Compressed Signal Representation for Inertial Sensor Signals

Christoph Amma
Karlsruhe Institute of Technology (KIT)
Kronenstraße 12
Karlsruhe, 76131, Germany
christoph.amma@kit.edu

Hannes Volk
Karlsruhe Institute of Technology (KIT)
Kronenstraße 12
Karlsruhe, 76131, Germany
hannes.volk@student.kit.edu

Tanja Schultz
Karlsruhe Institute of Technology (KIT)
Kronenstraße 12
Karlsruhe, 76131, Germany
tanja.schultz@kit.edu

UbiComp'13 Adjunct, September 8–12, 2013, Zurich, Switzerland.
ACM 978-1-4503-2215-7/13/09.

http://dx.doi.org/10.1145/2494091.2494140

Abstract
We present and evaluate a method to generate a compressed representation of multi-dimensional inertial sensor signals using a piecewise linear approximation. The representation can be computed on small sensor nodes and thus allows for a reduction of the amount of data that needs to be transmitted to the main processing node. On an existing gesture database, we present the compression rate that is reached and evaluate the quality of the representation in terms of the accuracy reached for gesture classification. We compare the results to our baseline system using a simpler approach for data reduction.

Author Keywords
signal compression, gesture recognition, inertial sensors

ACM Classification Keywords
I.5.4 [Pattern Recognition]: Applications; E.4 [Coding and Information Theory]: Data compaction and compression

Inertial sensors like accelerometers and gyroscopes have become widely used in ubiquitous and wearable computing to sense motion and orientation of objects, devices and humans. Characteristic patterns, e.g. gestures or activities, are often recognized using pattern matching techniques. For this purpose, the data is usually captured with a given sampling rate and the sensor samples are used directly as features for a classification system. This typically involves a high redundancy of information in the raw data and our hypothesis is that these signals can be compressed while retaining the relevant information for further recognition of characteristic motion. In this preliminary study, we investigate the impact of our proposed compression method on classification accuracy and compare it to a baseline compression method.

We propose a multidimensional signal representation using piecewise linear approximations and evaluate the quality of the representation in terms of signal compression and recognition performance with an existing state of the art gesture recognizer. We use a dataset from our airwriting system, which can recognize handwriting performed in the air by an inertial measurement unit placed on the back of the hand [1]. We use a subset of our data containing recordings of single characters written in the air. In total, the classification task comprises 26 classes, which is typical for many gesture recognition systems proposed recently. We use Hidden Markov Models for classification which have proofed to perform very well for these kind of problems [2, 1]. We therefore deem our results can be transfered to other similar tasks.

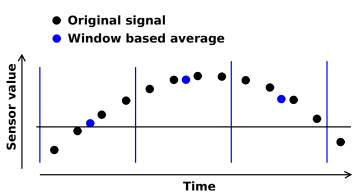

Figure 1: Illustration of the sliding window based average method used for the baseline system.

Signal representation

The signals we are dealing with in this study are six-dimensional inertial sensor signals from a triaxial accelerometer and a triaxial gyroscope.

Baseline system

For the baseline system we use a simple sliding window based approach to compress the signal. A non-overlapping sliding window is used and on each window the average signal value is computed (see figure 1), resulting in a six-dimensional vector for each window. By choosing the window length, the quality of the signal approximation can be adjusted against the compression rate. We will choose a range of compression rates to evaluate this representation.

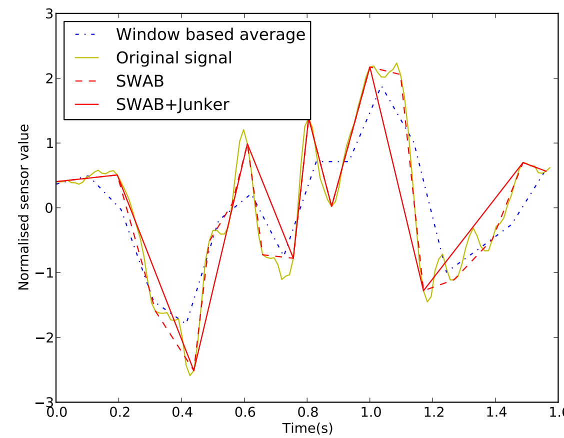

Figure 2: Results of the SWAB and SWABju algorithms for an example 1d-signal together with the baseline method for a compression rate of eight.

Multidimensional SWAB

Our proposed signal representation builds upon the SWAB algorithm by Keogh et al. [3]. The SWAB algorithm computes a piecewise linear approximation of a one-dimensional signal, i.e. the signal is approximated by a sequence of connected line segments and can be represented by the segment boundaries. It combines a bottom-up with a sliding window approach allowing for online operation while retaining a high quality approximation. The space and time complexity is linear and the algorithm can be easily implemented on small sensor nodes. Junker et al. [2] propose an extension to the

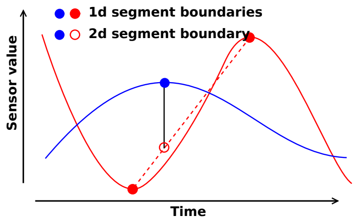

Figure 3: Illustration of the multi-dimensional version of SWAB for the 2d case. First 1d segments are created for each dimension, then 2d segments are created by interpolation between the segment boundaries of the remaining dimension.

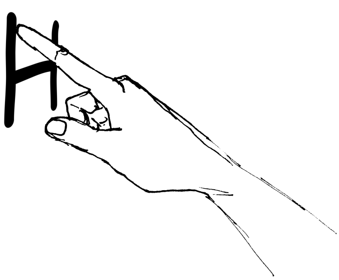

Figure 4: Recording setup.

SWAB algorithm that merges adjacent segments based on slope similarity of the linear regressions to generate a more compressed representation than SWAB alone. We use different angles between adjacent slopes as merging criteria to fine tune the compression rate of the algorithm. We refer to this algorithm as *SWABju*. Figure 2 depicts the results of the two variants of the algorithm.

To extend the algorithm to the multidimensional case, we compute a piecewise linear approximation for each dimension independently using the SWABju algorithm. We combine the segment boundaries of the individual dimensions to acquire the resulting multidimensional segmentation (*SWABmd*) by interpolating the values in the missing dimensions. Figure 3 illustrates the algorithm. The segmentation can be represented as a sequence of triples, where each triple corresponds to a segment boundary in one of the signal dimensions. A triple consists of the value of the segment boundary, the time index and the dimension in which it occurred. The time index can be given relative to the last triple to keep the absolute value low. For the calculation of the signal compression rate, we assume that the time index and the dimension together need at most as much bits as the values of the sensor samples which is in our case 16 bits. Therefore a triple can be expressed in 32 bits. The interpolation imposes a small delay on the reconstruction from the triples, since the algorithm needs to wait until a segment boundary appeared in all dimensions. There was enough activity in all six dimensions in our dataset, so that this delay could be neglected.

Dataset

We evaluate the features generated from the different representations within a gesture recognition task. We used data from our existing gesture database described in [1]. The database contains gestures of the 26 upper case characters of the Latin alphabet. Nine users contributed to our dataset, each of whom wrote 25 times the alphabet, which makes 650 recordings per person. Users wrote characters into the air as if they were writing on an imaginary blackboard in front of them. Figure 4 illustrates the recording setup. An inertial measurement unit containing a triaxial accelerometer and a triaxial gyroscope was attached to the back of the hand to sense hand motion. Originally data was recorded with a very high sampling rate of 819.2 Hz, but we downsampled it by a factor of 10 and thus all data we use in this study has a sampling rate of approximately 82 Hz, which is a more realistic setting. Consequently, the given compression rates are given relative to 82 Hz.

Feature extraction and classification

We use the compressed 6d signal vectors as features. For the baseline system, this is the average on each window, for the SWABmd, this are the segment boundaries. We use Hidden Markov Models (HMM) for classification, i.e. for each of the 26 characters, one HMM is trained.

Experiments and results

We perform a leave-one out cross-validation to evaluate the person-independent performance of the classifier. We report the average error rate over all persons. We evaluate the performance of the baseline system for compression rates (original data rate divided by compressed data rate) of 1 to 11. Depending on the merging criterion for adjacent segments, SWABmd reaches compression rates from 6 to 8 on our data set which equals 16.6 % to 12.5 % of the original data. Figure 5 depicts the results of our experiments. It shows the error rates plotted against the data compression ratio. The dashed line shows the results of the baseline system, the single points marks the error rates reached using the SWABmd method.

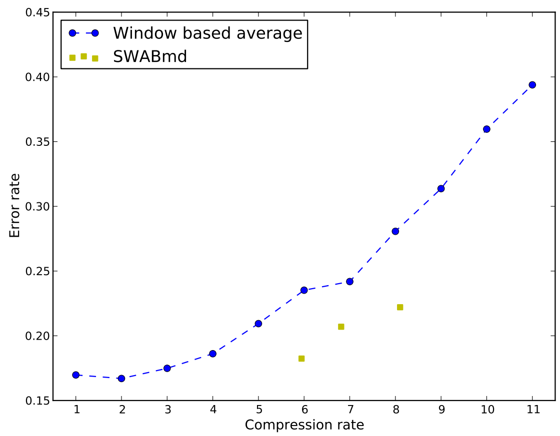

Figure 5: Error rates for the SWABmd method and the baseline system.

The baseline system performs best when using a compression rate of 2 (50 % of the original signal) reaching an error rate of 16.7 %. This means a number of 41 samples per second have to be transmitted. For greater window lengths and thus higher compression rates, the error rate increases. The SWABmd based system reaches an error rate of 18.2 % for a compression rate of 6 and an error rate of 22.2 % for a compression rate of 8. The baseline system has an error rate of 23.5 % and 28.0 % respectively for these compression rates.

Conclusion and future work

We propose a multidimensional signal representation for inertial sensor signals using a piecewise linear approximation based on the SWAB algorithm and its extension by Junker. The signal is represented as sequence of segment boundaries and can be computed online. We evaluate the representation within a gesture recognition task and compare it to our baseline system which uses a simple sliding window based average to represent the signal in a compressed form. Our method reaches compression rates of 6 to 8 in relation to the 82 Hz sampling rate of the original signal. We reach an accuracy of 18.2 % for the SWABmd method compared to 23.5 % for the baseline system at the same compression rate. We can conclude two main points from the presented study. If the size of transmitted data is critical, the SWABmd method offers a better compression versus accuracy ratio than the baseline method. From the evaluation of the baseline system we learn that a number of 30 to 40 feature samples per second is sufficient to reach the highest accuracy. Although our dataset can be considered typical for common gesture recognition systems, the results must be checked on other datasets, which we plan to do in the future.

References

[1] Amma, C., Georgi, M., and Schulz, T. Airwriting: A wearable handwriting recognition system. *Personal and Ubiquitious Computing* (2013).

[2] Junker, H., Amft, O., Lukowicz, P., and Tröster, G. Gesture spotting with body-worn inertial sensors to detect user activities. *Journal on Pattern Recognition 41*, 6 (2008), 2010–2024.

[3] Keogh, E., Chu, S., Hart, D., and Pazzani, M. Segmenting time series: A survey and novel approach. In *IEEE International Conference on Data Mining* (2001).