Chest-Mounted Inertial Measurement Unit for Pedestrian Motion Classification Using Continuous Hidden Markov Model

Ghazaleh Panahandeh, Nasser Mohammadiha, Arne Leijon, and Peter Händel
KTH Royal Institute of Technology, School of Electrical Engineering
Stockholm, Sweden
Email: {ghpa, nmoh, leijon, ph}@ee.kth.se

Abstract—This paper presents a method for pedestrian motion classification based on MEMS inertial measurement unit (IMU) mounted on the chest. The choice of mounting the IMU on the chest provides the potential application of the current study in camera-aided inertial navigation for positioning and personal assistance. In the present work, five categories of the pedestrian motion including standing, walking, running, going upstairs, and going down the stairs are considered in the classification procedure. As the classification method, the continuous hidden Markov model (HMM) is used in which the output density functions are assumed to be Gaussian mixture models (GMMs). The correct recognition rates based on the experimental results are about 95%.

I. Introduction

Micro-electromechanical system (MEMS) inertial sensors, as a lightweight and cheap motion capture system, are becoming as a standard feature of smart-phones and personal digital assistance for navigation and positioning applications. MEMS accelerometers and gyroscopes are the fundamental and primary motion capture sensors. MEMS inertial measurement units (IMUs) generally consist of three orthogonal gyroscopes and three orthogonal accelerometers. By integrating linear accelerations and rotational velocities provided by the IMU, a navigation system is capable of tracking the position, velocity, and attitude of a moving object. However, due to the integration drift, the position and the attitude must be periodically corrected, otherwise the system cannot be used over extended periods of time. The correction can be provided by adding the IMU sensor with a complementary sensor such as global positioning system (GPS) receiver for outdoor applications, vision sensors, or laser scanners mainly for indoor applications. Independent of the environment, the IMU sensor can additionally provide human motion information such as human gait or step detection [1], [2]. Foot-mounted pedestrian navigation, as a self-contained IMU inertial navigation system, is becoming a promising approach for both indoor pedestrian navigation and ambulatory gait analysis [1], [3]. To limit the positioning error accumulation, the estimated velocity is periodically updated to zero when the foot is flat on the ground plane. When the zero velocity updates (ZUPT) method is threshold-based both for the sensor values and time duration [3], the threshold values could differ significantly from walking to running, i.e., the IMU sensor values become larger,

Fig. 1: Designed Chest-mounted IMU (Chem-IMU). The black box indicates the IMU sensor mounted on the chest for a sample participant. The drawing axis on the IMU illustrates approximately the direction of the accelerometers and gyroscopes output signals.

and the zero velocity intervals become shorter in running. That is, ZUPT methods are relatively good only in walking scenarios. Hence, human activity recognition is an important issue not only in biomedical phenomena analysis but also in navigation applications. Additionally, foot-mounted ZUPT approaches require a fixed IMU mounted on the shoe that restricts the navigation solutions to be aided with additional sensors such as camera; since such combinations requires the sensor-to-sensor calibration.

Not to lose the cyclic nature of the human walking, and to avoid capturing additional forces which do not contribute in walking patterns' recognitions, we propose to mount the IMU on the chest as an alternative to the foot-mounted IMU; see Fig 1. Exploring the chest-mounted IMU (Chem-IMU) signals shows that these signals can be effectively analyzed to distinguish different human gaits. Additionally, the potential application of the vision-aided inertial navigation for pedestrian navigation can be provided by such structure [4]. In reality for pedestrian inertial navigation systems, we are dealing with human beings different activities such as standing, walking, running, etc. Except in normal walking or standing, the other motions may cause motion blurred images; so, the motion state of the system estimated trough the images cannot be trustable. To avoid such situations, the state of the moving subject needs to be feed-backed to the system. The prior knowledge about the pedestrian motion, which can be extracted from the IMU,
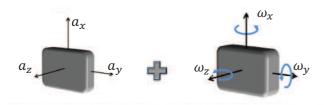
Fig. 2: Inertial measurement units (IMUs) consist of three orthogonal accelerometers and three orthogonal gyroscopes.
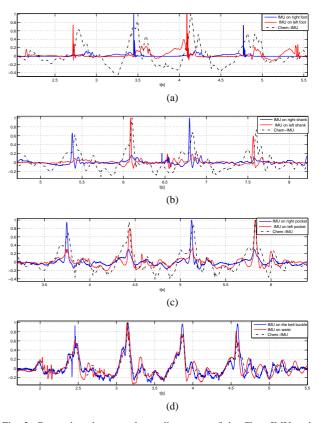

(a)

(b)

(c)

(d)

Fig. 3: Comparison between the cyclic pattern of the Chest-IMU and other body mounted IMUs: foots (a), shanks (b), pockets (d), on the waist close to L3 vertebra and the belt buckle(d).

can significantly improve the performance of the navigation solution. Especially if it can be aided with the step information, such as step frequency or step length information.

In this paper, we look at the problem of human motion classification based on data provided by a Chem-IMU. Although, number of researches have been devoted to human motion detection through discrete Hidden Markov Model (HMM) [5], [6], [7], rather less attention has been paid to the continuous HMM structures. Since the accelerations and rotations have continuous-valued representations in time-frequency domain, we propose to use a continuous HMM (HMM with continuous output density functions) rather than a discrete HMM. Gaussian mixture models (GMMs) are powerful in modeling any desired continuous distribution and are used for example in speech processing successfully. Therefore, in this paper HMM with GMMs as output density functions is trained and used for human motion classification.

The paper is organized as follows. The signal and sensor model of the IMU is described in Section II. The problem description and the proposed classification method is presented in Section III. Experimental validation and discussion are given in Section IV. Finally, the conclusion of the study is summarized in Section V.

II. SIGNAL AND SENSOR MODEL

The detection of the pedestrian gait, as mentioned before, has been done using the measurements from the cluster of three orthogonal gyroscopes and three orthogonal accelerometers cluster of the IMU mounted on the chest. Independent of the position of the IMU, the output measurement signals of the accelerometer $\mathbf{a}_k \in \mathbb{R}^3$ and the gyroscope $\omega_k \in \mathbb{R}^3$ clusters of the IMU, Fig. 2, are modeled as:

$$\mathbf{y}_k = \mathbf{x}_k + \mathbf{n}_k \in \mathbb{R}^6 \qquad (1)$$

where

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{a}_k \\ \omega_k \end{bmatrix}, \qquad\qquad \mathbf{n}_k = \begin{bmatrix} \mathbf{n}_k^{\mathbf{a}} \\ \mathbf{n}_k^{\omega} \end{bmatrix},$$

$\mathbf{n}_k^{\mathbf{a}} \in \mathbb{R}^3$ and $\mathbf{n}_k^{\omega} \in \mathbb{R}^3$ are noise vectors associated with the acceleration and angular rates, respectively, and k is the corresponding sample index. To describe complete characteristics of the motion (translation and rotation), both the accelerometers and the gyroscopes data are considered for the classification purpose.

Although the primary goal of this work is pedestrian motion classification through a wearable IMU, we are also interested in the potential application of the classifier when the IMU is aided with a secondary vision system for inertial navigation. As a result, IMU is considered to be placed on the chest. However, the proposed method is not restricted to the Chem-IMU setup, and it can be used for different placements of the IMU on the human body as far as the IMU signals present a cyclic nature corresponding to the mentioned human motions. Fig. 3 represents such cyclic patterns for different IMU placements. In this figure, the vertical axises indicate the normalized norm of the accelerometer output signals. Four different experiments were done in which the Chem-IMU signals were compared with the other IMU signals placed on different parts of the subject's body as follows: feet (a), shanks (b), pockets (c), on the waist close to L3 vertebra and on the belt buckle(d). In the normal walking scenario, each step is due to the movement of one of the legs while the other foot touches the ground. Hence to present totally four steps in Fig. 3 (a), (b), and (c) two IMUs were placed symmetrically on the right and left legs.

III. CHEM-IMU BASED PEDESTRIAN GAIT DETECTION

A. Problem Description

Applying the periodicity of the human walk, captured by the Chem-IMU, the pedestrian gait can be analyzed despite

the variation of the periodicity. In the proposed method, the following motion states for classification are considered:

- Standing, no rigid body acceleration. This state considers the case when the user is almost standing, however, no restriction for the rotation is considered. In other words:

$$\frac{1}{L}\sum_{k=1}^{L}\|\mathbf{a}_k\| \approx g, \tag{2}$$

where L is the length of the analysis window, g is the magnitude of the local gravity vector, and $\|\cdot\|$ indicates the norm operation.
- Walking.
- Running.
- Going down the stairs.
- Going upstairs.

Although only five category of the human physical activities are considered in this study, other physical activities such as bending, sitting down, standing up can also be considered as additional states in the proposed classification approach.

B. Proposed Solution

To correctly identify different actions from the IMU signals, it is essential to find proper transformations that are capable of describing specific nature of each action. That is the detection algorithm must be capable of dealing with noise as well as biases in the input signals. Therefore, the detection procedure mainly consists of three steps: 1) pre-processing 2) feature extraction, and 3) classification. These steps are briefly described in the following.

1) Pre-processing: The pre-processing module of the system is a band-pass filtering which reduces the measurement noise and removes the effect of IMU displacement on the chest for different subjects. The band-pass filter is a combination of a standard low-pass filter and a DC level cancelation.

2) Feature Extraction: Feeding the classifier with sufficient extracted features from the input signals is a critical stage in any pattern recognition problem. Therefore, the input signals must be properly mapped to another domain which can best represent the statistics of different patterns.

Time-frequency domain analysis is a powerful tool to study signals with different frequency behaviors for which this behavior is time variant. For human motion, the IMU output signals are not subject to major change within short windows up to $1 \sim 2$ seconds. Hence, we apply a discrete Fourier transform (DFT) to the signals with overlapped time-frames. The time-frame length should be chosen such that the signal can be considered stationary within each frame. The overlap length influences the update frequency of the classification result. Furthermore, a proper DFT length has to be considered which allows a good frequency resolution. The segmentation of the signal into short frames is done with multiplying a Hann window to the signal to reduce the edge effects of the windowing. N number of samples from the highest DFT coefficients of the IMU output signals are transferred to log-spectrum domain (by taking the logarithm of the magnitude-squared DFT coefficients). Then, the detected features for each

of the IMU signals are concatenated to make a $6N$-length feature vector for the current time-frame, that is represented by $\mathbf{z}_l$ in the following. The feature vector could include different subsets of features that come from different sources of information. These features can be a strong identifier specially when the energy expenditures are significantly different for instance from standing to walking and running. Since, the feature vector could include different subsets of features that come from different sources of information; the mean and the variance of the signals in each segments are explicitly concatenated in the feature vectors.

3) Classification: We model the 'human gait' with a continuous HMM; hence, a HMM which has five states, corresponding to five different motion states, is applied to classify the obtained feature vectors. The classification is done by applying the forward algorithm [8] to find the probability for each of the aforementioned states given all the current and previous output signals from the IMU, i.e., $p_{S_l|\mathbf{z}_1^l}\left(S_l = i \mid \mathbf{z}_1^l\right)$ where S_l denotes the hidden state in time-frame l, and $\mathbf{z}_1^l$ denotes the feature vectors measured from time-frame 1 till l, i.e. $\mathbf{z}_1^l = \{\mathbf{z}_1, \dots \mathbf{z}_l\}$. The most probable state can be considered as the output of the detection procedure.

C. Training HMM

As mentioned earlier, because of the continuous nature of the signals in the time-frequency domain, the state-conditional distributions are assumed to be GMM as:

$$f_{\mathbf{z}_l|S_l}(\mathbf{z}_l \mid S_l = i) = \sum_{m=1}^{M} w_{m,i} \mathcal{N}(\mathbf{z}_l; \mu_{m,i}, C_{m,i}), \tag{3}$$

where $\mathcal{N}(\mathbf{z}_l; \mu_{m,i}, C_{m,i})$ represents a multivariate Gaussian distribution with mean vector $\mu_{m,i}$ and covariance matrix $C_{m,i}$. Moreover, $\sum_{m=1}^{M} w_{m,i} = 1$.

The HMM training stage consists of estimating the HMM parameters $\lambda = \{A, q, \mu, C, w\}$ where A is the transition probability matrix with $a_{i,j} = p(S_{l+1} = j \mid S_l = i)$, and q is the initial state probability distribution with $q_j = p(S_1 = j)$. For this purpose, Baum-Welch algorithm [9] is used, which is an Expectation Maximization (EM) algorithm. The EM algorithm [10] is an iterative procedure to obtain the Maximum Likelihood (ML) estimates of the distribution-parameters that is carried on by iteratively maximizing the log-likelihood of the observations until a local maximum is archived. The EM approach is usually used when the optimization problem is difficult using only the observed variables. Hence, the so called hidden variables are considered to make the optimization tractable. In the current problem, the hidden variables are considered to be the hidden states of the HMM, S, and the GMM weights, w. To avoid the mentioned locally-optimality, usually the procedure is repeated for some different initial values, and the best result is kept at the end. In the current application, it makes sense to predefine the transition probability matrix A based on some sensible values of the tendency of humans to switch between different states. In principle, this matrix could be also obtained provided a lot of training data captured from many people's daily life; however, this is not

in the scope of the current study, and we do not consider it here. Hence, the Baum-Welch algorithm is simplified to estimate the GMM parameters given the training data for different states separately. Performing identical pre-processing and time-frequency feature extraction steps, as done for the classification, an EM algorithm is then applied to obtain w, μ, C separately for each state [9].

IV. EXPERIMENTAL VALIDATION AND DISCUSSION

The primary focus of this paper is to design and evaluate the performance of a Chem-IMU classification procedure. However, the proposed method is not dependent on the position of the IMU and can be applied in other structures such as foot-mounted IMU. Considering different feature extraction methods such as time domain features, frequency features, and time scale features, [11], the log-spectrum DFT coefficients and the signal mean and variance are used as the feature vectors. But as the ultimate goal of this work, we will investigate different time-frequency features for online applications. For instance, Fig. 4 provides an example of the obtained probabilistic descriptions of the extracted features for IMU signals. In this figure, the mean value of the GMM components, 2-dimensional GMM, for three different states including standing, walking, and running are presented. The length of the feature vectors is 60 (the first 10 log-spectrum DFT coefficients for each of the IMU outputs' signals).

To evaluate the performance of the proposed Chem-IMU pedestrian motion classification, users were asked to connect the IMU to their chest as illustrated in Fig. 1. A MicroStrain 3DM-GX2 IMU has been used in the experiments, which converts the raw voltage outputs of the motion sensors into digital format (in terms of A/D converter codes) with sampling rate of 250 Hz. The sensor module is directly connected to a Laptop (carried by the participants) via a USB port where only the three axis accelerometer and angular rate outputs of the sensor unit are used. The accelerometer output signals are then re-scaled into physical units of g ($1g = 9.80665m/s^2$). For the clarity, the DC level of the signals are kept in the following figures, although in the pre-processing module of the system the DC level of the signals are removed.

In order to obtain sufficient amount of data for training, ten subjects–three females and seven males of the hight of 165cm to 200cm–attended in the experiment. Independent of the position of the IMU on the chest, which affects the DC level of the signals, similar patterns were observed in the IMU output signals.

Fig. 5 shows the six IMU signals, three accelerations $[a_x, a_y, a_z]$ and three angular rotations $[\omega_x, \omega_y, \omega_z]$, when the user's state starts from walking to running and ends up with walking. As it can be seen two different states, with green dashed line, are detected in this test: in the beginning, the subject is walking and the detected state is depicted with level two; the next state is detected as running, depicted with level three ; finally, the last state that is walking is again detected with level two.
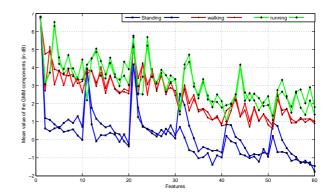

Fig. 4: The mean value of the 2-dimensional GMM components for three states: standing, walking, and running. Every ten features belong to a corresponding IMU output signals, $N = 10$.
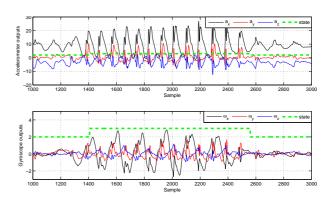

Fig. 5: IMU output signals, accelerometers and gyroscopes, when subject starts running from walking and then slows down the speed of running to walking and the corresponding three detected states are plotted over the both accelerometers and gyroscopes signals by green dashed line; walking: level two and running: level three.

Accordingly, Fig. 6 provides the detected states, green dashed line, over the six IMU signals, three accelerations $[a_x, a_y, a_z]$ and three angular rotation $[\omega_x, \omega_y, \omega_z]$. The user starts at the standing state, where the detected state is depicted with level one; then going down the stairs that is depicted with level four; then turning back which is detected as standing state; finally walking up the same stairs,which is depicted with level five.

Moreover, a supervised classification was done to evaluate the correct detection rate of the proposed classification approach. In this experiment, three different subjects were asked to perform the defined motions according to TABLE I. The average correct detection rate of the subjects for each of the motions are also reported in the TABLE I, which indicates promising performance, a probability of about 95% correct detection rate for the considered five states: standing, walking, running, walking up and going down the stairs. In practice, the correct detection rate must be better than such a number, since in the test, switching between the states were happening very fast. Experience show that more than 80% of the detected errors are happening in the transition from standing to the walking, running, going down the stairs and vice-versa, as it is shown in Fig. 7. The transition states which are incorrectly
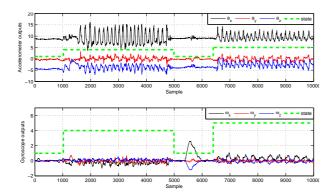
Fig. 6: IMU output signals, accelerometers and gyroscopes, when subject going downs the starts from standing state and then turns back to go up the same stairs the same stairs. Three different detected states are plotted over the both accelerometers and gyroscopes signals by green dashed line; standing: level one, going down the stairs: level four, and going upstairs: level five.

TABLE I: Average success rate for the five defined motion states .

Gait pattern	Standing	Walking	Running	Going down the stairs	Going upstairs
Average number of samples	≈2500	≈10000	≈3500	≈3100	≈3100
Average success rate	≈94%	≈93%	≈95%	≈99%	≈95%

detected are mainly due to the body configuration to start from one state to another state that is mainly detected as going upstairs; see Fig. 7.

V. CONCLUSION AND FUTURE WORK

In this paper, a Chem-IMU model has been proposed for pedestrian motion classification. Five different states are considered including standing, walking, running, going upstairs, and going down the stairs. Considering the continuous nature of the measured accelerations and rotations, a continuous HMM with GMM output density functions has been used in the classification procedure to analyze pedestrian motion. The performance of the proposed method has been evaluated by experimental data using the trained HMM model. The

probability of correct detection rate based on our experiments are about 95% that provides a promising application for navigation and personal digital assistance. Our future work in this area will include real time activity classification through studying the short segment of the IMU signals. Considering the cyclic nature of the step patterns, a circular HMMs can describe such patterns within one activity cycle.
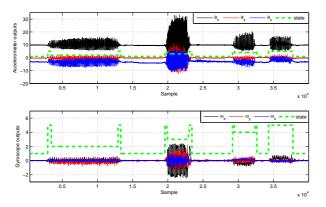

Fig. 7: IMU output signals, accelerometers and gyroscopes, for all the five defined motion states together with the detected states levels: standing: one, walking: two, running: three, going down the stairs: four, and going upstairs: five. The test is starting with standing → walking → standing → running → standing → going down the stairs → standing → going upstairs → standing.

REFERENCES

[1] Ö. Bebek, M. A. Suster, S. Rajgopal, M. J. Fu, X. Huang, M. C. Cavusoglu, D. J. Young, M. Mehregany, A. J. Van Den Bogert, and C. H. Mastrangelo, "Personal navigation via high-resolution gait-corrected inertial measurement units," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 11, pp. 3018–3027, 2010.

[2] H. Zhou and H. Hu, "Reducing drifts in the inertial measurements of wrist and elbow positions," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, pp. 575–585, Mar. 2010.

[3] E. Foxlin, "Pedestrian tracking with shoe-mounted inertial sensors," *IEEE Computer Graphics and Applications,*, vol. 25, pp. 38–46, Nov.-Dec. 2005.

[4] G. Panahandeh and M. Jansson, "Vision-aided inertial navigation using planar terrain features," in *IEEE Int. Conf. on Robot, Vision and Signal Processing*, Sep. 2011.

[5] S. K. Park and Y. S. Suh, "A zero velocity detection algorithm using inertial sensors for pedestrian navigation systems," *Sensors*, vol. 10, no. 10, pp. 9163–9178, 2010.

[6] W. Wan, H. Liu, L. Wang, G. Shi, and W. Li, "A hybrid HMM/SVM classifier for motion recognition using IMU data," in *IEEE Int. Conf. on Robotics and Biomimetics*, pp. 115–120, Dec. 2007.

[7] G. Shi, Y. Zou, Y. Jin, X. Cui, and W. Li, "Towards HMM based human motion recognition using MEMS inertial sensors," in *IEEE Int. Conf. on Robotics and Biomimetics ROBIO,*, pp. 1762–1766, Feb. 2009.

[8] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, pp. 257–286, Feb. 1989.

[9] J. Bilmes, "A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models," tech. rep., Univ. California, Berkeley, TR-97-021 Rep, 1997.

[10] A. P. Dempster, N. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Statist. Soc*, vol. 39, no. 1, pp. 1–38, 1977.

[11] M. Asgharioskoei and H. Hu, "Myoelectric control systemsa survey," *Biomedical Signal Processing and Control*, vol. 2, no. 4, pp. 275–294, 2007.