Biometric recognition by gait: A survey of modalities and features

Patrick Connor [*,a], Arun Ross [b]

[a] Stepscan Technologies Inc., 14 MacAleer Drive (Suite 1), Charlottetown, Prince Edward Island C1E 2A1, Canada
[b] Department of Computer Science and Engineering, Michigan State University, 428 South Shaw Lane, 3142 Engineering Building, East Lansing, MI 48824, USA

A B S T R A C T

The scientific literature on automated gait analysis for human recognition has grown dramatically over the past 15 years. A number of sensing *modalities* including those based on vision, sound, pressure, and accelerometry have been used to capture gait information. For each of these modalities, a number of *methods* have been developed to extract and compare human gait information, resulting in different sets of features. This paper provides an extensive overview of the various types of features that have been utilized for each sensing modality and their relationship to the appearance and biomechanics of gait. The features considered in this work include (a) static and dynamic (temporal) features; (b) model-based and model-free visual features; (c) ground reaction force-based and finely resolved underfoot pressure features; (d) wearable sensor features; and (e) acoustic features. We also review the factors that impact gait recognition, and discuss recent work on gait spoofing and obfuscation. Finally, we enumerate the challenges and open problems in the field of gait recognition.

1. Introduction

High'st queen of state, Great Juno, comes; I know her by her gait.
The Tempest (Act 4, Scene 1)

Gait is one of many physical and behavioral traits of an individual that can be used for recognizing a person. Gait refers to a person's manner of walking. A number of studies have established the potential of using gait information to distinguish between individuals. Gait has even been used in criminal cases to identify perpetrators based on their walking behavior. Besides recognizing people, gait patterns such as the Parkinsonian shuffle can be used to help identify and assess pathological conditions. Gait recognition has been defined as (Boyd and Little, 2005) "... the recognition of some salient property, e.g., identity, style of walk, or pathology, based on the coordinated, cyclic motions that result in human locomotion." The timeline and taxonomy depicted in Fig. 1 indicate the use of gait recognition in three main application areas: gait analysis, gait forensics, and gait biometrics. As also highlighted in this figure, the introduction of new enabling technologies has significantly advanced the objective study of gait.

Gait recognition and analysis has a rich history, and in recent years is being considered as a biometric recognition tool. Aristotle was the first person to take note of human and animal gaits, as recorded in his work "On the Gait of Animals" (Aristotle, 350 BC), written in 350 BCE. The study of locomotion was later taken up by Borelli, considered to be a father of biomechanics. He related animals to machines using

mathematics in his posthumous work (Borelli, 1679) that bears the same title as Aristotle's treatise. Wilhelm and Eduard Weber are recognized for their work, "Mechanics of the human walking apparatus" (Weber and Weber, 1836), where they proposed a number of hypotheses about human locomotion (some correct, some incorrect), including the pendulum-like behavior of the forward leg motion. Later that century, physiologist and inventor Etienne-Jules Marey of France developed cameras that could take multiple distinct photos per second exposed on the same negative, referred to as chronophotography. With this, Marey captured a variety of human and animal movements. In the same period, Eadweard Muybridge in the United States used a bank of 12 cameras to famously prove in 1878 that there is a point at which all of a horse's hooves are off the ground during a gallop. He also explored human gait and other human movements with his equipment.

Gait analysis and enabling technologies made great strides in the latter half of the 20th century. After World War II, Verne Inman and Howard Eberhart began their collaboration and later helped establish the Prosthetic Device Research Project at UC Berkeley, where they performed basic research on the biomechanics of human locomotion. Notably, their research led to a seminal paper in gait analysis offering a unifying theory of gait, which proposed that the body moves its center of gravity in a way that minimizes energy expenditure, as determined by specific joint movements (Saunders et al., 1953). Jacqueline Perry and David Sutherland contributed substantially to the developing field of clinical gait analysis, evolving the model of the gait cycle and separately setting up important laboratories in 1968 and 1974. Around
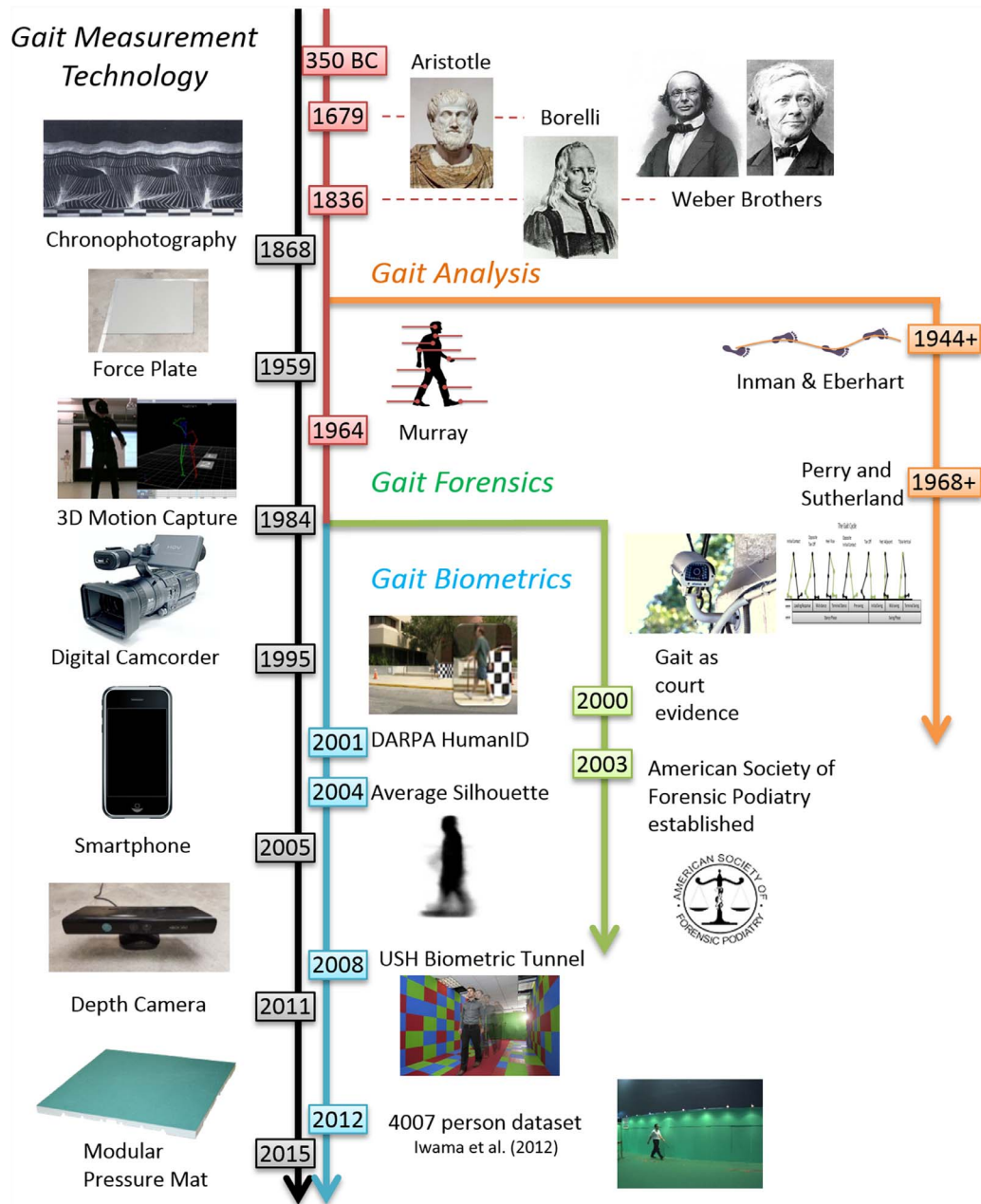
Fig. 1. A taxonomy and timeline of gait recognition. On the left side is shown a list of enabling technologies. On the right, gait recognition and its advancements over time are shown, dividing into three main areas: gait analysis, gait forensics, and gait biometrics. Some of the images used in this and other figures are from the Internet. URLs are provided in Appendix 1.

this time and afterward, video and force platforms to capture gait in its visual and force delivery forms, respectively, were being developed. The 3-axis force sensor underlying Kistler force plates was launched in 1969. It was used to capture the total XY and vertical forces delivered by the foot while walking. In 1984, the well-known commercial 3D motion capture technology business Vicon was founded. They introduced a video-based motion capture system that reflected infrared light off passive targets attached to key anatomical locations on the body. The reflected images were captured by an array of surrounding cameras to accurately localize the 3D positions of the points.

Gait as a means of discriminating between individuals was studied in the late 1960s and 70s. As part of Patricia Murray's gait analysis research, she studied normal gait using chronophotography (Murray, 1967; Murray et al., 1964). She identified features of normal gait that were consistent within an individual, but that varied between individuals. Adding to this, studies such as by Johansson (1973),

Cutting and Kozlowski (1977) and Cutting et al. (1978) demonstrated the *human ability* to recognize individuals by their gait expressed using point light displays. Such studies, beyond subjective experience and Shakespearean anecdotes, provide the basis for forensic and biometric gait recognition because they objectively affirm that *gait is sufficiently stable for a normal individual and distinctive between normal individuals.*

Gait has been used as forensic evidence in court. Videos of gait at crime scenes have been compared with gait videos of known suspects in cases in the UK and elsewhere. Notably, Hadyn Kelly of London, England, was perhaps the first to offer expert gait analysis testimony in court for a robbery case in 2000 (Gui, 2016). At the time of writing this paper, he runs a consultancy and training business called GaitForensics that specializes in this area. Forensic podiatry is a broader discipline that includes gait forensics. The American Society of Forensic Podiatry was established in 2003, and its President John DiMaggio in a 2014 interview (Daugherty, 2014) estimated that members of the society are
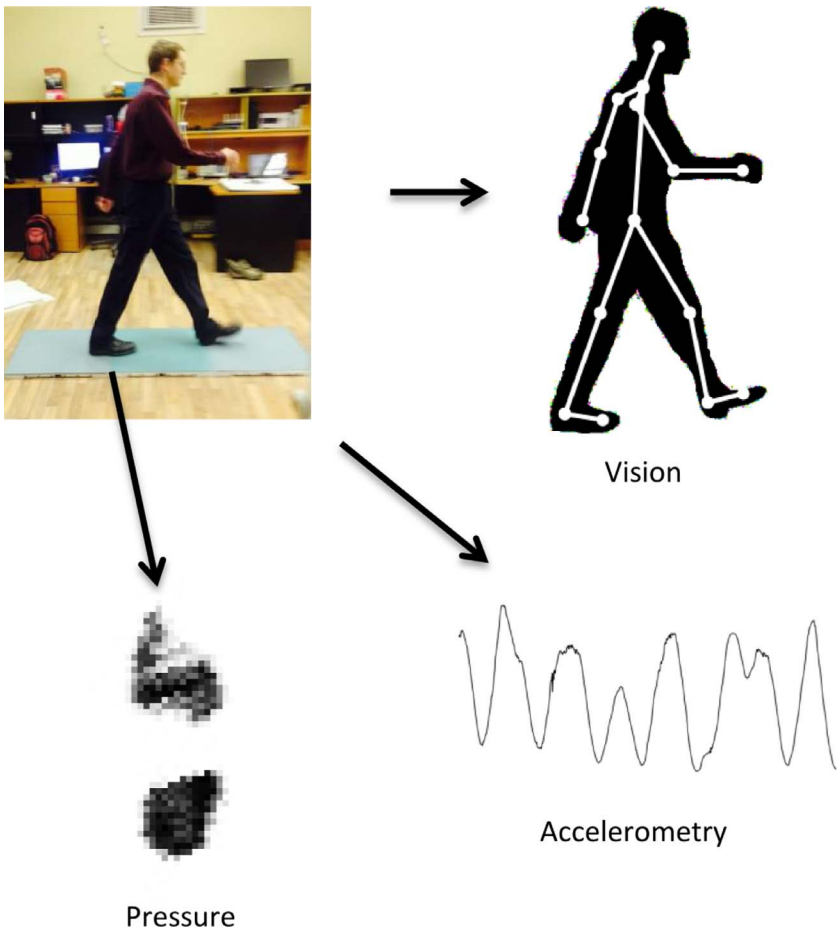
Fig. 2. Examples of sensor modalities used for gait recognition. Gait features can be extracted from (a) a sequence of visual images or video (often transformed into a silhouette or skeletal structure), (b) an underfoot pressure image sequence obtained using a pressure mat sensor, (c) an acceleration trace recorded by an accelerometer in a mobile or wearable device, or (d) an audio recording.

contacted on 30 to 40 cases per year, and that "he expects that number to grow as more police departments learn about the discipline".

Gait as a biometric cue began first with video-based analysis (Niyogi and Adelson, 1994) and, shortly afterward, by analysis of underfoot pressures (Addlesee et al., 1997). A more recent addition is the use of accelerometer data (Mäntyjärvi et al., 2005) from mobile devices such as smartphones. These three different sensory modalities for gait recognition are illustrated in Fig. 2. Table 1 outlines state-of-the-art features pertaining to each sensory modality. Early studies had few subjects, but provided proof of concept. In the early 2000s, at a time when digital video recording technology was becoming more accessible and advanced, DARPA established the HumanID at a distance research program, which boosted research in vision-based gait recognition. An important outcome was the creation of the HumanID gait challenge problem datasets (Phillips et al., 2002; Sarkar et al., 2005). The largest of these was collected at the University of South Florida (Tampa) and offered significant numbers of subjects with realistic environmental or walking condition variations (e.g., footwear, carrying briefcase, walking surface, elapsed time). It is the most commonly used dataset in the gait recognition literature to date. In vision-based gait recognition, an important observation made in 2004 was that the average of a person's silhouettes (centered within the image) from a video sequence is an effective way of representing their gait for recognition purposes (Han and Bhanu, 2004, 2006; Liu and Sarkar, 2004). It has since become a standard method against which new vision-based approaches are commonly compared. In 2006, the first book on vision-based biometric gait recognition was published by Mark Nixon (University of Southampton), Tieniu Tan (Chinese Academy of Sciences), and Rama Chellappa (University of Maryland), all researchers that have published substantially in the area. Shortly afterward, Mark Nixon established the University of Southampton's "biometric tunnel" (Seely et al., 2008) to simulate a controlled high-throughput environment such as an airport tunnel, to evaluate automatic gait, face, and ear recognition approaches. In terms of recognition accuracy, early studies on vision-

Table 1
State-of-the-art features by sensory modality.

Sensory odality	Features	References
Vision	Average silhouette or GEI	Han and Bhanu (2004, 2006)
	Histogram of gradient approaches	Liu et al. (2012), Hofmann and Rigoll (2012) and Hofmann et al. (2012)
	Joint position and trajectory-based	Goffredo et al. (2010), Świtoński et al. (2011a) and Yu and Zou (2012)
Underfoot	Wavelet packet features on GRF	Yao et al. (2010), Moustakidis et al. (2008) and Mason et al. (2016)
Pressures	Center of pressure features	Connor (2015)
	2D pressure images (barefeet)	Pataky et al. (2011) and Connor (2015)
Accelerometry	Gait cycle template matching approaches	Bours and Shrestha (2010) and Gafurov et al. (2010)
	Inner products (of multi-axis sensors)	Zhong and Deng (2014)
Audio	Mel-frequency cepstral coefficients	Shoji et al. (2004) and Geiger et al. (2013)

based approaches to gait recognition achieved high recognition rates, but were evaluated only on a few subjects. Eventually, rates reached above 90% with larger populations (e.g., 50 + subjects Hofmann et al., 2012; Liu et al., 2012; Seely et al., 2008). In recent years, a major focus has been to understand and account for the variations in gait between recording sessions (e.g., clothing, walking speed), to make gait recognition a robust biometric cue in real world scenarios. Databases for gait research have also grown with time. Recently, an indoor gait dataset involving 4007 participants was collected (Iwama et al., 2012). Identification accuracies ranging from the mid to high 90s were obtained on this dataset (El-Alfy et al., 2014; Iwama et al., 2012). Another recent development is the availability of video cameras that provide depth information such as the Microsoft Kinect, which has been used in a number of gait studies (e.g., Hofmann et al., 2012; Sivapalan et al., 2011). Such low cost sensors provide additional discriminative gait information that is easy to extract.

Although underfoot pressure sensors and accelerometers are potential modalities for gait recognition, yet they have been relatively less studied. Perhaps this is because they require some form of direct contact with the subject, whereas vision-based approaches can recognize people at a distance from the sensor. Recent technological advances may increase research in these areas. The Bio_Sole concept introduced by "autonomous_ID", a Canadian startup company, measures foot pressures using an insertable insole and emits a signal indicating whether or not the wearer is the "owner". This could be used as a smart ID badge at controlled access points. Vera-Rodriguez and Mason of Swansea University have studied and reported extensively on recognition using similar data (Vera-Rodríguez et al., 2013a), finding that single footsteps can be very distinctive. Stepscan Technologies Inc., another Canadian startup, has recently brought to market its high-resolution (5 mm) and modular pressure-sensitive floor tiles, providing sufficient floor coverage to capture underfoot pressure from multiple consecutive footsteps naturally and unobtrusively. Jung of the Korean university KAIST has repeatedly demonstrated the value of high-resolution pressure information for gait recognition (e.g., Cho et al., 2012; Jung et al., 2004). The remarkable rise of the smartphone (and even smartwatch), with its internal sensors, has increased the potential applicability of wearable sensors for security applications. This decade-old gait recognition from smartphone sensors was largely advanced by a group at Gjøvik University College, involving Gafurov, Derawi, and Bours who studied a variety of issues including gait covariates and gait impersonation or spoofing (Bours and Shrestha, 2010; Derawi et al., 2010b; Gafurov et al., 2007).

In this paper, we focus on gait recognition as a biometric cue. We do so by reviewing several gait recognition modalities and their features. By **modalities**, we refer to the different sensing modes including vision, underfoot pressure, and accelerometry. By **features**, we mean unique and distinguishing measures extracted from the raw data recorded by these modalities.

The paper proceeds as follows. In the next section, we briefly cover the biomechanics of gait and what it means for gait recognition. Following that, Section 3 reviews the various gait recognition modalities in the literature. In each of these we highlight important types of features and their correspondence to the biomechanics of gait. In Sections 4 and 5 we discuss the issues of gait covariates and spoofing. Section 6 wraps up the survey with a discussion on the challenges and open problems in gait recognition. The present survey differs from other gait biometric surveys (e.g., Boyd and Little, 2005; Makihara et al., 2015; Nixon and Carter, 2006; Sprager and Juric, 2015) by including more modalities and focusing much more on the content of feature sets than on feature reduction, classification, and fusion techniques. Instead of reporting only the most popular features, we cover many different and, perhaps, lesser known feature sets to capture as many distinctive aspects of gait as possible. It is hoped that doing so will guide future work to combine features that overcome the limitations of any one modality and arrive at practical systems that take advantage of the unobtrusive nature of gait recognition.

2. Understanding and characterizing gait

What characterizes's one's gait? Since we are interested here in identifying the collection of gait features that uniquely identify individuals, it is worth understanding the various factors that influence the *generation* of gait. Several studies indicate that people can deduce not only the identity of a person (Cutting et al., 1978; Johansson, 1973; Stevenage et al., 1999) but also their gender (Kozlowski and Cutting, 1977) based on their walk. Some gait recognition techniques can even distinguish between people of different ethnicity (Zhang et al., 2010), which might be explained by the fact that some gait analysis parameters have been shown to change with ethnicity (Al-Obaidi et al., 2003; Ryu et al., 2006). This could be a result of cultural or sociological influences, although it is clear that gait features are predominantly dictated by physical traits. Indeed, stride length, joint angle ranges, center of mass, etc., will be influenced by skeletal structure and musculature, as discussed further below. The act of walking is also influenced by emotional state (Roether et al., 2009), although features of gait induced by emotions will be transient and, therefore, of less interest here, except as a potential gait covariate (Janssen et al., 2008; Montepare et al., 1987).

Simple walking is a finely choreographed action, coordinating the timely activation of a large collection of muscles over a complex bone and joint structure to deliver an energy efficient form of locomotion. Diagrammed in Fig. 3, the field of gait analysis breaks the standard walking (gait) cycle (Whittle, 2007) into seven stages and two phases: the swing and stance phases. As the legs travel through this pattern, the arms swing in phase with the opposite leg, and the trunk moves up and down and side to side as it moves its center of mass toward the supporting limb to maintain balance. The foot itself is a complicated network of bones and muscles, and is the body's interface with the walking surface. The spatiotemporal pattern of forces delivered to this surface is the net effect of those invoked during gait. The style of delivery and one's foot shape can significantly define one's gait and its quality, as attested to by the prevalence of the orthotics industry.

Since gait largely proceeds from a mechanical basis, it would seem straightforward to suggest that people of similar gender, age, ethnicity, etc. will have gait similarities partly because of similarities in their physique. For example, men tend to be taller and more muscular than women and so have a longer stride length and higher natural walking speed (Öberg et al., 1993). Walking speed also declines with age (regardless of gender) due to a reduction in stride length rather than cadence (Öberg et al., 1993; Samson et al., 2001). Samson et al. (2001) determined that this was partly due to height and weight differences between young and old subjects. The differences in physique between ethnicities may also influence gait parameters. Al-Obaidi et al. (2003) found that young Kuwaiti men were slightly shorter and substantially heavier on average than young Swedish men. Yet, the Kuwaiti men also walked faster than the Swedes when asked to walk quickly and had a longer stride length regardless of walking speed. In this case, increased height does not account for increased stride length, but other physical or perhaps cultural features may be at play.

Gait can temporarily vary with changes in physical features or circumstances such as walking speed and load bearing. The length and volume of a foot can vary about 5 mm and 4.4%, respectively, from morning to evening (Nah, 1985; Nakajima et al., 2000), which may slightly affect weight distribution. Gait can also shift during the "warm up" period after a lengthy stationary period, such as rising after sleep. Walking speed affects gait (Jordan et al., 2007; Orendurff et al., 2004; Rosenbaum et al., 1994; Taylor et al., 2004) to the degree that a number of gait recognition approaches have been developed or adjusted to overcome it (e.g., Guan and Li, 2013; Kusakunniran et al., 2012; Lee et al., 2007; Yoo and Nixon, 2011; Yu and Zou, 2012). Adding a load can also modify gait parameters (Birrell et al., 2007;
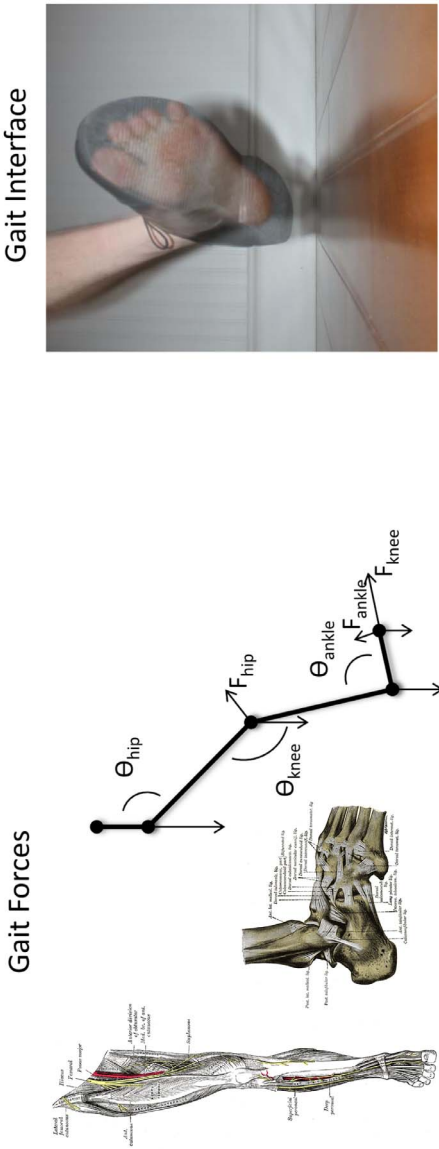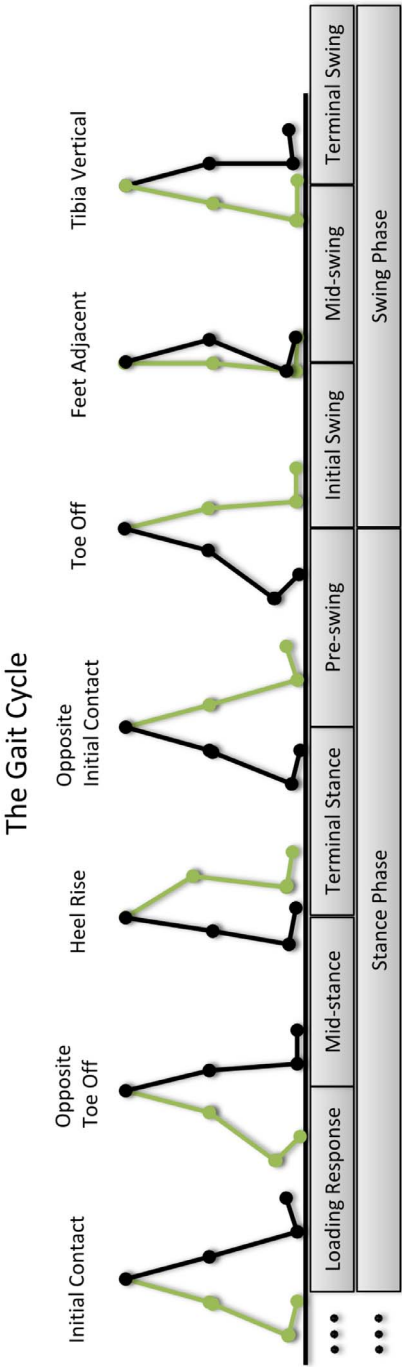
Fig. 3. The mechanical nature of gait. The gait cycle is divided into two phases made up of seven periods, demarcated by seven events with reference to a particular limb (in black), as shown in the upper panel. In the lower-left panel, muscles from the hip, leg, and foot are recruited at various times to provide the coordinated torque forces needed to generate this series of lower body movements and shift the center of gravity. The foot shape, footwear, and the walking surface influence gait by affecting the interface between gait forces and the ground. Anatomical images of the human leg and foot are taken from Gray's Anatomy, 20th edition.

Majumdar et al., 2010; Qu and Yeo, 2011) including their consistency (Hsiang and Chang, 2002), although small loads of 10% body weight or less do not appear to have a very large effect on gait recognition (Moustakidis et al., 2009; Sarkar et al., 2005). Adding a significant load encourages a reduction in walking speed (Diedrich and Warren Jr, 1998), so these two variables are related.

The apparent distinctiveness of the gait will be affected by its generative influences and our ability to reliably measure them. It is the goal of gait recognition, then, to capture the distinctive features of this biomechanical process, some of which may be observable from only certain views and by specific sensing modalities. Murray et al. (1964)

and Murray (1967) characterized gait using 20 specific parameters, many appearing to have within-subject consistency and between-subject variability. These included pelvic, chest, and ankle rotation, and the serial displacement of several body parts. Many gait movements are evident visually, given an appropriate view angle. The measurement of forces or pressures on the walking surface will capture the net effects of ongoing muscle contractions and the timing thereof. Wearable accelerometers will sense the linear and angular acceleration of limb segments where they are placed. These sensing modalities, however, have their limitations. For videos of gait, viewing angle and contrast/brightness conditions may change, and subject clothing can obscure or

change the gait appearance. For underfoot pressure sensors, footwear can redistribute the forces of the barefoot. For audio, the sound of the foot contacting the surface can be buried in background noise, and so on for other modalities. All of these covariates may at first seem to discourage the use of gait as a biometric cue. As will be shown, however, a number of these issues have been effectively addressed and will continue to be investigated.

3. Biometric gait recognition features

Based on the aforementioned commentary, it would seem that to capture the distinctiveness of gait, the various sensing modalities should capture biomechanical measures pertaining to the body's physical dimensions, its body part masses, or the time-varying muscle-generated forces applied during gait. In the past two decades, there have been a large number of such measures or *features* proposed and evaluated. In the next several sections, we review these in detail. Before doing so, we give a brief primer on common terminology and our reporting methodology.

Performance metrics: Most gait recognition approaches are designed to generate a match score that indicates the degree of similarity between two gait samples. This match score is then used to render a decision by the gait recognition system. In the *identification* mode of operation, the most commonly reported metric in the gait biometric literature is the identification rate or the classification rate (CR). It simply indicates how likely a random test or "probe" sample is associated with the correct person in the labeled "gallery" based on the match scores generated. In the *verification* mode of operation, the generated match score is compared against a threshold in order to determine if there is a match or not. The false acceptance rate or FAR denotes how likely a user claiming to be someone else is incorrectly accepted or "verified" as that alternate identity. It's companion, the false rejection rate or FRR, denotes how likely a user claiming to be himself is incorrectly rejected. Changing the matching threshold, results in a set of FAR and FRR values that can be plotted in a two-dimensional graph referred to as the Detection Error Tradeoff (DET) curve or the Receiver Operating Characteristic (ROC) Curve. A common summary measure called the equal error rate (EER) reports the point on the DET curve at which the FAR equals the FRR. More recently, however, reporting only EER values in the context of biometric performance has been discouraged (Jain et al., 2015). Identification can be viewed a multi-class problem, while verification can be formulated as a two-class problem (DeCann and Ross, 2012). For clarity and brevity, we primarily report CRs and EERs in this paper, although these measures must be carefully interpreted. There are metrics which we do not report including expected confusion (Bobick and Johnson, 2001) and the cumulative match characteristic (Phillips et al., 2000). For the latter, we merely report the rank-1 result as the CR.

Data: While many studies use data collected in their own labs, a number of openly available datasets from several modalities have been prepared for evaluating gait recognition techniques, especially for challenging covariates like walking speed and viewpoint. Where a result is based on a public gait database, we refer to it by name. We always report the number of subjects in the evaluated dataset because it indicates the significance of the results. So, studies with large number of subjects exhibiting high CRs, low FRRs, or low EERs are among the most effective (but see Jain et al. (2015) for guidelines involving biometric system evaluation). The number of gallery/probe instances per subject has less of an effect on these measures, but is relevant when considering statistical confidence on results such as given by the "rule of 30" (Doddington et al., 2000).

Feature reduction and classification methods: While our focus is on the extracted features of gait, results for each method are also influenced by the associated feature reduction methods and classifiers used. Because the present focus is on discerning how features relate to the biomechanics of gait and on capturing distinctive aspects of gait, the exact classification/verification rates are not as critical as recognizing effective versus ineffective features and distinct versus similar features. Nevertheless, we report feature reduction and classification methods in each study for the sake of completeness, by brief acknowledgment and often through their acronym only. For additional details on a specific method, the reader is directed to the associated paper for which the results are reported.
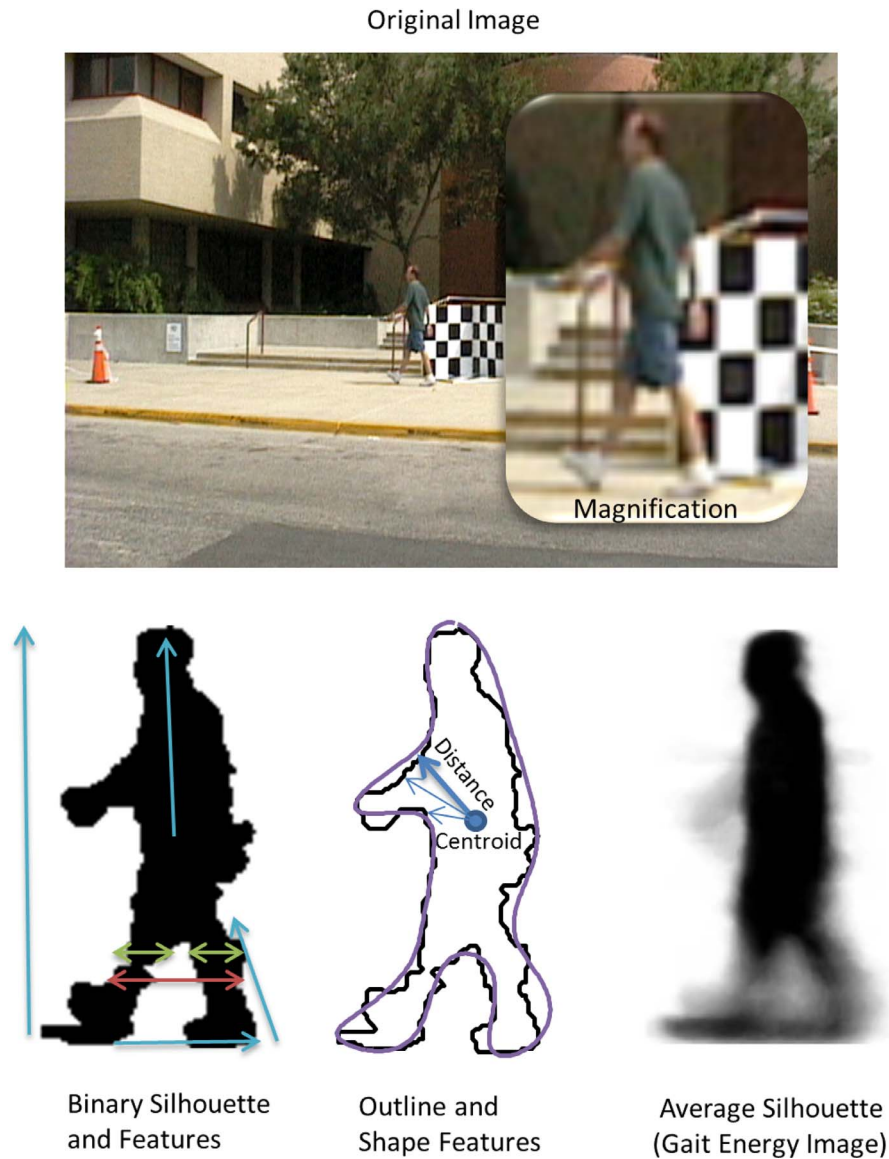
3.1. Gait features from video

Video is an obvious sensory means of perceiving gait. Sometimes referred to as the "machine vision" approach, it can be subdivided into two general categories: model-based and model-free. Model-based approaches usually attempt to fit a walking model (e.g., a multi-segment skeleton) to the video frames and then compute features based on the model parameters. Model-free approaches, however, extract features directly from the video imagery or silhouette images derived therefrom.

3.1.1. Model-free features: moving silhouettes

Early on in gait recognition research, it seemed appropriate to segment the walker's pixels from the background pixels, which resulted in the binary silhouette shown in Fig. 4. This is the most common data preprocessing step in vision-based gait recognition studies. Table 2 captures a summary of these studies, describing the feature sets extracted, their performance, the number of participants in the gallery, the study reference. These studies are also reviewed in the text of section. Murase and Sakai (1996) extracted silhouettes from video of gait and used Principal Components (PCs) to reduce the 3D (x, y, and time) data down to 16 features per time instance. They achieved 100% recognition accuracy (7 subjects, Linear Time Warping, 1NN). Huang et al. (1999) extended this approach to include canonical analysis (CA), which encouraged separation between the classes (i.e., individuals). They achieved 100% (6 subjects, 1NN by class means) on the same dataset, but had much improved the class separation. BenAbdelkader et al. (2001) computed a silhouette self-similarity matrix, that is, the pair-wise correlation between each silhouette image and all others in the same gait sequence. This achieved a CR of 93% (7 subjects, PCA, kNN). Collins et al. (2002) extracted, cropped, and scaled key frames from the image sequence that represented specific poses in the gait cycle. They evaluated these as features on several different gait databases, giving promising results: 87% (UMD outdoor dataset (Chalidabhongse et al., 2001), 55 subjects, 1NN), 93% (University of Southampton dataset (Nixon et al., 2001), 28 subjects, 1NN), and 100% (MIT dataset (Lee and Grimson, 2002), 24 subjects, 1NN). Liu and Sarkar (2006) normalized binary silhouette sequences according to a generic or common walking model they computed from many different subjects' sequences using a hidden Markov model (HMM). Like Huang et al. (1999), they also used a linear discriminant subspace to enhance between-subject differences in each key frame. When testing on the USF humanID gait challenge database (Sarkar et al., 2005) (hereafter referred to as the humanID gait database), they achieved substantially better performance than the dataset baseline when the walking surface condition was changed (60% over 40%, 122 subjects, 1NN) and when the carrying condition was changed (89% over 62%, 122 subjects, 1NN). Their algorithm also compared favorably on the CMU MoBo database (Gross and Shi, 2001) (84%, 25 subjects, 1NN), in which various walking speeds are present.

Body width or build: Several approaches have collapsed the silhouette in ways that appear to capture the person's width or build. Kale et al. (2002) extracted the silhouette width of each image row from five key frames in a video sequence and modeled this with an HMM. With this approach they achieved a CR of 89% (5 subjects) on Little and Boyd's data (Little and Boyd, 1998), 56% (43 subjects) on the UMD database, and as much as 70% (25 subjects) on the CMU MoBo database (training and testing on same pace data). Cuntoor et al. (2003) used a vector of silhouette widths (example shown in red in Fig. 4) of the leg

Original Image

Fig. 4. Gait features derived from silhouettes in video sequences. From an image sequence, a binary silhouette of the person walking is first extracted. Then features can be extracted from the sequence of silhouettes, such as body segment lengths and width features. The outline of the silhouette can also be extracted and characterized using a variety of methods. An average silhouette can be substituted for the sequence of binary silhouettes, although it has less temporal information. The original image and silhouette shown here are part of the humanID gait database. (For interpretation of the references to color in the text, the reader is referred to the web version of this article.)

Table 2
Representative model-free gait recognition feature sets.

Feature set	Performance	Subjects	Reference
Silhouette PCs	100% CR	7	Murase and Sakai (1996)
Silhouette key frames	87, 93, 100% CR	55, 28, 24	Collins et al. (2002)
Silhouette row widths (HMM)	89, 56, 70% CR	5, 43, 25	Kale et al. (2002)
Row width averages (GPCA)	0.9% EER	16	Cattin (2002)
Areas of partial silhouettes (CA)	75.4% CR	114	Foster et al. (2003)
Body part lengths and height	91–100% CR	18	Johnson and Bobick (2001)
Average silhouette or GEI	44–57% CR	71	Han and Bhanu (2004, 2006)
GEI from 2D-projected 3D silhouettes and multiple views	96.6–99.6% CR	103	Seely et al. (2008)
Chrono-gait image	51.3% CR	122	Wang et al. (2010) and Liu et al. (2012)
Gradient histogram energy image	79.8% CR	122	Hofmann and Rigoll (2012)
Depth silhouettes (radon trans., Krawtchouk moments, GA)	57.4% CR	122	Ioannidis et al. (2007)
Gait energy volumes	95% CR	25	Sivapalan et al. (2011)
Histogram of gradients on depth silhouettes	97.7% CR	176	Hofmann et al. (2012)
Fourier descriptors (FDs) of silhouette outline	85% CR	114	Mowbray and Nixon (2003)
Histogram of normal vectors on silhouette outline	97.5% CR	3141	El-Alfy et al. (2014)

region. These features performed well on the CMU dataset giving a 91% CR (25 subjects, PCA, k-means, HMM) but not as well on the UMD dataset giving a 31% CR (44 subjects). Instead of a person's horizontal coverage, (Cattin, 2002) computed "histogram" or width features from silhouettes as the sum of silhouette pixels in each row (depicted by green arrows in Fig. 4). Cattin took the mean, standard deviation, and the 2D Fourier transform of each row's width over a gait cycle. The mean performed best, achieving a very low EER of 0.9% (16 subjects, GPCA). The standard deviation and frequency domain features were also discriminative (EER of 4.3% and 7.8%, respectively). Hong et al. (2007) used the same histogram features with dynamic time warping (DTW) on the CASIA-A (former NLPR) database (Wang et al., 2003b) and achieved 96.25% correct classification (for the 0 and 45° viewpoints, 20 subjects, 1NN). Lee et al. (2007) explored a slightly different approach based on self-similarity of images to a specified key frame, which they referred to as a "shape variation-based frieze pattern". These features are computed by taking the difference between a silhouette key frame and the silhouette video frames and summing along the horizontal or vertical axis, thereby collapsing the differenced image. These features, along with a measure of their half gait cycle-symmetry, achieved an average CR of 85% on the CMU MoBo database (25 subjects, 1NN based on a special cost function). Foster et al. (2003) computed areas of specific parts of the silhouette (centered horizontal and vertical rectangles, bottom half, and full area) as features over time. They achieved between 16 and 30% recognition (114 subjects, CA, kNN) for individual areas and a high CR of 75.4% for the combined areas.

Physical dimensions: The silhouette has also been used to extract features of body appearance such as height and the distance between certain body parts, shown by blue arrows in the lower-left of Fig. 4. Johnson and Bobick (2001) extracted height, head-to-pelvis distance, maximum leg length, and stride length. For several different view perspectives, and when scaling parameters to a common view, they achieved classification rates between 91% and 100% (18 subjects, maximum likelihood classifier based on a Gaussian fit to each individual's data). BenAbdelkader et al. (2002) extracted height, stride length, and cadence from silhouettes. Together, these features achieved 47% and 65% CRs for the side and oblique views (41 and 17 subjects, 1NN), respectively. When the height information was removed, rates dropped to 18% and 51%, showing height to be a discriminative feature.

Average Silhouette: Another way in which the spatiotemporal binary silhouette has been transformed is by averaging it over a complete gait cycle, as shown in the lower-right panel of Fig. 4. When doing this, Liu and Sarkar (2004) found the average silhouette matching performance to be at par with the baseline algorithm on the humanID gait database (Phillips et al., 2002; Sarkar et al., 2005), which is based on similarity matching of the silhouette sequences between gallery and probe samples. The benefit, however, is that the average silhouette is far more efficient to use and store than a video sequence. Han and Bhanu (2004, 2006) dubbed the average silhouette (horizontally aligned and scaled) as the "Gait Energy Image" (GEI) and employed it on the humanID gait database confirming Liu and Sarkar's finding by getting an average of 44% (71 subjects, 1NN) for direct template matching, with the average baseline result being 42%. However, they also achieved an average of 57% (PCA, MDA) by including a statistical matching approach. The substantial improvement is due to better performance on the more challenging covariates. It is the inclusion of these covariates that makes the humanID gait database so challenging and resulting in lower CRs than most other databases. Tao et al. (2007) convolved average silhouettes with a bank of differently scaled and oriented Gabor filters to get a 4th-order tensor, and summed over one or two of the Gabor dimensions (either scale, direction, or both) to get 2D and 3D feature sets. They applied these to the humanID gait database and achieved a best average classification rate of 60.6% (122 subjects, GTDA). Seely et al. (2008) synthesized an average silhouette

(centered and normalized for size) for a side view that was computed from silhouettes extracted from multiple viewpoints. On its own, this gave a CR of 96.6% (103 subjects, 1NN). Without size normalization, the classification result of the centered data improved to 98.6%. They also concatenated the average silhouettes from three different views (side, front, and top) and achieved a further improvement in performance (99.6%). The average silhouette or GEI has become the baseline model-free feature representation against which new methods are usually compared. Iwama et al. (2012) introduced the OU-ISIR Large Population dataset of 4007 participants, they evaluated several different model-free features of which GEI gave the best CR of 94.24% (1NN, 3141 subjects, data from 4 different camera angles) and an EER of about 1.2%.

Other 2D image silhouette derivatives: Taking a cue from the GEI study, a number of related 2D image model-free representations have been proposed. Bashir et al. (2009) introduced the gait entropy image (GEnI), which gives each pixel an "uncertainty" value. The more stable a given pixel is across images in the gait cycle, the smaller the entropy value, and thus high values highlight areas of motion and the body boundary. It performed comparably to GEI on several datasets, but improved on several of their covariate cases. On the bag-carrying case of the CASIA-B gait database (Yu et al., 2006), it improved on GEI's 60.2% (124 subjects, GEI, PCA, MDA, 1NN) achieving a CR of 80.1% (GEnI, PCA, MDA, 1NN). On the coat-wearing case of the SOTON small dataset C (Shutler et al., 2002), it improved on GEI's 72.7% (11 subjects), reaching 81.8%, and on the briefcase-carrying case of the humanID gait database outperformed GEI's 62% (Han and Bhanu, 2006) yielding 82%. An improved result of 93% on this briefcase-carrying case (Liu et al., 2012, 122 subjects) was achieved using the Chrono-Gait Image (CGI) (Wang et al., 2010). To compute the CGI, the binary silhouettes in the sequence are given an RGB color according to their time stamp and then averaged to recapture some sense of the time sequence that is lost when taking the average silhouette from the sequence. Another related approach is the use of histograms of orientation gradients (HOG) (Dalal and Triggs, 2005; Liu et al., 2012), which has proven popular in other computer vision tasks. Here, the orientation gradient, or average direction (x,y) of intensity change is computed for each pixel. For many individual cells or areas of the image, a histogram of the gradients is created and normalized for contrast. In Liu et al. (2012), combining the HOG representations of the GEI and CGI images achieved an average CR of 59.4% (122 subjects, PCA, LDA, 1NN) on the humanID gait database. Hofmann and Rigoll (2012) proposed another histogram based representation. Their data source used the original RGB frames masked by the silhouette to show only the walkers' pixels. For each pixel, they computed the gradient orientations and generated 9-bin orientation histograms over non-overlapping 8x8 pixel regions that were normalized by neighboring histograms. The histograms were then averaged over a full gait cycle of images to produce the "Gradient Histogram Energy Image". They tested their method on the humanID gait database and achieved a high average CR of 79.8% (122 subjects, PCA, LDA, 1NN), substantially better than previous methods. Their most significant improvement was in the covariate cases where the probe videos differed from the gallery videos with respect to the following attributes: walking surface, shoe, view, and carrying of a briefcase. One possible reason for the overall improvement is the additional information embedded in the original imagery. It may also be partly due to the method capturing something of the details of the clothes. The method is quite ineffective on the clothing and time covariates, where the time factor is likely also affected by the change in clothing between recording sessions.

Depth: Instead of using monochromatic or color imagery, some studies use depth imagery, which can be captured using consumer motion capture devices such as Microsoft Kinect or industrial-level lidar systems (Galai and Benedek, 2015). Similar to Hofmann and Rigoll (2012), Ioannidis et al. (2007) assigned real values to the silhouette pixels. However, the silhouette values were based on the

distance from the center of mass and used the "geodesic" distance. From these silhouettes, they extracted generalized radon transform features, which take the sum of pixel values projected onto a line; several lines at various angles were used for this purpose. From these, they extracted Krawtchouk moments and used a genetic algorithm for classification. For the humanID gait database (Sarkar et al., 2005), they achieved an average CR of 57.4% (122 subjects, subsets A–G). Sivapalan et al. (2011) proposed to use average silhouette volumes, or "Gait Energy Volumes". They showed how volumes could be computed either from a depth camera such as the Microsoft Kinect or from two or more standard cameras by silhouette extraction and registration. Then they centered, scaled, and averaged these volumes over time. Using the CMU MoBo database, the gait energy volumes achieved a CR as high as 95% (25 subjects, PCA, MDA, 1NN), which was as good or better than GEI. Hofmann et al. (2012) computed a HOG representation using depth imagery. On their large dataset of 176 subjects recorded with a Microsoft Kinect (version 1), they achieved a 97.7% CR (normal gait condition, 1NN) which was substantially better than that of GEI at 69.3% and GEIs based on silhouettes extracted from depth data at 90.1%. This approach also outperformed gait energy volumes (Sivapalan et al., 2011), which achieved a CR of 79.6%.

Body shape: The body shape itself has been shown to be discriminative. Wang et al. (2003b) investigated the shape of silhouette outlines from several camera angles. The outline boundary points were translated into distances from the outline centroid, as illustrated in the lower-middle panel of Fig. 4, and PCA was used to reduce the dimensionality (15 features). For identification, they achieved a 94% CR (20 subjects, PCA, 1NN by class means) for the frontal view of the CASIA-A database. The authors also extracted several physical features (gait period, stride length, height, ratio of chest width to body height) which increased the CR of the side view from 75% to 82.5%. Wang et al. (2003a), using the same database, employed Procrustes shape analysis on the outline, rather than mere distance to the centroid. The approach achieved a best result (frontal view again) of only 90% this time (ENN). For the side view, however, the accuracy improved to 89% (from 75%) without the help of the physical dimension features employed before (see also Wang et al., 2002). Several studies have extracted Fourier descriptors (FDs) of the silhouette outline, which define the violet contour of the lower-middle panel of Fig. 4. Yu et al.'s (2004) use of key FDs on the SOTON gait database (Shutler et al., 2002) achieved a CR of 85.2% (114 subjects, 1NN). On the same database, Mowbray and Nixon (2003) also used FDs and achieved an 85% CR. Kusakunniran et al. (2011) used a novel "higher-order shape configuration" and Procrustes shape analysis. Testing on the CASIA-C gait database (Tan et al., 2006) (50 subjects, walking fast, normal, and slow), they evaluated various speed combinations of gallery and probe datasets and achieved CRs between 92% and 98% (1NN). El-Alfy et al. (2014) used a histogram of normal vectors (HoNV) to describe the shape of the silhouette boundary. They computed a set of vectors that are normal to the silhouette boundary, divided the image into a grid, made a histogram of the normal vectors within each grid cell and concatenated the histograms together. Finally, they averaged the histograms over all frames in a gait cycle. Testing on the OU-ISIR Large Population dataset, their method outperformed HOG and GEI, achieving a low EER of 0.58% (1NN, 3141 subjects) and a CR of 97.5% based on video from four different camera positions and angles.

### 3.1.2. Model-based features: body models fitted to video

Many different methods have been used to derive a skeletal or body model from video sequences. The goal is to identify key anatomical locations accurately and efficiently. Most approaches are complex and computationally expensive as they try to optimize the fit of some number of skeletal segments or volumes to the data. Alternatively, several gait recognition studies use the Microsoft Kinect and its API, which is one of several commercial marker-less motion capture systems becoming readily available. An example of this is shown in Fig. 5. After

fitting the skeletal model, regardless of the approach, the extraction of joint positions and angles is straightforward. Table 3 summarizes a number of features based on this type of information. Note that most model-based gait recognition approaches use their own distinct model-fitting approach, which would seem to complicate cross-comparisons between studies with similar model-based features.

Early video-based gait research involved fitting shapes to the imagery and using their parameters as features. Niyogi and Adelson (1994) extracted the side-view outline of a walker using four splines which capture the front contour, back contour, and two inside leg contours. With the positions traversed by these contours as features, they achieved 81% correct classification (5 subjects, 1NN). Little and Boyd (1995) used optic flow to fit two ellipses to the motion of the walker. Although they did not perform recognition, they computed the vertical and torque phase differences between the centroids of the ellipses and showed that these show some degree of similarity between walks of the same subject. In a subsequent study (Little and Boyd, 1998), they again fit two ellipses to the optical flow data and derived a series of phase features from them (the centroid and elongation, and inter-ellipse differences thereof). For a subject pool of 6, they achieved CRs as high as 95.2% (5 ellipse features, ENN). Lee and Grimson (2002) expanded upon this approach by fitting seven ellipses to various regions of the binary silhouette. From each region, the following attributes were extracted: centroid, elongation, and major axis orientation. Total height was also computed. From these attributes, the mean, standard deviation, and Fourier spectrum over all frames in a video sequence were used as features. For a reduced set (41 of 57) of the mean and standard deviations features on the within-day dataset of the CMU MoBo database gave 100% classification (25 subjects, 1NN). When the data in the gallery vs. probe was taken from different days, the CR plummeted to around 40%, however.

As in the silhouette-based features described earlier, features describing gait and specific body dimensions can also be derived from skeletal models fitted to gait imagery. Tanawongsuwan and Bobick (2003a) extracted stride length and cadence information from motion capture data. Transforming these for classification at a fixed walking speed, they achieved between 20% and 35% CRs (15 subjects, 1NN) across different walking speeds using stride length, cadence, or both. Preis et al. (2012) extracted step length and cadence among other features from Microsoft Kinect (version 1) motion capture sensor data. Step length and cadence together achieved a 55.2% CR (9 subjects, C4.5, Naive Bayes). Bhanu and Han (2003) fit a 12-segment skeletal model to the silhouette, and extracted segment lengths of the neck, torso, upper arm, forearm, thigh, calf and foot. Classification rates increased with the number of structural features to a maximum CR of 62% (8 subjects, 1NN). Bhanu and Han also extracted the mean and standard deviation of joint angles (neck, upper arm, forearm, leg, and thigh) and achieved a maximum CR of 72% when increasing the number of angular features. Combining both types of features gave an 83% CR. Wagg and Nixon (2004) used mean gait parameter profiles (e.g., average hip rotation angle over a gait cycle) to fit individuals' gait cycle to a multi-shape (ellipses and rectangles) model. With their model, they extracted 63 features per gait cycle, defining parameters for joint rotation models (45 features for hip, knee, and ankle) and body size parameters (18 features). Performing statistical analyses, they found that the lower knee width, ankle width, and cadence were the most discriminative features. Classification accuracy on the SOTON large gait database was 84% (115 subjects, 1NN) on indoor data and only 64% (115 subjects, 1NN) on outdoor data. Yoo and Nixon (2011) fit an 8-segment skeletal model (no arms) to digital video sequences on the same SOTON gait database. From this model, they extracted 10 types of features: height, cycle time, stride length, speed, average joint angles, variation of hip angles, correlation coefficient between the angles of the right and left legs, and the center coordinates over the hip-knee cycle. They achieved CRs of 96.7% (30 subjects, 1NN) and 84.0% (100 subjects, 1NN). Dikovski et al. (2014), also using the Microsoft
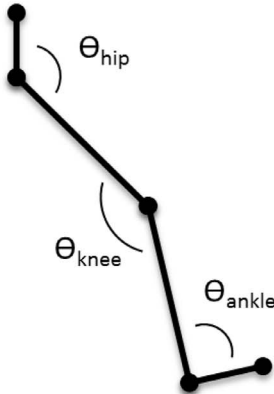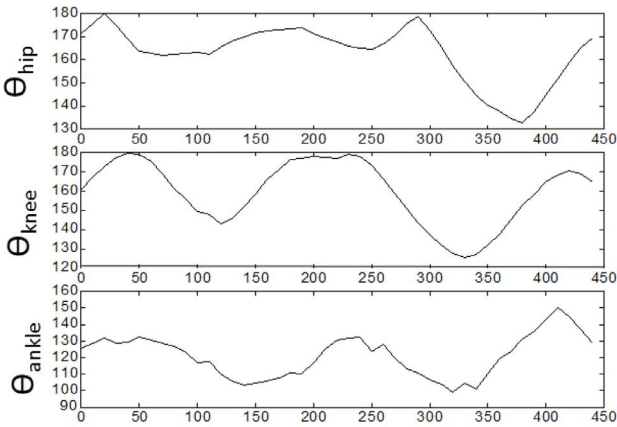
Skeleton Fitting (MS Kinect v2)
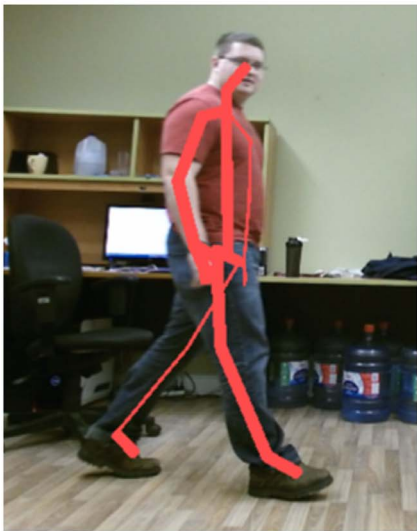
Joint Positions
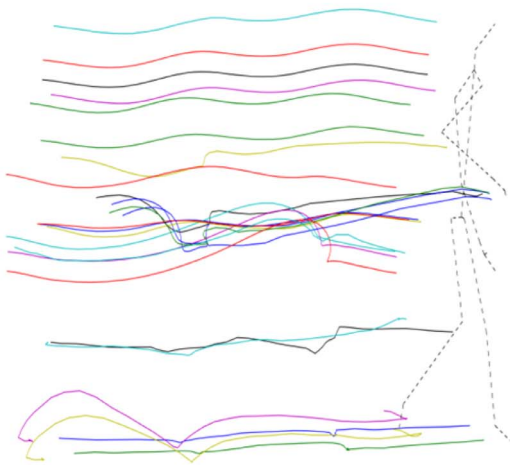
Joint Angles

Time (ms)

Fig. 5. Gait features derived from a skeletal model fitted to a video sequence. The image from a Microsoft Kinect (version 2) contains a 25-joint skeletal model fitted by the Microsoft Kinect SDK. Joint positions are recorded from this model over time, here shown with a moving average filter applied to reduce joint position noise. Joint angles over time can also be easily computed, from which additional features may be derived.

Table 3
Representative model-based gait recognition feature sets.

Feature set	Performance	Subjects	Reference
Fitting two ellipses	95.2% CR	5, 6	Little and Boyd (1998)
Fitting seven ellipses	100% CR	25	Lee and Grimson (2002)
Body part lengths and joint angles (12-segment model)	83% CR	8	Bhanu and Han (2003)
Body part lengths and joint rotation model parameters	84, 64% CR	115, 115	Wagg and Nixon (2004)
Joint angle trajectories	73% CR	18	Tanawongsuwan and Bobick (2001)
Joint position trajectories	96.9% CR	25	Świtoński et al. (2011a)
Fourier magnitude and phase on joint angle trajectories	79.4% CR	46	Ariyanto and Nixon (2011)

Kinect sensor (version 1), evaluated a collection of segment length and angular features (mean, max, min, and std of the time sequence data). They found that the segment length features (89% CR, 15 subjects, MLP) were more distinguishing than angular features (78%) and that lower-body features (89%) were more telling than upper-body features (81%). Ariyanto and Nixon (2011) extracted a model from 3D gait data (laser scan) to determine height, stride length, and footstep pose structural features. Height was most discriminative (41.3% 46 subjects, 1NN) followed by footstep pose (17.4%) and stride length (7.1%). The best combination of these features gave a maximum of a 56% CR (footstep pose plus height).

A key advantage of the model-based approach is that it provides the most direct description of locomotion by providing joint angle and position trajectories over time. Tanawongsuwan and Bobick (2001), using hip and knee angle trajectories from marker-based motion capture data, achieved a 73% CR (18 subjects, 1NN, DTW). Goffredo et al. (2010) fitted a skeletal model and, by assuming the walk was in a straight line, transformed the joint position trajectories to a

common (side) view. Initially testing on the SOTON database (Shutler et al., 2002; 2000), where the recording camera always has a side view, they used dynamic gait features (hip and knee angle traces) for recognition and achieved a 95.8% CR (20 subjects, kNN). Testing on a subset of the CASIA-B database (Yu et al., 2006) (65 subjects), where viewing angle is variable, they achieved a 73.6% CR. Although their approach achieved lower CRs than others (Bashir et al., 2008; Yu et al., 2006) when the gallery and probe camera angles were the same, they achieved 10s of percent higher CRs than the others when the gallery and probe camera angles were significantly different. In two studies, Świtoński et al. (2011a, 2011b) extracted human motion through a marker-based motion capture system. From these, they extracted gait paths for the head, both feet and their average position, both hands, and the base of the spine or root. The paths of the hands and feet were most discriminative, followed by the root and lastly the head. These paths as features achieved as high as a 96.9% CR (25 subjects, MLP). They also evaluated velocity and acceleration along the paths, finding these to be more discriminative than the raw position trajectories alone. In their second paper, using 22 angular trajectories (position, velocity, and acceleration), they achieved a CR of 98.0% (25 subjects, PCA, kNN). Yu and Zou (2012) fit a 3D, 18-segment skeleton to single-camera video of the CASIA-B database. For identification, thirteen joint position trajectories were used as features, for subjects in several clothing conditions. Accuracy was highest when subjects wore a thin coat (93%, 124 subjects, NPE, HCRF, SVM), slightly less with a thick coat (89%) and less still with a backpack (85%).

Fourier-based methods have been used in several studies to approximate the gait trajectories to good effect. Cunado et al. (1997) used the Hough Transform to determine the hip and knee angles from video and computed the magnitude and phase spectra of the Fourier transform from them. They found that the magnitude alone was not as discriminative (50%, 10 subjects, 1NN) as combining the magnitude and phase information (80%). A subsequent paper by Cunado et al. (2003) employed a more detailed gait model. Taking the magnitude and phase-weighted magnitude of the Fourier series describing the pattern of hip rotation, they achieved CRs of 80% and 100% (10 subjects, same data as previous study, 1NN), respectively. Yam et al. (2004) expanded on this further, computing the magnitude and phase features for the rotation of both the hip and the knee. These features achieved a maximum CR of about 85% (20 subjects, 1NN) for walking and above 90% for running. Bouchrika and Nixon (2007) investigated dynamic features of gait by fitting an elliptical Fourier descriptors model to gait video. In so doing, they extracted the joint locations as well as the movement. In recognition, they included static body parameters (stride, overall height, and height of body parts) and the hip and knee angular motions as described by the phase-weighted magnitudes of the Fourier descriptors. In total they achieved a 92% (20 subjects, kNN) CR. In a second paper (Bouchrika and Nixon, 2008b), they extracted joint angles over time using the same approach and then used feature subset selection to reduce the number of features used for classification. Their best result was 95.7% (20 subjects, kNN), derived from purely dynamic features of gait. Ariyanto and Nixon (2011) took the thigh and shin angles (sagittal and frontal plane) plus the center of hip rotation throughout the gait sequence and computed Fourier magnitude and phase features. From the abduction angle of the left thigh, right thigh, and left shin along with the center of the hip position, a 79.4% CR (46 subjects, DTW, 1NN) was achieved, which proved better than the segment length features of the same study (a best CR of 56%). The right thigh abduction angle, a frontal-view feature, was also particularly informative alone (about as effective as height and footprint pose together), perhaps reflecting the width of one's stance during gait. More recently, Kovač and Peer (2014) used gait trajectories (angles, phase differences, body part mass ratios, and distances between centers of mass) from an 8-segment (no arms) skeletal model. The trajectories were recast as Fourier descriptors and evaluated on the OU-ISIR gait database A (Tsuji et al., 2010) which contains recordings of different

walking speeds for each subject. With a transformation for walking speed invariance and fusion of various features, they achieved about 95% (25 subjects, PCA, 1NN) CRs when the gallery and probe walking speeds were within 2 km/h. Classification for running speeds (8–10 km/h) also had high classification rates (usually high 90s or 100%), so such features are at least as distinctive for running as for walking.

3.1.3. Correspondence to gait mechanics

Both model-based and model-free features tell us much about the physical structure of the walker. In the model-based approaches, the approximate length of each body segment is computed from joint positions deduced from the video. This appears to be the most accurate and direct way of determining the lengths and positioning of body segments, and can provide good recognition rates (Dikovski et al., 2014; Wagg and Nixon, 2004), although it requires a computationally expensive process to fit such models. The model-free, silhouette-based approaches provide information about the subject's approximate build and body shape (i.e., the distribution of mass), putting flesh on the body structure. Features that seem to capture this effectively are the width histogram techniques (Cattin, 2002; Hong et al., 2007) and Fourier descriptors of the silhouette outline (Mowbray and Nixon, 2003; Wagg and Nixon, 2004; Yu et al., 2004). These features, however, also include temporal information, so their effectiveness is not a reflection of body shape discriminability alone. Foster et al.'s (2003) effective area features, which also give a sense of build, saw a drop in CR from 75.4% (SOTON database, 114 subjects, CA, kNN) to 52.7% after removing the means of the areas over time, suggesting that there is both a constant (structural) component that is discriminative and a temporal component. The average silhouette or GEI nicely focuses on representing the non-temporal aspects of silhouette data, that is, a person's build. Lee et al. (2007) pointed out that the same average silhouette can be generated from a shuffled set of its source silhouette frames, and therefore must not contain much temporal information. This may explain why average silhouette studies (Liu and Sarkar, 2004; Veres et al., 2004) found that the (static) upper-body is equally or more discriminative than the (dynamic) lower-body, whereas a silhouette sequence-based study (Phillips et al., 2002) saw the lower-body as far more discriminative. Regardless, Seely et al.'s (2008) CR of 99.6% (103 subjects, 1NN) and Iwama et al.'s (2012) CR of 94.24% (3141 subjects, 1NN) demonstrates that the GEI, is a strongly discriminating representation. Depth imagery's information appears help to improve silhouette extraction quality (Hofmann et al., 2012), which could boost the effectiveness of the many dependent feature representations.

Model-based approaches most directly and efficiently represent gait dynamics. They do this by expressing movement as a set of joint position trajectories, rather than as a sequence of images. Movement is the outcome of forces applied by muscle groups to limbs and the body. Świtoński et al.'s (2011a) model-based study considered velocities and accelerations of specific joints and found these to be more informative than the mere shape of the movement (i.e., the joint positions). They concluded that, "… more important is how energetic the movements are rather than what is their shape.". Joint position or angle is an indirect function (a second order integral) of acceleration, which is a direct function of force and mass ($F = ma$). So, it seems appropriate that acceleration be more discriminative than position/angle, which is the cumulative (i.e., averaged) effect of a history of accelerations. Regardless, several model-based approaches (Bouchrika and Nixon, 2008b; Kovač and Peer, 2014; Yu and Zou, 2012) effectively used position or angular trajectories to achieve positive results.

Both model-based and model-free approaches give us information about the high-level features of gait, such as cadence and stride length (BenAbdelkader et al., 2002; Preis et al., 2012). It would seem, however, that the model-based approach is generally better suited for computing these, given that it provides a specific anatomical position for the heel, by which these measures are commonly determined.

Most studies observed gait from the side view, but other views

contribute significantly as well. Ariyanto and Nixon (2011) found that the right thigh abduction angle, a frontal-view feature, was particularly informative (about as effective as height and footprint pose together). Indeed, unlike a side view, a frontal view is not affected by self-occlusion. Seely et al. (2008) improved results slightly from 98.6% to 99.6% (103 subjects, 1NN) by including average silhouettes from three different views (side, front, and top). Several studies described above (BenAbdelkader et al., 2002; Collins et al., 2002; Hong et al., 2007; Wang et al., 2003b) found that the side view was less effective than either the front or oblique (45°) views for their features. In contrast, Świtoński et al.'s (2011a) model-based study found that the fore-aft and vertical dimensioned features were most discriminative and thus recommended a side view.

3.2. Gait features computed from underfoot force/pressure sensors

A key aspect of gait is the force placed on the ground by the foot. Using this information as a biometric is sometimes referred to as floor sensor-based gait recognition (Liu et al., 2009) or footstep recognition (Rodriguez et al., 2009). Although not directly based on force, Kennedy (1996) found that simple geometric features of barefoot ink impressions were highly unique among several thousand people, and such impressions have subsequently become a separate form of biometric (e.g., Kumar and Ramakrishnan, 2011; Kuragano et al., 2005; Uhl and Wild, 2008). In the same way that unique body size and shape will influence gait, so will a unique 3D foot size and shape influence gait, especially in terms of underfoot pressures. In the literature, underfoot pressure profiles have been sensed by a grid of on/off switches or by pressure sensor arrays which vary greatly in resolution and arrangement. At the coarsest resolution, a vertical ground reaction force profile is captured by a few sensors (a force plate), usually with a high sampling rate and large range of forces. As resolution increases, the pressures (or binary switch activation) applied by different areas of the foot are more precisely specified, giving the geometric shape, orientation, and COP of the foot, and thereby providing a spatially and temporally rich base for deriving distinctive features of gait.

3.2.1. Ground reaction force features
Ground Reaction Forces (GRFs) are the counteracting pressures placed on the foot as it comes in contact with the floor and can be delivered vertically, laterally, and fore-and-aft. When these pressures

Table 4
Representative ground reaction force (GRF) feature sets.

Feature set	Performance	Subjects	Reference
GRF profile key points features	93% CR	15	Orr and Abowd (2000)
HMM of GRF profile	91% CR	15	Addlesee et al. (1997)
PCs of GRF profile and its gradient	13% EER	41	Vera-Rodríguez et al. (2007, 2008)
DTW of GRF profile	4.0% FAR, 2.2% FRR	132	Derlatka (2013)
WP decomposition of GRF	98.2% CR	40	Moustakidis et al. (2008)
Mass computed from GRF	50% CR	43	Jenkins and Ellis (2007)

are measured over the time of a footfall, a GRF profile is acquired. Fig. 6 shows the vertical GRF profile's "M" shape, which corresponds to a peak force from the initial heel impact and another peak when the person is propelled forward by the same foot. In the lateral and fore-and-aft directions, there are corresponding events, although of lower forces overall. Table 4 provides a list of the most common types of gait recognition features used in this modality.

One intuitive and popular approach has been to extract features based on key points of the GRF profile and specific statistical measures. Orr and Abowd (2000) used 10 features based on vertical GRF profiles to identify individuals. These included mean profile value (i.e., approximate weight), standard deviation of the profile, length of the profile (footstep duration), total area under the profile curve, maximum heel-strike force and its location in time, minimum force and its time, and the maximum push-off force and its time (see Fig. 6). Together, these features gave a 93% recognition rate (15 subjects, 1NN). Suutala and Röning (2004) expanded on these features, reaching 31 distinctive features. They found 13 of these to be most descriptive, including all those in Orr and Abowd (2000) except the total area under the profile curve. Their results, however, were much poorer, achieving only 70.2% (11 subjects, Distinction-Sensitive LVQ). In a subsequent paper by Suutala and Röning (2005), they reworked the classification step and improved results to 79.2% (11 subjects again, MLP). Vera-Rodríguez et al. (2007, 2008) used two piezoelectric sensors (one for the heel and one for the toes) to capture pressure data that gave similar profiles to the gradient of the GRF profile signals seen in other work.
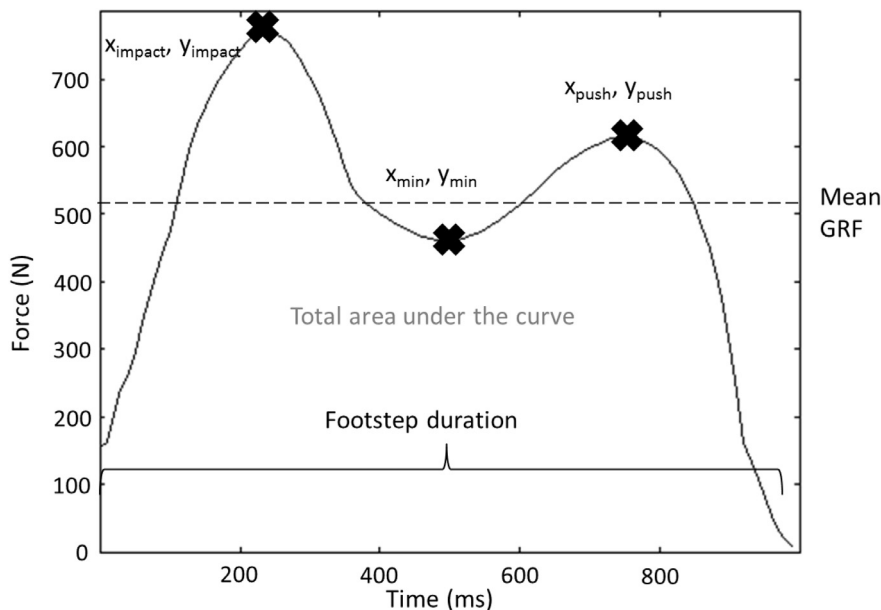

Fig. 6. A typical vertical GRF profile and some of its key point and statistical features. The characteristic M shape has a peak force for heel impact and push-off.

They extracted five key points based on maxima and minima in the two sensors' curves. For each key point, the magnitude and time were recorded (i.e., 10 features). Also, the differences in magnitude and time between key point pairs (i.e., 20 features) and the total profile area under the curve, length, norm, mean, and standard deviation for each sensor's profile were computed (i.e., 12 features). These 42 features achieved an EER of 16% (41 subjects, SVM). They reduced the features down to the 17 most informative ones using a greedy forward feature selection technique and achieved an EER of 12.5%. Selected features were five time- and six magnitude-based key point derived features, as well as the norm, area under the curve, and standard deviation for both sensors. Mason et al. (2016) also used GRF key points and their differences to achieve a low EER of 1.33% (10 subjects, forward feature selection, kNN). To identify key points when the peaks and valleys were difficult to distinguish, they calculated proxy positions by finding the largest area of a triangle made by 3 equally spaced points along the GRF curve and setting the key point to the second point of the triangle. This might account for their substantially better results than others using the same key points.

Instead of identifying key points and basic statistical measures of the GRF profile, a number of studies holistically characterize the overall profile. The earliest GRF biometric study, Addlesee et al. (1997), used an HMM fitted to vertical GRF profiles to identify individuals. Their best reported result was a CR of 91% (15 subjects). Stevenson et al. (2007) modeled each subject's footsteps with a "characteristic" HMM and achieved an EER of about 19% (8 subjects). Vera-Rodríguez et al. (2007, 2008) also extracted "holistic" features for each footfall, made up of the principal components of the heel and toe sensor signals and the GRF profile derived by integrating over time. This achieved slightly better performance (EER of 13%, 41 subjects, PCA, SVM) than using key point features (EER of 17%). Derlatka (2013) used DTW on GRF profiles (vertical, lateral, and fore-and-aft) and achieved an FAR of 4.0% and FRR of 2.2% (132 subjects, kNN). Later, Derlatka and Bogdan (2015a) broke the GRF stance phase into 5 subphases and achieved a maximum CR of 97.38% (200 subjects, DTW, ensemble of kNNs). Cattin (2002) used the power spectral density (PSD) of the GRF profile gradient as features and achieved an EER of 5.3% (20 subjects, GPCA). Using a more "realistic" data collection setup, they achieved slightly worse results (EER of 9.4%).

A few very effective GRF-based studies have employed wavelet packet (WP) features to characterize the GRF profile. Moustakidis et al. (2008) extracted WP features for all three GRF dimensions on a dataset containing walking speed and load carriage covariates. When the gallery and probe data sets contained examples of all covariates, the average CR was 98.2% (40 subjects, fuzzy-based feature selection, GK-SVM). The approach held up well when the gallery contained only data for the normal-speed plus nothing-carried case but probes were taken from all seven variations on walking speed and load carriage. The lowest CR was only 82.7%. Mason et al. (2016) followed a similar WP approach as Moustakidis et al. (2008), except that they used eight rather than three GRF components. Of the eight components, four were vertical measures from the four force plate sensors, two were measuring lateral forces from pairs of sensors, and two were fore-and-aft forces. They achieved a low EER of 1.28% (10 subjects, lege04 wavelet). Yao et al. (2010) used GRF profiles from three dimensions, each profile consisting of a sequence of 5 steps. They also extracted WP features and achieved CRs between 91% and 97% (103 subjects, fuzzy-based feature selection, SVM). An analysis of the misclassification showed that the remaining confusion was significantly due to the similar masses of certain subjects.

Nevertheless, mass as calculated from the GRF profile is discriminative. Yao et al. (2010) measured each subject's mass and normalized the GRF curve by it to remove its influence, and classification rates dropped by 25% or more. Addlesee et al. (1997) calculated the apparent mass as the average GRF over a single step and normalized the GRF curve by it to see classification rates plummet from 80 + % to 50% and below (15 subjects, HMM). Jenkins and Ellis (2007) evaluated body mass alone as a means of identification. They computed "steady state" body mass as an average of the sum of both foot pressures, not including the first and last quarters of the signal to reduce error. This single feature achieved a CR of 50% (43 child subjects, 1NN), whereas they found that Orr and Abowd, 2000 features on this data achieved only a 36% CR (1NN). Also, when the data was normalized by each subject's average standing force, Orr and Abowd's features only achieved a 17% CR.

3.2.2. Pressure mat features

Nakajima et al. (2000) provided the first investigation of mat-based features for biometrics, although the measurements were recorded from stationary subjects rather than moving ones. They normalized the position and orientation of a pair of footprints and used direct template matching to achieve a CR of 82.64% (10 subjects, 1NN). Since then, a number of studies have been conducted on pressure mats of various sizes and sensor resolutions, whether home-grown or commercial-off-the-shelf. Some involve simple arrays of on/off switches, while others measure actual pressures. Unlike the force plates that are typically used to measure GRF, pressure mats offer the ability to compute spatial features of the footstep as well as between-footstep features commonly used in gait analysis (e.g., stride length, toe-out angle). Mat-type sensors are also able to better separate the contributions of multiple individuals than the force plates of GRF studies (Morishita et al., 2002). Also, sufficiently large mats avoid the need to aim footfalls, and thus avoid the need for unnatural movement or explicit user cooperation. A sampling of the studies we review below are listed in Table 5, giving measures similar to previous tables but also indicating whether the participants walked barefoot or wore shoes. Barefoot walking in this context appears to be more discriminative than shod walking.

Arrays of binary switches provide useful, though basic, spatial information, as shown in Fig. 7. Middleton et al. (2005) evaluated three features extracted from such binary profiles: stride length, single step period, and the ratio of time spent on the heel vs. time spent on the toes. In total, these features gave an 80% CR (15 subjects, 1NN by class means). The ratio feature alone achieved a 60% CR. Yun et al.'s (2003,2005) features included the binary activation image summed over time, the computed location (centroid) of each step, and the number of sensors activated by each step. In Yun et al. (2003), the most discriminative of these were the computed footstep locations over a walk across the floor, achieving a CR of 92.8% (10 subjects, MLP). The raw (binary) data was also informative (84%). The resolution of the sensors in a follow up paper (Yun et al., 2005) was much higher

Table 5
Representative pressure mat feature sets.

| Feature set	Performance	Subjects	Footwear	Reference
Stride length, period, heel vs. toe period	80% CR	15	Bare	Middleton et al. (2005)
Footstep locations	92.8% CR	10	Unspecified	Yun et al. (2003)
Center of geometry path	91.4% CR	11	Bare	Jung et al. (2004)
Center of pressure path	≤ 1.3% EER	11	Both	Takeda et al. (2011)
Max pressure image	99.8% CR	104	Bare	Pataky et al. (2011)
Footprint energy image	4.2% FRR, 0.2% FAR	24	Shod	Cho et al. (2012)

Low-Resolution Switch Array Images
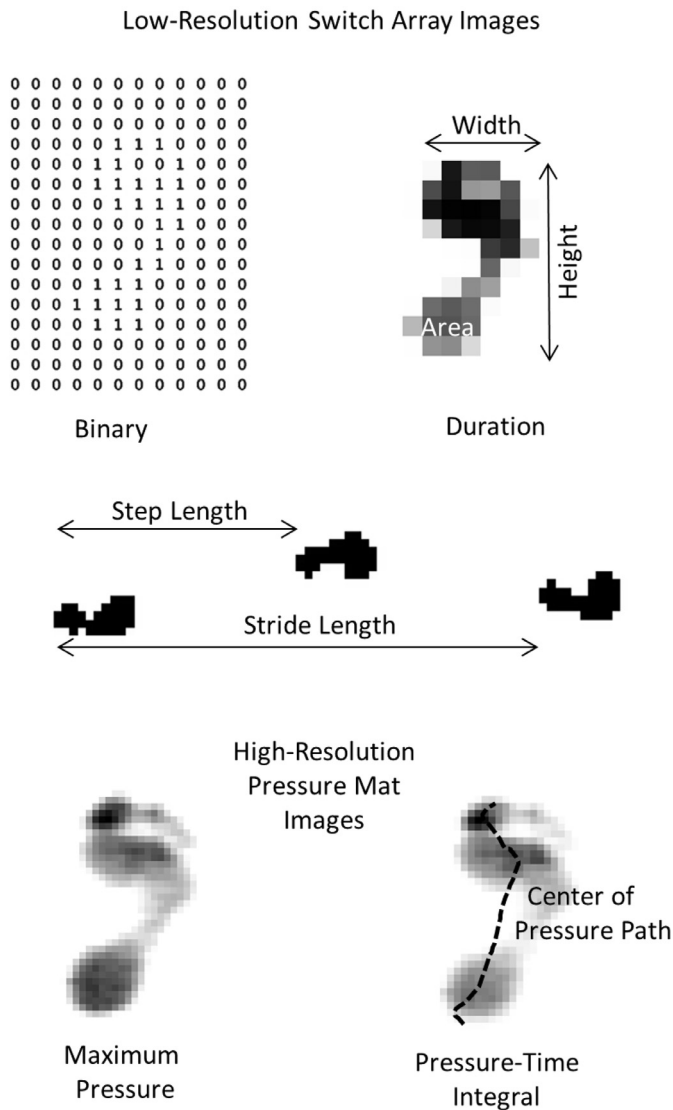
Fig. 7. Gait features derived from pressure-sensitive mats. The mats may use simple binary sensors, from which foot dimensions, step duration, and between-step information can be extracted. Sensor arrays that extract pressure values can also be used to compute the center of pressure and pressure images, such as the maximum pressure image and the pressure-time integral image.

(< 2 cm between sensors). They found that the computed location was again discriminative (89%, 10 subjects, MLP) and that adding the time of heel-strike and toe-off raised the CR to 96%. Suutala et al. (2008) used an irregularly shaped tile flooring of binary switches, each sensor covering a 10 cm × 10 cm area, to perform identification. Like Middleton et al. (2005), they acquired both binary and duration-based images from the switch array. Footstep-level features included footprint area, length, width, centroid, and the duration minimum, maximum, mean, and standard deviation. Footstep shape features based on line detection and Sobel gradient filters were also computed. Gait cycle-level features included time, Euclidean distance, and axis-aligned distance between adjacent footsteps. Based on footstep features, a CR of 64% was achieved (9 subjects, GP). Adding gait cycle-level features improved the CR to 81%. Overall, they determined that the between-footstep features were more informative than their low-resolution footstep features, although footstep shape was significantly discriminative.

High-resolution pressure-sensitive mats can add center of pressure (COP) or geometry features to the above repertoire, as illustrated at the bottom of Fig. 7, although the same can be done to a certain degree

with binary switch arrays. Jung et al. (2004) used the center of geometry as characterized by a hidden Markov model (HMM) based on data from a switch array and achieved a 79.6% CR (11 subjects). They also used aligned binary images of a pair of feet (which they call the "overlapped foot shape") and got a 91.4% CR. Adding the COP features to the binary image features increased the CR to 98.6%. Takeda et al. (2011) employed a high-resolution (5 mm × 7 mm) pressure mat to collect gait data, acquiring barefoot and shod data from 11 subjects. They extracted the normalized COP, which was the COP at nine equally spaced times along the full COP trajectory. They also extracted foot width, length, toe-out angle, total area, the areas of several zones of the foot, the normalized maximum pressure position over time (9 features), and the normalized total area over time (9 features). Using a fuzzy similarity metric, EERs for gallery and probe data with the same footwear condition were 1.3% (11 subjects) and below. EERs increased to between 12.5% and 28% when the gallery did not contain instances of the same footwear condition as the probe data. Qian et al. (2010) also extracted COP and key point features from a high-resolution pressure mat. From the COP curve, they computed the "turning point" of the lateral COP trajectory. They also computed the GRF profile by taking the pressures along the COP curve and extracted key point features, which were the maximum heel and toe peak pressure points, the minimum between them, and the end point. For each of these, they used the normalized position, pressure, curve length (distance along curve from the start of the COP to the key point position) and normalized time (relative to gait cycle length) as features. Finally, they included the mean pressure (i.e., their measure of mass) and stride length. Their best result, a 92.3% CR (11 subjects, FLD), included all of these features for a pair of left and right footsteps.

High-resolution pressure mats provide 2D pressure image information. For each X–Y pressure mat image location, Pataky et al. (2011) computed various pressure percentiles (100% or maximum pressure, 90%, 80%, 70%, 60%, 50%), the pressure-time integral (sum of pressures over footstep duration), the time to first contact (relative to heel strike), the time to maximum, and the contact duration. On their barefoot pressure data, the single most effective pressure image features were the maximum pressure and pressure time integral (both shown in Fig. 7), achieving as much as 99.0% and 99.6%, respectively (104 subjects, 1NN). This work also showed that careful foot alignment (orientation and translation) can have a significant impact on pressure image-based CRs, by as much as 6% in this study. The total best performance was 99.8% (104 subjects, LE, 1NN). Cho et al. (2012) also investigated the use of contact duration features, which they call the "footprint energy image", for shod gait recognition. Using contact duration images in the form of Jung et al.'s 'overlapped foot shape" (Jung et al., 2004), they achieved a 7.1% FRR and 0.3% FAR (24 shod subjects, kNN). Taking this further, they used the barefoot impressions of each subject to create weighted filters that focus recognition on the areas directly below the barefoot, rather than the below the shoe-only areas. This improved results to a 4.2% FRR and 0.2% FAR. Note that subjects wore different footwear from one another, but wore the same individual footwear for all recordings. Vera-Rodríguez et al. (2013a) expanded earlier GRF work by employing two medium-resolution pie-zoelectric sensor mats, one for each foot in a stride (88 sensors per small mat), known as the Swansea Footstep Biometric Database (Vera-Rodríguez et al., 2011; 2013a). They computed raw GRF profiles and "contour features", which capture the maximum positive and negative changes in pressure among the sensors at each time step. They also computed the pressure-time integral after a series of image processing steps. Under the various experimental conditions evaluated, they achieved an EER between 5–15%(40 subjects, PCA, SVM) for temporal and spatial features separately and a 2.5–10% EER for their fusion.

Conveniently most of the features that can be computed from force plates and binary switch arrays can also be computed from high-resolution pressure mats, the exception being features based on the lateral and fore-and-aft GRF dimensions. Connor (2015) compared many of the

above listed features for both barefoot and shod datasets from the CASIA-D gait database (Connor, 2015; Zheng et al., 2011, 2012). For barefoot data with two different walking speeds, the most useful features were primarily pressure image features and COP features. For shod data (one walking speed only), there were two scenarios: when the shoes between gallery and probe were the same or different. In the same-shod scenario, the pressure image features remained useful. In the different-shoe case, however, image features were generally poorer than features like step length, cadence, and GRF key point features. COP features remained among the most useful over all scenarios. Overall, Connor achieved high CRs of 99.8% (92 subjects, PCA, FFS, 1NN) and 99.5% (13 subjects) in the barefoot and different-shoe scenarios, respectively, when five footsteps of data were used for each probe. In the same-shoe case, rates reached 100% (15 subjects) with only one footstep needed for the probe.

3.2.3. Correspondence to gait mechanics

The underfoot pressure modality captures some static information about the walker. Overall mass gives rise to the average GRF value, and is static over the timescale of a gait recording. Mass appears to best computed from the average pressure over multiple footsteps (Jenkins and Ellis, 2007), possibly after trimming certain portions. Pressure mats capture the foot (or shoe) size and shape and a sense of the general load distribution tendencies, such as more heavily loading the heel or toes, etc. Pressure image features support this, collapsing the temporal dimension just as the average silhouette does in the visual modality. In fact, the contact duration image or footprint energy image is an average silhouette of foot contact with the pressure mat. Of these features, the PTI and maximum pressure features are most effective when the subject is barefoot (Pataky et al., 2011). Otherwise, duration images appear to be more useful (Cho et al., 2012), especially when the footwear is different between the gallery and probe (Connor, 2015).

Dynamic information, whether spatial or not, captures the time-varying forces placed by gait onto the walking surface. The subject first loads their heel with their weight and later launches off with the forefoot. The GRF curve captures a summarized, non-spatial force, regardless of whether it is characterized by key points or more generalized approaches. Of these, the generalized approaches, especially WP decomposition features appear to be among the most discriminating (Mason et al., 2016; Moustakidis et al., 2009; Yao et al., 2010). Pressure mats go a step further, effectively providing a large number of related mini GRF curves, which provide the shift of loading distribution or position with time. An effective way of representing this shift is the COP, which was found to be very helpful in both barefoot and shod contexts (Connor, 2015).

Pressure mats have the potential to provide the most accurate between-footstep measures. The stride-length, support base, toe-out angle, and cadence have shown here to be discriminative (Connor, 2015; Suutala et al., 2008; Yun et al., 2003), as also they were in the machine vision modality. Most between-footstep measures can be determined directly from foot positions, and a pressure mat's accuracy is limited only by its resolution (the current highest resolution is 5mm sensor spacing).

The dimension of the force features matters. The effective pressure mat features described above are based on forces in the vertical dimension only. This seems aligned with several GRF studies, as well. For example, Cattin (2002) evaluated the PSD of the vertical, lateral, and fore-and-aft dimensions and found the vertical dimension to be most discriminating (vertical EER of 5.3%, 20 subjects, GPCA, lateral EER of 6.6% and fore-and-aft EER of 10.5%). Moustakidis et al. (2009) and Yao et al. (2010) also found that the vertical dimension to be the most informative. They also suggested that the fore-and-aft dimension is more discriminating than the lateral dimension. In contrast, Mason et al.'s (2016) WP work found that more than half of the most discriminative components came from the fore-and-aft dimension and only a quarter came from the vertical GRF.

3.3. Gait features from accelerometry, audio, and other modalities

Besides the vision and underfoot-pressure-based modalities, gait as recorded from accelerometers, gyroscopes, microphones, and radar has also been shown to discriminate between individuals. Of these, the most prominent is the use of accelerometers which are common in modern mobile phones. This poses a slightly different use case than those considered so far, that is, the user cooperates with the biometric to some degree by carrying the phone. This modality is sometimes grouped with other wearable sensor gait recognition modalities such as gyroscopes (Ngo et al., 2014) and pressure sensitive insoles (Yamakawa et al., 2008). The most commonly mentioned application of this modality is for user authentication on such mobile devices (even smart watches (Johnston and Weiss, 2015)), for example, to lockout users with an unfamiliar gait. Anecdotal evidence proposes that it is also possible to identify individuals by the sound of their footsteps. Only a few studies have been conducted to test this hypothesis, but they have met with some success (de Carvalho and Rosa, 2010; Geiger et al., 2013; Shoji et al., 2004). Even Doppler-based approaches using modalities such as radar (Mobasseri and Amin, 2009; Otero, 2005), ultrasound Kalgaonkar and Raj (2007); Zhang and Andreou (2008), and wifi signals (Wang et al., 2016) have been used to recognize individuals by their gait.
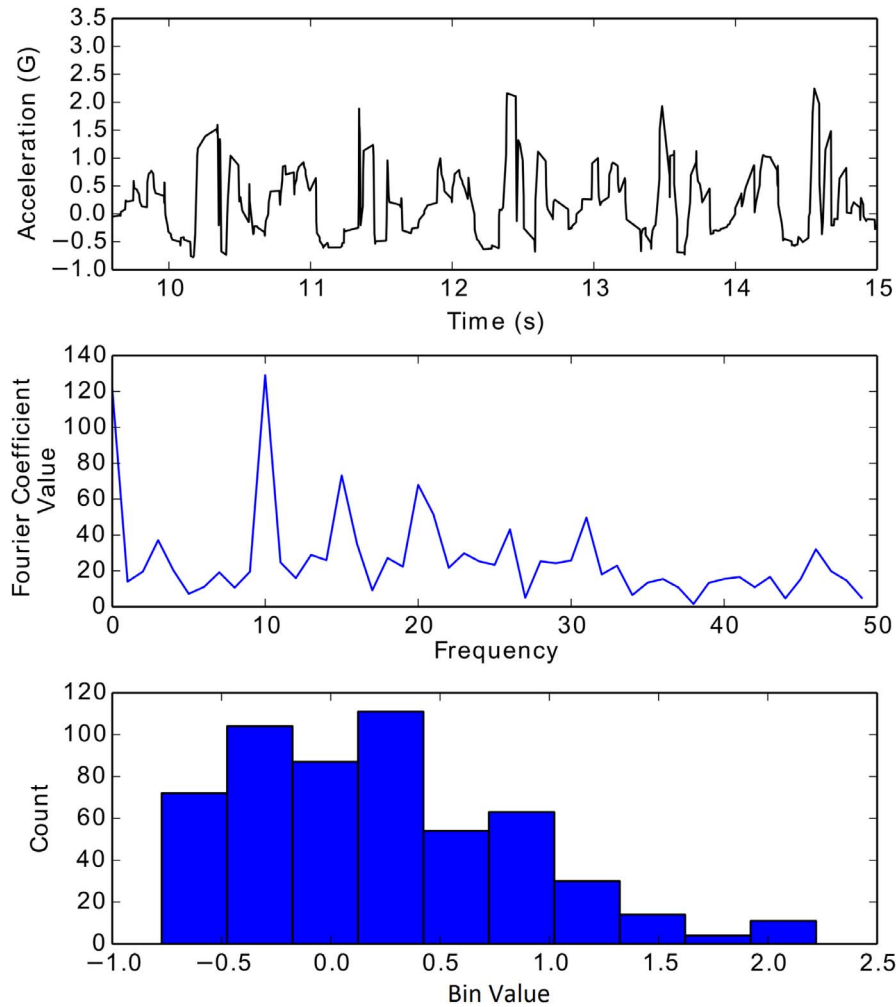
3.3.1. Accelerometry-based gait features

Acceleration, as opposed to velocity or position, is a direct function of the forces and masses involved in gait generation, since $F = ma$. It would therefore seem like acceleration would be a strong candidate for divulging discriminating features of gait. Table 6 captures a summary of different types of feature sets proposed for this modality. Unique to this table is the sensor placement location on the human body, which may affect performance.

The most straightforward approach to using acceleration for recognition, after preprocessing and gait cycle extraction, is to compare the accelerations of gallery gait cycles with probe gait cycles and to choose the closest match. Mäntyjärvi et al. (2005), an early acceleration gait recognition study, took this approach. Taking only the fore-and-aft and vertical dimensional acceleration signals, they composed left and right step templates as averages of all associated steps in the gallery gait

Table 6
Representative accelerometry feature sets.

Feature set	Performance	Subjects	Location	Reference
Direct acc. gait cycle matching	7% EER	36	Hip	Mäntyjärvi et al. (2005)
Direct matching, multiple templates	1.6–6.1% EER	30	Ankle	Gafurov et al. (2010)
DTW cycle matching	5.6% EER	21	Spine-base	Rong et al. (2007)
PCs of acc. gait cycle	1.6% EER	60	Hip	Bours and Shrestha (2010)
Acc. Fourier coefficients	10% EER	36	Hip	Mäntyjärvi et al. (2005)
Acc. value histograms	5% EER	21	Ankle	Gafurov et al. (2006)
Acc. inner product	6.8% EER	744	Pocket	Zhong and Deng (2014)
Acc. key points	95.8% CR, 2.2% EER	175	5 Places	Zhang et al. (2014)

Fig. 8. Gait features derived from accelerometry traces. *Upper Panel:* An accelerometry recording of five footsteps from a modern mobile phone. *Middle Panel:* Fourier coefficients of the recorded signal. *Lower Panel:* A ten-bin histogram of the acceleration values.

cycles for each subject. Then they used cross-correlation between the gallery templates and probe instances as a score by which to classify. This approach achieved an EER of 7% (36 subjects) and was superior to other features of the gait cycle they evaluated (discussed below). Using all three accelerometer dimensions, Rong et al.'s (2007) gait cycle matching approach achieved a slightly lower EER of 5.6% (21 subjects), employing DTW to normalize gait cycle lengths, compute average templates, and compute the matching metric. Derawi et al. (2010b) also used DTW to compare gallery cycle acceleration templates (8 to 15 templates per person) to probe gait cycles. The classification was based on the best cross comparison between an individual's group of gallery templates and the probe's gait cycles. This approach achieved an EER of 20.1% (51 subjects). Gafurov et al. (2010) also used a set of templates per person and performed cross comparisons, but used Euclidean distance as a measure of similarity rather than DTW. This achieved low EERs of between 1.6% and 6.1% (30 subjects) when the training and test instances involved a single type of shoes. In a second paper, Derawi et al. (2010a) altered their cycle comparison approach (Derawi et al., 2010b) to circularly shift the cycle for best fit combined with DTW and achieved a much improved EER of 5.7% (60 subjects). Bours and Shrestha (2010) made substantial gains by representing the average gait cycle template by its principle components, achieving an EER of 1.6% (60 subjects) that improved upon their best reported direct comparison approach that achieved 5.7% (based on Derawi et al. (2010a)).

Besides direct comparison of gait cycle signals, Fourier coefficients

and histograms of cycle signal values, as illustrated in Fig. 8, have also been evaluated. Mäntyjärvi et al. (2005) and Rong et al. (2007) found that Fourier coefficients were less discriminative than their direct gait cycle comparisons, yielding EERs of 10% (36 subjects, cross-correlation) and 21.1% (21 subjects, DTW), respectively. Regarding histograms, results were less conclusive. Mäntyjärvi et al. (2005) computed 10-bin histograms that were normalized by the length of the gait cycle signal and achieved an EER of 19%, much worse than their direct gait cycle comparison approach. Gafurov et al. (2006) derived histogram features in the same way as (Mäntyjärvi et al., 2005). Their results, however, were much better than Mäntyjärvi et al., yielding an EER of 5% (21 subjects). They suggest that the difference was due to a difference in the datasets' size and elapsed time between recording of gallery and probe data. It may also be due to the difference in sensor placement (their ankle placement, vs. Mäntyjärvi et al.'s hip placement), which noticeably alters the signal shape (Gafurov et al., 2006).

One challenge in using acceleration is that the orientation of the device (e.g., smart phone) may shift in transit or be carried differently on the person due to a change in clothing, situation, etc. Some studies have adjusted for this in part by this using the acceleration magnitude (square root of the sum of squares of each acceleration dimension) as the primary signal (e.g., Gafurov et al., 2007). Zhong and Deng (2014) proposed a similar but more general solution. They computed the inner product of a multi-dimensional acceleration or gyroscopic signal out of phase with itself to varying degrees. This mathematically eliminates the effect of rotation, and the acceleration magnitude used in other studies

Table 7
Representative acoustic feature sets.

Feature set	Performance	Subjects	Reference
MFCCs	60%, 42.3% CR	5, 155	Shoji et al. (2004) and Geiger et al. (2013)
Cadence	54% CR	5	Shoji et al. (2004)
Power spectral coefficients	52% CR	5	Shoji et al. (2004)
Signal energy	24.2% CR	155	Geiger et al. (2013)

is one form of this (a zero phase difference). From this convenient property, they generate a gait dynamics image, each row corresponding to a specific time difference panned over the length of the gait sequence. They show that these features are more effective than magnitude measures (the first row of their image), approximately doubling the identification rate from 32.5% to 66.3% (20 subjects) when the gallery and probe samples were taken on different days. On a large dataset (744 subjects, Ngo et al. (2014)), they achieved a low EER of 5.6% (fusion of accelerometer and gyroscopic images, 1NN). Even their 6.8% EER on acceleration data alone substantially outperformed the lowest EER of several other approaches on the same data set (13.5%, Ngo et al. (2014)).

Since extracting the gait cycle can be problematic, some approaches attempt to avoid it altogether. Zhang et al.'s (2014) approach automatically finds key points of the gait signature without having to identify individual gait cycles. They achieved a high CR of 95.8% (175 subjects, softmax, 1NN-like) and a low EER of 2.2% based on the accelerations from 5 sensors. Using a single sensor, they achieved a best CR of 73.4% (sensor at pelvis) and a best EER of 8.6% (sensor on thigh).

3.3.2. Acoustic gait features

Table 7 lists several types of feature sets used to automatically recognize people by the *sound* of how they walk. Shoji et al. (2004) recorded the audio of five subjects in slippers as they walked toward a microphone. From these recordings, they extracted three types of features as shown in Fig. 9: mel-frequency cepstral coefficients (MFCCs), cadence, and power spectral coefficients. Evaluated separately, these features achieved between 52% and 60% CRs (5 subjects, kmeans). When the cepstral and cadence features were combined, however, they reached a 100% CR. MFCCs and cadence were also evaluated by de Carvalho and Rosa (2010). In addition, they also considered cepstral coefficients, spectral envelope, and loudness as features. In this study, 15 subjects' data were divided into six groups according to footwear type (with overlap between groups). Here, the MFCCs appeared to be most effective, achieving as high as a 97.5% CR (4 subjects, kNN) in the "buskin" style footwear group and, when combined with the loudness of specific frequency bands, gave the highest CR in the all-footwear group (72.9%, 15 subjects, kNN). The MFCC and loudness features in the all-footwear group were separately able to achieve 55% and 59% CRs, respectively. The spectral envelope also performed relatively well, achieving a 46% CR when extracted using the Welch method. Geiger et al.'s (2013) acoustic gait recognition approach was evaluated on the TUM Gait from Audio, Image, and Depth (GAID) database (Hofmann et al., 2014). They evaluated energy, loudness, spectral (e.g., zero crossing rate, spectral roll off points, spectral flux, sharpness, etc.), and MFCC features. These features were collapsed into constant length vectors per recording using a variety of what the authors called "functionals", including mean, standard deviation, and 40 others. They found that the best features were the signal energy, loudness, spectral kurtosis, spectral skewness, spectral flux, and the first MFCC, together achieving a maximum CR of 51.9% (155 subjects, SVM). As separate groups, the MFCCs, spectral, and energy features achieved 42.3%, 33.2%, and 24.2% CRs respectively.

Instead of recording the footstep sounds passively, the Doppler-shift in an ultrasound transceiver has also been used to characterize gait for

recognition. Kalgaonkar and Raj (2007), performed a form of FM demodulation on the ultrasound reflections to essentially extract a map of the velocities of the walker. From the demodulated Fourier spectra, they computed 40 cepstral coefficients and achieved a CR of 91.7% (30 subjects, GMM). They also attempted recognition without the FM demodulation step (similar to features from a radar-based gait recognition study, Otero, 2005) and the CR dropped to 72.0%.

3.3.3. Correspondence to gait mechanics

As expressed earlier, acceleration information is a direct function of the combined forces and masses involved in gait generation, and captures primarily dynamic information. Studies that compare the gait cycle accelerations achieved among the lowest EERs, and were usually under 10%. To achieve this, however, a careful gait cycle segmentation process seems necessary, and representing the cycles by principal components appears to help substantially (Bours and Shrestha, 2010). It is tempting to think that audio as a gait biometric modality captures more information about a subject's footwear than about the subject themselves. In most of the studies, including the high CRs of Geiger et al. (2013), relative to the number of subjects, the footwear varied among the subjects. Yet, in Shoji et al.'s (2004) study, all five subjects wore slippers, and cepstral coefficients still achieved above-chance discrimination. Even in Geiger et al., there were 155 subjects and so there may have been some overlap among the subjects' footwear. Therefore, it is possible that the audio records a sense of the intensity of major transient underfoot events (e.g., impact force of the heel or the degree of slap beneath the forefoot). Audio can also indicate the cadence of the subject, based on the time between audible footsteps. The Doppler-shift based studies, whether from ultrasound or radar, record a sense of the subject limb and body movement velocities, which again indirectly reflects the forces and masses involved in gait generation.

Sensor location will affect the acceleration wave form (Gafurov et al., 2006). According to reviews of the acceleration literature (Derawi et al., 2010b; Youn et al., 2014), aside from the low EERs in Gafurov et al. (2010) which used ankle sensor placement, the waist or pants pocket is the most consistently discriminative sensor location. Gafurov et al. (2006) also evaluated the different acceleration dimensions, and reported that the fore-and-aft direction gave the lowest average EERs (over all footwear), but all dimensions give quite similar results. In contrast, Muaaz and Nickel (2012) noted that the vertical acceleration signal was more discriminative than the fore-and-aft and lateral dimensions.

4. Vulnerability to covariates in biometric gait recognition

Laboratory and restrictive experimental conditions can sometimes lead to recognition results that are optimistically biased. In real-world circumstances, there may be substantial variability between the gallery and probe samples of an individual, due to them being collected in different sessions. As diagrammed in Fig. 10, variations can include walking speed, clothing, footwear, load carriage, and walking environment. Such "covariates" are external to the subject and may mask or distort the sensing of the natural gait parameters or may impose a constraint on how people walk. Over time, a subject may also undergo internal physical changes due to injury, pregnancy, or advanced age, which can alter the underlying gait generation machinery. Although the literature is silent on recognition methods to address this latter concern, a few studies have demonstrated stability in recognition rates over several months. Below, we review the external covariates, one at a time, describing how each relevant modality is affected by them and solutions offered in the literature. Table 8 summarizes this by placing "Y" symbols in proportion to the effect a covariate has on a specific sensory modality. Note that there is less data for the underfoot pressure and the accelerometry modalities on which to base this summary.
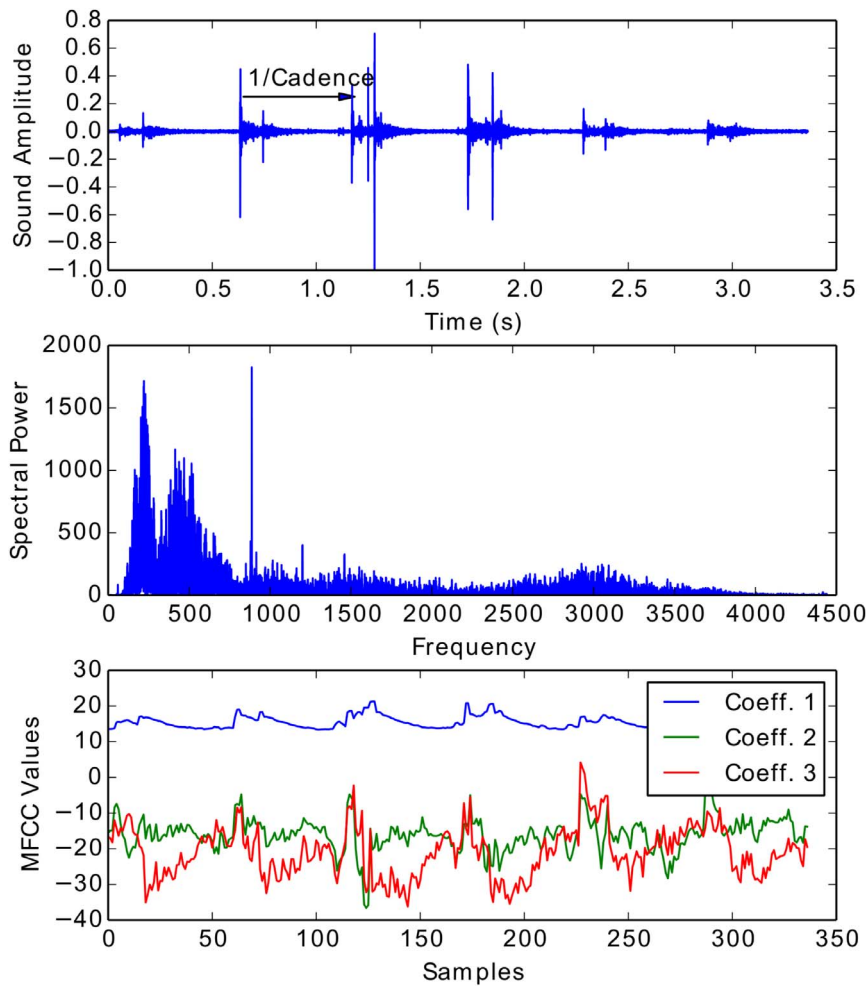
Fig. 9. Gait features derived from audio recordings of footsteps. *Upper Panel:* An audio recording of five footsteps as the walker approaches and passes by a microphone. *Middle Panel:* The spectral power of the recorded signal. *Lower Panel:* The first three mel-frequency cepstral coefficient signals.

4.1. Walking speed

Walking speed impacts recognition accuracy across all modalities. A change in speed is the result of a change in the average force or effort invoked during walking and leads to changes in all time-based features. Spatial features are also at risk of alteration. For example, it is known that step length, and thus maximum joint angles, will change relative to walking speed (Han, 2015; Tanawongsuwan and Bobick, 2003b). The peak values of ground reaction force also appear to vary with speed. Nilsson and Thorstensson (1989) showed that the m-shaped amplitudes shift with increasing speed, with higher amplitudes on the heel-strike and toe-off peaks and a lower trough between them. While humans have a preferred walking speed and tend to stay within a small range of this speed much of the time, an individual's walking speed can be influenced by internal motivation and external factors including footwear, load carriage, walking surface, and even the length of the walkway (Öberg et al., 1993). Walking speed can have a very negative impact on recognition rates. For example, Bouchrika and Nixon (2008a) specifically evaluated the effects of several covariates in a model-based approach and found that walking speed had a large effect, dropping rates by as much as 35% (10 subjects, ASFFS, 1NN). Similarly, Kusakunniran et al. (2011) achieved an average of 96% for a small difference in gallery-probe speed (1 km/h difference, OU-ISIR gait database, 25 subjects, 1NN) and 68% for a large difference (4 km/h) in gallery-probe speeds based on features of the silhouette outline. There are several general approaches taken to lessen the effect of this covariate. One approach is to transform features to suit the apparent walking speed. Tanawongsuwan and Bobick (2003a) modeled stride length vs. walking speed and cadence vs. walking speed with a linear model. Then, they used the models to transform probe stride lengths and cadences to match the gallery walking speed and, finally, performed classification. This approach achieved CRs about as high as when the gallery and probes had the same speed. Another strategy they attempted (Tanawongsuwan and Bobick, 2004) was to eliminate the arm swing and stretch the width of the leg region in the double support phase silhouettes relative to speed. This increased CRs from an average of about 63% to about 81% (24 subjects). Mason et al. (2016) evaluated a number of GRF normalization techniques, including one that linearly stretched the GRF profiles to have the same duration, which is normally correlated with walking speed. This helped improve results slightly across all applicable feature types they evaluated. Another way to address the walking speed covariate is to train with data from multiple walking speeds. Moustakidis et al.'s (2009) GRF recognition results saw CRs increase for fast and slow speeds from 84.0% to 98.3% and 88.7% to 98.3% (40 subjects, WP, FuzCoC, LDA), respectively, when training included fast and slow speed samples together with the usual normal speed samples. Lastly, features that are less sensitive to walking speed have been employed. For example, Guan and Li (2013) applied the Random Subspace Method (RSM) to Gabor-GEI features (Tao et al., 2007) that had been reduced using 2DPCA. Each RSM descriptor was the projection of the features onto a random subset of the significant 2DPCA eigenvectors, which gave each descriptor a different (random) perspective on the data. On the CASIA-C database, they achieved high CRs of 99.7% and 99.6% (153 subjects, 2DLDA) for slow and fast speed
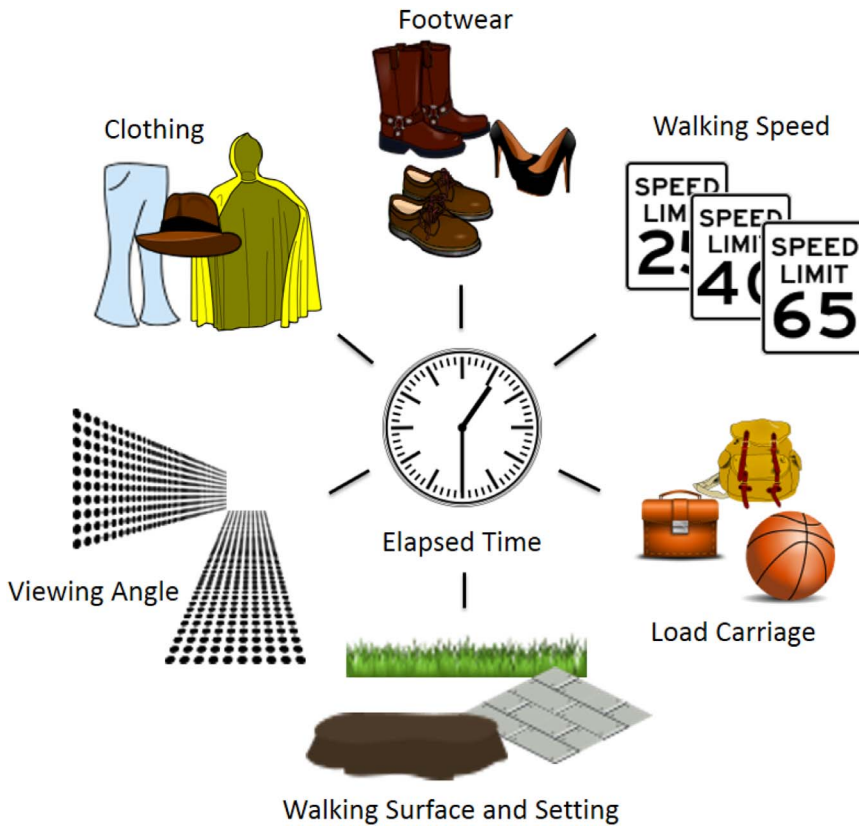
Fig. 10. There are a variety of gait covariates to consider, especially walking speed, which affects all gait sensory modalities and has a significant impact on recognition rates. Certain covariates have a larger impact in one modality than in another (e.g., clothing in vision-based vs. pressure-based modalities). Elapsed time may at first appear to be a highly influential covariate, but studies suggest that the other covariates that change with time contribute the most to identification errors, rather than a change in gait itself over time.

Table 8
Impact of covariates on various modalities.

Covariate	Model-free	Model-based	Pressure	Accelerometry
Walking speed	YY	YY	YY	?
Clothing	YY	Y		
Load carriage	Y	Y		?
Footwear	Y	Y	YY	YY
Viewing angle or placement	YY	Y		YY
Walking surface or setting	YY	YY	N/A	YY

probes given a "normal" speed gallery, substantially outperforming all previous methods. This improvement was not only due to RSM since their method without RSM (i.e., Gabor-GEI + 2DPCA + 2DLDA) achieved 97%. On the OU-ISIR-A dataset, where walking speeds ranged from 2 km/h to 7 km/h, the average CR for their method (train on one speed gallery and test on every speed) reached an impressive 98.05% (34 subjects).

4.2. Clothing

A person's clothing will change regularly and poses a significant challenge to vision-based and acoustic-based gait recognition, but perhaps has little to no effect on the other gait modalities. Clothing would seem to have very little effect on actual gait generation, except when it is particularly constricting or heavy (Bouchrika and Nixon, 2008a). However, it's shape can distort or occlude the appearance and sound of gait. Geiger et al.'s (2013) audio-based recognition study found that the wearing of shoe covers reduced CRs because of the audible noise generated by these objects during movement. Matovski et al. (2010) collected data on separate days, both times recording the same subjects walking while wearing their own clothes and while wearing a standardized clothing (overalls). The GEI-based CRs

were between approximately 18% and 70% (21 subjects, kNN) lower when the subjects wore their own clothes than when the subjects wore overalls, presumably because the subjects' clothes differed between days. According to a recent survey (Makihara et al., 2015), model-based methods are more clothing invariant than model-free methods. However, clothing changes still have a negative effect. In Bouchrika and Nixon's (2008a) model-based study, it was found that adding a normal coat did not reduce recognition rates much (5%), but that adding a trench coat that obscures leg appearance did reduce CRs by almost 30% (10 subjects, ASFFS, 1NN). In Yu and Zou's (2012) model-based study, the CR was highest when subjects wore a thin coat (93%, 124 subjects, NPE, HCRF, SVM), slightly less with a thick coat (89%) and less still with a backpack (85%). Several model-free approaches (Guan et al., 2012; Hossain et al., 2010; Islam et al., 2013) address this covariate by using groups of templates that focus on different parts of the silhouette. Then, templates which focus on familiar portions between gallery and probe samples can correctly identify the individual. These approaches were all evaluated on the OU-ISIR Treadmill database B, which included up to 32 clothing combination conditions for 68 subjects. The best CRs from this group of approaches were around 80%, which greatly improved over a simple GEI matching-based approach at 55%.

4.3. Load carriage

Recognition rates can drop when subjects carry an object, but this may be due more to a change in visual appearance than to a change in gait itself. In many situations, load carriage is common (e.g., airports, businesses, etc.) and therefore must be taken into account. The load itself will change the gait appearance and may also force the arm(s) to take a different pose than in natural gait (e.g., the ball-carrying covariate of the CMU MoBo database (Gross and Shi, 2001)). It will also add mass to the subject, which may affect gait accelerations. So far, it seems that the visual modality is affected the most, although, as far as covariates are concerned, the impact appears relatively small.

Bouchrika and Nixon (2008a) found that the addition of a hand bag or barrel-shaped gym bag dropped classification rates slightly (by less than 10%, 10 subjects, ASFFS, 1NN) and Sarkar et al.'s (2005) baseline results for the humanID gait database gave a 61% CR (122 subjects, Tanimoto-based similarity) for the briefcase covariate data, where the best CR was the shoe covariate at 78%. Underfoot pressures may be even less affected. Moustakidis et al.'s (2009) GRF study evaluated four carrying conditions (right hand, left hand, both, and backpack) at two different weights (5% and 10% of the subject's weight). Using WP features with a gallery that contained no loads (though multiple walking speeds), CRs for the load conditions averaged about 99% (40 subjects, WP, FuzCoC, LDA). As with other covariates, a natural approach to reducing the effect of load carriage is to seek features that are less affected by it. On the CMU MoBo database, Collins et al.'s (2002) approach of double and single support stance key frame comparisons achieved a high CR of 92% (25 subjects, 1NN) when the probes (but not the gallery) had the subject carrying the ball. Yet, a very similar key frame correlation approach in Veeraraghavan et al. (2009) only achieved 48% on the same dataset. A dynamic approach by Lee et al. (2007) based on self-similarity of images to a specified key frame achieved 77% (25 subjects, 1NN based on a special cost function). In the humanID gait database, the best results for the briefcase carrying covariate (93% (Liu et al., 2012), 122 subjects) were achieved using the Chrono-Gait Image (Wang et al., 2010). Another very effective approach (92%, 122 subjects, HOG-GEI, PCA, LDA) is the histogram of orientation gradients method (Liu et al., 2012), which captures the detailed shape information of the silhouette. If the negative effect of load carriage is due to a change in appearance, then strategies used to overcome the clothing covariate may be likewise effective.

4.4. Footwear

Changes in footwear is a challenge for all gait recognition modalities, but especially for the underfoot pressure modality. Although perhaps more stable from day-to-day than clothing, footwear can change regularly. Footwear can also affect gait generation. According to Lythgo et al. (2009), wearing shoes rather than walking barefoot was found to change gait by increasing walking speed (by 8 cm/s, 980 subjects, mostly children), stride length (11.1 cm), and support base (0.5 cm) while decreasing cadence (3.9 steps/min) and foot pose angle (0.1˚). Different types of footwear will define different interfaces between the foot and walking surface and thus differences in the distribution of pressure. Even the same shoe's pressure distribution may shift over a shoe's lifetime. The effect of footwear on recognition rates for the various modalities is mixed, but is generally low for vision-based approaches. Among the covariates of footwear, view, walking surface, and load carriage evaluated with the humanID gait database (Liu et al., 2012; Sarkar et al., 2005), the footwear covariate has had the least or second least impact of all covariates. Bouchrika and Nixon (2008a) found that some common footwear (boots, trainers) had little effect on recognition rates (less than a 10% drop, 20 subjects, ASFFS, 1NN), and thus gait generally. The exception to this was the use of flip-flop footwear, which had a drastic impact on recognition results, a drop of about 40%, which was perceived to be caused by the unfamiliarity of the subject group with them and the extra need to grasp the upper between the toes. Gafurov et al. (2010) evaluated the effect of different footwear in the accelerometry modality. All subjects wore four different pairs of shoes during the recording sessions. When gallery and probe instances involved only a single type of shoe, EERs were between 1.6% and 6.1% (30 subjects). When instances with all four footwear types were present, EERs increased substantially to between 16.4% and 23.6%. From the individual EERs for different pairs of footwear, the lighter pairs consistently gave lower EERs. In the underfoot pressures modality, footwear is a challenging covariate. Using barefoot data in the gallery and as probes can achieve very high CRs when high-resolution pressure mats are used (> 90% CR with large numbers of subjects (Connor,

2015; Pataky et al., 2011)). The more realistic scenario of footwear changing between gallery and probe acquisition sees substantially lower CRs. Takeda et al. (2011) evaluated recognition using the same footwear for all subjects–two pairs of slippers with different thicknesses. While EERs were under 2% (11 subjects, fuzzy similarity metric) when gallery and probe were recorded with the same slipper, EERs increased to 12% or more when the gallery and probe were recorded from different slippers. Connor (2015) achieved a very similar result for a variety of different footwear. In the same gallery-probe footwear case, EERs were 2.1% (15 subjects, PCA, FFS, 1NN) and below, whereas in the different footwear case, EERs rose to between 11.4% and 15.9% (13 subjects). A simple way of addressing this covariate, although perhaps impractical, is to include data from several different pairs of footwear for each subject in the gallery. When Takeda et al. included data from both types of slippers, EERs regardless of test slipper dropped to 2.1% and below. Another approach is to use features that are less sensitive to a change in shoes. Connor found that the types of underfoot pressure features that were most effective for the changing footwear case to be the COP (especially along the fore-and-aft axis), contact duration images, GRF key point features, and between-footstep features like cadence and step length. Cho et al. (2012) used barefoot impressions to provide a weighting filter that focused attention on areas directly below the barefoot, and improved results from 7.1% FRR and 0.3% FAR (24 shod subjects, kNN) to 4.2% FRR and 0.2% FAR.

4.5. Viewing angle

Changes in a sensor's "view" or orientation and placement with respect to the subject is particularly an issue for the vision and accelerometry modalities. In the underfoot pressures modality, the walker may cross the sensor surface at different angles, but the pressure image will be the same aside from a simple 2D image rotation. In contrast, the walker's orientation and distance to a camera, and the orientation and location of an accelerometer attached to an individual, may be different between gallery and probe. Having different views will lead to a larger variance in the feature values for the same individual, leading to larger overlap with others' data and a decrease in CRs. In the humanID gait database, where subjects walked in two different directions, the view covariate leads to some reduction in CRs, but only below the CR of the footwear covariate (Liu et al., 2012; Sarkar et al., 2005). However, the difference in viewing angle was such that "their view was approximately fronto-parallel" (Sarkar et al., 2005). In contrast, the CASIA-B gait database includes a series of eleven views covering a range of 180˚. On this database, Liu and Tan (2010) used GEI images as features for one of their baselines and included three spread-out views of data in the gallery. This yielded an average CR of 51.2% (124 subjects) across all views, whereas the average CR for probes with views in the gallery was 97.3%. There are a few possible ways to mitigate this covariate. One is to constrain the walking direction based on the setting. If, like the Biometric tunnel at the University of Southampton (Seely et al., 2008), there is a hallway or a narrow walking carpet or sidewalk to encourage a specific path, variability in view can be substantially reduced. Other solutions include either artificially reorienting the viewpoint or to using features less dependent on view. Seely et al. (2008) synthesized an average silhouette for a certain viewpoint from six cameras with other viewpoints and achieved a maximum CR of 99.6% (103 subjects, 1NN). Goffredo et al. (2010) estimated the view angle of the camera relative to the walker and "corrected" for it. With model-based features they achieved 73.6% on a subset of the CASIA-B database (65 subjects, only six view angles, kNN) and 95.8% on the SOTON small database (20 subjects). Hong et al. (2007) achieved a CR of 96.25% (20 subjects, 1NN) on the view-varying CASIA-A database by using a width-histogram style feature vector based on silhouettes. Liu and Tan (2010) radon transform-based approach, which took data from three viewing angles, was able to achieve an average CR of 90.7% (124 subjects, LDA) across all viewpoints, much better than the GEI-based approach at

51.2%. In accelerometry studies, the sensor orientation problem has been addressed by using either magnitude (Gafurov et al., 2007) or a phase-shifted image version thereof (Zhong and Deng, 2014), as described in Section 3.3.1. We are not aware of any published attempt to deal with variation in sensor location (e.g., hip, ankle), but between there being natural locations for certain devices and users being creatures of habit, devices will often be worn in a consistent location (e.g., smartwatch on the wrist, smartphone in a certain pocket).

4.6. Walking surface and setting

Changes in the setting, especially the walking surface, can have substantial effects on recognition performance in all modalities, except perhaps for underfoot pressures where the walking surface is constrained to be the sensor surface. From common experience, it would seem that gait, or at least walking speed, can be affected by walking surface (e.g., sand, snow) as well. Wagg and Nixon (2004) model-based approach achieved a CR of 84% (115 subjects) on indoor data and only 64% (115 subjects) on outdoor data. They explained that this drop was due to greater difficulty fitting their human model to the outdoor data. Muaaz and Nickel (2012) recorded subjects' accelerations while walking on a carpet, gravel, grass, and a grassed incline at a normal speed. Their base condition (normal speed on carpeted floor) achieved an EER of 16.3% (48 subjects, DTW), where classification was based on the similarity between a typical (of each subject) gait cycle and probe cycles using a majority voting approach. With gallery and probe instances having the same surface condition, they found that the carpet and gravel surfaces allowed for comparable recognition rates (EERs of 25.3% and 25.5%) but that grass and inclined grass walks made recognition more difficult (EERs of 35.9% and 36.8%). It may be that the grassy nature of the inclined surface is the primary cause for its increase in EER over a level carpeted surface rather than its inclined nature, since the two grassy EERs are very similar. The humanID gait database, where the grass surface is used in the gallery, saw a large drop in CRs when the probe was on concrete (46% lower than a change in footwear, Liu et al., 2012; Sarkar et al., 2005). A simple, though perhaps not always practical solution to avoid setting changes is to constrain the environment to be consistent in lighting, background, and walking surface. There may also be features that are less affected by walking surface. Liu and Sarkar (2006) normalized binary silhouette sequences improved results substantially over baseline results on the humanID gait database on the walking surface covariate test (from 32% to 57%, 122 subjects, 1NN). One setting-related issue that seems largely unaddressed in the vision-based gait recognition literature is occlusion, whether by other individuals or by objects. This will be a necessary issue to address if visual gait recognition is to be effectively used in busy or cluttered locations.

4.7. Elapsed time

The effect of elapsed time on gait recognition can appear drastic, but seems largely due to accompanying covariates. Matovski et al. (2010) hypothesized that performance drop was not only due to time, but a change in clothing, shoes, etc. that are different between recording sessions. Accordingly, they created a database with several elapsed periods (0, 1, 3, 4 , 5, 8 and 9 months) between gallery and probe instances. Using a GEI-based approach, they discovered that when clothing and footwear are standardized between recording sessions, recognition rates only dropped by 5% over time, part of which they attributed to a change in clothing worn under their standardized clothing (overalls). They also compared recognition rates between the standard and subjects own clothing and found a larger drop of between approximately 18% and 70% (21 subjects) depending on which dataset (overalls or subjects' clothing) is used as the gallery and the view direction(s) recorded. Using foot pressures, Pataky et al. (2011) compared data from barefoot subjects acquired between 1.5 and 5 years apart and

achieved a 94% recognition rate (gallery containing 104 subjects and probes from 10 subjects, LE, 1NN), a drop of about 6% from their same-day results with 104 subjects. In Jung et al. (2004), the gallery was recorded in the first month and the probes were recorded the next month, and they achieved a high CR of 98.6% (11 subjects). Vera-Rodríguez et al. (2013a) also used time elapsed data (collected over 16 months in separate sessions per user), the earliest data for the gallery and later recorded data for the probes. Their features were based on coarse underfoot pressure information and achieved low EERs between 2.5% and 10% (40 subjects, PCA, SVM). From these findings, it would appear that a year or so of time passing is among the least challenging of the covariates discussed here.

5. Gait spoofing and obfuscation

An important feature of a biometric is how robust it is against spoofing attacks. According to the various types of gait features discussed, a would-be impersonator should make their body shape appear similar to their target, and also walk in a similar manner. What few papers there are in the literature on gait spoofing suggests this to be a difficult task, even with extensive training. Gafurov et al. (2007) evaluated spoofing by comparing average gait cycles, where the gait cycle signal was the magnitude of gait accelerations over all dimensions. This approach achieved a 73.2% CR (100 subjects) and a 13% EER without spoofing attempts. They tested this system's vulnerability to incorrect verification when: 1) subjects briefly learned to mimic a friend/colleague with similar traits and 2) the claimed identity was chosen to be the one in the database with the greatest gait similarity. They found that intentional mimicry had a negligible effect on EERs (for 90 "attackers"), but that knowing the closest target in the database gave an attacker a somewhat higher likelihood of success (raising the EER to 25%, 100 subjects). Mjaaland (2009) and Mjaaland et al. (2011) studied gait spoofing using similar gait acceleration features, which achieved an EER of 6.2% (50 subjects, DTW) without spoofing attempts. In contrast to Gafurov's brief training, Mjaaland worked one-on-one with six would-be attackers over five one-hour training sessions each. In spite of this relatively extensive training regimen, which involved video and statistical feedback, none of the attackers were capable of consistently generating walks that would pass the verification process nor did they appear to improve over time. Hadid et al.'s (2012) spoofing evaluation was based on the average silhouette features of Veres et al. (2004). Accordingly, attacks were based on matching the subject's clothing and build. They found that wearing common clothing (overalls) and having a similar build increased the EER relative to the baseline (from 6% to 12% and 9%, respectively, 113 subjects, PCA, 1NN). The combination of attacks was slightly more effective, increasing the EER to around 18%. In a second paper (Hadid et al., 2013), with a different recognition approach, they achieved a lower EER of 14% (113 subjects, LBP-TOP, Gentle AdaBoost) on the same data under the combination of attacks. From the above studies, it seems that purposely and repeatedly circumventing a gait recognition system by imitation is either very difficult or impossible. In general, spoofing of gait could be approached in a visually similar way to that evaluated independently by Gafurov, Mjaaland, and Hadid. Basing features on the inaccessible underfoot pressures may further frustrate gait spoofing.

Can someone intentionally avoid being detected by a gait recognition system? Gafurov et al. (2007) found that when subjects mimicked others, the similarity to their own gait changed substantially (average CR drop of 50%, 90 subjects). Bouchrika and Nixon (2008a) studied footwear covariates and discovered that although boots and trainers had little effect on CRs, the wearing of flip-flops reduced them by 40% (20 subjects, kNN). These two findings would suggest that it is possible to willfully avoid detection by gait. The likelihood of this occurring might be reduced if sensors were implemented covertly or in a public location where a natural gait is likely to be used to avoid drawing attention to oneself.
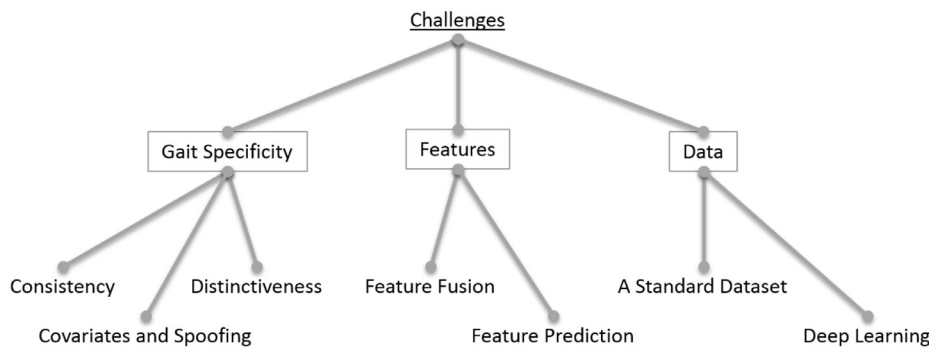
Fig. 11. Various challenges in biometric gait recognition can be categorized under three main topics: gait specificity, features, and data. Gait specificity refers to the challenge of distinguishing between individuals based on their distinctive gait attributes in the presence of noise (inconsistencies, covariates, and spoofing). Judiciously combining features across gait sensory modalities and with other biometric modalities (e.g., ear, iris, fingerprint) has the potential to enhance specificity and predict unseen gait features. Gait recognition is ripe for a new standard dataset that covers the gamut of gait sensory modalities, includes multiple recording sessions per subject, and is sufficiently large for data-hungry deep learning algorithms, which have had a significant impact in other pattern recognition domains.

6. Challenges in biometric gait recognition

There are a number of open problems and challenges in gait recognition that are yet to be thoroughly explored or addressed:

- assessing the distinctiveness of gait as a biometric cue;
- gauging the consistency of gait across time, emotional state, and pathological ailments;
- determining the effects of covariates and spoofing;
- effectively fusing gait features from multiple sensory modalities and with other biometrics;
- predicting the gait features in one sensing modality based on features from another modality;
- developing a comprehensive dataset that spans multiple modalities and covariates;
- exploring the role of deep learning in gait recognition.

These issues are categorized and diagrammed in Fig. 11 and discussed in the following sections.

6.1. The distinctiveness of gait

Aside from the observation of Murray et al. (1967) suggesting the existence of at least "twenty distinct gait components" (Nixon et al., 2010), it appears that little scientific work has been done to determine the actual distinctiveness of gait. The upper bound on gait recognition performance, therefore, remains unknown. Certainly there have been many papers on gait recognition that demonstrate improvement in performance on standard datasets. However, in some cases, it is possible that these approaches are dependent on information other than gait (e.g., clothing or other covariates that change between days, but testing is done with data captured the same day). Having an upper bound would help assess the role of ancillary information in optimistically improving the performance of "gait" recognition. A known bound would also help us determine the viability of using gait as a biometric in specific real-world applications. Perhaps a first step in this direction would be to collect data from a wide range of individuals using video, a marker-based motion capture system, and a pressure mat. From these, a variety of gait parameters could be computed to determine between- and within-subject variance besides agglomerative statistics.

6.2. The consistency of gait

The degree of stability or permanence of a biometric trait from the time of enrollment to the time of recognition is an important consideration. From Section 4.7 we saw that short periods of time (up to 5 years) appear to have a low impact on gait recognition performance. However, with extended periods of time this could change, due to the onset of advanced age (Maki, 1997; Winter et al., 1990), injury, change in weight, pregnancy (Foti et al., 2000) or pathology (e.g., Parkinson's

disease). While some of these conditions may be addressable, there is no known gait recognition approach that seeks to overcome such changes.

6.3. The effects of covariates and spoofing

In Section 4, we reviewed the growing body of work on gait covariates. Most of these, however, are focused on vision-based approaches that address the factors of walking speed and view angle. Covariates can have a large negative impact on recognition scores, and thus would appear to be the greatest single obstacle to the commercialization of gait recognition technologies. While Section 5 suggests that gait spoofing is difficult and potentially risky to the attacker, such studies are few in number. Additional investigation is warranted, especially across multiple sensory modalities to determine which is least vulnerable to obfuscation and spoofing whilst being robust to covariates.

6.4. The fusion of gait features across sensory modalities and with other biometrics

There are a number of examples of combining gait features from a single modality with other biometrics (e.g., Geng et al., 2010; Kimura et al., 2014; Muramatsu et al., 2013; Shakhnarovich et al., 2001; Vera-Rodríguez et al., 2012; Vildjiounaite et al., 2007; Zhou et al., 2007). Depending on the application (e.g., smartphone security), a particular combination of biometrics (e.g., face and gait from accelerometry) may be advantageous. By combining gait with other personal traits, the identity of an individual can be established at a distance in applications such as airport security.

In some controlled applications, the combination of features across multiple gait sensory modalities could be beneficial. There are a few studies that attempt this (Castro et al., 2015; Cattin, 2002; Derlatka and Bogdan, 2015b; Lee et al., 2013; Vera-Rodríguez et al., 2013b; Zheng et al., 2012) but, only two clearly show the benefit of combining features from multiple gait modalities. Cattin (2002) combined ground reaction force underfoot pressure features with "histogram" visual features to achieve a low equal error rate (EER) of 1.6%. However, Cattin achieved a still lower (better) EER of 0.9% using the histogram features alone. Lee et al. (2013) achieved a 51% classification rate (CR) using "gait energy image" visual features alone, 77% using underfoot pressure image features alone, and 85% using both in combination. However, their underfoot pressure features came from a different cohort of subjects than the visual gait data, so it is possible that the combination is synthetically unique, thereby unnaturally boosting results. Zheng et al. (2012) used underfoot pressure images to retrieve potential matching candidates from a database, which were further narrowed down using sophisticated visual features. Without the help of the underfoot pressure features, the visual features only achieved a 39% classification rate (88 barefoot subjects). Together, they achieved a high CR of 99%. However, they did not evaluate the effectiveness of underfoot pressure features alone, which are typically highly discriminative (in the barefoot case). In other studies (Connor, 2015;

Pataky et al., 2011), CRs of over 99% have been achieved on similarly large databases of barefoot subjects based on underfoot pressure features alone. Derlatka and Bogdan (2015b) used underfoot force over time to first verify a claimed identity and then rejected the positive verifications when any of a set of five measured body size parameters (e.g., height) did not fall within 5% of the system's stored values. Although this helped to reject several incorrectly verified cases, it rejected substantially more correct verifications. Castro et al. (2015) (see also Castro et al., 2016) added audio features to visual features, showing increases in classification rates between 3.1% and 9.4% (32 subjects, FV, PCA, SVM). Finally, Vera-Rodríguez et al. (2013b) combined gait energy image visual features extracted from the lower-body over a half gait cycle with ground reaction force and spatial underfoot pressure features. The visual features achieved an 8.4% EER (122 subjects) and the underfoot pressure features achieved a 10.7% EER. Together, they achieved a much improved 4.8% EER.

As a product of this review, we propose a fusion of features across modalities for future research. Note that some signals (e.g., acceleration) may be derived from other gait sensing modalities (e.g., machine vision model-based joint positions). Thus, the discussion below is not confined to specific sensors since similar information may be gleaned from other novel sensing modalities in the future.

It would appear that a combination of features from the visual and underfoot pressure modalities, with insights from accelerometry, will be the most discriminating. This is suggested because the visual and underfoot pressure views capture largely separate aspects of gait. Visual features capture a macro view of the body structure and movement, whereas underfoot pressure receives more detailed information about the interface between the foot and walking surface as well as the foot-level spatiotemporal forces at work. Besides offering complementary views of gait, a combination of these modalities may work better together because they are negatively impacted by different covariates. Visual features rarely focus on the shape and size of the foot and, thus, do not suffer as much from the footwear covariate as do underfoot pressure features. Also, underfoot pressure features are presumably oblivious to clothing (and perhaps to load carriage) that does not affect actual gait, and do not suffer from view point variations or occlusions. This latter quality makes multiple subject gait recognition a more viable possibility. Also, pressure mats can provide accurate foot location and gait cycle timing information for vision-based approaches that require gait cycle segmentation or assistance in determining view direction, walking speed, etc. Compared to model-based vision methods, accelerometry-based gait recognition methods appear to have evaluated various features and cycle matching approaches based on *acceleration information*. Several accelerometry-based studies, which have achieved results comparable to high-performing visual and underfoot-based features, were based on a single recording location. Since model-based methods implicitly have access to accelerations at multiple body locations, employing accelerometry-based approaches to model-based gait recognition may be quite helpful. Most video sequences, however, usually have a lower sampling rate (e.g., 30 frames/s) than accelerometry recordings (e.g., 100 Hz), but some high-performance video cameras can match this, if necessary.

Although potentially useful, the audio and Doppler-shift based features do not seem to add much to the combination of vision and underfoot pressures modalities. The footstep sound is generated by the impact of the foot with the ground and would seem to be a function of the underfoot pressure profile. It also seems very susceptible to noise, which is difficult to control in most practical applications. The Doppler-shift velocity features would seem to overlap with the limb velocity of model-based features or velocities extracted from accelerometry data.

A recommended combination of types of features extracted from video and high-resolution pressure mat data would include some or possibly all of the following:

- *Static features capturing structure and mass.* Body segment lengths

seem most accurately captured by model-based approaches. Averaged or key frame silhouette approaches, and possibly width histograms, seem like effective mechanisms for capturing the subject's relative mass and its distribution. The shape of the silhouette may capture additional build information. With silhouette-based features especially, approaches that address the clothing covariate will be helpful. High-resolution pressure mats can be used to extract the average dimensions of the foot, and the shape of the foot if barefoot. Pressure-based images can also provide a summary of underfoot mass distribution. Total body mass (including any load) can be calculated from GRF data (derivable from high-resolution pressure mats).

- *Dynamic features capturing forces and movement.* GRF curves report the change in overall vertical force over the footstep, and COP curves provide the underfoot trajectory. Contact duration images provide some footwear invariance and capture a spatiotemporal trace of the footstep. Model-based joint position trajectories relate gait movement and can be used to compute joint accelerations, which is where features from accelerometry-based studies may prove valuable.

- *Between-footstep features capturing high-level gait cycle information.* For this, certain high-resolution pressure mats can capture to within 5 mm and 0.01s of the location and timing of heel strike and thus give an accurate assessment of step length, step width, and cadence. It also has a direct view of the foot pose angle, relative to other footsteps on a path.

The inclusion of more gait cycle/footstep instances per classification should also help improve accuracy. This is one biometric trait where the use of additional information does not burden the user because it does not require explicit cooperation. Suutala and Röning (2005) showed an increase in recognition as the number of footsteps used to build a probe was increased. Using five footprints instead of one increased the CR from about 80% to 95% (11 subjects, MLP, product and sum rules). Connor (2015) used this strategy as well, bringing rates from 91.3% (one footprint) up to 99.8% (five footprints) in barefoot gait recognition (92 subjects, PCA, FFS, 1NN) and from 90.5% to 99.5% for shod gait recognition (13 subjects). Perhaps a similar approach might benefit vision-based approaches, such as taking multiple average silhouettes (one per gait cycle or half-gait cycle).

6.5. Cross-modality prediction of gait features

An interesting line of inquiry would involve predicting gait feature sets pertaining to one modality (e.g., vision), based on the features extracted from another modality (e.g., accelerometry data). This would entail developing models that can describe the correlation between features extracted from multiple modalities as well as determine the complementary nature of the associated feature sets. As alluded to earlier, the distinctive attributes of an individual's gait may be distributed across modalities and fusing these modalities can, therefore, provide a more comprehensive description of the individual's gait. While such a methodology may not have immediate practical applications, it would allow researchers to establish upper bounds on gait recognition accuracy and would help in understanding the information entropy of this biometric cue.

6.6. The need for a large multi-modality, multi-covariate standard dataset

Early in the growing field of gait recognition, DARPA's "HumanID at a Distance" research program spurred the community to gather appropriate real-world datasets, one of which was widely used by the vision-based gait recognition community. While this most popular dataset covers a variety of covariates and has 100+ subjects, it lacks the ability to move gait recognition from laboratory to commercialization. Large vision-based datasets (1000+ subjects) have been collected in

recent years, but the challenge is having subjects return for a second session, where real-world covariates come into play. A valuable contribution to the gait recognition community would be to generate a database spanning multiple gait sensory modalities, covering a wide range of covariates, and having large numbers of subjects that return for subsequent sessions after a prolonged time in order to facilitate longitudinal studies. This will be especially important for algorithms that are dependent on such large data sources such as those based on deep learning.

6.7. The role of deep learning in gait recognition

A trend in a number of pattern recognition domains is a move toward deep learning techniques, which avoid specifically handcrafting feature extraction methods by finding discriminating regularities in the raw data. Deep learning has begun to penetrate the gait recognition community in each of the major modalities (e.g., Gadaleta et al., 2016; Shiraga et al., 2016; Wolf et al., 2016; Wu et al., 2015; Zheng et al., 2010). For example, Shiraga et al. (2016) used a convolutional neural network (CNN) on a subset of the OU-ISIR Large Population dataset with the average silhouette or GEI as input. This approach consistently outperformed other methodologies compared in the paper in terms of cross-view EERs of between 1.2 and 2.5% (956 subjects, CNN), although their cross-view CRs ranging 80.4 to 94.8% were not always best. Each cross-view test held only one view in the gallery and all other views (different by a 10, 20, or 30° viewing angle) as probes. Zheng et al. (2010) used a deep learning-like approach to extract features from barefoot data captured with a high-resolution pressure mat. An FRR of 0.1% (88 subjects, DHSC) and FAR 0.9% was achieved. Using a CNN, Gadaleta et al. (2016) extracted 40 features from accelerometry and gyroscopic data that had been segmented into gait cycles. Their approach achieved FARs and FRRs of under 1% (9 subjects, CNN, PCA, SVM, SPRT) usually combining data from five or fewer gait cycles.

Just as deep learning has become the state of the art in several pattern recognition domains (e.g., speech recognition and object recognition), it may likewise become useful in gait recognition. The present survey has advocated using features from multiple modalities, and has loosely specified a complementary set of such features. Perhaps, rather than collecting a set of handcrafted features from a multimodal gait dataset, a deep learning approach may discover a more effective combination of multimodal features or provide a new and more discriminating view on gait data than do the features reviewed herein.

7. Summary

From a developmental standpoint, automatic gait recognition is still in its infancy compared to other advanced biometric cues such as face, fingerprint, iris and voice. Yet, it's unobtrusive quality makes it more attractive than other biometric traits for certain applications. In some cases, gait recognition could be used to narrow down the potential list of identities in a multi-biometric system thereby increasing throughput and decreasing the need for user interaction. However, a number of compelling challenges have stymied the adoption of this biometric cue in operational environments. Chief among them has been the adverse effect of covariates on gait recognition performance. While some of these effects, especially due to variations in walking speed and viewpoint, have been analyzed and studied in the context of vision-based gait recognition, their impact on other modalities has not been systematically studied. With an appropriate fusion of features across modalities, gait recognition may well escape its "infancy" and mature into a very practical, unobtrusive member of the industrial biometrics toolbox.

Appendix 1

Some of the images used in this article are from the Internet. The URLs for these images are provided below.Figure 1: http://goldberg.berkeley.edu/courses/S06/IEOR-QE-S06/images.htmlhttp://www.tonychai.com/mocap.html (Photo Credit: Tony Chai)https://commons.wikimedia.org/wiki/File:Sony_HDR-FX1E_20050423.jpghttps://www.ecs.soton.ac.uk/gait-tunnelhttps://commons.wikimedia.org/wiki/File:IPhone_1st_Gen.svg (Photo credit: Rafael Fernandez)https://commons.wikimedia.org/wiki/File:Aristotle_Altemps_Inv8575.jpghttps://commons.wikimedia.org/wiki/File:Giovanni_Alfonso_Borelli.jpghttps://commons.wikimedia.org/wiki/File:Wilhelm_Eduard_Weber_II.jpghttps://www.saw-leipzig.de/de/mitglieder/weberefhttp://www.publicdomainpictures.net/view-image.php?image=27277&picture=security-camera (Photo credit: Peter Griffin) http://www.am.sanken.osaka-u.ac.jp/BiometricDB/GaitLP.htmlhttps://www.theasfp.org/https://openclipart.org/Figure 3:https://commons.wikimedia.org/wiki/File:Gray827.pnghttps://commons.wikimedia.org/wiki/File:Gray355.pngFigure 10:https://openclipart.org/.

References

Addlesee, M.D., Jones, A., Livesey, F., Samaria, F., 1997. The ORL active floor. IEEE Pers. Commun. 4, 35–41.

Al-Obaidi, S., Wall, J.C., Al-Yaqoub, A., Al-Ghanim, M., 2003. Basic gait parameters: a comparison of reference data for normal subjects 20 to 29 years of age from Kuwait and Scandinavia. J. Rehabil. Res. Dev. 40 (4), 361–366.

Aristotle, 350 BC. On the Gait of Animals.

Ariyanto, G., Nixon, M.S., 2011. Model-based 3D gait biometrics. Proceedings of the 2011 International Joint Conference on Biometrics (IJCB). pp. 1–7.

Bashir, K., Xiang, T., Gong, S., 2008. Feature selection on gait energy image for human identification. Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008). pp. 985–988.

Bashir, K., Xiang, T., Gong, S., 2009. Gait recognition using gait entropy image. Proceedings of the 3rd International Conference on Imaging for Crime Detection and Prevention (ICDP 2009). IET, pp. 1–6.

BenAbdelkader, C., Cutler, R., Davis, L., 2002. View-invariant estimation of height and stride for gait recognition. Biometric Authentication. Springer, pp. 155–167.

BenAbdelkader, C., Cutler, R., Nanda, H., Davis, L., 2001. Eigengait: motion-based recognition of people using image self-similarity. Audio- and Video-Based Biometric Person Authentication. pp. 284–294.

Bhanu, B., Han, J., 2003. Human recognition on combining kinematic and stationary features. Audio- and video-based biometric person authentication. pp. 600–608.

Birrell, S.A., Hooper, R.H., Haslam, R.A., 2007. The effect of military load carriage on ground reaction forces. Gait Posture 26 (4), 611–614.

Bobick, A.F., Johnson, A.Y., 2001. Gait recognition using static, activity-specific parameters. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001). vol. 1. pp. I-423–I-427.

Borelli, G.A., 1679. On the gait of animals.

Bouchrika, I., Nixon, M.S., 2007. Model-based feature extraction for gait analysis and recognition. Computer Vision/Computer Graphics Collaboration Techniques. Springer, pp. 150–160.

Bouchrika, I., Nixon, M.S., 2008. Exploratory factor analysis of gait recognition. Proceedings of the 8th IEEE International Conference on Automatic Face & Gesture Recognition (FG'08). pp. 1–6.

Bouchrika, I., Nixon, M.S., 2008. Gait recognition by dynamic cues. Proceedings of the 19th International Conference on Pattern Recognition (ICPR 2008). pp. 1–4.

Bours, P., Shrestha, R., 2010. Eigensteps: a giant leap for gait recognition. Proceedings of the 2nd International Workshop on Security and Communication Networks (IWSCN). pp. 1–6.

Boyd, J.E., Little, J.J., 2005. Biometric gait recognition. Advanced Studies in Biometrics: Summer School on Biometrics. vol. 3161/2005. Springer, pp. 19–42.

de Carvalho, R.L., Rosa, P.F.F., 2010. Identification system for smart homes using footstep sounds. Proceedings of the 2010 IEEE International Symposium on Industrial Electronics (ISIE). pp. 1639–1644.

Castro, F.M., Marín-Jiménez, M.J., Guil, N., 2015. Empirical study of audio-visual features fusion for gait recognition. Computer Analysis of Images and Patterns. pp. 727–739.

Castro, F.M., Marín-Jiménez, M.J., Guil, N., 2016. Multimodal features fusion for gait, gender and shoes recognition. Mach. Vis. Appl. 1–16.

Cattin, P.C., 2002. Biometric Authentication System Using Human Gait. ETH, Zürich Ph.D. thesis.

Chalidabhongse, T., Kruger, V., Chellappa, R., 2001. The UMD Database for Human Identification at a Distance. Technical Report. University of Maryland.

Cho, Y.-O., So, B.-C., Jung, J.-W., 2012. User recognition using sequential footprints under shoes based on mat-type floor pressure sensor. Adv. Sci. Lett. 9 (1), 591–596.

Collins, R.T., Gross, R., Shi, J., 2002. Silhouette-based human identification from body shape and gait. Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition. pp. 366–371.

Connor, P.C., 2015. Comparing and combining underfoot pressure features for shod and unshod gait biometrics. Proceedings of the 2015 IEEE International Conference on Technologies for Homeland Security (HST).

Cunado, D., Nixon, M.S., Carter, J.N., 1997. Using gait as a biometric, via phase-weighted magnitude spectra. Audio- and Video-based Biometric Person Authentication. pp. 93–102.

Cunado, D., Nixon, M.S., Carter, J.N., 2003. Automatic extraction and description of human gait models for recognition purposes. Comput. Vision Image Understanding 90 (1), 1–41.

Cuntoor, N., Kale, A., Chellappa, R., 2003. Combining multiple evidences for gait recognition. Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASS 2003). vol. 3. pp. III–33.

Cutting, J.E., Kozlowski, L.T., 1977. Recognizing friends by their walk: gait perception without familiarity cues. Bull. Psychon. Soc. 9 (5), 353–356.

Cutting, J.E., Proffitt, D.R., Kozlowski, L.T., 1978. A biomechanical invariant for gait perception. J. Exp. Psychol. 4 (3), 357.

Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005). vol. 1. pp. 886–893.

Daugherty, S., 2014. Detectives turn to podiatry to help solve murder case. [Online; accessed 19-January-2017]. URL http://pilotonline.com/news/local/crime/detectives-turn-to-podiatry-to-help-solve-murder-case/article_6017ae34-fe72-56f3-85bb-8b4f08e7eb0a.html.

DeCann, B., Ross, A., 2012. Can a poor verification system be a good identification system? a preliminary study,. Proceedings of the IEEE International Workshop on Information Forensics and Security (WIFS). pp. 1–6.

Derawi, M.O., Bours, P., Holien, K., 2010. Improved cycle detection for accelerometer based gait authentication. Proceedings of the 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP). pp. 312–317.

Derawi, M.O., Nickel, C., Bours, P., Busch, C., 2010. Unobtrusive user-authentication on mobile phones using biometric gait recognition. Proceedings of the 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP). pp. 306–311.

Derlatka, M., 2013. Modified kNN algorithm for improved recognition accuracy of biometrics system based on gait. Computer Information Systems and Industrial Management. Springer, pp. 59–66.

Derlatka, M., Bogdan, M., 2015. Ensemble kNN classifiers for human gait recognition based on ground reaction forces. Proceedings of the 8th International Conference on Human System Interactions (HSI). pp. 88–93.

Derlatka, M., Bogdan, M., 2015. Fusion of static and dynamic parameters at decision level in human gait recognition. Pattern Recognition and Machine Intelligence. Springer, pp. 515–524.

Diedrich, F.J., Warren Jr, W.H., 1998. The dynamics of gait transitions: effects of grade and load. J. Mot. Behav. 30 (1), 60–78.

Dikovski, B., Madjarov, G., Gjorgjevikj, D., 2014. Evaluation of different feature sets for gait recognition using skeletal data from Kinect. Proceedings of the 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). pp. 1304–1308.

Doddington, G.R., Przybocki, M.A., Martin, A.F., Reynolds, D.A., 2000. The NIST speaker recognition evaluation–overview, methodology, systems, results, perspective. Speech Commun. 31 (2), 225–254.

El-Alfy, H., Mitsugami, I., Yagi, Y., 2014. A new gait-based identification method using local gauss maps. Proceedings of the 2014 Asian Conference on Computer Vision (ACCV) Workshops. pp. 3–18.

First use of forensic gait analysis evidence in court. 2016. [Online; accessed 19-January-2017]. URL http://www.guinnessworldrecords.com/world-records/first-use-of-forensic-gait-analysis-evidence-in-court/.

Foster, J.P., Nixon, M.S., Prügel-Bennett, A., 2003. Automatic gait recognition using area-based metrics. Pattern Recognit. Lett. 24 (14), 2489–2497.

Foti, T., Davids, J.R., Bagley, A., 2000. A biomechanical analysis of gait during pregnancy. JBJS 82 (5), 625.

Gadaleta, M., Merelli, L., Rossi, M., 2016. Human authentication from ankle motion data using convolutional neural networks. Proceedings of the 2016 Statistical Signal Processing Workshop. pp. 1–5.

Gafurov, D., Helkala, K., Søndrol, T., 2006. Gait recognition using acceleration from MEMS. Proceedings of the First International Conference on Availability, Reliability and Security (ARES 2006). pp. 1–6.

Gafurov, D., Snekkenes, E., Bours, P., 2007. Spoof attacks on gait authentication system. IEEE Trans. Inf. Forensics Secur. 2 (3), 491–502.

Gafurov, D., Snekkenes, E., Bours, P., 2010. Improved gait recognition performance using cycle matching. Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications Workshops (WAINA). pp. 836–841.

Galai, B., Benedek, C., 2015. Feature selection for lidar-based gait recognition. Proceedings of the International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM).

Geiger, J.T., Hofmann, M., Schuller, B., Rigoll, G., 2013. Gait-based person identification by spectral, cepstral and energy-related audio features. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2013). pp. 458–462.

Geng, X., Smith-Miles, K., Wang, L., Li, M., Wu, Q., 2010. Context-aware fusion: a case study on fusion of gait and face for human identification in video. Pattern Recognit. 43 (10), 3660–3673.

Goffredo, M., Bouchrika, I., Carter, J.N., Nixon, M.S., 2010. Self-calibrating view-invariant gait biometrics. IEEE Trans. Syst. Man Cybern. Part B 40 (4), 997–1008.

Gross, R., Shi, J., 2001. The CMU Motion of Body (MoBo) Database. CMU-RI-TR-01-18. Technical Report. Carnagie Mellon University.

Guan, Y., Li, C.-T., 2013. A robust speed-invariant gait recognition system for walker and runner identification. Proceedings of the 2013 International Conference on Biometrics (ICB). pp. 1–8.

Guan, Y., Li, C.-T., Hu, Y., 2012. Robust clothing-invariant gait recognition. Proceedings of the Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP). pp. 321–324.

Hadid, A., Ghahramani, M., Bustard, J., Nixon, M., 2013. Improving gait biometrics under spoofing attacks. Proceedings of the International Conference on Image Analysis and Processing (ICIAP 2013). Springer, pp. 1–10.

Hadid, A., Ghahramani, M., Kellokumpu, V., Pietikainen, M., Bustard, J., Nixon, M., 2012. Can gait biometrics be spoofed? Proceedings of the 21st International Conference on Pattern Recognition (ICPR). pp. 3280–3283.

Han, J., Bhanu, B., 2004. Statistical feature fusion for gait-based human recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004). vol. 2. pp. II–842.

Han, J., Bhanu, B., 2006. Individual recognition using gait energy image. IEEE Trans. Pattern Anal. Mach. Intell. 28 (2), 316–322.

Han, S., 2015. The influence of walking speed on gait patterns during upslope walking. J. Med. Imaging Health Inform. 5 (1), 89–92.

Hofmann, M., Bachmann, S., Rigoll, G., 2012. 2.5D gait biometrics using the depth gradient histogram energy image. Proceedings of the Fifth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS). pp. 399–403.

Hofmann, M., Geiger, J., Bachmann, S., Schuller, B., Rigoll, G., 2014. The TUM gait from audio, image and depth (GAID) database: multimodal recognition of subjects and traits. J. Vis. Commun. Image Represent. 25 (1), 195–206.

Hofmann, M., Rigoll, G., 2012. Improved gait recognition using gradient histogram energy image. Proceedings of the 19th IEEE International Conference on Image Processing (ICIP 2012). pp. 1389–1392.

Hong, S., Lee, H., Nizami, I.F., Kim, E., 2007. A new gait representation for human identification: mass vector. Proceedings of the 2nd IEEE Conference on Industrial Electronics and Applications (ICIEA 2007). pp. 669–673.

Hossain, M.A., Makihara, Y., Wang, J., Yagi, Y., 2010. Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control. Pattern Recognit. 43 (6), 2281–2291.

Hsiang, S.M., Chang, C., 2002. The effect of gait speed and load carrying on the reliability of ground reaction forces. Saf. Sci. 40 (7), 639–657.

Huang, P.S., Harris, C.J., Nixon, M.S., 1999. Recognising humans by gait via parametric canonical space. Artif. Intell. Eng. 13 (4), 359–366.

Ioannidis, D., Tzovaras, D., Damousis, I.G., Argyropoulos, S., Moustakas, K., 2007. Gait recognition using compact feature extraction transforms and depth information. IEEE Trans. Inf. Forensics Secur. 2 (3), 623–630.

Islam, M.S., Islam, M.R., Akter, M.S., Hossain, M.A., Molla, M.K.I., 2013. Window based clothing invariant gait recognition. Proceedings of the 2013 International Conference on Advances in Electrical Engineering (ICAEE). pp. 411–414.

Iwama, H., Okumura, M., Makihara, Y., Yagi, Y., 2012. The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition. IEEE Trans. Inf. Forensics Secur. 7 (5), 1511–1521.

Jain, A.K., Klare, B., Ross, A., 2015. Guidelines for best practices in biometrics research. Proceedings of the 8th IAPR International Conference on Biometrics (ICB). pp. 1–5.

Janssen, D., Schöllhorn, W.I., Lubienetzki, J., Fölling, K., Kokenge, H., Davids, K., 2008. Recognition of emotions in gait patterns by means of artificial neural nets. J. Nonverbal Behav. 32 (2), 79–92.

Jenkins, J., Ellis, C., 2007. Using ground reaction forces from gait analysis: body mass as a weak biometric. Pervasive Computing. Springer, pp. 251–267.

Johansson, G., 1973. Visual perception of biological motion and a model for its analysis. Percept. Psychophys. 14 (2), 201–211.

Johnson, A.Y., Bobick, A.F., 2001. A multi-view method for gait recognition using static body parameters. Audio- and Video-Based Biometric Person Authentication. pp. 301–311.

Johnston, A.H., Weiss, G.M., 2015. Smartwatch-based biometric gait recognition. Proceedings of the Seventh IEEE International Conference on Biometrics: Theory, Applications and Systems.

Jordan, K., Challis, J.H., Newell, K.M., 2007. Walking speed influences on gait cycle variability. Gait Posture 26 (1), 128–134.

Jung, J.-W., Lee, S.-W., Bien, Z., Sato, T., 2004. Person recognition method using sequential walking footprints via overlapped foot shape and center-of-pressure trajectory. IEICE Trans. Fundamentals Electron. Commun. Comput. Sci. 87 (6), 1393–1400.

Kale, A., Rajagopalan, A.N., Cuntoor, N., Kruger, V., 2002. Gait-based recognition of humans using continuous HMMs. Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition. pp. 336–341.

Kalgaonkar, K., Raj, B., 2007. Acoustic Doppler sonar for gait recognition. Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS 2007). pp. 27–32.

Kennedy, R.B., 1996. Uniqueness of bare feet and its use as a possible means of

identification. Forensic Sci. Int. 82 (1), 81–87.

Kimura, T., Makihara, Y., Muramatsu, D., Yagi, Y., 2014. Quality-dependent score-level fusion of face, gait, and the height biometrics. IPSJ Trans. Comput. Vision Appl. 6 (0), 53–57.

Kovač, J., Peer, P., 2014. Human skeleton model based dynamic features for walking speed invariant gait recognition. Math. Problems Eng. 2014.

Kozlowski, L.T., Cutting, J.E., 1977. Recognizing the sex of a walker from a dynamic point-light display. Percept. Psychophys. 21 (6), 575–580.

Kumar, V., Ramakrishnan, M., 2011. Legacy of footprints recognition–a review. Int. J. Comput. Appl. 35 (11), 9–16.

Kuragano, T., Yamaguchi, A., Furukawa, S., 2005. A method to measure foot print similarity for gait analysis. Proceedings of the 2005 International Conference on Computational Intelligence for Modelling, Control and Automation and the International Conference on Intelligent Agents, Web Technologies and Internet Commerce. vol. 2. pp. 816–822.

Kusakunniran, W., Wu, Q., Zhang, J., Li, H., 2011. Speed-invariant gait recognition based on Procrustes shape analysis using higher-order shape configuration. Proceedings of the 18th IEEE International Conference on Image Processing (ICIP 2011). pp. 545–548.

Kusakunniran, W., Wu, Q., Zhang, J., Li, H., 2012. Gait recognition across various walking speeds using higher order shape configuration based on a differential composition model. IEEE Trans. Syst. Man Cybern. Part B 42 (6), 1654–1668.

Lee, H., Lee, B., Jung, J.-W., Hong, S., Kim, E., 2013. Human biometric identification through integration of footprint and gait. Int. J. Control Autom. Syst. 11 (4), 826–833.

Lee, L., Grimson, W.E.L., 2002. Gait analysis for recognition and classification. Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition. pp. 148–155.

Lee, S., Liu, Y., Collins, R.T., 2007. Shape variation-based frieze pattern for robust gait recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'07).

Little, J., Boyd, J., 1995. Describing motion for recognition. Proceedings of the 1995 International Symposium on Computer Vision. pp. 235–240.

Little, J., Boyd, J., 1998. Recognizing people by their gait: the shape of motion. Videre 1 (2), 1–32.

Liu, L.-F., Jia, W., Zhu, Y.-H., 2009. Survey of gait recognition. Emerging Intelligent Computing Technology and Applications. With Aspects of Artificial Intelligence. Springer, pp. 652–659.

Liu, N., Tan, Y.-P., 2010. View invariant gait recognition. Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP 2010). pp. 1410–1413.

Liu, Y., Zhang, J., Wang, C., Wang, L., 2012. Multiple HOG templates for gait recognition. Proceedings of the 21st International Conference on Pattern Recognition (ICPR 2012). pp. 2930–2933.

Liu, Z., Sarkar, S., 2004. Simplest representation yet for gait recognition: averaged silhouette. Proceedings of the 17th International Conference on Pattern Recognition (ICPR 2004). vol. 4. pp. 211–214.

Liu, Z., Sarkar, S., 2006. Improved gait recognition by gait dynamics normalization. IEEE Trans. Pattern Anal. Mach. Intell. 28 (6), 863–876.

Lythgo, N., Wilson, C., Galea, M., 2009. Basic gait and symmetry measures for primary school-aged children and young adults whilst walking barefoot and with shoes. Gait Posture 30 (4), 502–506.

Majumdar, D., Pal, M.S., Majumdar, D., 2010. Effects of military load carriage on kinematics of gait. Ergonomics 53 (6), 782–791.

Maki, B.E., 1997. Gait changes in older adults: predictors of falls or indicators of fear? J. Am. Geriatr. Soc. 45 (3), 313–320.

Makihara, Y., Matovski, D.S., Nixon, M.S., Carter, J.N., Yagi, Y., 2015. Gait Recognition: Databases, Representations, and Applications. John Wiley and Sons, Inc.

Mäntyjärvi, J., Lindholm, M., Vildjiounaite, E., Makela, S.M., Ailisto, H.A., 2005. Identifying users of portable devices from gait pattern with accelerometers. Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005). vol. 2. pp. ii–973.

Mason, J.E., Traoré, I., Woungang, I., 2016. Machine Learning Techniques for Gait Biometric Recognition. Springer.

Matovski, D.S., Nixon, M.S., Mahmoodi, S., Carter, J.N., 2010. The effect of time on the performance of gait biometrics. Proceedings of the Fourth International Conference on Biometrics: Theory Applications and Systems (BTAS). pp. 1–6.

Middleton, L., Buss, A.A., Bazin, A., Nixon, M.S., 2005. A floor sensor system for gait recognition. Proceedings of the Fourth IEEE Workshop on Automatic Identification Advanced Technologies. pp. 171–176.

Mjaaland, B. B., 2009. Gait mimicking: attack resistance testing of gait authentication systems.

Mjaaland, B.B., Bours, P., Gligoroski, D., 2011. Walk the walk: attacking gait biometrics by imitation. Information Security. Springer, pp. 361–380.

Mobasseri, B.G., Amin, M.G., 2009. A time-frequency classifier for human gait recognition. Proc. SPIE Defense, Security, and Sensing. vol. 7306. 730628–730628–9

Montepare, J., Goldstein, S., Clausen, A., 1987. The identification of emotions from gait information. J. Nonverbal Behav. 11 (1), 33–42.

Morishita, H., Fukui, R., Sato, T., 2002. High resolution pressure sensor distributed floor for future human-robot symbiosis environments. Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems. vol. 2. pp. 1246–1251.

Moustakidis, S.P., Theocharis, J.B., Giakas, G., 2008. Subject recognition based on ground reaction force measurements of gait signals. IEEE Trans. Syst. Man Cybern. Part B 38 (6), 1476–1485.

Moustakidis, S.P., Theocharis, J.B., Giakas, G., 2009. Feature extraction based on a fuzzy complementary criterion for gait recognition using GRF signals. Proceedings of the

17th Mediterranean Conference on Control and Automation (MED'09). pp. 1456–1461.

Mowbray, S.D., Nixon, M.S., 2003. Automatic gait recognition via Fourier descriptors of deformable objects. Audio- and Video-Based Biometric Person Authentication. pp. 566–573.

Muaaz, M., Nickel, C., 2012. Influence of different walking speeds and surfaces on accelerometer-based biometric gait recognition. Proceedings of the 35th International Conference on Telecommunications and Signal Processing (TSP 2012). pp. 508–512.

Muramatsu, D., Iwama, H., Makihara, Y., Yagi, Y., 2013. Multi-view multi-modal person authentication from a single walking image sequence. Proceedings of the 2013 International Conference on Biometrics (ICB). pp. 1–8.

Murase, H., Sakai, R., 1996. Moving object recognition in eigenspace representation: gait analysis and lip reading. Pattern Recognit. Lett. 17 (2), 155–162.

Murray, M.P., 1967. Gait as a total pattern of movement: Including a bibliography on gait. Am. J. Phys. Med. Rehab. 46 (1), 290–333.

Murray, M.P., Drought, A.B., Kory, R.C., 1964. Walking patterns of normal men. J. Bone Joint Surg. 46 (2), 335–360.

Nah, G., 1985. Circadian Changes of Human Stature and Foot Volume for Young Adult Men. Korea Advanced Institute of Science and Technology Master's thesis.

Nakajima, K., Mizukami, Y., Tanaka, K., Tamura, T., 2000. Footprint-based personal recognition. IEEE Trans. Biomed. Eng. 47 (11), 1534–1537.

Ngo, T.T., Makihara, Y., Nagahara, H., Mukaigawa, Y., Yagi, Y., 2014. The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication. Pattern Recognit. 47 (1), 228–237.

Nilsson, J., Thorstensson, A., 1989. Ground reaction forces at different speeds of human walking and running. Acta Physiol. Scand. 136 (2), 217–227.

Nixon, M., Carter, J., Shutler, J., Grant, M., 2001. Experimental Plan for Automatic Gait Recognition. Technical Report. University of Southampton.

Nixon, M.S., Carter, J.N., 2006. Automatic recognition by gait. Proc. IEEE 94 (11), 2013–2024.

Nixon, M.S., Tan, T., Chellappa, R., 2010. Human Identification Based on Gait. vol. 4 Springer.

Niyogi, S.A., Adelson, E.H., 1994. Analyzing and recognizing walking figures in XYT. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'94). pp. 469–474.

Öberg, T., Karsznia, A., Öberg, K., 1993. Basic gait parameters: reference data for normal subjects, 10–79 years of age. J. Rehabil. Res. Dev. 30, 210–223.

Orendurff, M.S., Segal, A.D., Klute, G.K., Berge, J.S., Rohr, E.S., Kadel, N.J., 2004. The effect of walking speed on center of mass displacement. J. Rehabil. Res. Dev. 41 (6A), 829–834.

Orr, R.J., Abowd, G.D., 2000. The smart floor: a mechanism for natural user identification and tracking. Extended Abstracts on Human Factors in Computing Systems (CHI'00). pp. 275–276.

Otero, M., 2005. Application of a continuous wave radar for human gait recognition. Defense and Security. pp. 538–548.

Pataky, T.C., Mu, T., Bosch, K., Rosenbaum, D., Goulermas, J.Y., 2011. Gait recognition: highly unique dynamic plantar pressure patterns among 104 individuals. J. R. Soc. Interface. http://dx.doi.org/10.1098/rsif.2011.0430.

Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.J., 2000. The FERET evaluation methodology for face-recognition algorithms. IEEE Trans. Pattern Anal. Mach. Intell. 22 (10), 1090–1104.

Phillips, P.J., Sarkar, S., Robledo, I., Grother, P., Bowyer, K., 2002. The gait identification challenge problem: data sets and baseline algorithm. Proceedings of the 16th International Conference on Pattern Recognition. vol. 1. pp. 385–388.

Preis, J., Kessel, M., Werner, M., Linnhoff-Popien, C., 2012. Gait recognition with Kinect. Proceedings of the 1st International Workshop on Kinect in Pervasive Computing.

Qian, G., Zhang, J., Kidané, A., 2010. People identification using floor pressure sensing and analysis. IEEE Sens. J. 10 (9), 1447–1460.

Qu, X., Yeo, J.C., 2011. Effects of load carriage and fatigue on gait characteristics. J. Biomech. 44 (7), 1259–1263.

Rodriguez, R.V., Evans, N.W.D., Mason, J.S.D., 2009. Footstep recognition. Encyclopedia of Biometrics. Springer, pp. 550–557.

Roether, C.L., Omlor, L., Christensen, A., Giese, M.A., 2009. Critical features for the perception of emotion from gait. J. Vis. 9 (6), 15.

Rong, L., Jianzhong, Z., Ming, L., Xiangfeng, H., 2007. A wearable acceleration sensor system for gait recognition. Proceedings of the Second IEEE Conference on Industrial Electronics and Applications (ICIEA 2007). pp. 2654–2659.

Rosenbaum, D., Hautmann, S., Gold, M., Claes, L., 1994. Effects of walking speed on plantar pressure patterns and hindfoot angular motion. Gait Posture 2 (3), 191–197.

Ryu, T., Soon Choi, H., Choi, H., Chung, M.K., 2006. A comparison of gait characteristics between Korean and Western people for establishing Korean gait reference data. Int. J. Ind. Ergon. 36 (12), 1023–1030.

Samson, M.M., Crowe, A., De Vreede, P.L., Dessens, J.A.G., Duursma, S.A., Verhaar, H.J.J., 2001. Differences in gait parameters at a preferred walking speed in healthy subjects due to age, height and body weight. Aging. Clin. Exp. Res. 13 (1), 16–21.

Sarkar, S., Phillips, P.J., Liu, Z., Vega, I.R., Grother, P., Bowyer, K.W., 2005. The humanID gait challenge problem: data sets, performance, and analysis. IEEE Trans. Pattern Anal. Mach. Intell. 27 (2), 162–177.

Saunders, J.B.deC.M., Inman, V.T., Eberhart, H.D., 1953. The major determinants in normal and pathological gait. J. Bone Joint Surg. 35 (3), 543–558.

Seely, R.D., Samangooei, S., Lee, M., Carter, J.N., Nixon, M.S., 2008. The University of Southampton Multi-Biometric Tunnel and introducing a novel 3D gait dataset. Proceedings of the Second IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS 2008). pp. 1–6.

Shakhnarovich, G., Lee, L., Darrell, T., 2001. Integrated face and gait recognition from multiple views. Proceedings of the IEEE Computer Society Conference on Computer

Vision and Pattern Recognition (CVPR 2001). vol. 1. pp. I–439.

Shiraga, K., Makihara, Y., Muramatsu, D., Echigo, T., Yagi, Y., 2016. Geinet: View-invariant gait recognition using a convolutional neural network. Proceedings of the 2016 International Conference on Biometrics (ICB). pp. 1–8.

Shoji, Y., Takasuka, T., Yasukawa, H., 2004. Personal identification using footstep detection. Proceedings of the International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2004). pp. 43–47.

Shutler, J., Grant, M., Nixon, M.S., Carter, J.N., 2002. On a large sequence-based human gait database. Proceedings of the Fourth International Conference on Recent Advances in Soft Computing. pp. 66–72.

Shutler, J.D., Nixon, M.S., Harris, C.J., 2000. Statistical gait description via temporal moments. Proceedings of the 4th IEEE Southwest Symposium on Image Analysis and Interpretation. pp. 291–295.

Sivapalan, S., Chen, D., Denman, S., Sridharan, S., Fookes, C., 2011. Gait energy volumes and frontal gait recognition using depth images. Proceedings of the 2011 International Joint Conference on Biometrics (IJCB). pp. 1–6.

Sprager, S., Juric, M.B., 2015. Inertial sensor-based gait recognition: a review. Sensors 15 (9), 22089–22127.

Stevenage, S.V., Nixon, M.S., Vince, K., 1999. Visual analysis of gait as a cue to identity. Appl. Cogn. Psychol. 13 (6), 513–526.

Stevenson, J.P., Firebaugh, S.L., Charles, H.K., 2007. Biometric identification from a floor based PVDF sensor array using hidden markov models. Proceedings of the 2007 Sensors Applications Symposium Technology Conference (SAS).

Suutala, J., Fujinami, K., Röning, J., 2008. Gaussian process person identifier based on simple floor sensors. Smart Sensing and Context. Springer, pp. 55–68.

Suutala, J., Röning, J., 2004. Towards the adaptive identification of walkers: automated feature selection of footsteps using distinction sensitive LVQ. International Workshop on Processing Sensory Information for Proactive Systems (PSIPS 2004). pp. 14–15.

Suutala, J., Röning, J., 2005. Combining classifiers with different footstep feature sets and multiple samples for person identification. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005). vol. 5. pp. v–357.

Świtoński, A., Polański, A., Wojciechowski, K., 2011a. Human identification based on gait paths. Advances Concepts for Intelligent Vision Systems. Springer, pp. 531–542.

Świtoński, A., Polański, A., Wojciechowski, K., 2011b. Human identification based on the reduced kinematic data of the gait. Proceedings of the 7th International Symposium on Image and Signal Processing and Analysis (ISPA 2011). pp. 650–655.

Takeda, T., Kuramoto, K., Kobashi, S., Hata, Y., 2011. A challenge to biometrics by sole pressure while walking. Proceedings of the 2011 IEEE International Conference on Fuzzy Systems (FUZZ). pp. 1430–1435.

Tan, D., Huang, K., Yu, S., Tan, T., 2006. Efficient night gait recognition based on template matching. Proceedings of the 18th International Conference on Pattern Recognition (ICPR 2006). vol. 3. pp. 1000–1003.

Tanawongsuwan, R., Bobick, A., 2001. Gait recognition from time-normalized joint-angle trajectories in the walking plane. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001). vol. 2. pp. II-726–II-731.

Tanawongsuwan, R., Bobick, A., 2003. Performance analysis of time-distance gait parameters under different speeds. Audio- and Video-Based Biometric Person Authentication. pp. 715–724.

Tanawongsuwan, R., Bobick, A., 2003. A Study of Human Gaits Across Different Speeds. Technical Report. Georgia Tech.

Tanawongsuwan, R., Bobick, A., 2004. Modelling the effects of walking speed on appearance-based gait recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004). vol. 2. pp. II-783–II-790.

Tao, D., Li, X., Wu, X., Maybank, S.J., 2007. General tensor discriminant analysis and Gabor features for gait recognition. IEEE Trans. Pattern Anal. Mach. Intell. 29 (10), 1700–1715.

Taylor, A.J., Menz, H.B., Keenan, A.-M., 2004. The influence of walking speed on plantar pressure measurements using the two-step gait initiation protocol. Foot 14 (1), 49–55.

Tsuji, A., Makihara, Y., Yagi, Y., 2010. Silhouette transformation based on walking speed for gait identification. Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2010). pp. 717–722.

Uhl, A., Wild, P., 2008. Footprint-based biometric verification. J. Electron. Imaging 17 (1). 011016–011016–10

Veeraraghavan, A., Srivastava, A., Roy-Chowdhury, A.K., Chellappa, R., 2009. Rate-invariant recognition of humans and their activities. IEEE Trans. Image Process. 18 (6), 1326–1339.

Vera-Rodríguez, R., Evans, N.W.D., Lewis, R.P., Fauve, B., Mason, J.S.D., 2007. An experimental study on the feasibility of footsteps as a biometric. European Signal Process Conference (EUSIPCO). pp. 748–752.

Vera-Rodríguez, R., Lewis, R.P., Mason, J.S.D., Evans, N.W.D., 2008. Footstep recognition for a smart home environment. Int. J. Smart Home 2 (2), 95–110.

Vera-Rodríguez, R., Mason, J.S.D., Fierrez, J., Ortega-Garcia, J., 2011. Analysis of spatial domain information for footstep recognition. IET Comput. Vision 5 (6), 380–388.

Vera-Rodríguez, R., Mason, J.S.D., Fierrez, J., Ortega-Garcia, J., 2013. Comparative analysis and fusion of spatiotemporal information for footstep recognition. IEEE Trans. Pattern Anal. Mach. Intell. 35 (4), 823–834.

Vera-Rodríguez, R., Tome, P., Fierrez, J., Ortega-Garcia, J., 2012. Fusion of footsteps and face biometrics on an unsupervised and uncontrolled environment. SPIE Defense, Security, and Sensing. vol. 8371. pp. 83711U.

Vera-Rodríguez, R., Rubén, J., Fierrez, J., Mason, J.S.D., Ortega-Garcia, J., 2013. A novel approach of gait recognition through fusion with footstep information. Proceedings of the 2013 International Conference on Biometrics (ICB). pp. 1–6.

Veres, G.V., Gordon, L., Carter, J.N., Nixon, M.S., 2004. What image information is important in silhouette-based gait recognition? Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004). vol. 2. pp. II-776–II-782.

Vildjiounaite, E., Makela, S.M., Lindholm, M., Kyllonen, V., Ailisto, H., 2007. Increasing security of mobile devices by decreasing user effort in verification. Proceedings of the Second International Conference on Systems and Networks Communications (ICSNC 2007). pp. 80.

Wagg, D.K., Nixon, M.S., 2004. On automated model-based extraction and analysis of gait. Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition. pp. 11–16.

Wang, C., Zhang, J., Pu, J., Yuan, X., Wang, L., 2010. Chrono-gait image: a novel temporal template for gait recognition. Proceedings of the European Conference on Computer Vision (ECCV 2010). Springer, pp. 257–270.

Wang, L., Ning, H., Hu, W., Tan, T., 2002. Gait recognition based on Procrustes shape analysis. Proceedings of the 2002 International Conference on Image Processing. 3. pp. III–433.

Wang, L., Tan, T., Hu, W., Ning, H., 2003. Automatic gait recognition based on statistical shape analysis. IEEE Trans. Image Process. 12 (9), 1120–1131.

Wang, L., Tan, T., Ning, H., Hu, W., 2003. Silhouette analysis-based gait recognition for human identification. IEEE Trans. Pattern Anal. Mach. Intell. 25 (12), 1505–1518.

Wang, W., Liu, A.X., Shahzad, M., 2016. Gait recognition using wifi signals. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. pp. 363–373.

Weber, W., Weber, E., 1836. Mechanics of the Human Walking Apparatus.

Whittle, M.W., 2007. Gait Analysis: An Introduction, (fourth ed.). .

Winter, D.A., Patla, A.E., Frank, J.S., Walt, S.E., 1990. Biomechanical walking pattern changes in the fit and healthy elderly. Phys. Ther. 70 (6), 340–347.

Wolf, T., Babaee, M., Rigoll, G., 2016. Multi-view gait recognition using 3d convolutional neural networks. Proceedings of the IEEE International Conference on Image Processing (ICIP 2016). pp. 4165–4169.

Wu, Z., Huang, Y., Wang, L., 2015. Learning representative deep features for image set analysis. IEEE Trans. Multimedia 17 (11), 1960–1968.

Yam, C., Nixon, M.S., Carter, J.N., 2004. Automated person recognition by walking and running via model-based approaches. Pattern Recognit. 37 (5), 1057–1072.

Yamakawa, T., Taniguchi, K., Asari, K., Kobashi, S., Hata, Y., 2008. Biometric personal identification based on gait pattern using both feet pressure change. Proceedings of the World Automation Congress (WAC 2008). pp. 1–6.

Yao, Z.M., Zhou, X., Lin, E.D., Xu, S., Sun, Y.N., 2010. A novel biometric recognition system based on ground reaction force measurements of continuous gait. Proceedings of the Third Conference on Human System Interactions (HSI 2010). pp. 452–458.

Yoo, J.-H., Nixon, M.S., 2011. Automated markerless analysis of human gait motion for recognition and classification. ETRI J. 33 (2), 259–266.

Youn, I.H., Choi, S., Le May, R., Bertelsen, D., Youn, J.H., 2014. New gait metrics for biometric authentication using a 3-axis acceleration. Proceedings of the 11th IEEE Consumer Communications and Networking Conference (CCNC). pp. 596–601.

Yu, S., Tan, D., Tan, T., 2006. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. Proceedings of the 18th International Conference on Pattern Recognition (ICPR 2006). vol. 4. pp. 441–444.

Yu, S., Wang, L., Hu, W., Tan, T., 2004. Gait analysis for human identification in frequency domain. Proceedings of the First IEEE Symposium on Multi-Agent Security and Survivability. pp. 282–285.

Yu, T., Zou, J.-H., 2012. Automatic human gait imitation and recognition in 3D from monocular video with an uncalibrated camera. Math. Problems Eng. 2012 35 pages.

Yun, J., Lee, S., Woo, W., Ryu, J., 2003. The user identification system using walking pattern over the ubiFloor. Proceedings of International Conference on Control, Automation, and Systems. vol. 1046. pp. 1050.

Yun, J., Woo, W., Ryu, J., 2005. User identification using users walking pattern over the ubiFloorII. Computational Intelligence and Security. Springer, pp. 949–956.

Zhang, D., Wang, Y., Bhanu, B., 2010. Ethnicity classification based on gait using multi-view fusion. Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 108–115.

Zhang, Y., Pan, G., Jia, K., Lu, M., Wang, Y., Wu, Z., 2014. Accelerometer-based gait recognition by sparse representation of signature points with clusters. IEEE Trans. Cybern. 45 (9), 1864–1875.

Zhang, Z., Andreou, A.G., 2008. Human identification experiments using acoustic micro-Doppler signatures. Proceedings of the Argentine School of Micro-Nanoelectronics, Technology and Applications (EAMTA 2008). pp. 81–86.

Zheng, S., Huang, K., Tan, T., 2010. Translation-invariant representation for cumulative foot pressure images. arXiv preprint arXiv:1010.5426, 1–6.

Zheng, S., Huang, K., Tan, T., 2011. Evaluation framework on translation-invariant representation for cumulative foot pressure image. Proceedings of the 18th IEEE International Conference on Image Processing (ICIP 2011). pp. 201–204.

Zheng, S., Huang, K., Tan, T., Tao, D., 2012. A cascade fusion scheme for gait and cumulative foot pressure image recognition. Pattern Recognit. 45 (10), 3603–3610.

Zhong, Y., Deng, Y., 2014. Sensor orientation invariant mobile gait biometrics. Proceedings of the 2014 IEEE International Joint Conference on Biometrics (IJCB). pp. 1–8.

Zhou, X., Bhanu, B., Han, J., 2007. Human recognition at a distance in video by integrating face profile and gait. Face Biometrics for Personal Identification. Springer, pp. 165–181.