1-1-2008

Perception of human gestures through observing body movements

Amir S. Hesami

University of Wollongong, ahesami@uow.edu.au

Fazel Naghdly

University of Wollongong, fazel@uow.edu.au

David A. Stirling

University of Wollongong, stirring@uow.edu.au

Harold C. Hill

University of Wollongong, harry@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>

Part of the [Physical Sciences and Mathematics Commons](#)

Recommended Citation

Hesami, Amir S.; Naghdly, Fazel; Stirling, David A.; and Hill, Harold C.: Perception of human gestures through observing body movements 2008, 97-102.
<https://ro.uow.edu.au/infopapers/903>

Perception of human gestures through observing body movements

Abstract

A new approach to modelling and classification of human gait is proposed. Body movements are obtained using a sensor suit that records inertial signals that are subsequently modelled on a humanoid frame with 23 degrees of freedom (DOF). Measured signals include position, velocity, acceleration, orientation, angular velocity and angular acceleration. Using a range of concurrent features extracted from the sensor signals, a system using induced symbolic classification models, such as decision trees or rule sets, has been used for classification of identity. It is anticipated that this approach will also enable the identification of a variety of gestures. The feasibility of generating the identified behaviours in a humanoid robot will be explored. The approach is described and the characteristics of the algorithm are presented. The results obtained so far are reported and conclusions drawn.

Keywords

movements, human, body, perception, observing, gestures

Disciplines

Physical Sciences and Mathematics

Publication Details

S. Hesami, F. Naghdy, D. A. Stirling & H. C. Hill, "Perception of human gestures through observing body movements," in International Conference on Intelligent Sensors, Sensor Networks and Information Processing, 2008, pp. 97-102.

Perception of Human Gestures through Observing Body Movements

Amir Hesami¹, Fazel Naghdy¹, David Stirling¹, Harold Hill²

¹*School of Electrical, Computer and Telecommunication Engineering, University of Wollongong
Wollongong NSW 2522 Australia,*

sah376@uow.edu.au, fazel_naghdy@uow.edu.au, stirring@uow.edu.au,

²*School of Psychology, University of Wollongong
Wollongong NSW 2522 Australia, harry@uow.edu.au*

Abstract

A new approach to modelling and classification of human gait is proposed. Body movements are obtained using a sensor suit that records inertial signals that are subsequently modelled on a humanoid frame with 23 degrees of freedom (DOF). Measured signals include position, velocity, acceleration, orientation, angular velocity and angular acceleration. Using a range of concurrent features extracted from the sensor signals, a system using induced symbolic classification models, such as decision trees or rule sets, has been used for classification of identity. It is anticipated that this approach will also enable the identification of a variety of gestures. The feasibility of generating the identified behaviours in a humanoid robot will be explored. The approach is described and the characteristics of the algorithm are presented. The results obtained so far are reported and conclusions drawn.

1. INTRODUCTION

Human gait has been studied scientifically for over a century. Some researchers such as Marey E.J [1] 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. Braune and Fischer [2] used a similar approach to study human motion attaching light rods instead of white tapes to the limbs of an individual. Johansson [3] presented MLDs (Moving Light Displays; method of using markers attached to joints or points of interests) in psychophysical experiments and showed that humans can recognize the gaits associated with different activities such as walking, stair climbing, dancing etc. One particular challenge with many possible applications is identification of an individual from his/her biometric information. Various methods have been developed in response to this need, including fingerprinting and iris identification. Such methods have proved to be partially reliable. Studies in psychology suggest that it should also be possible to identify an individual through non-verbal gestures and body movements including the way that they walk. We will refer to this as the Dynamic Finger Print of individual, a

combination of individual gait and body motion. The primary aim of this study is to verify the Dynamic Finger Print hypothesis and develop a methodology to objectively measure it.

The approach builds on the work conducted over the last one hundred years on the study of human walking style. Cutting et al. [4] introduced a biomechanical invariant for gait. Michael et al. [5] used different participants to show the possibility of recognition a particular person among a group. The previous study shows that pattern of body movement can be adequate for identification of an individual [6] and [7]. Mounting incandescent bulbs to the joints in order to be able to study the gait, called cyclography technique, was implemented by Bernstein (1967). Based on two experiments, Stevenage et al. [8] concluded that human visual system and brain is sophisticated enough to identify six participants based on their gait under normal and adverse viewing conditions. Jaraba et al. [9] used a feature called centre of the control points and neural networks for individual recognition. He used 13 control points head, arms, elbows, hands, upper legs, knees, and feet and created a 2D matrix which was fed into a SOM (Self Organizing Maps). Using control points coordinates would not be a strong feature in recognition purposes. Cattin et al [10] used three force plates to measure the ground reaction force and a CCD camera in order to identify a human being. He used PCA to reduce feature's dimension, although 0.3% Equal Error Rate (EER) was achieved, it is not clear what characteristics have been used as PCA outputs for recognition. Lee and Grimson [11] based on the canonical view of a walking person which is perpendicular to the direction of walk, fitted 7 ellipses to 7 regions of the body and used their locations, orientations, and aspect ratio as classification features. Ekinici [12] used distance vectors and principal component analysis (PCA) for dimension reduction and individual identification. Campbell and Bobick [13] used phase space constraints to recognize some ballet moves. They developed techniques based on space curves were developed assuming the availability of 3D Cartesian tracking data to represent movements of ballet dancers. The system learned and recognized nine movements. They developed techniques based on space curves were developed assuming the availability of 3D Cartesian tracking data to represent movements of ballet dancers. The system

learned and recognized nine movements. Bregler [14] used coherence blob hypothesis and HMM to estimate a ‘hidden variable’ for each pixel in the image to determine which blob it belongs to. A three level framework for recognition of activities was described in which probabilistic mixture models for segmentation from low – level cluttered vide sequences were initially used. Lie et al. [15] used magnitude and phase spectra of horizontal and vertical movements of ankle as features, and used AdaBoost classifier to classify them for gait recognition purposes. Sagawa et al. [16] matches people gait image sequences in the frequency domain, by creating a volume by piling up the image sequences of walking and using Fourier transform to extract the frequency characteristics of the volume, and finding the similarities of two volumes. The data used in the study is obtained using a sensor suit Moven® developed by Xsens [17]. The inertial signals are measured and inferred through an inverse kinematic model of the human body that includes 23 degrees of freedom (DOF) model across the human body. The signals include position, velocity, acceleration, orientation, angular velocity and angular acceleration. In this stage of study, the focus is on recognizing the characteristics of walk. A set of motion primitives is extracted from the captured data as the basis of the identification algorithm. The methodologies developed will be later applied to other human movements such as hand gestures, running and etc. The paper reports on the work carried out so far. The data obtained from various candidates is analyzed to extract candidate motion primitives. The aim is to develop an algorithm for motion analysis which will be added to an image processing technique to extract data in real application using cameras. The paper is organized as follows. In Section 2 data acquisition is described. Section 3 covers the feature extraction and evaluation. In Section 4 experimental results using the models are presented, and finally the conclusion of the paper is presented in Section 5.

2. DATA ACQUISITION

Data is being acquired using an inertial movement suit, Moven®, 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 Figure 1 and Figure 2.

Fig. 1: Anatomical Landmarks for the Moven® suit.

Fig. 2: Body Segments defined in the Moven® suit.

In capturing human body motion, no external emitters or cameras are required. As explained by [18], mechanical trackers use goniometers which are worn by the user to provide joint angle data to kinematics algorithms for determining body posture. Full 6 DOF tracking of the body segments is achieved using connected inertial sensor modules. The body segment orientation, position, velocity, acceleration, angular velocity and angular acceleration are also estimated using this method. The kinematics data is saved in an MVNX file format which is 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 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. FEATURE EXTRACTION

The feature extraction section is the most important part 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 path of the individual's walking. As the most important organs in walking are legs, foots, arms and elbows, the extracted features are related to these body segment's movements. Features are extracted in a gait cycle for each individual. The gait cycle is a complete stride with both legs stepping, starts with right leg as shown in Fig.3. Recording period having a participant wearing the suit is shown in Fig. 4.

Fig. 3: Sample Gait Cycle.

Fig. 4: Recording the Body Motion.

The extracted features are listed below and shown as following:

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

Fig. 5: Foot Orientation Angle and Foot Angle.

Fig. 6: Knee Angle and Thigh Angle.

Fig. 7: Elbow Angle and Arm Angle.

Totally 12 features per individual were extracted. The extracted angles were used in radians.

4. EXPERIMENTAL RESULTS

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.

In order to examine the Dynamic Finger Print hypothesis, ten individuals wearing the Moven suit, undertook four

repetitions of a simple walking task. From these tasks, a range of features, across the individuals were collected and recorded as features for an identification trial. For this trial, the goal was to clearly identify an individual based on purely a combination of subtended joint angles. In addressing this recognition challenge, the rule induction system called See5 was used [19]. This system has induced symbolic classification models, such as decision trees or rule sets, based on a range of several concurrent features (attributes). The final decision trees and rule sets were created through adjustment of the various pruning options, such as setting the pruning confidence factor (CF) to 5%, and modification of the minimum number of objects to 2. 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 CF option affects the way that error rates are estimated and hence the severity of pruning. Values smaller than the default value which is 25% cause more of the initial tree to be pruned, while larger values result in less pruning. The minimum cases option specifies the minimum number of cases that will be maintained at the decision nodes, and essentially constrains the degree to which the induced model can fit the data. In order to get more reliable estimate of predictive accuracy f – fold cross validation is used. The cases in data file are divided into f blocks of roughly the same size and class distribution. For each block in turn, a classifier is constructed from the cases in the remaining blocks and tested on the cases in the hold – out block. In this way, each

case is used just once as a test case. The error rate of a classifier produced from all the cases is estimated as the ratio of the total number of errors on the hold – out cases to the total number of cases. Number of folds has been set to 10. Once a suitable classifier performance level has been revealed using cross validation, the resultant model is generated. The options and attributes are shown in Fig. 8.

```
Options:
  Rule-based classifiers
  Pruning confidence level 5%
  Test requires 2 branches with >= 25 cases

Class specified by attribute 'participant'

Read 3837 cases (13 attributes) from P1-10.data
```

Fig. 8: Classifier Attributes.

Summary of the error rates and number of rules per fold having the above attributes is shown in Fig.9. As shown above there are ten classes p1 – p10 which each class represents a participant. Participants undertaking the experiments were 5 males and 5 females between 18 to 40 years of age. According to Fig. 9, the average error rate achieved is 3% and number of rules is 23. Rules are shown in Fig.10.

Evaluation on training data (3837 cases):

Rules											
No	Errors										
23	117(3.0%)	<<									
(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	<-classified as	
400					6					(a):	class p1
	355				1					(b):	class p2
	2	323			19	9				(c):	class p3
			350							(d):	class p4
13				395						(e):	class p5
	4	3			432					(f):	class p6
6						351				(g):	class p7
							413	18		(h):	class p8
							23	291	13	(i):	class p9
									410	(j):	class p10

Time: 0.2 secs

Fig. 9: Classification errors.

Rule 1: (391/6, lift 9.3) Left Foot O > 1.124812 Right Thigh <= 0.4369857 Left Elbow <= 2.795459 Right Arm > 0.3421563 → class p1 [0.982]	Rule 2: (28/13, lift 5.0) Right Foot O > 1.260007 Left Foot O <= 1.124812 Right Elbow > 2.640656 Left Elbow <= 2.913522 → class p1 [0.533]	Rule 3: (268, lift 10.7) Left Foot O > 1.124812 Right Elbow > 2.901795 Left Elbow > 2.989989 → class p2 [0.996]	Rule 4: (38, lift 10.5) Right Foot O > 1.260007 Left Foot O <= 1.124812 Left Elbow > 2.913522 → class p2 [0.975]
Rule 5: (71/6, lift 9.7) Left Foot O > 1.124812 Right Foot <= 2.100667 Right Knee <= 2.866177 Right Elbow > 2.74591 Left Elbow > 2.795459 Left Arm > 0.1069464 Left Arm <= 0.1387282 → class p2 [0.904]	Rule 6: (26/11, lift 6.2) Right Foot O > 1.310526 Right Knee > 2.866177 Right Elbow <= 2.901795 Left Elbow > 2.795459 → class p2 [0.571]	Rule 7: (142, lift 10.8) Left Foot O > 1.124812 Right Knee <= 2.866177 Right Elbow <= 2.74591 Left Elbow > 2.795459 → class p3 [0.993]	Rule 8: (122, lift 10.8) Left Foot O > 1.124812 Right Thigh > 0.4369857 Right Elbow <= 2.901795 → class p3 [0.992]
Rule 9: (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 10: (31, lift 10.5) Left Foot O <= 1.124812 Right Thigh > 0.3605901 Right Elbow <= 2.640656 → class p3 [0.970]	Rule 11: (350, lift 10.9) Left Foot O <= 1.124812 Left Arm > 0.4538144 → class p4 [0.997]	Rule 12: (1810/1415, lift 2.1) Left Foot O > 1.124812 → class p5 [0.219]
Rule 13: (34, lift 8.5) Left Foot O > 1.124812 Right Thigh <= 0.4369857 Right Elbow <= 2.901795 Left Elbow <= 2.795459 Right Arm <= 0.3421563 → class p6 [0.972]	Rule 14: (166/4, lift 8.5) Left Foot O > 1.124812 Right Elbow > 2.74591 Right Elbow <= 2.901795 Left Arm <= 0.1069464 → class p6 [0.970]	Rule 14: (166/4, lift 8.5) Left Foot O > 1.124812 Right Elbow > 2.74591 Right Elbow <= 2.901795 Left Arm <= 0.1069464 → class p6 [0.970]	Rule 16: (666/328, lift 4.4) Left Foot O > 1.124812 Right Elbow <= 2.901795 Left Elbow > 2.795459 → class p6 [0.507]
Rule 17: (360/9, lift 10.5) Left Foot O <= 1.124812 Right Thigh <= 0.3605901 Right Elbow <= 2.640656 → class p7 [0.972]	Rule 18: (28, lift 8.6) Right Foot O <= 1.00804 Left Foot O > 0.1827743 Left Foot O <= 1.124812 Right Knee <= 2.993618 Right Elbow <= 2.838296 Left Elbow > 2.851741 Left Arm <= 0.4538144 → class p8 [0.967]	Rule 19: (28, lift 8.6) Left Foot O > 1.124812 Right Elbow > 2.901795 Left Elbow > 2.918272 Left Elbow <= 2.989989 → class p8 [0.967]	Rule 20: (380/23, lift 8.3) Right Foot O <= 1.00804 Left Foot O <= 1.124812 Left Knee > 2.257854 Right Elbow > 2.838296 Left Elbow > 2.851741 Left Arm <= 0.4538144 → class p8 [0.937]
Rule 21: (266, lift 11.7) Right Foot O > 0.1117351 Right Foot O <= 1.00804 Left Foot O > 0.1827743 Left Foot O <= 1.124812 Right Elbow > 2.640656 Left Elbow <= 2.851741 → class p9 [0.996]	Rule 22: (64/18, lift 8.4) Right Foot O <= 1.00804 Left Foot O <= 1.124812 Left Knee <= 2.257854 Right Elbow > 2.838296 Left Arm <= 0.4538144 → class p9 [0.712]	Rule 23: (2027/1617, lift 1.9) Left Foot O <= 1.124812 → class p10 [0.203]	Default class: p6

Fig. 10: Rules.

Each rule consists of a rule number which is used only to identify the rule, and statistics (n, lift x) or (n/m, lift x) that summarize the performance of the rule. N is the number of training cases covered by the rule and m, if it appears, shows how many of them do not belong to the class predicted by the rule. The rules accuracy is estimated by the Laplace ratio $(n-m+1)/(n+2)$. The lift x is the result of dividing the rule's estimated accuracy by the relative frequency of the predicted class in the training set. Each rule has one or more conditions that must all be satisfied if the rule 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 shown in brackets. All extracted features have been fed into the classifier whereas all of them have not been used in the rules. Number of times that each feature has been referred in the rules, and so the importance of each feature in classifying a person is shown in table 1.

TABLE 1: USAGE OF FEATURES

Feature	Number of times being referred	Percentage of usage/all features
Left Foot O	24	26.09%
Right Elbow	20	21.74%
Left Elbow	15	16.30%
Left Arm	9	9.78%
Right Foot O	8	8.70%
Right Knee	5	5.43%
Right Thigh	5	5.43%
Left Knee	2	2.17%
Right Foot	2	2.17%
Right Arm	2	2.17%
Left Foot	0	0.0%
Left Thigh	0	0.0%

According to Table.1 features Left foot and Left Thigh angle has not been used in classifier at all, and the two most important features are Left foot orientation angle and right elbow.

5. CONCLUSIONS

The primary aim of this study was to derive and test a Dynamic Finger Print for recognizing identity from human gait. The approach taken was to identify a person based on a combination of subtended angles at feet, knees, thighs, arms, and elbows. In this process 12 features were extracted. Using a decision tree, and converting trees into collection of rules called *rulesets*, 97% classifier accuracy was achieved. The targets were 5 males and 5 females between 18 and 40 years old, suggesting that the results should be reasonably generalizable. The same extracted features could also be used for gender classification, or even classification of different actions. In order to be able to use the described method in a real application, an image processing and computer vision front-end for data acquisition should be added to the system. The goal here was only to develop an algorithm and recognize appropriate features for identification purposes.

REFERENCES

- [1] Centre Nationale d'Art Moderne, E-J Marey 1830/1904: La Photographie Du Movement Paris Centre Georges Pompidou, Musee national d'art moderne, Paris.
- [2] W. Braune, O. Fischer, Der Gang Des Menschen/ The Human Gait, Springer, Berlin 1904.
- [3] G.Johansson, "Visual Perception of biological motion and a model for its analysis", Percept. Psychophys. 14(1973) 210-211.
- [4] James E.Cutting, Dennis R.Proffitt, and Lynn T.Kozlowski, A Biomechanical Invariant for Gait

- Perception. *Journal of Experimental Psychology: Human Perception and Performance* 1978, Vol.4, No. 3, 357 – 372.
- [5] Michael J. Richardson and Lucy Johnston, Person recognition from dynamic events: The kinematic specification of individual identity in walking style. *Journal of Nonverbal behavior*, 29(1), Spring 2005 @ Springer Science + Business Media Inc.
  - [6] Diane S. Berry, Stephen J. Misovich, Methodological Approaches to the study of social event perception. *Society for personality and social psychology*, 1994..
  - [7] Reuben M. Baron, Stephen J. Misovich, Dispositional Knowing From and ecological perspective, *Society for personality and social Psychology*, 1993..
  - [8] Sarah V. Stevenage, Mark S. Nixon and Kate Vince, Visual Analysis of gait as a cue to identity, *Applied cognitive psychology*, 13: 513 – 526 (1999).
  - [9] J. Elias Herrero – Jaraba, Carlos Orrite – Urnueka, David Buldain, and Armando Roy – Yarza, *Human Recognition by gait analysis using neural networks*, Springer – Verlag Berlin Heidelberg 2002.
  - [10] Philipp C. Cattin, Daniel Zlantik, and Ruedi Borer, Biometric System using Human Gait, *Institute of Robotics, Swiss Federal Institute of Technology (ETH)*.
  - [11] L. Lee and W. E. L. Grimson, Gait analysis for recognition and classification. *IEEE Conference on Face and Gesture Recognition*, 155 – 161, 2002.
  - [12] Murat Ekinci, A New approach for human identification using gait recognition, *Springer – Verlag Berlin Heidelberg* 2006.
  - [13] Lee Campbell and Aaron Bobick, Recognition of Human Body Motion Using Phase Space Constraints, *MIT Media Laboratory Perceptual Computing Section Technical Report No. 309*.
  - [14] Christoph Bregler, Learning and Recognizing Human Dynamics in Video Sequences, *IEEE Conf, Computer vision and Pattern recognition*, San Juan, Puerto Rico, June 1997.
  - [15] Agus Santoso Lie, Shuichi Enokida, Tomohito Wada, and Toshiaki Ejima, Magnitude and Phase Spectra of Foot Motion for Gait Recognition, A. Gagalowicz and W. Philips, *CAIP 2005, LNCS 3691*, pp. 390 – 397, 2006.
  - [16] Ryusuke Sagawa, Yasushi Makiyama, Tomio Echigo, and Yasushi Yagi, Matching Gait Image Sequences in the Frequency Domain for Tracking People at a Distance, *ACCV 2006, LNCS 3852*, pp. 141 – 150, 2006.
  - [17] Xsens Technologies B.V., *Moven®*, wireless inertial motion capture, www.xsens.com.
  - [18] Daniel Roetenberg, Henk Luinge, and Per Slycke. *Moven: “Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors”*, Xsens Technologies, December, 2007.
  - [19] RuleQuest Research: data mining tools, www.rulequest.com.