Learning gait component relationships by fusing logic and graphs using Markov Logic Networks

Ibrahim Venkat (Formerly K. Venkatasubramanian), Philippe De Wilde
School of Mathematical and Computer Sciences
Heriot-Watt University
Edinburgh UK.
vk52@hw.ac.uk

Abstract – *Gait recognition is a newly developing biometric which has potential to recognize people at a distance when application of other biometrics might not be feasible. We propose a new technique to represent and learn various gait component relationships using a recently developing statistical relational learning technique called Markov Logic Networks. Markov Logic Network is a robust statistical learning technique that fuses expressive first-order logic with probabilistic graphical models and prove to be efficient in handling noisy and uncertain data. Initially we derive component based pattern classifiers in the imaging domain using an automatic segmentation scheme and represent gait components and their relationships using first-order logic. Then we model and learn their characteristics using undirected graphs to finally classify gaits based on standard inference techniques. The proposed approach enables automatic gait recognition from low resolution videos and differs from conventional techniques which rely on manual markings on videos. We show that the proposed representation provide intuitive means to reason gait component relationships. Our results show that the proposed approach competes well with other state-of-the-art techniques.*

Keywords: Gait recognition, biometrics, logic-based fusion, Markov Logic Networks.

1 Introduction

Biometrics is the study of automated methods for recognizing people and a biometric is a physiological and/or behavioral characteristic that can be used to identify an individual. The field of computer vision perceives the usage of biometrics primarily as a form of identity access management and access control. Gait is considered more as a behavioral characteristic which aims to recognize people by processing walking patterns of human motion from video sequences. Usually in surveillance applications, it is difficult to acquire images from other biometrics such as face or iris at the resolution required for recognition. Though psychophysical studies such as [1] indicate that humans have the capability of recognizing people from even dull displays of gait, gait based biometric systems are still at the experimental stage. Contributions from various fields such as biomechanics, psychology, neuroscience and so on have been gaining momentum to improve gait recognition more feasible so that gait enabled biometrics can be deployed by surveillance and security control applications in scenarios where human recognition at a distance is more compelling.

Probabilistic graphical models are capable of compactly encoding complex distributions and provide intuitive means to handle uncertainty. Whereas First-order logic provides means to represent real world objects and their relationships using simple but efficient predicates and rules. The recently proposed Markov Logic Network (MLN) [2] fuses these two major paradigms in a single representation which has long been a goal of AI researchers. As such MLN serves as a robust technique to solve vital problems in statistical relational learning domains such as classification, clustering, object recognition and so on. The binary silhouette data which are considered as the primary features of Gait videos are subject to multiple uncertainty factors such as noise, low resolution, illumination and so on. We introduce a novel technique in this paper which deploys MLN to represent, interpret, learn gait component relationships and finally able to classify gaits.

We abbreviate the proposed **G**ait **R**ecognition **M**odel which uses **M**arkov **L**ogic **N**etwork as (**GRM-MLN**). The three-layer architecture employed by the GRM-MLN is shown in Figure 1. We briefly describe the three stages inherent in the proposed framework as follows:

1. Image Processing Layer (IPL)

   Initial image processing and component based classifications are performed at the image processing layer. Raw binary silhouettes which form the core input of the model are initially decomposed into three gait components namely right-half, left-half and the lower-part using basic image segmentation techniques after normalizing them using standard image processing techniques. Eigen-tensor based features [3] are extracted from these gait components from which component based recognition is performed.

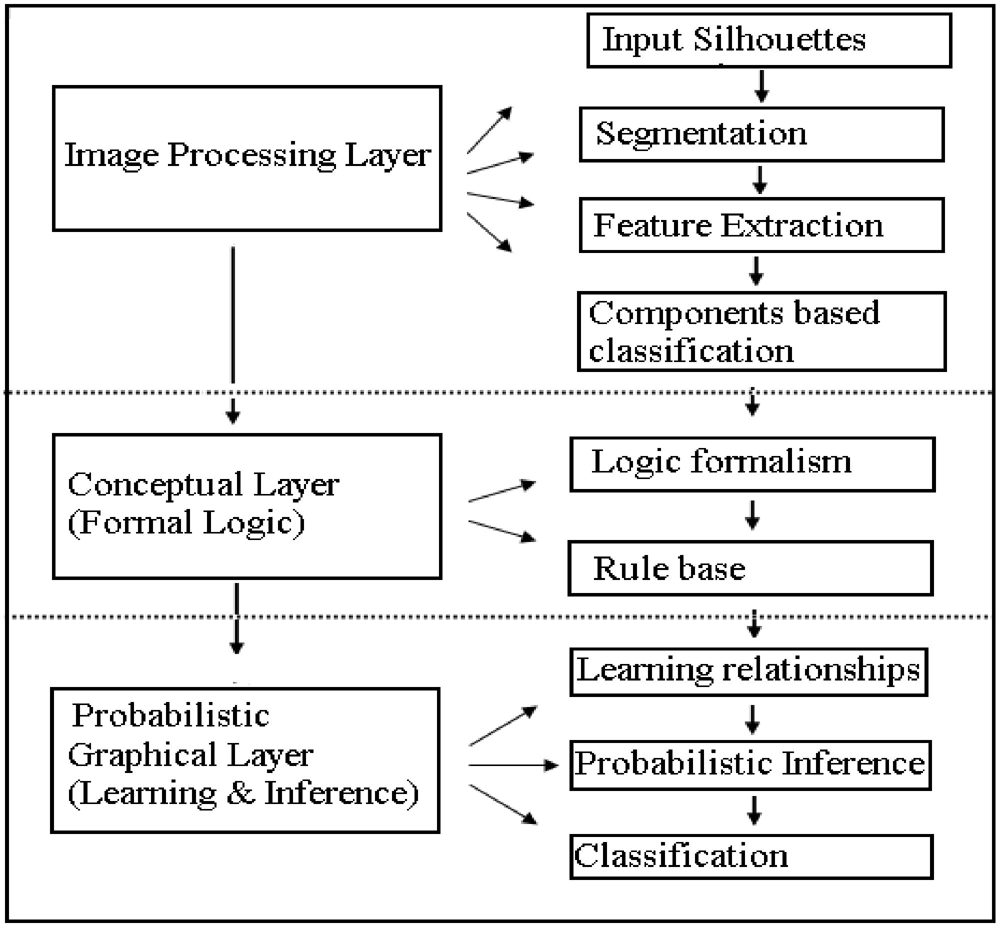

Figure 1: Three-Layer architecture of the proposed GRM-MLN

ii. Conceptual Layer (CL)

This layer is comprised of a set of predicate definitions (first-order formulas) and a rule-base governing how various gait components can be relatively combined. The information gained from the weak classifiers based on components based recognition from the IPL is transformed into the logical layer in terms of ground atoms.

iii. Probabilistic Graphical Layer (PGL)

Each of the gait components individually and/or collectively could recognize gaits. Undirected graphs employed by the GRM-MLN represent the dependencies between the components and the gaits which they recognize. The graph has a node for each variable, and the model has a potential function (weight) for each clique in the graph. From the training dataset the potentials encoded by the graphs receive possible groundings for the atomic formulas which are precisely governed by the rules in the rule-base of the logical layer. This enables GRM-MLN to learn characteristics about gait components and their relationships. A given probe is subjected to segmentation and component based classification is performed. The information obtained from these weaker classifiers are fed into the GRM-MLN as evidences and finally the most probable gait recognized for the given probe is determined.

The rest of the paper is organized as follows. Section 2 highlights related work relevant to component based object models and silhouette based gait approaches. Section 3 provides the theoretical constructs of GRM-MLN and Section 4 reports experimental results and comparisons. Finally Section 5 summarizes our conclusions and future avenues.

2 Motivation from related work

2.1 The trend of component based gait recognition approaches

Early medical studies [4] reports that there are 24 different components to human gait, and that, if all the measurements are considered, gait is unique as referred by Kale et al. [5]. Also lessons learnt from matured biometrics such as face [6, 7] infers that component based object models show improved recognition. Consequently in computer vision many papers attack the gait recognition problem using a component based reasoning. A human body component-based approach has been proposed by Boulgouris et al. [8]. Body components are manually labeled and weights have

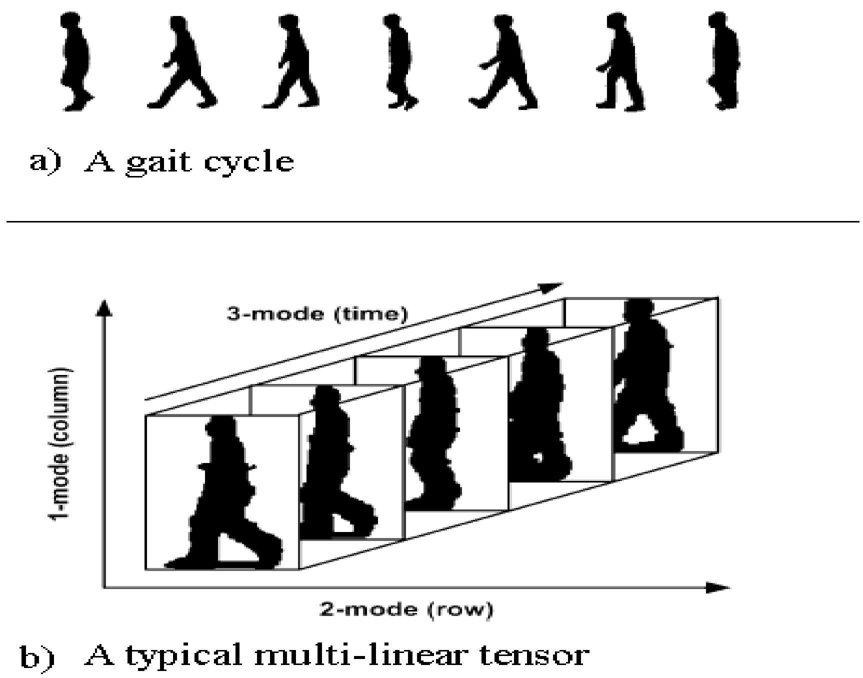

Figure 2: (a) A typical gait cycle representing a segment of a human walk (b) Frames containing silhouettes of a gait sample being stacked into a 3-mode multi-linear tensor (courtesy of Lu et al. [3])

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. Recently Li et al. [9] 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 recognition and they lead to performance improvement. Motivated by these studies the proposed GRM-MLN strategically uses a component based reasoning to recognize gaits. The automatic segmentation scheme what we have employed and the fusion of image processing, logic and graphical models significantly make GRM-MLN much different to approaches such as [8–11] which rely on manual markings or ground truth data that represent landmark co-ordinates. To the best of the authors knowledge, GRM-MLN is the ever first approach in biometrics literature to deploy the MLN technique for gait recognition.

2.2 Feasibility of silhouette based approaches for gait recognition

Gait recognition can be broadly divided into three categories namely temporal alignment-based, static parameter-based and silhouette shape-based [11]. 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. The third class of approaches rely on binary silhouettes of gait sequences. Silhouette-based gait recognition techniques are gaining much interest among current gait recognition approaches [12–14]. 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 [5]. Recent studies [15,16] have shown that silhouette shape has equal, if not more, recognition potential than gait kinematics as referred by Liu et al. [17]. Motivated by these approaches we too mainly use binary silhouettes.

3 GRM-MLN Formulation

We will describe the theoretical constructs of GRM-MLN in this section.

3.1 Gait component representation

A gait recognition algorithm typically consists of the following steps: algorithm to partition a gait sequence into normalized gait cycles, feature representation and extraction, and classification. A training or test sample is well defined in many object recognition (eg. 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 (See Figure 2 a). 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. In the gallery set, each subject's walking behavior is represented as several gait samples. However 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 linear time mode normalization technique proposed by [12]. We represent each gait cycle as a 3-mode multi-linear tensor after normalizing it, like the one shown in Figure 2 b).

GRM-MLN is a generic object recognition model and hence it can flexibly fit into any feature projection technique such as PCA or SVM. For demonstration sake we represent the silhouettes in terms of multi-linear tensors which has been recently applied by Lu et al. [3]. A tensor object can be defined as a multidimensional object (array), the elements of which are to be addressed by more than two indices [18]. An N^{th} order tensor denoted as $A \in \mathbb{R}^{I_1 \times I_2 \times \cdots I_n}$ can be expressed as

$$\mathbb{A} = \mathbb{S} \times_1 U^{(1)} \times_2 U^{(2)} \times \cdots \times_N U^{(N)} \tag{1}$$

and conversely

$$\mathbb{S} = \mathbb{A} \times_1 U^{(1)^T} \times_2 U^{(2)^T} \times \cdots \times_N U^{(N)^T} \tag{2}$$

It is addressed by N indices $i_n, n = 1, \cdots, N$, and each i_n addresses the n-mode of $\mathbb{A}$. $U^{(n)}$ is an orthogonal $I_n \times I_n$ matrix which can be written as

$$U^{(n)} = u^{(1)} u^{(2)} \cdots u^{(n)}_{I_n} \tag{3}$$

A rank-1 tensor $\mathbb{A}$ comprises of the outer product of N vectors $\mathbb{A} = u^{(1)} \circ u^{(2)} \circ \cdots \circ u^{(n)}$ which can be otherwise written as

$$A(i_1, i_2, \cdots, i_N) = u^{(1)}(i_1) \cdot u^{(2)}(i_2) \cdot \cdots \cdot u^{(N)}(i_N) \tag{4}$$

Each tensor X_m can be written as a linear combination of $p_1 \times p_2 \times \cdots \times p_N$ rank-1 tensors

$$U_{p_1 p_2 \cdots p_N} = u^{(1)}_{p_1} \circ u^{(2)}_{p_2} \circ \cdots \circ u^{(N)}_{p_N}, \tag{5}$$

where $U^{(n)} \in \mathbb{R}^{I_n \times P_n}$ and $P_n < I_n$. These rank-1 tensors are called as eigentensors. From the M training tensorial samples, $X_1, X_2, \cdots, X_M$ of gait silhouettes, eigentensors are extracted which captures most of the variation in the gaits from which a low-dimensional tensorial subspace is constructed. The objective is to determine the N projection matrices $U^{(n)}$ that maximize the total tensor scatter Ψ_y

$$U^{(n)} = \arg \max_{U^{(1)}, U^{(2)}, \cdots, U^{(N)}} \Psi_y, \tag{6}$$

where y is the projection of any input tensor sample over the tensorial subspace. Then the standard nearest neighborhood classifier principle is applied to compute similarity measures between the features of a test sample and training samples and gait recognition is performed.

3.2 Proposed automatic segmentation scheme

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. For a given silhouette frame $I(x, y)$ with width w and height h, its centre (x_c, y_c) can be computed by

$$(x_c, y_c) = \left(\frac{w}{2}, \frac{h}{2} \right) \tag{7}$$

With reference to their centre (x_c, y_c) we segment the gait samples into left and right halves. The lower-part of the gait samples intend to roughly capture the leg motion and are segmented 40% from the bottom portion of a whole silhouette. Such a simple automatic segmentation attempt to capture various gait part motion without relying on any manual markings or ground truth data. By finding the most optimal components of a gait using sophisticated segmentation techniques and optimization algorithms, better recognition accuracies could be achieved which deserve an exclusive study. We restrict the scope of this paper to show how a statistical relational technique can enhance gait recognition by reasoning with a few number of basic gait components without the need of any manual markings.

3.3 Statistical relational learning of gait components relationships using MLN

The segmented gait components and their relationships being real world object entities, we use first-order logic to precisely define the predicates and formulate the rule-base as follows. Since we use three components namely the right-half, left-half and the lower-part, basically three evidence predicates viz., $Lftbody(person)$, $Rgtbody(person)$ and $Lowerbody(person)$ are possible. Each component either individually or collectively might trigger the recognition of a gait which in turn leads to a series of component based recognition rules. For example the following rules intuitively represent how roughly lower-left-leg and lower-right-leg motion can influence the gait recognition process.

$$Lftbody(x) \wedge Lowerbody(x) => Recognize(x) \tag{8}$$

$$Rgtbody(x) \wedge Lowerbody(x) => Recognize(x) \tag{9}$$

Similarly all the possible rules (clauses) have been defined using first-order logic and a precise Knowledge Base (KB) of gait components and their relationships was formed. Each rule defined in the rule base corresponds to the event of a component based recognition mechanism. Initially at the Image Processing Layer (IPL), Eigen-tensor based feature space has been constructed for the various gait components from the training samples. By projecting the eigen-tensor

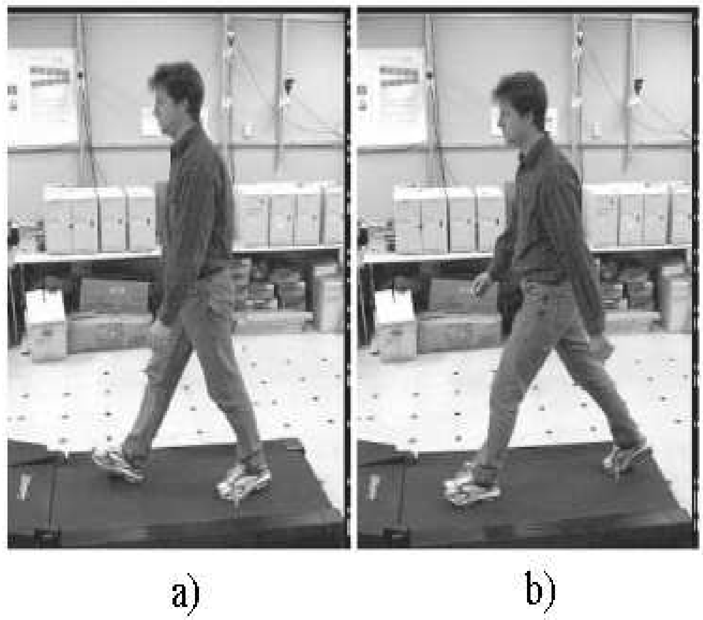

Figure 3: Typical samples from the CMU gait database showing the gait of a subject walking on a treadmill set in the middle of a room for the conditions of a) slow walk and b) fast walk

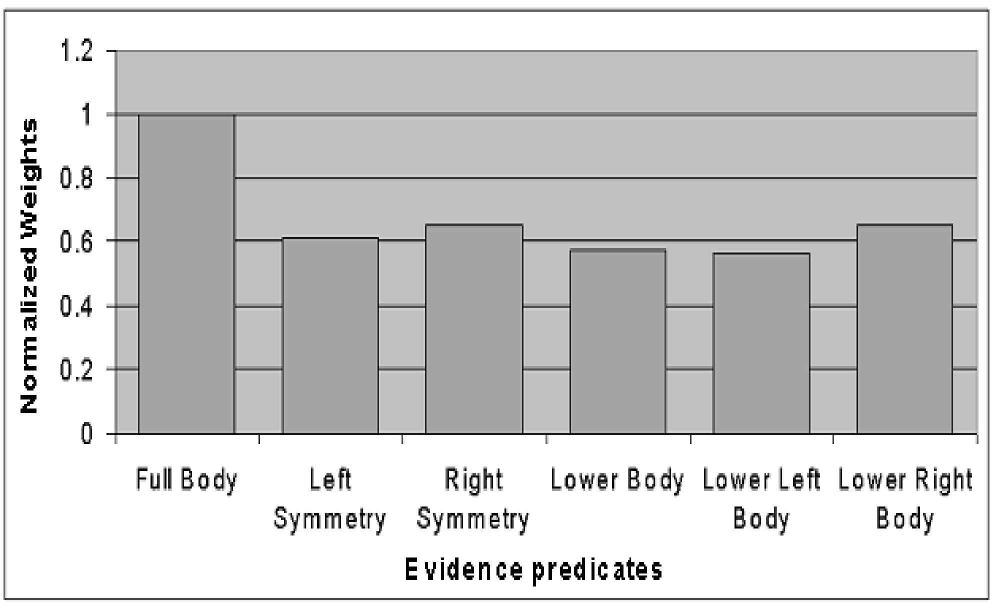

Figure 4: Learnt potentials of evidence predicates for training samples comprising of slow walk

feature of a gait component of a test sample over this feature space, recognition has been performed and the results are transformed into the Conceptual Layer (CL) in terms of ground atoms. In other words, in the event of a component or set of components have influenced the recognition of a gait, the corresponding rule in the KB will receive a grounding.

For an in-depth coverage of first-order logic the user is encouraged to read [21]. However, it is note worthy to brief some basic concepts relevant to *MLN* from [2] here. An MLN is obtained by attaching weights to the formulas (or clauses) in a first-order knowledge base, and can be viewed as a template for constructing ordinary Markov networks. Empirically several algorithms for MLN weight learning have been compared in terms of learning rates by by Lowd and Domingos [22]. Each possible grounding of a formula in the KB yields a feature in the constructed network. Inference is performed by grounding the minimal subset of

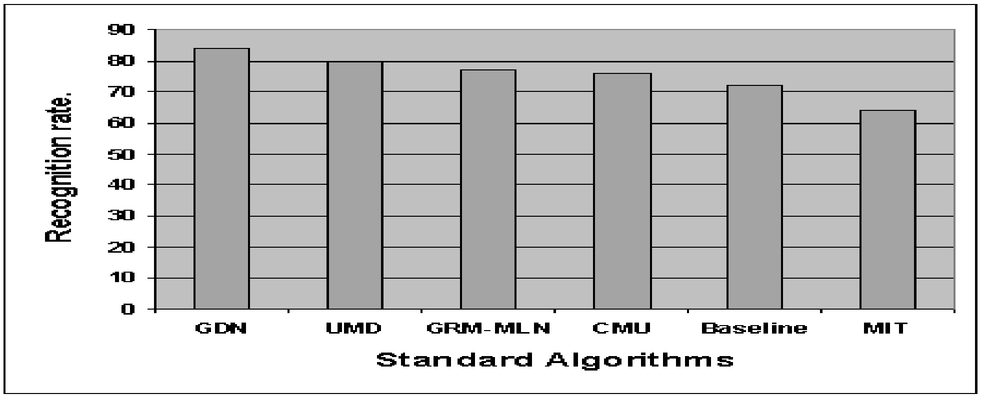

Figure 5: Comparison of GRM-MLN with GDN [11],UMD [10],CMU [15], Baseline [19] and MIT [20]

the network required for answering the query and running a Gibbs sampler over this subnetwork, with initial states found by the MaxWalkSat. Weights are learned by optimizing a pseudo-likelihood measure using the L-BFGS algorithm, and clauses are learned using the CLAUDIEN system. For a given probe, by observing the gaits triggered by its various components, we aim to query the most probable gallery instance with the aid of the learnt potentials readily available.

In AI, a *KB* technically represents a single large formula as the formulas in a KB are implicitly conjoined. A ground term is a term containing no variables. A ground atom or ground predicate is an atomic formula all of whose arguments are ground terms. A possible world assigns a truth value to each possible ground atom. A formula is satisfiable iff there exists at least one world in which it is true. The basic inference problem in first-order logic is to determine whether a knowledge base KB entails a formula F, i.e., if F is true in all worlds where KB is true (denoted by $KB \models F$). This is often done by refutation: KB entails F iff $KB \cup \neg F$ is unsatisfiable.

4 Experiments and discussion

We have used the CMU Mobo data set [23] which consists of gait sequences of subjects walking on a treadmill, positioned in the middle of a room. It contains six simultaneous motion sequences of 25 subjects (23 male, 2 female) walking on a treadmill. The 3CCD progressive scan images have a resolution of 640x480. Each subject is recorded performing four different types of walking: slow walk, fast walk, inclined walk, and slow walk holding a ball (to inhibit arm swing). For our experiments we have used gait silhouettes representing the slow walk and fast walk video sequences. More than 8000 images are captured per subject. Sample video frames for a typical subject for two typical conditions viz., slow walk and speed walk are shown in Figure 3. Each sequence is 11 seconds long, recorded at 30 frames per second. For training and testing we have used the slow walk and fast walk sequences respectively. The average walking speed of the treadmill was set to 2.06 miles per hour (mph) for capturing the slow walk gait sequences. For the fast walk this was set as 2.82 mph. The speed of the treadmill was adjusted to be at a comfortable walking speed for the subjects for both the slow walk and fast walk.

We have implemented GRM-MLN using an open source software called “Alchemy” offered by the *Statistical Machine Learning Group, University of Washington (http://alchemy.cs.washington.edu)*. The learnt potentials for the various gait components are shown in Figure 4. It can be seen that various gait components such as left, right and lower body motion contribute around 55% to 65% towards the overall gait recognition. Lower body motion without any further component based interpretation contributes to a recognition potential of about 57.5%. However the efficient fusion of lower body motion with left and right gait symmetries have enabled us to learn the recognition potentials of two vital components viz., lower-left and lower-right gait components. The average recognition potential of these components contributes to around 61% which is considerably better than the lower body motion alone. This illustrates a significant advantage of the fusion based mechanism deployed by GRM-MLN.

In the class of object recognition problems, identification is considered to be harder than verification [17]. Although gait can disclose more than identity, it is increasingly being applied to identification tasks [14]. 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 recognition (identification) problem which matches a given probe gait sequence against a set of gallery gait sequences, a one-to-many matching. GRM-MLN has yielded a recognition rate of 77%. Further we have compared the recognition rates of GRM-MLN with state-of-the-art gait recognizers with respect to the CMU dataset. From the bar-chart shown in Figure 5 it is seen that the proposed GRM-MLN algorithm competes well with other standard algorithms. The approach proposed by Veeraraghavan et al.,UMD [10], relies on landmark coordinates that describes specific shape configuration. Gait Dynamics Normalization (GDN) [11] performs gait recognition on manually created silhouettes. Though GRM-MLN yields recognition rates which are slightly lower than GDN and UMD, unlike them, it does not rely on any ground truth data. However by using more sophisticated segmentation algorithms and preprocessing techniques, better recognition rates can be achieved.

5 Conclusion and future work

We have proposed a simple but yet efficient statistical relational learning technique to reason and recognize gaits. This study shows how a simple component based gait reasoning approach can be coherently modeled using Markov Logic Networks. The proposed GRM-MLN has a natural generative semantics, which can establish dependencies between gait components and exploit these dependencies to successfully classify gaits. For a newly emerging biometric like gait, every piece of contribution such as GRM-MLN would be a milestone. For future avenues, we intend to extend this generic model to investigate other potential object recognition problems

References

[1] J. Cutting and L. Kozlowski, "Recognizing friends by their walk: Gait perception without familiarity cues," Bulletin of the Psychonomic Society, Tech. Rep., 1977.

[2] M. Richardson and P. Domingos, "Markov logic networks," *Machine Learning*, vol. 63, pp. 107–136, Feb 2006.

[3] H. Lu, K. Plataniotis, and A. N. Venetsanopoulos, "MPCA: Multilinear Principal Component Analysis of tensor objects," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18-39, no. 1, 2008.

[4] M. Murray, A. Drought, and R. Kory, "Walking patterns of normal men," *Journal of Bone and Joint Surgery*, vol. 46, pp. 335–360, 1964.

[5] A. Kale, A. Sundaresan, A. Rajagopalan, N. Cuntoor, A. Chowdhury, K. Volker, and R. Chellappa, "Identification of humans using gait," *IEEE Trans. on Image Processing*, vol. 13, pp. 1163–1173, 2004.

[6] B. Heisele, T. Serre, and T. Poggio, "A component-based framework for face detection and identification," *International Journal of Computer Vision*, vol. 74, pp. 167–181, 2007.

[7] H. Kanan and K. Faez, "Recognizing faces using adaptively weighted sub-gabor array from a single sample image per enrolled subject," *Image & Vision Computing*, vol. 28, no. 3, pp. 438–448, March 2010.

[8] N. Boulgouris and Z. Chi, "Gait recognition based on human body components," in *IEEE International Conference on Image Processing(ICIP)*, vol. 1, Sept 2007, pp. 353–356.

[9] X. Li, S. Maybank, S. Yan, D. Tao, and D. Xu, "Gait components and their application to gender recognition," *IEEE Trans. on Systems, Man and Cybernetics - Part C: Applications and Reviews VOL. 38, NO. 2, MARCH 2008*, vol. 38, pp. 145–155, 2008.

[10] A. Veeraraghavan, A. Chowdhury, and R. Chellappa, "Matching shape sequences in video with applications in human movement analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, pp. 1896–1909, Dec 2005.

[11] Z. Liu and S. Sarkar, "Improved gait recognition by gait dynamics normalization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 863–876, 2006.

[12] N. Boulgouris, K. Plataniotis, and D. Hatzinakos, "Gait recognition using linear time normalization," *Pattern Recognition*, vol. 39, pp. 969–979, 2006.

[13] M. Nixon, T. Tan, and R. Chellappa, *Human Identification Based on Gait*. Springer, 2006.

[14] C. Bauckhage, J. Tsotsos, and F. Bunn, "Automatic detection of abnormal gait," *Image and Vision Computing*, vol. 27, p. 108115, 2006.

[15] R. Collins, R. Gross, and J. Shi, "Silhouette-based human identification from body shape and gait," in *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition*, 2002, pp. 366–371.

[16] A. Veeraraghavan, A. Chowdhury, and R. Chellappa, "Role of shape and kinematics in human movement analysis," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, June 2004, p. 730737.

[17] Z. Liu and S. Sarkar, "Effect of silhouette quality on hard problems in gait recognition," *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics*, vol. 35, pp. 170–183, April 2005.

[18] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "On the best rank-1 and rank - $(r_1, r_2, \cdots, r_n)$ approximation of higher-order tensors," *J. Matrix Anal. Appl.*, vol. 21, pp. 1324–1342, 2000.

[19] S. Sarkar, P. Phillips, Z. Liu, I. Vega, P. Grother, and K.W.Bowyer, "The humanid gait challenge problem: Data sets, performance, and analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 162–177, 2005.

[20] L. Lee, "Gait analysis for classification," Ph.D. dissertation, Massachusets Inst. of Technology, 2003.

[21] S. Russel and P. Norvig, *Artificial Intelligence a modern approach*. Prentice Hall, 1995.

[22] D. Lowd and P. Domingos, "Efficient weight learning for markov logic networks," in *European Conference of Principles and Practice of Knowledge Discovery in Databases (ECMLPKDD-2007)*, 2007, pp. 200–211.

[23] R. Gross and J. Shi, "The CMU motion of body (mobo) database," Tech. Rep., 2001.

Ibrahim Venkat (Formerly Krishnamurthy Venkatasubramanian) received the B.Sc.(Mathematics) from Madurai Kamaraj University, South India, in 1989 and the M.Sc. (AI) from Universiti Malaysia Terengganu in 2006. After serving CSIR and TATA group, South India, he served as a Head of Computing in Kolej Univ. TATI Malaysia since 1996. Recently he has been associated with Universiti Sains Malaysia. He has been awarded several awards including Gold medals for his inventions and innovations. Currently he is doing Ph.D. in Computing at Heriot-Watt University on a meritorious James Watt Scholarship. His research interests include biometrics, computer vision and machine learning.