Gait Recognition: Databases, Representations, and Applications

YASUSHI MAKIHARA[1], DARKO S. MATOVSKI[2], MARK S. NIXON[2], JOHN N. CARTER[2], YASUSHI YAGI[1]

1) Osaka University, Osaka, Japan

2) University of Southampton, Southampton, UK

Abstract

There has been considerable progress in automatic recognition of people by the way they walk since its inception almost 20 years ago: there is now a plethora of technique and data which continue to show that a person's walking is indeed unique. Gait recognition is a behavioural biometric which is available even at a distance from a camera when other biometrics may be occluded, obscured or suffering from insufficient image resolution (e.g. a blurred face image or a face image occluded by mask). Since gait recognition does not require subject cooperation due to its non-invasive capturing process, it is expected to be applied for criminal investigation from CCTV footages in public and private spaces. This article introduces current progress, a research background, and basic approaches for gait recognition in the first three sections, and two important aspects of gait recognition, the gait databases and gait feature representations are described in the following sections.

Publicly available gait databases are essential for benchmarking individual approaches, and such databases should contain a sufficient number of subjects as well as covariate factors to realize statistically reliable performance evaluation and also robust gait recognition. Gait recognition researchers have therefore built such useful gait databases which incorporate subject diversities and/or rich covariate factors.

Gait feature representation is also an important aspect for effective and efficient gait recognition. We describe the two main approaches to representation: model-free (appearance-based) approaches and model-based approaches. In particular, silhouette-based model-free approaches predominate in recent studies and many have been proposed and are described in detail.

Performance evaluation results of such recent gait feature representations on two of the publicly available gait databases are reported: USF Human ID with rich covariate factors such as views, surface, bag, shoes, time elapse; and OU-ISIR LP with more than 4,000 subjects. Since gait recognition is suitable for criminal investigation applications of the gait recognition to forensics are addressed with real criminal cases in the application section. Finally, several open problems of the gait recognition are discussed to show future research avenues of the gait recognition.

Keywords

Gait, walking pattern, behavioural biometrics, forensics, appearance-based approach, model-based approach, silhouette

Time Period	#Subjects	Source	Rank-1 identification rate	Notes
1990's	~ 10	USC	95.2	
2000's	~ 120	USF HumanID, CASIA, SOTON Large Database	75% - 99%	Recognition rate depends on covariates
Recent (2010)	>300	SOTON Multimodal	95% - 100%	Includes time dependent covariates
Recent (2012)	> 4,000	OU-ISIR, LP	97.5%	No covariates

Table 1 - Progression of Gait Recognition Systems

Current Progress in Automated Recognition by Gait

Automatically recognising people by their gait via machine perception has made enormous progress since the 1990s. Then, research used databases with tens of subjects, showing that correct identification could be made. A decade later people progressed to databases of around 100 people, still showing that identification could be achieved. As computers became more powerful and memory became cheaper the databases moved into the thousands, and high recognition rates have been achieved. The most recent database is of more than 4000 subjects with a correct classification rate of 97.5%. Part of this progress has been due to the studies of the factors that affect the perception of the way people walk. This article reviews that progress and shows how the current position has arisen.

Background

Gait has some important advantages over other forms of biometric authentication. It can be acquired at a distance when other biometrics are obscured or the resolution is insufficient. It does not require subject cooperation and can be acquired in a non-invasive manner. It is easy to observe and hard to disguise as walking is necessary for human mobility. Gait can be acquired from a temporal sequence of images (e.g. a video) and may be partly acquired even from a single still image (e.g. a single still image at the double support phase may contain a stride as one of gait information).

Shakespeare made several references to the individuality of gait, e.g. in *The Tempest* [Act 4 Scene 1], Cares observes "*High'st Queen of state, Great Juno comes; I know her by her gait*", in Henry IV Part II [Act 2, Scene 3] "*To seem like him: so that, in speech, in gait, in diet, in affections of delight, in military rules, humours of blood, he was the mark and glass, copy and book*".

The aim of medical research has been to classify the components of gait for the treatment of pathologically abnormal patients (e.g. Parkinson's disease, degenerative hip disease, and idiopathic normal pressure hydrocephalus). Work (1) created standard movement patterns for pathologically normal people. Those patterns were then used to identify pathologically abnormal patients.

The biomechanics literature makes observations concerning identity: "A given person will perform his or her walking pattern in a fairly repeatable and characteristic way, sufficiently unique that it is possible to recognize a person at a distance by their gait" (2).

Psycho-physiological studies such as (3) and (4) have shown that humans can recognise friends and the sex of a person solely by their gait with 70-80% accuracy. Moreover, a biological motion study (5) has shown that several factors such as the sex, the body weight, feeling (e.g. nervous or relaxed), and emotion (e.g. happy or sad) clearly appear in their gaits even if the gaits are displayed as a simple biological motion (i.e. point light sources attached to the human joints). One early work highlighted the human ability to recognise people by gait to identify six individuals on the basis of their gait under conditions of simulated daylight, simulated dusk and via point-light displays (113). These and similar studies via human perceptions have inspired the use of gait as a biometric trait via machine perception.

Recently, there has been a rapid growth in the number of surveillance systems, aimed to improve safety and security. These systems are yet to include recognition capabilities and automatic gait recognition through machine perception (e.g. pattern recognition from CCVT footage) could be a most suitable choice. The primary aim of surveillance videos is to monitor people. However, the video data can be of a low quality (poor resolution, time lapse etc.) and the subject can try to conceal the more conventional biometrics. Nevertheless, such video can provide sufficient data for gait recognition technology and there is already research in using gait biometrics as forensic tools (6-8). Gait recognition could be employed at a border crossing or any high throughput environment. Gait contains very rich information and is considered to be unique. Studies have shown that gait can also be used to reveal a person's identity, gender (9-12), age (13-16), emotional state (17) etc.

Recognition by gait is one of the newest biometrics, since its development only started when computer memory and processing speed became sufficient to process sequences of image data with reasonable performance. The potential for gait recognition is great and hence there is a vast interest in computer vision and pattern recognition research in extracting and matching gait features.

Theory

A gait recognition system primarily consists of a computer vision system. A gait signature is created by extracting images of a walking subject which is then compared to the signatures of known subjects. Figure 1 shows an example of some of the basic steps in a gait recognition system.

Step 1 – Acquire a gait sample

Step 2 – Pre-processing

Step 3 – Find a suitable Gait Cycle
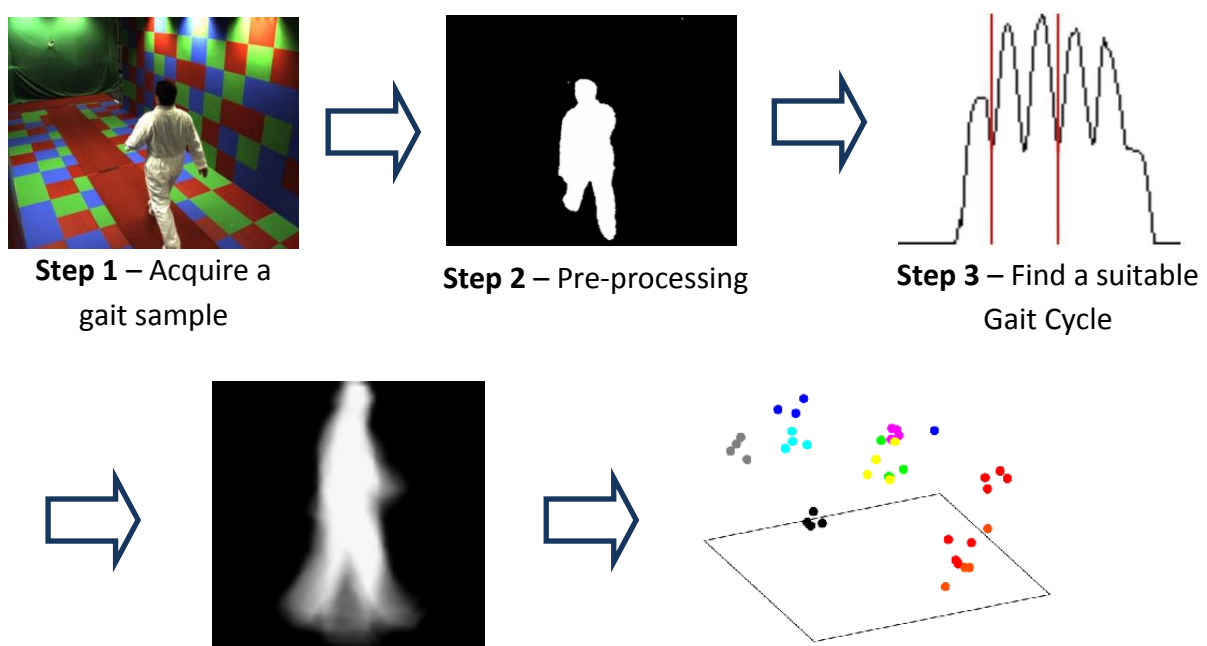
Step 4 – Create a gait
signature

Step 5 – Classification

Figure 1 - General steps of a gait recognition system

Step 1: Data can be acquired using a single or multiple cameras. If data is acquired using a single camera, recognition can be performed using a 2D gait signature shown in step 4 in Fig. 1. However, if multiple but synchronised cameras are used, the number of possibilities is greater. Examples of the usage of multiple-synchronised cameras include:

- Producing a 3D gait model and using it for recognition
- 3D information can be used to improve recognition for a 2D approach by producing a 2D signature with depth information (i.e. 2.5D gait signature).
- Achieving a view-invariant recognition. A gait signature from any view can be re-created using 3D data which can be mapped to a signature acquired in an outdoor environment where typically only a single and non-optimal view is available.

Some of the above properties can be also exploited if data is acquired using a single depth sensor. In particular, since an inexpensive depth sensor (e.g. Microsoft Kinect ™) is recently available, the depth-based gait signatures are promising choice for indoor and middle-range applications (e.g. access control and border control).

Step 2: An example of pre-processing step is background subtraction or background segmentation. The subject can be acquired easily and reliably in the form of silhouette by using chroma-keying if there are clear colour difference between the subject and the background. Background subtraction can measure the naturally occurring scene behind the walking subject using one of a plethora of computer vision techniques.

Step 3: As human gait is periodic, a gait sequence (sample) can consist of multiple gait cycles. Identifying the most suitable cycle can lead to better recognition rates. Signal processing techniques (e.g. peak detection or maximizing autocorrelation) can be applied to the foreground signal (e.g. sum of foreground pixels at the whole body or the lower body (23), aspect ratio of the foreground region (24), or a set of foreground pixels themselves as a multi-dimensional signal (25)) in the case of binary image. Moreover, a recent advanced signal processing technique (26) enables us to handle quasi-periodic gait image sequence and gait cycle detection by sub-frame order.

Step 4: There are number of approaches to produce a gait signature, some of which are described later. A baseline gait signature was proposed in (23). An example of a signature is shown in Step 4 in Fig. 1.

Step 5: A gait signature can be used directly within a classifier. Alternatively, features can be extracted from a signature and those features can be used for classification. Again, there is a selection of classification techniques; in the simplest case, a classifier such as k-nearest neighbour (using Euclidian distance) can be used. In a more elaborate case, a combination of principal component analysis (PCA) and the subsequent linear discriminant analysis (LDA) can be a reasonable

choice (27). In the most recent study, primal rank support vector machine (Primal RankSVM) (28) is employed to cope with general types of covariate factors in gait recognition (29).

Databases

A database can be collected for various purposes. Primary concerns include uniqueness and practicality. A database should contain enough subjects to allow for an estimate of inter- and intra-subject variation and for statistically reliable performance evaluation. The current databases contain smaller number of subjects compared to databases used to evaluate performance of other biometrics (e.g. face, fingerprint). However, there are databases that include covariate factors and application potential. Some of the most well-known databases together with some of their characteristics are shown in Table 2.

Name	Subjects	Sequences	Covariates	Viewpoints	Indoor(I)/ Outdoor(O)
CMU MoBo (30)	25	600	Y	6	I (Treadmill)
Georgia Tech (31)	15	268	Y	-	O
	18	20	Y	-	-
HID-UMD (32)	25	100	N	1	O
	55	222	Y	2	O
SOTON Small Database (33)	12	-	Y	3	I
SOTON Large Database (34)	115	2,128	Y	2	I/O
SOTON Multimodal (35)	>300	>5,000	Y	12	I
SOTON Temporal (36)	25	2,280	Y	12	I
USF HumanID (23)	122	1,870	Y	2	O
CASIA A (37)	20	240	Y	3	I
CASIA B (38)	124	1,240	Y	11	I
CASIA C (39)	153	1,530	Y	1	O
OU-ISIR, Treadmill A (40)	34	612	Y	1	I (Treadmill)
OU-ISIR, Treadmill B (41)	68	2,764	Y	1	I (Treadmill)
OU-ISIR, Treadmill C (42)	200	200	Y	25	I (Treadmill)
OU-ISIR, Treadmill D (43)	185	370	N	1	I (Treadmill)
OU-ISIR, LP (44)	4,007	7,842	N	2	I
TUM-IITKGP (45)	35	850	Y	1	O
TUM-GAID (46)	305	3,370	Y	1	O
WOSG (47)	155	684	Y	8	O

Table 2 - Details of some of the well-known gait databases

SOTON Large Database (34) was the first gait database to contain over 100 subjects, walking on a track indoor, a treadmill, and outdoor and who are observed from two views: side view and oblique view. Both SOTON Multimodal (35)[1] and SOTON Temporal (36) were collected at the biometric tunnel (48), which is described later. Particularly, SOTON Temporal contains the largest variations w.r.t. time elapse; a gait sequence was captured monthly for a year with controlled and uncontrolled clothing condition. It is therefore suitable for purely investigating the time elapse effect on the gait recognition by excluding clothing effect.

USF HumanID (23) is one of the most frequently used gait database since it contains over 100 subjects who walk along an ellipsoidal path outdoor, as well as contains a variety of covariates: view; surface; shoes; bag; and time elapse, although each covariate has at most two variations (e.g. concrete and grass for surface covariate, with and without bag for bag covariate). This database is suitable for investigating which covariate significantly impacts on the gait recognition performance. More details are found in the section of experimental results.

CASIA A (37) is an initial dataset of CASIA Gait Database and hence contains a limited number of subjects (i.e. 20) and view variations: front view; oblique view; and side view. On the other hand, CASIA B (38) is a frequently used gait database since it contains large view variations from front view (0 deg) to rear view (180 deg) with 18-degree interval. This database recorded ten sequences per subject: six normal sequences; two sequences with a long coat; two sequences with a backpack. This database is therefore suitable for evaluating cross-view gait recognition as well as investigating the clothing and carrying status effects on the gait recognition. CASIA C (39) contains the largest number of subjects (153) among the CASIA gait databases and each subject was captured by an IR camera at the outdoor night session, which is suitable for evaluating gait recognition as a surveillance task during night.

OU-ISIR Treadmill A (40) and B (41) contain the largest variations of speed from 2 km/h to 10 km/h with 1 km/h interval, and clothing up to 32 combinations, respectively, although the numbers of subjects are small (34 and 68, respectively). These databases are naturally suitable for evaluating cross-speed and clothing-invariant gait recognition. OU-ISIR Treadmill C (42) contains 200 subjects with 25 views: a combination of 12 azimuth views with 30 degree-interval and two tilt views; and one top view. OU-ISIR Treadmill D (43) contains 185 subjects with various degrees of fluctuations (e.g. different motions among gait cycles). Data acquisition setting and performance evaluations are summarized in (49). On the other hand, OU-ISIR LP (44) contains the largest number of subjects (over 4,000) with a wide age range from 1 year old to 94 years old also with an almost balanced gender ratio, although it does not contain any covariates. It is therefore suitable for estimating a sort of upper bound accuracy of the gait recognition with high statistical reliability. It is also suitable for evaluating gait-based age estimation.

TUM-IITKGP (45) contains some unique covariates other than those in USF HumanID: hands in pocket; dynamic occlusion (two people walking past); and static occlusion (two people just standing in the line of sight), which are possibly observed in a real situation. Although the number of subjects are not many, this database can be used as an initial benchmark for occlusion-handling gait recognition. TUM-GAID (46) is the first multi-signal gait database containing gait audio signals, RGB images, and depth images by Microsoft Kinect™ and also relatively large number of subjects (i.e. 305). This database contains some covariates: normal; backpack; shoes coating to test the robustness of gait audio signals; and time elapsed versions of them. Although this database is unsuitable for evaluating the gait recognition in wide-area surveillance scenarios (i.e. at a large distance) since the observation distance is relatively close (approximately 3.0 meter) due to the limitation of depth image acquisition requirement by Microsoft Kinect™, it is still useful for evaluating the multi-modal gait recognition in access control scenarios.

WOSG (47) contains 155 subjects with eight view variations. Whereas the other databases mainly focusing on the individual aspects (e.g. walking speed, carrying status, clothing), this database more

focuses on a less constrained environment which is likely to occur in an operational setting. More specifically, each of subject was captured with the short-wave-infrared spectrum (900 nm - 1,700 nm) in an active, outdoor environment, wherein environmental factors such as cloud cover (impacting illumination) and scene factors such motion artifacts due to trees or additional persons (impacting segmentation and tracking) exist.

In summary, challenging aspects of individual gait databases roughly fall into three: subject diversity, individual factors (e.g. clothing, speed, carrying status), and environmental or scene factors (e.g. view directions and illumination changes in outdoor). Moreover, the importance of these challenging aspects is highly dependent on application scenarios. For example, in an access control scenario, the environment is usually fixed and even appropriately designed for better authentication (e.g. indoor environment such as a corridor in front of door, relatively close observation distance, multi-modal information source such as an RGB camera and a depth camera, multi-modal biometric cues such as face, ear, the height as well as gait, are acceptable). Moreover, the individual factors are controlled to some extent since a user (client) wants to be successfully authenticated (i.e. cooperative). Consequently, the subject diversity is the most important aspect in the access control scenario and hence SOTON Multimodal (35), OU-ISIR LP (45), and TUM-GAID (46) are suitable for evaluating the access control scenario.

On the other hand, in a surveillance or criminal investigation scenario, the environment is usually unfixed and a subject such as a perpetrator or a suspect is uncooperative. Both the individual and environmental factors are therefore important as well as the subject diversity. SOTON Large Database (34) and USF HumanID (23) provide a variety of such individual and environmental factors, and hence they are suitable exploratory factor analysis in the surveillance and criminal investigation scenario. However, the variation of each factor is limited in SOTON Large Database (34) and USF HumanID (23), other databases are still useful when focusing on each factor (e.g. CASIA B (38) for view variations, OU-ISIR Treadmill A (40) for speed variations, OU-ISIR Treadmill B for clothing variation, and SOTON Temporal (36) for time elapse variations, WOSG (47) for environmental variations).

Approaches to Gait Recognition

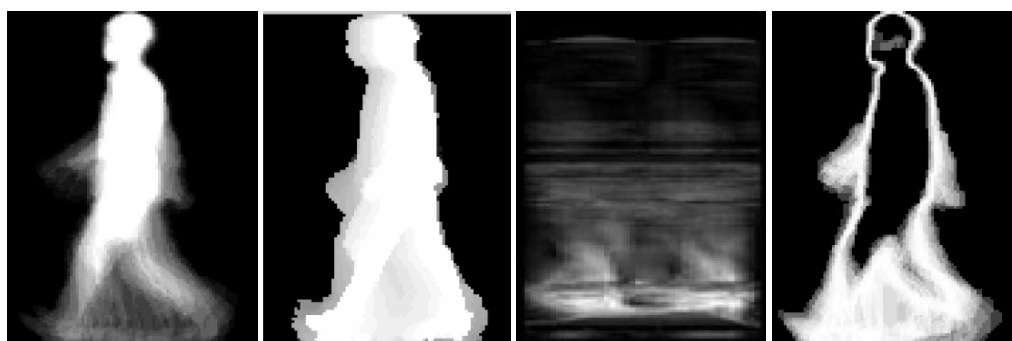| Model-free approach	Model-based approach
Spatiotemporal pattern (50)	Single oscillator (82)
Parametric eigen space (51)	Ellipsoidal fits (83)
Shape of motion (52)	Kinematic features (84)
PCA + canonical analysis (53)	Stride parameters (85)
Moments (54)	Human parameters (86)
Self-similarity plot (55)	Joint trajectories (87)
Key frame analysis (56)	Articulated model (88)
Frieze patterns (57)	Dual oscillator (89)
Unwrapped silhouette (58)	Linked feature trajectories (90)
Eigenspace sequences (24)	Stick model (91)
Relational statistics (58)	Layered deformable model (92)
Area (59)	Marionette mass-spring model (93)
Symmetry (60)	
Key poses (61)	
Hidden Markov model (62)	

Average silhouette (63) Gait style and content (64) Cubic higher-order local auto-correlation (65) Silhouette similarity (23) Fourier descriptor (66) Motion silhouette image (67) Gait energy image (27) Frequency-domain feature (25) Gabor filter (68) 3D distribution transform (69) Motion silhouette contour template (70) Shape variation-based frieze pattern (71) Flow field (72) Gait entropy image (73) Masked gait energy image (74) Gait curve (75) Gait flow image (76) Gait energy volume (77) Chrono-gait image (78) Gradient histogram energy image (22) Depth gradient histogram energy image (79) Depth-based gait feature (80) Local Gauss map (81)	

Table 3 - Approaches to Gait Recognition

The approaches to gait recognition can be divided in two main groups: *model-based* and *model-free* (*appearance-based*) (see Table 3). Model-based approaches use the human body structure and model-free methods use the whole motion pattern of the human body. Which approach is adopted depends on the acquisition conditions. Model-free approaches use the input images directly to produce a gait signature without fitting a model. These approaches can perform recognition at lower resolutions which makes then suitable for outdoor applications, where a subject can be at a large distance from the camera. Model-based approaches typically require higher resolution images of a subject to be able to fit the model accurately as well as relatively high computational cost.

We refer the readers to (94, 95) for surveys of model-based and model-free approaches at a relatively early research stage.

Model-Free Approaches

GEI   MSI   SVB frieze pattern   GEnI
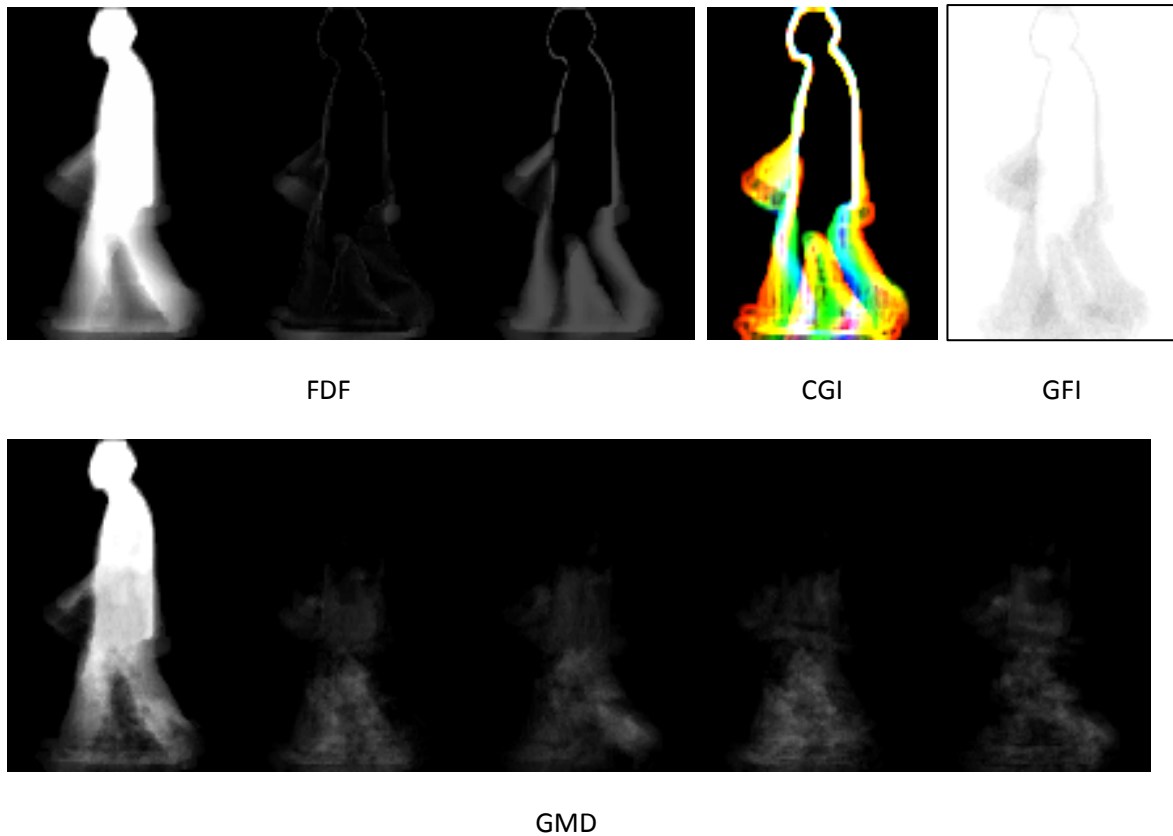
FDF · CGI · GFI

GMD

Figure 2 - Examples of model-free gait signatures

The model-free approaches derive the human silhouette by separating the moving object from the background. The subject can then be recognized by measurements that reflect shape and/or movement. The simplest approach is to directly compute the dissimilarity (e.g. Euclidean distance, cosine distance, Tanimoto distance) between two silhouette sequences (i.e. a probe and a gallery) in conjunction with phase synchronization by frame shifting, which is called a baseline algorithm (23) in the gait recognition community. Another simple approach is an averaged silhouette over a complete gait cycle (63) which is also known as gait energy image (GEI) (27) (see Fig. 2). Motion silhouette image (MSI) (67) is a similar representation to the GEI. The value of each pixel is computed as a function of motion of that pixel in the temporal dimension over all silhouettes that are part of a single gait cycle (see Fig. 2). Both the GEI and MSI are easy to compute but they are vulnerable to appearance changes of the human silhouette. Frieze pattern (57) represents the information contained in a gait sequence by horizontal and vertical projections of the silhouettes. Its extension, shape variation-based (SVB) frieze patterns (71) use key frame subtraction in order to mitigate the effects of appearance changes on the silhouette (see Fig. 2). The gait entropy image (GEnI) (73) is another example of a compact gait representation (signature). GEnI is computed by calculating the Shannon entropy for each pixel in a silhouette image sequence. Shannon entropy measures the uncertainty associated with a random variable.

The gait signatures for the approaches shown in Fig. 2 are usually used directly for classification. There are additional ways of extracting gait signatures without using a model. Some examples are described below:

- Niyogi and Adelson (50) represent the walking image sequence as an xyt-volume and fit the xyt pattern with a smooth spatio-temporal surface after change detection. The spatio-

temporal surface is expressed as a combination of a standard parameterized surface (i.e. a canonical walk) and a subject-specific deviation surface.

- Murase and Sakai (51) applied PCA to a set of silhouettes and then represents the silhouette sequence over a complete gait cycle as a closed curve in the eigenspace based on an idea of parametric eigenspace method (96), where a gait stance (phase) is handled as a parameter. Matching is done with linear time normalization so as to synchronize phases between a pair of a probe and a gallery silhouette sequences.

- Little and Boyd (52) derived a dense optical flow for each image sequence. Scale-independent scalar features of each flow, based on moments of the moving point characterize the spatial distribution of the flow. The periodic structure of these sequences of scalars is analysed. The scalar sequences for an image sequence have the same fundamental period but differ in phase, which is used as a feature for recognition of individuals by the shape of their motion.

- BenAbdelkader et al. (85) used background modelling to track the subject for a number of frames and extract a sequence of segmented images of the person. A self-similarity plot is computed via correlation of each pair of images in this sequence. For recognition, PCA is used to reduce the dimensionality of the plots. A k-nearest neighbour rule is used on the reduced space for classification.

- Another silhouette-based gait recognition technique using PCA has been proposed by Liang et al. (97). Eigenspace transformation based on PCA is applied to time-varying distance signals derived from a sequence of silhouette images to reduce the dimensionality of the input feature space. Supervised pattern classification techniques are performed in the lower-dimensional eigenspace for recognition.

- Hayfron-Acquah et al. (60) proposes a method for automatic gait recognition based on analysing the symmetry of human motion. The generalised symmetry operator is used to locate features according to their symmetrical properties rather than relying on the boarders of a shape. The symmetry operator is used on the optical flow image to produce a gait signature. For purposes of classification, the similarity differences between the Fourier descriptions of the gait signatures are calculated using Euclidean distance.

- Gait is a temporal sequence and can be modelled using hidden Markov models (HMM). The statistical nature makes the model relatively robust. The postures that an individual adopts are regarded as states of the HMM and are typical to that individual and provide means of discrimination in (98).

- Wang et al. (99) extract moving silhouettes by an improved background subtraction scheme. The extracted silhouette is unwrapped by contour following and it is represented in a complex number coordinate (i.e. horizontal and vertical axes are real and imaginary parts of a complex number) whose origin is aligned to the silhouette centroid. Thereafter, Procrustes shape analysis is applied to the unwrapped silhouette and finally mean Procrustes shape distance is used as similarity measure between a pair of silhouette sequences.

- Wang et al. (24) also extracts moving silhouettes in the same way as (99) and then extract a series of distances between each contour point and the silhouette centroid as a one-dimensional signal. The signal is then projected into the eigenspace by PCA and spatio-temporal correlation, which is extension from 2D image correlation to 3D correlation in the space and time domain, is used for matching in conjunction with linear time normalization.

- Vega and Sarkar (58) extract moving edge pixels and represent relations between all the pairs of such moving edge pixels in space of probability function (SoPF), where the domain is defined as a set of normalized distance and orientations between a pair of moving edge pixels. PCA is then applied to SoPF for dimension reduction and a sequence of the dimension-reduced SoPF is used for matching with Euclidean distance in conjunction with frame shifting or linear time normalization for phase synchronization.

- Kale et al. (62) use two different image features to directly train a HMM: the width of the outer contour of a binary silhouette (i.e. width vector); and the entire binary silhouette itself.

- Lee and Elgammal (64) introduce a bilinear decomposition of gait data into time-invariant gait-style and time-dependent gait-content factors. More specifically, they developed a generative model by embedding gait sequences into a unit circle and learning nonlinear mapping from a point on the unit circle to the silhouette space to enable us to synthesize temporally-aligned gait sequences. Given such synthesized gait data, an invariant gait style is separated from the time-dependent gait-content by the bilinear model and is used for recognition.

- Kobayashi and Otsu (65) employ a segmentation-free representation named cubic higher-order local auto-correlation (CHLAC). They first extract moving pixels by frame subtraction and then compute the summation of correlation between the moving pixels and a set of local spatio-temporal mask patterns (e.g. a cuboid of 3 pixels by 3 pixels by 3 frames) by scanning over the entire image and over a predefined number of frames. This results in a feature vector whose dimension is the same as the number of local spatio-temporal mask patterns. In a training stage, fisher discriminant analysis is learnt by a set of training samples and then Euclidean distance in the discriminant space is employed for dissimilarity measure in a matching stage.

- Mowbray and Nixon (66) regard a set of two-dimensional positions of silhouette contours over a complete gait cycle as a spatio-temporal periodic signal and hence extract a set of Fourier descriptors in the spatio-temporal domain. For the discrimination capability and dimension reduction purposes, only low-order Fourier descriptors are used for classification.

- Makihara et al. (25) first detect a complete gait cycle by maximizing the autocorrelation of a full spatio-temporal silhouette sequences rather than detecting peaks of areas or aspect ratio so as to cope with view variety in a more robust way. They apply one-dimensional Fourier analysis to the silhouette sequence over the complete gait cycle (i.e. a base period for Fourier analysis is the complete gait cycle) along the temporal axis for each pixel independently, and subsequently computes amplitude spectra of Fourier coefficients. They also use low-order Fourier coefficients (i.e. 0, 1, and 2) due to a similar reason to (66) and a set of such amplitude images is defined as a frequency-domain feature (FDF, see Fig. 2). Note that an image composed of 0-times elements correspond to the averaged silhouette (63) or GEI (27), while those of one-time and two-times frequency correspond to asymmetric and symmetric motions in case of a side-view image, respectively.

- Bashir et al. (72) compute the flow fields and decompose it into five components: stationary region; leftward motion; rightward motion; upward motion; and downward motion, which constructs a gait motion descriptor (GMD) (see Fig. 2). Since they consider that the downward motion does not contain much useful information for classification, they

compute the dissimilarity between the gait motion descriptors other than the downward motion components.

- Bashir et al. (74) introduce a masked GEI, which is a sort of modified versions of GEI, where a pixel with a small entropy (i.e. almost background or foreground during a complete gait cycle) is masked out (i.e. set to zero so as to make it out of consideration for matching). This operation mitigates the effect of clothing and carrying status variations to some extent since the large portion of dissimilarity due to such covariates is excluded. For example, assuming that a probe with a backpack has a definite foreground region within the backpack region and that a gallery without a backpack has a definite background region within the same region, the both regions are successfully masked out from dissimilarity computation since entropies for both regions are small.

- DeCann and Ross (75) firstly divide a silhouette into the left side and right side by a coronal plane and then extract a left-most and right-most silhouette contour for each height as a gait curve. The evolution of the gait curve across time is regarded as a spatio-temporal gait curve and it is averaged over a complete gait cycle. Through analysis on the averaged gait curve for backpack detection, the corrected gait curved without the backpack is estimated and used for matching.

- Lam et al. (76) extract an optical flow field from a silhouette sequence (i.e. optical flow only appears the silhouette boundary) and then activate pixels whose resultant magnitude of the optical flow is greater than a threshold so as to make a binary image (white: region without motion, black: region with motion). A sequence of such binary images are averaged over a complete gait cycle so as to construct a gait flow image (GFI, see Fig. 2).

- Wang et al. (78) points out that GEI and its variants loses important temporal information (gait stance or phase) in a gait sequence by simply averaging over a complete gait cycle and hence propose a template keeping the temporal information. More specifically, they first extract a silhouette contour and assign gait stance-dependent colour gradation for each contour (e.g. red for double support phase and blue for single support phase). A sequence of such coloured contours are averaged over a quarter gait cycle to generate a chrono-gait image (CGI, see Fig. 2).

- Whereas the most of the approach is based on the silhouette, Hofmann and Rigoll (22) exploits the edge gradient information inside the person silhouette via well-known form of histogram of oriented gradient (HOG) (100), and average them for each spatial cell over a complete gait cycle, which results in gradient histogram energy image (GHEI). To avoid the effect of background, they compute the edge gradient only for segmented silhouette and also for alpha-matte silhouette for better performance, which are named as s-GHEI and alpha-GHEI, respectively.

- El-Alfy et al. (81) captures a silhouette shape property by voting the normal vector of the smoothed silhouette contour into Gauss sphere (unit sphere). To enhance the locality in spatio-temporal domain, voting is done for individual spatio-temporal cells rather than an entire image sequences over a complete gait cycle, which leads to a histogram of normal vector (HoNV).

While all the above approaches are based on a normal RGB image or a grey-scale image, we can also employ a depth image captured by a commercial and inexpensive sensor (e.g. Microsoft Kinect™) which increases an amount of information as well as makes the foreground

segmentation easy. A couple of depth-based or three-dimensional volumetric representation-based approaches to the model-free analysis are introduced below.

- Ioannidis et al. (69) assigns depth-related data to silhouette sequences using two new transforms: the 3-D radial silhouette distribution transform; and the 3-D geodesic silhouette distribution transform. They extract three different features: two features based on the generalized radon transform (i.e. the radial integration transform and the circular integration transform); and a feature based on the weighted Krawtchouk moments, and they fuse them via a genetic algorithm in a matching stage.
- Sivapalan et al. (77) propose a three-dimensional version of averaged silhouette (63) or GEI (27) named gait energy volume (GEV). They firstly reconstruct a sequence of three-dimensional volumes by multi-view silhouettes and then spatially align them. Finally, a sequence of the registered three-dimensional volumes are averaged over a complete gait cycle, which corresponds to the GEV. They also show a variant of GEV which constructed from a frontal depth image with backside filling.
- Hofmann et al. (46) extend GHEI (45) into the depth domain. They compute a gradient of a depth image and average them for each spatial cell over a complete gait cycle, which results in depth gradient histogram energy image (DGHEI).
- Nakajima et al. (80) firstly reconstruct a depth image from a canonical view (not only azimuth and tilt angles but also the distance from camera to a subject) and then apply pixel-wise one-dimensional Fourier transform for temporal axis in analogous to (25). They extract three types of gait features as results of Fourier analysis: direct current elements; amplitude spectra; and phase information.

Model-Based Approaches

The advantages of the previous approaches (silhouette or features derived from it) are computational efficiency and simplicity. However, model-based approaches have greater invariant properties and are better at handling occlusion, noise, scale and rotation. Model-based approaches require a high resolution and are not yet very suitable for outdoor surveillance.

Model-based approaches incorporate knowledge of the shape and dynamics of the human body into the extraction process. These approaches extract features that fit a physical model of the human body. A gait model consists of shapes of various body parts and how those shapes move relative to each other (motion model). The shape model for a human subject can use ellipse to describe the head and the torso, quadrilaterals to describe the limbs and rectangles to describe the feet. Alternatively arbitrary shapes could be used to describe the edges of the body parts. The motion model describes the dynamics of the motion of the different body parts. Using a model ensures that only image data corresponding to allowable human shape and motion is extracted, reducing the effect of noise. The models can be 2 or 3 dimensional. Most of the current models are 2 dimensional, but deliver good results on databases of more than 100 subjects.

Some examples of model-based approaches are described below:

- Yam et al. (89) have used pendular motion and the understanding of biomechanics of human locomotion to develop two models: a bilateral symmetric and analytical model (employs the concept of forced couple oscillator). The gait signature is the phase-weighted magnitude of the Fourier description of both the thigh and knee rotation.
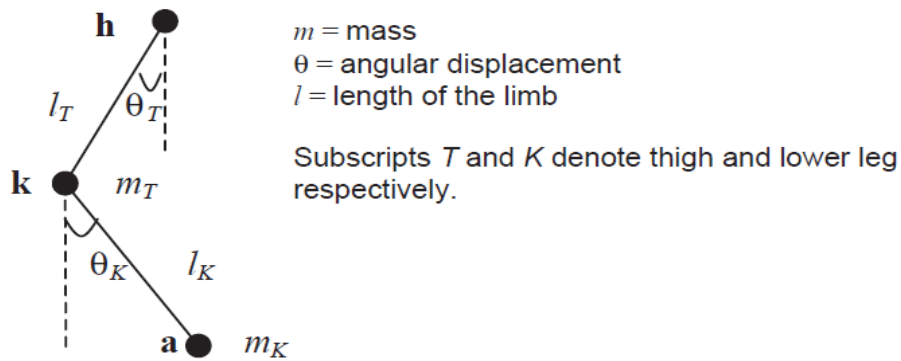

m = mass
θ = angular displacement
l = length of the limb

Subscripts T and K denote thigh and lower leg respectively.

Figure 3 - Example of a gait model – The dynamically coupled pendulum model (Reprinted from (89), Copyright (2004), with permission from Elsevier)

- Bouchrika and Nixon (101) have proposed a new approach to extract human joints. Spatial model templates for human motion are derived from the analysis of gait data collected from manual labelling. Motion templates describing the motion of the joints are parameterised using the elliptic Fourier descriptors.

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} X(t)S_x \\ Y(t)S_y \end{bmatrix}$$

where α is the rotation angle, S_x and S_y are the scaling factors across the horizontal and vertical axes respectively and $X(t)$ and $Y(t)$ are the Fourier summation. Hough transform is used in the feature extraction process.

- Wang et al. (87) have proposed an algorithm based upon the fusion of static and dynamic body information. The static body information is in a form of a compact representation obtained by Procrustes shape analysis. The dynamic information is obtained by a model based approach which tracks the subject and recover joint-angle trajectories of lower limbs. A fusion at the decision level is used to improve recognition results.
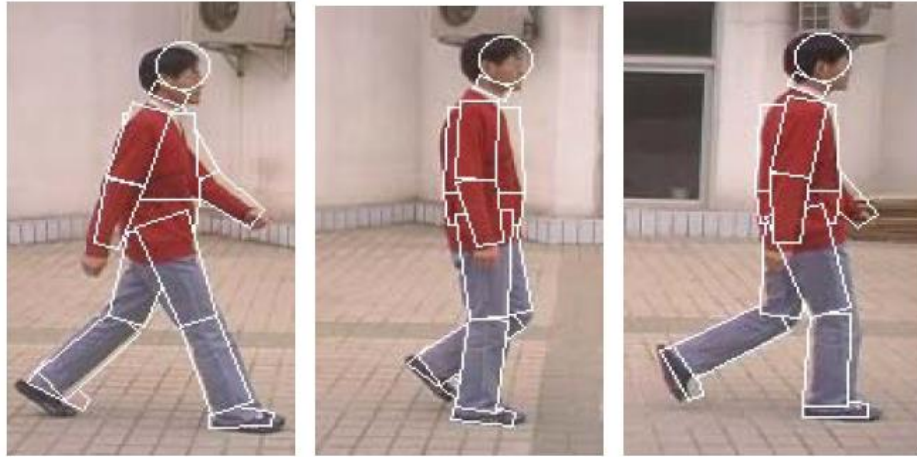
Figure 4 - Example of results (Copyright [2004] IEEE. Reprinted, with permission, from (87))

There have been moves towards developing 3D gait models. Examples of work in this fields are (102,103). Guoying et al. (103) use video sequences from multiple cameras to construct 3D human models. The motion is tracked by applying a local optimisation algorithm. The length of key segments is extracted as static parameters and the motion trajectories of lower limbs are used as dynamic features. Linear time normalisation is used for matching and recognition. 3D approaches are robust to changes in viewpoint and have a great potential. However at present, experiments only on small databases are possible mainly due to high computational requirements.

Gait is dependent on large number of parameters (joint angles and body segment size) which leads to complex models with many free parameters. Finding the best fit model for a subject leads to searching a high-dimensional parameter space. Therefore, there is a trade-off between the accuracy of the model (complexity) and computational cost. The models are often simplified based on certain acceptable assumptions, e.g. a system could assume constant walking speed. However, as computing power increases the problems arising of high complexity can be mitigated.

We finally discuss the suitability of the approaches described above in the context of the two application scenarios: access control and surveillance. In the access control scenario, since we can design the environment as we have already discussed in the section of databases, it is possible to capture a subject with relatively high image resolution (i.e. capturing at a close distance). Consequently, not only model-free approaches but also model-based approaches are suitable. Moreover, it is allowed to use a depth camera in this scenario, recent depth-based gait features (69, 77, 46, 80) are promising since they have possibilities for higher accuracy due to richer information than two-dimensional silhouette-based approaches such as GEI (27).

On the other hand, in the surveillance or criminal investigation scenario, the suitability of the approaches is heavily dependent on actually captured images. For example, in case where the spatial resolution of captured subject in the images is relatively high (e.g. convenience-store robbery) and where clothing conditions are different between a matching pair, the model-based approaches are suitable due to its clothing invariances. In addition, when observation views are different between the matching pair, the 3-D model-based approaches (102, 103) are useful. On contrary, when the spatial resolution is low, the model-free approaches are more suitable than the model-based approaches. Although the model-free approaches are essentially affected by the individual factors

such as view and clothing, SBV frieze pattern (71) may mitigate the effect of the clothing variation to some extent.

Experimental Results

The HumanID gait challenge problem (23) was set up to outline a baseline algorithm for gait recognition and propose a number of difficult experiments including various covariates for the existing gait matchers. The gallery set consists of 122 subjects walking on a grass surface recorded by a single camera.

Table shows the differences of the probe set compared to the gallery set for each of the challenge experiments.

Experiment	Probe	# of Subjects	Difference
A	Different camera view than gallery	122	View
B	Subjects wore different shoes	54	Shoe
C	Different camera view and different shoes	54	Shoe, View
D	Subjects walked on a different surface	121	Surface
E	Different shoes and different walking surface	60	Surface, Shoes
F	Different walking surface and different camera view	121	Surface, View
G	Different walking surface, different shoes and different camera view	60	Surface, Shoe, View

Table 3 – Some experiments comprising the HumanID gait challenge problem

The results in Figure 5 show the progress in gait recognition over a period of 2 years for the experiments shown in Table . We can see that relatively high identification rates are achieved against some of covariate factors such as views and shoe (Experiments A, B, and C), while that surface is still challenging problem (Experiments D, E, F, and G).
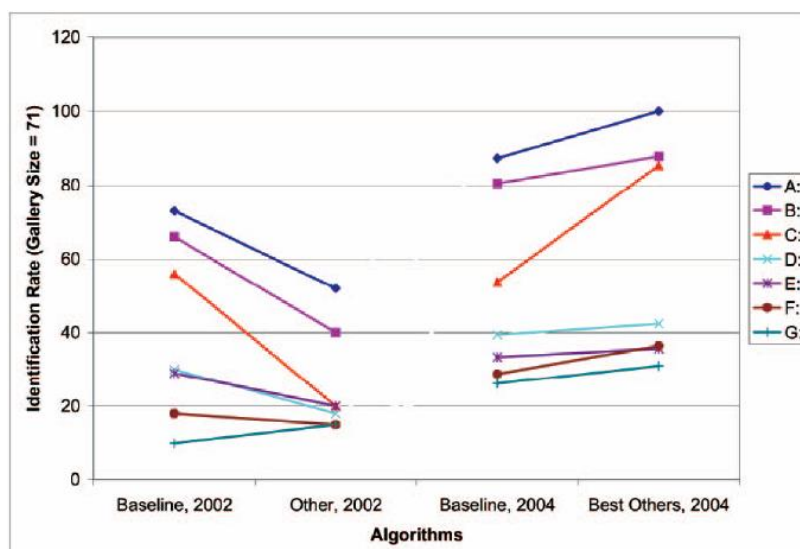
On the other hand, recent gait recognition progress from viewpoints of the number of subjects is reported in (44), where the world's largest gait database of more than 4,000 people has been constructed to enable statistically reliable performance evaluation of gait recognition, although covariates factors are missing. This database is therefore used to measure a sort of upper bound accuracy of the gait recognition under a relatively controlled environment. The database contains view transition from 55 degree (front-oblique) to 85 degree (almost side view) with 10 degree interval, and we show the results by fusion of all the views for eight recent approaches: FDF (25), GEI (27), GEnI (73), GFI (76), Masked GEI (74), CGI (78), GHEI (22) and HoNV (81). Figure 6 shows a receiver operating characteristics (ROC) curve in a verification scenario and a cumulative matching characteristics (CMC) curve in an identification scenario, and furthermore typical evaluation measures such as equal error rates (EER) and rank-1, rank-5, and rank-10 identification rates are summarized in Table 5. It turns out that the most recent approach (i.e. HoNV) achieves the best performance (e.g. 0.58% EER, and 97.5% rank-1 identification rate from 3,141 subjects) for both verification and identification scenarios.
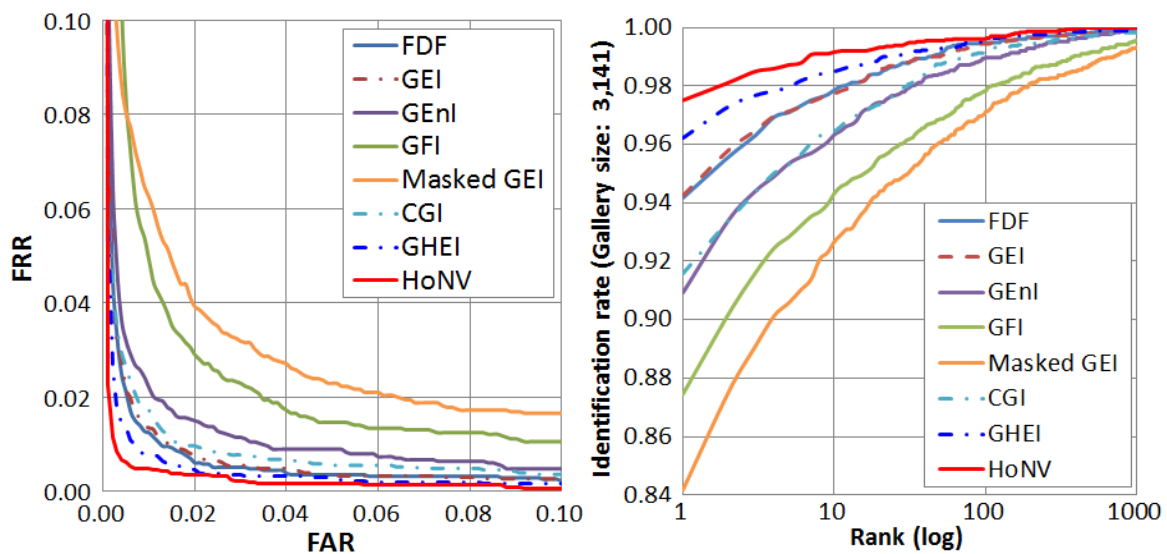

Figure 6 – ROC (left) and CMC (right) curves for recent benchmarks on OU-ISIR, LP.

Approach	GEI	FDF	GEnI	Masked GEI	GFI	CGI	GHEI	HoNV
EER	1.18	1.12	1.59	3.15	2.53	1.29	0.81	**0.58**
Rank-1	94.2	94.2	90.9	84.2	87.5	91.6	96.2	**97.5**
Rank-5	97.1	97.1	95.4	90.6	92.8	95.3	98.0	**98.7**
Rank-10	97.7	97.9	96.3	92.7	94.3	96.5	98.5	**99.2**

Table 5 – EER and identification rates [%] for recent benchmarks on OU-ISIR, LP.

As a general topic, it is important for fair comparison of approaches to gait recognition, to keep experimental protocols on the same dataset as well as the consistency of performance evaluation

measures, although some of previously published papers might have reported the experimental results with their own measures. In addition, a recent study on relating ROC and CMC via the biometric menagerie (111) concludes that the same ROC may produce different CMCs and that the difference is more prominent in soft biometrics such as gait, and hence it is highly recommended for gait recognition researchers to report both ROC in verification scenarios and CMC in identification scenarios and pick up some representative measures from them (e.g. EER from ROC, rank-n identification rates from CMC).

Application

Gait research is currently at an evaluation stage rather than an application stage. However, the potential for gait recognition is great. The complete unobtrusiveness without any subject cooperation or contact for data acquisition makes gait particularly attractive for identification purposes. It could be used in applications including forensics, security, immigration, surveillance.

Many surveillance systems capture only a low resolution video at varying lighting conditions, and gait recognition might be the only plausible choice for automatic recognition. A bank robber may wear a mask so you cannot see his face, wear gloves so you cannot get fingerprints, wear a hat so you cannot get DNA evidence – but they have to walk or run into the bank and they could be identified from their gait.

Gait recognition has been used as evidence for conviction in some criminal cases. One forensic study has already used gait biometrics to provide evidence for identification (6). In 2004, a perpetrator robbed a bank in Denmark. The Institute of Forensic Medicine in Copenhagen was contacted by the police to perform gait analysis, as they thought the perpetrator had a unique gait. The institute instructed the police to establish a covert recording of the suspect from the same angles as the surveillance recordings for comparison. The gait analysis revealed several characteristic matches between the perpetrator and the suspect. For example, both the perpetrator (to the left) and the suspect showed inverted left ankle (white arrow) during left leg's stance phase and markedly outward rotated feet. (see Fig. 7) The suspect was convicted of robbery and the court found that gait analysis is a very valuable tool (104).

Figure 7 - Bank Robbery (Larsen et al. 2008 (104))

Moreover, a gait verification system for this criminal investigation has been developed in (7). The system is equipped with a graphical user interface (see Fig. 8) so as that a criminal investigator, who may not be a researcher on gait recognition, can obtain gait-based person verification results through appropriate manual interventions: target specification; interactive silhouette extraction; and undesired region masking (e.g., helmet) if necessary. The system outputs a posterior distribution that a pair of walking image sequences are generated from the same person based on circumstance-dependent probability distribution functions of dissimilarity scores from the same person's pairs and different persons' pairs. The system is now under a trial use phase by National Research Institute of Police Science in Japan.
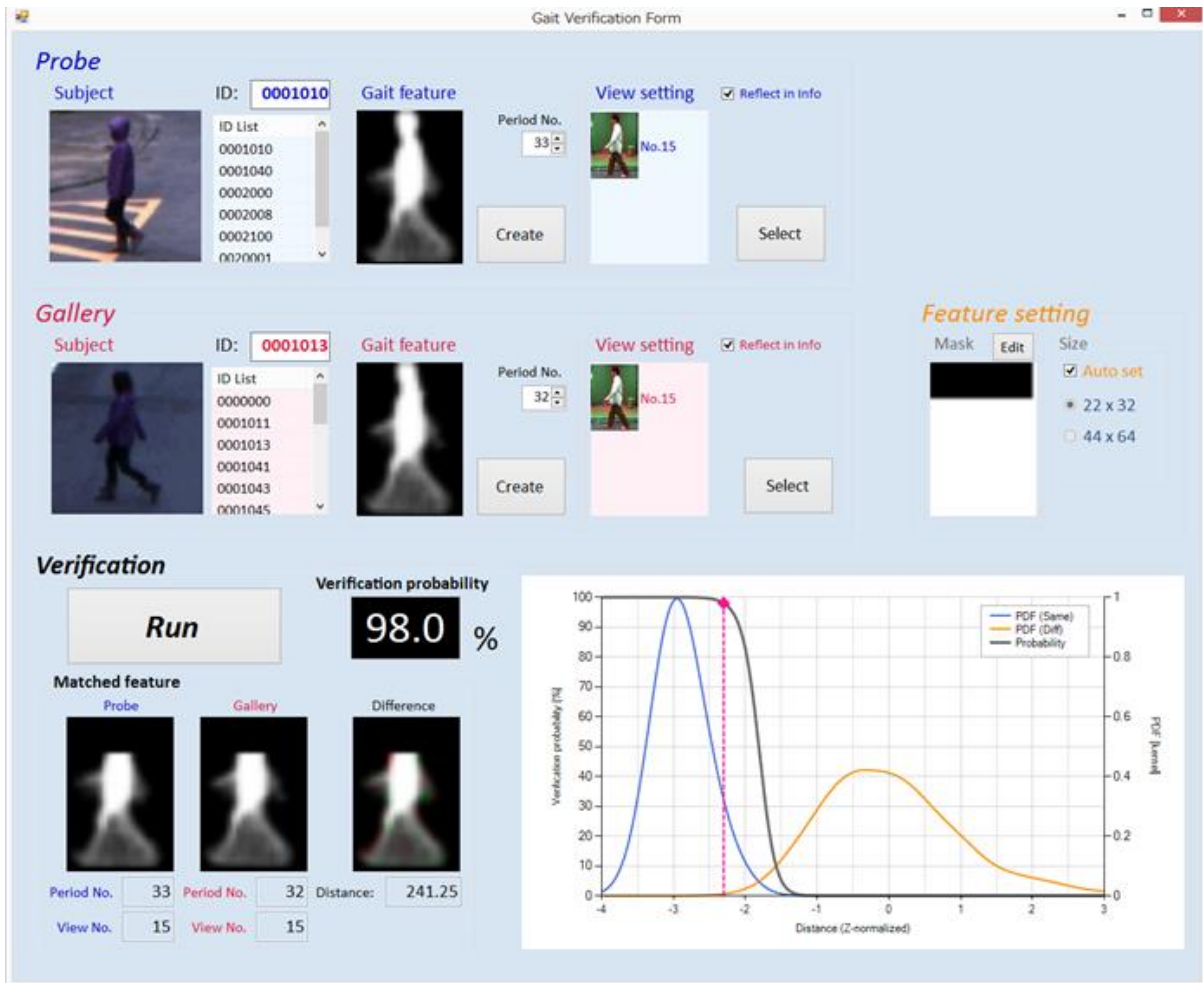
Figure 8 – Gait verification system

One system named the Biometric Tunnel (48) has led to the first live demonstration of gait as a biometric and could indicate a possible route for future deployment of the technology. The left side of Fig. 9 depicts the system. It consists of a simple corridor with 12 synchronised and fixed cameras. The subjects are asked to walk through the middle and the lighting and background are controlled to facilitate analysis. The right side of Fig. 9 shows the details of the arrangement. The system is designed with a high throughput environment in mind.
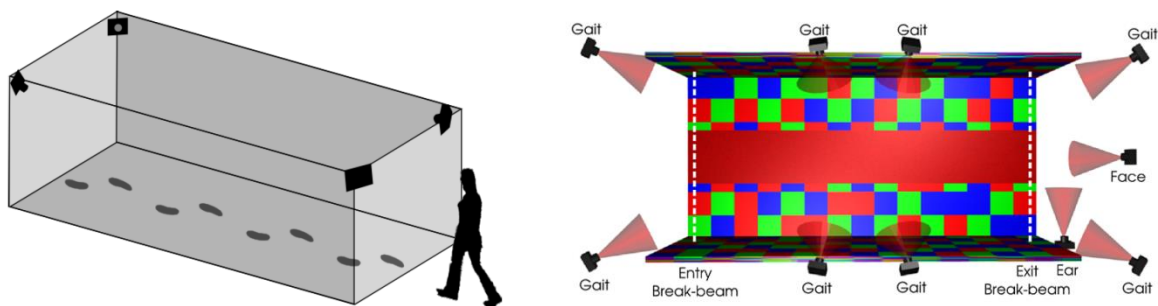

Figure 9 - The Biometric Tunnel

Open Problems

Although a large number of gait recognition algorithms have been reported, it is important to note that gait biometrics is still in its infancy. The majority of studies achieve good recognition rates on gallery and probe acquired in similar conditions. However, it is very challenging to extract gait features that are invariant to change in appearance as well as to conditions that affect a person's gait. Examples of things that can change and negatively affect the effectiveness of current gait algorithms are: change of clothing, shoe type, carrying a load and injuries/medical conditions. Clothing for instance can change the observed pattern of motion and make it difficult to accurately locate joint position. Furthermore, there are certain factors that are related to the environment and not the subjects themselves that can cause difficulties for current gait matchers. Examples of environmental confounding factors are: camera viewing angle, background and illumination.

Recently, studies have reported progress in solving some of the issues outlined above. A study by Goffredo et al. (105) describes a self-calibrating view-invariant gait recognition algorithm, while studies (25, 106) introduce training-based approaches in a framework of a view transformation model. Studies (41,74) have developed a clothing invariant gait matcher, while studies (17, 40, 108) are the most recent approaches focussed on speed variation.

In addition, while the most of the recent gait recognition approaches exploit model-free features, in particular, silhouette-based features such as GEI (27) due to its suitability to outdoor application, note that the success of the silhouette-based approaches heavily rely on the silhouette quality and that the silhouette quality may be degraded under outdoor scenes with challenging environmental factors including but not limited to illumination changes, dark illumination, cluttered background, motion blur, and image compression noise.

A few studies discuss the effect of the silhouette quality on the gait recognition accuracy. For example, Liu and Sarkar (18) proposed a method of silhouette cleaning based on population-HMM and concluded that the silhouette quality does not affect the gait recognition accuracy in a consistent way. Moreover, since GEI (27) is computed by aggregating the pixel-wise silhouette value over one gait cycle, it is claimed that the variance of pixel-independent zero-mean noise (i.e. expectation of the numbers of over-segmentation and under-segmentation per pixel over one gait cycle are the same) is inversely proportional to the number of images in one gait cycle based on statistics theory, i.e. the effect of the silhouette quality on GEI is limited.

On the other hand, it is obvious that the non zero-mean noise affects the silhouette-based features (e.g. GEI is significantly degraded if leg region silhouettes are always missing over one gait period by under-segmentation). In fact, Matovski et al. (112) demonstrate that the silhouette quality plays an important role in an applied automatic gait recognition by proposing a silhouette quality-dependent sequential verification procedure. Another direction is to improve the silhouette quality (i.e. silhouette refinement) considering spatial adjacency. For example, graph-cut segmentation (19) is a promising choice for better silhouette extraction and it is exploited for silhouette extraction in

conjunction with background subtraction and background update by colour transformation in (20). Moreover, since the observed colour at the boundary of the subject and the background is essentially a mixture of them, alpha matting (21) is considered for more accurate silhouette representation in (22).

Currently, gait recognition can deliver very high recognition rates in a constrained environment and if certain factors are controlled. A move towards developing algorithms invariant to change over time is needed. Furthermore, additional work is required to translate the research to outside environment and to explore how scalable it is. Attempts so far suggest that developing highly reliable gait-based human identification system in a real world application is, and will continue to be, very challenging. In the short term, some of the challenges associated with gait recognition can be addressed by fusing gait with other biometrics (109, 110).

List of Abbreviation

CASIA: The Institute of Automation of the Chinese Academy of Sciences
CCTV: Closed-circuit television
CGI: Chrono-gait image
CHLAC: Cubic higher-order local auto-correlation
CMC: Cumulative matching characteristics
CMU: Carnegie Mellon University
DGHEI: Depth gradient histogram energy image
DNA: Deoxyribonucleic acid
EER: Equal error rate
FAR: False acceptance rate
FDF: Frequency-domain feature
FRR: False rejection rate
GAID: Gait from audio, image and depth
GEI: Gait energy image
GEnI: Gait Entropy Image
GEV: Gait energy volume
GFI: Gait flow image
GHEI: Gradient histogram energy image
GMD: Gait motion descriptor
HMM: Hidden Markov model
HOG: Histogram of oriented gradient
HoNV: Histogram of normal vector
ID: Identification
LDA: Linear discriminant analysis
LP: Large population
MSI: Motion silhouette image
OU-ISIR: Osaka University, The Institute of Scientific and Industrial Research
PCA: Principal component analysis
RGB: Red green blue
ROC: Receiver operating characteristics
SoPF: Statistics of probability function

SOTON: The University of Southampton
SVB: Shape variation-based
SVM: Support vector machine
TUM: Technical University of Munich
UMD: The University of Maryland
USF: University of Southern California
WOSG: West Virginia University Outdoor SWIR Gait

Recommended Reading List

(1) Murray, M., Drought, A., and Kory, R., The journal of bone and joint surgery **46**(2)**,** 335, 1964.

(2) Winter, D.A., Biomechanics and Motor Control of Human Movement Ottawa ON, Canada, John Wiley & Sons, 2009.

(3) Cutting, J.E. and Kozlowski, L.T., Bulletin of Psychonomic Society **9**(5)**,** 353-356, 1977.

(4) Kozlowski, L.T. and Cutting, J.E., Perception & Psychophysics **21**(6)**,** 575-580, 1977.

(5) Troje, N.F., Journal of Vision **2**, 371-387, 2002.

(6) Bouchrika, I., Goffredo, M., Carter, J.N., and Nixon, M.S., Journal of Forensic Sciences **56**(4), 882-889, 2011.

(7) Iwama, H., Muramatsu, D., Makihara, Y., and Yagi, Y., IPSJ Trans. on Computer Vision and Applications **5**, 163-175, 2013.

(8) Lynnerup, N. and Larsen, P.K., IET Biometrics **3**(2), 47-54, 2014.

(9) Yoo, J., Hwang, D., and Nixon, M.S., Proc. of the 15th Int. Conf. on Advanced Concepts For Intelligent Vision Systems, 138-145, 2006.

(10) Huang, G. and Wang, Y., Proc. of the 8th Asian Conf. on Computer Vision **1**, 462-471, 2007.

(11) Li, X., Maybank, S.J., Yan, S., Tao, D., and Xu, D., IEEE Trans. on Systems, Man, and Cybernetics, Part C **38**(2), 145-155, 2008.

(12) Yu, S., Tan, T., Huang, K., Jia, K., and Wu, X., IEEE Trans. on Image Processing **18**(8), 1905-1910, 2009.

(13) Davis, J.W., Proc. of the 3rd Int. Conf. on Audio- and Video-based Biometric Person Authentication, 295-300, 2001.

(14) Begg, R.K., IEEE Trans. on Biomedical Engineering. **52**(5), 828-838, 2005.

(15) Lu, J. and Tan, Y.-P., IEEE Trans. on Information Forensics and Security **5**(4), 761-770, 2010.

(16) Makihara, Y., Okumura, M., Iwama, H., and Yagi, Y., Proc. of the 1st Int. Joint Conf. on Biometrics, 1-6, 2011.

(17) Livne, M., Sigal, L., Troje, N.F., and Fleet. D.J., Computer Vision and Image Understanding **116**(5), 648-660, 2012.

(18) Liu, Z. and Sarkar, S., IEEE Trans. of Systems, Man, and Cybernetics Part B: Cybernetics **35**(2), 170-183, 2005.

(19) Boykov, Y. and Funka-Lea, G., Int. Journal on Computer Vision **70**(2), 109-131, 2006.

(20) Makihara, Y. and Yagi, Y., Proc. of the 19th Int. Conf. on Pattern Recognition, 1-4, 2008.

(21) Levin, A., Lischinski, D., and Weiss, Y., IEEE Trans. on Pattern Analysis and Machine Intelligence **30**(2), 228-242, 2008.

(22) Hofmann, M. and Rigoll, G., Proc. of the 19th IEEE Int. Conf. on Image Processing, 1389-1392, 2012.

(23) Sarkar, S., Phillips, P.J., Liu, Z., Vega, I.R., Grother, P., and Bowyer, K.W., IEEE Trans. on Pattern Analysis and Machine Intelligence **27**(2)**,** 162-177, 2005.

(24) Wang, L., Tan, T., Ning, H. & Hu, W., IEEE Trans. on Pattern Analysis and Machine Intelligence **25**(12), 1505-1518, 2003.

(25) Makihara, Y., Sagawa, R., Mukaigawa, Y., Echigo, T., and Yagi, Y., Proc. of the 9th European Conf. on Computer Vision **3**, 151-163, 2006.

(26) Makihara, Y., Aqmar, M.R., Ngo, T.T., Nagahara, H., Sagawa, R., Mukaigawa, Y., and Yagi, Y., IEEE Trans. on Signal Processing **62**(8), 2066-2079, 2014.

(27) Han, J. and Bhanu, B., IEEE Trans. on Pattern Analysis and Machine Intelligence **28**(2), 316-322, 2006.

(28) Chapelle, O. and Keerthi, S., Information Retrieval **13**, 201-215, 2010.

(29) Martin-Felez, R. and Xiang, T., Proc. of the 12th European conf. on Computer Vision **1**, 328-341, 2010.

(30) Gross, R. and Shi, J, Tech. report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon Univ, 2001.

(31) Tanawongsuwan, R. and Bobick, A., Tech. report GIT-GVU-03-01, Georgia Institute of Tech, 2003.

(32) Chalidabhongse, T., Kruger, V., and Chellappa, R., Tech. report, Univ. of Maryland, 2001.

(33) Nixon, M.S., Carter, J.N., Shutler, J., and Grant., M., Tech. report, Univ. of Southampton, 2001.

(34) Shutler, J.D., Grant, M.G., Nixon M.S., and Carter J.N., Proc. of the 4th Int. Conf. on Recent Advances in Soft Computing, 66-71, 2002.

(35) Samangooei, S., Bustard, J., Nixon, M.S. and Carter, J.N., Multibiometrics for Human Identification; Bhanu, B and Govindaraju, V. , Cambridge University Press, 277-301, 2011.

(36) Matovski, D.S., Nixon, M.S., Mahmoodi, S., and Carter, J.N., Proc. of the 4th IEEE Int. Conf. on Biometrics: Theory, Applications and Systems, 2012.

(37) Wang, L., Ning, H., Hu, W., and Tan, T., Proc. of the 19th IEEE Int. Conf. on Image Processing **3**, 433-436, 2002.

(38) Yu, S., Tan, D., and Tan, T., Proc. of the 18th Int. Conf. on Pattern Recognition **4**, 441-444, 2006.

(39) Tan, D., Huang, K., Yu, S., and Tan, T., Proc. of the 18th Int. Conf. on Pattern Recognition **3**, 1000-1003, 2006.

(40) Tsuji, A., Makihara, Y., and Yagi, Y., Proc. of the 23rd IEEE Conf. on Computer Vision and Pattern Recognition, 717-722, 2010.

(41) Hossain, M.A., Makihara, Y., Junqiu, W., and Yagi, Y., Pattern Recognition **43**(6)**,** 2281-2291, 2010.

(42) Makihara, Y., Mannami, H., and Yagi, Y. Proc. of the 10th Asian. Conf. on Computer Vision, 975-986, 2010.

(43) Mori, A., Makihara, Y., and Yagi, Y., Proc. of the 20th Int. Conf. on Pattern Recognition, 2194-2197, 2010.

(44) Iwama, H., Okumura, M., Makihara, Y., and Yagi, Y., IEEE Trans. on Information Forensics and Security **7**(5), 1511-1521, 2012.

(45) Hofmann, M., Sural, S., and Rigoll, G., Proc. of the 19th Int. Conf. on Computer Graphics, Visualization and Computer Vision, 2011.

(46) Hofmann, M. Geiger, J., Bachmann, S., Schuller, B., and Rigoll, G., Journal of Visual Communication and Image Representation, Special Issue on Visual Understanding and Applications with RGB-D Cameras **25**(1), 195-206, 2014.

(47) DeCann, B., Ross, A., and Dawson, J., Proc. of SPIE Conf. on Biometric Technology for Human Identification X, 1-16, 2013.

(48) Seely, R., Samangooei, S., Lee, M., Carter, J., and Nixon, M., Proc. of the 2nd IEEE Int. Conf. Biometrics: Theory, Applications and Systems, 2008.

(49) Makihara, Y., Mannami, H., Tsuji, A., Hossain, M.A., Sugiura, K., Mori, A., and Yagi, Y., IPSJ Trans. on Computer Vision and Applications **4**, 53-62, 2012.

(50) Niyogi, S.A. and Adelson, E.H., Proc. of the 7th IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 469-474, 1994.

(51) Murase, H. and Sakai, R., Pattern Recognition Letter **17**, 155-162, 1996.

(52) Little, J. and Boyd, J., Videre: Journal of Computer Vision Research **1**(2)**,** 1-32, 1998.

(53) Huang, P.S., Harris, C.J., and Nixon, M.S., Artifitial Intelligence Engineering **13**(4), 359-366, 1999.

(54) Shutler, J.D. and Nixon, M.S., Proc. of the 12th Brithish Machine Vision Conf., 705-714, 2001.

(55) BenAbdelkader, C., Cutler, R., and Davis, L., Proc. of the 3rd Int. Conf. on Audio- and Video-Based Biometric Person Authentication, 284-294, 2001.

(56) Collins, R., Gross, R., and Shi, J., Proc. of the 5th IEEE Int. Conf. Face and Gesture Recognition, 366-371, 2002.

(57) Liu, Y., Collins, R., and Tsin, Y., Proc. of the 7th European Conf. on Computer Vision, 657-671, 2002.

(58) Vega, I.R. and Sarkar, S., IEEE Trans. on Pattern Analisys and Machine Intelligence **25**(10), 1323-1328, 2003.

(59) Foster, J.P., Nixon, M.S., and Prugel-Bennet, A., Pattern Recognition Letter **24**, 2489-2497, 2003.

(60) Hayfron-Acquah, J., Nixon, M.S., and Carter, J.N., Pattern Recognition Letters **24**(13)**,** 2175-2183, 2003.

(61) Zhao, G., Chen, R., Liu, G., Hua, L., Proc. of the 6th IEEE Int. Conf. on Face and Gesture Recognition, 23-28, 2004.

(62) Kale, A., Sundaresan, A., Rajagopalan, A., Cuntoor, N., Roy-Chowdhury, A., Kruger, V., and Chellappa, R., IEEE Trans. on Image Processing **13**(9)**,** 1163-1173, 2004.

(63) Liu, Z. and Sarkar, S., Proc. of the 17th Int. Conf. on Pattern Recognition, 211-214, 2004.

(64) Lee, C.-S. and Elgammal, A., Proc. of the 6th IEEE Int. Conf. on Face and Gesture Recognition, 147-152, 2004.

(65) Kobayashi, T. and Otsu, N., Proc. of the 17th Int. Conf. on Pattern Recognition **3**, 741-744, 2004.

(66) Mowbray, S.D. and Nixon, M.S., Proc. of the 1st IEEE Int. Conf. on Advanced Video and Signal Based Surveillance, 566-573, 2003.

(67) Lam, T.H.W. and Lee, R.S.T., Proc. of the 1st IAPR Int. Conf. on Biometrics, 612-618, 2006.

(68) Tao, D., Li, X., Wu, X., and Maybank, S., Proc. of the 18th IEEE Conf. on Computer Vision and Pattern Recognition **2**, 1670-1677, 2006.

(69) Ioannidis, D., Tzovaras, D., Damousis, I.G., Argyropoulos, S., and Moustakas, K., IEEE Trans. on Information Forensics and Security **2**(3), 623-630, 2007.

(70) Lam, T.H.W., Lee, R.S.T., and Zhang, D., Pattern Recognition **40**(9), 2563-2573, 2007.

(71) Lee, S., Liu, Y., and Collins, R., Proc. of the 19th IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1-8, 2007.

(72) Bashir, K., Xiang, T., and Gong, S., Proc. of the 20th British Machine Vision Conf., 1-11, 2009.

(73) Bashir, K., Xiang, T., and Gong, S., Proc. of the 3rd Int. Conf. on Imaging for Crime Detection and Prevention, 2009.

(74) Bashir, K., Xiang, T., and Gong, S., Pattern Recognition Letters **31**(13), 2052-2060, 2010.

(75) DeCann, B. and Ross, A., Proc. of the SPIE, Biometric Technology for Human Identification VII. 76670Q-76670Q-13, 2010.

(76) Lam, T.H.W., Cheung, K.H., and Liu, J.N.K., Pattern Recognition **44**(4), 973-987, 2011.

(77) Sivapalan, S., Chen, D., Denman, S., Sridharan, S., and Fookes, C., Prof. of the 1st Int. Joint Conf. on Biometrics, 1-6, 2011.

(78) Wang, C., Zhang, J., Pu, J., Wang, L., and Yuan, X., IEEE Trans. on Pattern Analysis and Machine Intelligence **23**(11), 2164-2176, 2012.

(79) Hofmann, M., Bachmann, S., and Rigoll, G., Proc. of the 5th IEEE Int. Conf. on Biometrics Theory, Applications, and Systems, Paper ID 172, 1-5, 2012.

(80) Nakajima, H., Mitsugami, I., and Yagi, Y. IPSJ Trans. on Computer Vision and Applications **5**, 94-98, 2013.

(81) El-Alfy, H., Mitsugami, I., and Yagi, Y., Proc. of ACCV 2014 workshop on Human Gait and Action Analysis in the Wild: Challenges and Applications, 2014.

(82) Cunado, D., Nixon, M.S., and Carter, J.N., Computer Vision and Image Understanding **90**(1), 1-41, 2003.

(83) Lee, L. and Grimson, W.E.L., Proc. of the 5th IEEE Int. Conf. on Face and Gesture Recognition, 155-162, 2002.

(84) Bhanu, B. and Han, J., Proc. of the 4th Int. Conf. Audio- and Video-Based Biometric Person Authentication, 600-608, 2003.

(85) Benabdelkader, C., Cutler, R., and Davis, L., Proc. of the 5th IEEE Int. Conf. on Automatic Face and Gesture Recognition, 267-272, 2002.

(86) Bobick, A.F. and Johnson, A.Y., Proc. of the 14th IEEE Computer Society Conf. on Computer Vision and Pattern Recognition **1**, 423-430, 2001.

(87)   Wang, L., Ning, H., Tan, T., and Hu, W., IEEE Trans. on Circuits and Systems for Video Technology **14**(2)**,** 149-158, 2004.

(88)   Wagg, D.K. and Nixon, M.S., Proc. of the 6th IEEE Int. Conf. on Face and Gesture Recognition, 11-16, 2004.

(89)   Yam, C., Nixon, M.S., and Carter, J.N., Pattern Recognition **37**(5)**,** 1057-1072, 2004.

(90)   Zhang, R., Vogler, C., and Metaxas, D., Proc. of the 17th IEEE Computer Society Conf. on Computer Vision and Pattern Recognition **2**, 342-349, 2004.

(91)   Zhao, G., Liu, G., Li, H., and Pietikainen, M., Proc. of the 7th Int. Conf. on Automatic Face and Gesture Recognition, 529-534, 2006.

(92)   Lu, H., Plataniotis, K.N., and Venetsanopoulos, A.N., EURASIP Journal Advanced Signal Process **2008**(62), 1-13, 2008.

(93)   Ariyanto, G. and Nixon, M.S., Proc. of the 5th IAPR Int. Conf. on Biometrics, 354-359, 2012.

(94)   Nixon, M.S., Tan, T.N., and Chellappa, R., Human Identification Based on Gait, New Yorg, Springer-Verlag, 2005.

(95)   Nixon, M.S. and Carter, J.N., Proc. of the IEEE **94**(11)**,** 2013-2024, 2006.

(96)   Murase, H. and Nayar, S.K., Int. Journal on Computer Vision **14**(1), 5-24, 1995.

(97)   Liang, W., Tieniu, T., Huazhong, N., and Weiming, H., IEEE Trans. on Pattern Analysis and Machine Intelligence **25**(12)**,** 1505-1518, 2003.

(98)   Sundaresan, A., Roychowdhury, A., and Chellappa, R., Proc. of the 10th IEEE Int. Conf. on Image Processing, 93-96, 2003.

(99)   Wang, L., Tan, T.N., Hu, W.M., Ning, H.Z., IEEE Trans. on Image Processing **12**(9), 1120-1131, 2003.

(100)   Dalal, N. and Triggs, B., Proc. of the 18th IEEE Computer Society Conf. on Computer Vision and Pattern Recognition **1**, 886-893, 2005.

(101)   Bouchrika, I. and Nixon, M., Proc. of the 8th IEEE Int. Conf. on Automatic Face and Gesture Recognition, 1-6, 2008.

(102)   Urtasun, R. and Fua, P., Proc. of the 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition, 17-22, 2004.

(103)   Guoying, Z., Guoyi, L., Hua, L., and Pietikainen, M., Proc. of the 7th IEEE Int. Conf. Automatic Face and Gesture Recognition, 529-534, 2006.

(104)   Larsen, P.K., Simonsen, E.B., and Lynnerup, N., Journal of Forensic Sciences **53**(5)**,** 1149-1153, 2008.

(105)   Goffredo, M., Bouchrika, I., Carter, J.N., and Nixon, M.S., IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics **40**(4)**,** 997-1008, 2010.

(106)   Kusakunniran, W., Wu, Q., Zhang, J., Li, H., and Wang, L. IEEE Trans. on Image Processing, **23**(2), 696-709, 2014.

(107)   Aqmar, M.R., Shinoda, K., and Furui, S., Proc. of the 20th IEEE Int. Conf. on Pattern Recognition, 2190-2193, 2010.

(108)   Guan, Y. and Li, C.-T., Proc. of the 6th IAPR Int. Conf. on Biometrics, 1-8, 2013.

(109)   Zhou, X. and Bhanu, B., Pattern Recognition **41**(3)**,** 778-795, 2008.

(110) Muramatsu, D., Iwama, H., Makihara, Y., and Yagi, Y., Proc. of the 6th IAPR Int. Conf. on Biometrics, 1-8, 2013.

(111) DeCann, B. and Ross, A., Proc. of the 6th Int. Conf. on Biometrics: Theory, Applications and Systems, 1-8, 2013.

(112) Matovski, D.S., Nixon, M.; Mahmoodi, S. and Mansfield, T., Proc. of the 21th Int. Conf. on Pattern Recognition, 3272-3275, 2012

(113) Stevenage, S., Nixon, M., Vince, K., Applied Cognitive Psychology 13, 513-526,1999