Chapter 1
A Survey of Using Biometrics for Smart Visual Surveillance: Gait Recognition

Imed Bouchrika

Abstract In spite of the increasing concerns raised by privacy advocates against the intrusive deployment of large scale surveillance cameras, the research community has progressed with a remarkable pace into the area of smart visual surveillance. The automation for surveillance systems becomes an obligation to avoid human errors and ensure an efficient strategy for tackling crimes and preventing further terrorist attacks. In this research article, we survey the recent studies in computer vision on gait recognition for covert identification and its application for surveillance scenarios and forensic investigation. The integration of biometric technologies into surveillance systems is a major step milestone to improve the automation process in order to recognize criminal offenders and track them across different places. The suitability of gait biometrics for surveillance applications emerges from the fact that the walking pattern can be captured and perceived from a distance even with poor resolution video as opposed to other biometric modalities which their performance deteriorates in surveillance scenarios.

Keywords Biometrics · Gait recognition · Visual surveillance · Re-identification

Introduction

The deployment of surveillance systems has become ubiquitous in our society regardless of the increasing concerns on personal privacy and data confidentiality. Although privacy advocates fear large scale intrusive surveillance of private individuals, there is a non-debatable consensus that the security and safety of citizens are considered the most rudimentary requirements to be ensured and guaranteed against the sharp increase of the uncountable crimes and terrorist attacks that took place during the last decade across Europe, North Africa and the Middle East. Law enforcement agencies can make use of surveillance footage for the safety of

I. Bouchrika (✉)
Faculty of Science and Technology, University of Souk Ahras, Souk Ahras, Algeria
e-mail: imed@imed.ws

P. Karampelas and T. Bourlai (eds.), *Surveillance in Action*,
Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-319-68533-5_1

our neighborhood and crime prevention. The use of surveillance technology should without doubt assist to lessen the risks and number of crimes by serving as a deterrent. Biometric technologies can be a major milestone to improve the automation process of visual surveillance in order to recognize criminal offenders and track them across different places. Gait defined as the way we walk, is considered recently as a more suited modality for people recognition in surveillance and forensic scenarios. This is because it can be captured non-intrusively and covertly from a distance even with poor resolution imageries. Although, the distinctiveness of gait cannot compare with traditional biometrics, considerable body of research emerged recently asserting the potency and merits of gait biometrics in surveillance applications [28]. In this chapter, we survey the recent studies on the use of computer vision methods for gait biometrics covert identification and its application for surveillance scenarios and forensic investigation.

Smart Visual Surveillance

Visual surveillance systems work based on mounting cameras at remote locations in order to transmit video streams which are stored and monitored at real time. The placement of video cameras is installed at sensitive places and public areas with video being sent to the control center where video feeds are viewed by security staff on multiple screens. The deployment of visual surveillance systems plays a critical role in gathering vital intelligence for law enforcement officers to help them track and detect suspicious individuals who had or about to commit crimes. Visual surveillance systems span to a number of applications from observing people, vehicles, animals and military targets. Originally, visual surveillance technologies were designed for human agents to observe protected remote spaces concurrently as well as to record video data for further offline analysis. In fact, observing surveillance video is a labor-intensive activity specially when a large number of cameras need to be controlled simultaneously. Nowadays, with the sudden increase of surveillance cameras that have exceeded 5 million cameras within the United Kingdom alone, massive amount of video data is being transmitted and recorded. The issue of whether it is watched, reviewed or analyzed is still questionable due to the lack of human resources, time and sheer volume of data. Consequently, criminal incidents and suspicious events may be missed and not noticed in time to prevent a true crime from taking place. For installations with a large number of cameras, it is obviously impossible for one or even several operators to watch all the cameras. Even in the unrealistic scenario that there is one security staff per camera, an empirical study published by Security Oz [42] argued that after observing 12 min of continuous video monitoring, the operator will often miss afterwards up to 45% of screen activities. After 22 min of watching, up to 95% is overlooked. Therefore, if surveillance video is not monitored carefully and acted upon, there is obviously an increased security risk for individuals and private facilities.

Because of the impossibility to watch and screen all surveillance videos all the times, the installation of surveillance systems is rendered limited in terms of effectiveness and efficacy. There is a significant and increasing demand to address such limitations that posed as a great challenge for researchers and industrial practitioners to provide solutions and theories to meet the overwhelming global security needs. According to Liu et al. [37], there are over 6000 research articles published in the literature since 1971 covering the area of intelligent visual systems and analytics. The concept floated to appearance in early 1970s but emerged to the research community during the 80s. The research progressed at faster pace since 2000 due to development of computational resources and availability of cameras at cheap prices. Most of the published papers in visual surveillance fall into one of the three aspects including hardware, software and surveillance applications. New smart visual surveillance systems are currently being developed to take video sequences, pre-process the data using computer vision techniques and then analyze the data to learn and detect interesting events and objects. For instance, vehicle license plates can be automatically recognized or virtual fences can be setup around critical facilities so that warnings can be raised in the event of an unauthorized access or intrusion. Elliott [17] has recently defined an Intelligent Video System (IVS) as:

> any video surveillance solution that utilizes technology to automatically, without human intervention, process, manipulate under/or perform actions to or because of either the live or stored video images

Smart or intelligent visual surveillance system is designed to collect, index and store video data allowing security and law enforcement agents to monitor, analyze and search for events or objects of interest at real time or previous times. Further, along with the sheer volume of online and offline video footage coming from multiple cameras, it becomes a necessity to develop efficient methods for the retrieval and analysis of video data based on semantic content. Security and military applications have been the main thrust in developing smart surveillance tools. From small security outpost, shopping malls to large cities with thousands of citizens. Recently, as the number of cameras has increased immensely by both governmental and private agencies for the protection of their citizens, employees and properties. Smart visual applications analyze and filter out the massive amount of data recorded from multiple continuous video feeds ensuring only appropriate alerts are presented for law enforcement officers. The automation for surveillance systems becomes an obligation to avoid human errors and ensure an efficient strategy for tackling crimes and preventing further terrorist attacks. The process of automation can be fulfilled either in a centralized fashion such that all processing is done within a single place where all video streams are sent. Alternatively, processing can be performed in a distributed way where even cameras can embed certain intelligence for detecting events or recognizing people. The automation of intelligent video systems can include simple tasks from people counting, pedestrians detection to further complex and intricate tasks as behavior analysis [33] and people identification.

Biometric Technologies

Biometrics is defined as the automated process of extracting and deriving descriptive measurements based on either the behavioral or physiological characteristics of the human body. Such measurements should distinguish an individual uniquely among other subjects. The derived biometric data is compared against a database of records to either verify or recognize the identity of an individual. The word *biometrics* is a composite word of two parts from the Greek language: *bios* means life meanwhile *metrics* refers to the verb “to measure”. Apart from the uniqueness condition, the biometric description should be universal and permanent. The universality factor implies that the biometric data can be acquired from all the population regardless of their gender, age, location or ethnicity. For the permanentness condition, it signifies that the biometric signature of an individual should stay the same throughout the different ages. Jain et al. [26] added further criteria that must be met for biometric systems including user acceptance, performance, unvulnerability and integration. As opposed to traditional identification or verification methods such as passports, passwords or pin numbers, biometrics cannot be transferred, forgotten or stolen and should be ideally obtained non-intrusively. Even though, people intuitively use some body characteristics such as face, gait or voice to recognize each other, it is proven a challenging task to extract and quantify such measurements into a biometric signature using computer vision methods.

The biometric measurements are based either on the physiological traits such as face, ear, fingerprint and DNA or based on the behavioral traits including gait, voice and signature. Figure 1.1 shows examples of both types of biometrics. Biometrics is now emerging in regular use being deployed in various applications such immigration border control, forensic systems, computer security and payment authentication. Biometric systems are sold mainly for the following purposes: physical access control, logging attendance and personal identification purposes. The choice of a specific biometric modality depends on the nature and requirements of the intended applications in addition to the availability of biometric features. Though in many cases, the combination of multiple biometric modalities may be needed to achieve the desired level of performance [46]. Fingerprints, iris and face are among the most popular physiological traits used successfully in commercial identification systems with fingerprint capturing over 50% of the market share [27]. The main reason for such popularity is the availability of large legacy databases which have been collected by law enforcement agencies all over the world [28]. There is no doubt that biometrics has a clear-cut benefits over the use of passwords or identification cards, there are still limitations that biometrics are vulnerable to such as spoofing, linkability attacks, evasion and database alteration [18, 28].

Biometric systems are setup to work either in identification or verification mode. For identification, a one-to-many matching process is conducted for newly acquired biometric data against all people already enrolled in the database in order to infer the identity of the subject whose matched biometric data exhibits the highest similarity value. For the case when the identification system is forced to infer the identity of

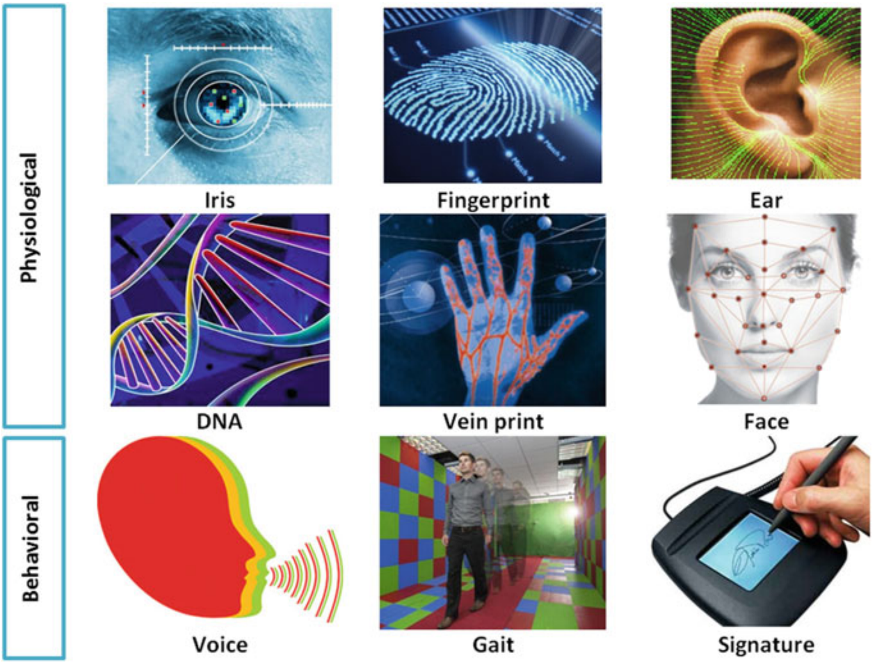

Fig. 1.1 Various types of biometric modalities

the unknown person from the list of people enrolled in the database, it is referred to as a *closed-set* biometric system. In contrast to the *open-set* biometric system when there can be an option to have a *reject* or *no match* response. For the verification mode, the system conducts a one-to-one match for the person's signature against a pre-recorded verified signature in the database to confirm the claimed identity. The identity claim is confirmed if the computed similarity score is above a preset threshold. Figure 1.2 shows an overview for a biometric system with the different components and steps involved from extracting the measurements to the matching process to infer the person identity.

For the history of biometrics, people have been recognizing each other based on voice, face or walking style for thousands of years. However, the first systematic basis for people identification dates back to 1858 when William Herschel recorded the handprint of each employee on the back of a contract whilst working for the civil service of India [7]. This was used as a way to distinguish staff from each other on payday. It was considered the first systematic capture of hand and fingerprints that was used primarily for identification purposes. For the use of biometrics in forensic analysis, Alphonse Bertillon who was a French police officer was the pioneer to use the first biometric evidence into the judicial system presented as anthropometric measurements of the human body to counter against repeat offenders who could easily change or fake their names and personal details. In 1879, he introduced the

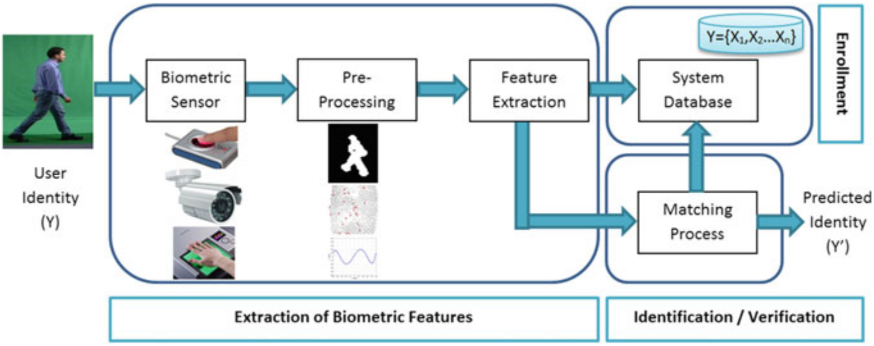

Fig. 1.2 Overview of a biometric system

bertillonage system such that every individual is identified through detailed records taken from their body anthropometric data and physical descriptions as they are impossible to spoof or change them. The system was adopted by the police authorities throughout the world. In 1903, the Bertillon system was scrutinized on the basis that anthropometric measurements are not adequate to differentiate reliably between people. This was fueled by the case of identical twins who have almost the same measurements using the Bertillon system. The first scientific research paper on the use of fingerprints for identification was published in 1880 by Henry Faults in Nature [28]. During 1890s, Sir Francis Galton and Edward Henry proposed separately classification systems for people recognition based on their fingerprints taken from all the ten fingers [2]. The characteristics that Galton described to identify people are still used today. The Galton system was adopted by New York state prison where fingerprints were utilized for the identification of apprehended inmates.

Gait Recognition

In spite of the fact that physiological biometric systems enjoy the merit of reliable identification rates, their use in surveillance systems is never considered due to the limitation of extracting them unintrusively. Interestingly, gait biometrics is reported by recent studies to be suitable for surveillance scenarios because it can be captured from a distance even with poor resolution imageries in contrast to other biometric modalities which require high resolution images. Moreover, the main potency of gait biometrics is the non-invasiveness property and therefore the person is not obliged to cooperate or interact with the acquisition hardware. The invasiveness property sets the identification procedure via gait ideal for cases where direct contact with the offender is not an option which is the case for all criminal cases. Moreover, a biometric signature constructed from the gait rhythmic motion pattern is considered the only likely identification method suitable for covert surveillance and reliably not prone to spoofing attacks and signature forgery. Consistently, many recent research studies concluded that gait recognition is more suitable for forensic science as other identification traits that can be related to the crime scene can be wiped out or concealed in contrast to the gait motion as the mobility of the person is a must for them to walk away from the crime scene. Recently, Lucas et al. [39] reported that a combination of eight anatomical measurements from the human body is enough to attain a probability to the order of 10^{-20} for a finding duplicate signature comparing such results to fingerprint analysis. Interestingly, one of the murder cases that attracted the media attention in the United Kingdom where a child was kidnapped and murdered, it was impossible for the investigating team to infer the identity of the killers from poor resolution surveillance video footage. The only inspiring method that could be used to reveal their identities in such delicate situation was the gait pattern as suggested by the research team from the University of Southampton [43]. The notion that people can be recognized by the way they walk has gained an increasing popularity and produced impacts on public policy and forensic practice by its take up by researchers at the Serious Organized Crime Agency.

In 1964, Murray et al. [41] performed the early experiments describing the standard gait pattern for normal walking people aimed at studying the gait pattern for pathologically abnormal patients. The medical investigations were conducted on sixty individuals aged between 20 and 65 years old. Each subject was instructed to walk for a repeated number of trials. For the collection of gait data, special markers were attached on every person. Murray et al. [41] suggested that human gait consists of 24 different components which make the gait pattern distinctive for every subject if all gait movements are considered. It was reported that the motion patterns of the pelvic and thorax regions are highly variable from one person to another. Furthermore, the study reported that the ankle rotation, pelvic motion and spatial displacements of the trunk embed the subject individuality due to their consistency at different experiments. In 1977, Cutting et al. [14] published a paper confirming the possibility of recognizing people by the way they walk via observing Moving Lights Displays (MLD) mounted on the joints positions. An MLD shown in Fig. 1.3 is a two-dimensional video of a collection of bright dots attached to the human body taken against a dark background where only the bright dots are visible in the scene. Different observers are asked to see the actors performing various activities. Based

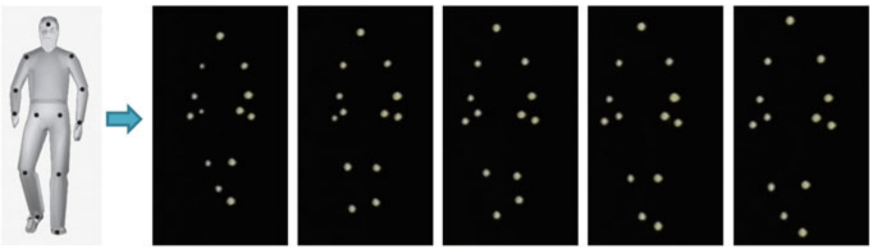

Fig. 1.3 Medical studies

on these experiments observers can recognize different types of human motion such as walking, jumping, dancing and so on [29]. Moreover, the observer can make a judgment about the gender of the performer [14], and even further identify the person if they are already familiar with their gait [14, 19]. Cutting argued that the recognition is purely based on dynamic gait features as opposed to previous studies which were confounded by familiarity cues, size, shape or other non-gait sources of information. Although, there is a wealth of gait studies in the literature aimed for medical use with only a few referring to the uniqueness nature of the walking pattern, none is concerned with the automated use of the gait pattern for biometrics. The gait measurements and results from the medical literature introduced by Cutting, Murray and Johansson are to be of benefit for the development of automated gait biometric systems using computer vision methods though the extraction of the gait pattern is proven complex.

Gait Recognition Methods

Vision-based system for people recognition via the way they walk, is designed to extract gait features without the need to use special sensors or reflective markers to assist the extraction process. In fact, all that is required is a video camera to stream images for vision-based software for processing. Marker-less motion capture systems are suited for applications where mounting sensors or markers on the subject is not an option as the case of outdoor surveillance applications. Typically, gait biometric system consists of two main components: (i) a hardware platform dedicated for data acquisition. This can be a single CCTV camera or distributed cameras network. (ii) A software platform for video processing and identification. The architecture of the software tier for gait biometrics is composed broadly of three main stages: (i) *detection and tracking of the pedestrian*: intra-camera tracking is performed to establish the correspondence of the same person across consecutive frames. (ii) *Feature extraction*: in order to estimate a set of measurements either related to the configuration of the whole body or the configuration of the different body parts in a given scene and tracking them over a sequence of frames. (iii) *Classification stage*: which involves matching a test sequence with an unknown label against a group of labelled references considered as the gallery dataset. Figure 1.4 shows the flow diagram for gait identification outlining the different subsystems involved in the process of an automated people recognition.

Much of the interest in the field of human gait analysis was limited to physical therapy, orthopedics and rehabilitation practitioners for the diagnosis and treatment of patients with walking abnormalities. As gait has recently emerged as an attractive biometric, gait analysis has become a challenging computer vision problem. Although, the distinctiveness of gait features cannot compare with traditional biometrics, it has proven to be a potential alternative for surveillance scenarios [28]. Many research studies have aimed to develop a system capable of overcoming the difficulties imposed by the extraction and tracking of biometric gait features. Various

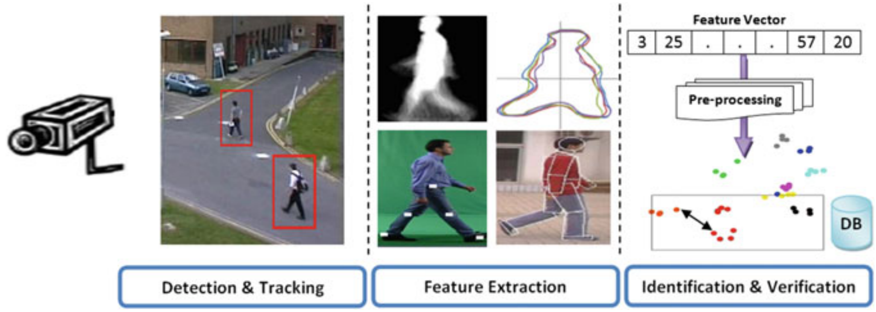

Fig. 1.4 Overview of gait biometric system

methods were surveyed in [43] and [56]. Based on the procedure for extracting gait features, gait recognition methods can be divided into two main categories which are model-based and appearance-based (model-free) approaches.

Appearance-Based Approaches

Appearance-based or model-free approaches for gait recognition do not need a prior knowledge about the gait model. Instead, features are extracted from the whole body without the need to explicitly extract the body parts. The majority of appearance approaches depends on data derived from silhouettes which are obtained via background subtraction. Appearance-based Method relies pivotally on statistical methods to reduce or optimize the dimensionality of feature space using methods such as Principal Component Analysis. In addition, advanced machine learning methods are usually applied as multi-class support vector machine and neural networks. Contentiously, recent investigations by Veres et al. [51] reported that most of the discriminative features for appearance-based approaches are extracted from static components of the top part of the human body whilst the dynamic components generated from the swinging of the legs are ignored as the least important information.

Silhouette-Based Methods

Silhouette-based methods work by separating walking people from the background. The simplest baseline method is to compute the similarity score between two synchronized silhouette sequences [47]. The *Gait Energy Image (GEI)* is another basic silhouette-based representation introduced by Han and Bhanu [21] in which gait signature is constructed through taking the average of silhouettes for one gait cycle. Experimental results confirmed that higher recognition rates can be attained to reach 94.24% for a dataset of 3,141 subjects [25]. However, such method performs poorly when changing the appearance. The *Motion Silhouette Image* (MSI) is a similar representation to GEI proposed by Lam et al. [34] where each pixel intensity is computed as a function of the temporal history of motion for the corresponding pixels across a complete gait cycle. Experiments conducted on the large SOTON gait dataset showed that 87% can be achieved. *Gait Entropy Image (GenI)* is a silhouette-based representation introduced by Bashir et al. [4] which is computed by calculating the Shannon entropy for each pixel achieving a correct classification rate of 99.1% on dataset of 116 subjects. The Shannon entropy estimates the uncertainty value associated with a random variable. Other similar representations include Motion Energy Image, Gait History Image, Frieze Patterns and Chrono-Gait Image.

For using the gait symmetric property, Hayfron-Acquah et al. [22] introduced a method for constructing a gait signature based on analysing the symmetry of human motion. The symmetry map is produced via applying the Sobel operator on the gait silhouettes followed by the Generalized Symmetry Operator. The symmetry map was evaluated on a dataset containing 28 people using the k-NN classifier, a high recognition rate of 96.4% was attained for the value of $k = 3$. There is a recent tendency to use model-free depth based representation using 3D sensors (Fig. 1.5). Sivapalan et al. [48] proposed the *Gait Energy Volume* descriptor (GEV) by extending the Gait Energy Image into 3D. The implementation for GEV was evaluated on the CMU MoBo database confirming that improvements can be attained over the 2D GEI version as well as fused multi-view GEI variant. Recently, there is a trend of employing Deep Learning and Neural Networks using Silhouette-based descriptors to account for the issue of view-invariance. Li et al. [36] proposed a gait representation called *DeepGait* using deep convolution features. Meanwhile, Wu et al. [55] utilized deep Convolution Neural Networks (CNNs) for gait recognition using the OU-ISIR gait dataset with a reported success rate of 94% for a population of 4,007 subjects. Zeng et al. [59] described a silhouette-based approach for view-invaraint gait biometrics using deterministic learning theory. Radial basis function (RBF) neural networks are used to locally approximate the gait dynamics from different view angles.

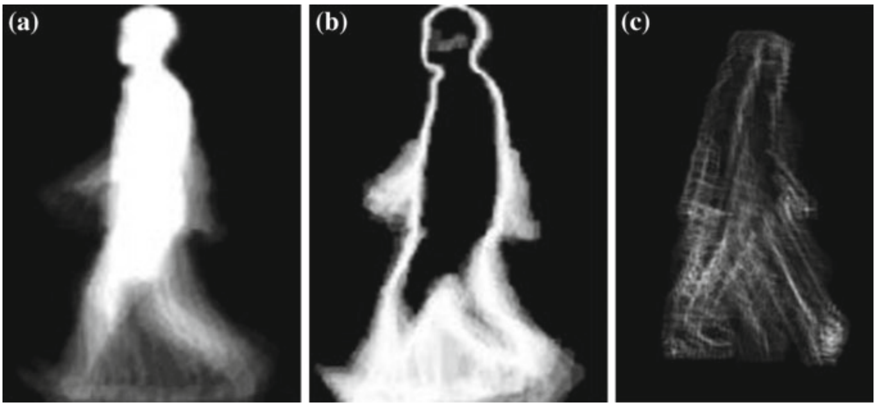

Fig. 1.5 Silhouette-based methods for gait recognition: **a** use gait energy image [21] **b** gait entropy image [4] **c** symmerty map [22]

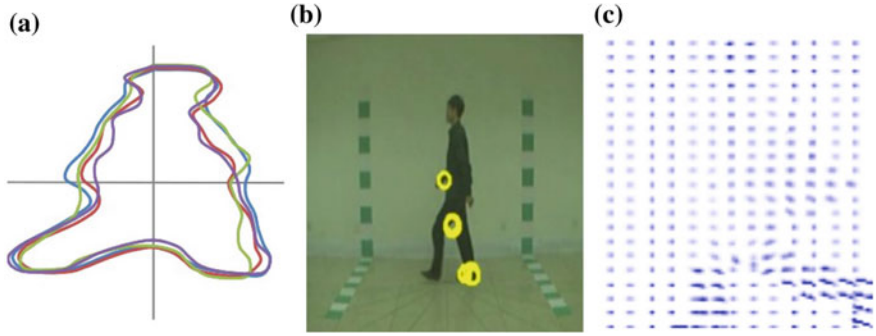

Fig. 1.6 Non-silhouette appearance-based methods for gait recognition: **a** procruste shape analysis [12] **b** STIP descriptors **c** optical flow

Non-silhouette Methods

As the accuracy of silhouette-based methods depend on the background segmentation algorithm which is not reliable for the case of real surveillance footage in addition to the sensitivity issue to varied appearances, a number of appearance-based methods have emerged recently that use instead interest-point descriptors. Kusakunniran et al. [32] proposed a framework to construct gait signature without the need to extract silhouettes. Features are extracted in both spatial and temporal domains using Space-Time Interest Points (STIPs) by considering large variations along both spatial and temporal directions at a local level. Bashir et al. [4] used the dense optical flow field computed using the method in [11] for each frame of the whole gait cycle to extract four different types of motion descriptors (Motion Intensity Image and four Motion Direction Images) based on the horizontal and vertical optical flow components and magnitude, their proposed experiments on the CASIA and SOTON gait datasets with the clothing and bag carrying covariates outperform previous reported studies. Hu et al. [23] described an incremental learning framework using optical flow for gait recognition. The Local binary pattern (LBP) operator is utilized to encode the textural information of optical flow (Fig. 1.6).

Model-Based Approaches

For the model-based approach, a prior model is established to match real images to this predefined model, and thereby extracting the corresponding gait features once the best match is obtained. Usually, each frame containing a walking subject is fitted to a prior temporal or spatial model to explicitly extract gait features such as stride distance, angular measurements, joints trajectories or anthropometric measurements. Although model-based approaches tend to be complex requiring high computational cost, these approaches are the most popular for human motion analysis due to their advantages [56]. The main strength of model-based techniques is the ability to extract detailed and accurate gait motion data with better handling of occlusion, self-occlusion and other appearance factors as scaling and rotation. The model can be either a 2 or 3-dimensional structural model, motion model or a combined model.

2D Structural Models

The structural model describes the topology of the human body parts as head, torso, hip, knee and ankle by measurements such as the length, width and positions. This model can be made up of primitive shapes based on matching against low-level features as edges. The stick and volumetric models are the most commonly used structural-based methods. Akita et al. [1] proposed a model consisting of six segments comprising of two arms, two legs, the torso and the head. Guo et al. [20] represented the human body structure by a stick figure model which had ten articulated sticks connected with six joints. Rohr et al. [45] proposed a volumetric model for the analysis of human motion using 14 elliptical cylinders to model the human body. Karaulova et al. [30] used the stick figure to build a hierarchical model of human dynamics represented using Hidden Markov Models. For the deployment of structural model-based methods for gait recognition, Niyogi et al. [44] was perhaps the pioneer in 1994 to use a model-based method for gait recognition. Gait signature is derived from the spatio-temporal pattern of a walking subject using a five stick model. Using a database of 26 sequences containing 5 different subjects, a promising classification rate of 80% was achieved. The recent trend of model-based approaches has shifted towards combining different cues including motion and 3D data in order to construct models able to handle the extraction of gait features (Fig. 1.7).

Motion Models

The motion model describes the kinematics or dynamics of the human body or its different parts throughout time. Motion models employ a number of constraints that aid the extraction process as the maximum range of the swinging for the low limbs. Cunado et al. [13] was the first to introduce a motion model using the Velocity Hough Transform to extract the hip angular motion via modeling human gait as a moving

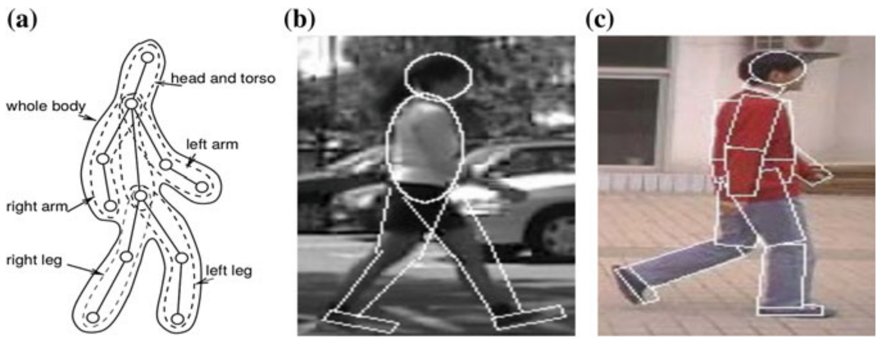

Fig. 1.7 Model-based approaches: **a** Karaulova [30]. **b** Wagg [52]. **c** Wang [53]

pendulum. The gait signature is derived as the phase-weighted magnitudes of the Fourier components. A recognition rate of 90% was achieved using the derived signature on a dataset containing 10 subjects. Yam et al. [57] modeled the human gait as a dynamic coupled oscillator which was utilized to extract the hip and knee angular motion via evidence gathering. The method was tested on a small dataset of 20 walking and running people, achieving a recognition rate of 91% based on gait signature derived from the Fourier analysis of the angular motion. Wagg et al. [52] proposed a new model-based method for gait recognition based on the biomechanical analysis of walking people. Mean model templates are adopted to fit individual subjects. Both the anatomical knowledge of human body and hierarchy of shapes are used to reduce the computational costs. The gait features vector is weighted using statistical analysis methods to measure the discriminatory potency of each feature. On the evaluation of this method, a correct classification rate of 95% is reported on a large database of 2,163 video sequences containing 115 different subjects (Fig. 1.7). Bouchrika et al. [9, 10] proposed a motion-based model for the extraction of the joints via the use a parametric representation of the elliptic Fourier descriptors describing the spatial displacements. The evaluation was carried out on a limited dataset containing 20 people from the Southampton dataset reporting a correct classification rate of 92% using the k-NN classifier.

3D Models

As most of the model-based methods are exploiting 2D images, there are recent work aimed for introducing 3-dimensional models for the extraction of gait features. 3D approaches for gait identification are known for their robustness of invariance to different viewpoints. Though, it is always a difficult task to acquire 3D data in surveillance data in addition to the high computational and financial costs involved. Recent studies on using 3D models include the work of Ariyanto et al. [3] who introduced a 3D approach using a marionette and mass-spring model to gait biometrics with 3D voxel gait data. The articulated human body is modeled using the stick-figure which emulates the marionettes' motion and joint structure. The stick-figure is composed of 11 nodes corresponding to the human joints including the head. Zhao [60] employed local optimization algorithms in order set up the 3D model from video sequences captured from multiple cameras. Features for the classification are derived from body segments including static parameters in addition to dynamic features for the motion trajectories of the limbs. Recently, Tang et al. [50] proposed an approach for gait partial similarity matching which is based on the assumption that 3-dimensional objects share similar view surfaces across different views. 3D parametric models for the human body are morphed by pose and shape deformation from a template model using 2-dimensional silhouette data. López-Fernández et al. [38] used 3D angular data to create a view-invariant gait biometric signature for people walking on unconstrained paths. Support vector machine is used for the classification process on two datasets containing 62 subjects in total. Kastaniotis et al. [31] used the Microsoft Kinect sensor to recognize people via a pose estimation process to extract the skeleton of a walking person. Based on a publicly available dataset of 30 people, a high recognition rate of 93.3% is reported.

Areas of Applications

Re-Identification

Due to the unprecedented surge for the number of crimes and terror attacks, the deployment of surveillance technology has become ubiquitous on our modern society. It is no wonder that many military, corporate and government agencies devoted a large amount of funding to research institutions to advance research for the deployment of biometric technologies within their surveillance systems to ease the management and monitoring for ensuring the safely of their citizens or assets. Due to the limitation inherited from using a single camera which can only cover a limited field of view, the deployment of interconnected set of cameras can becomes more prevalent in sensitive and crowded areas including shopping malls and airports. Assuming that pedestrians tracking within a single camera is adequately accurate and reliable, the re-identification issue is reduced thereafter to associating subjects of the same identity across different cameras. People re-identification is the task of tracking people regardless of their identity across different cameras [16]. The process to trace and follow the whereabouts of pedestrians and possibly producing a semantic description of their behaviour from visited places would be a major breakthrough to automated surveillance for law enforcement officers.

As tracking people within a single field of view can be performed reliably and robustly, it is a difficult task to have the tracking done within a network of cameras when it comes to solve the association problem for the same subjects seen at different places, different times from different cameras without overlapping views. The complexity for re-identification is mainly exacerbated from the variation of people appearances across different cameras [6]. This is besides other challenging factors including image quality, crowded scenes and occlusion. The articulated property of the human body is another prime reason for the appearance to subtly change continuously. Surveillance systems consisting of interconnected cameras cover large spatial areas with non-overlapping views to yield enhanced coverage. Therefore, the complexity of tracking people would proportionally increase with size of the area and number of cameras rending the process of re-identification and extremely resource intensive task temporally and spatially. There are a number of databases being setup and made publicly available for researchers to tackle the problem of people re-identification for surveillance applications. Bedagkar-Gala [6] published recently a survey on the different methods for people re-identification, performance metrics and datasets. The author classifies the techniques for people re-identification into three main categories: (i) Descriptor learning which is based on deriving the most discriminative characteristics on the fly or instead a learning phase is setup in order to derive a descriptive dictionary of features that better represent distinctively a person's appearance using bag-of-features approach. (ii) Distance metric learning aims to maximize the matching score between different identities regardless of the choice for appearance cues or color calibration using different types of features. (iii) Color Calibration approaches is based on modeling the color relationship across different cameras and update such model regularly for each camera. The color brightness transfer function (BTF) is employed to establish the association between the same people across different camera viewpoints.

Gait is an emergent biometrics which has attracted unprecedented interest from the research community. It enjoys the potency to overcome most of the drawbacks that other biometric modalities suffer from. Gait biometrics is deemed suitable for re-identification applications partly due to the fact that the gait motion can be recorded and extracted from a distance besides the non-invasive and less-intrusive property. For the process of markerless extraction of gait features from surveillance videos, a Haar-like template matching procedure is presented by Bouchrika [10] to locate the spatial positions of the lower limbs for a single walking person. The method is not dependent on foreground segmentation for the extraction of gait features. This is mainly because it is resource intensive and expensive process to perform background subtraction for real-time scenarios due to the task of updating the background model. The motion models are taken from the medical data reflecting the angular motion for the hip and knee within one full gait cycle. For the initial phase, the described approach builds a motion map from the change detection of the inter-frame differences. The drawback of depending on frame differencing is that the camera has to be stationary. Moving pixels belonging to moving people across successive frames are extracted with the emphasis to extract clear contour data. Bedagkar-Gala [5] combined the use of colors with gait features including Gait Energy Image and Frame Difference Energy Images in order to track 40 people across 9 different cameras taken from the SAIVT SoftBio dataset. Wei et al. [54] proposed a Swiss-system based cascade ranking model for re-identification of people in surveillance using five gait silhouette-based descriptors. The Swiss-system uses multi-feature ensemble learning where a series of rankers are applied to every pair of matches.

Forensic Analysis

Forensic gait analysis has been recently applied in investigations at numerous criminal cases as law enforcement officers have no option to identify the perpetrator using well-established methods as facial recognition or fingerprints. This is partly due to the fact that key biometric features such as the perpetrator's face can be obscured or veiled and the CCTV footage is deemed unusable for direct recognition whilst the perpetrators are usually filmed at a distance walking or running away from the crime scene. Gait experienced specialists are consulted to assist with the identification process of an unknown person by their walking pattern through a qualitative or quantitative matching process. This would involve examining the unique and distinctive gait and posture features of an individual. Subsequently, a statement is written expressing an opinion or experimental results that can be used in a court of law. Meanwhile, the practice of forensic podiatry involves examining the human footprint, footwear and also the gait pattern using clinical podiatric knowledge [15]. However, gait analysis performed by a podiatrist involves the recognition and comparison of nominal and some ordinal data without quantitative analysis using numerical forms of data [15]. Because of the rising profile of gait biometrics and forensic podiatry, gait is used numerously as a form of evidence in criminal prosecutions with the inauguration of the American Society of Forensic Podiatry in 2003. Recently, Iwama et al. [24] developed a software with a graphical user interface in order to assist non-specialists to match video sequences based on gait analysis. For the methods used for forensic gait analysis, Bouchrika et al. [8] classifies them into two major categories which are: *Descriptive-based* or *Metric-Based* approaches.

An incident of a bank robbery in 2004 was handled by the Unit of Forensic Anthropology at the University of Copenhagen [40]. The police observed that the perpetrator has a special gait pattern with a need to consult gait practitioners to assist with the investigation. The police were instructed to have a covert recording of the suspect walking pattern within the same angle as the surveillance recordings for consistent comparison. The gait analysis revealed that there are several matches between the perpetrator and the suspect as an outward rotated feed and inverted left ankle during the stance phase. Further posture analysis using photogrammetry showed that there is a resemblance between the two recordings including a restless stance and anteriour head positioning. There were some incongruities observed during the analysis including wider stance and the trunk is slightly leaned forward with an elevated shoulders. This is suspected to be related by the anxiety when committing a crime [35]. Based on the conducted analysis, a statement was given to the police regarding the identity however such methods are argued that they do not constitute the same level of confidence as well-established methods such as fingerprints. The findings were subsequently presented in court and the suspect was convicted of robbery whilst the court stressed that gait analysis is a valuable tool [35]. In a similar case handled by the Unit of Forensic Anthropology, a bank robbery was committed by two masked people wearing white clothing. The bank was equipped with several cameras capturing most of the indoor area. One of the camera showed one of the perpetrator walking rapidly from the entrance. The frame rate was low which left only few useful images showing the perpetrator gait. Based on experimental results showing the most discriminatory potency for the joints angles, Yang et al. [58] argued about the possibility of identification based on certain instances of the gait cycle using the observed angular swinging of the legs. Figure 1.8 shows the two discussed cases of the bank robberies handled by the forensic unit.

In a recent case handled by the Metropolitan Police of London [8], a number of crimes include physical assaults and burglary against pedestrians walking on a short pathway near a subway in one of the London suburb. The same crime was reported to occur numerous times in the same fashion and at the same place. The police officers strongly suspected it was carried out by the same members of an organized gang of youngsters aged between 17 and 20 years old. There are a number of CCTV cameras in operation at the crime scene. Two of them are pointing towards the entrances of the subway as shown in Fig. 1.9. Two other cameras are set to record both views of the walking pass next the subway. The police provided a set of videos in order to deploy gait analysis to find further information that would assist them in their investigation. CCTV footage from all cameras for the crime scene at two different

(a) (b)

Fig. 1.8 Forensic gait analysis using descriptive-based methods [35, 58]: **a** 2004 **b** 2012

(a) (b)

Fig. 1.9 Forensic gait analysis using metric-based methods: **a** 2011, UK case [8] **b** 2013, Australian case [49]

days was made available to the Image Processing Research group at the University of Southampton. The police provided another video of a suspect member of the gang being recorded whilst was being held at the police custody. The video was recorded at a frame rate of 2 frames per second and a resolution of 720×576 pixels. In one of the videos that was recorded on 4th April 2008, two members of the gang wore helmets to cover their faces and drove a scooter motorbike. A female pedestrian came walking through the subway where they followed her from behind on the walking path. When she entered the subway, one of them walked and snatched her bag violently using physical assault and even dragging her down on the ground. Afterwards they left away on a scooter. In a different CCTV footage recorded on the following day, the same crime was carried out with apparently the same looking perpetrators riding a scooter motorbike seen snatching a bag of another woman. The Instantaneous Posture Matching (IPM) [8] is proposed to conduct the forensic analysis which is based on computing the normalized distances for the joints positions between two video sequences on a specific window of frames. To access the measurement confidence, empirical study is conducted to explore the intra- and inter-match scores produced by Instantaneous Posture Matching on a larger gait dataset. Sun et al. [49] suggested the potential use of gait acceleration for crime analysis via the detection of heel strikes.

Conclusion

Because of the unprecedented number of crimes and terror attacks as well as the vital need to provide safer environment, a surge of concerns has emerged in many countries to ensure the safety of their citizens via the use of advanced surveillance technology. It is no wonder that many cities have deployed a number of biometric technologies within their surveillance systems to ease the management and monitoring from a large number of cameras. Biometrics can be of benefits not only for identify recognition, but it can play a vital role to enhance the automation process for surveillance systems. Gait is an emergent biometrics which has attracted unprecedented interest from the research community. In contrast to other biometric modalities which requires high resolution images, gait is suitable for covert recognition in surveillance scenarios mainly to the non-invasiveness property and therefore the person is not obliged to cooperate or interact with the acquisition hardware.

References

1. Akita K (1984) Image sequence analysis of real world human motion. Pattern Recognit 17(1):73–83
2. Arbab-Zavar B, Xingjie W, Bustard J, Nixon MS, Li CT (2015) On forensic use of biometrics. In: Handbook of digital forensics of multimedia data and devices
3. Ariyanto G, Nixon MS (2012) Marionette mass-spring model for 3d gait biometrics. In: 5th international conference on biometrics. IEEE, pp 354–359
4. Bashir K, Xiang T, Gong S (2009) Gait recognition using gait entropy image. In: 3rd international conference on crime detection and prevention, pp 1–6
5. Bedagkar-Gala A, Shah SK (2014a) Gait-assisted person re-identification in wide area surveillance. In: Asian conference on computer vision. Springer, pp 633–649
6. Bedagkar-Gala A, Shah SK (2014b) A survey of approaches and trends in person re-identification. Image Vis Comput 32(4):270–286
7. Berry J, Stoney DA (2001) The history and development of fingerprinting. Adv Fingerprint Technol 2:13–52
8. Bouchrika I (2017) Evidence evaluation of gait biometrics for forensic investigation. In: Multimedia forensics and security. Springer, pp 307–326
9. Bouchrika I, Nixon MS (2006) Markerless feature extraction for gait analysis. In: Proceedings of IEEE SMC chapter conference on advanced in cybernetic systems, pp 55–60
10. Bouchrika I, Carter JN, Nixon MS (2016) Towards automated visual surveillance using gait for identity recognition and tracking across multiple non-intersecting cameras. Multimed Tools Appl 75(2):1201–1221
11. Brox T, Bruhn A, Papenberg N, Weickert J (2004) High accuracy optical flow estimation based on a theory for warping. In: European conference on computer vision. Springer, pp 25–36

12. Choudhury SD, Tjahjadi T (2012) Silhouette-based gait recognition using procrustes shape analysis and elliptic fourier descriptors. Pattern Recognit 45(9):3414–3426
13. Cunado D, Nixon MS, Carter JN (2003) Automatic extraction and description of human gait models for recognition purposes. Comput Vis Image Underst 90(1):1–41
14. Cutting JE, Kozlowski LT (1977) Recognizing friends by their walk: gait perception without familiarity cues. Bulletin Psychon Soc 9(5):353–356
15. DiMaggio JA, Vernon W (2011) Forensic podiatry principles and human identification. In: Forensic podiatry. Springer, pp 13–24
16. Doretto G, Sebastian T, Tu P, Rittscher J (2011) Appearance-based person reidentification in camera networks: problem overview and current approaches. J Amb Intel Human Comput 2(2):127–151
17. Elliott D (2010) Intelligent video solution: a definition. Security 47(6)
18. Evans N, Marcel S, Ross A, Teoh ABJ (2015) Biometrics security and privacy protection. IEEE Signal Process Mag 32(5):17–18
19. Goddard NH (1992) the perception of articulated motion: recognizing moving light displays. PhD thesis, University of Rochester
20. Guo Y, Xu G, Tsuji S (1994) Understanding human motion patterns. In: Pattern recognition, conference B: computer vision & image processing, proceedings of the 12th IAPR international conference on 2
21. Han J, Bhanu B (2006) Individual recognition using gait energy image. IEEE Trans Pattern Anal Mach Intel 28(2):316–322
22. Hayfron-Acquah JB, Nixon MS, Carter JN (2003) Automatic gait recognition by symmetry analysis. Pattern Recognit Lett 24(13):2175–2183
23. Hu M, Wang Y, Zhang Z, Zhang D, Little JJ (2013) Incremental learning for video-based gait recognition with lbp flow. IEEE Trans Cybern 43(1):77–89
24. Iwama H, Muramatsu D, Makihara Y, Yagi Y (2012a) Gait-based person-verification system for forensics. In: IEEE fifth international conference on biometrics: theory, applications and systems (BTAS), pp 113–120
25. Iwama H, Okumura M, Makihara Y, Yagi Y (2012b) 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
26. Jain A, Ross AA, Nandakumar K (2011) Introduction to biometrics. Springer Science & Business Media
27. Jain AK, Kumar A (2012) Biometric recognition: an overview. In: Second generation biometrics: the ethical, legal and social context, pp 49–79
28. Jain AK, Nandakumar K, Ross A (2016) 50 years of biometric research: accomplishments, challenges, and opportunities. Patt Recognit Lett
29. Johansson G (1973) Visual perception of biological motion and a model for its analysis. Perception and psychophysics 14:201–211
30. Karaulova IA, Hall PM, Marshall AD (2000) A hierarchical model of dynamics for tracking people with a single video camera. In: Proceedings of the 11th british machine vision conference 1:352–361
31. Kastaniotis D, Theodorakopoulos I, Economou G, Fotopoulos S (2016) Gait based recognition via fusing information from Euclidean and Riemannian manifolds. Pattern Recognit Lett 84:245–251
32. Kusakunniran W (2014) Recognizing gaits on spatio-temporal feature domain. IEEE Trans Inf Forensics Secur 9(9):1416–1423
33. Ladjailia A, Bouchrika I, Merouani HF, Harrati N (2015) On the use of local motion information for human action recognition via feature selection. In: 4th international conference on electrical engineering (ICEE), 2015. IEEE, pp 1–4
34. Lam TH, Lee RS (2006) A new representation for human gait recognition: motion silhouettes image (MSI). In: International conference on biometrics. Springer, pp 612–618
35. Larsen PK, Simonsen EB, Lynnerup N (2008) Gait analysis in forensic medicine. J Forensic Sci 53(5):1149–1153

36. Li C, Min X, Sun S, Lin W, Tang Z (2017) Deepgait: a learning deep convolutional representation for view-invariant gait recognition using joint bayesian. Appl Sci 7(3):210
37. Liu H, Chen S, Kubota N (2013) Intelligent video systems and analytics: a survey. IEEE Trans Ind Inform 9(3):1222–1233
38. López-Fernández D, Madrid-Cuevas FJ, Carmona-Poyato A, Muñoz-Salinas R, Medina-Carnicer R (2016) A new approach for multi-view gait recognition on unconstrained paths. J Vis Commun Image Represent 38:396–406
39. Lucas T, Henneberg M (2015) Comparing the face to the body, which is better for identification? Int J Legal Med 1–8
40. Lynnerup N, Vedel J (2005) Person identification by gait analysis and photogrammetry. J Forensic Sci 50(1):112–118
41. Murray MP (1967) Gait as a total pattern of movement. Am J Phys Med 46(1):290–333
42. Nilsson F et al (2008) Intelligent network video: understanding modern video surveillance systems. CRC Press
43. Nixon MS, Tan TN, Chellappa R (2005) Human identification based on gait. Springer, New York, Inc. Secaucus, NJ, USA
44. Niyogi SA, Adelson EH (1994) Analyzing and recognizing walking figures in XYT. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition, pp 469–474
45. Rohr K (1994) Towards model-based recognition of human movements in image sequences. CVGIP Image Underst 59(1):94–115
46. Ross AA, Nandakumar K, Jain A (2006) Handbook of multibiometrics, vol 6. Springer Science & Business Media
47. Sarkar S, Phillips PJ, Liu Z, Vega IR, Grother P, Bowyer KW (2005) The humanID gait challenge problem: data sets, performance, and analysis. IEEE Trans Pattern Anal Mach Intel 27(2):162–177
48. Sivapalan S, Chen D, Denman S, Sridharan S, Fookes C (2011) Gait energy volumes and frontal gait recognition using depth images. In: International Joint Conference on, Biometrics (IJCB), 2011. IEEE, pp 1–6
49. Sun Y, Hare J, Nixon M (2016) Detecting acceleration for gait and crime scene analysis. http://eprints.soton.ac.uk/402396/
50. Tang J, Luo J, Tjahjadi T, Guo F (2017) Robust arbitrary-view gait recognition based on 3d partial similarity matching. IEEE Trans Image Process 26(1):7–22
51. Veres GV, Gordon L, Carter JN, Nixon MS (2004) What image information is important in silhouette-based gait recognition? In: Computer vision and pattern recognition, 2004, CVPR. IEEE computer society conference on, proceedings of the 2004. IEEE, vol 2, pp II–776
52. Wagg DK, Nixon MS (2004) On automated model-based extraction and analysis of gait. In: Proceedings of the sixth IEEE international conference on automatic face and gesture recognition, pp 11–16
53. Wang L, Ning H, Tan T, Hu W (2004) Fusion of static and dynamic body biometrics for gait recognition. IEEE Trans Circuits Syst Video Technol 14(2):149–158
54. Wei L, Tian Y, Wang Y, Huang T (2015) Swiss-system based cascade ranking for gait-based person re-identification. In: Twenty-ninth AAAI conference on artificial intelligence, pp 1882–1888
55. Wu Z, Huang Y, Wang L, Wang X, Tan T (2017) A comprehensive study on cross-view gait based human identification with deep CNNS. IEEE Trans Pattern Anal Mach Intel 39(2):209–226
56. Yam CY, Nixon M (2009) Gait recognition, model-based. In: Li S, Jain A (eds) Encyclopedia ofbiometrics. Springer, US, pp 633–639
57. Yam CY, Nixon MS, Carter JN (2004) Automated person recognition by walking and running via model-based approaches. Pattern Recognit 37(5):1057–1072
58. Yang SX, Larsen PK, Alkjær T, Simonsen EB, Lynnerup N (2013) Variability and similarity of gait as evaluated by joint angles: implications for forensic gait analysis. J Forensic Sci, pp 1556–4029

59. Zeng W, Wang C (2016) View-invariant gait recognition via deterministic learning. Neurocomputing 175:324–335
60. Zhao G, Liu G, Li H, Pietikainen M (2006) 3d gait recognition using multiple cameras. In: 7th international conference on automatic face and gesture recognition (FGR06). IEEE, pp 529–534