

Empirical Modelling of Human Gaits for Bipedal Robots

Matthew Field, David Stirling, Fazel Naghdy, Zengxi Pan

University of Wollongong, Australia

{mf91, stirring, fazel, zengxi}@uow.edu.au

Abstract

Modelling of human motion through a discrete sequence of motion primitives, retaining elements of skillful or unique motion of an individual is addressed. Using wireless inertial motion sensors, a skeletal model of the fluid human gait was gathered. The posture of the human model is described by nine sets of euler angles for each sample. An intrinsic classification algorithm known as Minimum Message Length encoding (MML) is deployed to segment the stream of data and subsequently formulate certain Gaussian Mixture Models (GMM) that contain a plausible range of motion primitives. The removal of certain less seemingly important modes has been shown to significantly affect the fluidity of a gait cycle. The approach is described and the outcomes so far are provided.

1 Introduction

Human motion is a ubiquitous element of life that has attracted widespread research, from the analysis of sporting techniques to designing humanoid robots. Gesture and posture recognition has been an increasingly popular research field due to enhanced image processing, camera resolution and new motion capture sensors. Motor skill encompasses the ability and quality in achieving a particular motion task. As such, motor skill can be divided into two parts, gross motor skill and fine motor skill. Gross motor skill refers to major translations of a body that define the motion and usually follow a pattern. Fine motor skill involves precise movements that entail accurate positioning in a delicate task.

In this work, the focus is on quantifying human motor skill by capturing motions with an array of inertial sensors and articulating the motions as a set of patterns. This will ultimately be directed towards a motion model and control scheme for a robot, which can be programmed through demonstration from motion capture information.

1.1 Human Motion Capture Data

There are a range of motion sensing technologies available for data capture with different research areas concentrating on particular systems depending on suitability. Optical motion capture systems employ either infrared emitting diodes (IRED) or reflective markers to determine dynamic positions of fixed points. They are popular in biomechanics due to the fixed point accuracy and reliability of results [Donati *et al.*, 2007]. However, the markers are prone to occlusion and the subject is confined to a small camera range.

Another popular approach is optical capture with a digital camera, which is very common in gesture recognition and gross motion tracking [Saboune *et al.*, 2005]. This method is very practical and increases the freedom of movement but is also affected by lighting conditions and is not suitable for monitoring complex behaviour.

Inertial sensors are becoming more practical due to the miniaturization to micro electromechanical systems (MEMS) allowing for easier mounting on the body. Solid state accelerometers have had some use in biomechanics experiments for estimating peak forces in jumping [Bisseling *et al.*, 2006], analysing gait patterns [Zijlstra *et al.*, 2003] and walking stability in the aged [Menz *et al.*, 2003]. MEMS sensors are more flexible and are suitable in arrays for more complex motion capture [Zhou *et al.*, 2006] providing accurate orientation and accelerometer data.

1.2 Perception of Motion

The perception of human motion and its structured representation requires a method that essentially compresses complex multidimensional data into a symbolic form. In order to describe the complex data stream it is divided into a set of motion primitives which can be inferred from inherent patterns in the motion data. Several approaches have been explored to achieve this symbolic representation.

Kahol [Kahol *et al.*, 2004] used an algorithm called Hierarchical Activity Segmentation to implement a dy-

dynamic hierarchical structure that describes the human body segments. The structure was characterized by demonstrator’s profiles from a naive Bayesian classifier. The approach successfully recognised new gestures. Hidden Markov Models (HMM) have also been extensively employed in gestural recognition with Green [Green *et al.*, 2003] using particle filtering to fine tune classification with blurred images. Saboune [Saboune *et al.*, 2005] employed a modified particle filtering algorithm named Interval Particle Filtering to analyse gait disorders from markerless motion capture. Inamura [Inamura *et al.*, 2004] also proposed a layered architecture encouraging symbolic representations of motion primitives with HMM.

In related work Sun [Sun *et al.*, 2006] applied a Minimum Message Length (MML) encoding algorithm to inertial data gathered from different hand movements. This formulated a Gaussian Mixture Model for each segment of the complex data, which defined each mode or motion primitive in a motion sequence. By training the model with pre-defined motions and refining the cluster sequences, further motions of the same style could be recognised with high accuracy.

1.3 Motion Transformation

The programming of humanoid robots with human motion capture data (HMCD) is a particularly challenging extension from motion recognition. Despite the difficulties due to the high degrees of freedom (DOF) and the discrepancies in structure there has been promising results. Lopez [Lopez *et al.*, 2006] showed through simulation that HMCD provided a much smoother robot gait than the mathematical-foot trajectory calculation. Naksuk [Naksuk *et al.*, 2005] was able to successfully transfer a walking motion to a stable simulation model by compensating the angular momentum at the center of mass (CoM). In a different direction, the motion data from one inertial sensor has been used as a feedback mechanism controlling galvanic vestibular stimulation (GVS). In this work a training mechanism was proposed to recreate the conditions of locomotor dysfunction found in astronauts after they arrive back to earth.

In the next section, the equipment and data preparation will be described. In Section 3 the basic concept behind the approach will be introduced. In Section 4, the different methods and ideas employed in the data analysis will be outlined. Section 5 will illustrate the results from segmentation and skill reduction. Finally, section 6 will discuss the implications of the results and future directions.

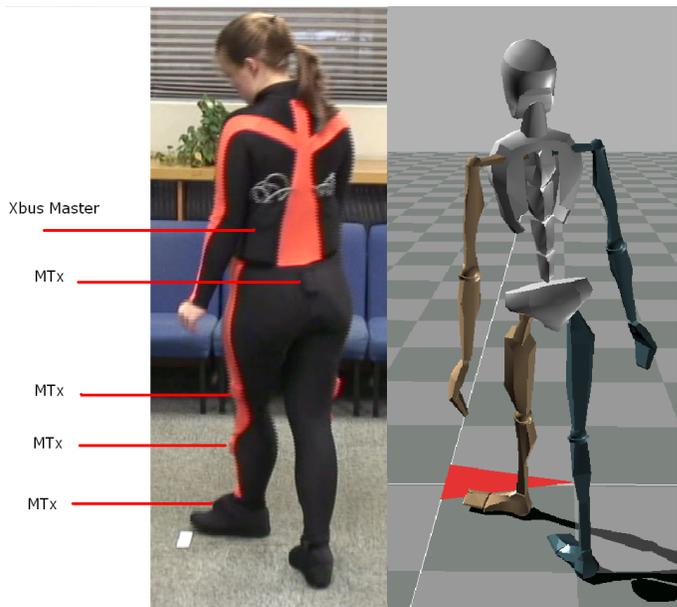


Figure 1: Motion capture model using Moven suit alongside the software output model.

2 Experimental Rig

The challenging goal of this work is to model human behaviour from an array of inertial sensors, resulting in an extensive motion database. The expression of locomotive behaviour in terms of motion primitives provides a symbolic data compression approach to recognize regular sequences in multiple data streams.

The sensors used in this work were a set of seven inertial measurement units, the MTx, which are embedded in a spandex body suit named Moven from Xsens Technologies as illustrated in Figure 1. Each sensor itself contains nine MEMS sensors to provide accurate 3D orientation to 2° RMS [Cutti *et al.*, 2006] and kinematics data from 3D acceleration and 3D rate of turn. With this array of orientation sensors mounted on fixed points on the body, an accurate kinematic model of the human body can be resolved. Although this equipment can provide accurate 3D skeletal models in real-time, unlike optical motion capture, it does require post-processing in the event of unknown contact points.

In this experiment only seven sensors for tracking the lower body were used. The system works via a Bluetooth wireless link allowing complete freedom of movement. The motion capture data is supplied in quaternion form for orientation and cartesian coordinates for position for each body segment. Other available kinematic data includes acceleration, velocity and angular acceleration and velocity all sampled with the position and orientation at 60Hz. For our data analysis the orientation was converted from quaternion form to Euler angles

of Roll, Pitch and Yaw because this form provides more immediate visual meaning and reduces the data dimensions from 36 variables to 27.

3 Approach

The underlying premise of this study is that human behaviour can be expressed explicitly as sets of motion primitives, which in different configurations display a skill. This notion of skill identifies the expertise or technique of a motion that is not readily modeled with classical mechanics. Mathematical models are commonly employed in analysis of postural stability and gait although these models often either localize the analysis or sacrifice complexity for a more generalized model. However, if a locomotive behaviour can be expressed in the segmented motion primitive form with sufficient accuracy it can be assumed that all human motion skills, from gross body movements to fine idiosyncratic behaviour can be captured.

The approach in this study is to segment the orientation euler angles of each major body part to describe an individual pose or motion primitive. Conceptually the segmented poses should vary in time throughout a particular motion and form a proportion of the complete behaviour. More abundant poses translate to the major movements of the limbs whereas the less abundant poses tend to be slight variations in these movements which translate to a motion style or technique. Furthermore, it is assumed that the removal of variant numbers of the least abundant poses results in a reduction in skill and fluidity of movement and further reduction results in a disjointed unnatural movement.

In order to judge the quality of a gait cycle, some integrated modeling technique is required to deal with the complex nature of the human form. The problem of analyzing dynamic human behaviour is the multidimensional and high degrees of freedom (DOF) of the human body. This complexity increases the mathematical computation exponentially with every additional DOF. However, in gait and posture research, there are a few integrated physical parameters which have been identified as crucial to maintaining balance and ensure a stable walking motion. The center of mass and zero moment point trajectories are often used to predict stability. In this study, these parameters are used to indicate the fluidity of gait for different skill levels.

4 Methodology

This section outlines the background theory relevant to the results of data analysis.

4.1 Machine Learning

An unsupervised classification method is required for partitioning unlabelled, complex data so that the model

fits the observations. Without a pre-defined model an appropriate number of motion modes cannot be determined. Through Gaussian Mixture Modelling and inductive Bayesian probability based segmentation, this problem can be overcome. The approach deployed in this work is Minimum Message Length encoding (MML).

Minimum Message Length encoding is, as the name implies, based on evaluating a models ability to compress a message expressing the data [Baxter, 1996]. Quantitatively, MML embodies 'Occam's razor' where the most complex model is automatically discarded despite the 'goodness of fit'. The algorithm first postulates a model of the data and evaluates it based on the estimated code length to express the model and the accuracy of the model. From information theory the minimum code length is $-\log P(E)$ where E is the evidence and from Bayes' theorem $P(E|H) \cdot P(H) = P(H \cap E)$ where H is a probable hypothesis. Maximising the probability $P(H \cap E)$ - that the evidence supports the hypothesis - is equivalent to minimising $-\log(P(E|H)) - \log(P(H))$ and thereby minimising the message length. The resultant modes are a trade off between accuracy and message compression.

4.2 Stability Criteria

Stability criteria for walking systems is a popular field of research with a wide range of techniques, however, they commonly revolve around a few key physical parameters. The center of mass (CoM), zero moment point (ZMP) and center of pressure (CoP). In this analysis, the CoM and ZMP are used the parameters are observed over a ground contact field.

Center of Mass

Bipedal stability criteria almost always involves tracking the center of mass. For the system to remain in a static position the projection of the center of mass must lie above the convex hull of the ground contact points. The human segment mass parameters are estimated using data from biomechanics literature [de Leva, 1996], which lists each weight as a percentage of the total and each CoM position as a percentage of total segment length from the proximal point. The result is then projected onto the horizontal plane.

$$x_G = \sum_{i=1}^n \frac{m_i x_i}{m} \quad (1)$$

$$y_G = \sum_{i=1}^n \frac{m_i y_i}{m} \quad (2)$$

where m_i is the mass and x_i and y_i the position of the i_{th} segment when there are n segments.

Tipping Moment

The tipping moment is a summation of all moments of the horizontal forces acting on the body which are centered at one point on the ground, commonly referred to as the zero moment point (ZMP). This parameter is extensively used in bipedal dynamic stability criteria, which state that the ZMP must lie within the convex hull of all ground contact points. It can be calculated with the following equations,

$$x_{zmp} = \frac{\sum_{i=1}^n m_i(\ddot{z}_i + g)x_i - \sum_{i=1}^n m_i\ddot{x}_i z_i}{\sum_{i=1}^n m_i(\ddot{z}_i + g)} \quad (3)$$

$$y_{zmp} = \frac{\sum_{i=1}^n m_i(\ddot{z}_i + g)y_i - \sum_{i=1}^n m_i\ddot{y}_i z_i}{\sum_{i=1}^n m_i(\ddot{z}_i + g)} \quad (4)$$

where x_i and \ddot{x}_i denotes position and acceleration of the i_{th} segment along the x-axis respectively.

Feet Projection

The purpose of the ground support calculation is to determine whether the body is balanced based on the condition that the CoM must lie over the support base. Assuming that a walking gait is described, the ground support can only come from within the convex hull of the feet positions. Bearing this in mind, the feet are tracked by an ellipse with the midpoints of the feet as the foci as illustrated in Figure 2. With this method the elliptical ground support can be transformed to a unit circle allowing a constant parameter space to compare the center of mass and tipping moment. It also provides a convenient form with which to analyse gait patterns. It, however, does not represent true ground support, rather (in a gait) the support for one foot and the intentional support for the other.

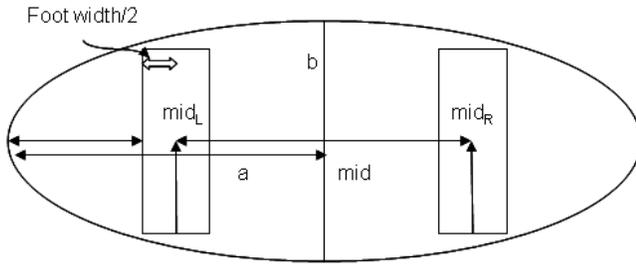


Figure 2: Derivation of ground contact ellipse.

The semi-major and minor axes can be found from

$$c = \frac{\sqrt{mid_L^2 - mid_R^2}}{2} \quad (5)$$

$$a = \frac{\sqrt{mid_L^2 - mid_R^2}}{2} + \frac{\sqrt{x_{lf}^2 - x_{rf}^2}}{2} + \frac{x_{fw}}{2} \quad (6)$$

$$b = a\sqrt{1 - \frac{c^2}{a^2}} \quad (7)$$

which gives

$$x = \frac{x_G - mid_x}{a} \quad (8)$$

$$y = \frac{y_G - mid_y}{b} \quad (9)$$

where x and y is the center of mass position transformed to a unit circle space.

5 Validation

In the initial experiments with the motion capture, a volunteer was fitted with the equipment and recorded a collection of gait cycles for analysis. The objective was to gather data that could be segmented into posture sets and evaluated at different segment resolutions. With different numbers of resulting modes or clusters a technique for measuring the gait performance was investigated. The volunteer performed several gait cycles in a consistent forward trajectory.

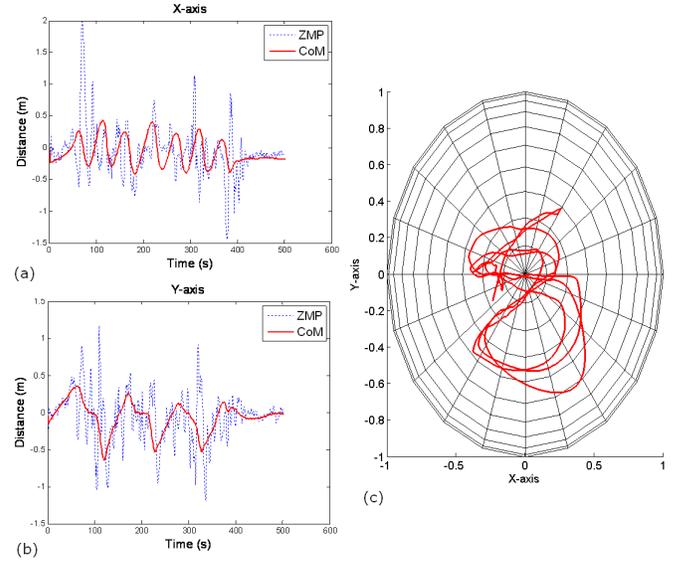


Figure 3: Raw data plots: (a)x-axis movement of CoM and TP, (b) y-axis movement of CoM and TP, (c) unit circle CoM plot.

5.1 Motion Capture

Using the orientation and acceleration output the CoM and TP were calculated with respect to the ground support as mentioned earlier. Figure 3(c) shows the CoM path on the unit circle and the Figures 3(a) and 3(b) show the CoM plotted against the TP for each axis.

5.2 Segmentation

The orientation data, including 27 euler angles, during this sequence was processed by the MML algorithm resulting in a significant number of modes or clusters. Each mode therefore represents the static posture at every time instant in the sequence. Due to the complexity of the human gait and the intention of quantifying subtle skills, the model retained all possible idiosyncratic modes. The algorithm found 49 unique poses to articulate the subject’s gait. Figure 4 indicates the proportion of each mode or the subject’s poses throughout the motion capture recording. Since each mode can be described by CoM position the sequence was recalculated at this reduced resolution and shown along with the TP for each axis in Figure 5.

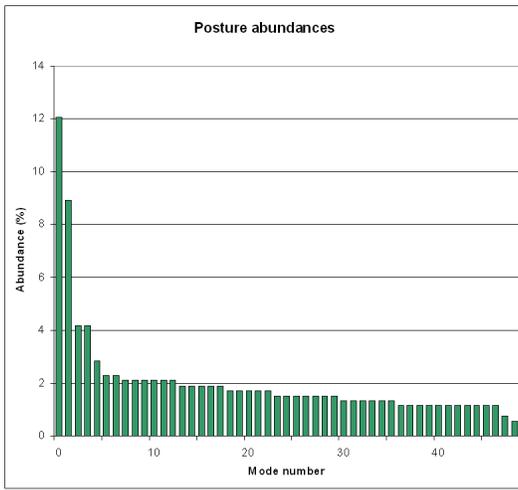


Figure 4: Mode abundances for the full motion sequence.

5.3 Skill Reduction

It is proposed that a collection of the least abundant clusters form the fine skills of the subject whereas the majority of the abundant clusters account for the gross motor skill. To remove one of the motion primitives the instances classified as that mode were replaced with closest temporal mode in an attempt to retain a more natural mode transition.

Figure 6 illustrates the reduction in skill level with the abundance graph nullifying the rejected motion primitives. Under the condition shown in the this graph with 30 modes left remaining, the CoM trajectory was recalculated and plotted with the ZMP. The CoM plot shows not only the reduced resolution of the motion sequence but also that the mode reduction had eliminated a brief but important part of the gait resulting in a substantially less fluid state transition.

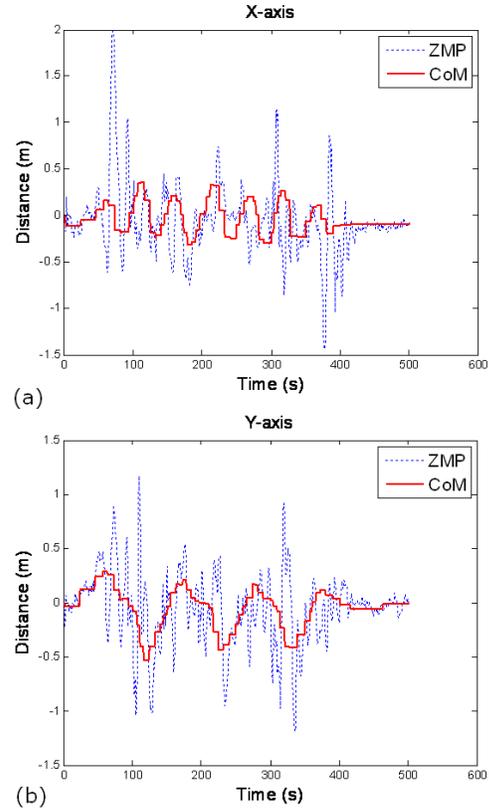


Figure 5: Result from segmented data (a) x-axis plot of CoM and TP, (b) y-axis plot of CoM and TP

6 Conclusions

This work illustrated that the human behaviour can be segmented into a sequence of motion primitives which directly correspond to a set of static poses. This provides a starting point for developing a structured hierarchy of motion primitives for synthesizing new behaviours. However, in order to generate dynamic sequences other influences such as inertia must also be accommodated within the motion primitives. This research is building towards a heuristic control scheme for a humanoid robot based on human demonstration, which also requires a detailed study of the structural transformation between human and humanoid.

The ideas provided in this paper are of relevance to the treatment of diseases such as Parkinson’s disease and the development of balance prostheses or aids for elderly patients. People with Parkinson’s disease may be able to control the gross motor skills but often gradually lose the influence of fine motor skills for precise movement. One could also use these techniques to diagnose locomotor disorders, analysing the effects on behaviour after rehabilitative treatment or for developing intelligent prosthe-

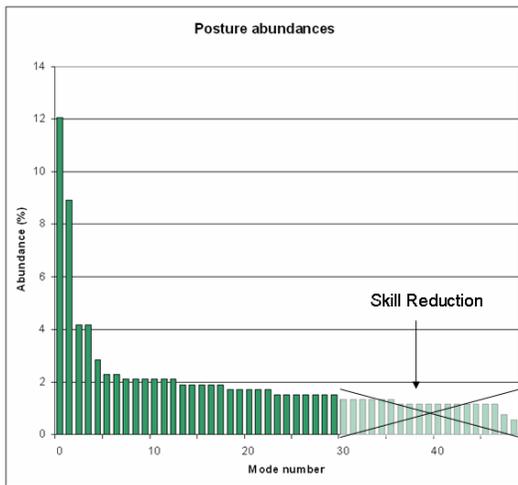


Figure 6: Reduced number of modes.

ses and potentially for Functional Electrical Stimulation (FES) feedback [Lauer *et al.*, 2005]. This is also relevant to sports science and the perception of sporting skills, which is not a parameter that is easily measured with traditional techniques.

Future work in this area will involve investigating the dynamic motion primitives by incorporating acceleration data or combined force vectors, and investigating a performance measure which responds to the reduction in fine motor skills.

Acknowledgments

Thanks to my supervisors and to the various motion capture volunteers.

References

- [Baxter, 1996] R. A. Baxter. Minimum Message Length Inference: Theory and Applications. *PhD thesis* Monash University, Australia, December, 1996.
- [Bisseling *et al.*, 2006] R. W. Bisseling, A. L. Hof. Handling of impact forces in inverse dynamics. *Journal of Biomechanics*, Volume 39, Issue 13, 2006, Pages 2438-2444.
- [Cutti *et al.*, 2006] Adrea Giovanni Cutti, Andrea Giovanardi, Laura Rocchi, Angelo Davalli. A simple test to assess the static and dynamic accuracy of an inertial sensors system for human movement analysis. In *Proceeding of the 28th IEEE EMBS Annual International Conference*, pages 1002–1009, New York City, USA, August 30th - September 3rd, 2006.
- [Donati *et al.*, 2007] M. Donati, V. Camomilla, G. Vannozzi, A. Cappozzo. Enhanced anatomical calibration

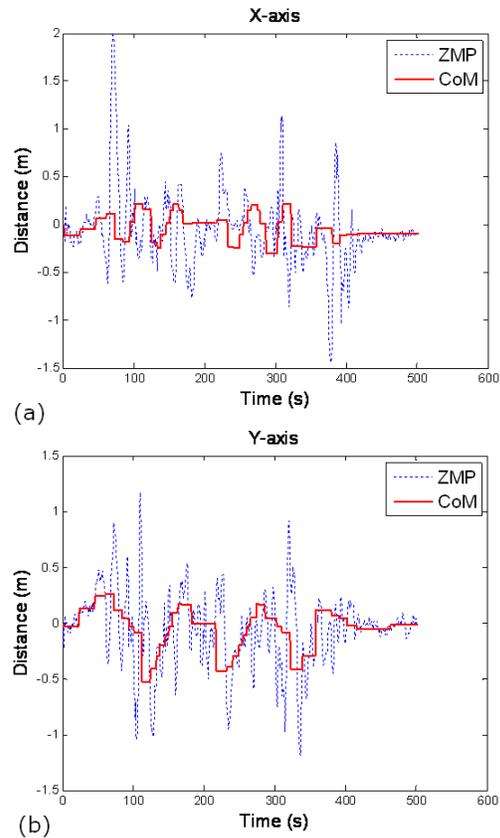


Figure 7: Result of a reduced number of modes (a) x-axis CoM and TP (b) y-axis CoM and TP.

in human movement analysis. *Gait and Posture*, May, 2007.

- [de Leva, 1996] Paolo de Leva. Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters. In *Journal of Biomechanics*, pages 1223–1230, Austin, Texas, August 1984. Kinesiology Department.
- [Green *et al.*, 2003] R. D. Green, L. Guan. Tracking Human Movement Patterns Using Particle Filtering. *Proceedings of the 2003 International Conference on Multimedia and Expo*, Vol. 3, 117-120, 2006.
- [Inamura *et al.*, 2004] T. Inamura, I. Toshima, H. Tanie. Embodied Symbol Emergence Based on Mimesis Theory. *International Journal of Robotics Research*, Vol. 23, No. 4-5, 363-377, 2004.
- [Kahol *et al.*, 2004] Kanav Kahol, Priyamvada Tripathi and Sethuraman Panchanathan. Automated Gesture Segmentation from Dance Sequences. In *Proceedings of the Sixth IEEE International Conference on Automatic Gesture and Face Recognition (FGR'04)*, Seoul, Korea, May 17th-19th 2004.

- [Lauer *et al.*, 2005] R. T. Lauer, B. T. Smith, R. R. Betz. Application of a Neuro-Fuzzy Network for Gait Event Detection Using Elettromyography in the Child With Cerebral Palsy. *IEEE Transactions on Biomedical Engineering*, Vol. 52, No. 9, 1532-1540, 2004.
- [Lopez *et al.*, 2006] A. Lopez, A. Melendez-Calderon. Human-inspired walking style for a low-cost biped prototype. *3rd IEEE LARS'06*, 141-148, 2006.
- [Menz *et al.*, 2003] H. B. Menz, S. R. Lord, R. C. Fitzpatrick. Age-related differences in walking stability. *Aged and Ageing*, Vol. 32, p137-142, 2003.
- [Naksuk *et al.*, 2005] N. Naksuk, G. Lee, S. Rietdyk. Whole-Body Human-to-Humanoid Motion Transfer. *Proceedings of 2005 5th IEEE-RAS International Conference on Humanoid Robotics*, 104-109, 2005.
- [Saboune *et al.*, 2005] J. Saboune, F. Charpillet. Markerless Human Motion Capture for Gait Analysis. *ArXiv Computer Science e-prints*, cs/0510063, October, 2005.
- [Sun *et al.*, 2006] Chao Sun, David Stirling, and Fazel Naghdy. Human Behaviour Recognition with Segmented Inertial Data. In *ARAA Australasian Conference on Robotics and Automation*, 1-9.
- [Zhou *et al.*, 2006] H. Zhou, H. Hu, Y. Tao. Inertial measurements of upper limb motion. *Med Bio Eng Comput*, 44:479-487, 2006.
- [Zijlstra *et al.*, 2003] W. Zijlstra, A. L. Hof. Assessment of spatio-temporal gait parameters from trunk acceleration during walking. *Gait and Posture*, Volume 18, 2003, p1-10.