

Received 19 December 2013; revised 15 August 2014; accepted 7 October 2014. Date of publication 30 October 2014;
date of current version 14 November 2014.

Digital Object Identifier 10.1109/JTEHM.2014.2365773

Automatic Detection and Classification of Unsafe Events During Power Wheelchair Use

JOELLE PINEAU¹, ATHENA K. MOGHADDAM^{1,2}, HIU KIM YUEN¹, PHILIPPE S. ARCHAMBAULT³,
FRANÇOIS ROUTHIER⁴, FRANÇOIS MICHAUD⁵, AND PATRICK BOISSY⁶

¹School of Computer Science, McGill University, Montréal, QC H3A 0G4, Canada

²Facebook Inc., Menlo Park, CA 94301, USA

³School of Physical and Occupational Therapy, McGill University, Montréal, QC H3A 0G4, Canada

⁴Department of Rehabilitation, Université Laval, Laval, QC G1V 0A6, Canada

⁵Department of Electrical Engineering and Computer Engineering, Université de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada

⁶Department of Surgery, Université de Sherbrooke, Sherbrooke QC J1K 2R1, Canada

CORRESPONDING AUTHOR: J. PINEAU (jpineau@cs.mcgill.ca)

This work was supported in part by the Discovery Grants Program, Natural Sciences and Engineering Research Council of Canada (NSERC), in part by the Fonds de Recherche du Québec Nature et Technologies (FQRNT), Regroupement Stratégique Ingénierie de Technologies Interactives en Réadaptation (INTER), in part by the Catalyst Grant through the Canadian Institutes of Health Research (CIHR)-Institute of Aging, and in part by CIHR through the CanWheel Team.

ABSTRACT Using a powered wheelchair (PW) is a complex task requiring advanced perceptual and motor control skills. Unfortunately, PW incidents and accidents are not uncommon and their consequences can be serious. The objective of this paper is to develop technological tools that can be used to characterize a wheelchair user's driving behavior under various settings. In the experiments conducted, PWs are outfitted with a datalogging platform that records, in real-time, the 3-D acceleration of the PW. Data collection was conducted over 35 different activities, designed to capture a spectrum of PW driving events performed at different speeds (collisions with fixed or moving objects, rolling on incline plane, and rolling across multiple types obstacles). The data was processed using time-series analysis and data mining techniques, to automatically detect and identify the different events. We compared the classification accuracy using four different types of time-series features: 1) time-delay embeddings; 2) time-domain characterization; 3) frequency-domain features; and 4) wavelet transforms. In the analysis, we compared the classification accuracy obtained when distinguishing between safe and unsafe events during each of the 35 different activities. For the purposes of this study, unsafe events were defined as activities containing collisions against objects at different speed, and the remainder were defined as safe events. We were able to accurately detect 98% of unsafe events, with a low (12%) false positive rate, using only five examples of each activity. This proof-of-concept study shows that the proposed approach has the potential of capturing, based on limited input from embedded sensors, contextual information on PW use, and of automatically characterizing a user's PW driving behavior.

INDEX TERMS Assistive technologies, wheelchairs, event detection, accelerometers, rehabilitation robotics.

I. INTRODUCTION

It is well recognized that mobility is an important factor for the social participation and quality of life of individuals. For people who live with locomotor impairments, mobility assistive devices such as manual wheelchairs, power wheelchairs (PW), scooters, and other motorized vehicles can be facilitators of mobility [1]. The increasing prevalence of chronic disease and the changing demographics of our society

will increase the use of mobility assistive devices such as PWs by older adults. Using a PW is a complex task whose efficiency and safety are modulated by factors such as individual capacity, wheelchair driving skills, design and technology features of the PW, environmental considerations, and interaction effects between these factors [2].

Although PW mobility has many potential benefits for users, PW incidents and accidents are not uncommon and

their consequences can be serious [3]–[5]. Users of PWs can have difficulty maintaining a supported seated posture when subjected to external forces [6], [7]. Oscillatory and shock whole body vibration (WBV) can be induced by ordinary obstacles or surfaces that PW users encounter every day [8] and are often linked to speed control of the PW. These incidents can be associated with unsafe behaviours combined with poor PW driving skills [9], [10]. PW driving skills are typically modulated by experience and joystick control [11] but can be improved through training [12], [13].

Few objective outcome measures used in the clinical assessment of PW mobility take into account all the factors influencing PW mobility and safety. PW driving skills are typically evaluated with performance-based measures during standardized driving tasks in controlled or real environments, and graded by an observer who will pass judgement on the individual's skill level [14]. Insufficient research has been published on the ability of these instruments to extrapolate to PW mobility and safety in everyday use. Indeed, while PW driving skills are an important factor in the safe operation of a PW, user behaviour and risk mitigation by the user during everyday use of the PW is as important.

In order to educate PW users on risk mitigation and improve their powered wheelchair driving skills, a better characterization of PW users' driving behaviour is required, in various real and ecological indoor and outdoor settings [15]. Ambulatory real life monitoring approaches of PW use with embedded sensors and datalogging are starting to emerge. Cooper et al. [16] determined the driving characteristics of PW users during unrestricted community activities by using a sensor attached to a PW and a custom-built datalogger. Speed, distance traveled, and the time that each subject's personal wheelchair was being driven were recorded for 24hr/d over approximately five days for each subject. Activity levels among an active group and a group of regular users were compared. Drivers of PW were most active during the afternoon and evening hours. Over the 5-day period of this study, there was little variation in the speed or distance driven per day across subjects with an average daily distance of 1.7 km. Another study by Sonenblum et al. [17] reported PW usage using instrumentation known as the Wheelchair Activity Monitoring Instrument (WhAMI). PW usage was logged electronically, and geo-location and interview data were used to attribute PW use to (1) in the home, (2) not in the home indoors, or (3) outdoors, over a sample of twenty-five non-ambulatory, full-time PW users. Distance wheeled, time spent wheeling, number of bouts, time spent in the wheelchair, and the percentage of time spent wheeling were measured to describe wheelchair use. Overall, most research into PW mobility using embedded sensors and datalogging have reported global metrics on use of PW, such as the total distance and time spent moving each day, and have not taken into account context specific information on PW. To improve our understanding of PW use, its potential inter-relationship with the environment and the driving skills of users, and ultimately its impact on the mobility of older

adults, better outcome measures of PW mobility need to be developed.

Advances in data mining methods combined with the availability of low cost sensing technologies provide new opportunities to automatically classify PW activity, and understand behaviours of PW users in natural environments [18]. In our previous work [20], we presented a method to automatically recognize and categorize five different driving activities during PW use, directly from sensor data. The method trains a support vector machine classifier, using time-delay embedding features of sensor-based data collected from a datalogging platform installed on a PW. The objectives of the present study are to: 1) test the robustness of the events recognition method with a dataset containing a larger spectrum of PW activities (35 instead of 5), 2) assess its performance to automatically classify *safe* vs. *unsafe* activities using a very limited training set (only five examples of each activity).

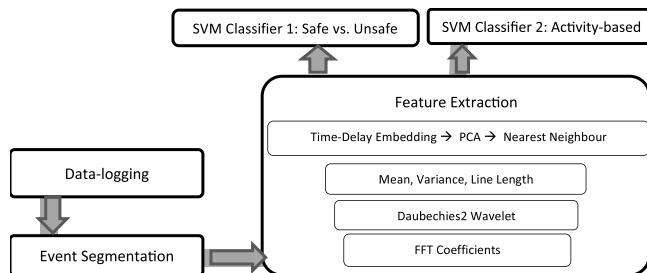


FIGURE 1. Classification process using different features.

II. METHODS AND PROCEDURES

A. OVERVIEW

Fig. 1 provides an overview of the multiple components of our end-to-end system for automatic detection and classification of PW driving activities. First, the data-logging device handles real-time recording of sensor data onboard the PW. Our analysis uses only the 3D accelerometer data. Second, event segmentation is applied to the recorded data to identify events of interest from the data-stream. Next, feature extraction computes a rich set of features from the events of interest. Finally, we apply the machine learning classifier to categorize the events according to type (safe vs. unsafe) and activity (35 different classes).

B. DATA-LOGGING

A proof of concept for a measurement approach to monitor real life use of PW driving was developed and is currently in use with a sample of PW users. The platform, called the Wireless Inertial Measurement Unit with GPS (WIMU-GPS) (Fig. 2), consists of a datalogger with embedded sensors and external sensors installed on the PW. The WIMU-GPS contains a number of sensors and components. In this paper, we consider only the data from the 3D accelerometer. The rationale for this choice is to limit the number of sensor inputs in order to explore the power of the proposed event detection and classification methods under a set of distinct and similar

driving tasks. Accelerometers are inexpensive, reliable, and unobtrusive, thus would be easy to install on any PW. In our case, the WIMU-GPS is installed on the base near the wheel, which takes only a few minutes and minimal tools to install.

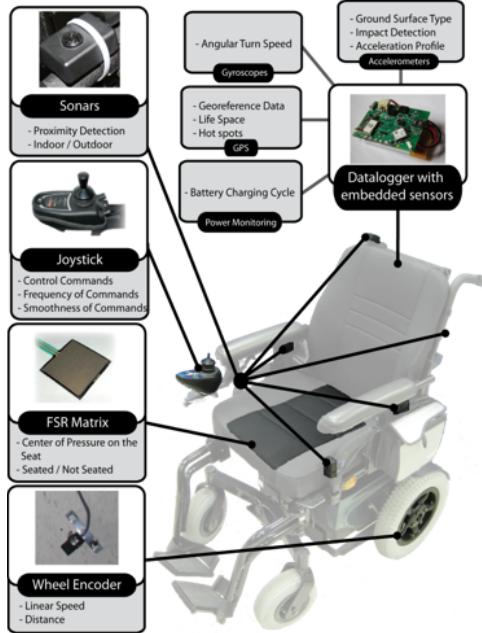


FIGURE 2. Overview of datalogging platform (WIMU-GPS).

Once the datalogging platform has been installed on the PW (an Oasis model from Orthofab), a wheelchair user (in our case a research assistant) is asked to perform 35 distinct activities, listed in Table 1. The activities are selected so that they cover most common wheelchair activities in different speeds, directions, and other conditions, following consultation with clinicians. Of the 35 activities, 15 are identified as being *safe* activities, and 20 are identified as *unsafe* a priori. Unsafe activities are defined as those activities where collisions with objects occurred at different speeds. The remaining activities are defined as the safe activities. The rational for the choice and segmentation of the driving tasks (safe or unsafe) is based on the assumptions that irrespective of the task performed, impacts (collisions) are detrimental to the safety of the user as they have difficulty maintaining a supported seated posture when subjected to external forces. As such, while manoeuvring an PW, collisions can occur from different point of the PW (forward motion, backward motion, side motion) at different speeds. Many of the safe activities selected (such as passing a step at high speed) can potentially be mistaken with the unsafe activities. This is done intentionally to ensure that our event classification process leads to robust classification of activities.

To form the dataset, each activity is repeated 30 times (each of these repetitions is considered an “event”). The 3D acceleration is recorded at the rate of 250 Hz. The 30 events for each activity are then concatenated to form a quasi-periodic signal for each activity. The 30 events are fur-

TABLE 1. PW activities in the dataset.

Activity Code	Activity Description	Type
DP1L	Rolling down 1 inch slope at low speed	Safe
DP1N	Rolling down 1 inch slope at normal speed	
DP1R	Rolling down 1 inch slope at high speed	
FS1L	Passing a step with 1" height and 4" width at low speed	
FS1N	Passing a step with 1" height and 4" width at normal speed	
FS1R	Passing a step with 1" height and 4" width at high speed	
FT1L	Crossing a 1" hole at low speed	
FT1N	Crossing a 1" hole at normal speed	
FT1R	Crossing a 1" hole at high speed	
FT3L	Crossing a 3" hole at low speed	
FT3N	Crossing a 3" hole at normal speed	
FT3R	Crossing a 3" hole at high speed	
MP1L	Climbing up 1" slope at low speed	
MP1N	Climbing up 1" slope at normal speed	
MP1R	Climbing up 1" slope at high speed	
IFCL	Frontal collision with a soft object at low speed	
IFCN	Frontal collision with a soft object at normal speed	
IFCR	Frontal collision with a soft object at high speed	
IFFL	Frontal collision with a fixed object at low speed	
IFFN	Frontal collision with a fixed object at normal speed	
IFML	Frontal collision with a moving object at low speed	
IFMN	Frontal collision with a moving object at normal speed	
IFMR	Frontal collision with a moving object at high speed	
ILCL	Side collision with a soft object at low speed	
ILCN	Side collision with a soft object at normal speed	
ILFL	Side collision with a fixed object at low speed	
ILFN	Side collision with a fixed object at normal speed	
ILML	Side collision with a moving object at low speed	
ILMN	Side collision with a moving object at normal speed	
IPCL	Posterior collision with a soft object at low speed	
IPCN	Posterior collision with a soft object at normal speed	
IPFL	Posterior collision with a fixed object at low speed	
IPEN	Posterior collision with a fixed object at normal speed	
IPML	Posterior collision with a moving object at low speed	
IPMN	Posterior collision with a moving object at normal speed	

ther separated into six cross-validation sets, each containing five repetitions of each activity. The reason for this is two-fold: it allows us to show that the system can be trained from relatively few examples (five events for each activity), and by having cross-validation sets we can get a confidence estimate on our results. All data used in this study was recorded from a single PW driver.

C. EVENT SEGMENTATION

The goal of this component is to extract events-of-interest from the input signal (without, at this stage, identifying the activity’s type). Segmentation of the input signal into events-of-interest during long-term recordings is an essential processing step to make the proposed approach viable. The segmentation step helps in increasing the accuracy of the subsequent classification steps. If we were to keep all the data for the classification steps, then we would face a substantial class imbalance problem, whereby the activity corresponding to standard driving (without an event-of-interest) would constitute a much larger portion of the dataset than the specific activities. This often makes the classification task unnecessarily difficult for the algorithms.

To segment the acquired data, accelerometer signals $S = (X, Y, Z)$ are partitioned into non-overlapping windows of length l , where l is the average duration of one event. For example, $(x_t, x_{t+1}, \dots, x_{t+l})$ forms the i th window’s x coordinate, where $i = \frac{t-1}{l} + 1$. Then for each data point $S_t = (x_t, y_t, z_t)$ in window w_i , the normalized acceleration

magnitude, β_t , is computed for each data point:

$$\beta_t = \sqrt{\left(\frac{x_t - \mu_{xi}}{\sigma_{xi}}\right)^2 + \left(\frac{y_t - \mu_{yi}}{\sigma_{yi}}\right)^2 + \left(\frac{z_t - \mu_{zi}}{\sigma_{zi}}\right)^2} \quad (1)$$

where $\mu_{si, s \in \{x, y, z\}}$ indicates the mean of over time window w_i .

$$\mu_{is} = \frac{\sum_{t \in w_i} s_t}{l}, s \in \{x, y, z\} \quad (2)$$

and $\sigma_{is}^2, s \in \{x, y, z\}$ indicates the variance of over the same window. In this case, variance is defined by the squared deviation of each signal value from its expected value or mean.

$$\sigma_{is}^2 = \frac{\sum_{t \in w_i} (s_t - \mu_{is})^2}{l}, s \in \{x, y, z\}. \quad (3)$$

Afterwards, $\beta(.)$ is smoothed by the autoconvolution function [25].

After the preprocessing step, accelerometer data points in S are clustered into two groups to differentiate between events-of-interest and baseline activity based on the $\beta(.)$ signal. Clustering is achieved using the K-means clustering algorithm [26], which partitions data into K disjoints subsets such that each data point belongs to the cluster with the nearest centroid (we assume $K = 2$). One cluster is composed of events-of-interest, to be further classified, and all the subsequent steps of our methodology are applied only to the data points that fall in this cluster. Data points in the other cluster are not considered for further analysis.

D. FEATURE EXTRACTION

One of the core components of the system is the extraction of appropriate features on which to perform automatic classification. We consider four different types of features, each of which focuses on different signal properties: time-delay embedding features, simple time-domain features, frequency-domain features, and wavelet transforms. One of the contributions of this paper is to provide a detailed comparison of their performance on a challenging dataset. We now present a brief mathematical description of each feature class.

1) SIMPLE TIME DOMAIN FEATURES

The first feature set considered contains simple characteristics of the time-series: mean, variance and line length on each of the three accelerometer coordinates. These time-domain signal properties are simple to measure and have shown good results in classification tasks of time-series [22]. The mean, variance and line length features are generated for windows, w_i of size l from the segmented data. Therefore, the feature vector for the i th window is $\{\mu_{xi}, \mu_{yi}, \mu_{zi}, \sigma_{xi}^2, \sigma_{yi}^2, \sigma_{zi}^2, ll_{xi}, ll_{yi}, ll_{zi}\}$, where μ and σ^2 are defined as in the previous section. The line length of a signal in a time window indicates the sum of absolute changes in the signal amplitude during that period [22]:

$$ll_{is} = \sum_{t \in w_i} |s_t - s_{t-1}|, s \in \{x, y, z\}. \quad (4)$$

2) FFT COEFFICIENTS

The second set of features considered is the magnitude of FFT coefficients for each of the accelerometer dimensions, x, y, z . Although it is more common to use the power spectrum of FFT coefficients, different trials showed that the magnitude of coefficients has better accuracy for classification of dynamic activities [23]. Using a Discrete Fourier Transform (DFT), we extract $\{F_{ix}^1, F_{ix}^2, \dots, F_{ix}^n, F_{iy}^1, \dots, F_{iy}^n, F_{iz}^1, \dots, F_{iz}^n\}$, the FFT feature vector corresponding to window w_i of the signal containing l data points, where F_{is}^k indicates the k th FFT coefficient of signal s :

$$F_{is}^k = \sum_{t \in w_i} s_t e^{\frac{2\pi j k}{l} t}, s \in \{x, y, z\}. \quad (5)$$

Usually, only the first n coefficients for each signal are computed, where n is chosen based on time constraints and desired classification accuracy; we use $n = 10$ throughout our analysis.

3) WAVELET TRANSFORM

In addition to FFT coefficients, a vector of wavelet transform coefficients is also considered. Using wavelet analysis, the original signal is decomposed into coefficients containing both temporal and spectral information. This information can be used to identify the point at which the activity type is changed. Studies have shown that wavelet coefficients extracted from accelerometer signals can provide useful information for activity analysis [23], [24].

We consider wavelet coefficients extracted using the Daubechies transform [32]. In the simplest case, the decomposition is as follows:

$$c(n) = h_0 s(2n) + h_1 s(2n+1) + h_2 s(2n+2) + h_3 s(2n+3) \quad (6)$$

$$d(n) = h_0 s(2n) - h_1 s(2n+1) + h_2 s(2n+2) - h_3 s(2n+3) \quad (7)$$

where the multipliers are:

$$h_0 = \frac{1 + \sqrt{3}}{4\sqrt{2}}, h_1 = \frac{3 + \sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3 - \sqrt{3}}{4\sqrt{2}}, h_3 = \frac{1 - \sqrt{3}}{4\sqrt{2}} \quad (8)$$

In our analysis, we consider the magnitude of the Daubechies 2 wavelets for the 3D accelerometer data, as recommended in previous work [23]. Therefore, the corresponding feature vector for w_i with length l is $\{|d_{ix}^1|, |d_{ix}^2|, \dots, |d_{ix}^n|, |d_{iy}^1|, \dots, |d_{iy}^n|, |d_{iz}^1|, \dots, |d_{iz}^n|\}$, where d_{is}^k indicates the detail coefficient at the k th level of decomposition of signal $s \in \{x, y, z\}$ in time window w_i . A proper value for n can be chosen based on time constraints and desired classification accuracy.

4) TIME-DELAY EMBEDDING FEATURES

The final set of features considered is based on a time-delay embedding of the accelerometer signal. The goal of

this method is to characterize unknown dynamical systems directly from sampled data [10], [11]. The time-delay embedding projects the dynamic properties of the time-series at time t in an m dimensional space $\tilde{S}_t^E = (S_t, S_{t-T}, \dots, S_{t-mT})$ where $T > 1$ is the sampling delay and $m > 1$ is an integer multiplying the sampling delay. The components of this state space form a coordinate system that captures the structure of the time-series [21].

Theoretically, we know that any time-delay structure can reconstruct the dynamics of the system if m is large enough and T does not have a conflict with any periodic property of the system [27]. But in practice, it remains a challenge to find the real dimension d and determine the proper sampling delay m necessary to capture the dynamics of the system. In general, two approaches are used to determine the embedding parameters. In the first one, statistical tools are used to find the parameters and the optimised values are used for further analysis [28]. Alternately, and as we do in this paper, one starts with the intended analysis and optimizes the result for m and T using a grid search over parameter space on a small validation dataset.

After generating the high dimensional model, principal component analysis is used to map the system to the lower dimensional space. In this method, each set of observations in the m -dimensional space, $\tilde{S}_t^E \in R^m$, is mapped to a p -dimensional space, $S_t^E \in R^p$, (with $p < m$) using an orthogonal transformation [29]. After mapping all data points to the new p -dimensional space,¹ the distance from each data point to its nearest neighbour in each (labelled) activity set is measured, and this measurement forms the time-delay embedding feature. In other words, if the dataset includes u different activities $\{A_1, \dots, A_u\}$ and S_{ti}^E indicates the p dimensional embedding for the i th activity at time t , the corresponding feature vector is $F(i, t) = \{|S_{ti}^E - S_{t'j}^E|\}_{j=1, \dots, u}$ where $t' = \text{Argmin}|S_{ti}^E - S_{t'j}^E|\}$. These pairwise distances are used as features for the subsequent classification step. For points that are members of the labelled set, the pairwise distance to their second nearest neighbour is used.

For our experiments, we build three separate time-delay embeddings (corresponding to recorded accelerometer data in the X, Y, and Z axes, respectively) for each of the 35 activity types. The parameters are the same for all embeddings, $m = 16$, $T = 0.1$ sec and $p = 5$.

E. CLASSIFICATION USING SVM CLASSIFIER

Given a dataset where each point is described by a feature vector $\{f(x_t^i), f(y_t^i), f(z_t^i)\}$ $i = 1, \dots, N$, supervised learning techniques can be applied to find a mapping from feature space to event identity. For this step, we use a Support Vector Machine classifier [34]. An SVM classifier constructs a set of hyperplanes in a high-dimensional space that classifies

¹The projected points are stored using a KD-Tree structure [30]. This structure is useful for fast searching because it organizes the data points spatially, and can respond to pairwise distance queries in time linear in the number of points and in the dimension of the tree.

the samples into two classes. The classifier is essentially a function mapping the feature vector $f(x, y, z)$ to the output set, $\gamma \in \{1, -1\}$. To construct these hyperplanes, the SVM uses an iterative training algorithm to minimize an error function, corresponding to the following optimization problem.

$$\begin{aligned} \min_{\omega, b, \zeta} \quad & \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n \zeta_i \\ \text{subject to} \quad & \gamma_i (\omega^T \phi(f_i) + b) \geq 1 - \zeta_i, \quad \zeta_i \geq 0 \end{aligned} \quad (9)$$

The feature vector is mapped to a higher dimensional space by ϕ and the classifier finds a separating hyperplane in this space. SVM classifiers are characterized by their kernel function, $K(f_i, f_j) = \phi(f_i)^T \phi(f_j)$. We use a basic polynomial kernel: $K(f_i, f_j) = (\alpha f_i^T f_j + r)^d$, $\alpha > 0$, with degree $d = 2$. In order to classify a dataset including multiple classes, multiclass SVM can be achieved by reducing a single classification problem into multiple binary classification problems. Our classifiers are trained using the LibSVM tool [34]. Hyper-parameters for the algorithms were selected using our previous study [20].

III. RESULTS

A. SEGMENTATION OF EVENTS

We begin by considering the segmentation of events-of-interest from the datastream. We applied the segmentation approach described above. For this step, we considered datapoints at the full time resolution (250 Hz). The magnitude signal β was generated over windows of size 7000 points (=28 sec), the average duration of each activity. Then β is smoothed by auto-convolving β_t and $\beta_{t+\delta}$ (where $\delta = 100$ points @ 250Hz). Finally, the datapoints in β were divided into two clusters using the K-Means algorithm. We applied a final smoothing step over the clustering output, removing cluster assignments with fewer than k consecutive points ($k = 1500$ points, selected by cross-validation). For most activities, we correctly segmented the 30 repetitions of each activity. Fig. 3a presents a sample segmentation, based on 30 repetitions of the DP1L (*Rolling down 1 inch slope at low speed*) activity. For a few of the activities, some of the events were segmented into two parts. In particular, for three of the activities (MP1L, MP1N, MP1R), each event-of-interest was segmented into two separate events, rather than a single one. This is not so surprising looking at the signal, as shown in Fig. 3b, where we observe that each event is characterized by a double-peak. For the other activity types (excluding MP1L, MP1N, MP1R), out of 960 events, our approach missed one event (0.1%) and over-counted 12 events (1%).

B. ACTIVITY CLASSIFICATION: SAFE VS. UNSAFE

The data points extracted via segmentation are further processed for feature extraction according to the four types of features outlined above. Features of the time domain, FFT coefficients, and wave transforms, were extracted using the Automated Feature Extraction and Selection toolbox (AFESt) [33]. Features for the time-delay

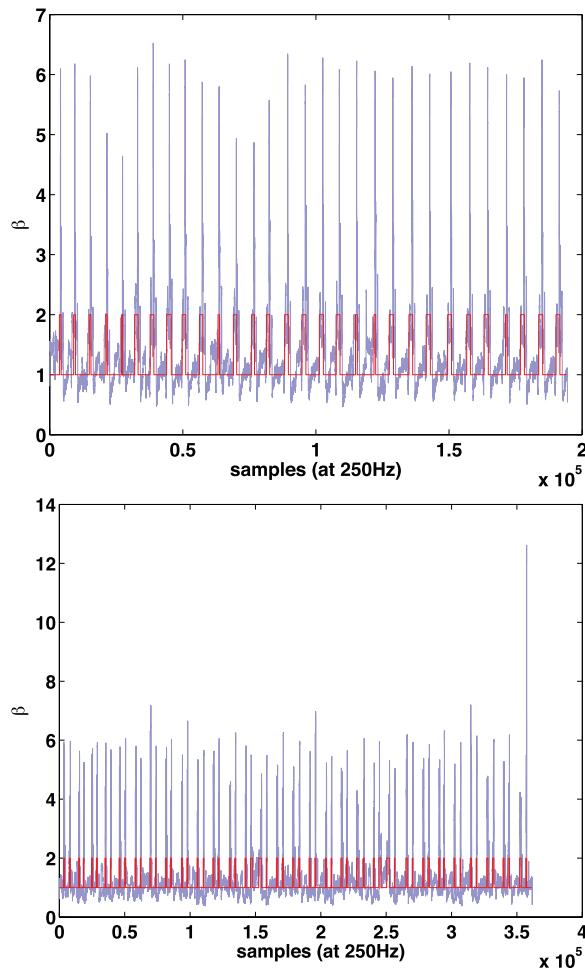


FIGURE 3. (a) Activity DP1L (*Rolling down 1 inch slope at low speed*), 30 repetitions, segmented into two groups: events-of-interest (red boxes) and background data (blue signal). (b) Activity MP1L (*Climbing up 1" slope at low speed*), 30 repetitions, segmented similarly; in this case the automatic segmentation yielded 60 events-of-interest.

embeddings were produced using our own code. Using these features vectors, we considered two types of classification tasks: binary classification to distinguish between Safe and Unsafe events, and per-type classification whereby events are categorized into the 35 types listed in Table 1. In this section we present the results of the Safe vs. Unsafe classification task. The results of the per-type classification task are described in the following section.

To train and evaluate the classifiers, the data corresponding to the events-of-interest were divided into two datasets: a training set that was used to train the algorithms, and included five trials (out of 30) for each activity; and a test set that was used to evaluate the performance of the algorithm, and included five different trials. We used a small training set as this is more realistic for eventual deployments where assessment of PW driving skills for a given PW user could include tasks needed to generate this training set. All reported results are the average of a K -fold cross validation ($K = 6$). Feature extraction, with all types of features, was done over windows of 1.6 sec, assuming a 1.596 sec (=one data point)

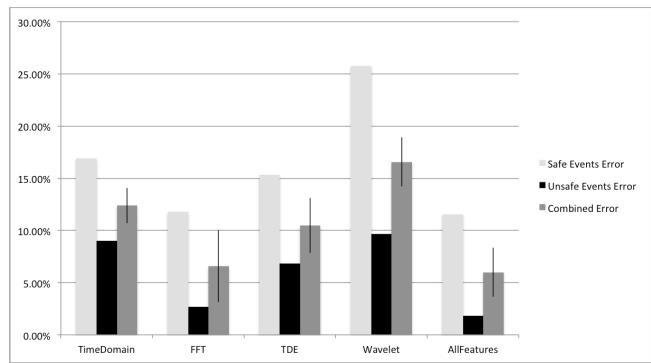


FIGURE 4. Average error rate (%) over all Safe / Unsafe events, using five different types of features: time-delay embeddings, time-based features, FFT coefficient magnitudes, wavelet transform, and all features. Error is defined as the number of false positives and false negatives over the total number of examples. Vertical bars indicate 95% confidence intervals.

overlap with the preceding window.

The average classification error to detect safe and unsafe activities using each type of features is shown in Fig. 4. We also examined the classification accuracy for each of the 35 different activity classes separately; results are presented in Fig. 5.

Using only simple time-domain features, our classification algorithm incorrectly classified 9% of unsafe events as being safe, and 17% of safe events were classified as unsafe, as shown in Fig. 4. We note from Fig. 5 that two out of 20 unsafe activities were classified as being safe more than 20% of the time. Using only FFT coefficients, results in Fig. 4 and Fig. 5 show that we can achieve a low classification error for unsafe events (3%). The FFT features also achieved the lowest error (33%) on the challenging IFMR (*Frontal collision with a moving object at high speed*) activity. However the safe DP1L (*Rolling down 1 inch slope at low speed*) activity was always misclassified as being unsafe. As shown again in Fig. 4, detection of events with time-delay embeddings misclassified 7% of unsafe events and 15% of safe events. The most difficult unsafe activity to classify was once again IFMR (*Frontal collision with a moving object at high speed*). Of the safe activities, the TDE features were the only ones to successfully classify DP1L (*Rolling down 1 inch slope at low speed*), with only 3% error. The average misclassification rate using only wavelet transforms, as shown in Fig. 4, was 10% for unsafe activities and 26% for safe activities. This is not competitive with any of the other feature families considered. We observe from Fig. 5 that classification errors for unsafe activities were localized to certain activity types, for example IFCR (*Frontal collision with a soft object at high speed*) and IFMR (*Frontal collision with a moving object at high speed*) for the unsafe activities. Finally, we considered the classification accuracy that can be achieved when combining all feature families within a single classifier. As shown in Fig. 4, with this combination, our classification algorithm achieved a misclassification rate of 2% for unsafe events and 12% for safe events. This was lower than with any

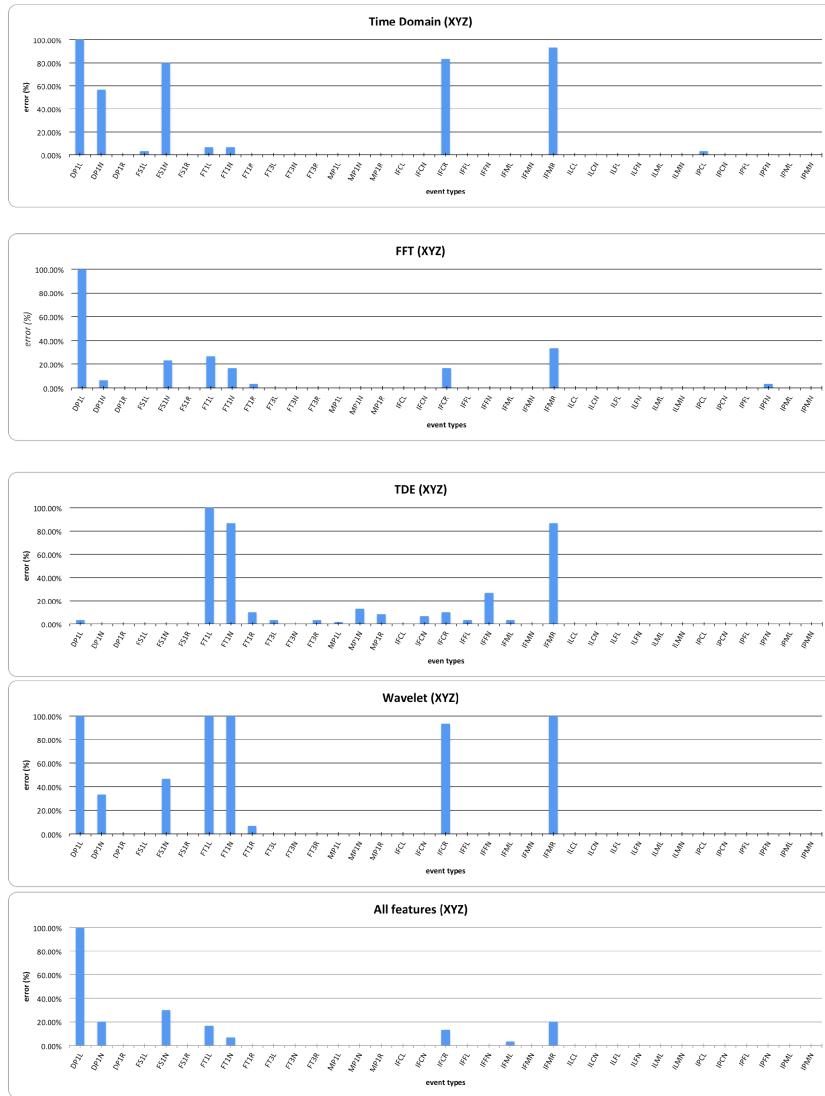


FIGURE 5. Safe vs. unsafe classification average error rate (%) per activity class, using 5 different types of features: time-based features, FFT coefficient magnitudes, time-delay embeddings, wavelet transforms, and all features. The average error rate for each event was calculated over six folds of cross-validation. In each fold, the training set included five examples of each activity.

single feature family, but not significantly better than with the FFT features. This may be due to the small size of the training set; in general, larger training sets are necessary when training classifiers with greater numbers of parameters (features in this case). This classifier outperformed all the others on most activities, though there were a number of activities for which the time-delay embedding features or FFT features performed better, notably IFCR (*Frontal collision with a soft object at high speed*).

Running paired t-tests comparing the misclassification rates obtained with the various feature sets confirmed that we could partition feature families into two sets: {FFT features, All features} vs. {Time-based features, Wavelet features}. Results obtained with any of the feature types in the first set were significantly better (at the $p = 0.05$ level) than those

obtained with any of the features in the second set, but not significantly better than those obtained within their set. There was no significance between time-delay embeddings and any of the other feature families.

We end this section by presenting results showing the impact of the training set size on the Safe/Unsafe classification performance. We focus on the FFT family of features and the AllFeatures set, which performed best in the results above. We observe in Fig. 6 that the Safe/Unsafe classification performance improved steadily until the 5 first training examples, and continued to improve but at a slower rate, up to 15 training examples for FFT features. In the case of the AllFeatures, improvements continued up to 25 training examples (the limit of our dataset, preserving a five example test) and may improve further with more data. This confirms

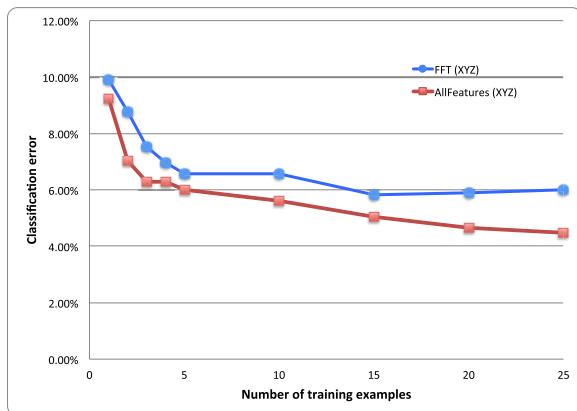


FIGURE 6. Classification error as a function of the size of the training set.

that moderate amounts of data can be sufficient to correctly classify events, but that more data can be useful, especially when considering a greater number of features.

IV. CONCLUSION

The paper presents a machine learning approach targeting the development of automated analysis tools to characterize the driving behavior of wheelchair users. The results presented constitute a proof-of-concept that the system can accurately detect unsafe activities using a variety of features and support-vector machine classification, with less than 2% error, and a relatively low rate of false positives (12%). This approach can be helpful in monitoring safe usage of the wheelchair under varied operating conditions. The number of false alarms can likely be reduced by leveraging richer sensor data (e.g., using audio or video data to better diagnose activities). Offering more training examples in a diversity of safe activities could also improve robustness.

We compared classification performance using different sets of features common in the time-series analysis literature: time-delay embeddings, time-domain characterization, frequency-domain features, and wavelet transforms. This was deemed necessary because these features have been often used in the literature, without reliable comparison between them. The results presented here show that the most robust performance was achieved when combining all features, however nearly as good performance was obtained by using FFT features alone.

Our approach could be helpful in diagnosing frequent driving mistakes in every day usage, without requiring a user to keep a detailed diary. One of the challenges in moving forward with the implementation of this technique onboard regular PWs is to validate the method using natural recording conditions. This poses logistics challenges in terms of obtaining accurate labeling of activities. The labels could however be assigned through post-processing of the recorded data, and corroboration through camera data. Furthermore, the current system was evaluated using activity recordings from a single user; we expect results to be relatively robust

to different users given the very narrow class of events considered, however this remains to be tested.

Acknowledgment

The authors gratefully acknowledge contributions on early versions of this work by Thérèse Audet, Jan Polgar and Jordan Frank.

REFERENCES

- [1] A.-L. Salminen, Å. Brandt, K. Samuelsson, O. Töytäri, and A. Malmivaara, "Mobility devices to promote activity and participation: A systematic review," *J. Rehabil. Med.*, vol. 41, no. 9, pp. 697–706, 2009.
- [2] A. Cook and J. Polgar, *Cook and Hussey's Assistive Technologies: Principles and Practice*, 3rd ed. New York NY, USA: Elsevier, 2008.
- [3] L. A. McClure *et al.*, "Wheelchair repairs, breakdown, and adverse consequences for people with traumatic spinal cord injury," *Archives Phys. Med. Rehabil.*, vol. 90, pp. 2034–2038, Dec. 2009.
- [4] S. Ummat and R. L. Kirby, "Nonfatal wheelchair-related accidents reported to the national electronic injury surveillance System," *Amer. J. Phys. Med. Rehabil.*, vol. 73, no. 3, pp. 163–167, 1994.
- [5] L. Worobey, M. Oyster, G. Nemunaitis, R. Cooper, and M. L. Boninger, "Increases in wheelchair breakdowns, repairs, and adverse consequences for people with traumatic spinal cord injury," *Amer. J. Phys. Med. Rehabil.*, vol. 91, no. 6, pp. 463–469, 2012.
- [6] R. A. Cooper, M. J. Dvorznak, T. J. O'Connor, M. L. Boninger, and D. K. Jones, "Braking electric-powered wheelchairs: Effect of braking method, seatbelt, and legrests," *Archives Phys. Med. Rehabil.*, vol. 79, no. 10, pp. 1244–1249, 1998.
- [7] T. A. Corfman, R. A. Cooper, S. G. Fitzgerald, and R. Cooper, "Tips and falls during electric-powered wheelchair driving: Effects of seatbelt use, legrests, and driving speed," *Archives Phys. Med. Rehabil.*, vol. 84, no. 12, pp. 1797–1802, 2003.
- [8] A. Fast, J. Sosner, P. Begeman, M. Thomas, and D. Drukman, "Forces, moments, and acceleration acting on a restrained dummy during simulation of three possible accidents involving a wheelchair negotiating a curb: Comparison between lap belt and four-point belt," *Amer. J. Phys. Med. Rehabil.*, vol. 76, no. 5, pp. 370–377, 1997.
- [9] W.-Y. Chen *et al.*, "Wheelchair-related accidents: Relationship with wheelchair-using behavior in active community wheelchair users," *Archives Phys. Med. Rehabil.*, vol. 92, no. 6, pp. 892–898, 2011.
- [10] M. M. LaBan and T. S. Nabity, Jr., "Traffic collisions between electric mobility devices (wheelchairs) and motor vehicles: Accidents, hubris, or self-destructive behavior?" *Amer. J. Phys. Med. Rehabil.*, vol. 89, no. 7, pp. 557–560, 2010.
- [11] G. U. Sorrento, P. S. Archambault, F. Routhier, D. Dessureault, and P. Boissy, "Assessment of joystick control during the performance of powered wheelchair driving tasks," *J. Neuroeng. Rehabil.*, vol. 8, p. 31, May 2011.
- [12] A. D. Mountain *et al.*, "Ability of people with stroke to learn powered wheelchair skills: A pilot study," *Archives Phys. Med. Rehabil.*, vol. 91, no. 4, pp. 596–601, Apr. 2010.
- [13] A. H. MacPhee, R. L. Kirby, A. L. Coolen, C. Smith, D. A. MacLeod, and D. J. Dupuis, "Wheelchair skills training program: A randomized clinical trial of wheelchair users undergoing initial rehabilitation," *Archives Phys. Med. Rehabil.*, vol. 85, no. 1, pp. 41–50, Jan. 2004.
- [14] R. J. Kirby. (2013). *Wheelchair Skills Test Version 4.2*. [Online]. Available: <http://www.wheelchairskillsprogram.ca/eng/testers.php>
- [15] H. Hoenig, P. Giacobbi, and C. E. Levy, "Methodological challenges confronting researchers of wheeled mobility aids and other assistive technologies," *Disab. Rehabil. Assistive Technol.*, vol. 2, no. 3, pp. 159–168, May 2007.
- [16] R. A. Cooper *et al.*, "Driving characteristics of electric-powered wheelchair users: How far, fast, and often do people drive?" *Archives Phys. Med. Rehabil.*, vol. 83, no. 2, pp. 250–255, Feb. 2002.
- [17] S. E. Sonenblum, S. Sprigle, F. H. Harris, and C. L. Maurer, "Characterization of power wheelchair use in the home and community," *Archives Phys. Med. Rehabil.*, vol. 89, no. 3, pp. 486–491, Mar. 2008.
- [18] H. G. Kang *et al.*, "In situ monitoring of health in older adults: Technologies and issues," *J. Amer. Geriatrics Soc.*, vol. 58, no. 8, pp. 1579–1586, 2010.

- [19] P. Boissy, H. Hamel, P. Archambault, and F. Routhier, "Ecological measurement of powered wheelchair mobility and driving performance using event-driven identification and classification methods," in *Proc. RESNA*, 2008.
- [20] A. K. Moghaddam *et al.*, "Mobility profile and wheelchair driving skills of powered wheelchair users: Sensor-based event recognition using a support vector machine classifier," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug./Sep. 2011, pp. 7336–7339.
- [21] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [22] D. E. Olsen, R. P. Lesser, J. C. Harris, W. R. S. Webber, and J. A. Cristion, "Automatic detection of seizures using electroencephalographic signals," U.S. Patent 5 311 876, Nov. 18, 1992.
- [23] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, and D. Howard, "A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 3, pp. 871–879, Mar. 2009.
- [24] M. Sekine, T. Tamura, T. Togawa, and Y. Fukui, "Classification of waist-acceleration signals in a continuous walking record," *Med. Eng. Phys.*, vol. 22, no. 4, pp. 285–291, May 2000.
- [25] A. M. Uludağ, "On possible deterioration of smoothness under the operation of convolution," *J. Math. Anal. Appl.*, vol. 227, no. 2, pp. 335–358, Nov. 1998.
- [26] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, Jun. 1967, pp. 281–297.
- [27] F. Takens, "Detecting strange attractors in turbulence," in *Dynamic Systems and Turbulence*. Berlin, Germany: Springer-Verlag, 1981, pp. 366–381.
- [28] T. Buzug and G. Pfister, "Optimal delay time and embedding dimension for delay-time coordinates by analysis of the global static and local dynamical behavior of strange attractors," *Phys. Rev. A*, vol. 45, pp. 7073–7084, May 1992.
- [29] I. T. Jolliffe, *Principal Component Analysis*. New York, NY, USA: Springer-Verlag, 2002.
- [30] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM*, vol. 18, no. 9, pp. 509–517, 1975.
- [31] C. K. Chui, *An Introduction to Wavelets*. San Francisco, CA, USA: Academic, 1992.
- [32] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA, USA: SIAM, 1992.
- [33] G. Saulnier, *Automated Feature Extraction and Selection Toolbox*. [Online]. Available: <http://www.cs.mcgill.ca/~gsauln/AFESt.html>
- [34] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011. [Online]. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>



JOELLE PINEAU is currently an Associate Professor with the School of Computer Science, McGill University, Montreal, QC, Canada, where she is also a member of the Centre for Intelligent Machines. She received the B.A.Sc. degree in engineering from the University of Waterloo, Waterloo, ON, Canada, in 1998, and the M.Sc. and Ph.D. degrees in robotics from Carnegie Mellon University, Pittsburgh, PA, USA, in 2001 and 2004, respectively. Her research focuses on developing models and algorithms for learning and decision-making in partially observable stochastic domains, and applying these results to complex problems in robotics and healthcare.



ATHENA K. MOGHADDAM was born in Iran in 1987. She received the B.Sc. degree in computer engineering from the University of Tehran, Tehran, Iran, in 2010, and the M.Sc. degree in computer science from McGill University, Montreal, QC, Canada, in 2013. Since then, she has been with Facebook Inc., Menlo Park, CA, USA, where she is currently a Software Engineer in Pages Growth and Monetization.



HIU KIM YUEN was born in Hong Kong. He received the bachelor's degree in actuarial analysis from the Chinese University of Hong Kong, Hong Kong, in 2008. He is currently pursuing the master's degree with the Reasoning and Learning Laboratory, Department of Computer Science, McGill University, Montreal, QC, Canada.



PHILIPPE S. ARCHAMBAULT is currently an Associate Professor with the School of Physical and Occupational Therapy, McGill University, Montreal, QC, Canada. He received the Ph.D. degree in neurosciences from the University of Montreal, Montreal, in 2003. His research interests include power wheelchair mobility, assistive technology, virtual reality, and rehabilitation. He is a member of the Rehabilitation Engineering and Assistive Technology Society of North America and the Society for Neuroscience.



FRANÇOIS ROUTHIER is currently a Researcher with the Centre for Interdisciplinary Research in Rehabilitation and Social Integration, Quebec City Rehabilitation Institute, Quebec City, QC, and an Associate Professor with the Department of Rehabilitation, Université Laval, Quebec City. He has been a Junior 1 Research Scholar FRQS since 2013. His research interests include wheeled mobility, wheelchair skills, social participation, activity measurement, and assistive devices impacts.



FRANÇOIS MICHAUD (M'90) received the bachelor's, master's, and Ph.D. degrees in electrical engineering from the Université de Sherbrooke, Sherbrooke, QC, Canada, in 1992, 1993, and 1996, respectively. After completing post-doctoral work at Brandeis University, Waltham, MA, USA, in 1997, he became a faculty member with the Department of Electrical Engineering and Computer Engineering, Université de Sherbrooke, and founded the IntRoLab, a research laboratory working on designing intelligent autonomous systems that can assist humans in living environments. His research interests are in architectural methodologies for intelligent decision-making and design of interactive autonomous mobile robots. He was the Canada Research Chair (2001–2011) in Mobile Robots and Autonomous Intelligent Systems, and is the Director of the Interdisciplinary Institute for Technological Innovation, Université de Sherbrooke. He is a member of the Association for the Advancement of Artificial Intelligence and the Order of Engineers of Quebec.



PATRICK BOISSY received the Ph.D. degree in biomedical sciences with a specialization in rehabilitation. He is currently an Associate Professor of Orthopaedic Services with the Department of Surgery, Université de Sherbrooke, Sherbrooke, QC, Canada, where he is also a Senior Research Scholar FRQS with the Research Centre on Aging and Interdisciplinary Institute of Technological Innovation. His research is focused on activity monitoring using wearable sensors and their use in clinical and telehealth applications.