Motion States Inference through 3D Shoulder Gait Analysis and Hierarchical Hidden Markov Models

Julien Collart, Robert Fitch and Alen Alempijevic
Centre for Autonomous Systems
University of Technology Sydney, 15 Broadway Ultimo, NSW 2007
julien.collart@student.uts.edu.au

Abstract
Automatically inferring human intention from walking movements is an important research concern in robotics and other fields of study. It is generally derived from temporal motion of limb position relative to the body. These changes can also be reflected in the change of stance and gait. Conventional systems relying on gait are usually based on tracking the lower body motion (hip, foot) and are extracted from monocular camera data. However, such data can be inaccessible in crowded environments where occlusions of the lower body are prevalent. This paper proposes a novel approach to utilize upper body 3D-motion and Hierarchical Hidden Markov Models to estimate human ambulatory states, such as quietly standing, starting to walk (gait initiation), walking (gait cycle), or stopping (gait termination). Methods have been tested on real data acquired through a motion capture system where foot measurements (heels and toes) were used as ground truth data for labeling the states to train and test the models. Current results demonstrate the feasibility of using such a system to infer lower-body motion states and sub-states through observations of 3D shoulder motion online. Our results enable applications in situations where only upper body motion is readily observable.

1 Introduction
Walking is considered to be one of the most common and important forms of human movement. Humans are known to constantly challenge their balance control system during daily life activities [Winter, 1995]. From a standing posture, they can initiate their motion, walk alternating the foot in contact with the ground, turn to avoid obstacles or reach specific targets, adapt their speed, stop their motion, or move backward. The style of walking of an individual, or gait [Kale et al., 2003; Lee and Grimson, 2002], can be defined as the overall pattern of bipedal locomotion mainly determined by the person’s body properties and behavioral patterns acquired over time [Subramanian et al., 2015]. It is a very complex behavior that requires coordination of the central nervous system, the muscles and the limbs [Sun, 2015]. The quantitative description of all mechanical aspects of ambulation is commonly referred to as gait analysis [Cappozzo, 1984] and entails measurement, analysis and assessment of the biomechanical features associated with the walking task [Best and Begg, 2006].

Gait analysis represents a particular interest to a diversity of fields such as physical medicine, psychological, surveillance, activity monitoring, athletic evaluation, and behavioral biometrics [Han, 2013; Zeng and Wang, 2015]. Gaits vary from one person to another in certain details such as their relative timing and magnitude [Wang et al., 2003], making them a promising unobtrusive biometric measurable at distance and not requiring interaction or body-invading equipment [BenAbdelkader et al., 2002; Wang et al., 2004; Johnson and Bobick, 2001]. However, gait has also been proved to be similar among human populations because it is determined by basic patterns of bipedal locomotion.

Current research on human motion tends to extract gait from lower-body motion, requiring an unobstructed view of the feet, hip, and legs. Such conditions are almost impossible to reproduce in environments where the human body, especially the legs, can be occluded by objects or other people in the scene. For example, environments with furniture obstructing the lower body or dense crowds in large open spaces such as airports. This paper proposes a new approach for inferring the gaits, states (gait initiation, cycles, termination) and sub-states, using 3D-motion of the shoulder extracted from people walking along a straight path. We demonstrate the possibility for such a system to infer the motion sub-states in both offline and online conditions by using foot measure-
ments as a ground truth for labeling the shoulder data used for training and testing individual Hidden Markov Models (HMMs), one per motion state, and the overall Hierarchical Hidden Markov Models (HHMM) used for inferring the motion states and sub-states.

We begin with a brief set of definitions of the motion states and sub-states required to be inferred from the shoulders in Section 2. Next, we describe the experiments performed to collect data used to validate our methods in Section 3. In Section 4, we present the methods and assumptions. Lastly, we evaluate the offline and online results given by each individual HMM and the overall HHMM in Section 5, and present conclusions in Section 6.

2 Background

In this section, we provide definitions of the motion states and sub-states used for this work. These states were selected based on the literature review on gaits to be used as a proof-of-concept. Alternative states may be substituted to fit a particular application.

2.1 Quiet standing

During quiet standing, the task is to keep the body’s center of mass (COM) safely within the base of support [Winter, 1995]. The COM and center of pressure (COP) positions are equal on the transverse plane when the body is motionless [Ellble et al., 1994]. When the person starts walking, muscular and gravitational actions are required to create the initial posture dynamic condition for moving forward [Mickelborough et al., 2004; Bruyneel et al., 2010]. Thus, for this work, a person whose 3D shoulder motion velocities are null will be considered to be quietly standing.

2.2 Gait initiation

Gait initiation is a transient movement between a quasi-static upright posture and steady-state gait [Lepers and Brenière, 1995]. It is a complex task requiring a variety of postural adjustments and shifts resulting in a forward step [Ellble et al., 1994; Assaiante et al., 2000]. It may be challenging for postural stability as the body must be accelerated forward from a stationary posture while simultaneously maintaining equilibrium within a smaller base of support due to shorter steps [Muir et al., 2014]. Two major parts constitute the gait initiation: (1) the anticipation phase, from onset to swing-foot toe-off, preceding and preparing the first step execution partly by creating an initial forward controlled pendulum fall [Assaiante et al., 2000], leaning the trunk forward to shift the COM [Brenière and Do, 1991]; and (2) the execution phase [Lepers and Brenière, 1995], from swing-foot toe-off to stance-foot toe-off [Laudani et al., 2006], representing the first step execution, programmed to adjust the progression of the COP to the COM velocity [Brenière and Do, 1991]. In this paper, we selected the anticipatory phase and the execution phase as gait initiation sub-states.

2.3 Gait cycle

Gait is a periodic movement with the period within walking called a gait cycle [Nowlan, 2009] and defined as the full cycle motion of a same leg: from the heel striking the ground to the ipsilateral heel striking again [Liu et al., 2004]. The cycle is composed of two major parts: (1) the stance phase (initial contact, loading response, midstance, and terminal stance) representing the actions of the first leg, and (2) the swing phase (pre-swing, initial swing, mid-swing, and terminal swing) representing the actions of the following one. We chose to use the support states as the gait cycle sub-states: the stance single support (mid-stance + terminal stance), the stance double support (pre-stance), the swing single support (mid-swing + terminal swing), and the swing double support (pre-swing).

2.4 Gait termination

Gait termination, representing the transition from steady-state gait to a “quiet” standing position [Jaeger and Vanitchatchavan, 1992; Novak et al., 2013] and implying returning the COM within the base of support, is more difficult to achieve than the gait initiation because of the need of the central neural system to predict the future position of the COM [Winter, 1995]. Although individuals can come to a stop at any time [Novak et al., 2013], studies have reported that the decision to stop has to be made at specific times during the gait to be able to terminate or take another step [Jaeger and Vanitchatchavan, 1992]: foot placement is the principal component to cease walking [Hase and Stein, 1998]. Individuals have to make an extra step when the signal to stop occurs late within the stance cycle of the trailing stance limb [Bishop et al., 2004]. In this paper, we selected the gait termination sub-states from the stance heel-on to the swing heel-off and from the swing heel-off to the stance heel-on.

3 Experimental data acquisition

This section details the experimental procedure used to capture and extract data for foot (heels and toes) and shoulder (positions and orientations) motion from 15 participants via a passive motion capture system.

3.1 Material

The environment used is a 10 m-diameter cylindrical room equipped with the Optitrack\footnote{More information regarding Optitrack available on official website: \url{http://optitrack.com}.} passive motion capture system capable of sub-20 \textmu m accuracy in optimal
conditions. It comes with Motive, which is software used to create rigid bodies from 3 (minimum amount of markers required to create a frame) infrared reflective markers placed on volunteers’ body parts (supposed to be rigid) to track their 3D positions and rotations over time (at 120 fps), record them, and finally export them into a spreadsheet.

3.2 Subjects

Fifteen healthy volunteers were randomly selected to perform the experiment: three females and twelve males from 21 to 32 years old. Participants were asked to keep their shoes on and rigid bodies were created through the software using the markers placed on their head, shoulders, heels, and toes (Figure 1).

3.3 Experimental protocol

Participants were asked to repeat the following exercise 40 times: stand quietly in an initial area, walk straight to a second area (outside the field of view in order to collect more gait cycles), re-enter the field of view with an ongoing motion, walk straight back to the initial area, stop, wait, and turn.

3.4 Data correction

Even though the capture system was calibrated prior to the experiment, a post-calibration correction was required to correct a subsisting floor orientation error: 3 infrared markers were placed on the ground to extract the gravity vector of the arena frame to rotate the experimental data. As some markers experienced occlusions, data with small time gaps in the acquisition were interpolated using a shape-preserving piecewise cubic interpolation of the points neighboring missing values. When the gaps were too large, data were discarded.

4 Methods

This section describes our framework designed to infer the motion states and sub-states, online or offline, using the shoulder 3D-motion measurements and the ground truth given by the feet combined, with a Hierarchical Hidden Markov Models (HHMM), a structured multi-level stochastic model, here represented as a Hierarchical Dynamic Bayesian Network (H-DBN) [Patel et al., 2012]. First, we introduce a method to determine the motion sub-states from individual Hidden Markov Models (HMM) designed for each motion state. Then, we give an overview on the design of the HHMM framework trained with the parameters extracted from the individual HMMs and evaluated with ground truth data.

4.1 Hidden Markov Models for offline inference

Any motion state can be represented as an individual HMM, a probabilistic model in which the states, here the sub-states of the current motion state, cannot be directly observed but inferred through measurements known as observations. It models $P(Q, O)$ for a sequence of observations, $O = \{o_1, o_2, ..., o_T\}$, and the corresponding sequence of hidden states, $Q = \{q_1, q_2, ..., q_T\}$. Each $q_t$ takes values in a finite set of hidden states $\{s_1, s_2, ..., s_N\}$ and each observation $o_t$ is from $\{v_1, v_2, ..., v_M\}$. The parameters defining an HMM, $\lambda = \{A, B, \pi\}$, are given by the state transition probabilities matrix $A$, where $a_{ij} = P(q_t = s_j | q_{t-1} = s_i)$, the observations probabilities matrix $B$, where $b_j(k) = P(o_t = v_k | q_t = s_j)$, and the initial state probabilities matrix $\pi$, with $\pi_i = P(q_1 = s_i)$. These parameters are found after training the model and are used to infer the states from a testing sequence.

Following the motion states and sub-states definition presented in the background section, we used the foot measurements to split the data (Figure 2) into motion datasets applied for training and testing: gait initiation begins when the stance foot heel-off for the first time and finishes when the stance foot toe-off for the second...
time; followed by the gait cycles; and finally, by the gait termination from the last stance foot heel-off to the last stance foot heel-on. However, due to the freedom people have to start and stop with either foot, we replaced the concept of stance and swing, that can vary between trials, by the foot side. Consequently, the gait initiation and termination can be split into: gait initiation with right foot (GIR), gait initiation with left foot (GIL), gait termination with right foot (GTR), and gait termination with left foot (GTL). Gait cycle (GCY) only requires one model as a cycle covers both the actions of both sides. It should be noted that in order to extract a template from the GCY training data, only the cycles starting with the same foot should be considered.

In this work, models used for ground truth rely on the foot measurements: right and left heel height (RH) (LH), right and left toes height (RT) (LT); and models used for proof of concept, use the shoulders measurements: heights (SH), Yaw angles (SY), Roll angles (SR), and transverse positions reoriented to be one-dimension only (SP). The sinusoidal-like shape of SH while walking straight (Figure 2) provides sufficient information for extracting the cycles using the peaks (2 cycles from this signal is equivalent to 1 gait cycle). Shoulder transverse position reorientation is made using the angle between the intended direction vector, computed with the full period for GIR, GIL, GTR, and GTL, or the gait cycle for GCY, and any reference direction vector defined a priori (here, the world frame x-axis).

Gaits have been proved to be similar among people because they are determined by basic patterns of bipedal locomotion. In this work, we designed a method to extract these basic patterns from every motion state to train the models. All the training data, coming from both foot and shoulder measurements, have been aligned and stretched using the extrema computed from RH, LH, RT, and LT, depending on the motion state to be trained. For GIR and GIL, from the beginning to the stance heel maximum (1-20%) to the swing heel maximum (40-60%) to the swing toe maximum (60-80%) to the end (80-100%). For GTR and GTL, from the start to the stance heel maximum (1-25%), to the stance toe maximum (25-50%), to the swing heel maximum (50-75%), to the end (75-100%). Finally, GCY was aligned from the beginning of the cycle to the swing toe maximum (1-20%), to the stance heel maximum (20-40%), to the stance toe maximum (40-60%), to the swing heel maximum (60-80%), to the end of the cycle (80-100%). Then, both training and testing data were standardized using a standard score on the entire period motion state period (except for GCY where the standardizing is made on each cycle), to remove any individual unique characteristics. Finally, sub-states periods were extracted from the feet measurements templates to label the shoulders ones (Figure 3).

For each motion states sub-state, we fit a mixture of $M$ Gaussians using K-means to estimate the initial parameters ($\mu$ and $\sigma$). Then, we improve these parameter estimates using Baum Welch’s Expectation Maximization (EM) for the full motion state. The outputs of this operation are the HMM parameters: a transition matrix, a mixture distribution, means vector, and covariance matrices. For each testing sample, the probability density function of a conditional mixture of Gaussians is evaluated at each time $t$ where the maximum probability provides the most likely sub-state and is added to a sequence vector. This sequence of sub-states is finally input to a Viterbi algorithm with the model parameters to find the most-probable path through the HMM.

### 4.2 Hidden Markov Models for online inference

In contrast to the offline detection, online inference poses some challenges and requires initial assumptions. First, we hypothesize that the orientation of the person is extractable from the quiet standing state. This information is necessary, yet no method was found for extracting the direction of travel from the ongoing motion of the shoulders in the transverse plane. In this work, we limit the inference solely to straight walking to utilise the intended direction of travel computed at the end of the previous cycle. The second assumption hypothesizes...
4.3 Overall Hierarchical Hidden Markov Model

An HHMM can use autonomous probabilistic models to represent its hidden states. They have a tree structure where each level represents a hierarchy. The states not emitting any observable symbol are called “internal states” and vertical transitions represent the activation of a sub-state by an internal state. A state at the hierarchy d \(|\text{root}|, 2, ..., D\) and time \(t\), is noted \(Q^d_t\). In this work, we use a 2-level hierarchy to differentiate the motion states \((d = 1)\) from the sub-states \((d = 2)\). An observations node, \(O_t\), provides the information to the system. And the termination state, \(F^d_t\), specifies the completion of a sub-HMM by returning 1 after \(Q^d_t\) has transited to its final state, 0 otherwise, to return the control back to the higher level. The HHMM parameters are given in the form of: a prior distribution and a transitions probability model defined at each levels, and an observation probability distribution model.

For the first level, the probability to start at any particular motion state \(j\) is equal: someone could enter the field of view with an ongoing motion. The transition matrix \(A^1\) is expressed by the time series shown in Figure 4. Then, given \(\pi^1\) and \(A^1\), the probabilities at the highest level can be defined as follow:

\[
P(Q^1_t = j|Q^1_{t-1} = i, F^2_{t-1} = f) = \begin{cases} \delta(i, j) & \text{if } F^2_{t-1} = 0 \\ A^1(i, j) & \text{if } F^2_{t-1} = 1 \end{cases}
\]

(1)

Where \(\delta(i, j)\) represents a probability threshold. In normal cases, \(\delta(i, j)\) is equal to 1 if \(i = j\), and 0 otherwise. No \(F^1\) is required at the highest level because the motion state inference never actually stops. The second level probability of being in a particular sub-state \(j\) while being in state \(k\) is given by:

\[
P(Q^2_t = j|Q^2_{t-1} = i, F^2_{t-1} = f, Q^1_t = k) = \begin{cases} A^2_k(i, j) & \text{if } F^2_{t-1} = 0 \\ \pi^2_k & \text{if } F^2_{t-1} = 1 \end{cases}
\]

(2)

with \(A^2_k\) and \(\pi^2_k\) respectively representing the transition distributions and prior probabilities of the sub-states (level 2) when the motion state, \(Q^1_t\), is \(k\). Finally, the observation model signifies the probability of seeing a specific observation conditioned on a discrete hidden state. We use both Gaussian and discrete observations:

\[
P(O_t|Q^2_t = i) = N(\mu, \Sigma)
\]

(3)

\[
P(O_t|Q^2_t = i) = C(i)
\]

(4)

The means and covariances matrices are given by the mixture of Gaussians trained for every individual HMM with online training data. \(C(i)\) is the probability to see the observation \(O_t\) when in a particular sub-state \(i\). Full sequences are used to train the full model. Expectation-Maximization with 10 iterations is used for extracting the final engine used for the global inference. For any testing sequence, the most probable explanation of the evidence is computed and compared with the data fully labeled.

5 Evaluation

In this section, we evaluate our methods on the data obtained from 15 people during the experiment. We start by providing a proof of concept regarding the inference of the motion sub-states through the shoulder measurements when the motion state is known and trained using an HMM. Then, we show that the inference can be performed online. Finally, we provide results regarding the motion states inference using an HHMM.

5.1 Step 1: Offline evaluation of individual HMMs

The main objective was to evaluate feasibility of inferring the motions states and sub-states. Table 1 shows the results of the offline inference of motion sub-states returned by the individual HMMs. We first trained and tested the individual models solely with foot measurements and a Gaussian mixture of 1 or 6 Gaussians to verify the extractability of the sub-states using the feet as this would constitute the ground truth. After data extraction (as presented in methods section) standardization was performed offline. Thereafter 80% of the data was used for training and 20% for the evaluation of the individual model. Results where computed in 3 different ways: individually (I), where only data extracted from
Table 1: Accuracy for the offline sub-states inference.

<table>
<thead>
<tr>
<th>Feature</th>
<th>GIR 1</th>
<th>GIR 2</th>
<th>GIL 1</th>
<th>GIL 2</th>
<th>GCY 1</th>
<th>GCY 2</th>
<th>GCY 3</th>
<th>GCY 4</th>
<th>GTR 1</th>
<th>GTR 2</th>
<th>GTL 1</th>
<th>GTL 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feet</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>GIR</td>
<td>93.90</td>
<td>89.40</td>
<td>93.40</td>
<td>84.80</td>
<td>85.80</td>
<td>93.80</td>
<td>75.40</td>
<td>85.70</td>
<td>94.70</td>
<td>93.60</td>
<td>98.10</td>
<td>93.90</td>
</tr>
<tr>
<td>GCY</td>
<td>89.40</td>
<td>93.40</td>
<td>84.80</td>
<td>93.80</td>
<td>75.40</td>
<td>85.70</td>
<td>94.70</td>
<td>93.60</td>
<td>98.10</td>
<td>93.90</td>
<td>98.10</td>
<td>97.30</td>
</tr>
<tr>
<td>GTR</td>
<td>85.80</td>
<td>93.80</td>
<td>75.40</td>
<td>85.70</td>
<td>94.70</td>
<td>93.60</td>
<td>98.10</td>
<td>93.90</td>
<td>98.10</td>
<td>97.30</td>
<td>98.10</td>
<td>97.60</td>
</tr>
<tr>
<td>GTL</td>
<td>75.40</td>
<td>85.70</td>
<td>94.70</td>
<td>93.60</td>
<td>98.10</td>
<td>93.90</td>
<td>98.10</td>
<td>97.60</td>
<td>98.10</td>
<td>97.60</td>
<td>98.10</td>
<td>97.60</td>
</tr>
</tbody>
</table>

Table 2: Accuracy for the online sub-states inference.

<table>
<thead>
<tr>
<th>Feature</th>
<th>GIR 1</th>
<th>GIR 2</th>
<th>GIL 1</th>
<th>GIL 2</th>
<th>GCY 1</th>
<th>GCY 2</th>
<th>GCY 3</th>
<th>GCY 4</th>
<th>GTR 1</th>
<th>GTR 2</th>
<th>GTL 1</th>
<th>GTL 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoulders</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Shoulders</td>
<td>57.30</td>
<td>55.80</td>
<td>73.50</td>
<td>61.00</td>
<td>77.30</td>
<td>65.90</td>
<td>64.00</td>
<td>70.10</td>
<td>64.00</td>
<td>74.40</td>
<td>40.60</td>
<td>97.00</td>
</tr>
<tr>
<td>Shoulders</td>
<td>80.80</td>
<td>71.70</td>
<td>79.00</td>
<td>68.90</td>
<td>74.80</td>
<td>58.10</td>
<td>64.00</td>
<td>70.10</td>
<td>64.00</td>
<td>74.40</td>
<td>63.50</td>
<td>99.20</td>
</tr>
<tr>
<td>Shoulders</td>
<td>67.20</td>
<td>50.00</td>
<td>71.00</td>
<td>59.80</td>
<td>76.30</td>
<td>58.90</td>
<td>64.00</td>
<td>70.10</td>
<td>64.00</td>
<td>74.40</td>
<td>63.50</td>
<td>99.20</td>
</tr>
<tr>
<td>Shoulders</td>
<td>63.60</td>
<td>36.00</td>
<td>83.10</td>
<td>78.50</td>
<td>71.50</td>
<td>65.80</td>
<td>62.80</td>
<td>76.70</td>
<td>62.80</td>
<td>84.80</td>
<td>84.10</td>
<td>97.60</td>
</tr>
<tr>
<td>Shoulders</td>
<td>57.30</td>
<td>55.80</td>
<td>73.50</td>
<td>61.00</td>
<td>77.30</td>
<td>65.90</td>
<td>64.00</td>
<td>70.10</td>
<td>64.00</td>
<td>74.40</td>
<td>40.60</td>
<td>97.00</td>
</tr>
<tr>
<td>Shoulders</td>
<td>80.80</td>
<td>71.70</td>
<td>79.00</td>
<td>68.90</td>
<td>74.80</td>
<td>58.10</td>
<td>64.00</td>
<td>70.10</td>
<td>64.00</td>
<td>74.40</td>
<td>63.50</td>
<td>99.20</td>
</tr>
<tr>
<td>Shoulders</td>
<td>67.20</td>
<td>50.00</td>
<td>71.00</td>
<td>59.80</td>
<td>76.30</td>
<td>58.90</td>
<td>64.00</td>
<td>70.10</td>
<td>64.00</td>
<td>74.40</td>
<td>63.50</td>
<td>99.20</td>
</tr>
<tr>
<td>Shoulders</td>
<td>63.60</td>
<td>36.00</td>
<td>83.10</td>
<td>78.50</td>
<td>71.50</td>
<td>65.80</td>
<td>62.80</td>
<td>76.70</td>
<td>62.80</td>
<td>84.80</td>
<td>84.10</td>
<td>97.60</td>
</tr>
<tr>
<td>Shoulders</td>
<td>57.30</td>
<td>55.80</td>
<td>73.50</td>
<td>61.00</td>
<td>77.30</td>
<td>65.90</td>
<td>64.00</td>
<td>70.10</td>
<td>64.00</td>
<td>74.40</td>
<td>40.60</td>
<td>97.00</td>
</tr>
<tr>
<td>Shoulders</td>
<td>80.80</td>
<td>71.70</td>
<td>79.00</td>
<td>68.90</td>
<td>74.80</td>
<td>58.10</td>
<td>64.00</td>
<td>70.10</td>
<td>64.00</td>
<td>74.40</td>
<td>63.50</td>
<td>99.20</td>
</tr>
<tr>
<td>Shoulders</td>
<td>67.20</td>
<td>50.00</td>
<td>71.00</td>
<td>59.80</td>
<td>76.30</td>
<td>58.90</td>
<td>64.00</td>
<td>70.10</td>
<td>64.00</td>
<td>74.40</td>
<td>63.50</td>
<td>99.20</td>
</tr>
<tr>
<td>Shoulders</td>
<td>63.60</td>
<td>36.00</td>
<td>83.10</td>
<td>78.50</td>
<td>71.50</td>
<td>65.80</td>
<td>62.80</td>
<td>76.70</td>
<td>62.80</td>
<td>84.80</td>
<td>84.10</td>
<td>97.60</td>
</tr>
</tbody>
</table>

Table 3: Accuracy for the online states inference using the HHMM.

<table>
<thead>
<tr>
<th>Feature</th>
<th>GIR 1</th>
<th>GIR 2</th>
<th>GIL 1</th>
<th>GIL 2</th>
<th>GCY 1</th>
<th>GCY 2</th>
<th>GTR 1</th>
<th>GTR 2</th>
<th>GTL 1</th>
<th>GTL 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoulders</td>
<td>57.30</td>
<td>16.20</td>
<td>75.70</td>
<td>39.80</td>
<td>38.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shoulders</td>
<td>65.50</td>
<td>64.20</td>
<td>77.60</td>
<td>29.50</td>
<td>30.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shoulders</td>
<td>63.60</td>
<td>36.00</td>
<td>76.90</td>
<td>27.60</td>
<td>60.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shoulders</td>
<td>69.20</td>
<td>65.10</td>
<td>82.90</td>
<td>11.60</td>
<td>32.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
a single person were used for training and testing, overall (O), where all people were used to train and test a single HMM, and Exception (E), where the models were trained using everyone except the single person used for testing. This detection has been made through 10 iterations using the data extracted from the trials of the 15 people presented in the experimental protocol. Results presented show a better accuracy in inferring the motion sub-states using 6 Gaussians in individual mode (slightly better than global). Using the same labels as the feet, we also extract the accuracy of the individual models using the shoulders. Results prove the possibility of inferring the sub-states with the shoulders when the measurements are processed offline with the most adapted standardization.

5.2 Step 2: Online evaluation of the individual HMMs

The possibility of inferring the sub-states using online measurements was evaluated. It is more challenging than offline inference due to limitation of using a generic standardization and lack of information on intended direction of travel. However, using the presented methods, online inference is still possible. Table 2 presents results of online detection accuracy when foot measurements are trained and tested offline with a Gaussian mixture of 6 Gaussians for each person to label the ground truth compared to the inference made with the models trained and tested online. The results highlight a better accuracy when using 6 Gaussians and training/testing the models for each person individually.

5.3 Step 3: Online evaluation of the HHMM

We used the previously trained online HMM parameters to train the HHMM. Because the HHMM is required to infer the motion states as well as the sub-states, it is trained and tested using the initial data, before being split, preprocessed as if they are coming from online shoulders measurements. Ground truth labels are given using the foot measurements extracted from step 2. 80% of the sequences are used for training the HHMM, and 20% for testing it. Results (Table 3) show a better accuracy using the training individually for each person. Whereas the accuracy seems to be correct regarding the detection of the gait initiation and cycles, it drops for the gait termination. These results could be explained by a lack of data used to train the models for each individual; accuracy when using the overall data is better than when the model is trained individually.

6 Conclusions and future work

This paper proposes a new approach for inferring motion states, such as gait initiation, cycles, and termination, and sub-states using Hierarchical Hidden Markov Models and provides a proof-of-concept by comparing the outputs from HMMs and HHMMs when foot measurements are used to train a model used as ground truth. Our work, which relies on upper body observations, is important because it enables applications of human motion state inference to new scenarios which were previously not possible because lower body observations could not be obtained.

At present the online estimation leverages data from all previous persons observed. In fact, people walk with similar, not exact gait patterns. In future work we aim to identify this subset of characteristics tied to groups of people and establishing online normalization of data. Another limitation of this preliminary framework is that we have defined only an initial common set of states and sub-states. In future work, we will focus on adapting the states and sub-states in order to evaluate the possibility of using such a framework for anticipating turns. Also, because of the bipedal patterns extracted from the shoulders, we are going to investigate the feasibility of creating a motion model to be integrated in tracking, while still based on shoulder motion.

References


