

Augmented Human State Estimation using Interacting Multiple Model Particle Filters with Probabilistic Data Association

Georgia Chalvatzaki, *Student Member, IEEE*, Xanthi S. Papageorgiou, *Member, IEEE*,
Costas S. Tzafestas, *Member, IEEE* and Petros Maragos, *Fellow, IEEE*

Abstract—The accurate human gait tracking is an important factor for various robotic applications, such as robotic walkers aiming to provide assistance to patients with different mobility impairment, social robot companions, etc. A context-aware robot control architecture needs constant knowledge of the user’s kinematic state to assess the patient’s gait status and adjust its movement properly to provide optimal assistance. In this work we present a novel human gait tracking approach that uses two Particle Filters (PFs) and Probabilistic Data Association (PDA) with an Interacting Multiple Model (IMM) scheme for a real-time selection of the appropriate motion model according to the human gait analysis and the use of the Viterbi algorithm for an augmented human gait state estimation. The gait state estimates also interact with the IMM as a prior information that drives the Markov sampling process, while the PDA ensures that the legs of the same person are coupled. The observation data in this work are provided by a Laser Range finder (LRF) mounted on a robotic assistant walker. A detailed experimental validation is presented using ground truth data from a motion capture system, which was used in real experiments with elder subjects who presented various mobility impairments. The validation analysis regards the algorithm’s accuracy, robustness to occlusions and clutter, and the gait state classification success, subject to the effect of different number of samples used in the PFs. The results for the elder subjects show the dynamics of the proposed algorithm to be used in a real-time application due to its efficacy to provide accurate and robust augmented human gait estimates with a small number of particles.

Index Terms—Human detection and tracking, human-centered robotics, physically assistive devices, automation in life sciences: biotechnology, pharmaceutical and health care, medical robots and systems.

I. INTRODUCTION

HUMAN detection and tracking is a common research topic, and refers to the accurate estimation of a human’s position w.r.t. various sensors. Surveillance, ambient living environments, mobile robot companions and robotic assistants

are some of the applications where human tracking is necessary, [1]–[3]. Cameras are commonly used for human activity detection, [4]. Recent research focus on using Laser Range Finder (LRF) sensors for detection, due to the simple two-dimensional representation of the world, the lower computational cost in extracting features and their faster frame rates.

Approaches for tracking people can be found in [5], where the authors presented a method for tracking multiple persons using data from an LRF mounted on a mobile robot, using a Joint Probabilistic Data Association Filter on the samples of a Particle Filter (PF), to associate measurements to the respective tracked persons; a pedestrian tracking method using multiple static LRFs is presented in [6], where a Kalman Filter (KF) tracks the legs using a human walking model inducing predefined constant velocities for the stance/swing phase. A single person tracker for a mobile robot is presented in [7], using a Multiple Hypothesis Tracking (MHT) framework that incorporates a switched dynamic walking model for the left/right leg in an extended KF.

A fusion of visual and laser data for a mobile service robot has been implemented in [8] using Unscented KF for human tracking. In [9], an approach for a MHT laser-based people tracking for a mobile robot is presented. This framework uses a KF for tracking legs separately and the MHT associates past with current tracks, [10]. A detection and tracking framework for a mobile service robot that follows and interacts with humans is proposed in [11]. A robotic walker is presented in [12], which has two LRFs mounted, each one for tracking the user’s legs separately using PFs. A robotic walker for Parkinson’s Disease patients is presented in [13], which is equipped with a small LRF that detects the legs positions. A legs tracking framework for a walker used for rehabilitation is described in [14].

More recently, [15] presented a method based on two parallel Interacting Multiple Model (IMM)-unscented KFs to track human positions and velocities for a companion robot. The IMM framework incorporates two kinematic models; one for turning and one for constant heading motion. The IMM method is commonly used for tracking targets that present changes in their motion, thus cannot be modeled by a constant motion model. It is also a framework closely linked to data association problems. Although most IMM approaches use a bank of filters that run in parallel to correctly track a target [15], the PF can be easily used in the IMM framework for target tracking [16]. IMM-PF implementations are also

Manuscript received: September, 10, 2017; Revised December, 11, 2017; Accepted December, 31, 2017.

This paper was recommended for publication by Editor Tamim Asfour upon evaluation of the Associate Editor and Reviewers’ comments. This research work has been partially supported by two EU-funded Projects: BabyRobot (H2020, grant agreement no. 687831) and I-SUPPORT (H2020-PHC-19-2014, grant agreement no. 643666).

The authors are with the School of Electrical and Computer Engineering, National Technical University of Athens, Greece (email: {gchal, xpapag}@email.ntua.gr, {ktzaf, maragos}@cs.ntua.gr).

Digital Object Identifier (DOI): see top of this page.

used for multiple target tracking with association probabilities [17], but also in the field of fault diagnosis and systems identification [18].

This paper presents a novel human gait tracking algorithm that uses two PFs to track the legs and Probabilistic Data Association (PDA) for correlating the two legs to the same subject. An IMM is constructed as a first-order Markov model using as states specific gait phases defined in the gait analysis literature. The IMM guides the selection of the appropriate motion model for the particles of each leg. Moreover, an augmented human state estimate, regarding the corresponding gait phase, is provided using the Viterbi algorithm. This human gait state estimate is also the prior that drives the Markov sampling from the IMM for the next motion model selection. The data used in this work are provided by a single LRF that is mounted on an assistant robotic walker. This algorithm can be used for tracking a human who presents pathological gait, i.e. it deviates from normal walking function, either when a person changes its gait due to fatigue or when another person walks closely by, then the selection of the correct motion model along with the PDA helps the efficient tracking of the same person.

This tracking approach interweaves with our aim to more general gait tracking applications, e.g. the augmented human gait state estimate could provide in real-time the appropriate information for gait analysis, in contrast with our previous work that propagated the legs state information to an HMM, [19], therefore needing more computational time. A real-time gait analysis could infer the appropriate user-adaptation of the control strategy for a robotic assistant walker in a context-aware robot control architecture, [20]. Also, the augmented human gait state estimation can provide the necessary information for human identification and therefore could be used to lock the tracking process to one person with a specific motion model, consequently it could be used for multiple people tracking and identification. Gait-based human identification has also been presented in [21], where Deep Learning methods on energy images of several subjects have been used. However, to the best of our knowledge, there is no other method in literature that performs such a tracking and identification task based on augmented human state estimation from laser data.

In this work, we validate our IMM-PFs with the PDA tracking algorithm (IMM-PDA-PF) using ground truth data from a motion capture system, which was exploited during experiments with real patients who used an assistant robotic walker. Specifically, we evaluate the performance of our algorithm regarding its tracking accuracy, its robustness towards occlusions and environmental clutter and finally its classification success on estimating the human gait state.

II. PROBLEM STATEMENT

The problem we aim to solve consists of performing efficient and robust leg tracking and also estimating the human gait state, i.e. to provide an augmented human state estimation (both estimates of the legs states and the respective gait phase). The estimated human gait state will also provide inference regarding the desired motion model of the leg tracking system. For the legs tracking, we aim to estimate the kinematic state

of the legs at each time instant, regardless of the cluttered environment or occlusion cases. Let $\mathbf{x}_{f,k}$ be the state of each leg, where $f=\{left, right\}$ at each time instant k . The state space incorporates the Cartesian position vector $\mathbf{p}_{f,k} = [x \ y]^T$ and the respective velocity vector $\mathbf{v}_{f,k} = [v_x \ v_y]^T$. Thus, the state vector of each leg is denoted as:

$$\mathbf{x}_{f,k} = [\mathbf{p}_{f,k}^T \ \mathbf{v}_{f,k}^T]^T \quad (1)$$

For the human gait state estimation, we need to estimate the current gait phase. The gait phase is a discrete parameter, defined as θ_k and is drawn from a set of discrete states $\Theta = \{s_1, \dots, s_M\}$, where M is the number of gait states. The transition between those states is governed by a first-order Markov chain, where the transition probability between two states is denoted as:

$$a_{ij} = P(\theta_k = s_j | \theta_{k-1} = s_i), \forall s_i, s_j \in \Theta \quad (2)$$

where $i, j = 1, \dots, M$.

Thus, our problem evolves into estimating the augmented human state $\{\mathbf{x}_{f,k}, \theta_k\}$, i.e. the posterior pdf $p(\mathbf{x}_{f,k}, \theta_k | \mathbf{y}_k)$, where \mathbf{y}_k is the measurement vector at time k . The legs' states estimates are computed by the marginal distribution of the posterior pdf:

$$p(\mathbf{x}_{f,k} | \mathbf{y}_k) = \sum_{i=1}^M p(\mathbf{x}_{f,k} | \theta_k = s_i, \mathbf{y}_k) \cdot p(\theta_k = s_i | \mathbf{y}_k) \quad (3)$$

Given this formulation the legs' tracking problem is a two-step iterative estimation problem:

- 1) *the prediction step*, where the state is propagated in time according to the dynamics $p(\mathbf{x}_{f,k} | \mathbf{x}_{f,k-1}, \theta_k)$, i.e. the probability density function describing the state evolution given by the state transition equation:

$$\mathbf{x}_{f,k} = g_{f,k}(\mathbf{x}_{f,k-1}, \theta_k, \mathbf{w}_f) \quad (4)$$

where $g_{f,k}$ is a possibly non-linear function describing the discrete time dynamics of the each leg given the human state θ_k and \mathbf{w}_f is the process noise modelled as white Gaussian with covariance matrix $\Sigma_{\mathbf{w}_f}$.

- 2) *the correction step*, where the state estimate is updated according to measurements \mathbf{y}_k . The observations are provided by a LRF that is located at the origin of the coordinate system and contains observations of the subject's legs along with environmental clutter. The observations that relate to each leg are described by the measurement equation:

$$\mathbf{y}_{f,k} = h_{f,k}(\mathbf{x}_{f,k}, \theta_k, \mathbf{n}_f) \quad (5)$$

where $h_{f,k}$ is a possibly non-linear function and \mathbf{n}_f is an independent & identically distributed measurement noise vector.

For the estimation of the human gait state, we must compute the marginal distribution of the posterior pdf:

$$p(\theta_k | \mathbf{y}_k) = \int p(\mathbf{x}_{f,k}, \theta_k | \mathbf{y}_k) d\mathbf{x}_{f,k} \quad (6)$$

For addressing this augmented estimation problem, we propose an IMM that describes the human gait states, but also imposes a certain motion model for the human legs. The IMM

Table I: Gait States Definition

s_1 =LDS	Left Double Support
s_2 =LS/RW	Left Stance/ Right Swing
s_3 =RDS	Right Double Support
s_4 =RS/LW	Right Stance/ Left Swing

Definition of the possible gait states of the IMM.

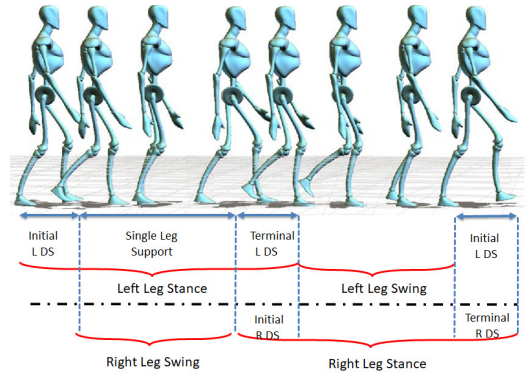


Figure 1: Human gait cycle representation regarding Single Leg or Double Leg Support (DS).

is incorporated into the leg tracking approach. In this approach, we solve the problem of tracking the two legs as interacting targets, i.e. we estimate the legs' states $\mathbf{x}_{f,k}$ at each time instant k , applying a PDA on the correction step of the filtering for associating each leg to the correct measurement vector $\mathbf{y}_{f,k}$ and keep the two legs coupled, in order to overcome problems of leg occlusions and cluttered environment. The human gait state estimate results from computing a joint likelihood for both legs and the use of the Viterbi decoding algorithm, exploiting the IMM's Markov chain that describes the gait states transitions.

III. IMM-PDA-PF AUGMENTED HUMAN STATE ESTIMATION

A. Interacting Multiple Model for Gait Tracking

The IMM used in this work is inspired by human gait analysis [22]. There are two main periods in gait cycle: the stance, when the foot is on the ground, and the swing when that same foot is no longer in contact with the ground and is swinging through, in preparation for the next foot strike. The stance period can be subdivided into three internal time intervals: the initial Double Support (DS), the single leg support and the terminal DS, Fig. 1. The initial DS begins with the Heel Strike (HS) and it is the time when both feet are on the ground. The single leg support is the period when only one leg is at stance while the opposite leg is swinging. The terminal DS begins with the HS of the contra lateral foot and continues until the original stance leg begins to swing. The gait cycle can be seen as an interacting model; when the one leg is in stance phase the other one is swinging.

With this reasoning, we have defined the first-order Markov model of Fig. 2 as the IMM for the human gait transition model. In Fig. 2 the human gait states are defined, i.e. the discrete states $s_i \in \Theta$, $i = 1, \dots, 4$ and the possible transitions a_{ij} between them are defined in (2). The gait states are described in Table I. Each state s_i refers to both legs and imposes a different motion model. Namely, each state s_i is characterized by a set of velocity Gaussian Mixture Models (GMMs), that alter the transition probability function $p(\mathbf{x}_{f,k}|\mathbf{x}_{f,k-1}, \theta_k = s_i)$, according to the dynamics that govern the state s_i . Thus, in the prediction step we use the imposed state transition model $p(\mathbf{x}_{f,k}|\mathbf{x}_{f,k-1}, \theta_k)$ to compute the predicted pdf: $p(\mathbf{x}_{f,k}, \theta_k|\mathbf{y}_{k-1})$:

$$p(\mathbf{x}_{f,k}, \theta_k|\mathbf{y}_{k-1}) = \int p(\mathbf{x}_{f,k}, \mathbf{x}_{f,k-1}, \theta_k|\mathbf{y}_{k-1}) d\mathbf{x}_{f,k-1} = \int \sum_{\theta_{k-1}} p(\mathbf{x}_{f,k}|\mathbf{x}_{f,k-1}, \theta_k) \cdot p(\theta_k|\theta_{k-1}) \cdot p(\mathbf{x}_{f,k-1}, \theta_{k-1}|\mathbf{y}_{k-1}) d\mathbf{x}_{f,k-1} \quad (7)$$

where $p(\theta_k|\theta_{k-1})$ is defined in (2) for the gait IMM and $p(\mathbf{x}_{f,k-1}, \theta_{k-1}|\mathbf{y}_{k-1})$ is the posterior joint state estimate at time $(k-1)$. At time k , a new observation \mathbf{y}_k comes, and then the

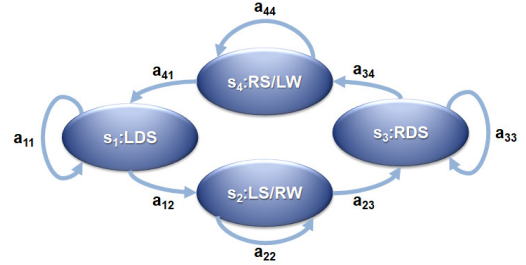


Figure 2: The gait IMM as a first-order Markov chain that represents the possible transitions for the human gait states (Table I).

posterior can be computed as:

$$p(\mathbf{x}_{f,k}, \theta_k|\mathbf{y}_k) = \frac{p(\mathbf{y}_k|\mathbf{x}_{f,k}, \theta_k) \cdot p(\mathbf{x}_{f,k}, \theta_k|\mathbf{y}_{k-1})}{\int \sum_{\theta_k} p(\mathbf{y}_k|\mathbf{x}_{f,k}, \theta_k) \cdot p(\mathbf{x}_{f,k}, \theta_k|\mathbf{y}_{k-1})} \quad (8)$$

where $p(\mathbf{y}_k|\mathbf{x}_{f,k}, \theta_k)$ is the gait state conditional likelihood function for the correction step.

B. Human Gait State Estimate

Because the transition model of the gait IMM refers to the movement of both legs, for the human gait state estimate $p(\theta_k|\mathbf{y}_k)$ defined in (6), we have to account both legs' estimated states $\mathbf{x}_{f,k}$. Thus, we define a human-centred state space defined as: $\mathbf{x}_k^H = \mathbf{x}_{left,k}^H - \mathbf{x}_{right,k}^H$ and compute the joint likelihood, i.e. (6) is re-written as

$$p(\theta_k|\mathbf{y}_k) = \int p(\mathbf{x}_k^H, \theta_k|\mathbf{y}_k) d\mathbf{x}_k^H \quad (9)$$

For estimating (9), we have defined GMMs regarding the joint distribution of the \mathbf{x}_k^H for each state s_i in the gait IMM and evaluate the conditional probabilities of being in state $\theta_k = s_j$, given the human-centred state \mathbf{x}_k^H and the measurements \mathbf{y}_k at time instant k computed as:

$$L_k = \begin{cases} p(\theta_k = s_j|\mathbf{x}_k^H, \mathbf{y}_k, \theta_{k-1} = s_i), & \text{for } j = \{i, i+1\} \\ 0, & \text{for } j \neq \{i, i+1\} \end{cases} \quad (10)$$

The results of (10) are used as emission probabilities L_k in the Viterbi algorithm with a one step back-propagation [23], i.e the human gait state estimate results as the Maximum a Posteriori (MAP) solution for the time interval $(k-1, k)$:

$$\hat{\theta}_k^{MAP} = \arg \max_{\mathbf{x}_{k-1,k}^H} (L_k) \quad (11)$$

The MAP estimate (11) is solved using optimization via the dynamic programming of the Viterbi algorithm. In the following section we will describe how the IMM is incorporated into a leg tracking method that used two PFs with Probabilistic Data Association (PDA-PF), [24].

C. IMM-PFs Probabilistic Data Association Legs Tracking

Our implementation incorporates two filters for estimating the position and velocity of each leg separately and associate them probabilistically. The particles represent samples of the posterior density distribution of the legs' states $\mathbf{x}_{left,k}$ and $\mathbf{x}_{right,k}$ at each time instant k for the left and right leg respectively. Each state constitutes of the Cartesian position and velocity along the axes. The implementation covers the basic particle filter methodology [25], including initialization, propagation in time, particles' weights update, resampling and posterior estimation.

Initialization: At the first time instant $k=1$, we initialize a set of N particles for each leg. Let the position of the n^{th} particle, for $n = 1, \dots, N$, be noted as: $\mathbf{p}_{f,k}^n = [x \ y]^T$ and its velocity as: $\mathbf{v}_{f,k}^n = [v_x \ v_y]^T$, where $f \in \{left, right\}$ is the label of each leg. The particles' states are denoted as:

$$\mathbf{x}_{f,k}^n = [\mathbf{p}_{f,k}^n \ \mathbf{v}_{f,k}^n]^T \quad (12)$$

Only for initialization we detect the initial positions of the legs with respect to the robotic platform inside an observation window and discriminate the left from the right leg using K-means clustering and circle fitting. The particles' positions are initialized to be equal to the detected positions. We also draw N samples for the legs' velocity from an initial zero-mean GMM distribution (we consider that both legs are still in front of the robot for initialization). The particles' weights $\omega_k^{f,n}$ of each leg are initialized equal to: $1/N$, with $n = 1, \dots, N$. The initial posterior estimate is approximated by the Minimum Mean Square Error:

$$\mathbf{x}_{f,k} = \sum_{n=1}^N \omega_{f,k}^n \cdot \mathbf{x}_{f,k}^n = [\mathbf{p}_{f,k} \ \mathbf{v}_{f,k}]^T \quad (13)$$

Regarding the human gait state estimation, only for initialization we deviate from (10), in the sense that we calculate the probabilities $L_1 = p(\theta_1 = s_j | \mathbf{x}_k^H)$, $\forall s_j \in \Theta$. The initial human gait state estimate θ_1 results from the maximum likelihood: $\hat{\theta}_1 = \arg \max(L_1)$.

IMM Particles' Propagation: At each time frame $k=2, \dots, T$ (where T is the total tracking time) the particles' states are propagated in time using the IMM and the state transition pdf $p(\mathbf{x}_{f,k} | \mathbf{x}_{f,k-1}, \theta_k)$. We use Markov Chain Monte Carlo sampling using the gait IMM transition model to draw N samples for the human gait state at time k , noted as θ_k^n , given the estimated gait state at time $(k-1)$, $\hat{\theta}_{k-1}$. Each gait state θ_k^n imposes a different motion model for the respective n^{th} particle of each leg according to certain velocity GMMs that describe the states s_i . The selected GMMs from the respective θ_k^n are centred around the previously estimated velocities $\mathbf{v}_{f,k-1}$. Let $\mathbf{v}_{f,k}^n$ be the n^{th} velocity sample drawn from the respective GMM of the state θ_k^n at time instant k . Then, the position of the n^{th} particle is propagated in time according to the equation:

$$\mathbf{p}_{f,k}^n = \mathbf{p}_{f,k-1} + \mathbf{v}_{f,k}^n \cdot \Delta t \quad (14)$$

where $\mathbf{p}_{f,k-1}$, is the estimated position vector of each leg for the $(k-1)$ time frame.

IMM-PDA Observation Likelihood computation: One of the main objectives is to correctly associate each leg with the respective cluster of laser points $\mathbf{y}_{f,k}$, where $f = \{left, right\}$. The problem, however, is that the legs do not have an explicit observation, but a cluster of laser points form the observation. The cluster of laser points that represents a leg has variable number of points, has a deformable shape and often contains outliers. Thus, inspired by the methods used in data association literature, [26], we considered the legs to be two interacting targets and we implemented a data association, to associate each leg to the most representative cluster of laser points, but also keep the formation of the two legs that belong to the same person by taking into consideration the IMM formulation.

At each time instant k , the particles' weights have to be updated according to the observations. The observations are the Cartesian positions of the laser points in the sagittal plane. In this implementation, we use an observation window for each leg, which is an experimentally defined rectangular area, centred around each particle, so that every sample $\mathbf{x}_{f,k}^n$ is associated with a different cluster of laser points, $\mathbf{y}_{f,k}^n$. Given the IMM formulation the particle weights depend on the gait state conditional observation likelihood:

$$\omega_{f,k}^n = \frac{p(\mathbf{y}_k | \mathbf{x}_{f,k}^n, \theta_k^n) \cdot p(\theta_k^n | \hat{\theta}_{k-1})}{\int \sum_k p(\mathbf{y}_k | \mathbf{x}_{f,k}^n, \theta_k^n) \cdot p(\theta_k^n | \hat{\theta}_{k-1})} \quad (15)$$

We treat each particle as a possible leg center and we expect the observations to be on the circular circumference of this center. We set the observation window centred on the n^{th} particle and we associate it with the corresponding observations, i.e. the laser points $\mathbf{y}_{f,k}^n$ detected inside the window. The observation likelihood that will provide the weight of the particles is computed based on three factors:

1. *The distribution of the laser points in the circular contour given the center (i.e. the respective particle):* In Fig. 3 an example of the circular representation of the legs from the laser points w.r.t. the laser scanner is presented. On the right of Fig. 3, there is a depiction of the detected laser points with black stars, the green and magenta circles are the circular representations of the right and left leg respectively. The labels $R0, R1, R2, R3, R4$ represent the segmentation of the circle into regions (the regions' boundaries are depicted with orange lines) based on which we have computed the observation likelihood for the IMM-PDA-PF. We have divided horizontally the circle into two semicircles. Laser points in the upper semicircle $R0$ do not contribute to the observation likelihood. The lower semicircle is split into *four* regions ($R1, \dots, R4$) of equal angle range. We have calculated the normal distribution of the Euclidean distances of the laser points of each region w.r.t. the corresponding center. Let \mathbf{d}_{R_m} be the vector of distances of the laser points w.r.t. the corresponding circle center for the R_m region, with $m \in 1, \dots, 4$. Thus, each region R_m is described by a normal distribution of the distances $\mathcal{N}(\mathbf{d}_{R_m} | \mu_{R_m}, \Sigma_{R_m})$, with μ_{R_m} the mean distance and Σ_{R_m} is the covariance matrix.

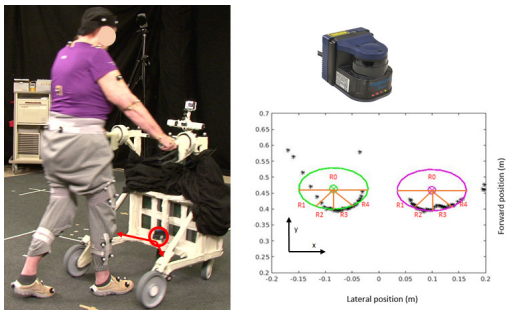


Figure 3: Example of the circular representation of the legs from the laser points w.r.t. the laser scanner. Left: a snapshot of a subject walking with the rollator while the Hokuyo LRF scans the walking area; Right: a presentation of the detected laser points, the circular representations of the legs and the regions based on which we compute the observation likelihood.

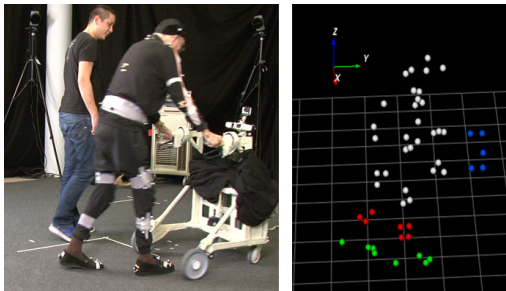


Figure 4: Left: A snapshot from the experimentation scene. A patient with his normal clothes wears a set of visual markers while walking with support of the passive rollator. Right: a representation of the visual markers from MOKKA visualization system.

2. *The number of laser points inside each observation window:* A normal kernel distribution, noted as λ_n for every particle with $n = 1, \dots, N$, describes the likelihood of the number of laser points detected on the circular contour that represents the leg.

3. *Association probability* We compute an association probability about the Euclidean distance between the two legs. The human legs are two interacting moving targets, and thus we introduce an association probability β_i , modelled by a Gamma distribution. This probability regulates the observation likelihood of the one leg w.r.t. the other; we account how probable is the current state of the n^{th} particle of the one leg w.r.t. the estimated position of the other leg at $(k-1)$. By inserting the association probability, we achieve to control the relative positions of the two legs, especially in cases of leg occlusions or cluttered environment, but also to assign the correct observations to each leg.

For the n^{th} particle of each leg we compute the observation likelihood using the following function:

$$p(\mathbf{y}_{f,k}^n | \mathbf{x}_{f,k}^n, \theta_k^n) = \beta_n \cdot \left[\lambda_n \cdot \sum_{m=1}^4 \pi_{R_m} \cdot \mathcal{N}(\mathbf{d}_{R_m} | \mu_{R_m}, \Sigma_{R_m}) \right] \quad (16)$$

We assume as π_{R_m} , the importance weights of the four regions, which were set experimentally so that the extreme regions $R1$ and $R4$, which often contain many outliers have less importance than the inner regions $R2$ and $R3$. All parameters have been experimentally defined. The weights are normalized

for all particles $n = 1, \dots, N$ according to:

$$\hat{\omega}_{f,k}^n = \omega_{f,k}^n \sum_{j=1}^N \omega_{f,k}^j \quad (17)$$

Resampling: For confronting the problems of weight degeneracy and sample impoverishment [25], at each time frame we check whether the effective sampling size $N_{eff} = 1 / \sum_{n=1}^N \hat{\omega}_{f,k}^n$ is less than the threshold $N_{thr} = N/2$. If so, we apply a random walk on the current particles' state providing new particles $^* \mathbf{x}_{f,k}^n$ given the sampled state θ_k^n . Then, we evaluate the weights of the new particles, according to the IMM-PDA Observation Likelihood computation, which provides the new weights: $^* \hat{\omega}_{f,k}^n$. Having the old pairs of particles and their weights $(\mathbf{x}_{f,k}^n, \hat{\omega}_{f,k}^n)$ and the new ones $(^* \mathbf{x}_{f,k}^n, ^* \hat{\omega}_{f,k}^n)$, we apply the Metropolis-Hastings algorithm, [27]. Based on this algorithm we can decide whether or not we have to replace the n^{th} pair $(\mathbf{x}_{f,k}^n, \hat{\omega}_{f,k}^n)$ with the new samples $(^* \mathbf{x}_{f,k}^n, ^* \hat{\omega}_{f,k}^n)$.

Posterior Estimation: For the posterior state estimate $p(\mathbf{x}_{f,k} | \mathbf{y}_{f,k})$ (from the marginalized posterior pdf of (3)), we apply maximum likelihood for finding the particle with the highest weight and collect the “best” particles, i.e. those having a weight greater or equal than 80% of the maximum weight:

$$s = \arg \max_i [\hat{\omega}_{f,k}^i \geq 0,8 \cdot \max(\hat{\omega}_{f,k}^s)] \quad (18)$$

where s is the index of the “best” particles, i.e. $s \in S \subseteq \{1, \dots, N\}$. In that way, we have a dynamic system, that leaves out particles that may track outliers and could contaminate the posterior estimation, and therefore provides smoother estimates. The posterior state estimate is then approximated by the weighted mean of the “best” particles:

$$p(\mathbf{x}_{f,k} | \mathbf{y}_{f,k}) = \left(\sum_s \mathbf{x}_{f,k}^s \cdot \hat{\omega}_{f,k}^s \right) / \left(\sum_s \hat{\omega}_{f,k}^s \right) \quad (19)$$

IV. EXPERIMENTAL ANALYSIS & VALIDATION

A. Experimental setup and data description

Experiments with real patients were conducted in Agaplesion Bethanien Hospital - Geriatric Center, under ethical approval by the ethics committee of the Medical Department of the University of Heidelberg. All subjects had signed written consent for participating in the experiments. The participants presented moderate to mild mobility impairment, according to clinical evaluation. The patients were wearing their normal clothes. For Ground Truth (GT), a set of markers from a VICON Motion Capture system was placed on certain areas of the subjects' body. In this work we employ data from a dataset of 23 patients aged over 65 years old. The subjects participated in a walking scenario having physical support of a rollator, where they had to walk in a room and make some turning manoeuvres to avoid obstacles. All patients performed the experimental scenarios under appropriate carer's supervision.

A snapshot of the experimentation scene with a subject walking supported by the robotic rollator while wearing a set of visual markers is shown on the left of Fig. 4; on the right, a representation of the markers from the MOKKA visualization system is provided. Two types of GT data are

needed for the validation of our work: a) the GT data of the tibia movement for validating the tracking accuracy, and b) the GT data of the gait phases for evaluating the classification. In Fig. 4, marked with red are the tibia markers and with blue the rollator markers, which are used for extracting the GT of the legs' movement w.r.t. the rollator. To extract GT data we apply a series of interpolation, cylinder fitting and smoothing algorithms for the positions, as well as simple differentiation for the velocities. Heel and toe markers (marked green) are also used to extract GT gait phases, based on an automatic gait phase detection algorithm thoroughly described in [28].

The extracted GT data from the 15 subjects constitute the training dataset (comprising approximately 60.000 frames, 75% of the whole dataset), dedicated to training the GMMs of the motion models, the pdfs that describe the gait states and the Transition Probability Matrix of the gait IMM of Fig. 2. In particular, 2 mixture-GMMs were used for modelling the DS phases and 5-mixtures for the stance/swing phase of the left and right legs in the IMM of Fig. 2. The data from the remaining 8 subjects formed the testing set, containing approximately 25.000 scanning frames, 25% of the whole dataset, or else ~ 12 minutes of walking. It is important to mention, that all the subjects had various mobility impairments, giving a rich dataset of different gaits, i.e. different gait speeds and variable gait phases durations. In this work we have used as training set a randomly selected group of subjects. In future work we plan to perform a systematic cross-validation study using different partitions of the training/testing dataset.

The laser data were provided by a Hokuyo rapid laser sensor UBG-04LX-F01, mounted on the robotic platform of Fig. 4 for the detection of the patients' legs. The laser sensor is placed at a height of about 40 cm from the ground in order to capture the motion of the subject's tibia.

B. Validation Strategy

Our validation strategy comprises the testing of our algorithm regarding: *i) its accuracy*, by computing its average position and velocity Root Mean Square Error (RMSE) w.r.t. the ground truth data from the visual markers, *ii) its robustness* evaluating the algorithm's behaviour in cases of leg occlusions, environmental clutter or loss of detection, by computing the percentage of the frames when the algorithm successfully tracked both legs to the total recorded frames (frames were hand-annotated), and *iii) the gait state classification* given the extracted ground truth gait states by assessing three measures. *a) The classification accuracy*: the quotient of the correct predictions to the total number of predictions and is a measure of how good the proposed model is, *b) the recall*: RE is a measure of how many actual positive observations were predicted correctly, *c) the precision*: PR is a measure of how many positive predictions were actual positive observations, and *d) the F1-score*: F1 is the weighted average of precision and recall,

$$RE = \frac{TP}{(TP+FN)}, PR = \frac{TP}{(TP+FP)}, F1 = \frac{2(RE \cdot PR)}{(RE + PR)} \quad (20)$$

where TP: True Positives, FN: False Negatives and FP: False Positives. We present a sensitivity analysis of the above metrics

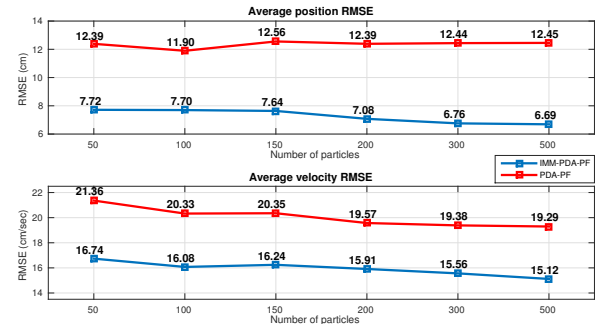


Figure 5: Tracking accuracy as the position and velocity average RMSE according to different number of particles of the proposed IMM-PDA-PF and the baseline PDA-PF w.r.t. GT.

w.r.t. the number of samples used in the PFs, to evaluate how the performance of the algorithm is affected by the number of particles. In this work, we evaluate the algorithm's performance for a range of particles from 50 to 500 samples. To evaluate the performance of the proposed method with respect to other state of the art approaches, we choose to compare the IMM-PDA-PF method with a recently published one, proposed in [24], which we here denote as the *baseline PDA-PF method*. We choose this method, following an extensive literature survey, as being the most relevant one for comparative evaluation purposes. This baseline method also employs PF tracking of legs using on-board LRF data, but uses a single motion model without any IMM-based human state estimation. We comparatively evaluate the performance of the two methods regarding their tracking accuracy and robustness for the same range of particles sets. We also combine the baseline PDA-PF tracking method with the human gait state estimate of Section III-B (i.e. without the proposed IMM framework), and provide comparative classification results on the aforementioned testing dataset (in terms of accuracy and F1-score only, due to space limitations).

C. Validation Results and Discussion

The validation results regarding the tracking accuracy are depicted in Fig. 5. The blue lines depict the average (position and velocity) RMSE of the proposed IMM-PDA-PF method, for an increasing number of particles, while the red lines on the same figures depict the respective comparative results of the baseline PDA-PF method. Inspecting the results of Fig. 5, we can deduce that, as far as the position tracking accuracy is concerned, the proposed IMM-PDA-PF method achieves significantly better results than the baseline PDA-PF method, for all sets of particles. In particular, when 50 particles are used the average position tracking RMSE for the IMM-PDA-PF is decreased by 38% (as compared to the baseline method) and this significant decrease is consistent for all different sizes of particle sets, achieving a value of 46% improvement in tracking accuracy for 500 particles (average RMSE of 6.69cm for the proposed method over 12.45cm for the baseline method). In average, over all particle sets of different sizes, the IMM-PDA-PF method achieves an ameliorated position tracking accuracy with a mean value of approximately 41% performance gain.

Table II: Comparative robustness evaluation

	particles					
Method	50	100	150	200	300	500
IMM-PDA-PF	98.49	99.83	100.00	100.00	100.00	100.00
Baseline PDA-PF	86.72	90.53	90.58	91.00	92.12	93.84

Robustness evaluation of the proposed IMM-PDA-PF vs. the baseline PDA-PF method according to different number of particles.

Regarding the velocity RMSE, the IMM-PDA-PF again achieves better results than the baseline approach for all particle sets, achieving an average RMSE decrease of approximately 20% w.r.t. the baseline. Generally, the relatively higher errors in velocity estimation can be explained by the computation of the GT velocity (simple differentiation of markers' positions) that induces random noise; also the laser clusters are highly deformable due to clothing further adding considerable noise to the observations. The particles' propagation motion model described in Eq. (14), favours the position estimation convergence; hence a higher order kinematic model should be used to minimize estimation error in velocity.

Table II presents a comparative evaluation of the robustness performance of the proposed IMM-PDA-PF method in contrast to the baseline PDA-PF tracking algorithm. Evaluating the cumulative results from the eight subjects of the test set, we can observe that even with as few as 50 particles the IMM-PDA-PF method achieves highly robust tracking performance (98.49% success rate), while for 150 particles or more there are apparently no false detections or tracking failures (100% tracking success rate). On the other hand, the baseline PDA-PF method achieves a mediocre performance (in terms of robustness) when using a small set of 50 particles (a low 86.72% tracking success rate), while even with 500 particles it still exhibits considerable failure rate, tracking successfully only 93.84% of all the frames of the eight subjects. The results of Table II clearly show that the proposed tracking algorithm can efficiently handle problems of leg occlusions and environmental clutter, i.e. noise or another person's legs in the proximity of the tracked person, using much fewer particles than the baseline PDA-PF. It suffices to mention that even in cases when tracking losses occurred (sample sizes of 50 or 100 particles), the algorithm, due to the PDA and the motion model selection from the IMM, managed in all cases to quickly re-detect the subject's legs and continue the tracking process.

Table III presents the classification results for the human gait state estimation, regarding the four states of the Markov model in Fig. 2. We aim to evaluate the ability of the proposed IMM scheme to reliably predict the human gait phases in real-time, hence providing an augmented human state estimation. Inspecting the results, we can notice high mean accuracies (over 90%) for the proposed IMM-PDA-PF implementations of 100 particles and over, with a peak value of 94% accuracy for 500 particles. Embarking on the classification problem, we have to check the recall and precision results. All recall results are approximately over 70% with maximum recall 88% at 300 and 500 particles. Regarding precision, the results are over 60% for all implementations, presenting an ascending pattern w.r.t. the increasing number of particles, reaching an 80% precision at 500 particles. It is also important to observe

and separately analyse the results for individual classes (i.e. gait phases), since it seems that gait states s_1 and s_4 are more easily recognized than states s_2 and s_3 ; indeed, the last two refer to the DS phases that correspond to very short time intervals and are hence difficult to capture.

The F1-score helps assessing the overall model performance. According to Table III the tracking implementations with 150 particles or more achieve F1-scores over 70%, with the best score of 82% classification success achieved for 500 particles. For comparison, Table IV depicts the classification results (accuracy and F1-scores) achieved when the baseline PDA-PF tracking algorithm is used (instead of the proposed IMM scheme) in combination with the human gait state estimation. It is evident that both the accuracy and F1-scores are much lower than the proposed method; e.g. for 500 particles the IMM-PDA-PF achieves 51% better classification performance (according to the F1-score) than the baseline. All these results lead to the conclusion that the gait IMM-PDA-PF scheme, proposed in this paper, is a robust method that can provide accurate and real-time augmented human gait state estimation.

Summing up, we can claim that the IMM-PDA-PF succeeded to perform accurate and robust tracking of the humans legs with a small number of particles, in contrast to the baseline PDA-PF, i.e. thus reducing the computational load for a real-time implementation. The contribution of the IMM in the selection of the motion model during gaiting is particularly related to the fewer number of particles. We present strong evidence that the proposed methodology can provide efficient human gait state estimates with small sets of particles. Even with as few as 150 particles the proposed algorithm can achieve efficient augmented human state estimation, constituting a novel approach that provides real-time tracking of humans with various mobility impairments and a potential tool for on-line gait analysis and identification. It is however important to augment the training dataset with more pathological gait data and to test new GMMs in order to achieve better classification results.

V. CONCLUSIONS & FUTURE WORK

We introduce a novel gait tracking algorithm for an augmented human state estimation that uses two PFs to track the legs and PDA to associate them, an IMM for the selection of the appropriate motion model and a gait state estimation using the Viterbi algorithm. The human gait state estimate drives the selection of the motion model from the Markov model of the gait IMM. We utilize data from an LRF mounted on a robotic assistant platform designed for mobility impaired subjects, constituting a non-invasive approach using a non-wearable device. We validate the performance of our human gait tracking algorithm regarding its accuracy, robustness to occlusions and clutter, and its classification abilities according to different number of particles. The GT data used in this work were extracted from visual markers.

The experimental results show that the proposed gait tracking method has the ability to track users with various mobility impairments, providing accurate and robust estimates even with a small number of particles, which is very important for a

Table III: Classification results (IMM-PDA-PF)

Accuracy							
gait state	particles	50	100	150	200	300	500
	s_1		89.78	91.73	93.92	94.16	93.45
s_2		86.37	89.78	89.54	90.02	94.78	94.12
s_3		88.56	89.29	88.08	87.35	92.31	98.04
s_4		88.08	90.75	92.46	91.48	90.88	94.12
mean		88.20	90.39	91.00	90.75	92.88	94.12
Recall							
gait state	particles	50	100	150	200	300	500
	s_1		60.87	78.26	89.29	87.10	88.00
s_2		84.62	86.75	84.57	91.28	89.86	77.27
s_3		57.89	61.54	55.88	51.02	94.74	100.00
s_4		75.26	79.78	83.33	80.77	80.50	95.00
mean		69.66	76.58	78.27	77.54	88.28	88.07
Precision							
gait state	particles	50	100	150	200	300	500
	s_1		29.79	38.30	53.19	57.45	52.38
s_2		80.49	87.80	90.24	82.93	97.79	94.44
s_3		41.51	45.28	35.85	47.17	40.91	80.00
s_4		99.32	99.32	98.64	100.00	99.22	100.00
mean		62.78	67.68	69.48	71.89	72.58	79.72
F1-score							
gait state	particles	50	100	150	200	300	500
	s_1		40.00	51.43	66.67	69.23	65.67
s_2		82.50	87.27	87.32	86.90	93.66	85.00
s_3		48.35	51.27	43.68	49.02	57.14	88.89
s_4		85.63	88.48	90.34	89.36	88.89	97.44
mean		64.12	69.61	72.00	73.63	76.34	82.12

Gait state classification results (accuracy, recall, precision, and F1-score) using the IMM-PDA-PF method with an increasing number of particles.

Table IV: Classification results (PDA-PF)

Accuracy							
gait state	particles	50	100	150	200	300	500
	s_1		77.69	78.81	79.16	79.11	77.10
s_2		76.56	79.11	76.41	80.65	75.48	78.72
s_3		74.69	79.50	77.93	77.15	74.99	76.66
s_4		76.12	75.58	77.93	76.75	75.43	76.66
mean		76.26	78.25	77.81	78.42	75.75	77.88
F1-score							
gait state	particles	50	100	150	200	300	500
	s_1		43.90	45.86	42.18	48.43	39.27
s_2		63.17	67.58	65.81	69.17	62.06	65.06
s_3		38.42	52.06	48.86	45.81	43.58	47.46
s_4		56.94	54.48	57.01	57.45	53.91	59.25
mean		50.61	55.00	53.47	55.22	49.71	54.23

Gait state classification results (accuracy and overall F1-score) for the baseline PDA-PF method with an increasing number of particles.

real-time application. The classification results seem promising and could be incorporated in a real-time gait analysis system for a user-adaptive context aware robot controller.

Our ongoing research comprises the development of different motion models that would stochastically describe different classes of mobility impairment, an on-line user classification according to the selected motion model and specific gait features, the adaptation of the number of particles according to the type of walking and the real time implementation and evaluation of the algorithm.

REFERENCES

[1] O. Brdiczka et al., "Detecting human behavior models from multimodal observation in a smart home," *IEEE Trans. on Automation Science and Engineering*, 2009.
 [2] J. Pineau et al., "Towards robotic assistants in nursing homes: Challenges and results," *Robotics and Autonomous Systems*, 2003.
 [3] T. Fong et al., "A survey of socially interactive robots," *Robotics and Autonomous Systems*, 2003.

[4] J. Zhou and J. Hoang, "Real time robust human detection and tracking system," in *IEEE Conf. on Computer Vision and Pattern Recognition-Workshops*, 2005.
 [5] D. Schulz et. al, "Tracking multiple moving targets with a mobile robot using particle filters and statistical data association," in *IEEE Int'l Conf. on Robotics and Automation*, 2001.
 [6] H. Zhao and R. Shibasaki, "A novel system for tracking pedestrians using multiple single-row laser-range scanners," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 35, no. 2, pp. 283–291, March 2005.
 [7] G. Taylor and L. Kleeman, "A multiple hypothesis walking person tracker with switched dynamic model," in *Proc. of the Australian Conf. on Robotics and Automation*, 2004.
 [8] N. Bellotto and H. Hu, "Vision and laser data fusion for tracking people with a mobile robot," in *2006 IEEE International Conference on Robotics and Biomimetics*, Dec 2006, pp. 7–12.
 [9] K. O. Arras, S. Grzonka, M. Luber, and W. Burgard, "Efficient people tracking in laser range data using a multi-hypothesis leg-tracker with adaptive occlusion probabilities," in *2008 IEEE International Conference on Robotics and Automation*, May 2008, pp. 1710–1715.
 [10] M. Luber, G. D. Tipaldi, and K. O. Arras, "Better models for people tracking," in *2011 IEEE International Conference on Robotics and Automation*, May 2011, pp. 854–859.
 [11] H. Kim, W. Chung, and Y. Yoo, "Detection and tracking of human legs for a mobile service robot," in *2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, July 2010, pp. 812–817.
 [12] T. Ohnuma, G. Lee, and N. Y. Chong, "Particle filter based feedback control of jaist active robotic walker," in *2011 RO-MAN*, July 2011, pp. 264–269.
 [13] W. H. Mou, M. F. Chang, C. K. Liao, Y. H. Hsu, S. H. Tseng, and L. C. Fu, "Context-aware assisted interactive robotic walker for parkinson's disease patients," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2012, pp. 329–334.
 [14] M. Martins, A. Frizzera, R. Ceres, and C. Santos, "Legs tracking for walker-rehabilitation purposes," in *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechanics*, Aug 2014, pp. 387–392.
 [15] D. Lee et al., "Parallel interacting multiple model-based human motion prediction for motion planning of companion robots," *IEEE Trans. on Automation Science and Engineering*, vol. 14, no. 1, 2017.
 [16] S. McGinnity and G. W. Irwin, "Multiple model bootstrap filter for maneuvering target tracking," *IEEE Trans. on Aerospace and Electronic Systems*, 2000.
 [17] C. Kreucher, K. Kastella, and A. O. Hero, "Multitarget tracking using the joint multitarget probability density," *IEEE Trans. on Aerospace and Electronic Systems*, 2005.
 [18] X. Wang and V. L. Syrmos, "Interacting multiple particle filters for fault diagnosis of non-linear stochastic systems," in *2008 American Control Conference*, 2008.
 [19] X. Papageorgiou et. al, "Experimental validation of human pathological gait analysis for an assisted living intelligent robotic walker," in *Biorob 2016*.
 [20] G. Chalvatzaki et. al, "Towards a user-adaptive context-aware robotic walker with a pathological gait assessment system: First experimental study," in *IROS 2017*.
 [21] Z. Wu, Y. Huang, L. Wang, X. Wang, and T. Tan, "A comprehensive study on cross-view gait based human identification with deep cnns," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 39, no. 2, pp. 209–226, Feb 2017.
 [22] J. Perry, "Gait Analysis: Normal and Pathological Function", 1992.
 [23] L. R. Rabiner, "Readings in speech recognition," A. Waibel and K.-F. Lee, Eds. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1990, ch. A tutorial on hidden Markov models and selected applications in speech recognition, pp. 267–296.
 [24] G. Chalvatzaki et. al, "Comparative experimental validation of human gait tracking algorithms for an intelligent robotic rollator," in *ICRA 2017*.
 [25] M. Arulampalam et. al., "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Trans. Signal Processing*, 2002.
 [26] A. Tchango et. al., "Tracking multiple interacting targets using a joint probabilistic data association filter," in *FUSION 2014*.
 [27] Z. Chen, "Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond," McMaster University, Tech. Rep., 2003.
 [28] G. Chalvatzaki et. al, "Estimating double support in pathological gaits using an hmm-based analyzer for an intelligent robotic walker," in *RO-MAN 2017*.