Int J Comput Vis (2011) 91: 7–23
DOI 10.1007/s11263-010-0362-6

Robust Gait Recognition by Learning and Exploiting Sub-gait Characteristics

Ibrahim Venkat · Philippe De Wilde

Received: 21 July 2009 / Accepted: 25 May 2010 / Published online: 18 August 2010

Abstract Gait recognition algorithms often perform poorly because of low resolution video sequences, subjective human motion and challenging outdoor scenarios. Despite these challenges, gait recognition research is gaining momentum due to increasing demand and more possibilities for deployment by the surveillance industry. Therefore every research contribution which significantly improves this new biometric is a milestone. We propose a probabilistic sub-gait interpretation model to recognize gaits. A sub-gait is defined by us as part of the silhouette of a moving body. Binary silhouettes of gait video sequences form the basic input of our approach. A novel modular training scheme has been introduced in this research to efficiently learn subtle sub-gait characteristics from the gait domain. For a given gait sequence, we get useful information from the sub-gaits by identifying and exploiting intrinsic relationships using Bayesian networks. Finally, by incorporating efficient inference strategies, robust decisions are made for recognizing gaits. Our results show that the proposed model tackles well the uncertainties imposed by typical covariate factors and shows significant recognition performance.

Keywords Gait recognition · Biometrics · Human motion analysis · Bayesian Network · Machine learning

1 Introduction

Biometrics is the study of automated methods for recognizing people, and a biometric is a physiological or behavioral characteristic that can be used to identify a person. A key objective of machine vision researchers is to build automated recognition systems that can compete and eventually surpass human visual intelligence. The importance of biometric technology as a tool for addressing national security challenges has been the focus of federally supported research agencies such as the Defense Advanced Research Projects Agency (US-DARPA) and the Intelligence Advanced Research Projects Activity (US-IARPA). The pressing need of the Intelligence Community for reliable biometric recognition performance has been expanding since 9/11. This expansion is expected to evolve beyond the scope of today's access control and verification applications that operate within tightly controlled conditions. Identifying individuals at a distance by observing their walking behavior when other features such as face and hand geometry are not clearly visible is a common task for the human visual system. Processing this human behavior by machines to identify individuals, which is termed gait recognition, is an ongoing investigation. In this study we address the gait recognition problem under unconstrained scenarios using a probabilistic sub-gait interpretation model.

Literature shows that in the early 1970's medical studies have first tried to treat gait as a discriminating trait (Johansson 1973). Only in the early 1990's was gait recognition addressed by machine vision researchers (Nixon et al. 2006). Biometrics such as fingerprint and iris are considered direct signatures of the physiology of an individual. Obtaining them is intrusive by nature. On the other hand, gait biometrics deal with the behavioral aspect, specifically the pattern of shape and motion in a video of a walking individual (Liu and Sarkar 2006). Like a biometric based on face recognition, it is unobtrusive in nature. Mature biometrics such as fingerprint, face and iris have crossed the experimental stage. But gait recognition is still in its infancy. Approaches

I. Venkat (✉) · P. De Wilde
Department of Computer Science, School of Mathematical and Computer Sciences, Heriot-Watt University, Riccarton, EH14 4AS, Edinburgh, UK
e-mail: vk52@hw.ac.uk

from researchers in various fields such as psychophysics, cognitive science and machine vision have tried to meet the various challenges thrown up by gait recognition. Recent advances in computing technology, especially high-speed processors, bulk storage and memory resources, are enabling gait as a practicable biometric right now, although the idea has emerged decades ago (Nixon and Carter 2006). Gait recognition can be broadly divided into three categories namely temporal alignment-based, static parameter-based and silhouette shape-based (Liu and Sarkar 2006). The first category considers both shape and dynamics and treats gait sequences as time series-based patterns. The second category characterizes human motion based on parameters such as stride length, speed etc. Approaches that use static body parameters such as ratio of sizes of various body parts need to be much improved to be practical. Silhouette-based gait recognition techniques are gaining much interest among current gait recognition researchers (Boulgouris et al. 2006; Nixon et al. 2006; Bauckhage et al. 2006). The reason is that they do not need any further information such as color, texture or grey-scale metrics. Intuitively the silhouette, which represents the binary map of walking humans, forms a robust feature to represent gait. This is because it captures the motion of most of the body parts (Kale et al. 2004). Recent studies Collins et al. (2002), Veeraraghavan et al. (2004) have shown that silhouette shape has equal, if not more, recognition potential than gait kinematics as referred by Liu and Sarkar (2005). Motivated by these approaches we too mainly use binary silhouettes.

In the class of object recognition problems, identification is considered to be harder than verification (Liu and Sarkar 2005). Although gait can disclose more than identity, it is increasingly being applied to identification tasks (Bauckhage et al. 2006). We don't address the verification problem where a single probe is matched with a single gallery, a one-to-one match. Rather we focus on the identification problem which matches a given probe gait sequence against a set of gallery gait sequences, a one-to-many matching.

Our PRObabilistic Sub-gait Interpretation Model which we abbreviate as (PROSIM) is based on a fundamental insight about human pattern matching and memory. While reasoning with objects which are prone to uncertainties, in our case visual processing of gaits, humans are often able to notice similarities between sub-gaits and gaits. When we see a person at a distance, we may notice a particular pattern of arm-swinging or hip movement as a characteristic of the whole walking gait of that person. A sub-gait is defined by us as part of the silhouette of a moving body. A formal definition of sub-gait is given in (2)–(6) of Sect. 3.1. First a set of sparse components or sub-gaits of the cluttered gait pattern is perceived, this is the *probe*. These are then matched to a bulk set of gait patterns, the *gallery*, that are remembered. This reasoning based on similarity mapping is processed in such a way to reveal inherent conditional independencies between gaits. In our study we intend to scientifically represent these independencies using Bayesian Networks (BN). BNs serve as fundamental tools in tackling uncertainty problems as they characterize intuitive notions of human reasoning. In other words, PROSIM employs BNs to find out and learn intrinsic sub-gait mappings that naturally exist in gait patterns. We derive robust probabilistic decisions by exploiting the mappings established. We have identified three potential sub-gaits among the possible sub-gaits of a gait silhouette, by experimental evaluation. Selecting potential sub-gaits is based on how significantly they contribute to the recognition mechanism of gaits. We will provide details in Sect. 4.2.

We briefly present PROSIM's architecture with the aid of the flow diagram shown in Fig. 1. Firstly we decompose the gallery silhouettes into sub-gaits and subject them to an appropriate feature extraction process to construct a low dimensional feature space. PROSIM is a generic model which could be applied to any feature space (subspace) projection technique, such as PCA or SVM. For demonstration sake we have used the recently proposed MPCA feature space (Lu et al. 2008). PROSIM further learns subtle sub-gait characteristics using a novel modular training scheme introduced in this study. Also using standard machine learning procedures, PROSIM estimates the parameters of the BN. All these preliminary activities are performed off-line to make minimal use of computing resources. Secondly the probe silhouettes are decomposed into similar sub-gaits and their extracted features are projected onto the feature space. Then gaits are shortlisted with the aid of similarity mapping-based reasoning. The intrinsic relationships between the sub-gaits and gaits are represented intuitively using BNs. Finally gaits are recognized by exploiting these relationships using robust probabilistic inference techniques.

The rest of the paper is organized as follows. Section 2 briefly describes papers related to our research problem and the relevance of Bayesian Network-based approaches to uncertainty problems. Section 3 presents the theoretical constructs of PROSIM. Section 4 discusses experimental results and evaluates its performance. Section 5 concludes our research.

2 Related Work

2.1 Motivation from Component-Based Object Models

Object recognition techniques can be classified as holistic or component-based. Holistic (global) approaches treat the entire object as one unit and are characterized by the lack of a-priori decomposition of the image into semantically meaningful components. Component-based models subdivide the object under study into components, then process

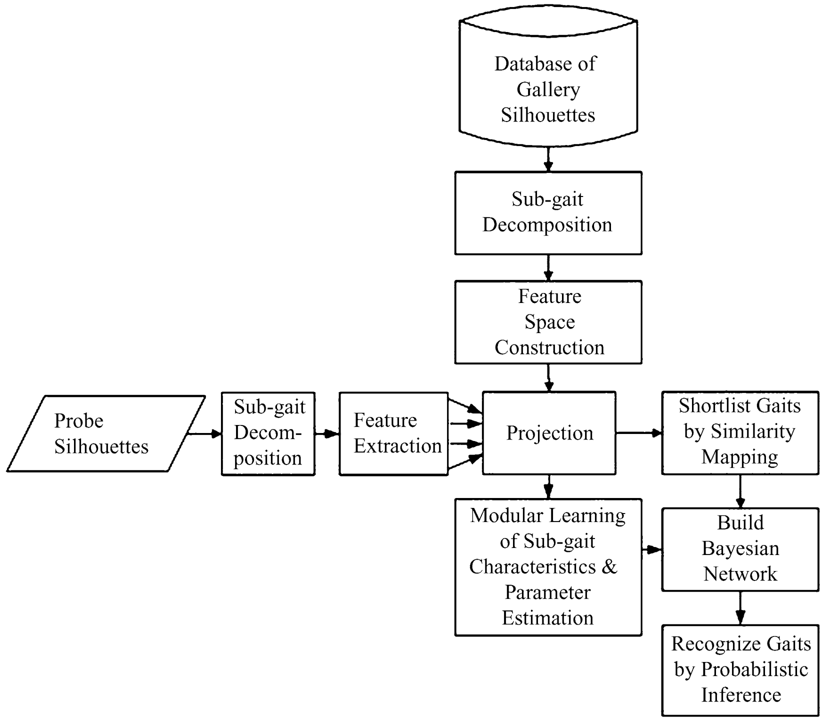

Fig. 1 Framework of the proposed probabilistic model

and manipulate these components, to finally classify them based on one-to-many and many-to-one mappings. Due to the intricacies of real world scenarios, the focus has shifted from holistic approaches to representations of individual object parts linked by structural information, with richer contextual descriptions of object configurations (Bergtholdt et al. 2009). Pioneering medical studies (Murray et al. 1964; Murray 1967) infer that human gait is composed of movements of many body parts as referred by Kale et al. (2004). This indicates that gait recognition performance can be enhanced by a component-based representation and analysis. We will survey some typical gait-based component models here. Recently a human body component-based approach has been proposed (Boulgouris and Chi 20007). Body components are manually labeled and weights have been assigned to them based on their contribution to recognition performance. By combining the results of various body components into a common distance metric, improved recognition performance has been achieved. In Lee and Grimson (2002), the silhouettes of gait sequences are subdivided into seven regions, fitted into ellipses and a set of moment related features are computed. This component-based approach shows improved recognition and gender classification using a small dataset captured at indoors. In Bauckhage et al. (2006) a method to establish homeomorphisms between 2D lattices and binary silhouettes is proposed. This method provides a robust vector space embedding of segmented body silhouettes. Feature vectors obtained from this scheme show improved detection of abnormal gait. Li et al. (2008) have proposed a component based approach by segmenting silhouettes into seven components, namely head, arm, trunk, thigh, front-leg, back-leg and feet. The effectiveness of these components for gait recognition and gender recognition has been analyzed. The approach relies on manually selected control points. These studies show that component-based algorithms have been attempted for gait and can lead to performance improvement.

Performance of the newly emerging gait recognition can be significantly improved by getting insight into related object recognition studies. A component-based person detection system has been proposed by Mohan et al. (2001) to detect humans in cluttered scenes. Initially example-based detectors are trained to find whether components of the human body are present in the specified geometric configuration. After ensuring this, a pattern classifier finally performs the recognition task. The authors have highlighted that the system is robust to covariate factors such as low resolution and occlusion due to the component-based approach. Early holistic approaches (Hallinan et al. 1999; Turk and Pentland 1991) used the intensity pattern of the whole face as input and modeled the photometric variation by linear combination of the eigenfaces. Although these well-known methods captured some geometric and photometric variations, they are limited in handling large-scale structural variations due to their fixed topology and holistic assumptions (Xu et al. 2008). Component-based algorithms have been proven to be superior for recognizing faces under unconstrained scenarios prone to local distortions, imprecise localizations, occlusions and variations in expression (Kim et al. 2005;

Martinez 2002). Heisele et al. (2007) compare holistic versus component-based approaches to perform face detection and identification in the presence of pose and illumination variance. Their experimental results clearly show that their component-based approach is much superior than the global approach. Abeni et al. (2006) compare an exclusive unified face recognition approach with a hybrid component-based approach. They show that the component-based system significantly outperforms the holistic system. These studies motivates us to consider an important phenomenon. That is performance degradation factors such as viewing angle (pose), occlusion, illumination that are inherent in typical object recognition scenarios (in our case gait recognition), could be mitigated well by component-based approaches.

2.2 Bayesian Models to Address Uncertainties

A Bayesian network (or a belief network) is a probabilistic graphical model that represents a set of variables and their probabilistic independencies. Bayesian models have been applied in various applications that work in unconstrained and realistic environments and they are now the mainstay of the AI research field known as "reasoning under uncertainty" (Jensen and Nielsen 2007). Bayesian networks have been proposed as tools to systematically analyze how humans make judgments under uncertainty (Krynski and Tenenbaum 2007). Intille and Bobick (2001) have proposed a Bayesian framework to recognize highly structured, multi-person action amidst multiple sources of visual perceptual uncertainties. Dahyot et al. (2004) have suggested a Bayesian approach inspired by probabilistic principal component analysis to detect objects subject to cluttered backgrounds coupled with occlusions. Tong et al. (2007) have proposed a Bayesian model to recognize facial expressions amidst uncertainties such as occlusions, pose and illumination variation. The proposed model is capable of representing the relationships among different facial action units in a coherent and unified hierarchical structure, accounting for the uncertainties in the recognition process and providing principled inference. The performance of facial action unit recognition algorithms gets affected by errors encountered during the feature extraction and face alignment process due to uncertainties such as occlusions. The authors claim that their Bayesian model could compensate these errors, by exploiting intrinsic relationships among the facial action units. Kemp et al. (2006) have presented a Bayesian model of inductive reasoning that combines causal reasoning with similarity-based reasoning and shown that it accounts well for human inferences about the properties of biological species. In a way their model exploits intrinsic similarity relationships between objects and their features to make predictions. Zhou et al. (2006) have proposed a Bayesian framework based on a simple human intuition which assumes that all humans have a head and two legs and each leg is jointed at the knee. A 2D articulated model which is a crude approximation to a real walker is fitted to gait silhouettes. The gait images were manually labeled to find out sections of gait cycles. The objective of this approach is to determine the likelihood of the image given the model. The authors claim that their approach tackles well uncertainties such as occlusion and noise. Such Bayesian studies inspired us to design PROSIM, which employs Bayesian Networks to recognize gaits under unconstrained scenarios by learning and interpreting sub-gait characteristics.

3 PRObabilistic Sub-gait Interpretation Model (PROSIM)

3.1 Sub-gait Segmentation

Segmenting specific body components such as head, torso, arms and legs demands manual labeling. However, manual labeling may not guarantee accurate marking of the body components on video sequences. This is because of factors such as low-image quality due to overall intensity, occlusion of feet when walking on grass, similarity of dark skin tones of some subjects with the background, occlusion of the arms due to various viewing angles, and the presence of dark or baggy clothing (Liu and Sarkar 2005). We intend to avoid such manual labeling and at the same time utilize the information from those body components. Hence we strategically segment the silhouettes into sub-gaits viz., Upper Gait (U), Mid Gait (M), Lower Gait (L), LeFt Gait (LF) and Right Gait (R). We will represent the set of sub-gaits by $S = \{U, M, L, LF, R\}$.

By manipulating the binary files that represent silhouettes, we compute the bounding rectangle that encompasses a silhouette and resize them to a standard dimension of 64×44 pixels. A typical silhouette frame of a gait video sequence and its sub-gaits are shown in Fig. 2. We define these sub-gaits using the language of mathematical morphology which is widely used to represent and describe image semantics (Gonzalez and Woods 2002). For a given silhouette frame $I(x, y)$ with width w and height h, its centre (x_c, y_c) can be computed by

$$(x_c, y_c) = \left(\frac{w}{2}, \frac{h}{2}\right). \tag{1}$$

Then the sub-gaits U, M and L can be defined as

$$\begin{aligned} U(I(x, y)) = \Big\{ &(x, y) | x_c - \frac{w}{2} \le x \le x_c + \frac{w}{2}, \\ &y_c + \frac{h}{2} \le y \le y_c + \frac{h}{2} - h\epsilon_1 \Big\}, \end{aligned} \tag{2}$$

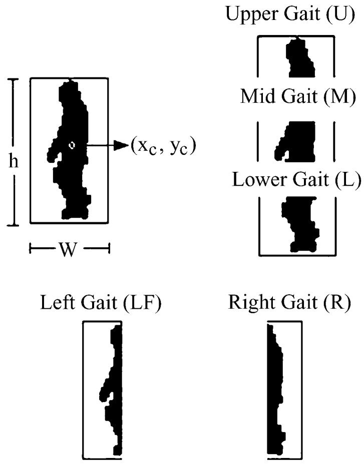

Fig. 2 A typical silhouette and its sub-gaits

$$M(I(x,y)) = \Big\{(x,y)|x_c - \frac{w}{2} \le x \le x_c + \frac{w}{2},\quad y_c + \frac{h}{2} - h\epsilon_1 < y \le y_c + \frac{h}{2} - h(\epsilon_1 + \epsilon_2)\Big\}, \tag{3}$$

$$L(I(x,y)) = \Big\{(x,y)|x_c - \frac{w}{2} \le x \le x_c + \frac{w}{2},\quad y_c + \frac{h}{2} - h(\epsilon_1 + \epsilon_2) < y \le y_c - \frac{h}{2}\Big\}. \tag{4}$$

For the sub-gait definitions above, the heights of each of the sub-gait segments are distinct and determined by constants ϵ_1 and ϵ_2. Values of these constants were chosen based on rough estimates performed on the mean silhouette of the gallery set. The left and right sub-gaits viz., LF and R, which are segmented from the centre are just a function of width (w) and do not require extra constants. Hence their definitions are straight forward as follows:

$$LF(I(x,y)) = \Big\{(x,y)|x_c - \frac{w}{2} \le x \le x_c,\quad y_c + \frac{h}{2} \le y \le y_c - \frac{h}{2}\Big\}, \tag{5}$$

$$R(I(x,y)) = \Big\{(x,y)|x_c < x \le x_c + \frac{w}{2},\quad y_c + \frac{h}{2} \le y \le y_c - \frac{h}{2}\Big\}. \tag{6}$$

In this paper we describe a procedure to recognize gaits by representing and interpreting sub-gait characteristics using a reasonable probabilistic framework. Finding optimal sub-gait dimensions such as the optimal height of the sub-gaits might further improve recognition performance. Such operational motivation factors needs further scrutiny of rigorous iterative experiments, exploration of advanced image segmentation and optimization techniques which deserve another dedicated study.

3.2 Formulation of the Probabilistic Framework

We employ BNs to establish intrinsic similarity mappings between the sub-gaits and the gaits. Each node of a BN has a set of probable values for each variable which are known as belief states. These belief states are propagated between nodes of the BN effectively. A BN maps intrinsic relationships that are inherent in a domain in terms of parent and child nodes. It is capable of learning these relationships and storing the belief states of a given domain in the form of Conditional Probability Tables (CPT). By manipulating these belief states, the state of a particular node can be queried from other nodes with the aid of probabilistic inference techniques. In our case we would like to query the belief state of a gait sequence by observing the probabilities entailed by its sub-gaits.

Though gait motion is periodic in nature, various sub-gaits contain different information about the gait they constitute. Owing to this variation, all the sub-gaits will not have the same probability of influencing the gait to be recognized. Therefore each sub-gait of a gait will have different belief states and this varies from subject to subject as the walking style of individuals varies. The more unique features a sub-gait contains, the more strength it will have to influence the recognition of the gait to which the sub-gait belongs. We define the strength of a sub-gait S_i which crucially contributes to the recognition of the gait G_p as *Influence Strength* and denote it as Z_{ip}. This leads to an important hypothesis. That is the gait pattern being recognized by observing the sub-gait of a probe gait sequence X will be more similar to the corresponding gallery gait pattern, X's gallery set, if the magnitude of the influence strength is high. Hence influence strengths are closely associated to similarities and have an impact over recognition accuracy. By observing a sub-gait, if the suspect is ranked first or within a reasonable range, then the influence strength of that sub-gait tends to be high. Else if the ranking is beyond this range, the influence strength will be low.

Let the sub-gaits of a probe gait sequence be represented by $S = \{U, M, L, LF, R\}$. Let $G = \{G_1, G_2, G_3, \ldots, G_n\}$ represent the gallery gait sequences of n subjects. Let k be the total number of sub-gaits. Suppose that a probe gait's sub-gait, say $S_i \subset S$ for any $1 \le i \le k$, has led to the recognition of a set of gaits $\{G_p, G_q, G_r\} \subset G$, where p, q and r represent unique integers between 1 to n. Let Z_{ip}, Z_{iq} and Z_{ir} represent the corresponding influence strengths. Since S_i is influencing the recognition of $\{G_p, G_q, G_r\}$, we draw edges from S_i to the elements of $\{G_p, G_q, G_r\}$, resulting in a typical Directed Acyclic Graph (DAG) as shown in Fig. 3.

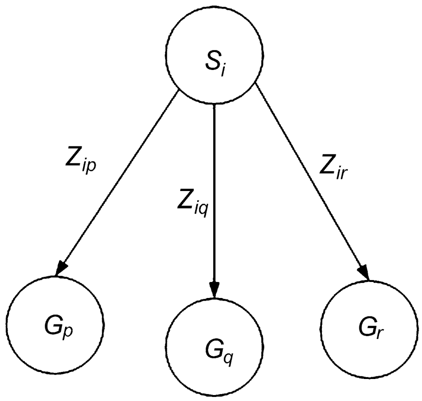

Fig. 3 DAG that establishes similarity mappings between sub-gait to gaits

This way we establish mappings from the set of sub-gaits to the subset of gaits that are activated. Conceptually these gallery gait patterns which are consequences of observing the sub-gait S_i, will be nearly similar to the probe gait pattern. As the mappings established by the DAG are based on similarities we call them similarity mappings.

In the DAG shown in Fig. 3 each gait is conditionally independent of the other gaits given its parent. That is

$$I_P(\{G_p\}, \{G_q, G_r\}|S_i), \qquad I_P(\{G_q\}, \{G_r, G_p\}|S_i)$$

and

$$I_P(\{G_r\}, \{G_p, G_q\}|S_i).$$

We denote independence of variables by I_P. This can be precisely written in the following general form

$$P(G_j|G_c, S_i) = P(G_j|S_i), \quad j = 1, \dots, n,\ i = 1, \dots, k, \tag{7}$$

where $G_c = G \setminus G_j$. Let the DAG shown in Fig. 3 be called D and its underlying probability distribution be called P. Then (D, P) satisfies the Markov condition provided by (7), as each element of D is conditionally independent of the set of all its non-descendents given the set of parents. Such a Markov conditioned DAG leads to what is known as a Bayesian Network (BN) by definition, see (Neapolitan 2003). The graphical nature of PROSIM helps us to visualize the abstract intrinsic similarity mappings (relationships) that exists in the gait domain. This is a consequence of mapping S_i to $\{G_p, G_q, G_r\}$.

3.3 Learning Sub-Gait Characteristics

3.3.1 Learning Parameters

We will call the parent nodes of the proposed BN the prior belief states of the sub-gaits. The gaits influenced by the sub-gaits are the child nodes of the BN. In this probabilistic framework we will infer the belief state of the gaits conditional on the sub-gaits, in order to recognize the gaits. Dirichlet density functions are widely used in Bayesian statistics as they provide intuitive means in representing prior beliefs which can be updated gradually by observing evidence, Neapolitan (2003). The prior belief states of the proposed BN can be quantified using the following Dirichlet density function

$$\rho(f_1, f_2, f_3, \dots, f_{r-1}) = \frac{\Gamma(N)}{\prod_{k=1}^{r} \Gamma(a_k)} f_1^{a_1-1} f_2^{a_2-1} \cdots f_r^{a_r-1}, \tag{8}$$

where $f_1, f_2, \dots, f_{r-1}$ are values taken on by random variables $F_1, F_2, \dots, F_{r-1}$, $0 \le f_k \le 1$. The variable f_r is defined by $\sum_{k=1}^{r} f_k = 1$, $N = \sum_{k=1}^{r} a_k$ and $a_1, a_2, a_3, \dots, a_r$ are integers ≥ 1.
The gamma function used in (9) is computed by

$$\Gamma(x) = (x-1)!, \quad x \in \mathbb{N}. \tag{9}$$

The prior belief states of the parameters which are the fundamental building blocks of the BN are updated by a machine learning procedure called parameter estimation. Out of several such procedures available, Maximum Likelihood Estimation (MLE) and Bayesian Estimation are considered most often by researchers (Duda et al. 2001). MLE has been recommended by Myung (2003) as it has many optimal properties in estimation including asymptotic consistency and unbiased nature. MLE demands large training samples. Fortunately as the BN can be realized through large samples available in the gait domains, MLE will converge to precise estimates enabling the distribution of the parameters to be normal. Consequently many of the inference methods in statistics such as Chi-square test, Akaike information criterion (Akaike 1974) and Bayesian information criteria (Schwarz 1978) are developed based on MLE. Equation (7) reveals that the Markov condition has been satisfied by the probability distribution entailed by the DAG of the proposed PROSIM. Hence we have

$$P(G|S_i) = \prod_{j=1}^{n} P(G_j|S_i), \quad i = 1, \dots, k. \tag{10}$$

Recall from Sect. 3.2 that G represents the n gait silhouette sequences in the gallery set. We mathematically define the influence strength Z_{ij} of a sub-gait S_i as

$$Z_{ij} = (n - \ell)/n, \tag{11}$$

where ℓ is the rank in which the gait G_j is being recognized by the sub-gait S_i. The objective of MLE is to estimate the unknown parameters Z_{ij} $(j = 1, \dots, n; i = 1, \dots, k)$ that best agree with the gallery samples. MLE of Z_{ij} is by definition the value $\hat{Z}_{ij}$ that maximizes $\ln P(G|S_i)$, the log-likelihood of the parameter set Z_{ij} with respect to the gallery set G. $\hat{Z}_{ij}$ can be computed by

$$\hat{Z}_{ij} = \arg\max_{Z_{ij}} \ln P(G_j|S_i). \tag{12}$$

To be a maximum, the shape of the log-likelihood function, $\ln P(G_j|S_i)$, should be convex in the neighborhood of $\hat{Z}_{ij}$ which can be checked by computing the second derivatives of the log likelihoods. Finally a set of necessary conditions for the maximum likelihood estimate for Z_{ij} can be obtained from the set of n equations

$$\sum_{i=1}^{k} \nabla_{Z_i} \ln P(G_j|S_i) = 0, \quad j = 1, \dots, n, \tag{13}$$

where the gradient operator ∇_{Z_i} is given by

$$\nabla_{Z_i} \equiv \begin{pmatrix} \frac{\partial}{\partial Z_{i1}} \\ \frac{\partial}{\partial Z_{i2}} \\ \vdots \\ \frac{\partial}{\partial Z_{in}} \end{pmatrix}, \quad i = 1, \dots, k. \tag{14}$$

3.3.2 Proposed Modular Training Scheme

We will describe a novel modular training scheme employed by PROSIM here. A training or test sample is well defined in many object recognition (e.g. face, iris recognition) problems. For example, a face or an iris image is considered as a sample without any further partitions. However, the definition of a gait sample is subjective and not so precisely defined. Usually a gait sample is represented in terms of gait cycles (either full, multiple or partial cycles). A gait cycle begins when one foot contacts the ground and ends when the same foot contacts the ground again. Thus, each cycle begins at initial contact with a stance phase and proceeds through a swing phase until the cycle ends with the next initial contact of the limb. Prior to factoring the gait samples into modules, we have constructed the sub-gaits data sets from the gallery data sets by applying the sub-gait segmentation scheme formulated in Sect. 3.1. In the gallery set, because each subject's behavior is represented as several gait samples due to variations in walking speed, the number of frames per sample will be different. A suitable time mode normalization algorithm can be applied to normalize the gait samples to have a unique number of frames. We have normalized the number of frames in each sample by applying the time mode normalization technique proposed by Lu et al. (2008). We intend to decompose the normalized sub-gait samples into compact modules and train PROSIM to learn the intrinsic relationships between these modularized sub-gaits. The proposed modular training scheme enables PROSIM to represent and learn subtle walking patterns of human gaits. Figure 4 shows the modular scheme applied to a typical sub-gait. For example's sake we have shown the scheme for a lower sub-gait. We initially modularize all training samples into two subsets namely A and B. An even number of samples is split 50–50, an odd number the closest integer partition to 50–50. Gait subsamples of modules A and B represent how the subjects walk during the first part and second part of a walking segment. We further modularize these subsets into AB and BA which will have mixtures of walking samples from A and B together. That is AB will have some samples from the first half of A and B and BA will have some samples from the second half of A and B. Finally we modularize AB and BA into tiny modules viz., AB_1, AB_2, BA_1 and BA_2. That is each of these tiny modules represent about a quarter of a sub-gait sample. Mathematically we can model this modular scheme as follows:

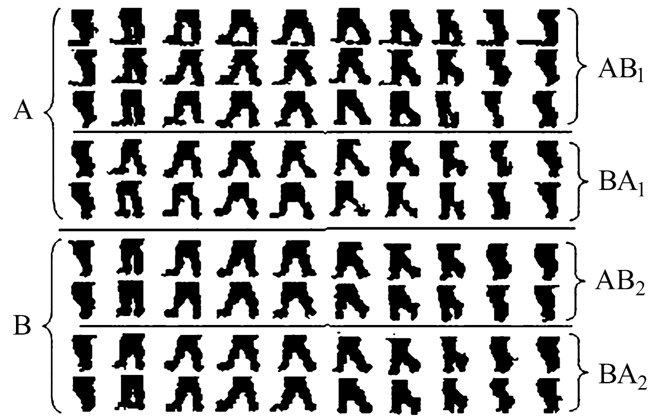

Fig. 4 Modular scheme of a typical sub-gait

Let the gallery set of sub-gait (or gait) silhouettes of a subject say U be represented by d gait samples. Let the ith sub-gait sample be denoted by u_i, where $1 \le i \le d$. We wish to modularize U such that

$$U = AB_1 \cup BA_1 \cup AB_2 \cup BA_2, \tag{15}$$

where

$$AB_1 = \sum_{i=1}^{a} u_i; \qquad BA_1 = \sum_{i=a+1}^{b} u_i; \tag{16}$$

$$AB_2 = \sum_{i=b+1}^{c} u_i; \qquad BA_2 = \sum_{i=c+1}^{d} u_i. \tag{17}$$

The indices a, b and c of (16) and (17) can be computed as

$$a = \lceil d/4 \rceil; \quad b = a + \frac{d-a}{3}; \quad c = b + \frac{d-a}{3}. \tag{18}$$

Table 1 CPT showing belief states of subtle sub-gait relationships learned from the proposed modular training scheme for some typical subjects. The sub-gait operators $L(\cdot)$ and $LF(\cdot)$ have been defined in (4) and (5)

Sub-gaits	Learned belief states of typical subjects				
	G_5	G_{10}	G_{15}	G_{20}	G_{25}
$L(A)$	0.97	0.86	0.75	0.93	0.86
$L(B)$	0.79	0.84	0.73	0.79	0.84
$L(A)\ L(B)$	0.78	0.99	0.92	0.99	0.99
$LF(A)$	0.94	0.86	0.89	0.95	0.86
$L(A)\ LF(A)$	0.97	0.86	0.68	0.79	0.86
$L(B)\ LF(A)$	0.78	0.97	0.67	0.99	0.97
$L(A)\ L(B)\ LF(A)$	0.93	0.58	0.69	0.91	0.58
$LF(B)$	0.92	0.87	0.86	0.97	0.87
$L(A)\ LF(B)$	0.97	0.78	0.83	0.99	0.78
$L(B)\ LF(B)$	0.97	0.66	0.85	0.97	0.66
$L(A)\ L(B)\ LF(B)$	0.56	0.58	0.33	0.77	0.58
$LF(A)\ LF(B)$	0.75	0.99	0.92	0.99	0.99
$L(A)\ LF(A)\ LF(B)$	0.93	0.58	0.69	0.91	0.58
$L(B)\ LF(A)\ LF(B)$	0.58	0.73	0.39	0.72	0.73
$L(A)\ L(B)\ LF(A)\ LF(B)$	0.97	0.78	0.83	0.99	0.78

Obviously the tiny modules defined in (16) and (17) can be appropriately merged to yield

$$AB = AB_1 \cup AB_2; \qquad BA = BA_1 \cup BA_2. \tag{19}$$

Modules can be combined using the following rules

$$AB \cap A = AB_1; \qquad BA \cap A = BA_1; \tag{20}$$

$$AB \cap B = AB_2; \qquad BA \cap B = BA_2. \tag{21}$$

We will show shortly how the proposed modular scheme enables us to relate the various sub-gaits and learn subtle walking patterns that are inherent in a subject's walking behavior. We perceive that the intrinsic relationships that exist between the modularized sub-gaits contribute significantly in governing the gait patterns. The BN employed in PROSIM learns the belief states of these relationships systematically from the sub-gait data sets using the MLE approach outlined in Sect. 3.3.1. The learned belief states are stored in the form of Conditional Probability Tables (CPTs). For k sub-gaits and m modules, the BN yields a CPT comprising of $2^{k*m} - 1$ number of rows. A typical CPT for the case of two sub-gaits L and LF whose samples are factored into two subsamples A and B is shown in Table 1. By combining various sub-gait modules we can reveal intrinsic characteristics of gait patterns. For example the CPT entry $L(A)LF(B)$ intends to reveal the belief state of "*left leg sub-gait pattern*" for a portion of a walking sequence. The sub-gait operators $L(\cdot)$ and $LF(\cdot)$ have been defined in (4) and (5). Trivially $L(A) \cap LF(B) = L(LF(AB))$. When more combinations of sub-gaits and subsamples are involved, the interpretation needs a few more steps. For example a typical CPT entry and its interpretation are as follows:

$$\begin{aligned} &L(A) \quad L(B) \quad LF(A) \\ &\quad = L(A) \cap L(B) \cap LF(A) \\ &\quad = L(AB) \cap LF(A) \\ &\quad = L(LF(AB_1)) \because \text{Eq.}(20) \end{aligned}$$

Similar logical reasoning can be extended to interpret any other entry in the CPT. The conditional probabilities in Table 1 give a measure of the strength of sub-gait relationships. For example, referring to the first column in the table, we observe that the conditional probabilities $P(G_5|L(A))$, $P(G_5|L(A), LF(A))$ and $P(G_5|L(A), LF(B))$ are higher. This reveals the fact that the gait motion of the subject G_5 is highly characterized by these intrinsic sub-gait relationships. Whereas $P(G_{15}|L(A), L(B), LF(B))$ and $P(G_{15}|L(B), LF(A), LF(B))$ (middle column of the table), being low indicate that G_{15} is poorly characterized by these sub-gait modules. We will shortly see how robust probabilistic decisions can be made by interpreting and exploiting these subtle relationships.

3.3.3 Gait Score Formulation

In this section we will propose a formula that will aid to decide on the most probable gaits. The Bayesian Network (BN) generated by PROSIM for a typical probe gait, whose gallery representation is $G20$, is shown in Fig. 5. We have made use of the resources provided by Murphy (2007) to build the BN. The parent nodes viz., *LA*, *LB*, *LFA*, *LFB*, *RA* and *RB* represent the sub-gaits. An edge from a sub-gait to a gait indicates that, the sub-gait has influenced the gait. Also the gaits (e.g. $G22$, $G40, \ldots$) influenced by the sub-gaits are represented by the child nodes. By this way PROSIM establishes similarity mappings from the set of sub-gait modules to the subset of gaits. Recall from Sect. 3.2 that the mappings established by the BN are based on similarities. Naively the more sub-gaits mapping to a gait, roughly infers that the gait has more chances of being similar to the probe. For example, in Fig. 5, the gait $G20$ has been mapped by maximum number of sub-gaits and as such could have more chances of being recognized in the first rank. But however as each sub-gait has a different recognition potential, this intuition is not adequate enough to make a robust decision. That is, it is not always necessary that a winner gait will always have more mappings. We will propose a gait score formula to make a robust decision.
Let C number of gaits be shortlisted by PROSIM from the huge gallery set, as a consequence of the similarity mappings. The formula which we intend to formulate will yield a score for each of the C gaits. This score will aid to finally rank list the shortlisted gaits. By exploiting the graph-

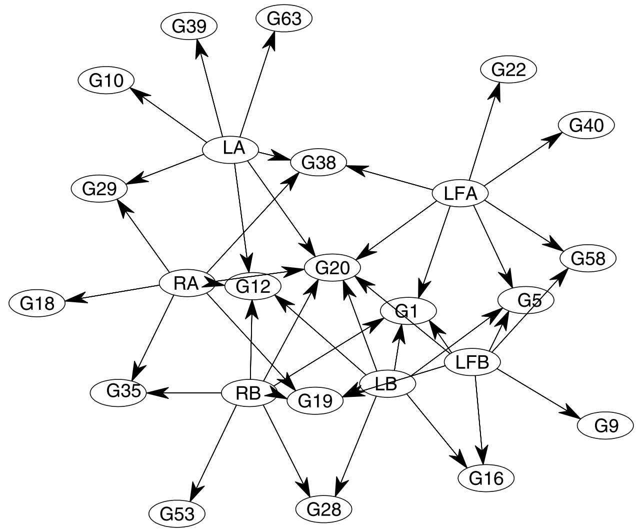

Fig. 5 A typical Bayesian Network showing similarity mappings between the sub-gait modules and the gaits influenced by them. $G20$ is the gallery representation of the probe (typically chosen from the USF dataset) being observed

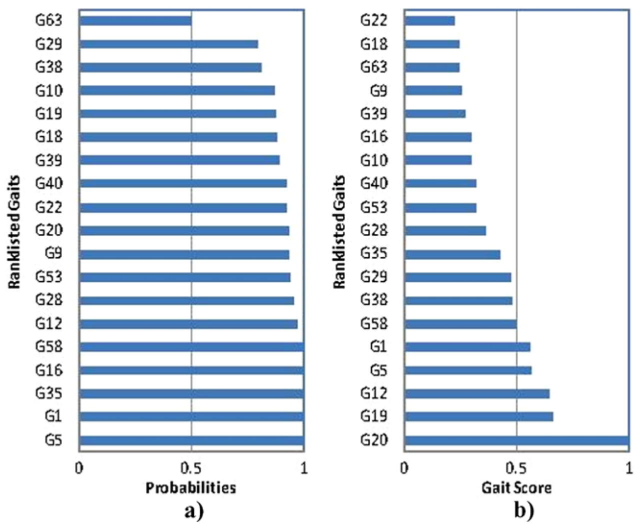

Fig. 6 Predicting winner gaits using a robust gait score function

ical structure of the BN, the probability distribution over $G_m, 0 < m \leq C$ can be computed by

$$P(G_m) = \sum_{S_i} P(S_i)P(G_m|S_i), \tag{22}$$

where $P(S_i)$ is the prior probability of sub-gaits and $P(G_m|S_i)$ is the probability of a gait given the condition that some sub-gaits (or a sub-gait) has influenced it. The probability distribution of the short listed gaits are shown in the bar chart of Fig. 6a). We see that the probability of the gait $G20$ falls some where in the middle. Also the probabilities of gaits $G5, G1, G35, \ldots, G53, G9$ which are larger than $G20$, fetch nearly similar values without much discrimination between them. This reveals that mere probabilities are not adequate enough to make a meaningful decision. To counter this problem we consider the well known rule of thumb given by Russel and Norvig (1995) which emphasizes that "Probability theory and utility theory together constitute decision theory". By utilizing the crucial influence strengths Z defined in Sect. 3.2 to weigh the prior probability of sub-gaits, logically the gait score could yield meaningful results if it is a function of the following two factors:

(i) The probability distribution of the gaits and their sub-gaits.
(ii) Prior probabilities of sub-gaits duly weighed by their corresponding influence strengths.

Consolidating the above factors, the gait score μ of an mth gait can be computed using

$$\mu(G_m) = \sum_{S_i} P(S_i)P(G_m|S_i) + \sum_{S_i} Z_{im} P(S_i). \tag{23}$$

The probe gait sequence (for the typical case), had been subjected to two covariate factors namely surface and view. The variations caused by surface and view are considered as one of the hard problems in gait recognition which impose vast uncertainty to the recognition process. With the aid of the gait score, (23), gaits have been rank listed as shown in Fig. 6(b). The chart shows that the gait score discriminates the winner gaits well; $G20$, which is the gallery representation of the observed probe, clearly stands out compared to Fig. 6(a). Thus the above formulation of face score aids PROSIM to make robust decisions under uncertainties.

3.3.4 *Robustness to Common Variations*

Some common uncertainties encountered in the process of gait recognition are caused due to variations present in challenging outdoor environments such as view, surface, shoe,

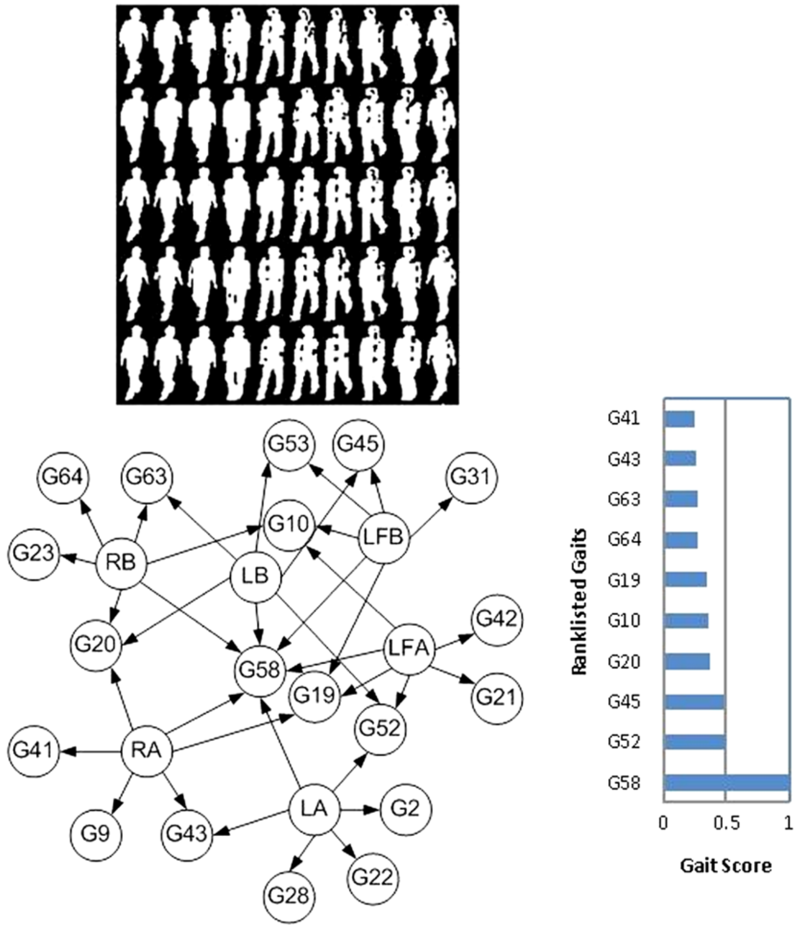

Fig. 7 A scenario that depicts the recognition process of a probe gait (typically chosen from CASIA dataset) with a typical viewing variation of 18°. Normalized silhouettes of gait sequences of the probe, the associated Bayesian Network generated by PROSIM and the bar chart of first ten winner gaits being recognized are shown

missing body components and so on. The experimental results of PROSIM's robustness against these uncertainties will be presented in Sect. 4. Here we will analyse the effect of uncertainties caused by two typical variations viz. view and missing body components. The scenario of a typical probe gait whose gallery representation is $G58$ has been subjected to viewing variations of 18° and 162° as shown in Figs. 7 and 8 respectively. The Bayesian Network (BN) generated by PROSIM (Figs. 7 and Fig. 8) helps us to analyse how this uncertainty affect the recognition mechanism, in particular the relationships between the gaits and sub-gaits.

As the viewing variation of 18° is relatively small and probably other variations being less severe, all the sub-gaits of $G58$ collectively contribute to the recognition process as seen in Fig. 7. Further the silhouettes are noisy due to factors such as similarity of colors of the subject and the background, varying illumination caused by the operating environment and so on. Despite these variations, $G58$ has been successfully recognized as a winner gait as shown in the bar chart.

Body components such as head, arms and some portion of the torso are missing in most of the normalized silhouette sequences shown in Fig. 8. A huge viewing variation of 162° along with the complexity of missing body components, obviously causes more uncertainty and consequently $G58$ has been degraded from rank 1 to rank 2 as shown in the bar chart. Interestingly when the gait of a subject is viewed from 162°, the left body motion is more visible than the right body motion. This is intuitively reflected by the sub-gait to gait relationships captured by the BN shown in Fig. 8. Specifically the right sub-gaits, *RA* and *RB*, have not contributed to the recognition of $G58$. However these sub-gaits played their role when the viewing angle was 18° as seen in Fig. 7. The proposed framework enables us to visualize such interesting relationships that exists between gaits and sub-gaits. We see that sub-gaits *RA* and *RB* lack to provide evidence due to uncertainties in the scenario. However PROSIM grasps information by accumulating evidences from other sub-gaits. By manipulating the available evidences (*LA*, *LB*, *LFA* and *LFB*) and the learned belief states from the stored CPTs, PROSIM is still able to recognize $G58$ reasonably well (in second rank).

The gait samples of a subject is represented in terms of normalized gait cycles which is comprised of a set of silhouettes. Some of the samples might have silhouettes with missing parts (weak samples). Within a sample the uncertainty caused by silhouettes with missing parts will be com-

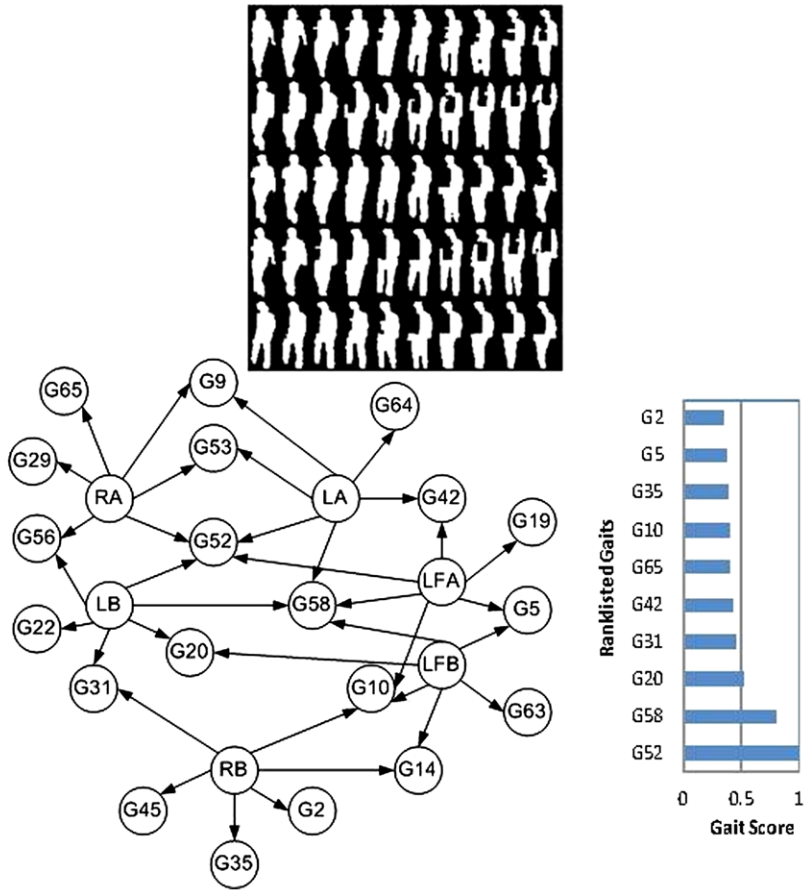

Fig. 8 A scenario that depicts the recognition process of a probe gait (typically chosen from CASIA dataset) with a typical viewing variation of 162°. Normalized silhouettes of gait sequences of the probe with missing body components, the associated Bayesian Network generated by PROSIM and the bar chart of first ten winner gaits being recognized are shown

Table 2 Experimental notation and description with compliance to human identification in USF HumanID data sets

Probe set	A	B	C	D	E	F	G
Capturing condition	GAL	GBR	GBL	CAR	CBR	CAL	CBL
Covariate factors	View	Shoe	Shoe, view	Surface	Surface, shoe	Surface, view	Surface, shoe, view

pensated by the ones which are complete. Furthermore, as we decompose the samples into compact modules (please see Sect. 3.3.2), the modules which have more good samples would compensate the uncertainty caused by modules that contain silhouettes with missing parts. For example a module of the left sub-gait (*LFA*) might fail to provide evidence or provide less evidence (influence strength could be weak due to weak samples). However the other module of the left sub-gait (*LFB*) or modules of other sub-gaits might provide sufficient evidence to mitigate the uncertainties imposed by the weak module.

4 Experimental Results

4.1 Data Set and Experimental Design

We have used the University of South Florida (USF) HumanID gait challenge data set (Sarkar et al. 2005) and the multi-view gait dataset offered by Chinese Academy of Sciences (CASIA 2006) to evaluate PROSIM and compared it with the state-of-the-art gait recognition algorithms. The USF data set which was collected on typical outdoor environment, consists of 122 subjects comprising of 1870 video sequences. The gait challenge baseline algorithm (Sarkar et al. 2005) as well as very recent algorithms such as Veeraraghavan et al. (2009) consider seven standard experimental probe sets, the details of which are tabulated in Table 2. The seven probe sets, A to G, are designed to perform a range of experiments in the order of increasing difficulties. The abbreviations of various capturing conditions in the table viz., C, G, A, B, L, and R refers to Concrete surface, Grass surface, shoe type A, shoe type B, Left view and Right view respectively.

4.2 Identifying Potential Sub-gaits

Recall from Sect. 3.1 that we have defined five sub-gaits ($k = 5$) viz., Upper Gait (U), Mid Gait (M), Lower Gait

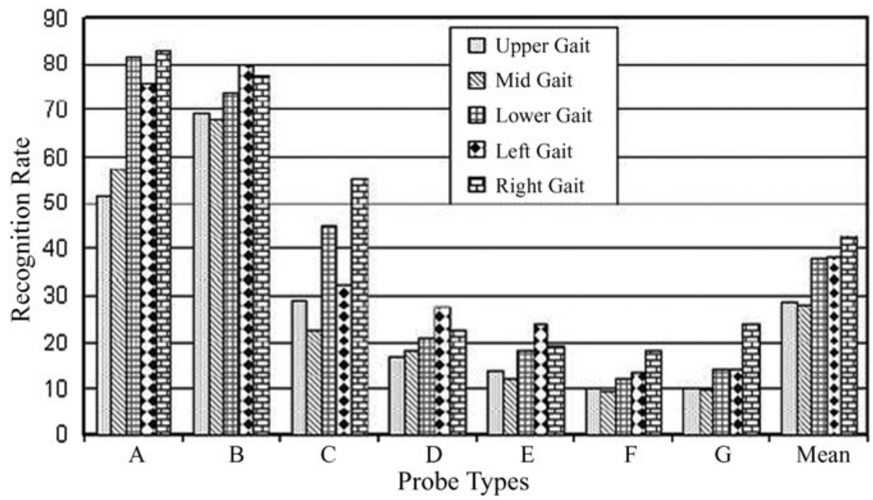

Fig. 9 Recognition potential of sub-gaits for the HumanID gait challenge data

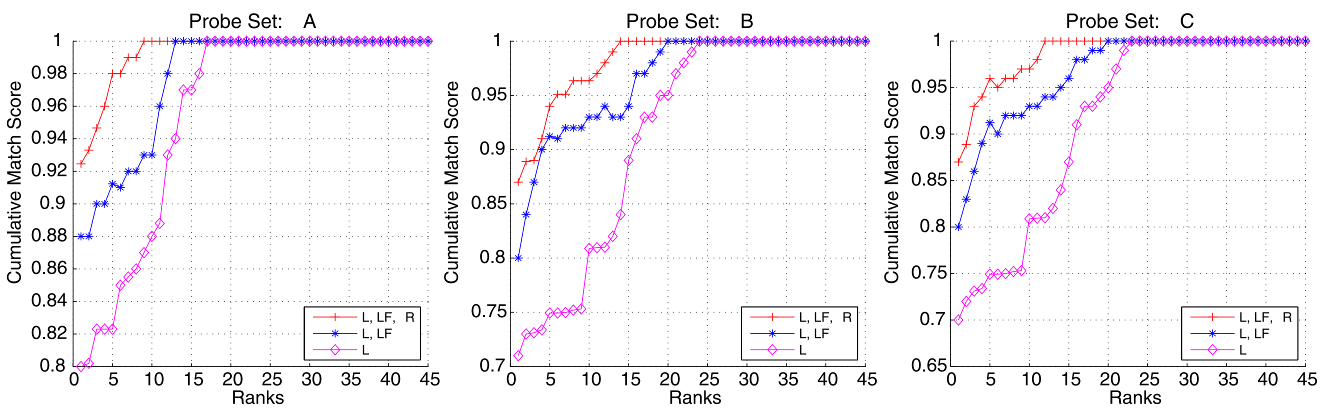

Fig. 10 CMC response of PROSIM with respect to HumanID gait challenge experiments A, B and C

(L), Left Gait (LF) and Right Gait (R). Also recall from Sect. 3.3.2 that the size of the Bayesian Network tends to grow exponentially as the number of sub-gaits (i.e. the parameters) increase. This will in turn demand more computing resources. Hence prior to parameter estimation, strategically selecting just a few potential sub-gaits would enable PROSIM to be computationally feasible. In this section, we will identify such potential sub-gaits based on their recognition power. We have computed recognition rates for all of the sub-gaits for the seven core experiments using the approach proposed by Lu et al. (2008), the results of which are shown in Fig. 9. The mean performance of all the experiments shown at the right most end of Fig. 9 justifies that the sub-gaits L, LF and R have higher recognition potential than U and M. Hence we will only employ these potential sub-gaits in the subsequent experiments.

4.3 Performance Evaluation

We have experimented PROSIM with the HumanID gait challenge experiments by gradually increasing the number of sub-gaits. A Cumulative Match Characteristic (CMC) curve (Moon and Phillips 2001) shows various probabilities of recognizing an individual depending on how similar their measurements are to that of others in the gallery. The rank 1 point on the CMC curve is the nearest-neighbor recognition performance. The CMC graphs of these experiments are shown in Figs. 10, 11 and 12. Initially by considering the lower sub-gait alone (i.e. L), mean recognition rates of about 60% and 67% have been yielded by PROSIM respectively for the rank 1 and rank 5 performance. Then by combining two potential sub-gaits (i.e. $L + LF$), this improved to about 67% and 82%. Finally by considering all the three potential sub-gaits (i.e. $L + LF + R$), the mean recognition rates have been considerably improved to about 75% and 90% for rank 1 and rank 5 performance respectively. These experimental results clearly show that when all the potential sub-gaits are used, PROSIM achieves maximum recognition performance.

Further we subject PROSIM to the HumanID gait challenge experiments using USF dataset and compared it against the following state-of-the-art gait recognition algorithms:

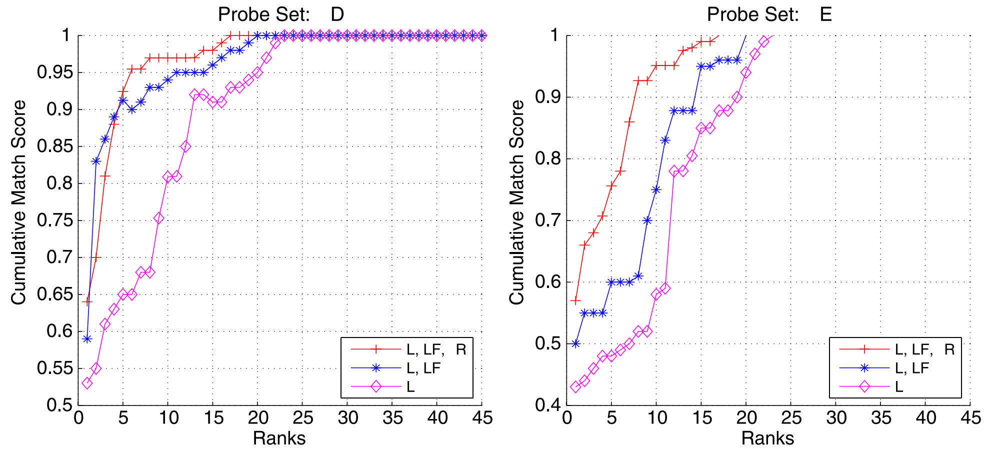

Fig. 11 CMC response of PROSIM with respect to HumanID gait challenge experiments D and E

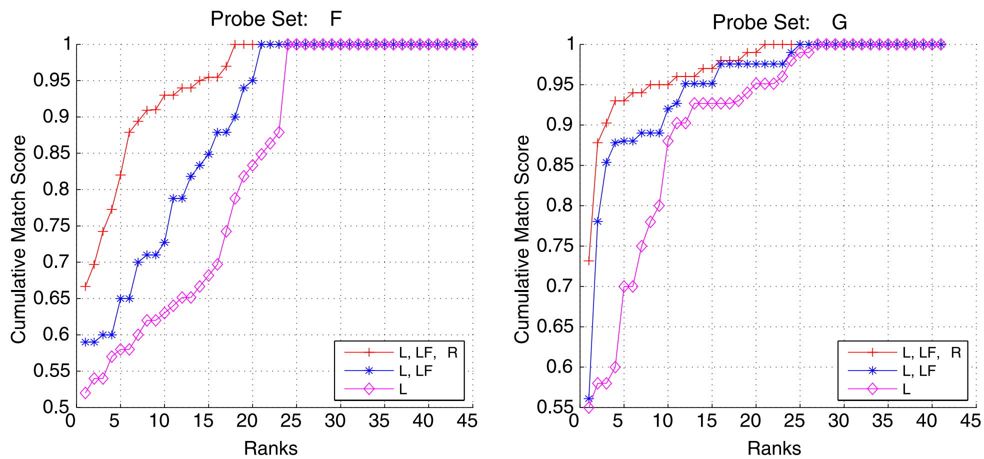

Fig. 12 CMC response of PROSIM with respect to HumanID gait challenge experiments F and G

i. Baseline (Sarkar et al. 2005)
ii. HMM—Hidden Markov Model (Kale et al. 2004)
iii. DATER—Discriminant Analysis with Tensor Representation (Yan et al. 2005)
iv. DTW/HMM—Dynamic Time Warpring/HMM (Veeraraghavan et al. 2004)
v. ETGLDA—Eigen Tensor Gaits based on Linear Discriminant Analysis (Lu et al. 2008)
vi. GEI—Gait Energy Image (Han and Bhanu 2006)
vii. LTN—Linear Time Normalization (Boulgouris et al. 2006)
viii. MR—Matrix Representation (Xu et al. 2006)
ix. NTWN—Nonlinear Time-Warp Normalization (Veeraraghavan et al. 2009)
x. pHMM—Population Hidden Markov Model (Liu and Sarkar 2006).

We have experimented PROSIM with two modes of recognition experiments. Initially we used the conventional experimental setting proposed by Sarkar et al. (2005) where training was done with a limited gallery set (capturing condition was fixed as Grass, Shoe Type A and Right Camera). Recognition tests were performed with various probe sets described in Table 2. We refer to this conventional recognition experiment as PROSIM-a. Very recently Veeraraghavan et al. (2009) have shown that improved recognition rates can be achieved by using multiple samples for training. They proposed a round-robin recognition experiment in which one of the challenge sets was used as test while the other seven were used as training examples. The process was repeated for each of the seven challenge sets. We refer this experiment as PROSIM-b. The rank 1 and rank 5 performance comparisons with state-art-of-the-art gait recognition algorithms are shown as bar charts in Figs. 13 and 14 respectively. Though PROSIM-a competes fairly with other algorithms, it is not as significant as PROSIM-b due to the restricted mode training. We see that PROSIM-b outperforms other algorithms in majority of the tests. Recognition rates of 75.3% and 89.6% achieved by PROSIM-b respectively for rank 1 and rank 5 performance, on an average of all the seven gait-challenge experiments, justifies the robustness of the proposed approach.

4.3.1 *Experiments with the CASIA Dataset B*

In this section we will investigate the generalization capability of PROSIM with the large multi-view CASIA dataset B

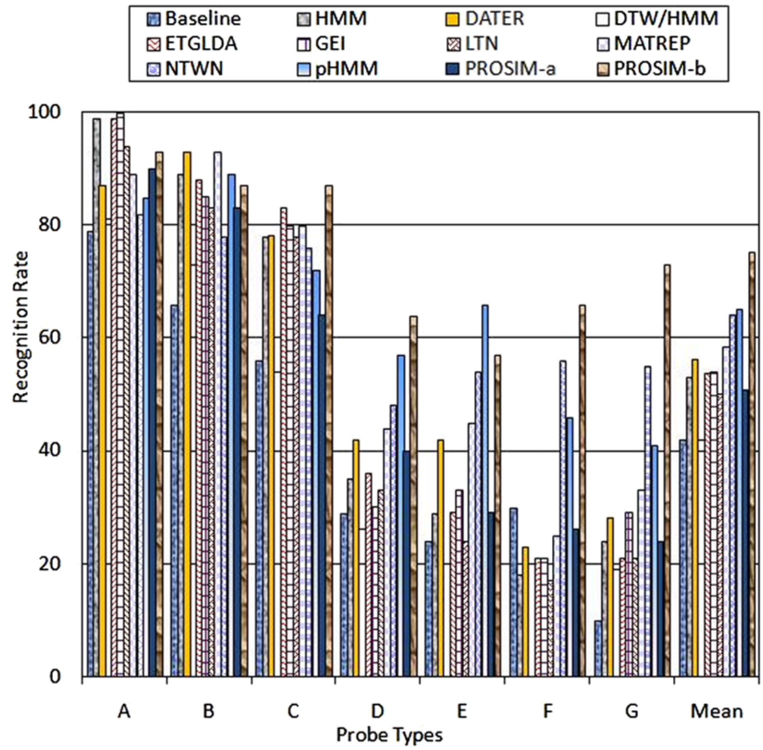

Fig. 13 PROSIM Vs. state-of-the-art gait recognizers: Rank 1 Performance

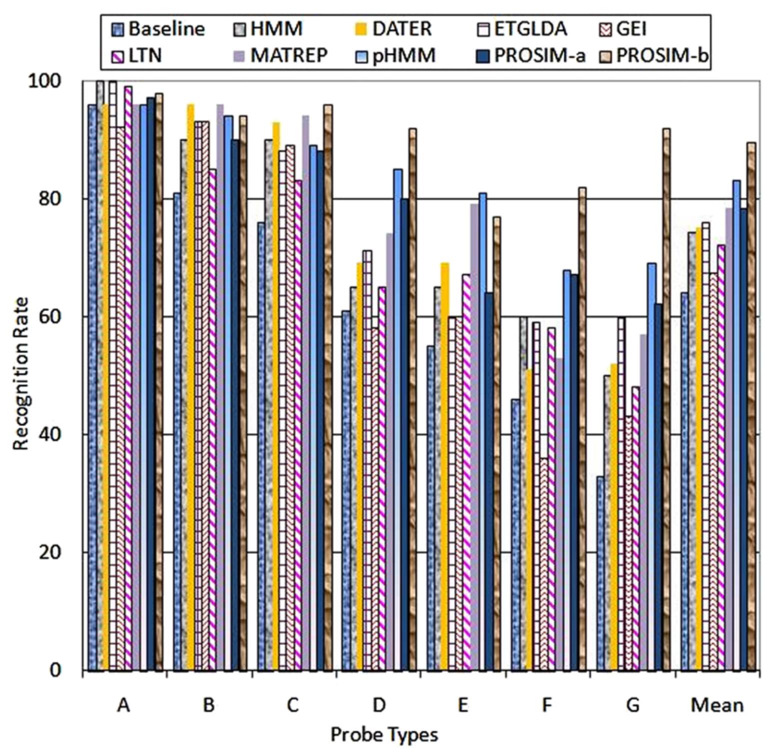

Fig. 14 PROSIM Vs. state-of-the-art gait recognizers: Rank 5 Performance

which contains gait sequences of 124 subjects captured from 11 viewing angles. There were totally 10 gait sequences for each subject (6 normal + 2 with a coat + 2 with a bag) for each of the 11 views. The dataset (CASIA 2006) enables us to experiment the effect of the following co-variate factors.

(i) View (Camera angles were varied from 0° to 180° in increments of 18°)
(ii) View and clothing ((i) + Subjects walked by covering them with a long coat)

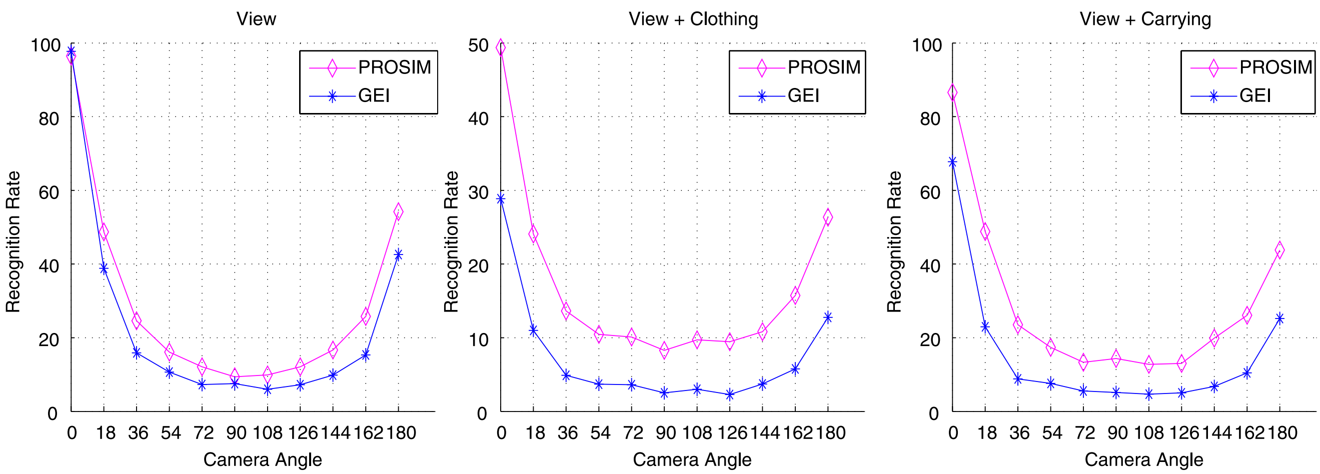

Fig. 15 PROSIM compared with GEI using CASIA dataset B

(iii) View and carrying ((i) + Subjects walked by carrying a bag).

The first four sequences (normal) were used for training and the remaining were used for testing. Yu et al. (2006) has implemented the GEI algorithm using the CASIA dataset. We have compared our results with the GEI algorithm which are shown in Fig. 15. When tested by varying the carrying condition alone (i.e. for the same view), PROSIM and GEI yielded a recognition rate of about 87% and 68% respectively. When tested by varying the clothing condition alone, PROSIM and GEI yielded a recognition rate of about 50% and 29% respectively. This indicates that clothing is a tough test as the occlusion caused by long coat (most of the body parts are occluded by a long coat) imposes vast uncertainty to the recognition process. For a small viewing variation of 18°, PROSIM and GEI yields a recognition rate of about 49% and 39% respectively. However when viewing is varied extremely (trained with 0° and tested with 90°) coupled with clothing variation the recognition rate has been degraded to 8.3% and 2.5% respectively by PROSIM and GEI. We see that PROSIM shows improved results especially over cloting and carrying conditions.

5 Conclusion

We have identified potential sub-gaits and discovered interesting sub-gait characteristics within the gait domain. The novel Probabilistic Sub-gait Interpretation Model (PROSIM) introduced in this work does not require manual labeling of body components. Further the proposed modular training scheme enables PROSIM to learn subtle gait patterns. The graphical nature of PROSIM aids to intuitively visualize intrinsic sub-gait relationships and demonstrates how these sub-gaits collectively contribute to the recognition process. With the aid of few potential sub-gaits PROSIM reports a reliable recognition performance and competes well with the state-of-the-art gait recognizers. PROSIM is a generic model which can be fitted to suit any subspace technique. Our results show that extreme viewing angle variations coupled with change of clothing remains to be the toughest test among the experiments we have performed.

An interesting avenue for future directions could be "The proposed model does not have direct dependencies among parts and does this detract from the power of the modeling?" We have applied Bayesian Networks (which use directed edges) in the proposed framework to exploit the conditional independence properties that exists between gaits and their sub-gaits to achieve robust gait recognition. Such independence assumptions reduce the number of parameters in the model, and therefore making the model computationally feasible for real time applications. However setting dependencies among parts could be modeled using undirected links. Graphical models such as Markov networks (Pearl 1997) which use undirected graphs can be employed to capture dependency among various sub-gaits. In this regard, it will be an interesting avenue in the future to apply undirected graphical models, to investigate the impact of dependencies between sub-gaits and ultimately how they would influence the gait recognition process. Further we intend to apply the proposed approach to a wide range of object recognition problems in the future.

Acknowledgements The authors would like to thank Prof. Sarkar and his team, University of South Florida for providing the gait challenge data sets. Portions of the research in this paper use the CASIA Gait Database collected by Institute of Automation, Chinese Academy of Sciences. We also thank Dr. Haiping, University of Toronto for providing gait resources and valuable advice. We thank Dr. Murphy and his team, University of British Columbia, for providing the Bayes Net Toolbox. This research is supported by James Watt Scholarship awarded by the School of Mathematical and Computer Sciences, Heriot-Watt University. We also thank Kolej University TATI and Universiti Sains Malaysia for their support. We thank the editor and anonymous reviewers for providing useful suggestions.

References

Abeni, P., Baltatu, M., & Alessandro, R. D. (2006). User authentication based on face recognition with support vector machines. In *Proc. of the Canadian conference on computer and robot vision*. Los Alamitos: IEEE Computer Society.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723.

Bauckhage, C., Tsotsos, J., & Bunn, F. (2006). Automatic detection of abnormal gait. *Image and Vision Computing*, *27*, 108–115.

Bergtholdt, M., Kappes, J., Schmidt, S., & Schnörr, C. (2009). A study of parts-based object class detection using complete graphs. *International Journal of Computer Vision*.

Boulgouris, N., & Chi, Z. (2007). Gait recognition based on human body components. In *IEEE international conference on image processing (ICIP)* (Vol. 1, pp. 353–356).

Boulgouris, N., Plataniotis, K., & Hatzinakos, D. (2006). Gait recognition using linear time normalization. *Pattern Recognition*, *39*, 969–979.

CASIA (2006). Gait database offered by Chinese Academy of Sciences. http://www.sinobiometrics.com.

Collins, R., Gross, R., & Shi, J. (2002) Silhouette-based human identification from body shape and gait. In *Proceedings of IEEE international conference on automatic face and gesture recognition* (pp. 366–371).

Dahyot, R., Charbonnier, P., & Heitz, F. (2004). A Bayesian approach to object detection using probabilistic appearance-based models. *Pattern Analysis Applications*, *7*(3), 317–332. DOI:10.1007/s10044-004-0230-5.

Duda, R., Hart, P., & Stork, D. (2001). *Pattern classification*. New York: Wiley.

Gonzalez, R., & Woods, R. (2002). *Digital image processing*. New York: Prentice Hall.

Hallinan, P., Gordon, G., Yuille, A., Giblin, P., & Mumford, D. (1999). *Two and three dimensional patterns of the face*. A.K. Peters.

Han, J., & Bhanu, B. (2006). Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(2), 316–322.

Heisele, B., Serre, T., & Poggio, T. (2007). A component-based framework for face detection and identification. *International Journal of Computer Vision*, *74*, 167–181.

Intille, S., & Bobick, A. (2001). Recognizing planned, multiperson action. *Computer Vision and Image Understanding*, *81*, 441–445.

Jensen, F., & Nielsen, T. (2007). *Bayesian networks and decision graphs*. Berlin: Springer.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 201–211.

Kale, A., Sundaresan, A., Rajagopalan, A., Cuntoor, N., Chowdhury, A., Volker, K., & Chellappa, R. (2004). Identification of humans using gait. *IEEE Transactions on Image Processing*, *13*, 1163–1173.

Kemp, C., Shafto, P., Berke, A., & Tenenbaum, J. (2006). Combining causal and similarity-based reasoning. In *Proceedings of the twentieth annual conference on neural information processing systems*. The Neural Information Processing Systems (NIPS) Foundation.

Kim, J., Choi, J., Yi, J., & Turk, M. (2005). Effective representation using ICA for face recognition robust to local distortion and partial occlusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(12), 1977–1981.

Krynski, T., & Tenenbaum, J. (2007). The role of causality in judgment under uncertainty. *Journal of Experimental Psychology*, *136*(3), 430–450. DOI:10.1037/0096-3445.136.3.430.

Lee, L., & Grimson, W. (2002). Gait analysis for recognition and classification. In *Proceedings of IEEE international conference on automatic face and gesture recognition* (pp. 155–162).

Li, X., Maybank, S., Yan, S., Tao, D., & Xu, D. (2008). Gait components and their application to gender recognition. *IEEE Transactions on Systems, Man and Cybernetics—Part C: Applications and Reviews*, *38*(2), 145–155.

Liu, Z., & Sarkar, S. (2006). Improved gait recognition by gait dynamics normalization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*, 863–876.

Liu, Z., & Sarkar, S. (2005). Effect of silhouette quality on hard problems in gait recognition. *IEEE Transactions on Systems, Man and Cybernetics—Part B: Cybernetics*, *35*, 170–183.

Lu, H., Plataniotis, K., & Venetsanopoulos, A. N. (2008). MPCA: Multilinear Principal Component Analysis of tensor objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(1), 18–39.

Martinez, A. (2002). Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(6), 748–763.

Mohan, A., Papageorgiou, C., & Poggio, T. (2001). Example-based object detection in images by components. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(4), 349–361.

Moon, H., & Phillips, P. J. (2001). Computational and performance aspects of PCA-based face-recognition algorithms. *Perception*, *30*, 303–321.

Murphy, K. (2007). *Software for graphical models: a review*. Tech. rep., International Society for Bayesian Analysis (ISBA) Bulletin.

Murray, M. (1967). Gait as a total pattern of movement. *American Journal of Physical Medicine*, *46*, 290–332.

Murray, M., Drought, A., & Kory, R. (1964). Walking patterns of normal men. *Journal of Bone and Joint Surgery*, *46*, 335–360.

Myung, J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*, 90–100. DOI:10.1016/S0022-2496(02)00028-7.

Neapolitan, R. (2003). *Learning Bayesian networks*. New York: Prentice Hall.

Nixon, M., & Carter, J. (2006). Automatic recognition by gait. *IEEE Special Issue on Biometrics: Algorithms & Applications*, *94*, 2013–2024.

Nixon, M., Tan, T., & Chellappa, R. (2006). *Human identification based on gait*. Berlin: Springer.

Pearl, J. (1997). *Probabilistic reasoning in intelligent systems*. San Mateo: Morgan Kaufmann.

Russel, S., & Norvig, P. (1995). *Artificial Intelligence a modern approach*. New York: Prentice Hall.

Sarkar, S., Phillips, P., Liu, Z., Vega, I., Grother, P., & Bowyer, K. W. (2005). The humanid gait challenge problem: data sets, performance, and analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*, 162–177.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464.

Tong, Y., Liao, W., & Ji, Q. (2007). Facial action unit recognition by exploiting their dynamic and semantic relationships. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(10), 1683–1699.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*(1), 71–86.

Veeraraghavan, A., Chowdhury, A., & Chellappa, R. (2004). Role of shape and kinematics in human movement analysis. In *Proceedings of IEEE conference on computer vision and pattern recognition* (Vol. 1, pp. 730–737).

Veeraraghavan, A., Srivastava, A., & Chowdhury, K. (2009). Rate-invariant recognition of humans and their activities. *IEEE Transactions on Image Processing*, *18*(6), 1326–1339.

Xu, D., Yan, S., Tao, D., Zhang, L., Li, X., & Zhang, H. J. (2006). Human gait recognition with matrix representation. *IEEE Transactions on Circuits and Systems for Video Technology*, *16*(7), 896–903.

Xu, Z., Chen, H., Zhu, S., & Luo, J. (2008). A hierarchical compositional model for face representation and sketching. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, *30*(6), 955–969.

Yan, S., Xu, D., Yang, Q., Zhang, L., Tang, X., & Zhang, H. (2005). Discriminant analysis with tensor representation. In *Proc. IEEE conf. comput. vision pattern recognit.*, 2005 (pp. 526–532).

Yu, S., Tan, D., & Tan, T. (2006). A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In *The 18th international conference on pattern recognition (ICPR)*.

Zhou, Z., Bennett, A. P., & Damper, R. (2006). A Bayesian framework for extracting human gait using strong prior knowledge. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*(11), 1738–1752.