# On the Statistical Properties of Body-Worn Inertial Motion Sensor Data for Identifying Sensor Modality

**Philipp M. Scholl**
Embedded Systems,
University of Freiburg
psholl@ese.uni-freiburg.de

**Kristof van Laerhoven**
Ubiquitious Computing,
University of Siegen
kvl@eti.uni-siegen.de

**ABSTRACT**

Interpreting datasets containing inertial data (acceleration, rate-of-turn, magnetic flux) requires a description of the datasets itself. Often this description is unstructured, stored as a convention or simply not available anymore. In this note, we argue that each modality exhibits particular statistical properties, which allows to reconstruct it solely from the sensor's data. To investigate this, tri-axial inertial sensor data from five publicly available datasets were analysed. Three statistical properties: *mode*, *kurtosis*, and *number of modes* are shown to be sufficient for classification - assuming the sampling rate and sample format are known, and that both acceleration and magnetometer data is present. While those assumption hold, 98% of all 1003 data points were correctly classified.

**ACM Classification Keywords**

H.3.0 Information Storage and Retrieval: General

**Author Keywords**

Sensor Modality Identification; Inertial Data; Accelerometer; Gyroscope; Magnetometer; Activity Recognition;

## INTRODUCTION

To correctly interpret sensor data, reliable information about the data itself is required: sample (and frame) format, recording rate, number of axis, position at the observed body and sensor modality need to be known. This *meta-information* is often stored along-side the data itself, either in (semi-)structured external file, as a header of the data or as a well-known convention. Misinterpretation, resulting from missing or incorrect meta-data, has a strong influence on a subsequent application's performance. If such meta-data is not in a machine-interpretable format, slow and cumbersome recovery by a human expert becomes necessary. Here, we investigate to which extent the *sensor modality* can be recovered from invariant statistical properties of the sensor data itself. We assume that sample format, number of axis, and sample rate are known beforehand, but scale, and other calibration factors are not, and show that a sensor's modality can be (automatically) verified, or its meta-data rendered redundant.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

*ISWC '17,* September 11–15, 2017, Maui, HI, USA

© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ISBN 978-1-4503-5188-1/17/09...$15.00

DOI: https://doi.org/10.1145/3123021.3123048

Providing structured meta-data enables datasets to be picked by search engines [4]. In the absence or partial availability of this data, an automatic identification of sensor modality provides a basic starting point that otherwise would require manual inspection by an expert. Webcrawlers could refrain from *correctly* specified meta-data. Opportunistic sensing [7], i.e. situations where the sensor type is not known beforehand, would be another application area. While proper data curation practises could alleviate these situations, and are arguably more straightforward, an error in such manually defined data is often found much later. Manually reconstructing meta-data is then hard to scale to large data collections, as inspection by a human expert is required. A second system, which identifies modality from data directly, could at least provide additional safety checks. Such kind of quality control allows to (automatically) check if datasets were correctly documented or if there might be errors in the data collection, for example when uploading into a public dataset repository.

Activity Recognition applications are build with assumptions about the data retrieved from wearable inertial sensors. Properties, like placement variations or body locations are assumed, even though they directly influence the recognition performance. Kunze et.al. [6] have shown such influence, and provide several techniques to mitigate those effects. Namely, using location independent features, adding location to the classification task or estimating location from long-running recordings [6]. Similarly to the last option, we look at the properties of different sensor modalities over longer time periods to extract the sensor modality. Whether sensor data arose from an accelerometer, gyroscope or a magnetometer is usually stored as non-standardized meta-data, but also strongly influences the recognition performance if incorrect. Hammerla et.al. [2] introduced the empirical cumulative distribution function (ECDF) as a mean to capture the statistical properties of acceleration data, while also serving as a feature reduction method. Inspired by this, we characterize *invariant* statistical features that capture the properties of the inertial sensor modality. A system to correlate known datastreams to unknown ones and subsequently propagate their meta-data is described in [3]. In contrast our proposal does not require prior knowledge in the form of known sensor data, i.e. we present a ruleset that can be readily applied. However, our proposal is limited to inertial sensor data.

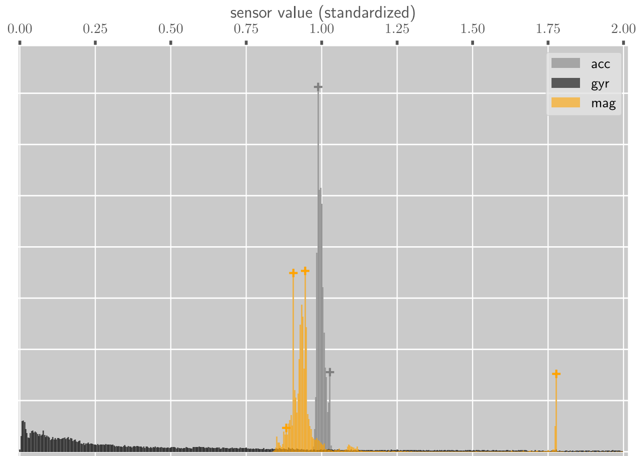


**Figure 1. Example histogram of three inertial data distributions of a CMU Kitchen dataset. The concentration of the gyroscope data around zero, as well as the concentration of the acceleration data around its mean, and the larger number of modes for magnetometer data is clearly visible. Identified modes on the distribution are highlighted.**

### DATASETS

Community provided datasets which included all three inertial sensor modalities, optionally mounted at different body positions were selected. These were converted into a common data format [11] to simplify their usage:

**CMU Kitchen** [8] contains inertial data of multiple body locations, including arms, legs and the back. Even though two inertial capture systems were recorded, only the wireless one, recording at 125Hz, was used. To balance with the other datasets, only a subset of 3h and 26 participants was extracted.

**ICS Forth ADL** [5] contains motion data from the thigh, ankle, torso and wrists. The measurement was taken at 50Hz. 15 participants were recorded for a total of 4.5h executing activities of daily living.

**Pamap2** [9] contains motion data of 8 participants executing activities of daily living at the hand, chest and ankles. In total 8h were recorded. Inertial data was recorded at 100Hz. Two acceleration (at different scales), one magnetometer and a gyroscope stream were used.

**Opportunity** [10] contains a whole-body inertial motion recording of daily living activities. A subset of 4 participants (with video recordings) contributed 15 data points each. In total 8h of data recorded at 30Hz was investigated.

**mHealthDroid** [1] recorded 12 activities of daily living. Shimmer nodes sampled at 100Hz provided the inertial data used in this paper. In total 6.5h were analyzed.

### PRE-PROCESSING

Each of the sensors differ in various aspects, which requires a few transformations prior to sensor identifications. In order to simplify the overall analysis, only the *magnitude* of sensor readings is used instead of its vector form. This is achieved by applying the L2-norm to each sensor reading, which also renders all subsequent calculations *orientation-independent*. Since sensor streams might be scaled differently, e.g. the two acceleration streams in the pamap dataset were recorded with 6g and 16g range, standardization is required. Dividing by the mean, i.e. *standardizing the scale* allows recordings to be compared. Additionally data was *lowpass-filtered* with a cutoff frequency of 2Hz. With the final assumption that the sensor is *most commonly at rest* on the body, we can now look at the properties of the transformed sensor streams, which we denote as $d$.

### GYROSCOPE

When the human body is at rest, no rotation is measurable. The rate-of-turn of a limb, measured by the gyroscope, is therefore most commonly near zero. This fact can be used to identify this sensor by the following rule:

$$mode(d) \approx 0 \Leftrightarrow gyr \tag{1}$$

Expressed differently, if the most common magnitude (mode) of sensor data is near zero, the sensor is a gyroscope and vice versa. This, however, only holds if the gyroscope data was baseline corrected.

### ACCELEROMETER

Due to being at rest, the accelerometer's mode of magnitude corresponds to the strength of earth's gravitational field. Designating the field strength with $g = 9.81\,\mathrm{m\,s^{-1}}$, we can formulate $mode(d_{acc}) \approx a * g$, where $mode(d_{acc})$ is the most commonly measured value, and $a$ an unknown scale factor applied to the data. If $a$ would be known, accelerometer data could be readily identified by comparisons to earth's gravitation. However, since data was lowpass filtered, the mean magnitude of acceleration corresponds to $g$ as well, i.e. $\tilde{d}_{acc} \approx a * g$. Due to standardization, we can formulate a rule for acceleration:

$$acc \Rightarrow mode(d) \approx 1 \tag{2}$$

Applying this rule to the scatter depicted in Figure 2 reveals why this is only a necessary condition; magnetometer data also fulfills this condition. A sufficient condition can be formulated for a subset of the overall accelerometer data, when including the kurtosis:

$$acc \Leftrightarrow mode(a) \approx 1 \text{ and } Kurt(d) > \alpha \tag{3}$$

Standardization is crucial for this condition, and relies on the assumption that the sensor is constantly accelerated by earth's gravitation. Other accelerations, due to limb movement for example, are only transient. Datasets which mostly contain strong movements, e.g. running or stirring as exemplary activities from the analysed data, will likely break this assumption. This is however tested with the mHealth, parts of the Opportunity and the Pamap dataset, which all contain sequences of strong, continuous motion.

### MAGNETOMETER

Figure 2 shows that some magnetometer readings exhibit a mode and kurtosis that is indistinguishable from accelerometer data. However, fluctuations in the measured magnetic field are more distinct than fluctuations of the gravity field. The respective distribution therefore is not uni- but multi-modal.

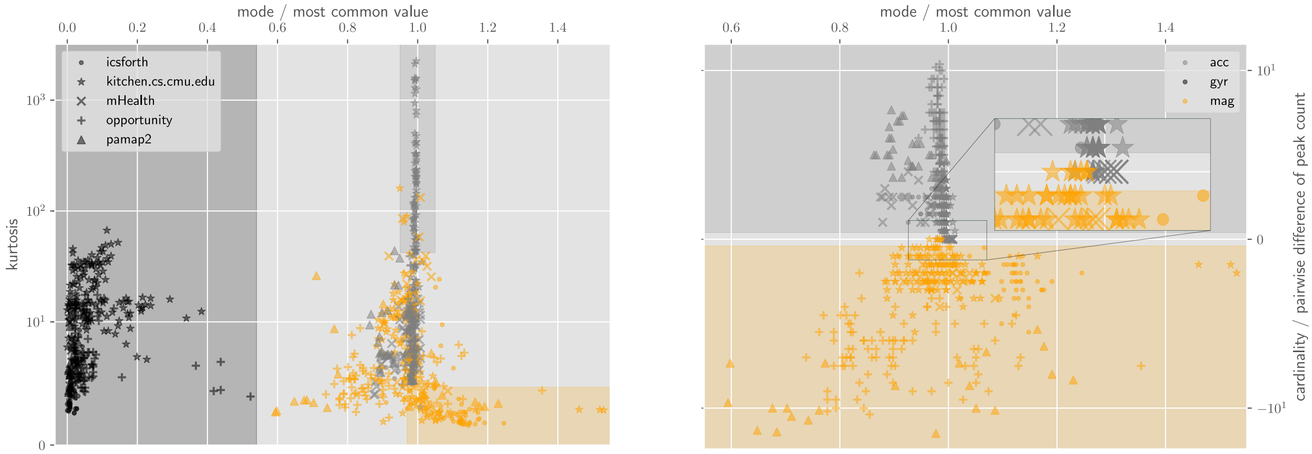


**Figure 2. Scatter plots of two possible feature sets for sensor modality detection. One feature is the mode of the histogram (512 equal-sized bins), i.e. the most common value. The second feature is either the kurtosis of the data, or the difference between the mean number of modes at the same limb and number of modes of one sensor stream. The left hand side shows that not all cases can be identified with mode and kurtosis only. The mode count difference provides a better indication, with the necessity to assume that both a magnetometer and accelerometer stream is present. Decision thresholds are shown as highlighted layers.**

This means there are multiple peaks, while the accelerometer distribution is rather "smooth" (cf. Fig. 1). A mode larger than the mean (or 1 in the standardized dataset), and a smaller kurtosis can indicate this:

$$mag \Leftarrow Kurt(d) < \beta \text{ and } mode(d) \geq \gamma \tag{4}$$

Whether such strong fluctuations are contained in the dataset depends on the experiment's condition. By proper choice of $\beta$ a subset of magnetometer data can be sufficiently identified. The smaller kurtosis can be explained due to the fact that magnetometer is often further spread out, and does not exhibit a strong concentration point. In contrast, acceleration data has a strong concentration and its kurtosis is higher.

## ACCELEROMETER VS. MAGNETOMETER

The question remains whether sensor streams, which fulfil none of the necessary conditions (3) nor (4) can still be identified. More directly, when it is not possible to decide between acceleration or magnetic flux based on kurtosis and mode alone. One observation that can be made about these cases, as well as the already identifiable cases, is that the number of modes for magnetometer is larger than the ones for acceleration data. Estimation of number of modes is achieved by adequately parameterized *peak detection* on the histogram. For a given stream, we designate the number of modes with $p$, as a shorthand for the number of peaks. However, streams have to be compared pair-wise, i.e. magnetometer and accelerometer must have observed the same motion. Let $\tilde{p}$ designate the mean number of modes of correlated sensor streams, then we can formulate the following condition:

$$d \Leftrightarrow \begin{cases} acc, & \text{if } \tilde{p} - p > .5 \\ mag, & \text{if } \tilde{p} - p < -.5 \\ unknown, & \text{otherwise} \end{cases} \tag{5}$$

Combined with condition (1) this allows to identify all sensor modalities, iff a correlated magnetometer and acceleration stream is to be distinguished.

## IDENTIFICATION RULESET

When only employing the sufficient conditions (3) and (4), we call this the *partial* ruleset. Together with the pair-wise condition (5), we can form a *full* ruleset:

$$d \Leftrightarrow \begin{cases} gyr, & \text{if } m < .5 \\ acc, & \text{else if } k \geq 42 \text{ and } .95 < m \leq 1.05 \\ mag, & \text{else if } k < 4.3 \text{ and } .97 < m \\ acc, & \text{else if } \tilde{p} - p > .5 \\ mag, & \text{else if } \tilde{p} - p < -.5 \\ unknown, & \text{otherwise} \end{cases} \tag{6}$$

where $m = mode(d)$ designates the mode of the data, $k = Kurt(d)$ its kurtosis, $p$ the total number of modes and $\tilde{p}$ the mean number of peaks of correlated data streams contained in one dataset.

Prior to applying above condition the data needs to be lowpass filtered, to exclude all frequencies above 2Hz. To reduce scaling effects, a standardization, by dividing by the mean of each stream was applied. The mode is determined from a histogram of 512 equal-sized bins, ranging from 0-2. Peak detection parameters were set to a minimum peak height of $.01 * m$, minimum distance of $5 bins$ and a minimum neighbor difference of $.008 * m$. These constants, as well as the decision thresholds in (6) were empirically determined.

## RESULTS AND LIMITATIONS

In total 1003 streams with durations ranging from 7min to 1h were analyzed. All three inertial sensor modalities are included, mostly positioned at the lower arm (61%), the upper body (20%) and the legs (19%). Data is scaled differently for each included dataset, showing that the proposed ruleset is independent of particular scale. Similarly, the sampling rates for each dataset differ. The *full* ruleset allows to identify 98%

Table 1. **Confusion matrices for sensor modality identification with *full* (left-hand) and *partial* (right-hand) ruleset. The full ruleset fails to identify $2\%$ of the analysed streams, but correctly identifies them for further manual inspection.**

| | - | acc | gyr | mag | | - | acc | gyr | mag |
|---|---|---|---|---|---|---|---|---|---|
| - | | 16 | | 6 | - | | 276 | | 205 |
| acc | | 339 | | | acc | | 78 | | 6 |
| gyr | | | 324 | | gyr | | | 324 | |
| mag | | | | 318 | mag | | 1 | | 113 |

of all cases, while 2% remain for manual inspection. If streams can not be compared pair-wise, the *partial* ruleset can still identify 51% of all cases, of those less than 1% are wrongly classified, while the remaining require manual inspection.

One could argue that, since threshold and features were designed from, and tested on the same set of data points, the proposed ruleset will not generalize to unseen streams and datasets, i.e. do we observe an over-fitted solution to this classification task? This could be answered by maximizing the classification score by a search of parameters (lowpass cutoff frequency, peak detection parameters, thresholds of (6) ...) on leave-one-dataset-out splits. In the worst case, there is no choice of parameters that performs equally well across all splits, i.e. there is no generalizing set of parameters - best case, a *single* set of parameters which performs well across all splits is found. Figure 2 shows that even when leaving out one datasets from training, points from another set lie next to the decision boundary. However, not all parameters are chosen based on these data points alone (in contrast to what a machine learning approach would do): (1) the mode threshold is based on the insight that gyroscope data is concentrated near zero, (2) the pair-wise peak threshold follows the observation that the magnetometer distribution exhibits more modes. This is the case for 98% of the observed data points. The latter observation has examples in multiple datasets, as is visible in Figure 2, partially ruling out an over-fit. A cross-validated *automatic* choice of parameters would reveal if the opposite was true, in a formal way. Here we merely report a single set of parameters that worked - but a better choice of parameters that maximizes the decision boundaries may well be possible.

A limitation of this work is the "critical mass", i.e. how many minutes of inertial data are required to make a decision about the sensor modality. The full dataset was used each time for feature computation. Varying this parameter would yield insights into the size of this mass, however was not attempted to avoid over-estimating the quality of the decision. Furthermore, standardization by dividing by the mean can be problematic if the sensor was asymmetrically driven into saturation. For example when the magnetometer was exposed to unipolar magnetic interference. In such cases, the mode could be nearer to zero yielding an incorrect classification. A possible solution could be to filter outliers.

## CONCLUSION
By applying the conditions formulated in equation (6) on an inertial sensor data stream it is possible to classify its modality, if the sensor was worn on the human body. Addtionally to estimating the modality from sensor data, this also allows to control the quality of meta-data, check the quality of data itelf, to opportunistically sense from unknown sensors, and to ease the interpretation of unstructured meta-data. To decide whether a sensor was a gyroscope will most probably generalize, as the mode of its distribution is a strong indicator (c.f. Figure 2). However, the decision between acceleration and magnetometer is more challenging, and as shown, only the pair-wise comparison of peak count provides a clear indication. This rests on the presence of environmental fluctuations on the magnetometer data, which might be smaller when less movement is involved. Given those assumptions, the ruleset correctly classifies 98% of 1003 streams in 5 different human motion datasets.


## ACKNOWLEDGMENTS AND REPRODUCIBILITY
We thank the authors of the datasets for providing them for public use. Software, dataset and results for reproduction are available at `https://github.com/pscholl/imustat`.


## REFERENCES

1. Oresti et.al. Banos. 2014. mHealthDroid: a novel framework for agile development of mobile health applications. In *International Workshop on Ambient Assisted Living*. Springer.
2. Nils Hammerla and Reuben Kirkham. 2013. On Preserving Statistical Characteristics of Accelerometry Data using their Empirical Cumulative Distribution. In *International Symposium on Wearable Computers*.
3. Melanie Hartmann, Alexander Bauer, and Ulf Blanke. 2013. Method and system for sensor classification. US Patent 7062320. (June 2013).
4. Google Inc. 2017. Google Developer Guides to Dataset Curaton. (2017). `https://developers.google.com/search/docs/data-types/datasets`
5. Katerina et.al. Karagiannaki. 2016. A benchmark study on feature selection for human activity recognition. In *ACM Ubicomp: Adjunct*. ACM.
6. Kai Kunze and Paul Lukowicz. 2014. Sensor placement variations in wearable activity recognition. In *2014 International Symposium on Wearable Computers*.
7. Marc et.al. Kurz. 2012. The OPPORTUNITY Framework and Data Processing Ecosystem for Opportunistic Activity and Context Recognition. 2 (2012).
8. De la Torre et.al. 2008. Guide to the Carnegie Mellon University Multimodal Activity (cmu-mmac) Database. In *Robotics Institute*.
9. Attila Reiss, Gustaf Hendeby, and Didier Stricker. 2013. Confidence-based multiclass AdaBoost for physical activity monitoring. In *International Symposium on Wearable Computers*. ACM.
10. Daniel et.al. Roggen. 2010. Collecting complex activity datasets in highly rich networked sensor environments. In *Int. Conf. Networked Sens. Syst.*
11. Philipp M Scholl and Kristof Van Laerhoven. 2016. A multi-media exchange format for time-series dataset curation. In *ACM Ubicomp: Adjunct*. ACM.