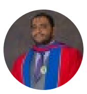
Robust Motion Mode Recognition for Portable Navigation Independent on Device Usage

CONFERENCE PAPER · MAY 2015

DOI: 10.1109/PLANS.2014.6851370

4 AUTHORS, INCLUDING:



Mostafa Elhoushi
InvenSense Inc.

Michael J Korenberg
Queen's University

Aboelmagd Noureldin
Royal Military College of Canada

Robust Motion Mode Recognition for Portable Navigation Independent on Device Usage

Mostafa Elhoushi[1,2], Jacques Georgy[1], Michael Korenberg[2], Aboelmagd Noureldin[2,3]
[1]Trusted Positioning Inc., Calgary, Canada
[2]Electrical and Computer Engineering Department, Queen's University, Kingston, Canada
[3]Electrical and Computer Engineering Department, Royal Military College, Kingston, Canada

***Abstract*— Portable navigation has become increasingly prevalent in daily activities. The need for accurate user positioning information, including a person's location and velocity, when using a portable device (such as a cell phone, tablet, or even a smart watch) is growing in various fields. Knowing the user's mode of motion or conveyance allows appropriate algorithms or constraints, related to each mode, to be used to estimate a more accurate position and velocity. The modes covered in this paper are walking, running, cycling, and land-based vessels (including vehicles, truck, buses, and trains which include light rail trains and subways). The work discussed in this paper involves the use of sensors – with and without Global Navigation Satellite Systems (GNSS) signal availability - in portable devices to help recognize the mode of motion for an arbitrary user, an arbitrary use case - whether the device is held in the hand, in the pocket, or at the ear, etc. -and an arbitrary orientation of the device.**

Keywords*— *activity recognition; mode of motion; inertial sensors; walking; cycling; running, land-based vessel

I. Introduction

While Global Navigation Satellite Systems (GNSS) signals allow for highly accurate navigation in open-sky areas, they are weak or absent, indoors, in urban canyons, and in underground locations [1]. For situations where GNSS signals are low or have no signal availability, portable navigation solutions have an alternative to use Micro-Electro-Mechanical Systems (MEMS) sensors (typically accelerometers, gyroscopes, magnetometers, and barometers) and fuse their readings in a navigation algorithm to try to obtain the most accurate navigation results [2]. However, MEMS sensors suffer from high error rates [3], and navigation solutions encounter other problems such as unknown heading and unknown misalignment.

Knowing a user's mode of motion is valuable to improve navigation results especially when GNSS signals are not available [4]. Different modes of motion can use distinctive navigation algorithms and particular motion constraints optimal for each mode. For example, pedestrian dead reckoning (PDR) algorithms [5] are used when walking mode is detected. Drastic errors in the results may occur if PDR is used when the mode of motion is actually driving. The calculated velocity shall be much lower than the actual velocity. In this case, the erroneous position calculations shall result in estimating that the user has traveled a short distance when they have actually traveled a longer distance. Knowledge of the mode of motion or conveyance may help in constraining the navigation algorithm outputs, e.g., refusing velocities above a threshold limit if cycling mode is detected, and accepting more frequent steps if running mode is detected.

II. Background

Much of the previous work in mode of motion recognition dealt with tethered portable devices [6], i.e., portable devices are attached to a specific known part of the body presenting a less challenging solution. Other work used multi-sensors attached to several parts of the body [7]. Additional work only dealt with one or two use cases [8]; mainly, handheld cases or dangling while walking. The significance of the work illustrated in this paper is that the use case and orientation of the portable device is arbitrary; therefore, the recognition system is much more robust.

Previous work mainly used motion sensor readings from accelerometers [9], which measure acceleration, and gyroscopes, which measure angular rotation [10]. Some have used other motion sensors such as magnetometers [11][12], which measure magnetic field intensity, and barometers [13], which estimate altitude deduced from air pressure and temperature measurement. Further work employed light sensors [14], audio sensors [15][16], and humidity sensors [16]. However, the drawback of the readings of latter sensors is that they may vary for the same motion mode if done in different places or at different times of the day. Heart beat rate readings were employed by some, but they also depend on many other factors other than the activity being performed by a person. Some research depended on velocity derived from GNSS signals [17], which, although may be accurate, will not work indoors or in urban canyons.

It is theoretically impossible to predict mode of motion given readings from a set of sensors at one epoch or sample of time, since any mode of motion is characterized by a series of signal values over a specific period of time [10]. Therefore, signal readings from any sensor need to be windowed, i.e., grouped into groups of N samples, where N is an arbitrary integer, as shown in Fig. 1, and the values in each window are processed and analyzed in order to predict the most likely motion mode performed during the time of the window. Long windows have the benefit of having more data to analyze but their disadvantage is that it is more likely to have a transition from one mode of motion to another within the same window time, and therefore navigation based upon the prediction result will be less accurate.

978-1-4799-3320-4/14/$31.00 © 2014 IEEE

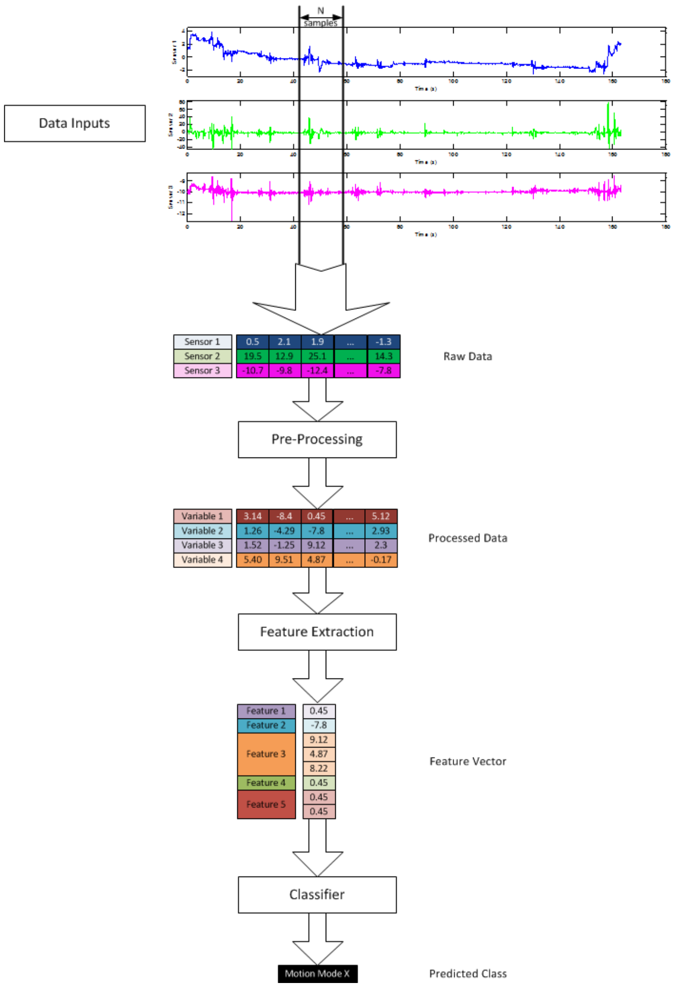


Fig. 1. Overview of motion mode recognition process.

Raw readings from most sensors are usually not sufficient to predict mode of motion, and therefore need some pre-processing to get variables which make more sense. For example, the readings from an accelerometer along the *x*-axis of a measuring device, give little indication of the type of motion, especially if the direction of motion is along another axis. What has more meaning, is the norm of acceleration readings along the three axes: x, y, and, z, or the levelled acceleration along the vertical axis or horizontal plane. We shall refer to such an operation to convert sensor readings to more meaningful variables as *pre-processing*.

The next step done in most of the related work involves *feature extraction*: extracting features from variables within each window to obtain data which can distinguish one motion mode from another. A feature may consist of a single value as mean or variance, while another feature may be composed of a vector of values, as in Fourier transform [4] or binned distribution [8]. All the features extracted are arranged into a vector, commonly known as feature vector, to be inputted into the next stage.

The next stage is *classification* [18], where the most probable motion mode which occurred within the window is predicted by a classifier given the feature vector of the window. Motion mode recognition has been mainly done in literature using pattern recognition or machine learning methods. In machine learning, a classifier is built by training it over feature vectors from a large set of training data. Each classifier such as decision tree [19], support vector machine [12], or neural networks [20], has its own training algorithm. Training is generally a time consuming offline process which usually uses huge amounts of data. Once a classifier model is built, it can be used within the motion mode recognition process to evaluate any feature vector from any window of data.

III. Methodology

The methods employed in this research work at each step of the motion mode recognition process are explained in the following sections.

A. Data Inputs

The work done in this paper is based on using a prototype of a portable device, which can be embodied as a smartphone, tablet, smart watch, or smart glasses that contains:

- An accelerometer triad: measuring accelerations or specific forces, f_x, f_y, and f_z, along each of the three orthogonal axes at a sampling rate of 20 Hz.
- A gyroscope triad: measuring angular rotation rates, ω_x, ω_y, and ω_z, along each of the three orthogonal axes at a sampling rate of 20 Hz.
- A magnetometer triad: measuring magnetic field intensities, B_x, B_y, and B_z, along each of the three orthogonal axes at a sampling rate of 1 sample every 0.15 seconds.
- A barometer: measuring atmospheric pressure, P, and estimating barometric height, $\tilde{h}$, at a rate of 10 Hz.
- A GPS receiver: estimating position and velocity using pseudo range and Doppler measurements when direct lines of sight to at least 4 GPS satellites are available, at a rate of 1 Hz.

B. Pre-Processing

The readings from the sensors are processed to calculate an estimate at a rate of 20 Hz of the following variables:

- leveled vertical acceleration, a_{up} [21]: the net acceleration of the portable device along the axis normal to the surface of the Earth. In small acceleration scenarios, it can be estimated using the following equation:

$$\boldsymbol{a}_{up} = \left(\frac{(\boldsymbol{a}-\bar{\boldsymbol{a}})\cdot\bar{\boldsymbol{a}}}{\bar{\boldsymbol{a}}\cdot\bar{\boldsymbol{a}}}\right)\bar{\boldsymbol{a}} \quad (1)$$

where $\boldsymbol{a}$ is the vector representing the three readings of the accelerometer triad, $\bar{\boldsymbol{a}}$ is its average across a window and is an estimate of the gravity vector, and $a_{up} = |\boldsymbol{a}_{up}|$,

- magnitude of leveled horizontal plane acceleration, a_h [21]: magnitude of the acceleration component along the plane parallel to the surface of the Earth, and is equivalent to:

$$\boldsymbol{a}_h = \boldsymbol{a} - \boldsymbol{a}_{up} \quad (2)$$

where $a_h = |\boldsymbol{a}_h|$, and

- norm of angular rotation components, $|\omega|$: the square root of the sum of squares of angular rotation components after removing their biases, where the biases are either obtained by static calibration or by filtering:

$$|\omega| = \sqrt{(\omega_x - b_x)^2 + \left(\omega_y - b_y\right)^2 + (\omega_z - b_z)^2} \quad (3)$$

In this work such variables are calculated within a fusion process (using a Kalman filter or particle filter) and therefore the above equations may be tuned or optimized. The fusion process involves readings from the different sensors mentioned in Section III.B. Therefore, it is expected that the three variables listed in this section and the features extracted from them mentioned in the next section will be more accurate when GNSS signals are available.

C. Feature Extraction

After processing the raw sensor signals to obtain the three variables mentioned above, features are then extracted from them in each window to obtain a feature vector. Each feature is the result of a certain mathematical operation performed upon the values of one or more variables across each window.

The features extracted from each of the three variables: leveled vertical acceleration, magnitude of leveled horizontal plane acceleration, and norm of angular rotation components, can be categorized into several categories:

- statistical features:
  - mean [22][8],
  - mean of absolute values [8],
  - median [23],
  - mode,
  - variance [24],
  - standard deviation [24][8],
  - 75th percentile [25],
  - inter-quartile range [25],
  - average absolute difference [8], and
  - binned distribution [8],
- energy, power, and magnitude features:
  - energy [22],
  - sub-band energies [26],
  - sub-band energy ratios [26], and
  - signal magnitude area [23],
- time-domain features:
  - zero-crossing rate [4][27] and
  - number of maximum peaks [28][8],
- frequency-domain features:
  - absolute values of short-time Fourier transform [26],
  - power of short-time Fourier transform [4],
  - power spectral centroid [27],
  - average of continuous wavelet transform at various approximation levels [29],
  - frequency domain entropy [22], and
  - frequency and amplitude of the most 4 contributing frequency components obtained using spectral fast orthogonal search [30],
- other:
  - cross-correlation between leveled vertical and horizontal acceleration components [26].

D. Classification

The classification method employed in this research work is decision tree. A decision tree consists of a sequence of inequality checks on elements of a feature vector. Depending on the result of each check, either another inequality check is made or the most probable motion mode is decided on.

IV. EXPERIMENT

To build the classification model, a vast amount of data was collected from different users of various ages, genders, heights, weights, and motion dynamics. Experiments were carried out using portable devices such as smartphones, tablets, smartwatches, and smart glasses. To ensure a robust recognition module, data was logged for each motion mode using many use cases, with most use cases covered in various orientations, as shown in Table I:

- For walking and running: different use cases were demonstrated by each user while in different modes of motion, including: handheld in any orientation or tilt, hand still by side of body, dangling, in pocket (including side pant pocket and back pant pocket, jacket pocket, suit pocket) in different orientations, on ear, in belt holder, in an arm band, attached to the chest, leg, or wrist, in a backpack, or in a purse. Some of the trajectories were done indoors. Other trajectories were done outdoors, with some in open sky for all the time, and others suffering from GNSS signal blockage for some time. Also, some trajectories had a portion of their times indoors and a portion outdoors.
- For cycling trajectories: different use cases covered included chest, arm, leg, pocket, belt, wrist/watch, backpack, mounted on thigh, attached to bicycle, and bicycle holder (in different locations on bicycle).
- The land-based vessel trajectories included the following uses cases: pocket, belt, ear, handheld, wrist/watch, glasses/head mounted, and backpack.

TABLE I. USE CASES COVERED FOR EACH MOTION MODE

Use Case	**Motion Mode**			
	Walking	***Running***	***Bicycle***	***Land-based Vessel***
Handheld	✓			✓
Hand Still by Side	✓			✓
Pocket	✓	✓	✓	✓
Ear	✓	✓		✓
Belt Holder	✓	✓	✓	✓
Dangling	✓	✓		
Arm Band	✓	✓	✓	
Chest		✓	✓	
Leg		✓	✓	
Thigh			✓	
Wrist/ Smart watch	✓	✓	✓	✓
Backpack	✓	✓	✓	✓
Purse	✓			
Laptop Bag	✓			
Bicycle Handle			✓	
Jacket Clipped			✓	
Bicycle Holder			✓	
Car Dashboard				✓
Car Drawer				✓
Car Box between Seats				✓
Car Holder				✓
On Seat				✓
Goggle/ Smart Glasses				✓

Fig. 2. Photos of data collection experiments: walking [top left], sitting in subway [top centre], standing in streetcar [top right], running [bottom right], cycling [bottom centre], and walking [bottom right].

  - The use cases covered sitting (in all vessel platforms), standing (in different types of trains and buses), and on platform (such as on seat in all vessel platforms, on a car holder, on the dashboard, in a drawer, or in a box between the car's seats).
  - Data was collected in various cars of different models, in trucks, in trains, light rail trains, streetcars, subways, public transit buses, and city-to-city buses.
  - Trajectories covered various places and various cities. Crowded downtown areas, wide uptown areas, on highways, or in quiet neighborhoods were aspects in the trajectories.

Fig. 2 shows photos of some of the data collection experiment trajectories. The total number of trajectories was more than 1000, and the total time of the various motion modes was about 200 hours.

V. RESULTS

The experiments or trajectories were categorized as either training trajectories or evaluation trajectories. Features extracted from training trajectories were used to generate the classification model, and then the classification model was tested against the features extracted from the evaluation trajectories. The results of classifying the evaluation

TABLE II. CONFUSION MATRIX OF THE EVALUATION OF CLASSIFIER

Actual Motion Mode	**Predicted Motion Mode**			
	Walking	***Running***	***Bicycle***	***Land-based Vessel***
Walking	96.5%	2.2%	1.1%	0.3%
Running	0.4%	99.4%	0.2%	0.0%
Bicycle	1.4%	1.9%	92.0%	4.8%
Land-based Vessel	0.3%	0.0%	8.4%	91.2%

trajectories showed an overall accuracy of 94.77%. The confusion matrix [18], which shows the proportion of correct and incorrect predictions of each motion mode, is shown in Table II.

We notice from the confusion matrix that Running is the motion mode with highest accuracy rate – reaching 99.4% - indicating that the high frequency dynamics of leveled vertical acceleration makes it unique compared to other motion modes. Walking has the second highest classification rate, with some misclassification with Running when the user is walking at a high speed.

On the other hand, about 8% of Land-based Vessel motion mode is misclassified as Bicycle, since the motion dynamics of Land-based Vessel when at relatively low speeds have some similarity to the Bicycle.

Bicycling has the second lowest accuracy rate of about 92%. Detecting Bicycle motion mode is challenging, since some use cases result in signal dynamics similar to Land-based Vessel. This happens when the portable device is mounted to the bicycle handle or placed in the user's chest or backpack. In the other use cases the signal dynamics are similar to Walking or Running, depending on the cycling rate. When the portable device is in the user's pocket or attached to the user's thigh or leg, decreased mis-classification with walking or running may happen. In the former case, the leveled vertical acceleration is expected to have fewer peaks, while in the latter case, more peaks are expected, similar to that resulting from steps in pedestrian motion. In such case, more peaks are expected due to the portable device's motion being affected by the pedalling of the user, and therefore higher misclassification with Walking and Running is expected.

VI. CONCLUSION

This research has successfully resulted in a recognition module that has differentiated between walking, running, cycling, and land-based vessel motion modes. The recognition process is robust because it is based on users with various characteristics that performed tests in a variety of orientations and use cases.

Future work may include reducing the misclassification between Bicycle mode of motion and Land-based Vessel mode of motion. The work may be extended to cover more ways of conveyance including planes and marine vessels. Also, a challenging motion mode would be to detect walking within a moving land-based vessel, e.g., walking within a moving train or walking within a moving bus.

REFERENCES

[1] P. D. Groves, *Principles of GNSS Inertial and Multi-Sensor Integrated Navigation Systems - GNSS Technology and Applications*. Boston, USA: Artech House, 2008, p. 522.

[2] A. Noureldin, T. B. Karamat, and J. Georgy, *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*. Berlin Heidelberg: Springer, 2013, p. 313.

[3] P. Aggarwal, Z. Syed, A. Noureldin, and N. El-Sheimy, *MEMS-Based Integrated Navigation*. Boston, USA: Artech House, 2010, p. 213.

[4] S. Saeedi, X. Zhao, and Z. Sayed, "Context Aware Mobile Personal Navigation Using Multi-level Sensor Fusion," in *Proceedings of the 24th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2011)*, 2011, pp. 1394–1403.

[5] C. Fischer, P. Talkad Sukumar, and M. Hazas, "Tutorial: implementation of a pedestrian tracker using foot-mounted inertial sensors," *IEEE Pervasive Comput.*, pp. 1–19, 2012.

[6] Z. Sun, X. Mao, W. Tian, and X. Zhang, "Activity classification and dead reckoning for pedestrian navigation with wearable sensors," *Meas. Sci. Technol.*, vol. 20, no. 1, Jan. 2009.

[7] N. Kern, B. Schiele, and A. Schmidt, "Multi-sensor activity context detection for wearable computing," in *Ambient Intelligence*, E. Aarts, R. W. Collier, E. van Loenen, and B. de Ruyter, Eds. Berlin Heidelberg: Springer, 2003, p. pp 220–232.

[8] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explor. Newsl.*, vol. 12, no. 2, pp. 74–82, 2011.

[9] N. Ravi, N. Dandekar, P. Mysore, and M. Littman, "Activity recognition from accelerometer data," in *Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence - Volume 3*, 2005, pp. 1541–1546.

[10] O. D. Lara and M. a. Labrador, "A Survey on Human Activity Recognition using Wearable Sensors," *IEEE Commun. Surv. Tutorials*, pp. 1–18, 2012.

[11] S.-W. Lee and K. Mase, "Activity and location recognition using wearable sensors," *Pervasive Comput. IEEE*, vol. 1, no. 3, pp. 24–32, Jul. 2002.

[12] K. Altun and B. Barshan, "Human activity recognition using inertial/magnetic sensor units," in *HBU'10 Proceedings of the First international conference on Human behavior understanding*, 2010, pp. 38–51.

[13] J. Lester and T. Choudhury, "A practical approach to recognizing physical activities," *Lect. Notes Comput. Sci. Pervasive Comput.*, pp. 1–16, 2006.

[14] P. Lukowicz, H. Junker, and M. Staeger, "WearNET: A distributed multi-sensor system for context aware wearables," in *UbiComp 2002: Ubiquitous Computing*, L. Borriello, Gaetano and Holmquist, Ed. Springer Berlin Heidelberg, 2002, pp. 361–370.

[15] J. Pan, "Sensor-Based Abnormal Human-Activity Detection," *IEEE Trans. Knowl. Data Eng.*, vol. 20, no. 8, pp. 1082–1090, Aug. 2008.

[16] T. Choudhury, G. Borriello, S. Consolvo, D. Haehnel, B. Harrison, B. Hemingway, J. Hightower, P. "Pedja" Klasnja, K. Koscher, A. LaMarca, J. A. Landay, L. LeGrand, J. Lester, A. Rahimi, A. Rea, and D. Wyatt, "The Mobile Sensing Platform: An Embedded Activity Recognition System," *IEEE Pervasive Comput.*, vol. 7, no. 2, pp. 32–41, Apr. 2008.

[17] P. Groves, Z. Jiang, H. Martin, and K. Voutsis, "Context Detection, Categorization and Connectivity for Advanced Adaptive Integrated Navigation," in *Proceedings of the 2013 International Technical Meeting of The Institute of Navigation*, 2013.

[18] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 2nd Editio. San Diego, CA: Elsevier Academic Press, 2003, p. 689.

[19] R. E. Guinness, "Beyond Where to How : A Machine Learning Approach for Sensing Mobility Contexts Using Smartphone Sensors," in *Proceedings of the 2013 International Technical Meeting of The Institute of Navigation*, 2013.

[20] X. Zhao, S. Saeedi, N. El-Sheimy, Z. Syed, and C. Goodall, "Towards Arbitrary Placement of Multi-sensors Assisted Mobile Navigation System," in *Proceedings of the 23rd International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2010)*, 2010, pp. 556 – 564.

[21] D. Mizell, "Using gravity to estimate accelerometer orientation," in *Proceedings of the 7th IEEE International Symposium on Wearable Computers*, 2005, pp. 252–253.

[22] L. Bao and S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing*, 2004, pp. 1–17.

[23] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *Inf. Technol. Biomed. IEEE Trans.*, vol. 10, no. 1, pp. 156–167, Jan. 2006.

[24] M. Susi, V. Renaudin, and G. Lachapelle, "Motion mode recognition and step detection algorithms for mobile phone users," *Sensors*, vol. 13, no. 2, pp. 1539–62, Jan. 2013.

[25] K. Frank, M. Nadales, and P. Robertson, "Reliable Real-Time Recognition of Motion Related Human Activities Using MEMS Inertial Sensors," in *ION GNSS 2010*, 2010.

[26] M. Susi, D. Borio, and G. Lachapelle, "Accelerometer signal features and classification algorithms for positioning applications," in *Proceedings of the 2011 International Technical Meeting of The Institute of Navigation*, 2011, pp. 158–169.

[27] J. Yang, "Toward physical activity diary: motion recognition using simple acceleration features with mobile phones," in *Proceedings of the 1st international workshop on Interactive multimedia for consumer electronics*, 2009, pp. 1–10.

[28] K. Kunze and P. Lukowicz, "Dealing with sensor displacement in motion-based onbody activity recognition systems," in *Proceedings of the 10th international conference on Ubiquitous computing - UbiComp '08*, 2008, p. 20.

[29] T. M. E. Nijsen, R. M. Aarts, P. J. M. Cluitmans, and P. a M. Griep, "Time-frequency analysis of accelerometry data for detection of myoclonic seizures.," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 5, pp. 1197–203, Sep. 2010.

[30] M. Korenberg, "A robust orthogonal algorithm for system identification and time-series analysis," *Biol. Cybern.*, vol. 276, pp. 267–276, 1989.