

Symbolic Modelling of Dynamic Human Motions

David Stirling, Amir Hesami, Christian Ritz,
Kevin Adistambha and Fazel Naghdy
*The University of Wollongong
Australia*

1. Introduction

Numerous psychological studies have shown that humans develop various stylistic patterns of motion behaviour, or dynamic signatures, which can be in general, or in some cases uniquely, associated with an individual. In a broad sense, such motion features provide a basis for non-verbal communication (NVC), or body language, and in more specific circumstances they combine to form a Dynamic Finger Print (DFP) of an individual, such as their gait, or walking pattern.

Human gait has been studied scientifically for over a century. Some researchers such as Marey (1880) attached white tape to the limbs of a walker dressed in a black body stocking. Humans are able to derive rich and varied information from the different ways in which people walk and move. This study aims at automating this process. Later Braune and Fischer (1904) used a similar approach to study human motion but instead of attaching white tapes to the limbs of an individual, light rods were attached. Johansson (1973) used MLDs (Moving Light Displays; a method of using markers attached to joints or points of interests) in psychophysical experiments to show that humans can recognize gaits representing different activities such as walking, stair climbing, etc. The Identification of an individual from his/her biometric information has always been desirable in various applications and a challenge to be achieved. Various methods have been developed in response to this need including fingerprints and pupil identification. Such methods have proved to be partially reliable. Studies in psychology indicate that it is possible to identify an individual through non-verbal gestures and body movements and the way they walk.

A new modelling and classification approach for spatiotemporal human motions is proposed, and in particular the walking gait. The movements are obtained through a full body inertial motion capture suit, allowing unconstrained freedom of movements in natural environments. This involves a network of 16 miniature inertial sensors distributed around the body via a suit worn by the individual. Each inertial sensor provides (wirelessly) multiple streams of measurements of its spatial orientation, plus energy related: velocity, acceleration, angular velocity and angular acceleration. These are also subsequently transformed and interpreted as features of a dynamic biomechanical model with 23 degrees of freedom (DOF).

This scheme provides an unparalleled array of ground-truth information with which to further model dynamic human motions compared to the traditional optically-based motion capture technologies. Using a subset of the available multidimensional features, several

successful classification models were developed through a supervised machine learning approach.

This chapter describes the approach, methods used together with several successful outcomes demonstrating: plausible DFP models amongst several individuals performing the same tasks, models of common motion tasks performed by several individuals, and finally a model to differentiate abnormal from normal motion behaviour.

Future developments are also discussed by extending the range of features to also include the energy related attributes. In doing so, valuable future extensions are also possible in modelling, beyond the objective pose and dynamic motions of a human, to include the intent associated with each motion. This has become a key research area for the perception of motion within video multimedia, for improved Human Computer Interfaces (HCI), as well as its application directions to better animate more realistic behaviours for synthesised avatars.

2. Dynamic human motions used in bodily communication

Bodily communication or non-verbal communication (NVC) plays a central part in human social behaviour. Non-verbal communication is also referred to as the communication without words. Face, hands, shrugs, head movements and so on, are considered as the NVC. These sorts of movements are often subconscious and are mostly used for:

- Expressing emotions
- Conveying attitudes
- Demonstrating personality traits
- Supporting verbal communication (McNeil, 205)

Body language is a subset of NVC. Body language is used when one is communicating using body movements or gestures plus, or instead of, vocal or verbal communication. As mentioned previously these movements are subconscious, and so many people are not aware of them although they are sending and receiving these all the time. Researchers have also shown that up to 80% of all communications is body language. Mehrabian (1971) reported that only 7% of communication comes from spoken words, 38% is from tone of the voice, and 55% comes from body language.

A commonly identified range of NVC signals have been identified (Argyle, 1988) such as:

- Facial expression
- Gaze and pupil direction
- Posture
- Non-verbal vocalizations
- Clothes, and other aspects of appearance
- Bodily contact
- Gesture and other bodily movements
- Spatial behaviour
- Smell

In addition to this as Argyle described the meaning of a non-verbal signal can be different from sender or receiver's points of view. To a sender it might be his emotion, or the message he intends to send and to the receiver can be found in his interpretation. Some NVC signals are common among all the different cultures where some others might have different meanings in different cultures. According to Schmidt and Cohn (2002) and Donato et al. (1999) there are 6 universally recognized facial expressions:

- | | |
|------------|-------------|
| 1. Disgust | 2. Fear |
| 3. Joy | 4. Surprise |
| 5. Sadness | 6. Anger |

But there are other emotions that could be recognized through body movements including anxiety, nervousness, embarrassment, lying, aggression, boredom, interest, tiredness, defensive, curiosity, agreement, disagreement, and even some states such as thinking and judging. Some emotions are expressed as a sequence of movements, so one will need to use prior or posterior information from movements in order to be able to recognize such specific emotions.

2.1 Body parts and related emotions

Certain movements of one body part often need to be associated with the movements of various other parts in order to be interpreted as an emotion. Table 1 details a basic list of the parts that one is able to acquire data from their movements and the emotions related to those movements are described.

member	movement	interpretation
<i>head</i>	<i>lowering</i>	<i>defensive or tiredness.</i>
	<i>raising</i>	<i>interest, visual thinking.</i>
	<i>tilting</i>	<i>interest, curiosity.</i>
	<i>oscillating up & down</i>	<i>agreement.</i>
	<i>oscillating left & right</i>	<i>disagreement.</i>
<i>arms</i>	<i>touching</i>	<i>thinking.</i>
	<i>expanding</i>	<i>aggression</i>
<i>hands</i>	<i>crossing</i>	<i>anxiety</i>
	<i>holding behind</i>	<i>lying, self confidence</i>
	<i>palms up or down</i>	<i>asking</i>
	<i>rubbing together</i>	<i>extreme happiness.</i>
<i>neck</i>	<i>repetitive movements</i>	<i>anxiety, impatience..</i>
	<i>touching</i>	<i>fear.</i>
<i>shoulder</i>	<i>raised</i>	<i>tension, anxiety or fear.</i>
	<i>lowered</i>	<i>relax</i>
<i>chest</i>	<i>rubbing</i>	<i>tension and stress.</i>
<i>belly</i>	<i>Rubbing or holding</i>	<i>tension</i>
<i>legs</i>	<i>standing with feet together</i>	<i>anxiety</i>
	<i>crossing</i>	<i>tension and anxiety</i>
	<i>repetitive movements</i>	<i>anxiety, impatience</i>
<i>thighs</i>	<i>touching</i>	<i>readiness</i>
<i>feet</i>	<i>curling</i>	<i>extreme pleasure</i>
	<i>stamping</i>	<i>anger and aggression</i>
	<i>moving</i>	<i>anxiety, impatience, lying</i>

Table 1. Noted emotions for associated body movements (Straker, 2008).

These interpretations are acquired from different psychological researches through different web sites and dissertations. Interpretation would clearly depend on cultural and other context.

Table 1 infers a highly complex multidimensional space in which a human body can relay emotional expressions as various spatial articulations at any point in time. This together with any associated temporal sequence surrounding an observed postural state, combine to provide an extremely challenging context in which to capture and further model the

dynamics of human motions. A rich array of initial, contributory intentions further obfuscate matters. The decidedly successful analysis of facial micro expressions by Ekman and others (Ekman, 1999) has proven insightful for identifying the underlying emotions and intent of a subject. In a related but possibly more prosaic manner, it is the intended to establish three basic goals from the analysis and modeling of dynamic motions of a human body, these are to:

1. develop a sufficient model of dynamic finger printing between several individuals
2. model distinctive motion tasks between individuals
3. formulate a model to identify motion pretence (acting) as well as normal and abnormal motion behaviours

Successfully achieving some or all of these goals would provide invaluable outcomes for human behavioural aspects in surveillance and the detection of possible terrorism events as well as medical applications involving dysfunction of the body's motor control.

3. Motion capture data

Given the three distinct task areas it became prudent to utilise, were ever possible, any existing general motion capture data that may be available, as well as record specific motion data that addressed more specific task needs. To this end the Carnegie Mellon University (CMU) Motion Capture Database (2007) has been utilized explore the second goal, that is to investigate plausible models for the identification of distinctive motion tasks between individuals. This database was created with funding from NSF EIA-0196217, and has become a significant resource providing a rich array of motion behaviours that have been recorded over a prolonged period. Alternatively, the first and last goal objectives require more specific, or specialised captured motion data. For these areas, a motion capture system based on a network array of inertial wireless sensors, as opposed to the more traditional, optical multiple camera based system.

3.1 Inertial motion capture

Data recorded from this technology is being acquired using an inertial movement suit, Moven® from Xsens Technologies, which provides data on 23 different segments of the body kinematics such as position, orientation, velocity, acceleration, angular velocity and angular acceleration as shown in Fig. 1.

In capturing human body motion no external emitters or cameras are required. As explained by Roetenberg et al. (2007) mechanical trackers use Goniometers which are worn by the user to provide joint angle data to kinematic algorithms for determining body posture. Full 6DOF tracking of the body segments are determined using connected inertial sensor modules (MTx), where each body segment's orientation, position, velocity, acceleration, angular velocity and angular acceleration can be estimated. The kinematics data is saved in an MVNX file format which is subsequently read and used, using an intermediate program coded in MATLAB.

Using the extracted features, a DFP (Dynamic Finger Print) can be generated for each individual. DFP is used to identify the individual or detect departure from his/her expected pattern of behaviour. Using this comparison, it is possible to find the smoothness or stiffness of the movement and find out if the person is concealing an object. In order to recognize identity of an individual, different measurements will be made to extract the unique

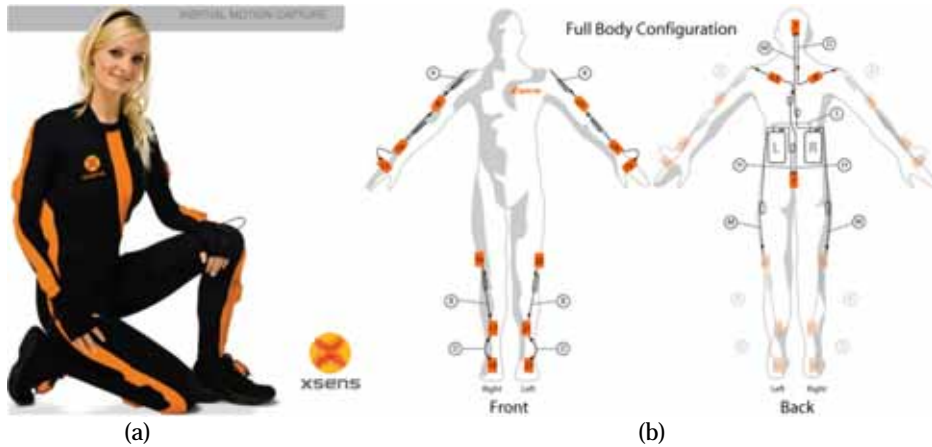


Fig. 1. Inertial Motion Capture: (a) Moven®, light weight latex motion suit housing a network of 16 MTx inertial sensors (b) distribution of MTx sensors including the L and R aggregation and wireless transmitter units— adapted from (Xsens Technologies, 2007).

Dynamic Finger Print (DFP) for that individual. The data produced by the suit consists of kinematics information associated with 23 segments of the body. The position, velocity, acceleration data for each segment will be then analyzed and a set of feature of derived will be used in classification system.

3.2 Feature extraction

The determination/selection and extraction of appropriate features is an important aspect of the research. All the classification results would be based on the extracted features. The features should be easy to extract and also must contain enough information about the dynamics of the motion. The selected features should be independent of the location, direction and trajectory of the motion studied. In the case of a sequence of walking motions (or gait) it would be reasonable to deduce that the most decisive/important facets to consider would be the legs, feet and arms. Features are extracted in a gait cycle for each individual. The gait cycle is a complete stride with both legs stepping, starting with the right leg as shown in Fig. 2. A typical recording session of a participant wearing the suit is shown in Fig. 3.

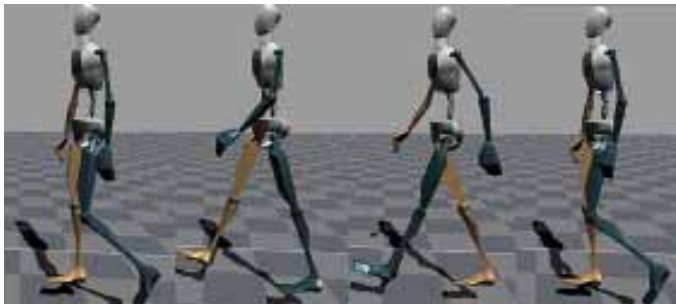


Fig. 2. A sample gait cycle: as received from the wireless inertial motion suit and animated on a 23 DOF avatar within the Moven Studio™ software.

The data produced by the Moven system is stored in rich detail within an MVNX (Moven Open XML format) file which contains 3D position, 3D orientation, 3D acceleration, 3D velocity, 3D angular rate and 3D angular acceleration of each segment in an XML format (ASCII). The orientation output is represented by quaternion formalism.



Fig. 3. Recording of the Body Motions; on average, each participant walked between ground markers, white to black, and return in some seven seconds.

The extracted features chosen are the subtended angles of the following body elements:

- Left and Right Foot Orientation,
- Left and Right Foot,
- Left and Right Knee,
- Left and Right Thigh,
- Left and Right Elbow,
- Left and Right Arm.

In total 12 features per individual was extracted, were each angle is given in radians. The location and interpretation of these features is illustrated on the animated motion avatar in Fig. 4.

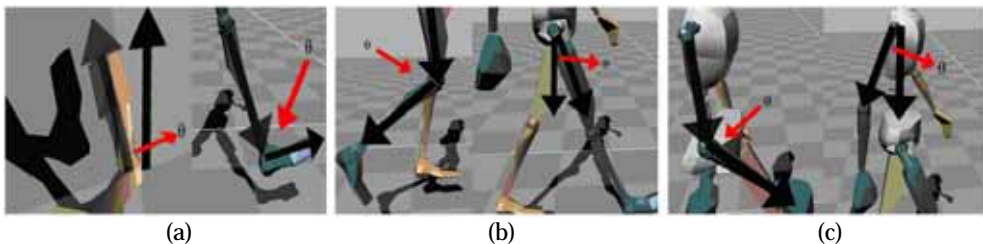


Fig. 4. Selected features annotated of the Moven avatar; (a) Foot Orientation Angle and Foot Angle, (b) Knee Angle and Thigh Angle (c) Elbow Angle and Arm Angle.

An example plot combining all of the 12 selected features, for five participants (p6-p10), can be seen in Fig. 5. These have been concatenated together for comparison; the extent of each individual is delineated by grey vertical lines—each individual marking some 3 to 4 gait cycles in-between. This amounted to some 3 to 4 seconds for a subject to walk from one marker to the other, and for a sample rate of 120Hz this equates to some 360 to 480 captured data frames per person.

One can readily appreciate several various differences in gait amongst these participants—such as the marked variations in angular extent of foot orientations (Left Foot O, Right Foot O), and their associated temporal behaviour. Despite this array of other differences the leg

period of each remains approximately similar as their variation of height is not significant, nor the distance each travelled between the markers during the recording sessions.

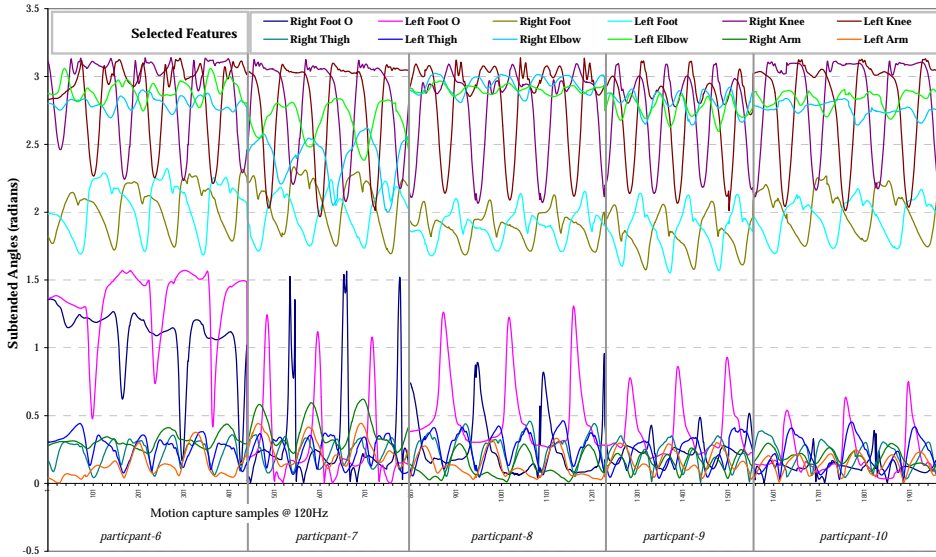


Fig. 5. Temporal trends for the 12 selected features across participants p6—p10.

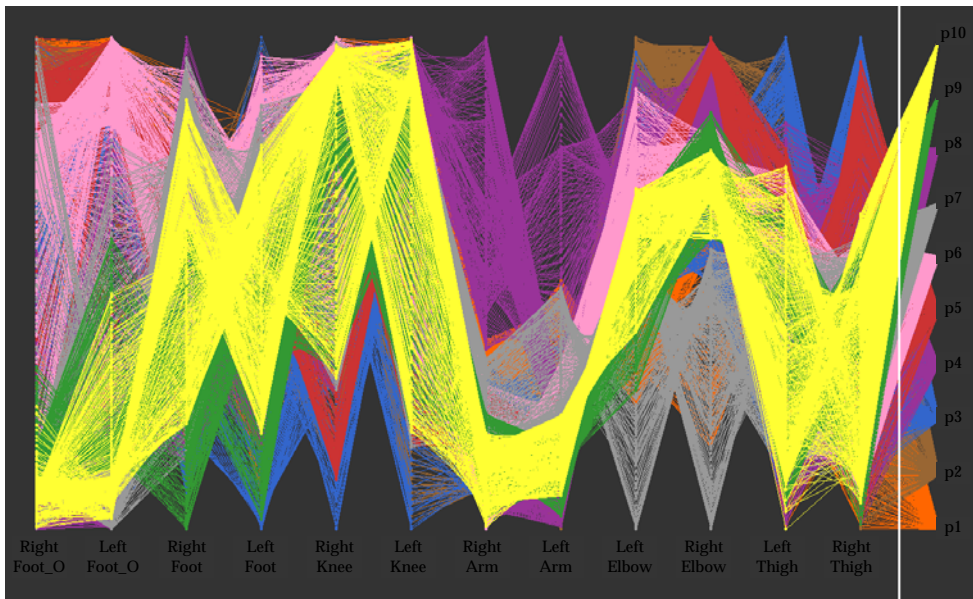


Fig. 6. Parallel Coordinate Plot: providing visualisation of all selected features, for all participants (p1-p10) —covering here, 3837 data frames.

Although there degrees of diversities between the trends in Fig. 5 of all selected features, one may still remain unconvinced that a set of dynamic finger prints ultimately exists, and if so how could they possibly be reliably extracted? Part of this difficulty arises from observing the distinct feature dissimilarities as a function of time. A more pragmatic approach would be to transform these into alternative domains such as FFT or Wavelets. However, an alternative to either of these might be to visualise the features through a Parallel Coordinate Plot (PCP), as illustrated in Fig. 6, in order to explore the multivariate data without the coupling effect of time.

The PCP of Fig. 6 obtained via a visualisation tool Ggobi (Cook and Swayne, 2007), here, arranges a series of parallel coordinates axes, one for each feature, scaled to represent the normalised range of each. The right-most axis of this plot further provides a numerically ordered array of the 10 participants. Every frame of the motion capture data, although constrained to the 12 selected features, is represented by a distinct line that intersects each feature coordinate axis at an appropriate (normalised) value. By colour coding (brushing) the data frames for each participant, one can more readily appreciate potentially unique signatures of profile patterns (or DFP) across the combined feature space. In comparison, both Fig. 5 and Fig. 6 are derived from the same data; however the participants in the former are essentially contrasted with each other (but only half of these for clarity) in the temporal domain. However, in the latter case of Fig. 6 all participants are explicitly compared with each other solely in the feature domain, which also reveals strong visual evidence for the existence of motion signatures amongst the various individuals.

4. Symbolic modelling of DFP

The principal benefit of symbolic machine-learning (modelling), as opposed to other approaches such as physical modelling (or knowledge-driven modelling), is that it is essentially an empirical, or data driven, modelling process which endeavours to represent only the patterns of relationships or process behaviours (here human movements). Hence, it is readily able to cope with significantly higher dimensionality of data. Non-symbolic machine learning approaches, such as artificial neural networks also address such problems, but lack the major benefits offered by symbolic modelling —these being the transparency of learnt outcomes or patterns, plus an adaptive process of the model structure to scale to accommodate data. These abilities are necessary in order to critique and understand patterns and knowledge that may be discovered.

In order to examine the Dynamic Finger Print hypothesis, the ten individuals wearing the Moven suit, undertook four repetitions of a simple walking task. From these tasks, the selected features, across the individuals were collected and recorded for an identification trial. For this trial, the goal was to clearly identify an individual based purely on a combination of the subtended joint angles. In addressing this recognition challenge, the machine learning, rule induction system known as See5 (RuleQuest, 2007) was used. This system, being a supervised learning algorithm was utilised to induce symbolic classification models, such as decision trees, and or rule sets, based on the range of chosen features (attributes), including a priori known classes. The final decision trees and rule sets were created through adjustment of the various pruning options, but primarily through the (major) pruning control for the minimum number of cases option (M).

Essentially a large tree is first grown to fit the data closely and then pruned by removing parts that are predicated to have relatively high error rate. The pruning option, M, is

essentially a stopping criterion to arrest the expansion formation of a decision tree and any associated rule set derived from it. It specifies the minimum number of cases that are required before any leaf classification node is formed and essentially constrains the degree to which the induced model can fit the data. In order to obtain a more reliable estimate of the predictive accuracy of the symbolic model *n-fold* cross validation is used as illustrated in Fig. 7.

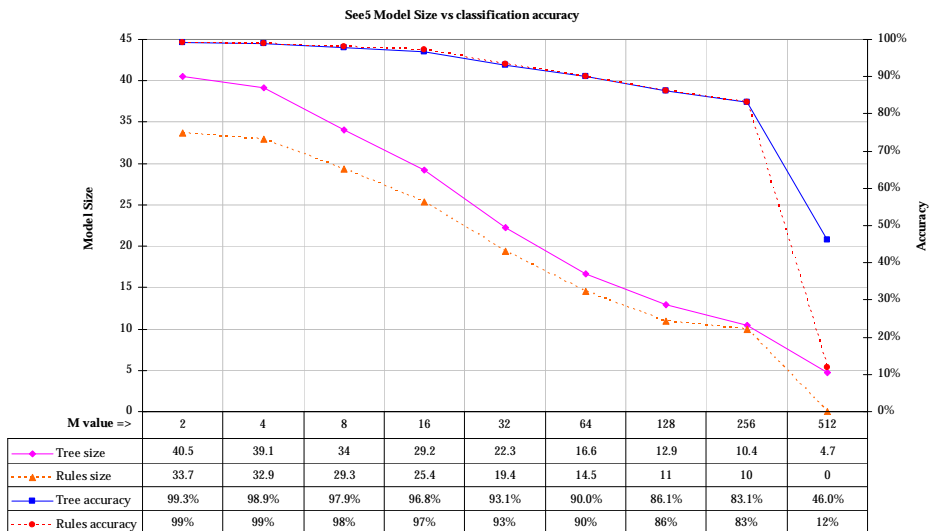


Fig. 7. Model size and accuracy variations as measured by 10-fold cross validation.

The cases in the feature data file are divided into *n-blocks* of approximately the same size and class distribution. For each block in turn, a classifier model is induced from the cases in the remaining blocks and tested on the cases in the hold-out block. In this manner, every data frame is used just once as a test case. The error rate of a See5 classifier produced from all the cases is then estimated as the ratio of the total number of errors on the hold-out cases to the total number of cases (See5, 2002). Here, the number of folds has been set to 10.

As can be seen in Fig. 7 there is a nonlinear trade-off between model size and accuracy. Given that the intended use of the model can be guided as to the most dominant factor. Which at the two extremes can be either; a greater generalisation with a reduced model size or, alternatively, a larger, more sensitive model that is less likely to produce misclassifications. The objective in this task was to model potential motion signatures, and as an example we have chosen a model size that generally reflects a 90~95% accuracy, here M=64. Once a suitable classifier performance level has been identified using the cross validation trends, the resultant model is generated as illustrated by the rule set model in Fig. 8.

For this task we are seeking to establish an individual motion signature for all participants, thus there are ten classes p1-p10. Participants undertaking the experiments were 5 males and 5 females between 18 to 40 years of age. According to Fig. 8, the average error rate achieved is some 6.8% and number of rules is 18.

Rule 1: (1119/728, lift 3.3) Left Foot O > 1.124812 Right Elbow <= 2.901795 => class p1 [0.350]	Rule 2: (296/28, lift 9.7) Left Foot O > 1.124812 Right Elbow > 2.901795 Left Elbow > 2.918272 => class p2 [0.903]	Rule 3: (66/28, lift 6.2) Right Foot O > 1.260007 Left Foot O <= 1.124812 Right Elbow > 2.640656 => class p2 [0.574]	Rule 4: (225/3, lift 10.7) Left Foot O > 1.124812 Right Foot <= 2.100667 Right Knee <= 2.866177 Right Elbow > 2.901795 Left Elbow > 2.795459 Left Arm > 0.1387282 => class p3 [0.982]
Rule 5: (191/21, lift 9.6) Left Foot O > 1.124812 Right Foot <= 2.100667 Right Knee <= 2.866177 Right Elbow <= 2.901795 Right Arm <= 0.2898046 => class p3 [0.886]	Rule 6: (65/25, lift 6.7) Right Foot O > 1.053137 Left Foot O <= 1.124812 Right Elbow <= 2.640656 => class p3 [0.612]	Rule 7: (350, lift 10.9) Left Foot O <= 1.124812 Left Arm > 0.4538144 => class p4 [0.997]	Rule 8: (395, lift 9.4) Left Foot O > 1.124812 Right Elbow > 2.901795 Left Elbow <= 2.918272 => class p5 [0.997]
Rule 9: (224/15, lift 8.1) Left Foot O > 1.124812 Right Knee > 2.866177 Right Elbow > 2.901795 Left Elbow > 2.795459 => class p6 [0.929]	Rule 10: (188/15, lift 8.0) Left Foot O > 1.124812 Right Elbow <= 2.901795 Right Elbow > 2.795459 Right Arm > 0.2898046 Left Arm <= 0.1387282 => class p6 [0.916]	Rule 11: (80/13, lift 7.2) Right Foot O > 1.00804 Right Foot O <= 1.260007 Left Foot O <= 1.124812 Right Elbow > 2.640656 => class p6 [0.829]	Rule 12: (615/311, lift 4.3) Left Foot O > 1.124812 Right Elbow <= 2.901795 Right Arm <= 0.3535621 => class p6 [0.494]
Rule 13: (326, lift 10.7) Right Foot O <= 1.053137 Left Foot O <= 1.124812 Right Elbow <= 2.640656 => class p7 [0.997]	Rule 14: (838/435, lift 4.3) Right Foot O <= 1.00804 Left Foot O > 0.1827743 Left Foot O <= 1.124812 Right Elbow > 2.640656 Left Arm <= 0.4538144 => class p8 [0.481]	Rule 15: (295/16, lift 11.1) Right Foot O <= 1.00804 Left Foot O > 0.1827743 Left Foot O <= 1.124812 Right Elbow > 2.640656 Left Elbow <= 2.852491 Left Arm <= 0.4538144 => class p9 [0.943]	Rule 16: (169/28, lift 9.7) Right Foot O <= 1.00804 Left Foot O <= 1.124812 Left Knee <= 3.004622 Right Elbow > 2.838296 Left Elbow <= 2.879424 Left Arm <= 0.4538144 => class p9 [0.830]
Rule 17: (302, lift 9.3) Right Foot O <= 1.00804 Left Foot O <= 0.1827743 Right Elbow > 2.640656 Left Arm <= 0.4538144 => class p10 [0.997]	Rule 18: (228/28, lift 8.2) Right Foot O <= 1.00804 Left Foot O <= 1.124812 Right Elbow <= 2.838296 Left Elbow > 2.852491 Left Arm <= 0.4538144 => class p10 [0.874]	Default class: p6	
Evaluation on training data (3837 motion frames)			
Decision Tree		Rules	
Size	Errors	No	Errors
18	270(7.0%)	18	261(6.8%)

Fig. 8. An example motion signature model for participants, p1–p10.

Each rule in Fig. 8 consists of an identification number plus some basic statistics such as (n , lift x) or (n/m , lift x) these, in fact, summarize the performance of each rule. Here, n , is the number of training cases covered by the rule and m , where it appears, indicates how many of the cases do not belong to the class predicted by the rule. The accuracy of each rule is estimated by the Laplace ratio $(n - m + 1)/(n + 2)$. The lift x , factor is the result of dividing a rule's estimated accuracy by the relative frequency of the predicted class in the training set. Each rule has one or more antecedent conditions that must all be satisfied if the rule consequence is to be applicable. The class predicted by the rule is shown after the conditions, and a value between 0 and 1 that indicates the confidence with which this prediction is made is here shown in square brackets (See5, 2002).

The overall performance of the signature model can be readily observed in the confusion matrix of Fig. 9 which details all resultant classifications and miss-classifications within the trial. The sum of values in each row of this matrix represents the total number of true motion frames that are derived from the associated participant (p1–p10). Any off-diagonal values in Fig. 9 represent miss-classification errors, such as 13 motion frames of participant p5 was very similar to those exhibited by p2. Here an ideal classifier would register only diagonal values in Fig. 9.

All extracted features were available to the induction algorithm as it constructed its various classifier models, however not all of these were ultimately utilised in the final rules. For

example considering the model of Fig. 8, the number of times that each feature has been referred in the rules, which reflects its importance in classifying a person, is shown in Table 2. According to Table 2 the features, Left Foot, plus the, Left Thigh and Right Thigh, angles have not been used in classifier at all, and the two most important features are angle of the Left Foot Orientation and that of the Right Elbow.

(p1)	(p2)	(p3)	(p4)	(p5)	(p6)	(p7)	(p8)	(p9)	(p10)	<= classified as
384	15				7					p1
	306	21			29					p2
		311			42					p3
			350							p4
	13			395						p5
		3			436					p6
6		25				326				p7
	28						347	28	28	p8
								327		p9
								16	394	p10

Fig. 9. Confusion matrix analysis of the motion signature model for participants p1–p10.

Feature	Usage	Percentage of usage of all features
Left Foot O	18	26.1%
Right Elbow	17	24.6%
Right Foot O	9	13.0%
Left Arm	8	11.6%
Left Elbow	8	11.6%
Right Arm	3	4.3%
Right Knee	3	4.3%
Right Foot	2	2.9%
Left Knee	1	1.4%
Left Foot	0	0%
Right Thigh	0	0%
Left Thigh	0	0%

Table 2. Usage of features, highlighting three redundant attributes.

Although we had originally included all of the apparently, seemingly important bodily attributes, the induced model has found these, Left to be redundant. These leads to an obvious suggestion of not manually selecting or limiting the range of available attributes, but rather allow the algorithm to choose an appropriate sub-set of these. This in fact is one of the specific approaches employed in Section 5.

Ultimately the various rules in such classifiers all define specific hyper-cubes within the multidimensional feature space. As an example, four rules from an initial version of the signature model are overlaid on a 2-dimensional projection of the 12-dimensional feature space. This was observed in some preliminary data visualisation work carried out on the motion data using Ggobi (Cook & Swayne, 2007). Using projection pursuit visualization, the rotating projection was paused whenever a significant 2D segmentation could be observed. Here, in Fig. 10 one can clearly identify participants 9 and 10, and also conceptualize four hyper-cubes encompassing the array of these points (motion data frames) with rules 13, 14, 23 and 24.

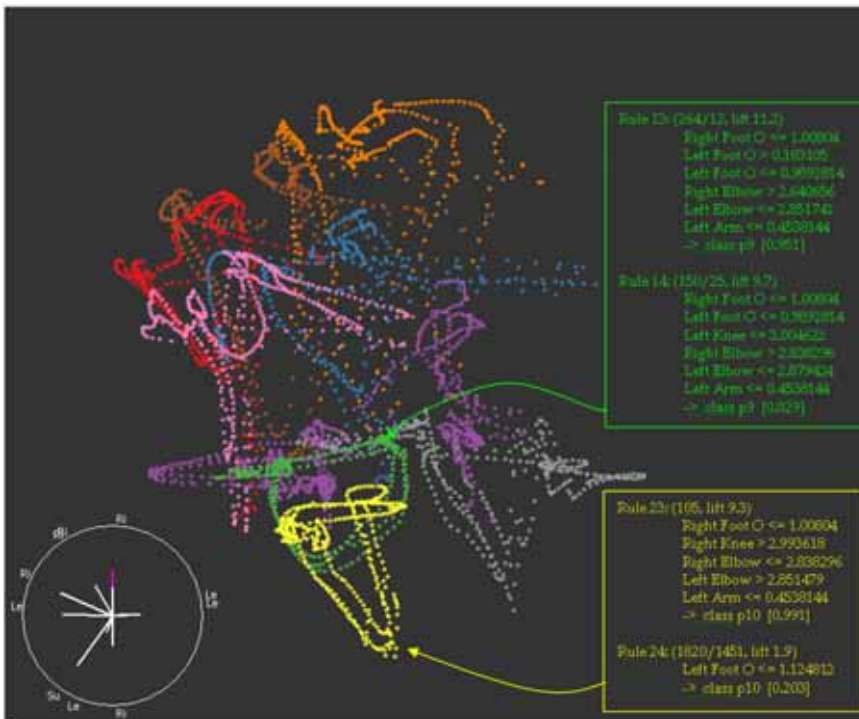


Fig. 10. Selective symbolic model rules identifying participants 9 and 10 with a 2D projection of the 12 dimensional feature-space.

The primary aim of this study was to identify a person based on a combination of subtended angles at the feet, knees, thighs, arms, and elbows. In this process 12 features were extracted and using a decision tree and converting this into a rule set classifier 93.2% accuracy was achieved. The participants were 5 males and 5 females between 18 to 40 years of age, indicating that the results obtained were not dependent to specific characteristics of participants. The extracted features could also be used in gender classification, or even different motion classifications. In order to be able to use the described method in a real application, an image processing and computer vision section for data acquisition should be added to the system. The goal in this section is only to test the hypothesis that a plausible

signature model to recognize specific individuals could be developed from an appropriate set of features.

5. Symbolic modelling of distinctive motion tasks

This section progresses the development of symbolic modelling to see if it can be used to model various distinctive tasks of human movement skills. As mentioned in Section 3, the CMU Motion Capture Database (2007) offers a significant array of general motions, which would take a considerable period of time to replicate. This data, however, is freely available from the Carnegie-Mellon Motion Capture Database, in the Acclaim ASF/AMC format (CMU Motion Capture Database, 2007).

The data consists of motion capture sequences for various activities such as sports, walking, running, dancing, and nursery rhyme actions. These are captured at a rate of 120 frames per second. For each frame, the optically inferred x, y and z axis rotation for each bone of the body are recorded with respect to the degrees of freedom available for the bone, e.g. the upper arm (humerus) has x, y and z rotations while the forearm (radius) has only x-axis rotation from the elbow.

In total, there are 28 bones in the model as shown in Fig. 11, with the 29th bone (root point) representing the rotation and translation of the whole body. This root point serves as the

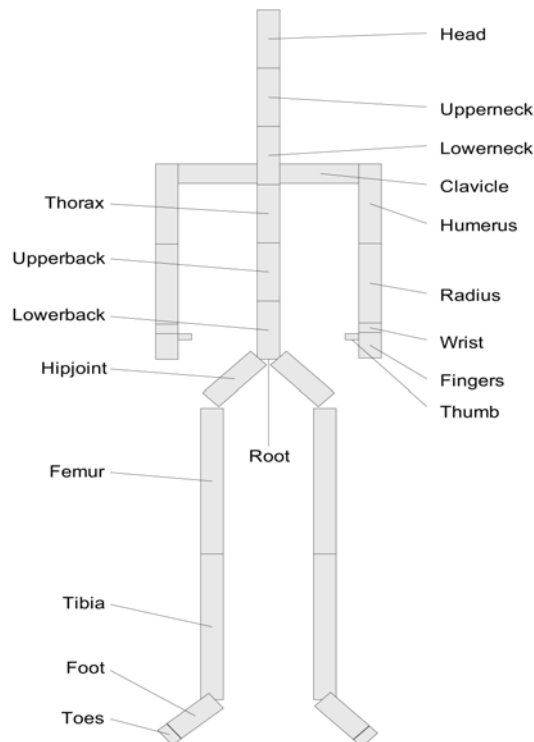


Fig. 11. Names and locations of the bones as per the CMU database used in this work.

point of origin for the whole skeleton and is situated between the lower back, left hip joint and right hip joint, as illustrated in Fig. 11. A plot showing an example of the dataset is shown in Fig. 12. In Fig. 12, the x-axis represents the frame number of the motion and the y-axis represents the degree of rotation applied to each bone in the skeleton. Fig. 12 shows the x, y, and z axis rotation of the lower back bone for two walking motions and a golf swing. For the purposes of this work, four types of motions consisting of walking, running, golf swing and golf putt were used. The motions were chosen to provide visually similar motions (walking and running), visually dissimilar movements but which utilised a similar set of bones (golf swing and golf putt).

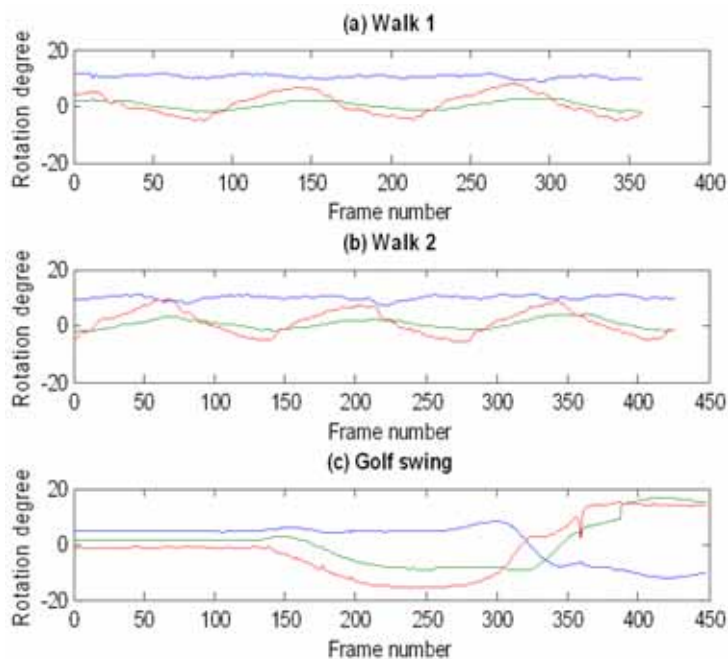


Fig. 12. The plots of x, y, and z axis rotations of the lower back bone of two walking motions and a golf swing with different lengths. Each curve represents rotation of the back bones in the skeleton vs. time.

5.1 Symbolic motion classification using see5

In this section, multiple experiments in developing symbolic models of the motion data using a See5 decision trees were performed where the M value was increased by power of 2 up to 32,768. For each experiment, the size of the decision tree, rule set and the average classification accuracy of each (which was confirmed by 10-fold cross-validation) were recorded. An example of the resulting decision tree for M=8 is shown in Fig. 13. In Fig. 13, a motion is classified by first looking at the root node of the tree, which contains a threshold decision about the left humerus, x-axis rotation. If the condition is not true, then the next node visited specifies that the left wrist, y-axis rotation be examined. Continuing down the

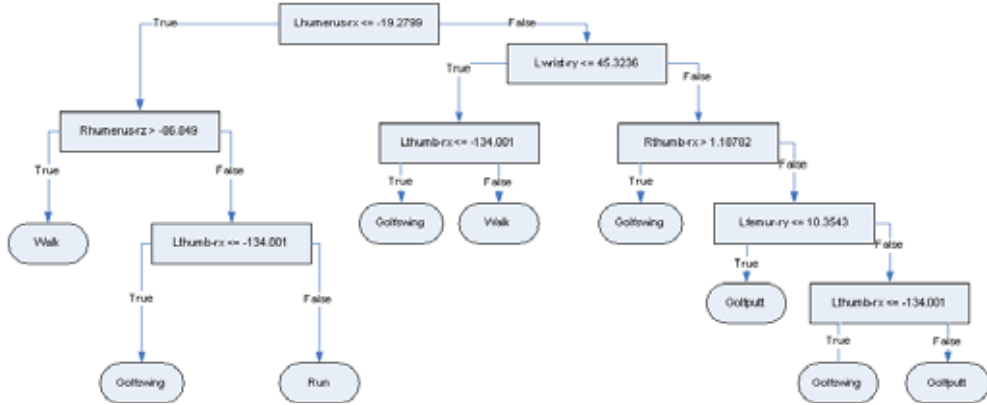


Fig. 13. Symbolic motion decision tree for: walk, run, golfswing and golfputt, using M=8.

tree to one of the leaf nodes, a data frame of a motion can be classified as a golf swing, golf putt, walk or run motion. It can be readily observed in Fig. 13 that to in order to classify these four motion classes, only seven bone tracks out of a possible 62 in the motion data base are actually used, and that these seven are the most important features for differentiating between the four motion classes.

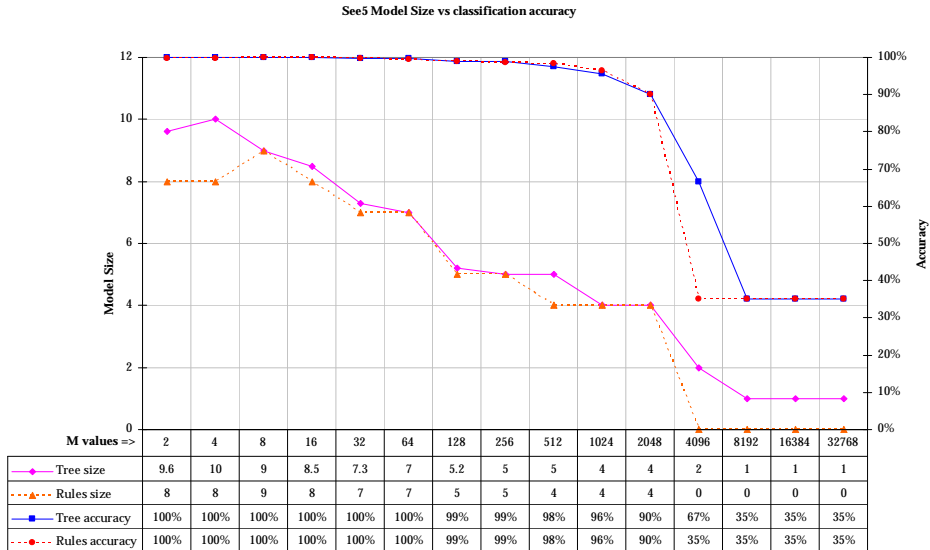


Fig. 14. Symbolic Model size and accuracy variations as measured by 10-fold cross validation for four motion classes (walk, run, golfswing, and golfputt).

From the graph presented in Fig. 14, the tree in Fig. 13 would perform classification with 99.9% accuracy per-frame, which results in 100% accuracy in motion classification. Plots of the M value vs. tree size vs. classification accuracy are shown in Fig. 14.

It is evident that in Fig. 14 that, there is a knee point in the graph approximately where $M=1024$, beyond which the classification accuracy begins to decrease significantly i.e., for $M=1024$ and $M=2048$, classification accuracies are 96% and 90%, respectively. A typical confusion matrix for such models is illustrated in Fig. 15. In Fig. 14 there is a further observed knee point at around $M=2048$, after which for greater values of M the accuracy rate again drops significantly (67% for $M=4096$ and 35% for $M=8192$).

<i>(golfswing)</i>	<i>(golfputt)</i>	<i>(walk)</i>	<i>(run)</i>	<i><= predicted as</i>
4463	4			golfswing
1	2507			golfputt
		6616		walk
			1608	run

Fig. 15. Typical confusion matrix of the motion model ($M=128$) for golfswing, golfputt, walk and run.

It is also of note that parameters of $M=2$ up to $M=32$ yields almost 100% classification results. Fig. 14 also shows that $M=8$ for this dataset provides the best classification performance (99.95%), where using smaller M values was not observed to improve classification performance. Using $M=8$, the resulting decision tree is relatively small with 17 nodes and seven bone motion tracks in total. Hence for the purpose of this work, experiments were performed using decision tree generated with $M=8$.

5.2 Symbolic modelling of normal and abnormal motion behaviours

In order to investigate the concept of being able to detect normal and abnormal motion behaviours, a further series of experiments, again involving the Moven inertial motion suit were designed. In this context individuals were asked to carry a back pack with a 5kg weight in it. From these tasks, the same range of features (as used in Section 3) was used again, for the various individuals undertaking the trial.

For this trial, the goal was to clearly identify if a person is carrying a weight or not. However, in addition to this each participant was invited to subtly disguise their gait on occasions of their choosing, informing the investigators at the end of any recording trial if they had do so. Thus motion data was collected for individual walking gaits that were influenced, or not, by an unfamiliar extraneous weight and also, or not, by a deliberate concealing behaviour of the participant. Again symbolic models of these motion behaviours were induced using the See5 algorithm (RuleQuest, 2007) from the participants using various combinations of subtended joint angles. The algorithm formulates symbolic classification models in the form of decision trees or rule sets, based on a range of several concurrent features or attributes. The model development process followed the same procedure previously discussed in the pervious sections.

For this particular work it was decided to formulate two parallel classifiers to identify both the gender of an individual as well as attempting to deduce if the individual was in fact carrying a weight. The layout of the system is shown in Fig. 16.

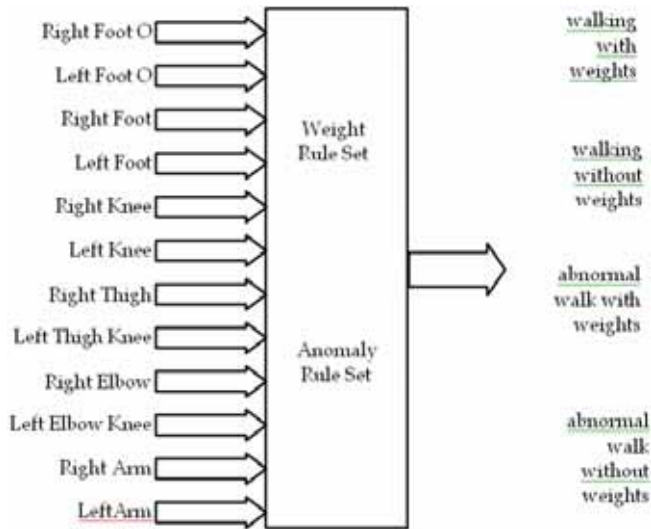


Fig. 16. Symbolic model proposal to identify: a weight induced gait anomaly; or an abnormally motion arising from some premeditated disguise.

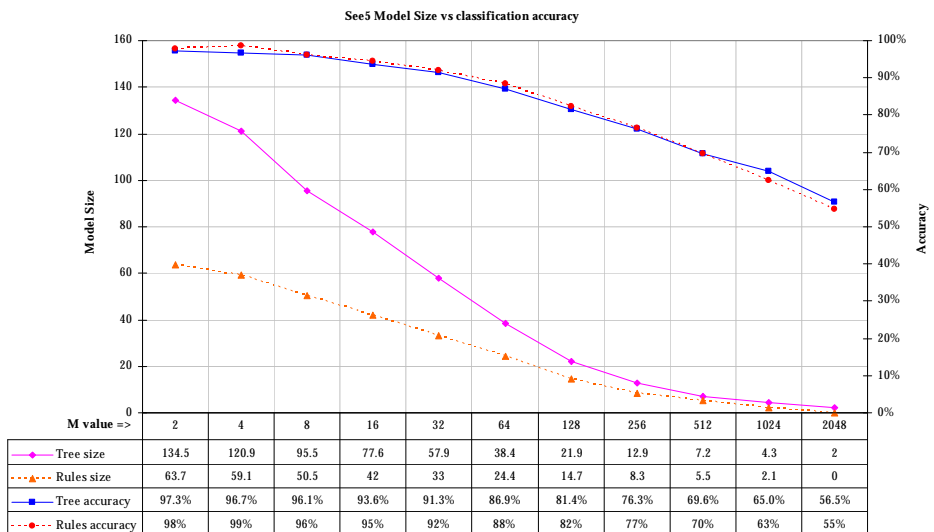


Fig. 17. Symbolic Model size and accuracy variations as measured by 10-fold cross validation for detecting weight induced gait anomalies.

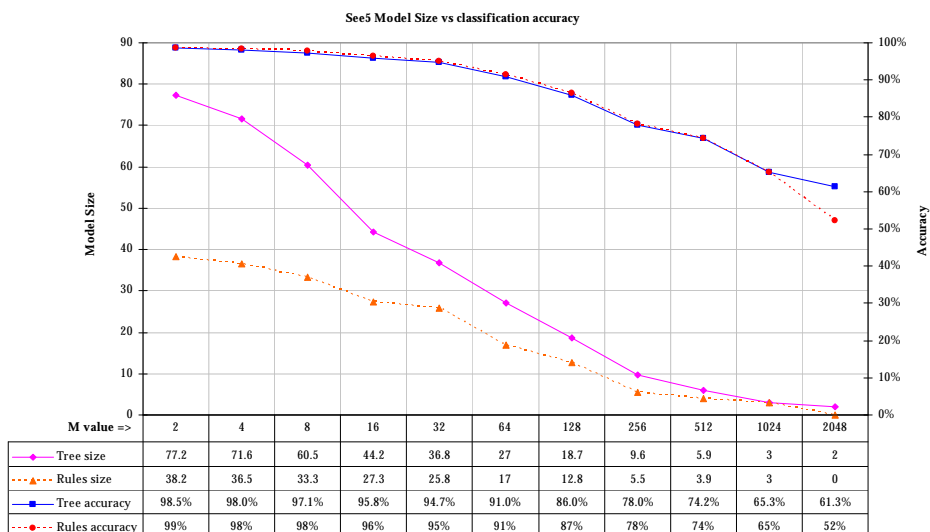


Fig. 18. Symbolic Model size and accuracy variations as measured by 10⁻fold cross validation for detecting disguised gait related motion behaviours.

Motion data for all 12 subtended joint angles was used in both rule sets in an attempt to classify disguised motion behaviours, and or, individuals that may be carrying an extraneous weight. As in the Sections, 3 and 5, a series of plausible models were firstly analyzed as illustrated in Figs. 17 and 18, before their appropriate formal forms were realized as illustrated in Figs 19 and 20.

The participants undertaking these motion experiments were 4 males and 5 females between 18 to 40 years of age. The primary aim of this study was to identify if a person is carrying an object and maybe concealing the object under his clothes based on a combination of the subtended angles at their feet, knees, thighs, arms, and elbows. In this process, again 12 features were extracted and using decision tree and rule set classifier models, more than 87% accuracy was achieved for detecting individuals carrying an extraneous weight, and an accuracy of at least 89% was also achieved in detecting unnatural (pretense) in gait motions.

6. Conclusion

The results from Section 4 and Section 5 clearly support all of the three objectives discussed at the end of Section 2. These being to firstly; develop a plausible model for dynamic finger printing of motion data between individuals. Secondly, investigate a model that could to also identify distinctions between various motion tasks, and finally to formulate a model to identify motion pretence, or acting, as well as normal and (physically induced) abnormal motion behaviours.

<p>Rule 1: (182, lift 2.2) Right Foot O <= 1.282073 Left Foot O <= 1.515643 Right Elbow > 3.057076 => WithWeights [0.995]</p>	<p>Rule 2: (479/6, lift 2.2) Right Foot O > 1.282073 Left Arm <= 0.1065039 => WithWeights [0.985]</p>	<p>Rule 3: (216/11, lift 2.1) Left Foot O <= 0.2262519 Left Foot <= 2.165244 Right Elbow > 2.876776 Left Elbow > 2.802877 => WithWeights [0.945]</p>	<p>Rule 4: (400/21, lift 2.1) Left Foot O <= 1.515643 Right Elbow > 3.027218 Left Elbow <= 3.027386 Right Arm <= 0.2105112 => WithWeights [0.945]</p>
<p>Rule 5: (175/26, lift 1.9) Left Elbow <= 2.605653 Left Arm <= 0.3550883 => WithWeights [0.847]</p>	<p>Rule 6: (475/74, lift 1.9) Right Foot O <= 0.05478672 => WithWeights [0.843]</p>	<p>Rule 7: (146/25, lift 1.8) Right Foot O > 0.05478672 Left Foot O > 1.515643 Left Elbow > 3.02983 Right Arm > 0.1455698 => WithWeights [0.824]</p>	<p>Rule 8: (260/49, lift 1.8) Right Foot O <= 1.282073 Right Elbow <= 2.399449 Left Elbow <= 2.802877 => WithWeights [0.809]</p>
<p>Rule 9: (528/139, lift 1.6) Left Foot O > 0.3129601 Left Foot O <= 1.515643 Right Knee <= 2.996378 Right Elbow <= 3.057076 Left Elbow <= 2.802877 Left Arm <= 0.3550883 => WithWeights [0.736]</p>	<p>Rule 10: (1092/294, lift 1.6) Right Foot O > 1.282073 Left Foot O <= 1.515643 Left Elbow <= 3.027386 Left Arm <= 0.3550883 => WithWeights [0.730]</p>	<p>Rule 11: (297/80, lift 1.6) Right Foot O <= 1.282073 Left Foot O > 1.422201 Left Foot O <= 1.515643 Left Elbow > 2.802877 Right Arm <= 0.3095023 => WithWeights [0.729]</p>	<p>Rule 12: (141, lift 1.8) Left Foot O > 1.515643 Left Elbow > 3.02983 Right Arm <= 0.1455698 => WithOutWeights [0.993]</p>
<p>Rule 13: (259/7, lift 1.8) Right Foot O <= 1.282073 Left Foot O > 1.422201 Right Arm > 0.3095023 => WithOutWeights [0.969]</p>	<p>Rule 14: (267/17, lift 1.7) Right Foot O > 1.282073 Right Elbow > 3.027218 Right Elbow <= 3.083863 Right Arm > 0.2105112 Left Arm > 0.1065039 => WithOutWeights [0.933]</p>	<p>Rule 15: (556/77, lift 1.6) Right Foot O > 0.05478672 Right Foot O <= 1.282073 Left Foot O > 1.422201 Right Elbow <= 3.057076 Left Arm > 0.1278452 => WithOutWeights [0.860]</p>	<p>Rule 16: (6287/2659, lift 1.1) Right Foot O > 0.05478672 => WithOutWeights [0.577]</p>
Default class: WithOutWeights			
Evaluation on training data (6762 motion frames)			
Decision Tree		Rules	
Size	Errors	No	Errors
26	926(13.7%)	16	819(12.1%)

Fig. 19. An example motion model for detecting subjects carrying an additional 5kg weight.

<p>Rule 1: (145, lift 2.1) Right Arm > 0.6420745 Left Arm <= 0.432619 => Abnormal [0.993]</p>	<p>Rule 2: (344/12, lift 2.0) Right Foot O <= 0.9345771 Left Foot O > 0.206346 Left Foot O <= 1.478905 Right Knee <= 3.012932 Right Arm > 0.2752825 Left Arm <= 0.432619 => Abnormal [0.962]</p>	<p>Rule 3: (298/12, lift 2.0) Right Elbow <= 2.419674 Left Arm > 0.432619 => Abnormal [0.957]</p>	<p>Rule 4: (246/12, lift 2.0) Right Elbow > 2.905675 Right Elbow <= 3.027221 Left Elbow <= 3.013905 Left Arm <= 0.432619 => Abnormal [0.948]</p>
<p>Rule 5: (395/59, lift 1.8) Right Foot O > 1.3187 Right Foot O <= 1.408431 Left Foot O > 0.206346 Right Elbow <= 3.027221 Right Arm > 0.2752825 => Abnormal [0.849]</p>	<p>Rule 6: (4467/2094, lift 1.1) Left Foot O > 0.206346 => Abnormal [0.531]</p>	<p>Rule 7: (144, lift 1.9) Left Thigh > 0.5184123 Left Arm <= 0.432619 => Normal [0.993]</p>	<p>Rule 8: (655/80, lift 1.7) Left Foot O <= 0.206346 => Normal [0.877]</p>
<p>Rule 9: (371/58, lift 1.6) Right Knee > 2.892806 Right Elbow > 2.419674 Left Arm > 0.432619 => Normal [0.842]</p>	<p>Rule 10: (3128/1117, lift 1.2) Right Arm > 0.1541919 Left Arm <= 0.432619 => Normal [0.643]</p>	<p>Rule 11: (202/73, lift 1.2) Left Foot > 1.776459 Right Elbow <= 3.032689 Right Arm <= 0.1541919 => Normal [0.637]</p>	
Default class: Normal			
Evaluation on training data (5122 motion frames)			
Decision Tree		Rules	
Size	Errors	No	Errors
19	606(11.8%)	11	554(10.8%)

Fig. 20. Motion model for detecting subjects manifesting disguised gait motion behaviours.

Motion capture data of human behaviour is necessarily by its nature highly complex and dynamic. Alternative approaches often seek to avoid where ever possible the so-called "curse of dimensionality" (Bellman, 1957) by developing methods to reduce this dimensionality to a tractable lower number of dimensions. Whilst these methods made succeed to various degrees they essentially smother or aggregate out fine detail and various nuances of motion behaviours.

In contrast, the application of symbolic machine learning is able to readily cope with the multidimensional properties of motion data, as evidenced by the example models developed in the previous sections. In effect, an appropriate (symbolic and inductive) DM algorithm will structure and or adjust numerous internal relationships between all of the input features that relate to and support the corresponding output, thereby avoiding, or significantly mitigating, the "curse of dimensionality".

However, whilst such models were often pruned significantly, which may also reduce the domain dimensionality the models address, this process always provides a transparent view of any resultant rules, patterns—often leading to new discovered knowledge. Thus the developer is able to readily critique and further explore various properties and consequences, often through a visualization process, that an individual element of existing or discovered knowledge poses in relation to any reduction in a models resolution (Asheibi, 2009).

Apart from this, motioning the induced symbolic patterns also provides a diagnostic ability guiding the often cyclic and interactive nature of applying machine learning in general. Previous other studies have validated this approach by combining together with unsupervised mixture modelling for gait recognition (Field et al., 2008)(Hesami et al., 2008).

The premise of this proposed work is that all humans have, by the stage of adolescence (or maturity) developed various stylistic signatures or patterns of motion behaviour that can be typically (uniquely) associated with an individual. These become (fundamentally) imprinted as patterns within the central nervous system (CNS) and govern everyday motions such as walking gaits, various gesticulations and other dynamic movements (trunk rotations) of an otherwise static body (Cuntoor et al., 2008). As is obvious, much of these motions can be unconsciously affected or modulated by underlying emotions (Dittrich et al., 1996) or by some conscious intent in order to conceal one's true identity.

In particular, the highly coupled nature of such complex data provides numerous opportunities for the discovery of actionable knowledge patterns, which in turn can be adapted for abnormal motion detection and tracking in two-dimensional (2D) video streams.

It is conjectured that the study of these dynamic (spatiotemporal) multidimensional manifestations will facilitate a new approach to anomaly pattern detection for human motions. By employing (symbolic) machine learning and other related data mining techniques, on a comprehensive range of motion capture trials, it is envisioned that a unique ontology ("structure or science of being", or taxonomy) of such manifest anomaly patterns could be formulated. This would provide a valuable resource structure of (manifest) pattern relationships. Amongst other future goals this research should address is that the motion ontology framework should be utilized to facilitate the derivation of various 2D images and silhouette maps to be subsequently utilized in video pattern analysis for anomaly identification and ultimately tracking.

7. References

- Argyle, M. (1988). *Bodily Communication*, Routledge publishing.
- Asheibi, A., Stirling, D. & Sutanto, D. (2009). Analysing Harmonic Monitoring Data using Supervised and Unsupervised Learning, Paper ID: TPWRD-00394-2007. *IEEE Trans. On Power Delivery*, vol 24, No 1, January.
- Bellman, R. (1957). *Dynamic Programming*, Princeton Univ. Press, Princeton, New Jersey.
- Braune, W., Fischer, O. (1904). *Der Gang Des Menschen/The Human Gait*, Springer, Berlin.
- Cook, D., and Swayne, D. F., (2007). *Interactive and Dynamic Graphics for Data Analysis With R and GGobi*, Springer, www.springerlink.com/content/978-0-387-71761-6.
- Donato, G.; Bartlett, M.S.; Hager, J.C.; Ekman, P. & Sejnowski, T.J. (1999), Classifying facial actions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 21, Issue 10, Oct. 1999 Page(s): 974 – 989.
- Ekman, P. (1999). Facial Expressions, In T. Dalgleish and T. Power (Eds.), *The Handbook of Cognition and Emotion*. Pp. 301-320. Sussex, U.K.: John Wiley & Sons, Ltd.
- Field, M., Stirling, D. Naghdy, F. & Pan, Z. (2008). Mixture Model Segmentation for Gait Recognition, *Symposium on Learning and Adaptive Behaviours for Robotic Systems, LAB-RS '08. ECSIS, 6-8 Aug. 2008* Page(s): 3 - 8.
- Hesami, A., Stirling, D., Naghdy, F. & Hill, H. (2008). Perception of Human Gestures through Observing Body Movements , *Proc. of the International Conference on Intelligent Sensors, Sensor Networks and Information Processing, ISSNIP2008, Sydney, Dec.*
- Johansson, G. (1973). Visual Perception of biological motion and a model for its analysis, *Percept. Psychophys.* 14 (1973) 210-211.
- McNeill, D. (2005). *Gesture and Thought*, The University of Chicago.
- Marey, E.J. (1904). *Centre Nationale d'Art Moderne, E-J Marey 1830/1904: La Photographie Du Movement* Paris Centre Georges Pompidou, Musee national d'art moderne, Paris.
- Mehrabian, A. (1971). *Silent messages*, Wadsworth publishing, Belmont, California.
- CMU Motion Capture Database. (2007). *CMU Graphics Lab Motion Capture Database*, <http://mocap.cs.cmu.edu/>.
- Cuntoor, N., Yegnanarayana B. & Chellappa, R. (2008) Activity Modeling Using Event Probability Sequences, *IEEE Transactions on Image Processing*, Volume: 17, Issue: 4, April 2008.
- Dittrich, W., Troscianko T, Lea S. & Morgan D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25,727-738.
- Roetenberg, D., Luinge, H. & Slycke, P. (2007). *Moven: Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors*, Xsen Technologies, December, 2007.
- RuleQuest Research (2007). *Data mining tools*, www.rulequest.com.
- Schmidt, K. & Cohn, J. (2002). Human facial expressions as adaptations: Evolutionary questions in facial expression. *Yearbook of Physical Anthropology*, 44:3-24.
- See5. (2002). *See5: An Informal Tutorial*, RuleQuest Research, www.rulequest.com.

Straker, D. (2008) . Changing Minds: in Detail, Published by Syque Press (October 2008).
Xsens Technologies (2009). Moven®/MVN®, wireless inertial motion capture
www.xsens.com.