gait recognition algorithms. With excellent performance and high computational efficiency, the proposed system is very suitable for deployment in gait-based recognition scenarios.

The rest of the paper is organized as follows: the proposed system is outlined in Section II. The extraction of side-view sub-sequences is briefly described in Section III. The construction of the proposed gait representations is presented in Sections IV, V, and VI. Section VII describes the combination of the two representations in a single framework. Experimental results are presented in Section VIII, and, finally, conclusions are drawn in Section IX.

II. System Overview

The first module of a gait recognition system is the input and pre-processing of the original raw frames. This module mainly includes two consecutive processes: *initial background subtraction* and *silhouette extraction*. The first process generates a foreground map that is re-scaled in order to yield a normalized silhouette of the walking subject. In the present work, the initial background subtraction step is based on the algorithm presented in [20], with the difference that, in order to reduce the computational cost, we use fixed rather than adaptive thresholds.

The extraction of silhouettes is followed by feature extraction. The gait feature extraction process comprises two steps: *side-view extraction*, and *gait template construction*. These are graphically depicted in Figure 1. Specifically, we first determine the number of frames in one cycle using the method in [21]. Then we introduce a method for determining the first frame of a side-view cycle, based on which the side-view cycle can be segmented out of the whole sequence. Finally, we construct two new gait representations by using the side-view cycle, and

fuse them at the decision stage. These processes are described in more detail in the ensuing sections.

III. SIDE-VIEW EXTRACTION

It has been shown that side-view silhouettes contain more gait information than silhouettes taken from other view angles [16]. In order to exploit this fact, we developed a side-view extraction methodology that detects side-view gait cycles. The assumption made is that there is at least one, or almost one, side-view (or near side-view) cycle in each gait sequence. In case no near side-view is expected to be presented, view-invariant approaches in the spirit of [22], [17], [23] would be more appropriate to achieve recognition efficiency (at the cost of higher complexity).

In order to locate, within the gait sequence, one cycle that only, or mainly, includes side-view frames, we first determine the number of frames within one gait cycle. Several approaches [21], [24], [25] have been developed so far for such determination. In our system, we use the approach in [21]. It must be noted that the detection of the number of frames is usually very reliable but, in any case, it does not normally have a significant impact on the performance on the final system. After the number T of frames in a gait cycle is determined, we locate the starting frame of the side-view cycle. To this end, we propose an algorithm that is based on the variation of the subject's apparent height in the frames comprising the gait sequence. If the subject is walking in a front-to-parallel direction (i.e., showing his/her side-view in the frame), the height of his/her silhouette in consecutive frames will change far less than when he/she is walking in other directions. Therefore, the beginning of a side-view cycle will be followed by T frames in which the variance of the subject's height H_B (before normalization) will be minimum. This means that the starting frame F of a side-view cycle will be determined by:

$$F = \underset{Z}{\operatorname{argmin}} \sqrt{\frac{1}{T} \sum_{n=Z}^{Z+T} \left(H_B^n - \overline{H}_B \right)^2} \tag{1}$$

where n is the frame index. The above process is visualized in Figure 2, which shows the apparent height of the subject in a sample sequence. The determination of H_B, on which the above equation relies, is usually easy, except in cases of very cluttered background. Inaccurate calculation of H_B will have a measurable performance impact only in the random path scenario.
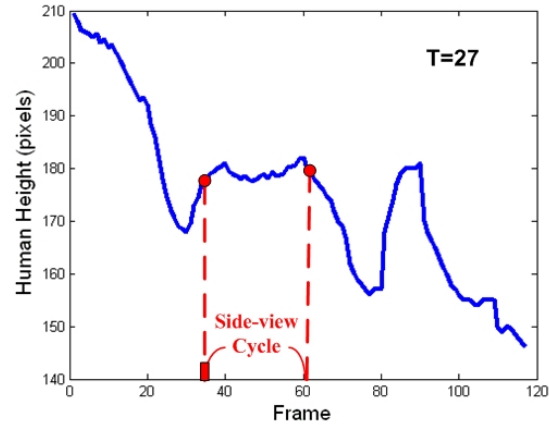
Fig. 2. Side-view cycle partitioning based on the subject's apparent height in each frame. The size T of a gait cycle in this sequence has been determined to be equal to 27. As seen, the variance of the subject's apparent height is minimum between frames 35 to 61. This indicates that frame 35 is the starting frame of the most suitable side-view cycle in this sequence.
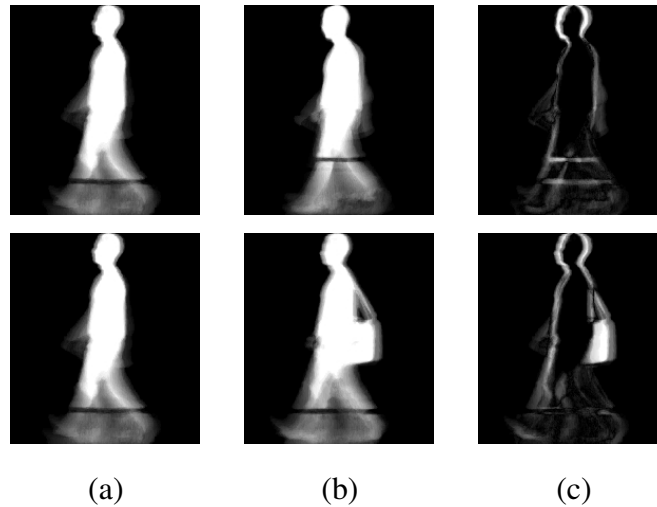

Fig. 3. Averaged silhouettes of the same subject: (a) Gallery template, (b) Probe template, (c) Difference between Gallery and Probe templates.

In general, the impact on performance is negligible in case the detected first frame is within $T/10$ frames of the actual first frame. It must be emphasized that the above approach calculates the part of the gait sequence with the minimum variance and, therefore, this calculation is insensitive to the camera height (an exception would be a camera observing from the top, in which case no height variation would be reliably observable).

IV. Feature Extraction

The averaged silhouette representation [5], widely known as Gait Energy Image (GEI) [6], is an efficient representation that captures gait information from all frames together, instead of relying on information from each frame separately. Shape information is captured in the upper body part (i.e., head and torso area) of the GEI, while dynamic information is contained mainly in the lower body part (i.e., the leg area). However, the averaged silhouette exhibits poor performance in some cases. One such case is shown in the top row of Figure 3, where recognition would be difficult due to the fact that the same subject is walking with his head up and his torso straight in the gallery set, but keeps his head down and leans forward in the probe set. Another example is shown in the bottom row of Figure 3, where the subject walks without / with a bag in the gallery / probe set. In the above two cases, the difference, especially along the contour of the silhouettes, between the same subject's gallery and probe averaged silhouette templates is large. This is due to the fact that, when the person was leaning forward or carrying a bag, the body shape changed and the gravity centres of the silhouettes were shifted.

Although in strict lab conditions, such situations may be uncommon (as subjects are requested to walk according to a strict experimental protocol), in more practical situations they are rather common. Therefore, in most practical cases, possible differences, e.g., due to heavy clothing or carried objects, will introduce variability that can prevent recognition. At the same time, however, these differences may contain discriminative information, e.g, the subject's walking style. In order to remove the effect of structural differences, while retaining the useful gait information, we develop two gait representations based on the average frame: the Shifted Energy Image (SEI) and the Gait Structural Profile (GSP).

V. Gait representation using the Shifted Energy Image (SEI)

Even if a specific person's walking style may be different when he/she walks on two different occasions, the general shape and the size of this person's body components remain the same and maintain their discriminating capacity. However, the position of the gravity centre of the body may change due to the existence of factors like carried objects, heavy clothing, fatigue, etc. Consequently, due to such variations, the direct comparison of averaged silhouette templates can lead to erroneous decisions. For this reason, we propose a new representation based on
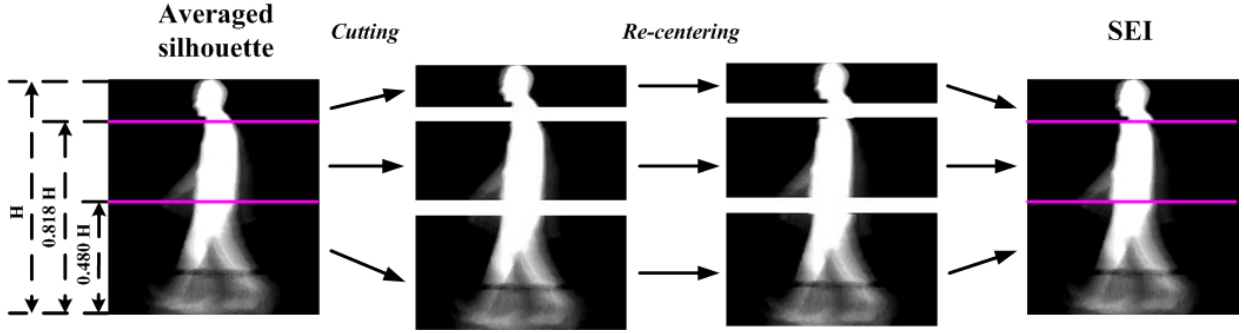
Fig. 4.   The construction of SEI.

the averaged silhouette. Due to its robustness to structural variations, the new feature exhibits enhanced discriminating performance.

The construction of the new representation is graphically presented in Figure 4. First, since all the silhouettes are scaled and their heights are normalized, it is mainly the horizontal component of the gravity centre that will be affected by structural variations. This consideration motivated us to divide the averaged silhouette into three predefined areas, *head*, *torso* and *legs*. The reason why three and no more areas were chosen was the expectation that most appearance changes due to heavy clothing or carried objects would mainly affect one of the above areas and, therefore, could be mitigated by its re-centering. In the experimental results section we confirm the validity of this assumption by performing tests using additional segments. According to the anatomical model in [26], the heights of a human's shoulder and his / her pelvis are respectively equal to 81.8% and 48% of the total body height. Therefore, if the height (in pixels) of the scaled silhouette is H, the head area includes the first $0.182H$ rows of the silhouette, and the torso and legs include the next $0.338H$ and $0.48H$ rows respectively.

After the areas are defined on an averaged silhouette, their horizontal centres are calculated. Specifically, if the number of foreground pixels in the head area $\mathcal{A}_h$ is N_h, and the intensity of pixel (x, y) in the averaged silhouette is denoted by $I(x, y)$, the new horizontal centre x_{c_h} of the head area can be calculated as:

$$x_{c_h} = \frac{1}{N_h} \sum_{(x,y) \in \mathcal{A}_h} x \cdot I(x, y) \tag{2}$$

The new horizontal centres of the torso (x_{c_t}) and the leg (x_{c_l}) areas can be trivially calculated in a
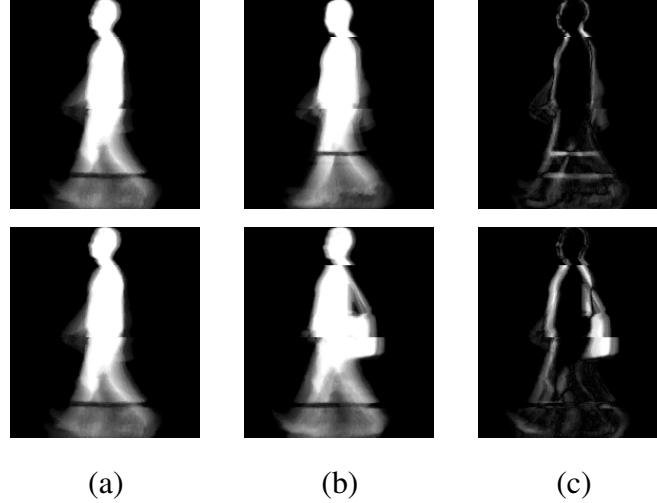
Fig. 5.   SEIs of the same subject: (a) Gallery SEI, (b) Probe SEI, (c) Difference between Gallery and Probe SEIs.

similar manner. Subsequently, the three trunks are shifted according to their centres separately, in order to obtain the Shifted Energy Images (SEI). If the horizontal centre of the averaged silhouette is x_{c_I}, we denote the SEI of the head, torso and leg area by $\mathbf{Q}_h$, $\mathbf{Q}_t$ and $\mathbf{Q}_l$ respectively, and they can be calculated as:

$$\begin{aligned}
\mathbf{Q}_h(x,y) &= I\left(x + (x_{c_h} - x_{c_I}), y\right), & 0.818H \leq y < H \\
\mathbf{Q}_t(x,y) &= I\left(x + (x_{c_t} - x_{c_I}), y\right), & 0.48H \leq y < 0.818H \\
\mathbf{Q}_l(x,y) &= I\left(x + (x_{c_l} - x_{c_I}), y\right), & 0 \leq y < 0.48H
\end{aligned} \tag{3}$$

An example of the impact of the SEI in the evaluation of the difference between different gait sequences of the same subject is shown in Figure 5. As seen, the difference between different gait sequences of the same subject is considerably smaller than the one calculated using averaged silhouettes (see Figure 3). This is particularly noticeable in the head area, where the gravity centre has undergone a very noticeable displacement.

A. Segment partitioning insights

In the above analysis, our method was presented using three segments determined based on an anthropometric model. However, the number of segments used could be higher. In order to assess the performance of our method for higher numbers of segments, we explored ways in which each average silhouette can be partitioned into horizontal segments. To this end, we
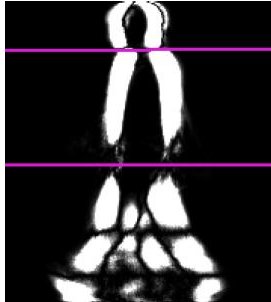
Fig. 6. Energy of the difference between the two global templates.

calculated two global templates, i.e., templates that are calculated based on all gait sequences in our training set. The first template is the average of all silhouettes in our training set. The second template is calculated by re-centering on a row-by-row basis of the average silhouette for each subject in our training set, followed by averaging over all subjects. The calculation of the two templates is meant to highlight the silhouette areas that are mostly affected by the re-centering as well as disclose possible areas that are robust to the re-centering operation.

The squared difference between the the two global templates is shown in Figure 6. As seen, there are several low-energy horizontal lines on the difference template that appear to act as boundaries between wide horizontal segments that are mostly affected by re-centering. Interestingly, these segments are remarkably similar to the ones determined based on the anthropometric model discussed earlier. This similarity is evident by comparing the segments shown in Figure 6 to those shown in Figure 4. This conclusion offers concrete supportive evidence about the validity of the semantic segments determined based on the model. In the experimental results section, we include additional experimental results and analysis for SEI construction based on three or more segments.

B. Dimensionally Reduction for SEIs

The final SEI template is formed using the shifted segments described above. Since not all SEI coefficients are important for recognition, we perform Linear Discriminant Analysis (LDA) in order to concentrate the discriminatory information on a few subspace coefficients [27]. Linear discriminant analysis provides a systematic methodology for the calculation of a subspace in which recognition is easier. In particular, a matrix $\mathbf{W}$ is calculated such that the Fisher's criterion
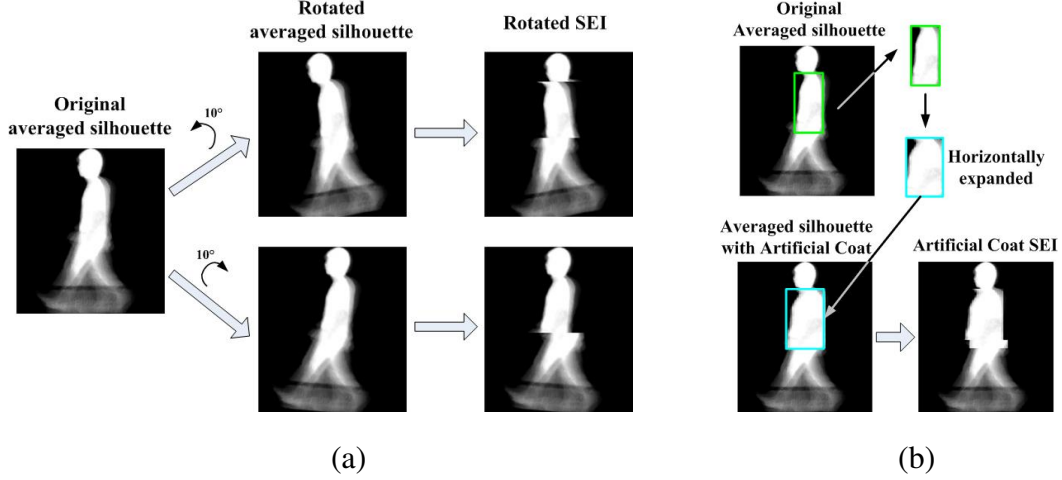
Fig. 7. (a) Construction of rotated SEI, (b) Construction of SEI with artificial coat.

is maximized:

$$J(\mathbf{W}) = \frac{|\mathbf{W} \cdot \mathbf{S}_B \cdot \mathbf{W}^T|}{|\mathbf{W} \cdot \mathbf{S}_W \cdot \mathbf{W}^T|} \tag{4}$$

where $\mathbf{S}_B$ is the between-class scatter matrix, and $\mathbf{S}_W$ is the within-class scatter matrix. In order to calculate $\mathbf{S}_B$ and $\mathbf{S}_W$, we use the method in [28] to construct feature vectors from the templates. If the dimensions of a template $\mathbf{Q}$ are $M \times N$, it is converted into an $I \times 1$ vector $\mathbf{q}$, where $I = M \times N$. Let N_g denote the number of classes (i.e., the number of different subjects) in the Gallery set, and N_s denote the total number of feature vectors from all N_g classes. Trivially, $N_s = \sum_{i=1}^{N_g} N_i$, with N_i denoting the number of feature vectors of the ith class. The within-class scatter $\mathbf{S}_W$ is defined as:

$$\mathbf{S}_W = \sum_{i=1}^{N_g} \sum_{\mathbf{q} \in \mathcal{C}_i} (\mathbf{q} - \mathbf{m}_i)(\mathbf{q} - \mathbf{m}_i)^T \tag{5}$$

where $\mathcal{C}_i$ is the ith class, $\mathbf{q}$ is one of the feature vectors of the ith class, and $\mathbf{m_i} = \frac{1}{N_i} \sum_{\mathbf{q} \in \mathcal{C}_i} \mathbf{q}$, is the mean feature vector of the ith class. The between-class scatter $\mathbf{S}_B$ is defined as:

$$\mathbf{S}_B = \sum_{i=1}^{N_g} N_i (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T \tag{6}$$

where $\mathbf{m}_i$ is the mean feature vector of the ith class, and $\mathbf{m} = \frac{1}{N_g} \sum_{i=1}^{N_g} \mathbf{m}_i$, is the global mean feature vector.

The application of LDA relies on the availability of more than one training feature vectors (SEIs) for each class. However, we would like to be able to use our proposed algorithm even
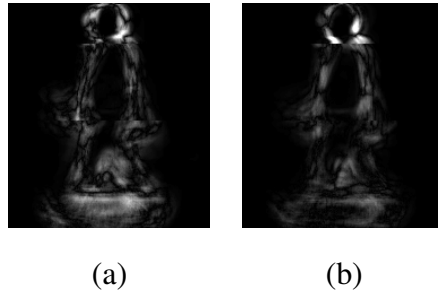
(a) (b)

Fig. 8. Eigenvectors corresponding to the two largest eigenvalues. (a) corresponds to the largest eigenvalue, (b) corresponds to the second largest eigenvalue.

in cases where only one full gait cycle is available. This is why we resort to constructing three SEIs for each subject based on a single full gait cycle and its two constituent half-cycles. Since the three SEIs are constructed based on a single gait cycle, the resultant within-class scatter matrix may not be very appropriate for use in the application of LDA. In order to deal with this problem, we introduce within-class variance using the three SEIs constructed earlier. As a result, we develop two novel approaches for the construction of artificial templates by simulating two conditions: 1) *the subject is leaning forward when walking*, 2) *the subject is wearing a coat*. Additional variability could be used by using different view-direction variations, but we chose not to take this approach as our focus is on dealing with real structural variations rather than appearance changes that are due to viewpoint variations.

The first condition can be simulated by rotating the original averaged silhouette by r degrees using its gravity centre as the centre of rotation. In our experiments, we used $r = 10°$ as this is close to the maximum inclination that would be realistically expected when walking. Using smaller values would gradually erode the efficiency of LDA. For reasons of symmetry, we also create a rotated template using the opposite rotation $(-r)$. Subsequently, two new SEIs are constructed based on the rotated averaged silhouettes using equation (3). This construction is shown in Figure 7(a).

The existence of a coat can be simulated by taking the following steps, which are graphically described in Figure 7(b): 1) In the torso area of the SEI, we locate pixels, (x_l, y_l) and (x_r, y_r), that indicate the front and back boundaries of the torso. 2) We extract the torso area between x_l to x_r and expand the extracted segment horizontally by 50% so that it resembles a coat.

Although smaller expansion factors are conceivable, expanding by considerably less than 50% would diminish the efficiency of LDA to handle different clothing and the performance achieved would be close to that exhibited by a system not using LDA. 3) We use the artificially expanded area in place of the extracted area. 4) We construct a new SEI by using the averaged silhouette with artificial coat.

The combination of the SEI construction and the subspace projection allows the selection of the most discriminative features of an SEI. Using LDA, a SEI can be represented as a short feature vector of projection coefficients derived using the eigen-SEIs of the LDA. Figure 8 shows the first two eigen-SEIs of a subspace projection.

VI. STRUCTURAL GAIT REPRESENTATION USING THE GAIT STRUCTURAL PROFILE (GSP)

The SEI, presented in the previous section, distorts the structural coherence of averaged silhouettes in order to achieve improved insensitivity to a variety of changes. However, structural information can be useful in gait recognition and this is why we follow an additional approach that can capture body geometry without any distortion. This approach is based on vectors defined by the gravity centre of the body as well as the centres of a few selected areas in each averaged silhouette. These vectors have been shown [29], [30] to contain high discriminatory power. In the present paper, instead of extracting information from each of the frames separately, we propose a gait representation based on vectors calculated directly from the averaged silhouette. This new representaion will be, henceforth, called the Gait Structural Profile (GSP) and it is graphically explained in Figure 9.

Unlike the method in [30], the present method does not require the availability of manually-labelled silhouettes. Instead, initially the averaged silhouette is segmented into four predefined body areas - *head*, *torso*, *left leg* and *right leg* - based on anatomical knowledge of the human body [26] and subsequently the gravity centre of each body area is calculated. If the gravity centre of the entire body and the *head* are denoted $\mathbf{g}$ and $\mathbf{g}_h$ respectively, the GSP vector for the head area is calculated as:

$$\mathbf{p}_h = \mathbf{g}_h - \mathbf{g} \tag{7}$$

i.e., it represents the co-ordinates of the centre of the head area in the co-ordinate system that has as origin the gravity centre of the entire body. The GSPs of the other body areas are trivially calculated in a similar way.
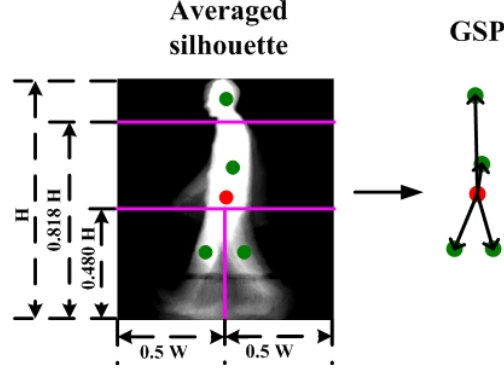
Fig. 9. The construction of GSP.

When the GSP distance between two subjects is calculated, the distances of the gallery and the probe subjects are calculated separately for each area α, i.e.,

$$d_{GSP_\alpha} = \|\mathbf{p}_\alpha^{gallery} - \mathbf{p}_\alpha^{probe}\| \tag{8}$$

where $\|\cdot\|$ denotes Euclidean distance and the subscript α denotes the relevant body area with $\alpha \in \{head,\ torso,\ left\ leg,\ right\ leg\}$.

After the distances are calculated for each of the four body parts, they are combined into a single dissimilarity measure D_{GSP} using weighting. Specifically, since the discriminatory capabilities of the four body areas are different, we develop a weighting method based on the *within* and *between* variances of the gallery subjects. The within-variance $\sigma^2_{within,\alpha}$ of $\mathbf{p}_\alpha$ is calculated as:

$$\sigma^2_{within,\alpha} = \frac{1}{N_g} \sum_{i=1}^{N_g} \sum_{\mathbf{p}_a \in \mathcal{C}_i} \|\mathbf{p}_\alpha - \overline{\mathbf{p}}_{\alpha,i}\|^2 \tag{9}$$

where $\mathbf{p}_a \in \mathcal{C}_i$ represents instances of the ath body area in the ith class, N_g is the number of different subjects (i.e., classes) in the gallery and $\overline{\mathbf{p}}_{\alpha,i}$ is the average GSP vector for area α of subject i. The between-variance of the $\mathbf{p}_\alpha$, denoted as $\sigma^2_{between,\alpha}$, is calculated as:

$$\sigma^2_{between,\alpha} = \frac{1}{N_g \cdot (N_g - 1)} \sum_{\substack{i=1 \\ i \neq f}}^{N_g} \sum_{f=1}^{N_g} \|\overline{\mathbf{p}}_{\alpha,i} - \overline{\mathbf{p}}_{\alpha,f}\|^2 \tag{10}$$

where $\overline{\mathbf{p}}_{\alpha,i}$ and $\overline{\mathbf{p}}_{\alpha,f}$ are the average GSP vectors of subject i and f for $i \neq f$.

Apparently, a large within-variance leads to large False Reject Rates whereas a small between-variance leads to large False Accept Rates. Our objective is to ensure the robustness of the system by minimizing the impact of both cases above. To this end, the importance of the αth area is taken into account by utilizing a weight calculated as:

$$w_{GSP_\alpha} = \frac{\sigma^2_{between,\alpha}}{\sigma^2_{within,\alpha}} \tag{11}$$

where α denotes the head (h), torso (t), left leg (ll), and the right leg (rl) area. Weights calculated as above offer a clear performance improvement over using equal weights.

Finally, the total distance for the GSP representation is defined as:

$$D_{GSP} = \mathbf{w}^T_{GSP} \cdot \mathbf{d}_{GSP} \tag{12}$$

where

$$\mathbf{w}_{GSP} = [w_{GSP_h}, w_{GSP_t}, w_{GSP_{ll}}, w_{GSP_{rl}}]^T$$

$$\mathbf{d}_{GSP} = [d_{GSP_h}, d_{GSP_t}, d_{GSP_{ll}}, d_{GSP_{rl}}]^T$$

In the above analysis, we used four body areas. Alternative choices about the number of assumed body areas could have been made. Related experimental results and analysis, showing the impact of alternative choices to recognition performance, are included in section VIII-A2 showing the impact of alternative choices to recognition performance.

VII. Feature Combination

As is obvious from our discussion so far, the gait representation proposed in this paper consists in the division of the averaged silhouette into predefined areas, the extraction of characteristic vectors for these areas and the construction of new templates based on re-centered areas. This approach allows improved performance without disregarding important structural information.

As seen in the previous sections, the distances among the SEI templates and among the GSP vectors are calculated separately. In order to achieve better recognition performance, we combine these distances. Since D_{SEI} and D_{GSP} are not directly comparable, we first normalize them using the min-max normalization method. Because the discriminatory power of the two features is different, weights for the normalized distances are calculated using the mean-variance method. This method yielded superior results in comparison to other normalization methods [31]:

$$w_{SEI} = \frac{m_{SEI}}{\sigma_{SEI}}, \quad w_{GSP} = \frac{m_{GSP}}{\sigma_{GSP}} \tag{13}$$

where m_{SEI} (m_{GSP}) and σ_{SEI} (σ_{GSP}) are the mean Euclidean distances and the standard deviations between different gallery subjects for the SEI (GSP) representations.

The final distance for the combination of the two features is:

$$D = w_{SEI} \cdot \widetilde{D}_{SEI} + w_{GSP} \cdot \widetilde{D}_{GSP} \tag{14}$$

where $\widetilde{D}_{SEI}$ and $\widetilde{D}_{GSP}$ are the normalized distances using the SEI and GSP representations respectively. This distance is used for making recognition decisions in our final system, named *SEIS*, based on the *SEI* representations as well as on the extracted *Structural* features.

VIII. EXPERIMENTAL RESULTS

By using the pre-processing and gait representations that were introduced in the present paper, we developed a complete gait recognition / authentication system that can work in real time. Our system is based on a camera from Point Grey Research[2] for video capturing. The entire system is implemented in C++ and can run on a laptop with Intel Core 2 Duo 2.66GHz CPU, 4 GB RAM and 64-bit Windows Vista Operating System.

The computational efficiency of our proposed method is presented in Table I, in which the average running time is shown for enrolment and recognition using the GEI, SEI and SEIS methods respectively. The pre-processing stage (i.e., initial background subtraction and silhouette extraction) is common for all methods and for this reason it is omitted. It should be noted, however, that the pre-processing is taking place in parallel with video capturing and, therefore, the required preprocessing time (3-4 seconds) is not additional to the time needed for the rest of the calculations.

The enrolment process includes side-view extraction and gait template construction, and the recognition process includes side-view extraction, gait template construction and decision making. As seen the additional computations needed by SEI in comparison to GEI are minor. Our complete SEIS system is more complex, but still fast enough to operate close to real time.

---

[2]http://www.ptgrey.com/

TABLE I

COMPUTATIONAL EFFICIENCY OF THE PROPOSED GAIT RECOGNITION SYSTEM (IN SECONDS).

Method	Enrolment			Recognition			
	Side-view extraction	Template construction	Total	Side-view extraction	Template construction	Comparison and decision making	Total
GEI	$0.4s$	$0.1s$	$0.5s$	$0.2s$	$0.1s$	$1.3s$	$1.6s$
SEI	$0.4s$	$0.2s$	$0.6s$	$0.2s$	$0.2s$	$1.5s$	$1.9s$
SEIS	$0.4s$	$0.3s$	$0.7s$	$0.2s$	$0.2s$	$2.5s$	$2.9s$

A. Sensitivity to variations in the degrees of freedom

In order to evaluate the sensitivity of our algorithm to variations in the number of SEI segments or GSP body areas, we conducted comparative experiments that are presented below.

1) Body-part segmentation for SEI: According to the anatomical model and our proposed SEI feature extraction method, there are limited ways of forming body segment sets, a) three segments: head-torso-legs, b) four segments: head-torso-thighs-shins, c) five segments: head-chest-abdomen-thighs-shins. We evaluate the above segment sets using the ACTIBIO database (a detailed description of the database is provided later in this section), and the results are shown in table II. As seen, the best performance is achieved by using the three-segment set (i.e., head, torso, and legs) although the performance differences in case more segments are used are generally marginal.

2) Body-part combinations for GSP: Similar to SEI construction, more than one segment-set formations can be applied to GSP extraction. In order to assess our method, we tried the following: a) 4 segments: head - torso - left leg - right leg, b) 6 segments: head - torso - left thigh - right thigh - left shin - right shin, c) 8 segments: head - left torso - right torso - left thigh - right thigh - left shin - right shin. We extracted GSP features based on the above three combinations, and combined each GSP with the three-segment SEI (i.e., the best performing SEI) using the fusion method presented in section VII. The combined results are shown in table III. As seen, the best performance is achieved by using four segments (i.e., head, torso, left leg and right leg), although in several cases the performance differences are small. The results show the intrinsic efficiency of the approach taken and highlight its general insensitivity to variations

TABLE II

RANK 1 AND RANK 5 RECOGNITION RESULTS FOR SEI USING DIFFERENT NUMBERS OF BODY SEGMENTS.

	Rank 1			Rank 5		
segments	3	4	5	3	4	5
Probe A	**29**	28	25	**58**	57	54
Probe B	**16**	15	11	**45**	42	40
Probe C	42	**47**	38	76	**77**	76
Probe D	**42**	36	40	**82**	**82**	81
Probe E	78	80	**81**	**92**	89	88
Probe F	**49**	45	45	**60**	55	58
Probe G	30	29	**32**	52	52	**55**
Average	**41**	40	39	**66**	65	65

TABLE III

RANK 1 AND RANK 5 RECOGNITION RESULTS FOR GSP COMBINED WITH SEI, WHERE GSP IS USING DIFFERENT NUMBERS OF BODY SEGMENTS.

	Rank 1			Rank 5		
segments	4	6	8	4	6	8
Probe A	**29**	**29**	**29**	**58**	**58**	**58**
Probe B	**19**	16	16	**47**	45	44
Probe C	48	48	**52**	76	76	**78**
Probe D	**41**	**41**	38	**82**	**82**	81
Probe E	78	78	78	**93**	92	92
Probe F	**52**	45	48	**63**	60	60
Probe G	**29**	**29**	**29**	**52**	**52**	**52**
Average	**42**	41	41	**67**	66	66

of the free parameters.

B. *Experimental evaluations using three different gait databases*

Our proposed algorithm is tailored to gait recognition that is robust to body shape changes. For the experimental evaluation of our method, we use the CMU database [18], the CASIA database [19] and the gait database that was recorded in the framework of the ACTIBIO

project (Unobtrusive Authentication Using ACTIvity Related and Soft BIOmetrics)[3]. In the above databases, there are test gait sequences in which the walking person is carrying a bag or wearing a coat, which facilitates the demonstration of the benefits of our approach.

C. Experimental evaluation using the CMU MoBo database

The CMU database is an established database and in the present work we use it in order to compare our proposed method with other methods in the literature. In the CMU MoBo database, 25 walking subjects were captured under three conditions, i.e., *fastwalk*, *slowwalk*, and *walking with a ball*, and from six view angles, i.e., *east*, *southeast*, *south*, *southwest*, *northwest* and *north*. Based on the CMU database, the best performing method that uses side-view sequences is the one using pHMMs [9], and this is why we include it in our comparison. In addition, we compare with the popular averaged silhouette template in [5], which is widely known as GEI [6], as well as with the multiview method introduced in [32]. It should be noted that the purpose of this comparison is to comparatively assess the efficiency of gait representations in their simplest form, and therefore *no LDA was applied in combination with any of the representations in the comparison*.

In our experiments, we used the *fastwalk* sequence as the the gallery (reference) set while *slowwalk* was used as the probe set. We only used the side-view sequences for the proposed gait template construction. Results are reported in terms of cumulative match scores. Essentially, we measure rank-*n* performance, i.e., the percentage of tests in which the correct subject appears in the top *n* matches [33]. The results for the proposed algorithm and the other three existing methods are tabulated in Table IV. It can be seen that the proposed algorithm outperforms the other two methods, based on the averaged silhouette (GEI) and the pHMM. Moreover, despite the fact that the proposed system detects and uses only side-views, it still rivals the recognition performance of the system that combines multiple views [32]. This provides further evidence about the robustness of the proposed system.

D. Experimental evaluation using the CASIA database

The CASIA database is one of the largest publicly available gait databases. It consists of three data sets, which are captured in different environments. In our experiments, we use Dataset B,

---

[3]For more information please see http://www.actibio.eu/

TABLE IV

RECOGNITION RATES FOR THE PROPOSED AND THREE OTHER EXISTING METHODS ON THE CMU DATABASE.

Method	Recognition Rate (%)			
	Rank 1	Rank 3	Rank 5	Rank 7
GEI [6]	84	92	92	96
pHMM [9]	84	-	-	-
Multi [32]	92	96	96	100
SEI (non-LDA)	88	96	96	100
SEIS (non-LDA)	**92**	**96**	**100**	**100**

TABLE V

RECOGNITION RATES FOR THE PROPOSED AND TWO OTHER STATE-OF-THE-ART METHODS ON THE CASIA DATABASE.

Probe	Rank 1 Recognition Rate (%)				
	CAS [19]	GEI [6]	SEI	SEI(LDA)	**SEIS**
Set A	98	99	99	99	**99**
Set B	33	60	42	63	**64**
Set C	52	22	63	72	**72**

because it has a large number of subjects and the walking conditions are close to the ones that our system aims to work in. There are 124 subjects in CASIA Dataset B, and the gait data were captured from 11 views. In addition, three walking conditions are included, namely, normal walking (*"nm"*), walking in a coat (*"cl"*) and walking with a bag (*"bg"*).[4]

For the experimental evaluation based on CASIA database, we followed the protocol described in [19]. Specifically, we use sequences *"nm-01"*, *"nm-02"*, *"nm-03"* and *"nm-04"* as the gallery (reference) set, sequences *"nm-05"* and *"nm-06"* as probe (test) Set A, sequences *"cl-01"* and *"cl-02"* as probe Set B, and sequences *"bg-01"* and *"bg-02"* as probe Set C.

Rank 1 recognition rates (i.e., correct classification rates) using our method, as well as two other existing methods, are tabulated in Table V. The CAS (Chinese Academy of Sciences) method is based on GEI, and its recognition results for all 11 views were initially reported in [19]. In that work, it was concluded that the best results were achieved using the side view (i.e., 90° view-angle), which is in agreement with the assumptions made in our proposed method.

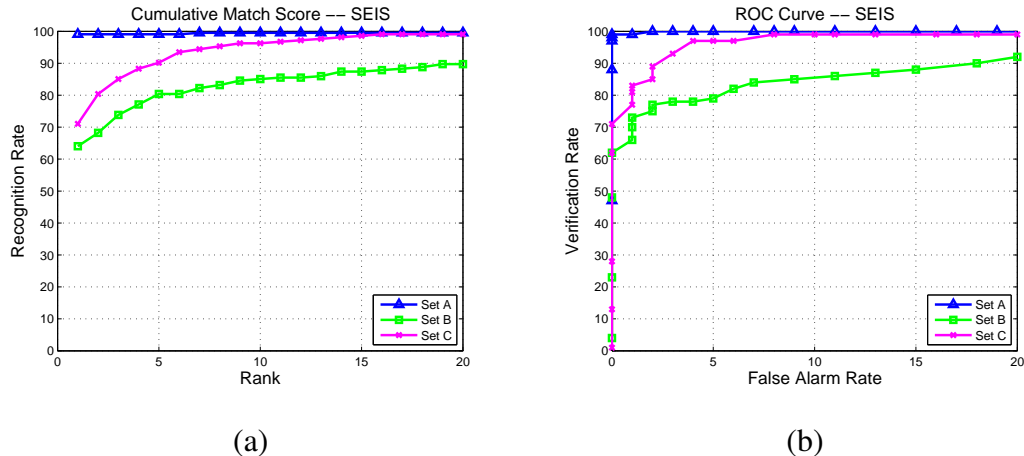[4]http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp

Fig. 10. Experimental results on the CASIA database using the proposed SEIS algorithm (i.e., combination of SEI and GSP): (a) Recognition results, (b) Verification results.

For this reason, we only compare our results to the side-view CAS results. As seen in Table V, the recognition rates achieved using SEI without LDA are superior to those achieved using the CAS method (i.e., using simple GEI) in all probe sets. That clearly demonstrates the advantage of our proposed SEI template over the conventional GEI template. The method developed at the University of California at Riverside [6] is also based on GEI but it additionally deploys Principal Component Analysis (PCA) and Multiple Discriminant Analysis (MDA) for dimensionality reduction. In our proposed method, we apply LDA on the SEI for dimensionality reduction, therefore, it is reasonable to compare the results of SEI with LDA to that method. As seen in Table V, our SEI-LDA method consistently outperforms the two other state-of-the-art methods in all probe sets. Finally, further improvement is achieved when SEI is combined with structural information. The resultant SEIS algorithm is the best-performing method, outperforming all others in Table V, and its full recognition and verification results are presented in Figure 10.

E. Experimental evaluation using the ACTIBIO database

The ACTIBIO database was compiled in 2009 in the framework of the ACTIBIO project[5] and is probably the newest such database. In this database, there are 28 subjects who are walking in an indoor environment. Each subject is not necessarily walking in straight lines but instead

---

[5]http://www.actibio.eu/

TABLE VI

Eight different sets in the ACTIBIO database.

Set	Day	Rep.	Subjects	Condition
Gallery	1	1	28	Normal walking
Probe A	1	1	28	Carrying a bag
Probe B	1	1	27	Wearing a coat
Probe C	1	1	28	Wearing slippers or socks only
Probe D	1	1	28	Walking diagonally
Probe E	1	2	28	Normal walking
Probe F	2	1	27	Normal walking
Probe G	2	2	27	Stop for a while

he / she is walking casually along random paths. Therefore, this database is very suitable for the experimental assessment of our proposed algorithm. For each subject, there are up to eight gait sequences, captured on two recordings with the second recording taking place one month after the first. Two repetitions were captured in each day and several walking conditions were recorded in each repetition. A full description of each recorded condition is presented in Table VI. For each subject, one of the eight sequences is used as a gallery (reference) sequence and the others are probe (test) sequences. Figure 11 shows sample frames taken from two sequences of the same subject. As seen,

- in both sequences, the subject is not walking in a strictly straight line, i.e., the view-angle of the frames change during the walking period,
- the body is not always kept erect; in sequence 1 (upper row), the subject is walking with his head inclined downwards, while in sequence 2 (lower row), the same subject is looking straight ahead when walking,
- in sequence 2, the subject is carrying a bag.

In our experiments, we used our side-view partitioning algorithm in order to extract one side-view cycle out of each gait sequence. Subsequently, we constructed our proposed gait representations using the extracted side-view cycle.

First, we assessed the performance of the proposed SEI representation in comparison with the GEI. To this end, we constructed the GEI and SEI separately *without* applying Linear
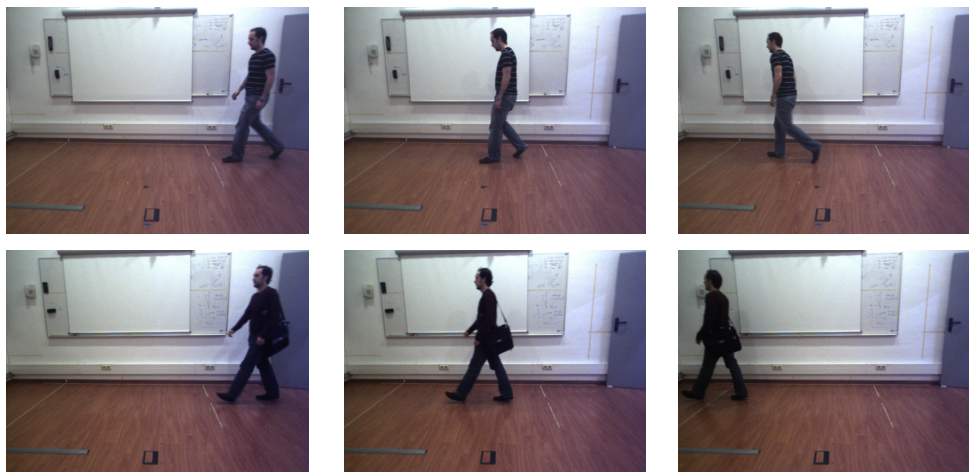
Fig. 11.   Sample frames from the ACTIBIO database.

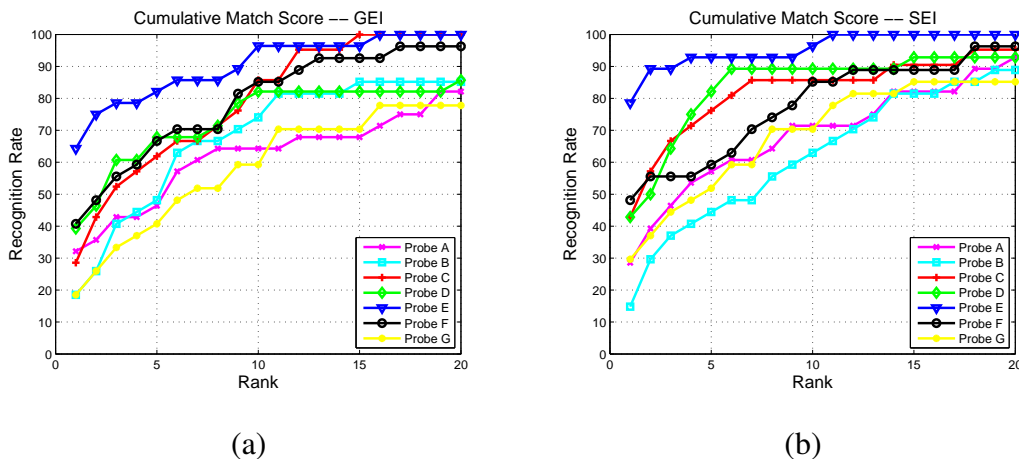(a)   (b)

Fig. 12.   Recognition rates on the ACTIBIO database using: (a) GEI, (b) SEI. No LDA was used with either of the tested algorithms.

Discriminant Analysis. The recognition results of using these two representations are shown in Figure 12. The Rank 1 and Rank 5 performance is also summarized in the first two columns of Table VIII. Furthermore, in Figure 13, we report verification results in terms of Receiver Operating Characteristic curves (ROC). The ROCs report the probability of positive recognition of an authorized subject versus the probability of granting access to an unauthorized subject. *As seen in all comparisons, the SEI outperforms the GEI at almost every rank for all probe sets, clearly demonstrating the advantage gained when the SEI is used.*
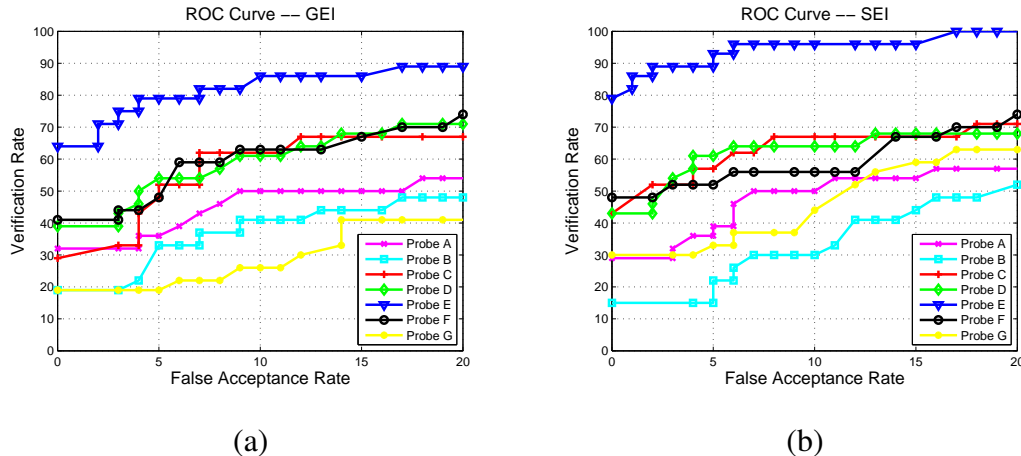
Fig. 13. Verification rates on the ACTIBIO database using: (a) GEI, (b) SEI. No LDA was used with either of the tested algorithms.

TABLE VII

AVERAGE RECOGNITION AND VERIFICATION RATES ON THE ACTIBIO DATABASE USING THE SEI-LDA METHOD WITH DIFFERENT TRAINING SETS .

Training sets	Recognition Rate (%)		Training sets	Verification Rate (%)	
	Rank 1	Rank 5		FAR 5%	FAR 10%
I	45	77	I	50	61
II	53	82	II	55	67
III	**59**	**88**	**III**	**63**	**75**

Secondly, we applied LDA on the original SEIs. The subspace projection transformed the original SEIs into short vectors of $N_g - 1$ coefficients, where N_g is the number of subjects in the gallery, i.e., 28 in this case. In order to improve the efficiency of the calculation of the appropriate subspace, as explained in Section V-B, we created artificial gait subjects by using rotation of gait half-cycles as well as introducing artificial clothing. In order to assess the efficiency of the artificial training sets, several experiments are conducted using different training sets:

- *Training sets I:* They are using only the original templates that are created from the full gait cycle and the half-cycles.

- *Training sets II:* The artificial sets created by rotation are included.

- *Training sets III:* Both artificial sets with rotation and extra clothing, are included.
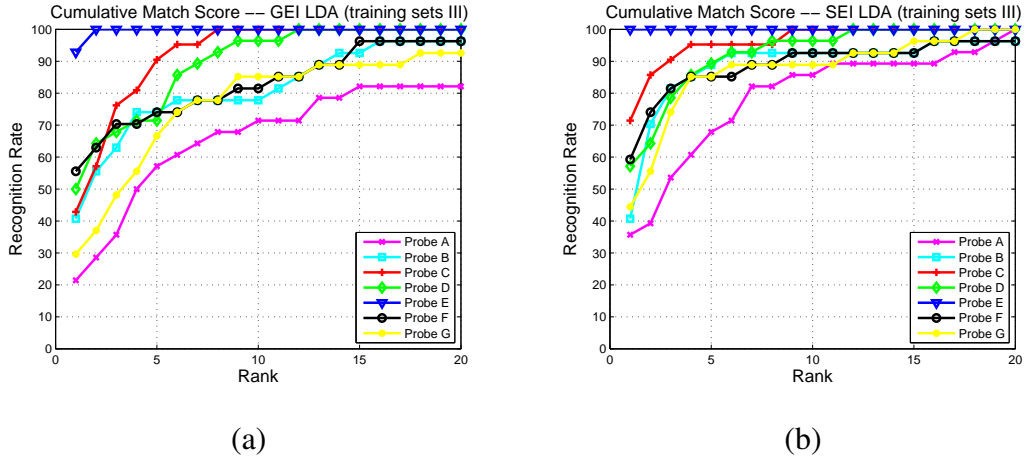
Fig. 14. Recognition rates on the ACTIBIO database: (a) GEI with LDA using training sets with artificial rotation and clothing, (b) SEI with LDA using training sets with artificial rotation and clothing.
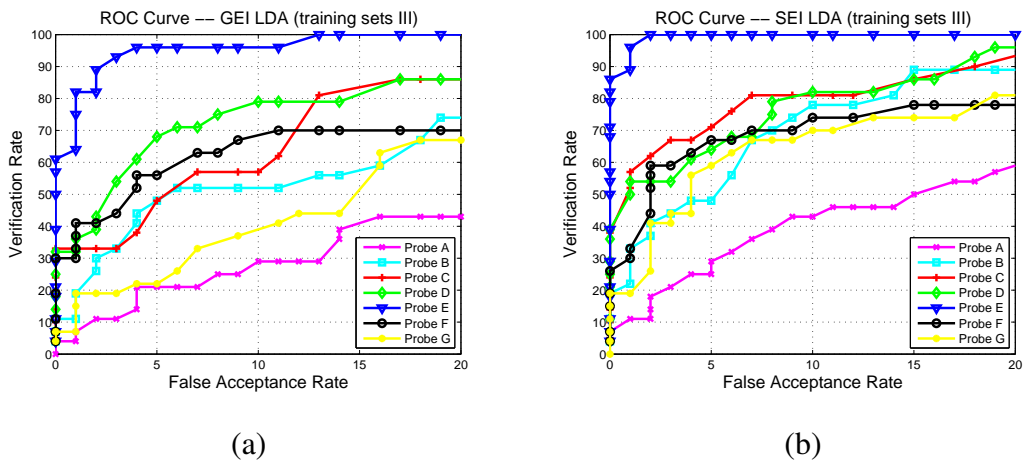

Fig. 15. Verification rates on the ACTIBIO database: (a) GEI with LDA using training sets with artificial rotation and clothing, (b) SEI with LDA using training sets with artificial rotation and clothing.

The recognition and verification results of the above three experiments are shown in Table VII. As seen, the results using rotated training sets are superior to those obtained using original templates only. The combination of both rotated sets and artificial clothing sets yields even better results. This proves that the rotation and artificial clothing training sets improve the system performance. For the purpose of comparison, the recognition and verification results by applying LDA on the GEI using the *training sets III* are also provided in Figure 14 and 15. As seen, the approach using SEI with LDA still outperforms the approach using GEI with LDA.
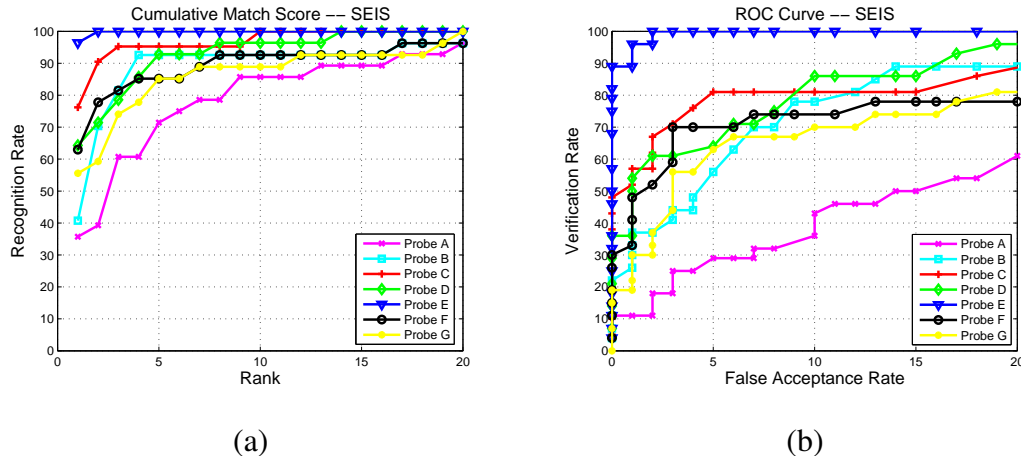
Fig. 16. Experimental results on the ACTIBIO database using the proposed SEIS algorithm (i.e., combination of SEI and GSP): (a) Recognition results, (b) Verification results.

TABLE VIII

RANK 1 AND RANK 5 RECOGNITION RESULTS ON THE ACTIBIO DATABASE.

Probe	Rank 1 Recognition results					Probe	Rank 5 Recognition results				
	GEI	SEI	GEI(LDA)	SEI(LDA)	**SEIS**		GEI	SEI	GEI(LDA)	SEI(LDA)	**SEIS**
Probe A	32	29	21	36	**37**	Probe A	47	58	58	69	**71**
Probe B	19	16	41	41	**41**	Probe B	49	45	75	89	**92**
Probe C	29	42	42	71	**77**	Probe C	61	76	90	96	**96**
Probe D	39	42	50	58	**64**	Probe D	69	82	72	89	**92**
Probe E	64	78	92	**100**	97	Probe E	82	92	100	100	**100**
Probe F	41	49	57	60	**62**	Probe F	67	60	75	85	**85**
Probe G	19	30	30	44	**57**	Probe G	40	52	67	85	**85**
Average	35	41	48	59	**62**	Average	59	66	77	88	**89**

Our final and most important experiment is based on the combination of the SEI-LDA with GSP, as described in Section VII. The recognition and verification results for the resulting method using the SEI and Structural features (SEIS) are shown in Figure 16.

For the sake of comparison, for all above approaches, the recognition results at Rank 1 and Rank 5 and the Equal Error Rates (EER), i.e., the error rate when FAR equals FRR, are presented in Table VIII and IX respectively. From these results, the following conclusions can be drawn:

- Compared to the GEI, a clear advantage of the SEI is shown. On average, the SEI yields 6% and 7% recognition rate improvements at rank 1 and rank 5 respectively. Additionally,

TABLE IX

EQUAL ERROR RATE (EER) ON THE ACTIBIO DATABASE.

Probe	Equal Error Rate (%)				
	GEI	SEI	GEI(LDA)	SEI(LDA)	**SEIS**
Probe A	39	36	32	**29**	31
Probe B	30	35	24	15	**15**
Probe C	24	25	10	10	**10**
Probe D	25	27	18	11	**11**
Probe E	15	7	2	**0**	1
Probe F	23	25	22	17	**15**
Probe G	37	26	26	15	**15**
Average	28	26	19	14	**14**

the EER of the SEI is lower than that of the GEI.

- After the proposed LDA method is applied, improvements on both GEI and SEI are achieved. In this case, SEI still outperforms GEI by a significant margin. Specifically, when using LDA, on average, the SEI gives 11% higher recognition rate than GEI at both rank 1 and rank 5. Moreover, our method achieves a lower EER.

- The combination of SEI-LDA and GSP, i.e., the SEIS method, achieves even better rank 1 and rank 5 recognition rates although these additional improvements are occasionally marginal.

For reasons of completeness, we also tested our method using low-quality silhouettes from the USF database. Direct comparison between the GEI and the SEI (without application of LDA) on the average brings the performance of the two methods within two percentage points of each other. This is due to the fact that the segment displacement of SEI becomes largely inaccurate as a result of the very low quality of the extracted outdoor silhouettes and, therefore, the expected performance gains materialize in part. Another conclusion is that, in case of low quality silhouettes, the final SEIS method improves on the recognition accuracy of SEI marginally. This should be expected as the discriminatory capacity of the GSP component of the final system is low in the presence of excessive noise. These results underline the fact that the clear performance gains achieved by our method in conditions of medium- or good-quality silhouettes are coupled with very robust performance, which is directly competitive to conventional methods, in cases

of low (or very low) quality silhouettes.

In conclusion, our best-performing SEIS approach, combining template shifts and structural features, generally outperforms all other methods in our extensive experimental assessment. This improved performance of our system comes with the valuable additional convenience of real-time operation in casual walking situations.

IX. CONCLUSIONS

We presented a very efficient, novel gait recognition system. The proposed system is based on two novel gait representations, the Shifted Energy Image and the Gait Structural Profile, that are robust to appearance variations such as those arising from heavy clothing or carried objects. Furthermore, a novel method is introduced for the simulation of walking conditions and the generation of artificial subjects that are used for the application of Linear Discriminant Analysis. At the decision stage, fusion of the two representations was conducted. Extensive experimental assessment using one traditional and two newer databases showed that our proposed system exhibits superior performance in comparison to current state-of-the-art methods for gait recognition.

ACKNOWLEDGEMENT

REFERENCES

[1] X. Huang and N. V. Boulgouris, "Gait recognition for random walking patterns and variable body postures," in *IEEE Internarional Conference on Acoustics, Speech, and Image Processing*, March 2010, pp. 1726–1729.

[2] X. Huang and N. V. Boulgouris, "Gait recognition using linear discriminant analysis with artificial walking conditions," in *IEEE International Conference on Image Processing*, September 2010, pp. 2461–2464.

[3] J.E. Cutting and L.T. Kozlowski, "Recognizing friends by their walk: gait perception without familiarity cues," *Bulletin Psychonometric Soc.*, vol. 9, no. 5, pp. 353–356, 1977.

[4] J. Dugelay and F. Matta, "Physiological and behavioural approaches for person recognition using videos," in *1st workshop on Image Processing Theory, Tools and Applications*, Sousse, Tunisia, Nov. 2008, pp. 1–6.

[5] Z. Liu and S. Sarkar, "Simplest representation yet for gait recognition: Averaged silhouette," in *Int. Conf. on Pattern Recognition*, Aug. 2004.

[6] J. Han and B. Bhanu, "Individual recognition using gait energy image," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 28, no. 2, pp. 316–322, 2006.

[7] K. Bashir, T. Xiang, and S. Gong, "Feature selection on gait energy image for human identification," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Apr. 2008, pp. 985–988.

[8] A. Kale, A. Sundaresan, A.N. Rajagopalan, N. Cuntoor, A.K. Roy-Chowdhury, V. Krueger, and R. Chellappa, "Identification of humans using gait," *IEEE Trans. Image Processing*, vol. 13, no. 9, pp. 1163–6173, Sep. 2004.

[9] Z. Liu and S. Sarkar, "Improved gait recognition by gait dynamics normalization," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 28, no. 6, pp. 863–876, 2006.

[10] D. Ioannidis, D. Tzovaras, I.G. Damousis, S. Argyropoulos, and K. Moustakas, "Gait recognition using compact feature extraction transforms and depth information," *IEEE Trans. on Information Forensics and Security*, vol. 2, no. 3, pp. 623–630, Sep. 2007.

[11] P.J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K.W. Bowyer, "The gait identification challenge problem: Data sets and baseline algorithm," in *Proc. Int. Conf. Pattern Recognition*, Quebec City, Canada, Aug. 2002, pp. 385–388.

[12] H. Lu, K.N. Plataniotis, and A.N. Venetsanopoulos, "A full-body layered deformable model for automatic model-based gait recognition," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, 2008.

[13] M. Goffredo, J.N.Carter, and M.S. Nixon, "Front-view gait recognition," in *IEEE Second Int. Conf. on Biometrics: Theory, Applications and Systems*, Washington D.C., USA, Sep. 2008.

[14] O. Barnich and M. Van Droogenbroeck, "Frontal-view gait recognition by intra- and inter-frame rectangle size distribution," *Pattern Recognition Letters*, vol. 30, no. 10, pp. 893–901, Jul. 2009.

[15] I.F. Nizami, S. Hong, H. Lee, S. Ahn, K. Toh, and E. Kim, "Multi-view gait recognition fusion methodology," in *3rd IEEE Conf. on Industrial Electronics and Applications*, Singapore, Jun. 2008, pp. 2101–2105.

[16] S. Yu, D. Tan, and T. Tan, "Modelling the effect of view angle variation on appearance-based gait recognition," *Computer Vision - ACCV*, vol. 3851, pp. 807–816, 2006.

[17] A. Lopez-Mendez, C. Canton-Ferrer, and J.R. Casas, "Virtual view appearance representation for human motion analysis in multi-view environments," in *European Signal Processing Conference (EUSIPCO)*, Aalborg, Denmark, August 2010, pp. 959–963.

[18] R. Gross and J. Shi, "The CMU Motion of Body (MoBo) database," Tech. Rep. CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Jun. 2001.

[19] S. Yu, D. Tan, and T. Tan, "A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition," in *18th Int. Conf. on Pattern Recognition*, 2006, pp. 441–444.

[20] X. Huang and N.V. Boulgouris, "Robust object segmentation using adaptive thresholding," in *IEEE Int. Conf. on Image Processing*, San Antonio, Texas, USA, Sep. 2007.

[21] S. Sarkar, P.J. Phillips, Z. Liu, I.R. Vega, P. Grother, and K.W. Bowyer, "The human id gait challenge problem: data sets, performance, and analysis," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 27, no. 2, pp. 162–176, Feb. 2005.

[22] C. Canton-Ferrer, J. R. Casas, and M. Pardas, "Human model and motion based 3d action recognition in multiple view scenarios," in *European Signal Processing Conference (EUSIPCO)*, Florence, Italy, September 2006.

[23] A. Kale, A. Chowdhury, and R. Chellappa, "Towards a view invariant gait recognition algorithm," in *Proc. IEEE Conf. on Advanced Video and Signal Based Surveillance*, 2003.

[24] A. Sundaresan, A.K. Roy Chowdhury, and R. Chellappa, "A hidden markov model based framework for recognition of humans from gait sequences," in *Proc. Int. Conf. Image Processing 2003*, Barcelona, Spain, Sep. 2003, pp. 14–17.

[25] N.V. Boulgouris, K.N. Plataniotis, and D. Hatzinakos, "Gait recognition using dynamic time warping," in *Proc. IEEE Int. Symp. Multimedia Signal Processing*, Siena, Italy, Sep. 2004, pp. 263–266.

[26] D.A. Winter, *The Biomechanics and Motor Control of Human Movement*, John Wiley & Sons, 2 edition, 1990.

[27] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*, John Wiley & Sons, Inc., 2001.

[28] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Maui, HI, Jun. 1991, pp. 586–591.

[29] A.F. Bobick and A.Y. Johnson, "Gait recognition using static, activity-specific parameters," in *Proceedings of IEEE Computer Vision and Pattern Recognition*, 2001, vol. 1, pp. 423–430.

[30] X. Huang and N.V. Boulgouris, "Model-based human gait recognition using fusion of features," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Taipei, Apr. 2009, pp. 1496–1472.

[31] A. Ross, K. Nandakumar, and A.K. Jain, *Handbook of Multibiometrics*, Springer, 2006.

[32] X. Huang and N.V. Boulgouris, "Human gait recognition based on multiview gait sequences," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, 2008.

[33] P.J. Phillips, H. Moon, S. Rizvi, and P. Raus, "The feret evaluation methodology for face recognition algorithms," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 10, pp. 1090–1104, 2000.

PLACE PHOTO HERE

Xiaxi Huang received the BEng degree in Control Science and Engineering from Zhejiang University, China, in 2004, the MSc degree in Electrical and Electronic Engineering from University of Bath, UK, in 2005, and the PhD degree in Electronic Engineering from King's College London in 2010, with research interests in digital image/video processing and gait biometrics based on computer vision.

PLACE PHOTO HERE

Nikolaos V. Boulgouris (S'96 M'04 SM'09) is a Senior Lecturer with the Department of Electronic and Computer Engineering at Brunel University, UK. From December 2004 to August 2010 he served as a Lecturer and a Senior Lecturer at King's College London, UK. From 2003 to 2004, he was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering of the University of Toronto, Canada. He received the Diploma and the Ph.D. degrees from the Electrical and Computer Engineering department of the University of Thessaloniki, Greece, in 1997 and 2002, respectively.

Dr. Boulgouris is an Associate Editor for the IEEE Transactions on Image Processing. He served as an Associate Editor for the IEEE Signal Processing Letters between 2007 and 2011. He was a co-editor of the book Biometrics: Theory, Methods, and Applications that was published by Wiley - IEEE Press in 2009. Dr. Boulgouris is a Senior Member of the IEEE.