Unsupervised Motion Artifact Detection in Wrist-Measured Electrodermal Activity Data

Yuning Zhang, Maysam Haghdan, and Kevin S. Xu

ISWC 2017

COLLEGE OF ENGINEERING
THE UNIVERSITY OF TOLEDO

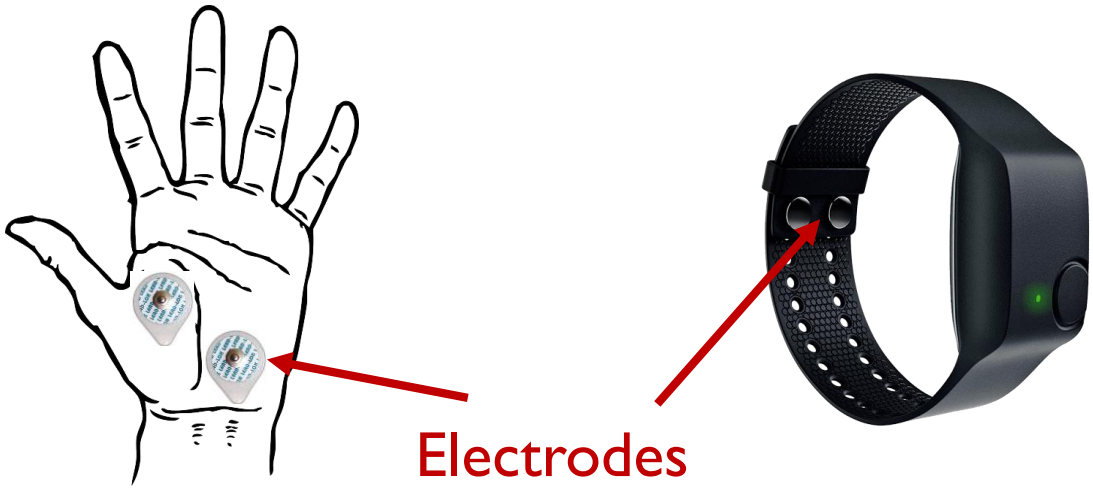
Wearables and Physiological Data

- Wearables allow collection of physiological data in minimally invasive manner
  - Heart rate via photoplethysmogram (PPG)
  - Electrodermal activity (EDA) via galvanic skin response (GSR)
- Data collected from wearables often contain many motion artifacts (MAs)
- Our focus: detection of MAs in wrist-measured EDA data

THE UNIVERSITY OF TOLEDO 1872

Electrodermal Activity (EDA)

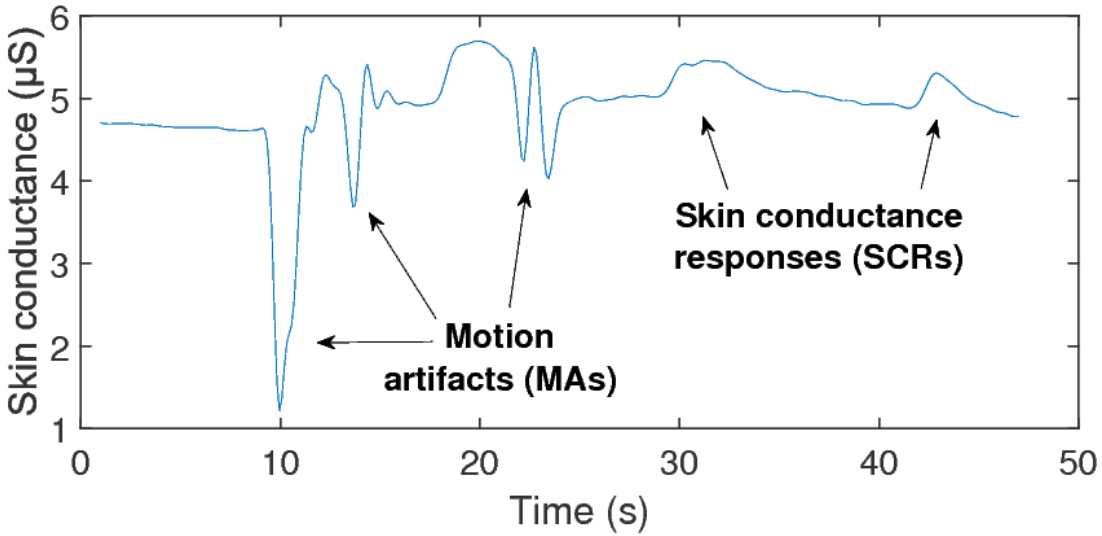
- Reflects the emotional and sympathetic responses of a person
  - Applications include estimation of stress levels, seizure detection, analysis of mental health disorders, etc.
- Measured via galvanic skin response (GSR)
  - Change in electrical conductance between a pair of electrodes touching the skin
  - Usually referred to as skin conductance (SC)



THE UNIVERSITY OF TOLEDO 1872

Motion Artifacts (MAs) in EDA

- Caused by change in amount of contact between the skin and the electrodes
  - Possible causes: motion, rotation, bumping the wearable, etc.
- MAs might be misidentified as skin conductance responses (SCRs)
  - SCRs and MAs can both generate a peak in SC
- Identification of the SC portions that contain MAs becomes extremely important as a pre-processing step



THE UNIVERSITY OF TOLEDO 1872

Prior Work on MAs in EDA

- MA suppression: pass entire SC signal through a smoothing filter (Sano and Picard, 2013; Hernandez et al., 2014; Chen et al., 2015)
  - Distorts SC signal including SCRs
- Redundancy: use two independent EDA sensors and compare the two (Hedman, 2010)
  - Impractical for general use
- MA detection: automatically detect portions of SC with MAs using machine learning (Taylor et al., 2015)
  - Trained supervised machine learning algorithms on a small EDA data set collected in a lab environment
  - Requires lots of human effort to label MAs!

Our Contributions

- Apply 8 different machine learning algorithms for MA detection: 5 supervised and 3 unsupervised
  - Supervised: classification
  - Unsupervised: anomaly detection
- Evaluate on 1 lab-based and 1 real-world EDA dataset totaling ~23 hours of data
- Examine the usefulness of the accelerometer data in identifying MAs

THE UNIVERSITY OF TOLEDO 1872

Datasets Used

- UT Dallas Stress (UTD) Data (13 hours lab-collected)
  - 20 college students subject to 3 types of stress: physical, cognitive, and emotional
- Alan Walks Wales (AWW) Data (10 hours real-world)
  - Collected by Alan Dix while he walked around Wales for ~3 months
  - We extracted 5 hours of walking data and 5 hours of resting data
- We use both EDA and 3-axis accelerometer (ACC) data collected by Affectiva Q sensor

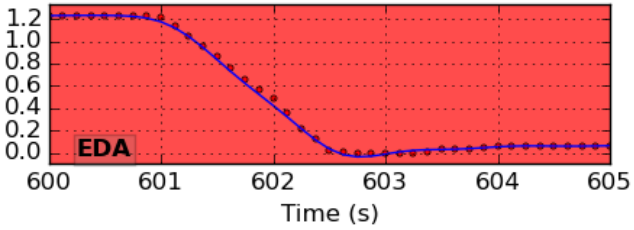
THE UNIVERSITY OF TOLEDO 1872

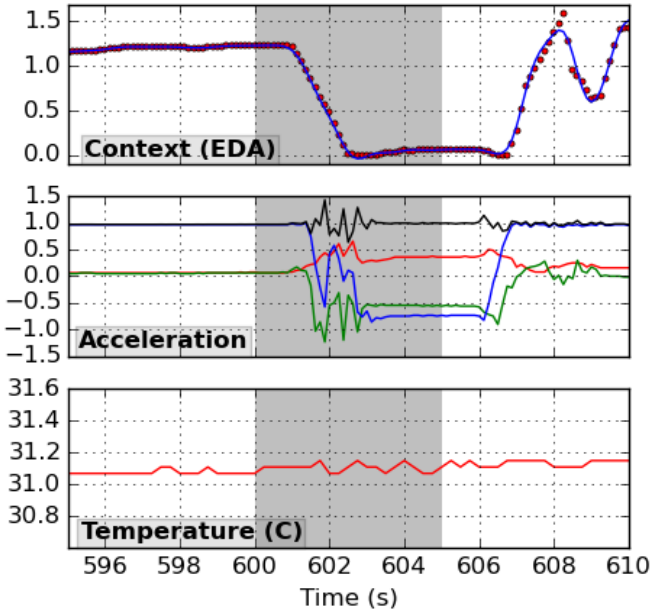
Methods Used

- Split EDA and ACC data into 5-second time windows and compute a variety of statistical features
- Feed feature set into machine learning algorithms
  - 5 supervised algorithms for binary classification:
    - Support vector machines
    - k-nearest neighbor classifiers
    - Random forests
    - Logistic regression
    - Multilayer Perceptron
  - 3 unsupervised algorithms for anomaly detection:
    - 1-class support vector machines
    - k-nearest neighbor distances
    - Isolation forests

Experiment Set-Up

- 3 experts use EDA Explorer software to label windows as MA or clean following a common set of criteria
  - Combine labels by majority vote
- Evaluate on 3 different feature sets separately:
  - ACC features only, EDA only, and ALL (ACC+EDA)
- Code and data available at https://github.com/IdeasLabUT/EDA-Artifact-Detection





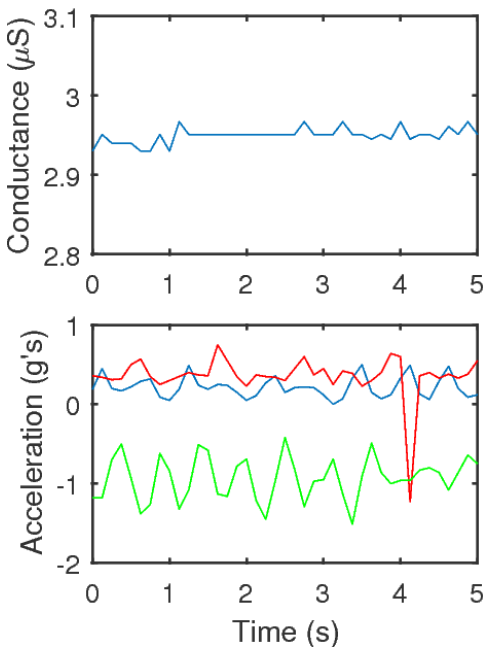
THE UNIVERSITY OF TOLEDO 1872

Overview of results

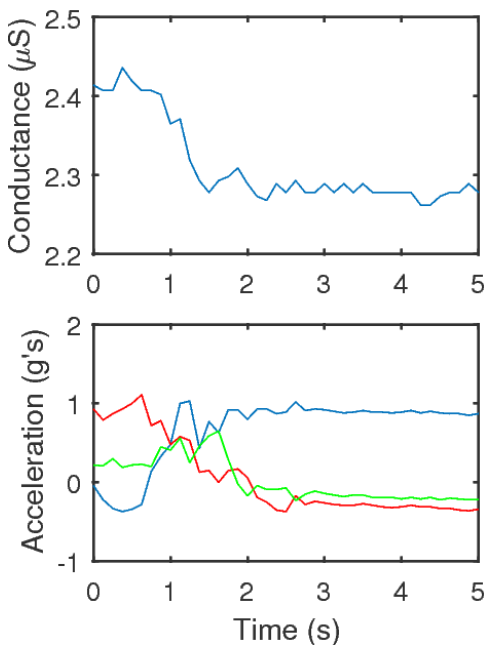
- Using ACC only features results in poor predictor: AUCs around 0.6-0.8
- Using EDA only or ALL (ACC+EDA) features results in much stronger predictor: AUCs around 0.8-0.95
- Unsupervised algorithms perform very competitively and sometimes better than supervised!
  - Both for in-sample and out-of-sample prediction
  - Best predictor on AWW data is unsupervised kNN distances (AUC of 0.90 on resting and 0.85 on walking)
- Using ALL features provides minimal benefit compared to EDA only
  - AUCs of supervised algorithms improve by 0.4%
  - AUCs of unsupervised algorithms decrease by 4.3%

Why ACC Appears Not to Help

- On many occasions, sudden changes in ACC do not affect EDA at all!
  - Motion that may not affect contact between electrodes and skin
- Maybe chosen features are not good for MA detection using ACC



Example where using ACC leads to incorrect prediction



Example where using ACC leads to correct prediction

THE UNIVERSITY OF TOLEDO 1872

Summary

- Found unsupervised ML algorithms to be highly competitive with supervised algorithms for automatically detecting MAs in EDA
  - Highly accurate MA detection is possible without requiring significant human effort for labeling data!
  - Evaluation on ~23 hours of data in both lab- and real-world settings
- Accelerometer data does not appear to be very helpful in MA detection
  - Perhaps due to poor feature construction
- Next steps: incorporate output of ML algorithms into automatic SCR detection algorithms

THE UNIVERSITY OF TOLEDO 1872