

## Research Article

# Human Skeleton Model Based Dynamic Features for Walking Speed Invariant Gait Recognition

**Jure Kovač and Peter Peer**

*Faculty of Computer and Information Science, University of Ljubljana, Tržaška 25, 1000 Ljubljana, Slovenia*

Correspondence should be addressed to Jure Kovač; [jure.kovac@fri.uni-lj.si](mailto:jure.kovac@fri.uni-lj.si)

Received 1 October 2013; Revised 19 December 2013; Accepted 23 December 2013; Published 21 January 2014

Academic Editor: Yue Wu

Copyright © 2014 J. Kovač and P. Peer. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Humans are able to recognize small number of people they know well by the way they walk. This ability represents basic motivation for using human gait as the means for biometric identification. Such biometrics can be captured at public places from a distance without subject's collaboration, awareness, and even consent. Although current approaches give encouraging results, we are still far from effective use in real-life applications. In general, methods set various constraints to circumvent the influence of covariate factors like changes of walking speed, view, clothing, footwear, and object carrying, that have negative impact on recognition performance. In this paper we propose a skeleton model based gait recognition system focusing on modelling gait dynamics and eliminating the influence of subjects appearance on recognition. Furthermore, we tackle the problem of walking speed variation and propose space transformation and feature fusion that mitigates its influence on recognition performance. With the evaluation on OU-ISIR gait dataset, we demonstrate state of the art performance of proposed methods.

## 1. Introduction

Psychological studies showed that humans have small but significant ability to recognize people they know well by their gait. This ability has encouraged the research for using gait as the means of biometric identification. Early studies on Point Light Displays (PLD) [1], which enable isolated study of motion by removing all other contexts from observed subjects, confirmed this ability.

Commonly used biometrics based on fingerprints, face, iris, and so forth have two obvious deficiencies. They perform badly at low image resolutions and need active user participation. Gait on the other hand does not suffer from these deficiencies. It can be captured with ordinary equipment without individual's awareness or even consent. The main deficiencies of such biometrics are the unknown level of uniqueness and covariate factors that change gait characteristics. These can be external (changes of view, direction, or speed of movement, illumination conditions, weather, clothing, footwear, terrain, etc.) or internal (changes due to illness, injuries, ageing, pregnancy, etc.). Problems are also caused by uncertain measurements, occlusions, and the

use of noninvasive acquiring techniques (without sensors or markers). All these negatively influence the recognition performance in real-life environment, which is still too weak for efficient use in biometry.

Methods can be categorized into two main groups. Model based approaches [2–5] build the model of human body or its movement in 3D and acquire gait parameters from this model (e.g., step dimensions, cadence, human skeleton, body dimensions, locations and orientations of body parts, joint kinematics, etc.). The methods of this group mostly focus on gait dynamics and less on appearance of individuals, which makes them more resistant to problems like changes of view and scale but in general do not achieve as good results as methods that also consider appearance. Furthermore, such methods are computationally demanding and especially susceptible to problems like occlusions.

Model-free approaches [6–9] acquire gait parameters by performing measurements directly on 2D images, without adopting specific model of human body or motion. Feature correspondence in consecutive images is obtained by prediction of speed, shape, texture, and color. They mostly use geometric representations like silhouettes, optical flow,

joint trajectories, history of movement, and so forth. The methods do not rely only on gait dynamics, but also measure the individual during movement—with it they also take appearance of individual into consideration. Methods are therefore less sensitive to covariate factors that result in variations of gait dynamics (e.g., ageing, illness, and walking speed change) but more susceptible to factors that result in changes of appearance (e.g., clothing, obesity, hairstyle, etc.), changes of view, and direction of movement.

Although several gait recognition methods demonstrate impressive performance under controlled (in-lab) environment setups [2, 3, 5–7, 9, 10], the use of gait recognition in real-life application is still limited, mostly because of covariate factors that influence individual's gait and therefore make recognition task more difficult (e.g., view changes, walking speed changes, occlusions, etc.). Nevertheless, examples of real-life applications using gait analysis exist. Authors in [11, 12] demonstrate how monitoring gait motion parameters of residents in the senior housing facility can detect anomalies in resident's movements [11] and also discriminate and recognize facility residents and visitors [12]. Such monitoring can be used for fall risk assessment, detection of health problems, and monitoring of patients during rehabilitation. Although authors do not specifically handle covariate factors, they propose an efficient way for eliminating walking samples that do not conform to constraints posed by gait analysis methods.

However, often in other real-life scenarios only limited number of individual's walking samples are available (e.g., security cameras), which can also be short, contain only a few steps, and are influenced by several previously mentioned covariate factors. Under such circumstances these covariate factors must be dealt with in order to make walking samples useful for gait analysis. Our work focuses on variations of walking speed, since it represents one of the major covariate factors that affect gait recognition performance, is almost always present in real-life environment, and therefore requires special attention. Several approaches handling changing walking speed exist in the literature. The most outstanding are summarized below, but, as opposed to our work, none of them are model based and none of them uses solely gait dynamics for recognition task.

Authors in [8] researched the influence of walking speed changes to recognition performance based on cadence and step length and suggested the improvement by silhouette normalization. Authors proposed a stride normalization of double-support gait silhouettes based on a statistical relation between the walking speed and the stride. They used baseline algorithm [7] on only five silhouettes of gait cycle (two single-support images and three double-support images) for recognition and discarded the other still informative images.

Furthermore authors in [13] used geometrical transformations to apply walking speed normalization to averaged silhouette [6] and Probabilistic Spatiotemporal Model (PSTM) [10] and demonstrated how negative effects of walking speed changes can be mitigated to improve recognition performance.

Authors in [14] proposed a HMM-based time-normalized gait feature extraction with standard gait poses and tested it

on the slow and fast walking data. The method does, however, not consider spatial changes (e.g., stride changes).

Authors in [15] introduced a spatiotemporal Shape Variation-Based Frieze Pattern (SVB frieze pattern) representation for gait, which captures motion information over time and represents normalized frame difference over gait cycles. A temporal symmetry map of gait patterns is constructed and combined with vertical/horizontal SVB frieze patterns for measuring the dissimilarity between gait sequences.

Authors in [16] proposed an approach based on Dynamic Time Warping (DTW), which uses a set of DTW functions to represent the distribution of gait patterns using uniform and wrapped-Gaussian distributions.

Authors in [17] proposed a three-way ( $x$ -,  $y$ -, and time-axis) method of autocorrelation that effectively extracted spatio-temporal local geometric features to characterize motions called Cubic Higher-order Local Autocorrelation (CHLAC). It is relatively robust against variations in walking speed, since it only uses the sums of local features over a gait sequence and, thus, does not explicitly use the phase information of the gait. Researchers have assumed that walking speed does not change much within or across gait sequences.

Authors in [18] separated static and dynamic features from gait silhouettes by fitting a human model and then created a factorization based speed transformation model for the dynamic features using a training set for multiple persons on multiple speeds. The model can transform the dynamic features from a reference speed to another arbitrary speed.

Authors in [19] propose a new descriptor named Higher-order derivative Shape Configuration (HSC), which can generate robust feature when body shape changes due to varying walking speed. Procrustes shape analysis was used for gait signature and HSC is able to retain discriminative information in the gait signatures, while it is still able to tolerate the varying walking speed. They upgraded the method by introducing a Differential Composition Model (DCM) [20], which differentiates different effects caused by walking speed changes on various human body parts.

Human body skeleton model is proved to be an effective tool for representing human motion and was therefore adapted by several model based gait recognition approaches [3, 4]. Similar as authors in [3], we acquire gait signature by segmenting 2D human skeleton from silhouette images and then use this model to further extract motion parameters. On the other hand, authors in [4] first acquire individual's gait's characteristics by the principle of deformable template matching and then use view decomposing principle of general viewing angle and prior constraints from general knowledge of human body to impose gait characteristics to generic 3D skeleton model.

Authors in [3, 4] both assume that subjects are walking with constant (i.e., normal) speed. Moreover, authors in [3] do not handle any covariate factors at all, while authors in [4] achieve some degree of view invariance by reconstructing a 3D body model. Both works mainly use static gait parameters like subject's height, gait frequency, stride length, and even walking speed itself, which are all highly discriminative under unchanged, normal-walking conditions, but on the

other hand very susceptible to, for example, perspective deformations caused by view change and gait dynamics transformations caused by walking speed change. Although dynamic gait parameters are also used, they are not fully utilized for the recognition task. Authors in [3] use only means and standard deviations of main joint angles, while authors in [4] only use motion trajectories of main joints. Both are similarly susceptible to covariate factors as static features.

As opposed to both of these works our focus is on dynamic motion parameters only. We observe time series of changing joint angles, angle phase differences, mass ratios of different body parts, distances of body parts from body center, and so forth, through the entire gait sequence. These are less sensitive to covariate factors that greatly affect the appearance of individual's gait.

We argue that although dynamic features are claimed to be less discriminative than static features, they still contain enough discriminator power to achieve comparable recognition performance to appearance based methods, and are at the same time more resistant and easily adoptable to handle covariate factors that exist in real-life scenarios. Our main contributions are as follows:

- (i) feature space transformation based on statistical model of different walking speeds to compensate for walking speed changes in dynamic gait features;
- (ii) feature fusion scheme that enables the use of per feature classifiers as a weak classifiers that are fused into the final distance based classifier;
- (iii) image moments based cycle detection stage that enables almost perfect gait phase alignment among different video sequences, which is crucial for distance based classification on time series signal based features.

The remainder of this paper is organized as follows. In Section 2 we will introduce the procedure for extracting human skeletons from silhouette images. Section 3 will discuss the details of acquiring gait signals and motion parameters from skeleton models and introduce the proposed feature space transformation, which helps mitigate the effects of changed walking speed on recognition performance. In Section 4 we will describe classification procedure based on proposed feature fusion scheme and in Section 5 we will present the results and analysis of performed experiments together with comparison to related walking speed invariant state-of-the-art methods. Finally, we will provide some directions for future work in the conclusion.

## 2. Skeleton Segmentation

Input video must first be processed to acquire silhouettes of walking subjects. We use Gaussian mixture based background subtraction to acquire motion foreground, which corresponds to subjects's silhouettes that are further processed by morphology operations to improve their quality. Since

most of the gait datasets already provide extracted silhouettes, we use standard procedures for silhouette extraction when required and do not pay special attention to motion segmentation procedure as this is another active research area (see [21] for recent advancements). Therefore, we assume that silhouettes of decent quality are either provided or acquired as the input to our method. It means that extracted silhouettes must conform to the constraints posed by relatively complex skeleton segmentation procedure. For example, walking subjects, which are far away from the camera, might appear too small in the video to effectively distinguish human body parts (limbs, head, etc.). Despite the silhouette quality constraint posed in this work, we still provide methods for handling low quality and noisy silhouettes in Section 4.5.

We adopt the skeleton extraction procedure from [3], since their method uses simple segmentation steps for deriving skeletons based on single view 2D motion sequences, which turned out to supply sufficient gait information for gait based identification. Our segmentation procedure contains the following steps.

First, we extract silhouette contour from every image of the sequence. Then calculate body height and segment lengths based on average anatomical properties of human body [22]. For a body height  $H$ , the initial estimate of the vertical position of the neck, shoulder, waist, pelvis, knee, and ankle was set to be  $0.870H$ ,  $0.818H$ ,  $0.530H$ ,  $0.480H$ ,  $0.285H$ , and  $0.039H$ , respectively.

Next, we divide silhouette into individual body parts (head, neck, torso, waist, upper legs, bottom left, right leg, and left and right feet) based on calculated segment lengths (Figure 1(a)). For each body segment we then find left most and right most contour points and calculate their middle (Figure 1(b)). After that, we fit lines through this points using Hough transformation and chose the line with largest value in accumulation field. Such fitting is more robust to errors as opposed to linear regression used in [3] (Figure 1(c)).

While upper body part segmentation is straightforward, the lower body part (thighs and shins) is much harder to extract and must therefore be specially handled. First we detect shin bones. Especially during single support phase, only one segment is detected in lower body part. Therefore, both legs are represented by single body segment. In this case, left and right most contour points are used as left and right bone angles. In the case of double support, each shin is represented by its own body segment and middle points of both segments are used, similar to those in upper body part. The thigh bone angles are then derived from torso lower midpoint and starting points of both left and right shins. Finally, we compute bone endpoints from bone starting points, derived angles and segment length estimations (Figure 1(d)).

Extracted bones form a simplified (no arms, shoulders, etc.) skeleton of human body for single silhouette frame. Raw gait signature is then acquired by detecting skeletons on all frames of the input video. Such signature enables direct extraction of bone and joint angles, which form time-series signals that can be used for recognition.

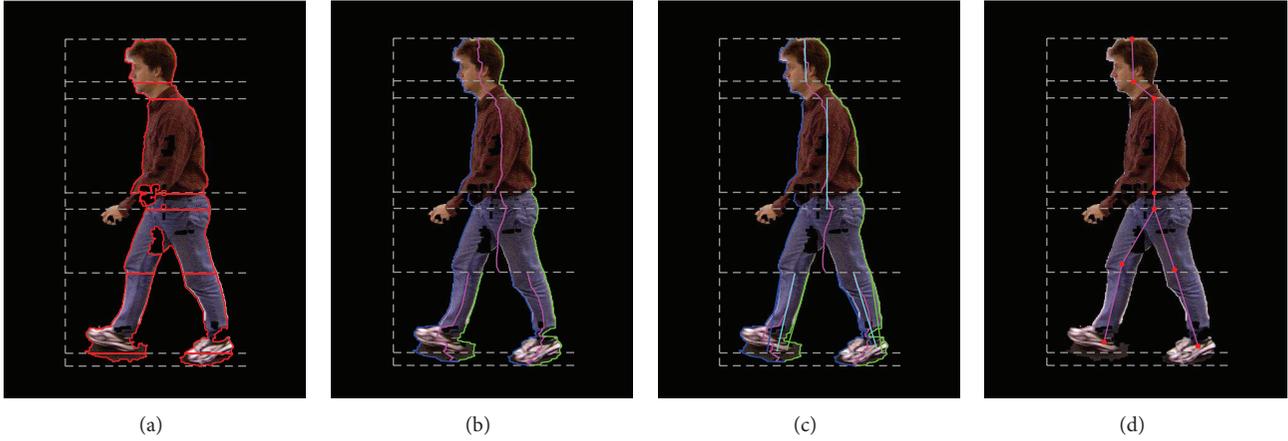


FIGURE 1: Skeleton segmentation algorithm in steps.

### 3. Gait Signal Extraction

Several gait signals are extracted from acquired skeleton signatures:

- (i) body part angles: head, neck, torso, left and right thighs, and left and right shins;
- (ii) joint locations: left and right knees and ankles;
- (iii) masses and centres of mass (COM) for the whole silhouette and individual body parts;
- (iv) silhouette width and height.

All these so-called raw signals represent the basis for the derived signals, which are actually used in recognition system later. By definition, joint angles are measured as one joint relative to the other, so relative joint angles are derived from extracted raw angle values. At normal walking, torso is almost vertical, so relative hip angle (thigh bone angle) is the same as extracted value, while relative knee angle is calculated as

$$\theta_{\text{knee}} = \theta_{\text{thigh}} - \theta_{\text{shin}}. \quad (1)$$

Phase difference of thighs and shins describe the correlation among left and right thighs or left and right shins and can be very specific for walking subjects:

$$\theta_{\text{tphs}} = \theta_{\text{thigh}_1} - \theta_{\text{thigh}_2}. \quad (2)$$

Body part mass ratio represents the ratio between masses of specific body parts and the mass of entire body:

$$Q_{m_i} = \frac{m_i}{m}. \quad (3)$$

Body part center of mass distance represents the distances of center of mass of specific body parts to the center of mass for the entire body:

$$D_{ci} = \sqrt{(c_{i,x} - c_x)^2 + (c_{i,y} - c_y)^2}. \quad (4)$$

The total of 64 feature signals are extracted and later used in classification.

**3.1. Image Moments Based Motion Parameters.** Image moments for binary images and its derivatives, especially COM, turned out very useful for gait motion parameters estimation. Their use for estimating gait frequency, gait cycle starts, and gait phases (double, single support) is described below.

Image moments are defined as

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y), \quad (5)$$

where  $ij$  is the order of the moment and  $I(x, y)$  is the image value at point  $(x, y)$ . Since binary silhouette images contain values either 0 or 1, its COM can be computed as

$$\mathbf{C} = \left[ \frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right], \quad (6)$$

where  $M_{10}$  and  $M_{01}$  are the sums of all  $x$  and  $y$  coordinates of all nonzero image elements, respectively, and  $M_{00}$  is the number of all non-zero image elements also referred to as mass. Both COM  $\mathbf{C}$  and mass  $M_{00}$  are also useful as gait features.

**3.1.1. Gait Cycle Detection.** Walking sequences are not synchronized among subjects and can start at any phase of the gait cycle. To be able to compare the cycles later, we must define at which phase gait cycle should start. Double support is the easiest to detect, therefore, we set all cycles to start at double support.

First, we determine gait frequency based on frequency analysis of extracted signals. The most appropriate signals for gait frequency detection are vertical movement of COM, body mass and silhouette width. The main property of these signals is that the signal has its peaks at double support (max) and single support (min) phases of gait cycle. When subject is observed from side view, COM is at its lowest point (highest  $y$  value), and subject's mass is at its highest point, since hands and legs are spread and occupy more space than at single support, and the same goes for subject's width. Among these subject's width is the most sensitive to noisy silhouettes and

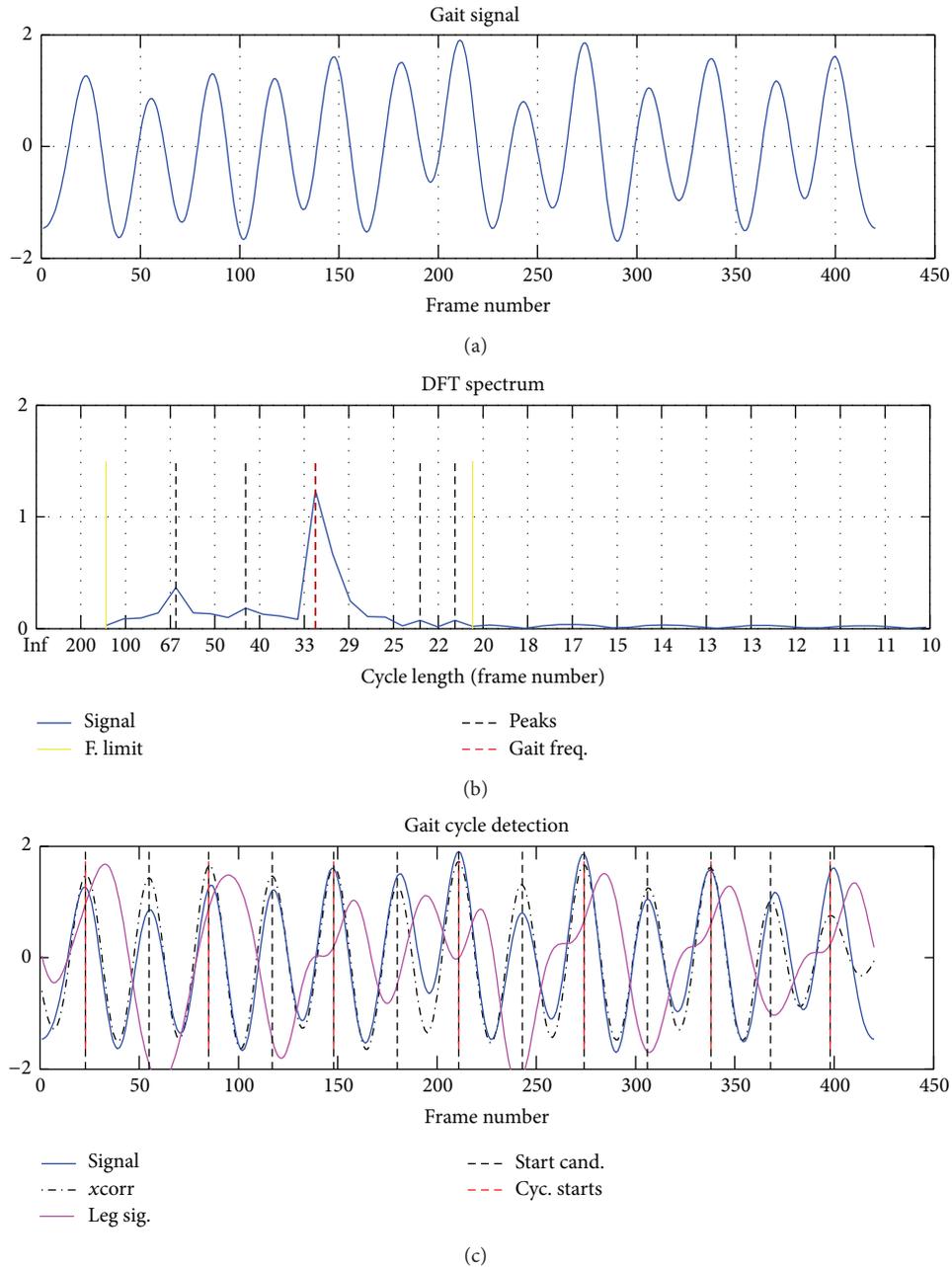


FIGURE 2: Gait frequency analysis: (a) shows COMy signal with peaks at double support phases (half gait cycle); (b) shows DFT spectrum with one outstanding frequency indicating half gait cycle length; (c) shows gait cycle starts detection, where left leg double support is detected by leg signal value at main signal autocorrelation peaks.

also the least reliable, since it is severely affected by hand movements and is therefore the least appropriate. COM is the most resistant to noisy silhouettes, but a bit less stable in finding double support peaks than subject’s mass. Therefore, the best way is to use mass when dealing with high quality silhouettes and COM when silhouettes are noisy.

We analysed the signal with Discrete Fourier Transform (DFT) to determine frequency spectra. Two frequencies with highest amplitudes stand out (Figure 2).  $F_s$  represents one half of the gait cycle, which usually has the highest

amplitude since it corresponds to subject’s steps, which are most common in the signal.  $F_g$  has slightly lower amplitude and represents the gait frequency (two steps):  $F_g \approx 2F_s$ .

**3.1.2. Left Leg Double Support Detection.** To be able to compare these signals later, it is essential to detect the same double support for every subject (e.g., the one with left leg in front), so that cycles of all the subjects start with the same leg in front (e.g., left) and are therefore roughly aligned. Merely

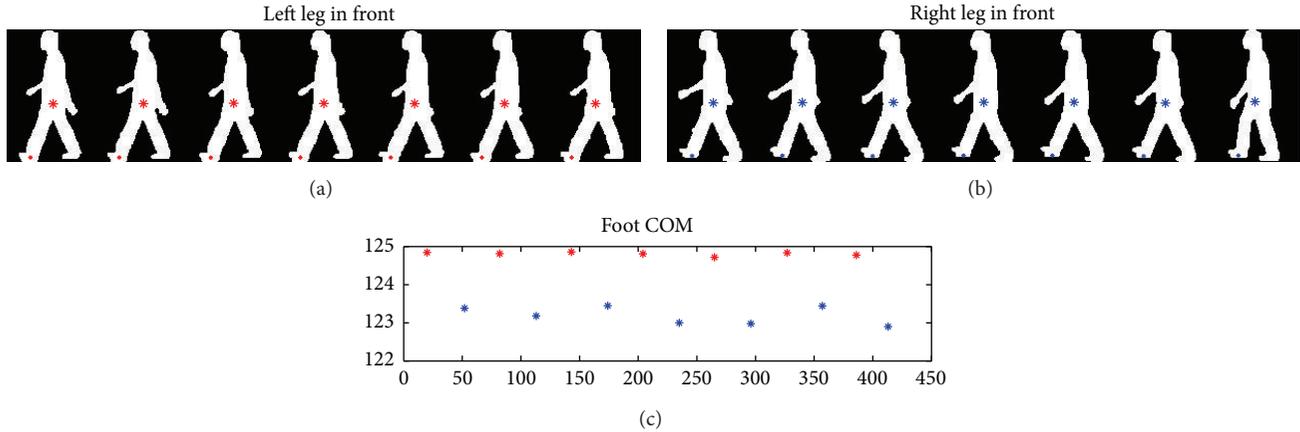


FIGURE 3: Double support detection: (a) shows double support when left leg is in front, (b) shows double support when right leg is in front. Red and blue markers indicate a location of body and leg COMs. (c) demonstrates a clear gap in COM $y$  value between left and right leg double support.  $x$  axis denotes time in frames, and  $y$  axis denotes  $y$  coordinate of COM.

switching the legs at the later stage does not help here, because all other features are also affected.

We solve this by exploiting constraint posed by perspective projection. It can be observed that because of perspective projection when looking at walking subject from side view, the leg closer to the camera appears a bit lower in the image (Figure 3). That is when left leg is in front, it appears lower than when the right leg is in front. We use this property when analysing another signal. The  $y$  coordinate of front leg foot COM is observed at double support phases. When the leg closer to the camera (e.g., left) is in front, the value of COM $y$  of the foot is higher than foot COM $y$  value of the leg that is farther away from the camera.

Finally, we derive cycle starts based on estimated gait frequency and detected left leg double support phases, which are used as initial estimates for cycle length and start of the first cycle. So obtained signal part is then cross-correlated against entire signal and its peaks represent starts of gait cycles (Figure 2(c)).

**3.1.3. Leg Crossover Detection.** One of the problems with the extracted signature represents the crossing of the legs. When observing subject from side view, which is most commonly the case, the legs are crossed during single support phase. Previously described algorithm for skeleton extraction cannot detect the crossing, instead it always tracks the front leg as leg number 1 and rear leg as leg number 2. When legs get crossed, the leg that was previously in the front (e.g., left leg) now goes to the back and vice versa.

Authors in [3] use physical constraints (e.g., foot does not move forward when in contact with the floor) to detect the crossover. Such constraint gets broken, for example, when walking on a treadmill. Therefore, we use vertical movement of COM, which turns out to be more appropriate for the given task.

It can be observed that single support phase location corresponds to local minima in COM $y$  or mass signal. This is similar as double support detection by local maxima in

the same signal described earlier. Gait signature is then adopted by softly switching left and right thigh angle signals at detected leg switches locations. Softly we denote interpolation through small time frame window during single support. Shin angles and joint locations are adopted according to changed thigh angle values afterwards (see Figure 4).

**3.2. Signal Postprocessing.** Signals in gait signature are prone to erratic detections, which manifest as spikes in gait signals. The spikes are detected by applying moving median filter through entire signal. Moving median enables local detection of spikes.

First, we compute the difference  $d(t)$  between signal  $s(t)$  and moving median  $s_m(t)$ . Then spikes are detected according to

$$d(t) = \|s(t) - s_m(t)\|,$$

$$s(t) = \begin{cases} \text{NaN} & d(t) > \alpha * 1.4826 * s_{m_\sigma}(t), \\ s(t), & \text{else,} \end{cases} \quad (7)$$

where  $s(t)$  is the value of the signal at time  $t$ ,  $\alpha$  is a parameter determining the size of the spike,  $s_{m_\sigma}$  is median deviation, which is similar to standard deviation, and except deviation is calculated based on median mean instead of average mean. Scaling median mean by factor 1.4826 gives median-absolute mean, which is on average equal to standard deviation for Gaussian distributions. When spike is detected, signal values are set to undefined NaN.

For example, in the case of bone angle signals, we detect spikes for both thighs and both shins and then interpolate the missing values with spline interpolation. After that, bone structures (joints and angles) must be adopted accordingly (e.g., changing angle of left thigh also affects left shin position). Moreover, all the signals are additionally filtered by trigonometric polynomials used when building gait descriptor (see Section 4.2).

As human gait is cyclic in nature, all gait signals are then sliced to smaller pieces based on found cycle starts. This

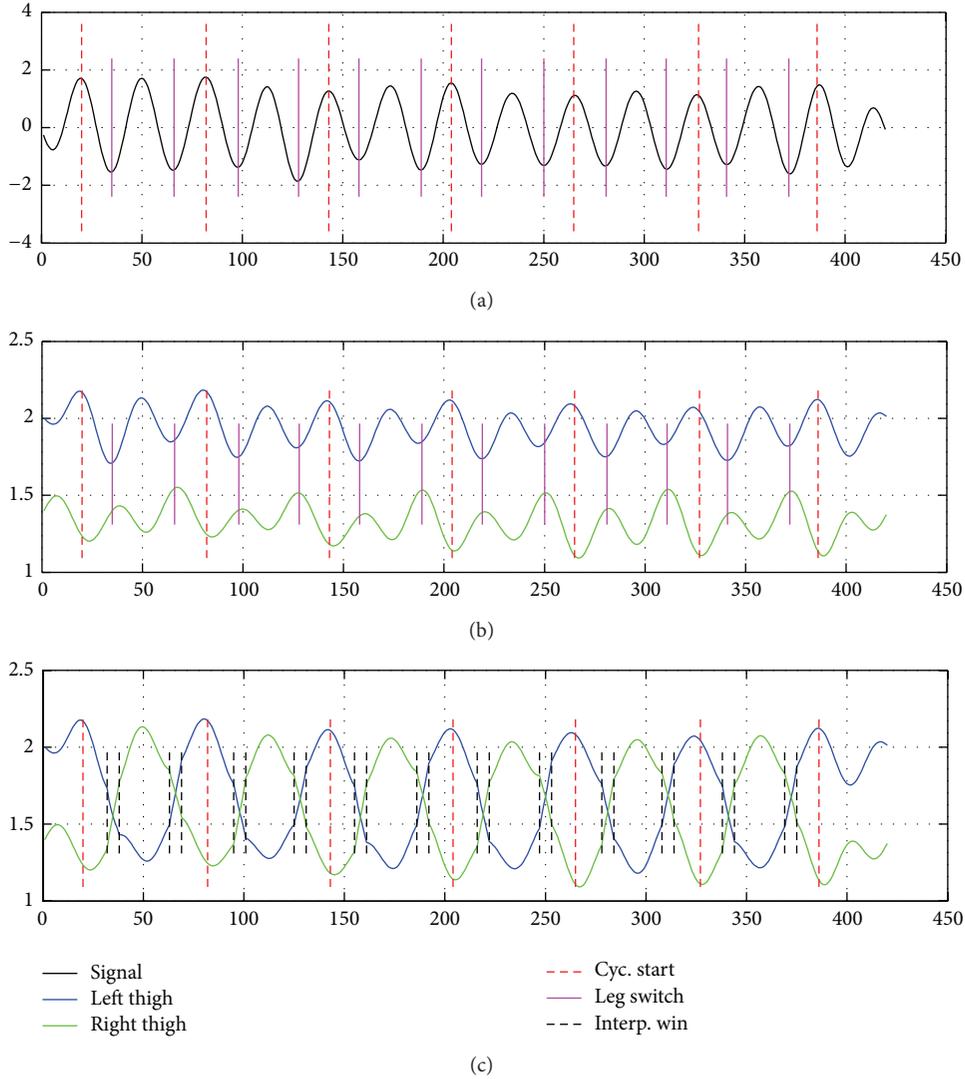


FIGURE 4: Leg crossover detection: (a) shows COM y signal (full-black line), detected cycle starts (dashed-red vertical lines), and leg crossover locations (full-magenta vertical lines). (b) shows front (blue) and rear (green) thigh angle signals as extracted from skeleton. (c) shows thigh angle signals interpolated into left (blue) and right (green) thigh angles.  $x$  axis denotes time in frames and  $y$  axis represents normalized signal values.

also implies that gait cycles of the same subject are similar under unchanged conditions. For acquired signals this also means that the next cycle should start with approximately the same values as previous cycle and also that cyclic signal should end with approximately the same values as it has started [10]. Because of noisy segmentation this is not always the case. Therefore, signals are adopted to correspond to this assumption. We make the signals cyclic by finding the difference between cycle start and cycle end and interpolating the entire signal to negate this difference:

$$S_i = S_i - \frac{2i - (L - 1)}{2(L - 1)} * D, \quad (8)$$

where  $S$  is the part of the signal corresponding to some specific gait cycle,  $S_i$  is the value at point  $i$  of that gait cycle,  $L$

is the gait cycle length, and  $D$  is the difference between cycle start and cycle end.

Also, different people walk with different speeds, which affect gait frequency and with that the duration of the gait cycle. If gait is sampled by the same sampling frequency (e.g., camera frame rate), cycle lengths differ in the number of acquired samples in one gait cycle. To be able to compare gait cycles of different subjects (and also gait cycles of the same subject acquired with different sampling frequencies), gait cycle signals must be resampled to the same length by interpolation and decimation.

Each gait cycle now contains only parts of the original signals corresponding to this gait cycle. These signal parts are called features (see examples in Figure 5). Each feature corresponds to one gait cycle of some extracted signal. Figure also demonstrates the difference in inter- and intraclass

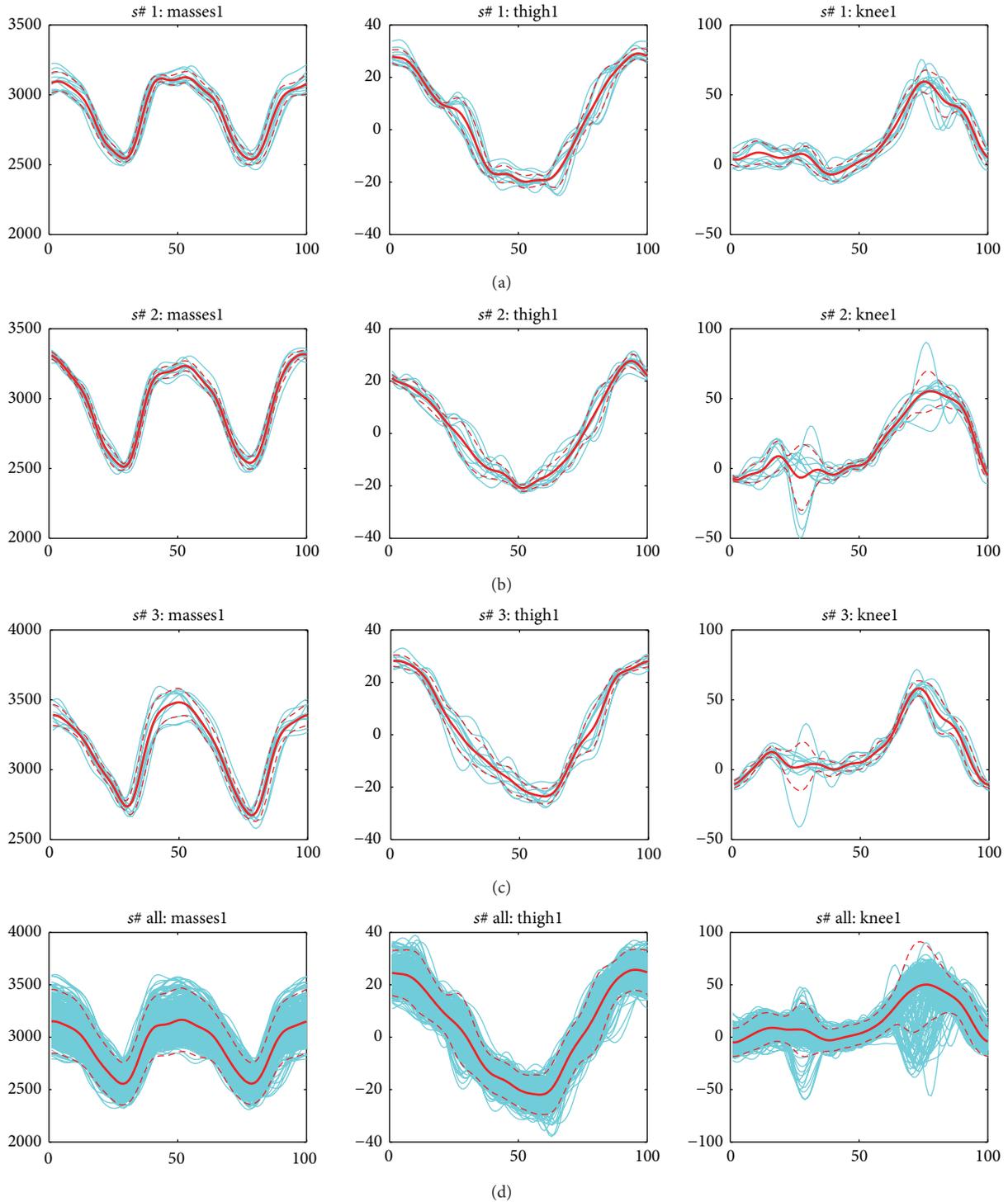


FIGURE 5: Example of extracted features: image shows extracted mass, thigh1, and knee1 signals (in columns) of several gait cycles for three subjects (rows (a), (b), and (c)) denoted by  $s\# 1$ ,  $s\# 2$  and  $s\# 3$ . Blue lines indicate signals, red-thick lines indicate signals mean, and red-dashed lines indicate intraclass variance. Row (d) shows mean and variance (interclass) of all the signals for all subjects ( $s\#$  all).

variance. It is evident that the latter is smaller, which is good for recognition purposes.

**3.3. Feature Space Transformation.** Changes in walking speed present one of the major covariate factors with negative impact on recognition results. When a person changes

walking speed, dynamic features (e.g., stride length and joint angles) are changed, while static features (e.g., thigh and shin lengths) remain unchanged. From top row of Figure 6, it can be seen, how single subjects's features differ for different walking speeds (red—fast walk, blue—slow walk). As we only deal with dynamic features, we have to compensate for such changes.

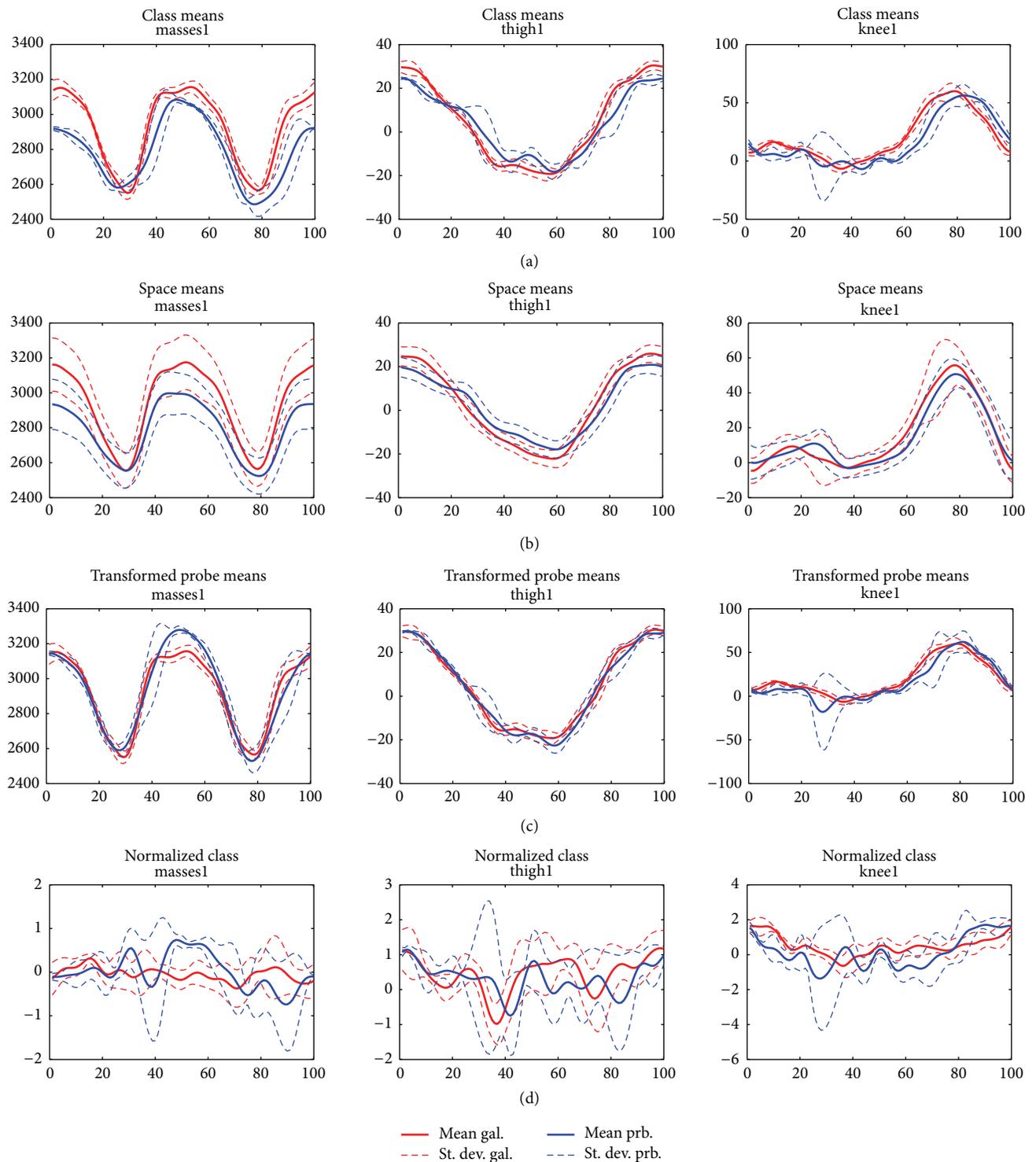


FIGURE 6: Feature space transformation: (a) single subject's features for fast walk (red) and slow walk (blue). (b) Slow and fast walk group mean and st. dev. (c) Slow walk features transformed to fast walk space. (d) Remaining differences after transformation. Note that y axis in rows (a), (b), and (c) shows absolute (measured) values, while row (d) shows differences after normalization, which therefore appear relatively small compared to measured values. x axis denotes the percent of gait cycle.

We perform a feature space transformation based on statistical model of feature space for gallery and probe groups, which contain subjects walking with different speeds. First, we calculate the main statistical measures for each group belonging to certain walking speed, based only on a few randomly chosen subjects from that group. Then we derive group mean and standard deviation (Figure 6 2nd row) and transform the feature signals of all the samples (all subjects, all gait cycles) in that group by

$$\vec{s} = \frac{\vec{s} - \vec{\mu}_g}{\vec{\sigma}_g}, \quad (9)$$

where  $\vec{s}$  is the signal vector of single feature of single gait cycle sample,  $\vec{\mu}_g$  is the mean vector of the same feature for target group, and  $\vec{\sigma}_g$  is the vector of standard deviations for the same group. All operations are element wise.

Such transformation removes major group space deviations caused by varying walking speed and retains only interclass differences based on which subjects can be discriminated (Figure 6, 4th row). The transformation is performed on both gallery and probe groups and subjects within these groups are now comparable. For a demonstration, probe features of slow walk are transformed by inverse transformation to fast walking feature space of the gallery (Figure 6, 3rd row). It is evident that the signals are closer together after transformation.

## 4. Classification

**4.1. Phase Alignment.** Classification is done by comparing the probe subjects with the subject stored in the gallery. The comparison is based on Euclidean distance. To make this comparison feasible, gait and probe features must be perfectly aligned. This is roughly achieved in cycle detection phase; however, cycle starts can still differ by a few frames, therefore, more precise alignment is required before classification.

We perform phase alignment by finding the offset of two subject's gait cycles based on a few selected features (COMy, mass, and width). We shift one subject's gait cycles in the range of few frames and calculate the distances of selected features for these shifts. The one with minimal distance is chosen and probe cycles are then shifted accordingly:

$$o = \arg \min_{i=-10}^{i=10} d(\vec{F}_1, \vec{F}_2), \quad (10)$$

$$\vec{F}_1 = [\vec{f}_{c_y,1}, \vec{f}_{m,1}, \vec{f}_{w,1}],$$

$$\vec{F}_2 = [\vec{f}_{c_y,2}, \vec{f}_{m,2}, \vec{f}_{w,2}],$$

where  $o$  is the offset,  $d$  is distance, and  $\vec{F}_j$  is the feature vector of subject  $j$  composed of feature vectors  $\vec{f}_{i,j}$  of specific features.  $c_y$ ,  $m$ , and  $w$  denote COMy, mass, and width. It is important to limit the shifting to small ranges (e.g.,  $[-10 \dots 10]$ ), since larger ranges can result in shifting the gait cycle to the next double support (with the wrong leg in front).

Prior to classification all subjects in the gallery are aligned to one chosen subject (e.g., first) and also all probes are aligned to the same subject.

**4.2. Gait Descriptor.** Only after such alignment we can compose gait descriptors for all cycles of all subjects. First, gait cycle signals of all acquired features are approximated by trigonometric polynomial coefficients. Periodic signals can be approximated by:

$$S(x) = a_0 + \sum_{n=1}^N a_n \cos(nx) + i \sum_{n=1}^N b_n \sin(nx), \quad (11)$$

where  $a_n$  and  $b_n$  are the polynomial coefficient acquired by DFT. We use these coefficient to form the descriptor for each feature. Such approximation serves as a low-pass filter, where higher frequencies are filtered out and, in the same time, greatly reduces feature size, so it also reduces the dimensionality of features. If acquired feature size is 100 elements, we can safely use  $N \approx 10$ , which gives approximately 20 coefficients. Gait descriptors of individual features

$$\vec{d}_i = [a_{0_i}, \dots, a_{N_i}, b_{1_i}, \dots, b_{N_i}], \quad (12)$$

are then stacked together to form a descriptor of the entire gait cycle:

$$\vec{D}_{c_i} = [\vec{d}_{i_1}, \vec{d}_{i_2}, \dots, \vec{d}_{i_{N_f}}], \quad (13)$$

where  $\vec{D}_{c_i}$  is subject's descriptor of  $c_i$ th gait cycle and  $\vec{d}_{i_1} \dots \vec{d}_{i_{N_f}}$  are feature descriptors of  $N_f$  different features.

Acquired descriptors are still too long for further processing. For example, in our case of 64 features each containing 20 coefficients, descriptor length is 1280. Dimensionality is further reduced by PCA to retain approximately 99% of variance, which ends up in approximately 100 PCA coefficients.

Furthermore, all gallery descriptors are processed by LDA to achieve better discrimination between gallery models.

**4.3. Distance Based Classification.** Now we calculate the distances between each probe descriptor (gait cycle) and all descriptors of all the models in the gallery. Distance between two subjects is defined as

$$D(p, g_m) = f_c(d(\vec{p}_i, \vec{g}_{m_j})), \quad i \in P, j \in G_m, \quad (14)$$

where  $p$  and  $g_m$  denote probe and gallery subjects,  $P$  and  $G_m$  number of descriptors (corresponding to gait cycles) for each subject, and  $d(\cdot)$  the distance between two descriptors.  $f_c$  is a criteria function, denoting which distance to choose, among distances formed by several probe descriptors. Most commonly mean, minimum and median are used. We choose median as it best reduces the importance of deviated distances (outliers) that might result from measurement errors. We then calculate  $D(p, g_m)$  for all gallery models to obtain a vector of distance scores:

$$\vec{D}_p = \{D(p, g_1), \dots, D(p, g_M)\}, \quad (15)$$

where  $M$  is the number of subjects in the gallery. Since we use Cumulative Match Characteristics (CMC) to represent classification results, the vector is sorted and ranks are assigned to the models based on calculated distances. The model with the smallest distance is chosen as best (rank 1) match.

**4.4. Feature Fusion Scheme.** Different features have different discriminative abilities. Therefore, it is natural to assume that using these features separately and assigning their influence by weighting could improve classification results. To demonstrate this point, we employ feature fusion scheme [23], where each feature can be seen as a separate classifier and final result is formed by fusing the results of all the classifiers. Such process can additionally help in the case of missing or erratically measured features. For example, in the case of occlusions only a few features (like missing parts of the legs, etc.) are affected and negative effects can be diminished by classification based on the remaining features, which over-vote the problematic ones.

First, we perform distance based classification as described earlier in Section 4.3 for each feature separately. We compose gait descriptors based on single feature and perform classification. PCA is not required in this case since trigonometric coefficients sufficiently diminish the dimensionality to make LDA feasible. Each feature forms a vector of distance scores  $\vec{D}_{p,f}$ , which is first normalized to interval  $[0, 1]$  to remove feature distance bias using min-max norm. Then we combine all the score vectors of all the features to a single fused distance score:

$$\vec{S}_p = \frac{1}{F} \sum_{i=1}^{i=F} (w_i \|\vec{D}_{p,i}\|_{\min-\max}), \quad (16)$$

where  $F$  is the number of features and  $w_i$  are feature weights. Please note that  $\|\cdot\|_{\min-\max}$  denotes vector normalization, which results in normalized vector, rather than vector norm, which results in a scalar value. Finally, the so-obtained distance score vector  $\vec{S}_p$  is sorted and ranks are assigned for all gallery models.

In this work we obtain weights experimentally by evaluating the discriminative power of the features on the basis of intra-/interclass feature variance. Class corresponds to a subject. The intraclass feature variance  $\sigma_{s,f}^2(X)$  is the variance of all the signal samples (rows of  $X$ ) of specific feature  $f$  for single subject  $s$ . Sample  $x_i \in X$  is the signal of the feature  $f$  during single gait cycle. Interclass feature variance  $\sigma_f^2$  is the variance measured on all the samples of all subjects for the same feature (see examples on Figure 5). The ratio of average (through subjects) intra class feature variance and inter class variance,

$$r_\sigma = \frac{\overline{\sigma_{s,f}^2}}{\sigma_f^2} \quad (17)$$

is an indicator of feature discrimination power: the smaller the ratio is, the bigger the discrimination power of the feature can be expected. For our experiments, the variances were calculated on training subjects subset.

Undoubtedly, the process of assigning weights could be more sophisticated using some training technique. However, classification performance turned out to be relatively insensitive to small variations in weights, therefore chosen weight assignment is sufficient for demonstrating the effect of feature classifier fusion on classification performance.

**4.5. Noise, Occlusions, and Outliers.** Due to sensitive nature of skeleton segmentation process, several types of errors caused by noise sensitive video capturing process, poor quality silhouette extraction algorithm, or scene objects occluding recorded subject may occur. We deal with these errors on several different layers of our method.

Most common are small erratic measurements of bone and joint positions, which are due to noisy silhouettes, for example, random artifacts appearing on particular frames of the video, holes between legs get filled in knee or ankle area, and so forth. This type of errors only occurs on a few consecutive frames and then disappears and might later reappear in some other areas. The result of such errors is noisy signals obtained from segmented skeleton. The noise can be effectively handled by spike detection, elimination, and interpolation as described in Section 3.2 and additional filtering by trigonometric polynomial coefficients as described in Section 4.2.

Larger disruptions can be caused by occluding objects in the scene, missing larger parts of silhouettes, body parts, and so forth. When these errors are short-lived, most of their effects are already eliminated by noise handling procedures. On the other hand, longer presence of occluding objects might result in erratic measurements of several features for entire or even several gait cycles. In the case of missing or badly measured features, bad effects are eliminated by feature fusion (Section 4.4), which acts as classifier fusion where single classifier (feature) adds only minor contribution to final score, while most of the score is formed by the rest of the unaffected features. In the cases when entire gait cycles are affected, such cycles appear as outliers in probe/gallery distance calculation process (Section 4.3) and the effects of problematic distances are eliminated by calculating the median of the distances of all available gait cycles.

## 5. Experimental Results and Discussion

We tested our methods on OU-ISIR gait database [24], more precisely dataset A, which is composed of high quality gait silhouette sequences of 34 subjects from side view, walking on a treadmill with speed variation from 2 km/h to 10 km/h at 1 km/h interval. For each subject and each speed two sequences are provided, one for the gallery and the other for the probe.

In our cross-speed test, we performed several experiments, where gallery subjects belong to one speed and matching probe subjects belong to the other speed. The results are given as rank 1 correct classification rate in Cumulative Match Characteristic (CMC) to enable comparison with other works, as CMC is also mostly used by other authors. Such metric is highly dependant on the size of the gallery and in order to enable fair comparison to other works, the number of subjects must be the same. 25 subject were used for identification and 9 were used for training previously described feature space transformation. Such split was designed in [18] and adopted by other researchers working on speed-invariant gait recognition as a method benchmark.

TABLE 1: Recognition performance (%) results of basic experiments without feature fusion and space transformation for walking speeds.  $G$  and  $P$  denote the speed group of gallery and probe, respectively.

$G/P$	2 km/h	3 km/h	4 km/h	5 km/h	6 km/h	7 km/h
2 km/h	84	80	44	24	16	16
3 km/h	36	80	72	28	28	12
4 km/h	44	68	92	84	52	28
5 km/h	40	36	68	100	88	64
6 km/h	16	24	28	68	92	92
7 km/h	28	20	28	56	76	96

TABLE 2: Recognition performance (%) results of experiments with feature fusion and without space transformation for walking speeds.

$G/P$	2 km/h	3 km/h	4 km/h	5 km/h	6 km/h	7 km/h
2 km/h	88	88	96	72	56	52
3 km/h	68	92	96	84	76	52
4 km/h	72	80	96	92	88	72
5 km/h	56	72	96	100	100	96
6 km/h	56	64	76	96	100	100
7 km/h	36	56	52	80	92	100

TABLE 3: Recognition performance (%) results of experiments without space transformation for running speeds.

$G/P$	No fusion			Feature fusion		
	8 km/h	9 km/h	10 km/h	8 km/h	9 km/h	10 km/h
8 km/h	96	92	92	96	96	96
9 km/h	92	100	100	100	100	100
10 km/h	76	92	96	92	96	96

The experiments were split to four sets to illustrate the impact of feature fusion and space transformation on recognition results. Additionally, we split the speeds to walking (2 km/h to 7 km/h) and running (8 km/h to 10 km/h), because running is so much different to walking in terms of gait dynamics that it is regarded as different action (in analogy to action recognition field) and the cross-matching is quite difficult.

First, we perform basic method evaluation, without feature fusion and without space transformation. The results in Table 1 are quite modest and rarely reach 90% even when gallery and probes are in the same speed group. It can be noticed that higher speeds yield better results. We assume that this is because higher speeds reveal more specifics (e.g., consider running), intraclass variation is smaller, and also dynamic features are easier (less noisy) for capturing at higher speeds, when subject's silhouettes are more spread at double support phases, more information is revealed.

Next, we perform experiments with previously described feature fusion scheme without space transformation. The results in Table 2 show noticeable improvement. For easier comparison see also test result summary in Figure 7.

Results for higher speeds (i.e., running speeds) with and without feature fusion are given separately in Table 3. These also confirm our high speed observation.

TABLE 4: Recognition performance (%) results of experiments without feature fusion and with space transformation for walking speeds.

$G/P$	2 km/h	3 km/h	4 km/h	5 km/h	6 km/h	7 km/h
2 km/h	92	84	84	72	44	32
3 km/h	68	88	84	72	60	48
4 km/h	72	92	96	88	88	64
5 km/h	64	76	100	100	96	92
6 km/h	40	60	72	88	100	96
7 km/h	36	36	48	76	96	92

TABLE 5: Recognition performance (%) results of experiments with feature fusion and with space transformation for walking speeds.

$G/P$	2 km/h	3 km/h	4 km/h	5 km/h	6 km/h	7 km/h
2 km/h	92	96	92	92	60	68
3 km/h	80	96	100	84	76	64
4 km/h	80	92	100	96	96	72
5 km/h	80	92	100	100	96	100
6 km/h	76	92	100	100	100	100
7 km/h	72	60	80	96	100	100

TABLE 6: Recognition performance (%) results of experiments with space transformation for running speeds.

$G/P$	Space transformation			Space transformation + fusion		
	8 km/h	9 km/h	10 km/h	8 km/h	9 km/h	10 km/h
8 km/h	96	96	96	100	100	100
9 km/h	88	100	100	100	100	100
10 km/h	92	96	96	100	100	100

The results of the next experiments where we use feature space transformation without and with feature fusion are given in Tables 4 and 5. The space transformation again improves recognition performance in both cases (with and without fusion). In the case, where gallery and probes are taken from the same speed (0 km/h speed change) as shown in result summaries in Table 7 and Figure 7, the differences are not so drastic. Nevertheless, space transformation reveals small improvement and together with feature fusion reaches almost 100%. This improvement is due to the effect achieved by removing the means of the whole speed group from specific subject's features, and since only inter class differences remain in the signals, only those are modelled in descriptors.

Finally, the results of space transformation for running are given in Table 6, and in the case of 18 both improvements reach 100% in all the experiments.

Table 7 and Figure 7 show average recognition performance for cross-speed walking and running experiments. Basic results are given in column 1, feature fusion results are given in column 2, space transformation results are given in column 3, and results with both improvements are given in column 4. Also, the average performance for different degree of speed changes (from 0 km/h to 5 km/h) and the performance gains for feature fusion and spatial transformation are given. Feature gain is calculated as a

TABLE 7: Average recognition performance (%) of cross-speed walking and running.

	Basic	Feature	ST	ST + feature	Feature gain	ST gain	ST + feature gain
Walk	53.0	79.1	74.9	88.3	26.1	9.2	35.3
Run	92.9	96.9	95.6	100.0	4.0	3.1	7.1
$\Delta = 0$ km/h	92.9	96.4	95.6	98.7	3.6	2.2	5.8
$\Delta = 1$ km/h	73.2	90.8	89.2	98.0	17.6	7.2	24.8
$\Delta = 2$ km/h	44.0	83.0	79.0	95.0	39.0	12.0	51.0
$\Delta = 3$ km/h	28.7	65.3	61.3	82.0	36.7	16.7	53.4
$\Delta = 4$ km/h	16.0	55.0	42.0	62.0	39.0	7.0	46.0
$\Delta = 5$ km/h	22.0	44.0	34.0	44.0	22.0	0.0	22.0

TABLE 8: Recognition performance (%) comparison to other works handling cross-speed walking gait identification.

Test scenario	Our method	SN [8]	HMM [14]	SVT [18]	HSC [19]	DCM [20]
Small speed change	96	/	84	90	96	98
Large speed change	68	34	/	58	68	82
Whole dataset	88.33	/	/	/	85.67	92.44

difference between feature fusion results (2nd column) and basic results (1st column), while space transformation gain is calculated as a difference between the results of both improvements (column 4) and results with only feature fusion (column 2). Such calculation indicates how much gain does space transformation bring on top of feature fusion. Overall gain (column 7) is calculated as a difference between the results of both improvements (column 4) and basic results (column 1). It can be noticed that performance gain reaches the peak at 3 km/h degree of speed change and then starts falling back down to 22% at the largest degree of speed change 5 km/h. This indicates the deficiency in both improvements, which do not cope well with severe speed changes. Nevertheless, the average performance gain for cross-speed walking recognition is 35%.

*5.1. Comparison to Other Works.* For cross-speed walking gait recognition authors in [18] designed two gait recognition tests on OU-ISIR-A dataset to enable comparison with other works handling speed changes. We compare our results to the following state-of-the-art gait recognition methods, which all focus on dealing with walking speed change: HMM-based time normalized (HMM) [14], Stride Normalization (SN) [8], Silhouette Volume Transformation (SVT) [18], High-order Shape Configuration (HSC) [19], and Differential Composition Model (DCM) [20]. The results of HMM and SN methods are based on 25 and 24 subjects from other datasets, while SVT, HSC, and DCM are based on the same OU-ISIR-A dataset used also in our experiments. Similar test scenarios are designed for OU-ISIR-A dataset. For small speed change HMM uses 3.3 km/h and 4.5 km/h, while OU-ISIR-A uses similar 3 km/h and 4 km/h speed change. For large speed change SN uses the speed change between 2.5 km/h and 5.8 km/h, which approximately corresponds to 2 km/h and 6 km/h speed change in OU-ISIR-A dataset.

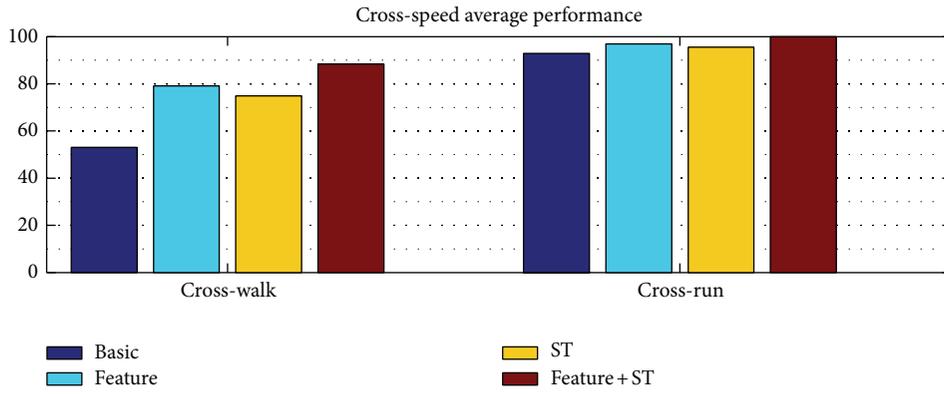
Results in Table 8 show method performance for both previously described gait tests and also the average method performance on the entire OU-ISIR-A dataset (test scenario:

whole dataset). It also gives performance for methods that published these results. It can be seen that our method was outperformed only by DCM method, especially in the scenario with larger speed changes (4 km/h or more). Nevertheless, our method demonstrates state-of-the-art performance by using model based gait features describing gait dynamics, which is greatly affected by walking speed changes and discarding any possible appearance based identification clues that other silhouette based methods benefit from. It can be noticed that no other model based methods can be found among these results, merely because existing model based methods rarely match the performance of nonmodel based techniques even with no covariate factors present. To our best knowledge, our method is also the first model-based method handling walking speed changes and together with the use of dynamic features represents a novel contribution to gait analysis research field.

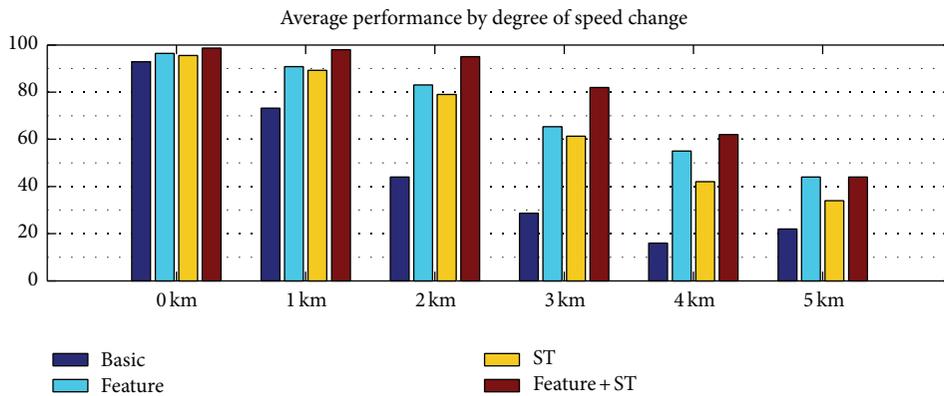
## 6. Conclusion

In this paper we presented a skeleton model based gait recognition system using only features that describe gait dynamics. We described solutions for solving specific problems, that is, two-stage feature alignment (image moments based coarse alignment together with distance based fine alignment) that is crucial for successful comparison of time-series features like ours. We addressed a problem of walking-speed variation and proposed feature fusion and space transformation approach that successfully mitigate its negative effects on recognition performance. Moreover, our features can be understood by human (medical, kinesiological) expert and can handle walking speed changes by undemanding training stage that also has a human understandable interpretation. Such property gives our method the perspective for being used also in other fields like medicine, kinesiology, and sports.

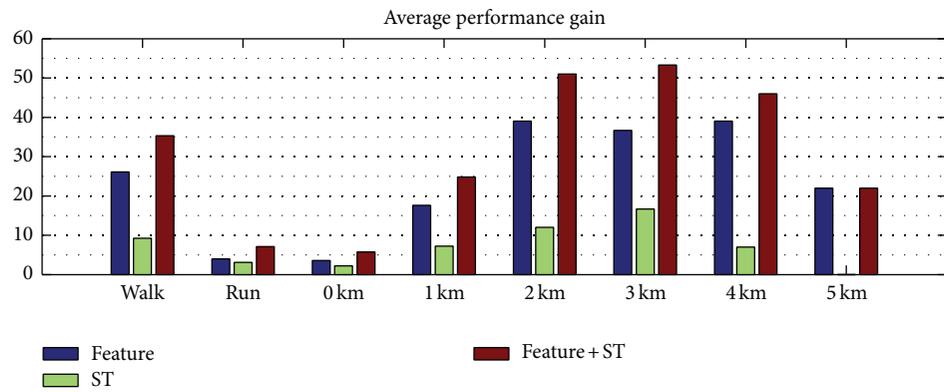
We evaluated the performance of proposed methods on OU-ISIR gait database and proved that model based system concentrating on dynamic features only can demonstrate



(a)



(b)



(c)

FIGURE 7: Recognition performance (%) results summary: (a) average performance of cross-speed walking and running, (b) average performance by degree of speed change (in km/h), and (c) performance gains for feature fusion and space transformation.

state-of-the-art performance and can easily find its place alongside other appearance based state-of-the-art methods. To the best of our knowledge, our method is the first model-based method handling walking speed changes efficiently, that is, comparable to state-of-the-art results. Although OU-ISIR gait database contains relatively small number of 34 subjects, especially in benchmark form as designed by [18], it provides the broadest range of walking speeds and is up-to-date the most appropriate database for studying walking speed effects on human gait. Moreover, general

recognition ability of similar model-based approaches was already addressed by other authors (e.g., [3]), where correct classification rate of 84% was achieved on 100 subjects at unchanged conditions.

Nevertheless, there are further problems that need to be addressed in the future. First, the problem of time series feature alignment should be circumvented by using a distance metric that is starting point invariant. Furthermore, the performance of larger (4 km/h and up) speed change should be investigated and improved. This could be achieved by more

sophisticated feature selection technique on one hand and on the other by finding stronger models of walking speed spaces, which could also tackle the problem of cross-walker-runner identification.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

Operation is partly financed by the European Union, European Social Fund.

## References

- [1] G. Johansson, "Visual motion perception," *Scientific American*, vol. 232, no. 6, pp. 76–88, 1976.
- [2] G. Ariyanto and M. S. Nixon, "Model-based 3D Gait biometrics," in *Proceedings of the International Joint Conference on Biometrics (IJCB '11)*, pp. 354–359, October 2011.
- [3] J.-H. Yoo and M. S. Nixon, "Automated markerless analysis of human Gait motion for recognition and classification," *ETRI Journal*, vol. 33, no. 2, pp. 259–266, 2011.
- [4] T. Yu and J.-H. Zou, "Automatic human Gait imitation and recognition in 3D from monocular video with an uncalibrated camera," *Mathematical Problems in Engineering*, vol. 2012, Article ID 563864, pp. 1–35, 2012.
- [5] 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, Article ID 261317, pp. 1–14, 2008.
- [6] Z. Liu and S. Sarkar, "Simplest representation yet for Gait recognition: averaged silhouette," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR '04)*, vol. 4, pp. 211–214, August 2004.
- [7] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer, "The humanID Gait challenge problem: data sets, performance, and analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162–177, 2005.
- [8] R. Tanawongsuwan and A. Bobick, "Modelling the effects of walking speed on appearance-based Gait recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, pp. II783–II790, July 2004.
- [9] Z. Wang, X. Sun, L. Sun, and Y. Huang, "Manifold adaptive kernel semisupervised discriminant analysis for Gait recognition," *Advances in Mechanical Engineering*, vol. 2013, pp. 1–12, 2013.
- [10] M. Peternel and A. Leonardis, "Visual learning and recognition of a probabilistic spatio-temporal model of cyclic human locomotion," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR '04)*, pp. 146–149, August 2004.
- [11] W. Stone and M. Skubic, "Capturing habitual, in-home Gait parameter trends using an inexpensive depth camera," in *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5106–5109, 2012.
- [12] T. Banerjee, J. M. Keller, and M. Skubic, "Resident identification using kinect depth image data and fuzzy clustering techniques," in *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5102–5105, 2012.
- [13] J. Kovac and P. Peer, "Transformation based Walking speed normalization for Gait recognition," *KSII Transactions on Internet and Information Systems*, vol. 7, no. 11, pp. 2690–2701, 2013.
- [14] Z. Liu and S. Sarkar, "Improved Gait recognition by Gait dynamics normalization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 6, pp. 863–876, 2006.
- [15] S. Lee, Y. Liu, and R. Collins, "Shape variation-based frieze pattern for robust Gait recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, June 2007.
- [16] A. Veeraraghavan, A. Srivastava, A. K. Roy-Chowdhury, and R. Chellappa, "Rate-invariant recognition of humans and their activities," *IEEE Transactions on Image Processing*, vol. 18, no. 6, pp. 1326–1339, 2009.
- [17] T. Kobayashi and N. Otsu, "Three-way auto-correlation approach to motion recognition," *Pattern Recognition Letters*, vol. 30, no. 3, pp. 212–221, 2009.
- [18] A. Tsuji, Y. Makihara, and Y. Yagi, "Silhouette transformation based on walking speed for Gait identification," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 717–722, June 2010.
- [19] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li, "Speed-invariant Gait recognition based on procrustes shape analysis using higher-order shape configuration," in *Proceedings of the 18th IEEE International Conference on Image Processing (ICIP '11)*, pp. 545–548, September 2011.
- [20] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li, "Differential composition model," *Systems, Man, and Cybernetics*, vol. 42, no. 6, pp. 1654–1668, 2012.
- [21] L. Zappella, X. Llado, and J. Salvi, *New Trends in Motion Segmentation, Pattern Recognition*, Edited by P.-Y. Yin, InTech, 2009.
- [22] D. A. Winter, *The Biomechanics and Motor Control of Human Movement*, John Wiley & Sons, 1990.
- [23] X. Huang and N. Boulgouris, "Model-based human Gait recognition using fusion of features," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, pp. 1469–1472, 2009.
- [24] Y. Makihara, H. Mannami, A. Tsuji et al., "The OU-ISIR Gait database comprising the Treadmill dataset," *IPSN Transactions on Computer Vision and Applications*, vol. 4, pp. 53–62, 2012.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

