Philips Wearable Sensing Technologies: relevant Philips publications

General

1. Coppola, G., Conference keynote "Smart Wearable systems: vision and challenges", Wearable Technologies Conference 2008, July 07 2008, Munich.
2. Coppola, G., Conference presentation "Philips Optical Heart Rate Measurement", Sports and Technology Conference, 29-30 October 2008, Eindhoven.
3. Coppola, G., Conference presentation "Smart Wearable Systems: Trends and Challenges", Printed Electronics Europe 2009, 7-8 April 2009, Dresden.
4. Coppola, G., Conference presentation "Wearable monitoring technologies for healthcare and wellness", Medica Medicine + Sports Conference, 21 November 2013, Düsseldorf.
5. Haakma, R., Conference presentation "Self-management of health and disease - Vision on the future", EIT Innovation Day, 25 November 2015, Eindhoven.

Sleep

1. Long, X., Fonseca, P., Foussier, J., Haakma, R., and Aarts, R.M., "Using dynamic time warping for sleep and wake discrimination," *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI'12)*, pp. 886-889, Hong Kong, Jan. 2012.
Awarded 'first runner-up' (second price) in the student paper competition of BHI'12.

Abstract - In previous work, a Linear Discriminant (LD) classifier was used to classify sleep and wake states during single-night polysomnography recordings (PSG) of actigraphy, respiratory effort and electrocardiogram (ECG). In order to improve the sleep-wake discrimination performance and to reduce the number of modalities needed for class discrimination, this study incorporated Dynamic Time Warping (DTW) to help discriminate between sleep and wake states based on actigraphy and respiratory effort signal. DTW quantifies signal similarities manifested in the features extracted from the respiratory effort signal. Experiments were conducted on a dataset acquired from nine healthy subjects, using an LD-based classifier. Leave-one-out cross-validation shows that adding this DTW-based feature to the original actigraphy- and respiratory-based feature set results in an epoch-by-epoch Cohen's Kappa agreement coefficient of $\kappa = 0.69$ (at an overall accuracy of 95.4%), which represents a significant improvement when compared with the performance obtained without using this feature. Furthermore it is comparable to the result obtained in the previous work which used additional ECG features ($\kappa = 0.70$).

2. De Bruijn, R., Møst, E.I.S., Raymann, R.J.E.M., Haakma, R. & Markopoulos, P., "What is good sleep? Subjective Sleep Evaluation by Laymen", poster at the *14th edition of the annual international clinical symposium 'Update@kempenhaeghe.nl'*, Heeze, March 2012, as well as at the *12th Congress of the European Sleep Research Society (ESRS '12)*, Paris, September 2012.

Abstract: Most sleep quality measures take a scientific and/or clinical perspective, targeting quantification of sleep related behavior and/or clinically relevant sleep anomalies. However, it is not known to what extent existing sleep quality questionnaires characterize lay concepts of sleep quality.
The aim of this study is to identify how laymen describe their sleep quality intuitively, using a sentence stem completion methodology. Preliminary results indicate that people's own description of their ideal sleep does not necessarily match the subscales mentioned in standard sleep questionnaires.

3. Long, X., Fonseca, P., Haakma, R., Aarts, R.M., and Foussier, J., "Time-frequency analysis of heart rate variability for sleep and wake classification," *12nd IEEE International Conference on BioInformatics and BioEngineering (BIBE'12)*, pp. 85-90, Cyprus, Nov. 2012.
Awarded the first price in the student paper competition of BIBE 2012.

Abstract - This paper describes a method to adapt the spectral features extracted from heart rate variability (HRV) for sleep and wake classification. HRV series can be derived from electrocardiogram (ECG) signals obtained from single-night polysomnography (PSG) recordings. Traditionally, the HRV spectral features are extracted from the spectrum of an HRV series with fixed boundaries specifying bands of very low frequency (VLF), low frequency (LF), and high frequency (HF). However, because they are fixed, they may fail to accurately reflect certain aspects of autonomic nervous activity, which in turn may limit their discriminative power when using HRV spectral features, e.g., in sleep and wake classification. This is in part related to the fact that the sympathetic tone (partially reflected in the LF band) and the respiratory activity (modulated in the HF band) will vary over time. In order to minimize the impact of these differences, we adapt the HRV spectral boundaries using time-frequency analysis. Experiments conducted on a dataset acquired from 15 healthy subjects show that the discriminative power of the adapted HRV spectral features are significantly increased when classifying sleep and wake. Additionally, this method also provides a significant improvement of the overall classification performance when used in combination with some other (non-spectral) HRV features.

4. Long, X., Foussier, J., Fonseca, P., Haakma, R., and Aarts, R.M., "Respiration amplitude analysis for REM and NREM sleep classification," *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'13)*, pp. 5017-5020, Osaka, Japan, Jul. 2013.

Abstract - In previous work, single-night polysomnography recordings (PSG) of respiratory effort and electrocardiogram (ECG) signals combined with actigraphy were used to classify sleep and wake states. In this study, we aim at classifying rapid-eye-movement (REM) and non-REM (NREM) sleep states. Besides the existing features used for sleep and wake classification, we propose a set of new features based on respiration amplitude. This choice is motivated by the observation that the breathing pattern has a more regular amplitude during NREM sleep than during REM sleep. Experiments were conducted with a data set of 14 healthy subjects using a linear discriminant (LD) classifier. Leave-one-subject-out cross-validations show that adding the new features into the existing feature set results in an increase in Cohen's Kappa coefficient to a value of $\kappa = 0.59$ (overall accuracy of 87.6%) compared to that obtained without using these features (κ of 0.54 and overall accuracy of 86.4%). In addition, we compared the results to those reported in some other studies with different features and signal modalities.

5. Foussier, J., Fonseca, P., Long, X., and Leonhardt, S., "Automatic feature selection for sleep/wake classification with small data sets," *International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC2013)*, pp. 178-184, Barcelona, Spain, Feb. 2013.

Abstract - This paper describes an automatic feature selection algorithm integrated into a classification framework developed to discriminate between sleep and wake states during the night. The feature selection algorithm proposed in this paper uses the Mahalanobis distance and the Spearman's ranked-order correlation as selection criteria to restrict search in a large feature space. The algorithm was tested using a leave-one-subject-out cross-validation procedure on 15 single-night PSG recordings of healthy sleepers and then compared to the results of a standard Sequential Forward Search (SFS) algorithm. It achieved comparable performance in terms of Cohen's kappa ($\kappa = 0.62$) and the Area under the Precision-Recall curve ($AUC_{PR} = 0.59$), but gave a significant computational time improvement by a factor of nearly 10. The feature selection procedure, applied on each iteration of the cross-validation, was found to be stable, consistently selecting a similar list of features. It selected an average of 10.33 features per iteration, nearly half of the 21 features selected by SFS. In addition, learning curves show that the training and testing performances converge faster than for SFS and that the final training-testing performance difference is smaller, suggesting that the new algorithm is more adequate for data sets with a small number of subjects.

6. Fonseca, P., Long, X., Foussier, J., and Aarts, R.M., "On the impact of arousals on the performance of sleep and wake classification using actigraphy," *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'13)*, pp. 6760-6763, Osaka, Japan, Jul. 2013.

Abstract - We evaluated the impact of arousals on the performance of actigraphy-based sleep/wake classification. Using a dataset of 15 healthy adults and a threshold optimized for this task we found that the percentage of sleep epochs with activity counts above that threshold was significantly larger in epochs with and following arousals. We also found that 41.1% of all false positive classifications occurred in these epochs. Finally, we determined that excluding these epochs from the evaluation led to a maximum precision increase of 17.2%. Considering wake detections in those epochs as correct led to a maximum precision increase of 31.3%. We concluded that unless arousals can be automatically identified or at least distinguished from wake, the performance of actigraphy-based sleep/wake classifiers is limited by their presence.

7. Foussier, J., Fonseca, P., Long, X., Misgeld, B., and Leonhardt, S., "Combining HRV features for automatic arousal detection," *Computing in Cardiology (CinC'13)*, Zaragoza, Spain, Sep. 2013.

Abstract - Arousals are vital for sleep as they ensure its reversibility. However, an increased amount of arousals might indicate sleep disturbances or disorders. Since arousal events are similar to wake states but much shorter than the standard annotation epoch length of 30 s, they degrade sleep staging classification performance. Arousals are also related to physiological activities, such as cardiac activation, thus making the detection in a less disturbing way than with polysomnographies in sleep laboratories possible. Therefore, we analyzed 72 features derived from the heart rate variability (HRV) of 15 whole-night polysomnographic ECG recordings to quantify cardiac activation during sleep. After calculating the Mahalanobis distance (MD), ranking the best uncorrelated features and performing MANOVA, we show that combining multiple features increases the discriminative power (MD = 1.56, $\chi^2 = 33117$) to detect arousals during the night compared to the best single feature (MD = 1.16, $\chi^2 = 16633$). A linear mixed model is used to show between-subject effects and to validate the significance of each feature based on Wald test statistics.

8. Foussier, J., Long, X., Fonseca, P., Misgeld, B., and Leonhardt, S., "On the relationship of arousals and artifacts in respiratory effort signals," *IFMBE International Conference on Health Informatics*, vol. 42, pp. 31-34, Vilamoura, Portugal, Nov. 2013.

Abstract - Arousals are vital and thus very important for healthy sleep. Physiologically they manifest in cardiorespiratory signals or body movements. Generally, the acquisition of such signals is much easier than with the standard electroencephalogram (EEG). In this work we visually analyzed respiratory effort (RE) signals acquired with a respiratory induction plethysmography sensor (RIP) during whole-night polysomnography (including sleep stage and arousal annotations done with EEG) and annotated the artifacts. Artifacts are present when a change or distortion of the respiratory signal occurs. In total, the data from 15 subjects were acquired in two different sleep laboratories. The performance of detecting arousals only with the use of artifacts was evaluated. Since arousal and artifact sections are not always aligned in time, arousals have been widened by detection windows of 15 s and 30 s around it. If one artifact is present within this window the arousal was marked as detected. Median detection rates using this new approach of 69.81%, 77.36% and 83.02% were achieved for the original arousals on 15 s and 30 s window expansion, respectively. It is shown that in average 40.7% of the artifacts belong to the wake state, reducing the capability of detecting arousals that occur by definition only during sleep. During sleep, much more artifacts than arousals are present in the rapid eye movement (REM) stage, which is related to the fact that respiration is much more irregular during REM than during non-REM sleep and thus leading to increased artifacts.

9. Long, X., Fonseca, P., Haakma, R., Aarts, R. M., and Foussier, J., "Spectral boundary adaptation on heart rate variability for sleep and wake classification," *International Journal on Artificial Intelligent Tools*, 23(3):1460002, 2014.

Abstract - A method of adapting the boundaries when extracting the spectral features from heart rate variability (HRV) for sleep and wake classification is described. HRV series can be derived from electrocardiogram (ECG) signals obtained from single-night polysomnography (PSG) recordings. Conventionally, the HRV spectral features are extracted from the spectrum of an HRV series with fixed boundaries specifying bands of very low frequency (VLF), low frequency (LF), and high frequency (HF). However, because they are fixed, they may fail to accurately reflect certain aspects of autonomic nervous activity which in turn may limit their discriminative power, e.g. in sleep and wake classification. This is in part related to the fact that the sympathetic tone (partially reflected in the LF band) and the respiratory activity (modulated in the HF band) vary over time. In order to minimize the impact of these variations, we adapt the HRV spectral boundaries using time-frequency analysis. Experiments were conducted on a data set acquired from two groups with 15 healthy and 15 insomnia subjects each. Results show that adapting the HRV spectral features significantly increased their discriminative power when classifying sleep and wake. Additionally, this method also provided a significant improvement of the overall classification performance when used in combination with other HRV non-spectral features. Furthermore, compared with the use of actigraphy, the classification performed better when combining it with the HRV features.

10. Long, X., Fonseca, P., Foussier, J., Haakma, R., and Aarts, R.M., "Sleep and wake classification with actigraphy and respiratory effort using dynamic warping," *IEEE Journal of Biomedical and Health Informatics*, 18(4): 1272-1284, 2014.

Abstract - This paper proposes the use of dynamic warping (DW) methods for improving automatic sleep and wake classification using actigraphy and respiratory effort. DW is an algorithm that finds an optimal non-linear alignment between two series allowing scaling and shifting. It is widely used to quantify (dis)similarity between two series. To compare the respiratory effort between sleep and wake states by means of (dis)similarity, we constructed two novel features based on DW. For a given epoch of a respiratory effort recording, the features search for the optimally aligned epoch within the same recording in time and frequency domain. This is expected to yield a high (or low) similarity score when this epoch is sleep (or wake). Since the comparison occurs throughout the entire-night recording of a subject, it may reduce the effects of within- and between-subject variations of respiratory effort, and thus help discriminate between sleep and wake states. The DW-based features were evaluated using a Linear Discriminant classifier on a data set of 15 healthy subjects. Results show that the DW-based features can provide a Cohen's Kappa coefficient of agreement $\kappa = 0.59$ which is significantly higher than the existing respiratory-based features and is comparable to actigraphy. After combining the actigraphy and the DW-based features, the classifier achieved a κ of 0.66 and an overall accuracy of 95.7%, outperforming an earlier actigraphy- and respiratory-based feature set ($\kappa = 0.62$). The results are also comparable with those obtained using an actigraphy- and cardiorespiratory-based feature set but have the important advantage that they do not require an ECG signal to be recorded.

11. Fonseca, P., Aarts, R.M., Foussier, J., and Long, X., "A novel low-complexity post-processing algorithm for precise QRS localization," *SpringerPlus*, 3: 376, 2014.

Abstract - Precise localization of QRS complexes is an essential step in the analysis of small transient changes in instant heart rate and before signal averaging in QRS morphological analysis. Most localization algorithms reported in literature are either not robust to artifacts, depend on the sampling rate of the ECG recordings or are too computationally expensive for real-time applications. This paper proposes a localization algorithm based on the intersection of tangents fitted to the slopes of R waves detected by any QRS detector. Despite having a lower complexity, this algorithm achieves comparable trigger jitter to more complex localization methods without requiring the data to first be upsampled. It also achieves high localization precision regardless of which QRS detector is used as input. It is robust to clipping artifacts and to noise, achieving an average localization error below 5 ms even for recordings where the signal was severely degraded. Finally, it increases the accuracy of

template-based false positive rejection, allowing nearly all mock false positives added to a set of QRS detections to be removed at the cost of a very small decrease in sensitivity.

12. Goelema, M.S., Haakma, R. & Markopoulos, P. "Does being monitored during sleep affect people on a cognitive and behavioural level?", *7th International Conference on Health Informatics (ICHI '14)*, Angers, March 2014, pp. 27-33.

Abstract: Nowadays it is possible to monitor behavior or physiological features with specially-made devices that make self-monitoring an accessible and simple activity. Unknown is the effect these wearable devices may have on people's lives and this also applies to the area of sleep monitoring devices. The aim of this preliminary study is to address the extent to which sleep monitoring devices affect people on a cognitive and behavioral level. Four participants aged from 34 to 60, filled out a sleep diary for three consecutive weeks and wore in the latter two weeks a sleep monitoring device. Adjustments on a cognitive and behavioral level were observed, but this was probably due to participating in this study and completing the sleep diary as was indicated by the participants. Since the market for self-monitoring devices is rapidly developing and more accessible for lay people, it is important to investigate the reactive outcomes of these devices as they may have consequences for people who have a high adherence to self-control. Moreover, the knowledge about self-monitoring will improve which will lead to better interventions carried out by, for example, sleep coaches.

13. Goelema, M.S., Haakma, R. & Markopoulos, P. "The variation between nights should be taken into account when investigating the relationship between subjective and objective sleep measurements". SLEEP 2014 28, A98 – A98.

Abstract: Objective sleep measurements are expected to be predictors of the subjective sleep quality. The study investigates whether higher correlations between subjective and objective sleep measures will be found when the difference between two nights is considered rather than one specific night. Moderate to high correlations between SSA scores and physiological measures were found when considering the difference between the nights of the measurements in healthy controls and in the physiological illness group. However, for the mental disorder group the subjective sleep quality was mostly correlated to the duration of REM sleep on the second night.

14. Long, X., Haakma, R., Goelema, M.S., Weysen, T., Fonseca, P., Foussier, J., and Aarts, R.M., "Self-dissimilarity of respiratory effort across sleep states and time," *SLEEP*, 37(Abstract Suppl.), p. A36, Minneapolis, MN, May 2014.

Abstract: Respiratory activity strongly associates with sleep states. For instance, respiration is more regular during deep sleep compared with wakefulness. When awake, the respiratory regularity and the measurement of respiratory effort would be influenced by motion artifacts or other external factors. We therefore tested the hypothesis that the self-dissimilarity of respiratory signal morphology within a subject differs between sleep states, which would in turn help separate them. Moreover, the self-dissimilarity between two periods of respiratory signals might be in accordance with their time difference, which was investigated for each state. The conclusion is that sleep states can be differentiated using respiratory self-dissimilarity expressing the signal morphology which is usually evoked by the autonomic activity, the alternation of ventilation control or other external factors such as will or body movements. The lower self-dissimilarity score in short term implies the inclusion of nonrandom components of respiration which might be explained by less influence of body movements, presence of consciousness or memory of breathing control.

15. Long, X., Fonseca, P., Haakma, R., Foussier, J., and Aarts, R.M., "Automatic detection of overnight deep sleep based on heart rate variability: a preliminary study," *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'14)*, pp. 50-53, Chicago, IL, Aug. 2014.

Abstract: This preliminary study investigated the use of cardiac information or more specifically, heart rate variability (HRV), for automatic deep sleep detection throughout the night.
The HRV data can be derived from cardiac signals, which were obtained from polysomnography (PSG) recordings. In total 42 features were extracted from the HRV data of 15 single-night
PSG recordings (from 15 healthy subjects) for each 30-s epoch, used to perform epoch-by-epoch classification of deep sleep and non-deep sleep (including wake state and all the other sleep stages except deep sleep). To reduce variation of cardiac physiology between subjects, we normalized each feature per subject using a simple Z-score normalization method by subtracting the mean and dividing by the standard deviation of the feature values. A correlation-based feature selection (CFS) method was employed to select informative features as well as removing feature redundancy and a linear discriminant (LD) classifier was applied for deep and non-deep sleep classification. Results show that the use of Z-score normalization can significantly improve the classification performance. A Cohen's Kappa coefficient of 0.42 and an overall accuracy of 81.3% based on a leave-one-subject-out cross-validation were achieved.

16. Long, X., Haakma, R., Aarts, R.M., Fonseca, P., and Foussier, J., "Between-laboratory and demographic effects on heart rate and its variability during sleep," *Workshop on Smart Healthcare and Healing Environments (SHHE), European Conference on Ambient Intelligence (AmI'14),* , Eindhoven, The Netherlands, Nov. 2014.

Abstract: This work statistically analyzed the between-laboratory and demographic effects on cardiac activity during sleep in a multi-laboratory sleep monitoring study. A multilevel analysis was employed to evaluate these effects on two cardiac parameters expressing heart rate and heart rate variability. Results show that the two parameters vary between sleep stages and from subject to subject while there are no differences found between laboratories. In addition, these two parameters exhibit differently for subjects varied in their demographics.

17. Long, X., Foussier, J., Fonseca, P., Haakma, R., and Aarts, R.M., "Analyzing respiratory effort amplitude for automated sleep stage classification," *Biomedical Signal Processing and Control*, 14: 197-205, 2014.

Abstract: Respiratory effort has been widely used for objective analysis of human sleep during bedtime. Several features extracted from respiratory effort signal have succeeded in automated sleep stage classification throughout the night such as variability of respiratory frequency, spectral powers in different frequency bands, respiratory regularity and self-similarity. In regard to the respiratory amplitude, it has been found that the respiratory depth is more irregular and the tidal volume is smaller during rapid-eye-movement(REM) sleep than during non-REM (NREM) sleep. However, these physiological properties have not been explicitly elaborated for sleep stage classification. By analyzing the respiratory effort amplitude, we propose a set of 12 novel features that should reflect respiratory depth and volume, respectively. They are expected to help classify sleep stages. Experiments were conducted with a data set of 48 sleepers using a linear discriminant (LD) classifier and classification performance was evaluated by overall accuracy and Cohen's Kappa coefficient of agreement. Cross validations (10-fold) show that adding the new features into the existing feature set achieved significantly improved results in classifying wake, REM sleep, light sleep and deep sleep (Kappa of 0.38 and accuracy of 63.8%) and in classifying wake, REM sleep and NREM sleep (Kappa of 0.45 and accuracy of 76.2%). In particular, the incorporation of these new features can help improve deep sleep detection to more extent (with a Kappa coefficient increasing from 0.33 to 0.43). We also revealed that calibrating the respiratory effort signals by means of body movements and performing subject-specific feature normalization can ultimately yield enhanced classification performance.

18. Long, X., Fonseca, P., Aarts, R.M., Haakma, R., and Foussier, J., "Modeling cardiorespiratory interaction during sleep with complex networks," *Applied Physics Letters*, 105(20): 203701, 2014.

Abstract: Human sleep comprises several stages including wake, rapid-eye-movement sleep, light sleep, and deep sleep. Cardiorespiratory activity has been shown to correlate with sleep stages due to the regulation of autonomic nervous system. Here, the cardiorespiratory interaction (CRI) during sleep is analyzed using a visibility graph (VG) method that represents the CRI time series in complex networks. We demonstrate that the dynamics of the interaction between heartbeats and respiration can be revealed by VG-based networks, whereby sleep stages can be characterized and differentiated.

19. Long, X., Yang, J., Weysen, T., Haakma, R., Foussier, J., Fonseca, P., and Aarts, R.M., "Measuring dissimilarity between respiratory effort signals based on uniform scaling for sleep staging," *Physiological Measurement*, *35*(12): 2529-2542, 2014.

Abstract: Polysomnography (PSG) has been extensively studied for sleep staging, where sleep stages are usually classified as wake, rapid-eye-movement (REM) sleep, or non-REM (NREM) sleep (including light and deep sleep). Respiratory information has been proven to correlate with autonomic nervous activity that is related to sleep stages. For example, it is known that the breathing rate and amplitude during NREM sleep, in particular during deep sleep, are steadier and more regular compared to periods of wakefulness that can be influenced by body movements, conscious control, or other external factors. However, the respiratory morphology has not been well investigated across sleep stages. We thus explore the dissimilarity of respiratory effort with respect to its signal waveform or morphology. The dissimilarity measure is computed between two respiratory effort signal segments with the same number of consecutive breaths using a uniform scaling distance. To capture the property of signal morphological dissimilarity, we propose a novel window-based feature in a framework of sleep staging. Experiments were conducted with a data set of 48 healthy subjects using a linear discriminant classifier and a ten-fold cross validation. It is revealed that this feature can help discriminate between sleep stages, but with an exception of separating wake and REM sleep. When combining the new feature with 26 existing respiratory features, we achieved a Cohen's Kappa coefficient of 0.48 for 3-stage classification (wake, REM sleep and NREM sleep) and of 0.41 for 4-stage classification (wake, REM sleep, light sleep and deep sleep), which outperform the results obtained without using this new feature.

20. Goelema, M.S., Long, X. & Haakma, R. "Correlations between overnight breathing rate variation and subjective sleep quality scores. *Sleep-Wake Research in the Netherlands*, 25: 60-63, 2015.

Abstract: The relationship between objective sleep parameters, derived from polysomnography (PSG), and subjective sleep quality, obtained from questionnaires, has been researched thoroughly in the past. However, inconsistent results were found, the study outcomes differ to which extend above-mentioned variables are correlated and only a few objective parameters were related to the subjective sleep experience. The most profound association was between wake time and subjective sleep quality (r = -.59). Correlations between other objective measures, such as respiratory parameters, and subjective sleep quality have not yet been analyzed. We expect that a stable sleep, seen in, for example, a low breathing rate variation overnight, is indicative for a good sleep quality rating. In this preliminary work, we investigated whether respiratory parameters are related to subjective sleep quality the next morning. Breathing rate and its variation were found to be correlated with subjective sleep quality rating. The association between the breathing rate variation and the SSA score was more profound for women and seen to a greater extent for the first night in the elderly group. However, these correlations were not as high as we expected.

21. Goelema, M.S., Willems, M.M., Haakma, R. & Markopoulos, P. "On the reactivity of sleep monitoring with diaries and actigraphy". *9th International Conference on Health Informatics (ICHI '16), Rome, February 2016, p 240–247.*

Abstract: The declining costs of wearable sensors have made self-monitoring of sleep related behavior easier for personal use but also for sleep studies. However, it could be that wearing such devices impacts people's perception of their sleep or the very behavior that is being measured. Two small-scale field studies about the effects of sleep monitoring on a cognitive and a behavioral level are discussed. In the first study, we examine the self-monitoring effects of wearing a sleep monitoring device and in a second study the effects of keeping a sleep diary. The method, in both studies, was designed to be as open as possible in order to focus on the effects of sleep monitoring where participants are not given a goal, motivation or feedback. In both studies participants became more aware of their sleeping routine, but changing a sleeping habit was found challenging because of other priorities. Some behavioral modifications were observed, for example, differences in total sleep time and bedtimes were found (compared to a non-monitoring week and a monitoring week). Nevertheless, what the causes are of these changes remains unclear. It is important to know what the effects may be of sleep monitoring as the outcomes may already have an effect on the participant behavior which could cause researchers to work with data that do not represent a real life situation. In addition, the self-monitoring may serve as an intervention for facilitating healthier sleeping habits.

22. Long, X., Arends, J.B., Aarts, R.M., Haakma, R., Fonseca, P., and Foussier, J., "Time delay between cardiac and brain activity during sleep transitions," *Applied Physics Letters*, 106(14): 143702, 2015.

Abstract: Human sleep consists of wake, rapid-eye-movement (REM) sleep, and non-REM (NREM) sleep that includes light and deep sleep stages. This work investigated the time delay between changes of cardiac and brain activity for sleep transitions. Here, the brain activity was quantified by electroencephalographic (EEG) mean frequency and the cardiac parameters included heart rate, standard deviation of heartbeat intervals, and their low- and high-frequency spectral powers. Using a cross-correlation analysis, we found that the cardiac variations during wake-sleep and NREM sleep transitions preceded the EEG changes by 1–3 min but this was not the case for REM sleep transitions. These important findings can be further used to predict the onset and ending of some sleep stages in an early manner.

23. Long, X., Haakma, R., Leufkens, T.R.M., Fonseca, P., and Aarts, R.M., "Effects of between- and within-subject variability on autonomic cardiorespiratory activity during sleep and their limitations on sleep staging: a multilevel analysis," *Computational Intelligence and Neuroscience*, 2015: 583620, 2015.

Abstract: Autonomic cardiorespiratory activity changes across sleep stages. However, it is unknown to what extent it is affected by between- and within-subject variability during sleep. As it is hypothesized that the variability is caused by differences in subject demographics (age, gender, and body mass index), time, and physiology, we quantified these effects and investigated how they limit reliable cardiorespiratory-based sleep staging. Six representative parameters obtained from 165 overnight heartbeat and respiration recordings were analyzed. Multilevel models were used to evaluate the effects evoked by differences in sleep stages, demographics, time, and physiology between and within subjects. Results show that the between- and within-subject effects were found to be significant for each parameter. When adjusted by sleep stages, the effects in physiology between and within subjects explained more than 80% of total variance but the time and demographic effects explained less. If these effects are corrected, profound improvements in sleep staging can be observed. These results indicate that the differences in subject demographics, time, and physiology present significant effects on cardiorespiratory activity during sleep. The primary effects come from the physiological variability between and within subjects, markedly limiting the sleep staging performance. Efforts to diminish these effects will be the main challenge.

24. Fonseca, P., Long, X., Radha, M., Haakma, R., Aarts, R.M., and Rolink, J., "Sleep stage classification with ECG and respiratory effort," *Physiological Measurement*, 36(10): 2027-2040, 2015.

Abstract: Automatic sleep stage classification with cardiorespiratory signals has attracted increasing attention. In contrast to the traditional manual scoring based on polysomnography, these signals can be measured using advanced unobtrusive techniques that are currently available, promising the application for personal and continuous home sleep monitoring. This paper describes a methodology for classifying wake, rapid-eye-movement (REM) sleep, and non-REM (NREM) light and deep sleep on a 30 s epoch basis. A total of 142 features were extracted from electrocardiogram and thoracic respiratory effort measured with respiratory inductance plethysmography. To improve the quality of these features, subject-specific Z-score normalization and spline smoothing were used to reduce between-subject and within-subject variability. A modified sequential forward selection feature selector procedure was applied, yielding 80 features while preventing the introduction of bias in the estimation of cross-validation performance. PSG data from 48 healthy adults were used to validate our methods. Using a linear discriminant classifier and a ten-fold cross-validation, we achieved a Cohen's kappa coefficient of 0.49 and an accuracy of 69% in the classification of wake, REM, light, and deep sleep. These values increased to kappa = 0.56 and accuracy = 80% when the classification problem was reduced to three classes, wake, REM sleep, and NREM sleep.

25. Rolink, J., Kutz, M., Fonseca, P., Long, X., Misgeld, B., and Leonhardt, X., "Recurrence quantification analysis across sleep stages," *Biomedical Signal Processing and Control*, 20: 107-116, 2015.

Abstract: In this work we employ a nonlinear data analysis method called recurrence quantification analysis (RQA) to analyze differences between sleep stages and wake using cardio-respiratory signals, only. The data were recorded during full-night polysomnographies of 313 healthy subjects in nine different sleep laboratories. The raw signals are first normalized to common time bases and ranges. Thirteen different RQA and cross-RQA features derived from ECG, respiratory effort, heart rate and their combinations are additionally reconditioned with windowed standard deviation filters and ZSCORE normalization procedures leading to a total feature count of 195. The discriminative power between Wake, NREM and REM of each feature is evaluated using the Cohen's kappa coefficient. Besides kappa performance, sensitivity, specificity, accuracy and inter-correlations of the best 20 features with high discriminative power is also analyzed. The best kappa values for each class versus the other classes are 0.24, 0.12 and 0.31 for NREM, REM and Wake, respectively. Significance is tested with ANOVA F-test (mostly $p < 0.001$). The results are compared to known cardio-respiratory features for sleep analysis. We conclude that many RQA features are suited to discriminate between Wake and Sleep, whereas the differentiation between REM and the other classes remains in the midrange.

26. Goelema, M.S., Long, X., and Haakma, R., "Gender effect found in the association between overnight breathing rate variation and reported sleep quality scores," *SLEEP*, 38(Abstract Suppl.), p. A60, Seattle, WS, Jun. 2015.

Abstract: The relationship between objective sleep parameters, derived from polysomnography (PSG), and subjective sleep quality has been researched thoroughly in the past. Yet, correlations between other objective measures, such as respiratory parameters, and subjective sleep quality have not been analyzed. We expect that a stable sleep, seen in, for example, a low breathing rate variation overnight, is indicative for a good sleep quality rating. The association between breathing rate variation and the SSA score was more profound for females. However, these correlations were not as high as we expected. If future research can find a strong relationship between other objective sleep parameters and subjective sleep quality ratings, this would mean that predictions can be made about how someone has slept.

27. Long, X., Haakma, R., Fonseca, P., Aarts, R.M., Goelema, M.S., and Rolink, J., "What causes the differences in cardiac activity within and between subjects during sleep?" *SLEEP*, 38(Abstract Suppl.), p. A63, Seattle, WS, Jun. 2015.

Abstract: It is known that cardiac activity varies across sleep stages. However, it has not been quantitatively investigated in what aspects the cardiac activity is influenced by within-/between-subject differences. The differences can be caused by many factors such as subject demographics, time and (cardiac) physiology. We hypothesize that these factors affect the cardiac activity during sleep. Therefore, we try to quantify these effects leading to cardiac variations within and between subjects, which can be potentially used to help separate sleep stages. Demographics, time and within-/between-subject physiological differences have significant effects on cardiac activity during sleep. The major effects come from the differences within and between subjects in physiology, accounting for > 80% of total variance (except sleep stage). Practically, for cardiac-based sleep staging, the main challenge is to reduce these within-/between-subject differences.

28. Long, X., Haakma, R., Rolink, J., Fonseca, P., and Aarts, R.M., "Improving sleep/wake detection via boundary adaptation for respiratory spectral features," *37th International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'15)*, pp. 374-377, Milan, Italy, Aug. 2015.

Abstract: In previous work, respiratory spectral features have been successfully used for sleep/wake detection. They are usually extracted from several frequency bands. However, these traditional bands with fixed frequency boundaries might not be the most appropriate to optimize the sleep and wake separation. This is caused by the between-

subject variability in physiology, or more specifically, in respiration during sleep. Since the optimal boundaries may relate to mean respiratory frequency over the entire night. Therefore, we propose to adapt these boundaries for each subject in terms of his/her mean respiratory frequency. The adaptive boundaries were considered as those being able to maximize the separation between sleep and wake states by means of their mean power spectral density (PSD) curves overnight. Linear regression models were used to address the association between the adaptive boundaries and mean respiratory frequency based on training data. This was then in turn used to estimate the adaptive boundaries of each test subject. Experiments were conducted on the data from 15 healthy subjects using a linear discriminant classifier with a leave-one-subject-out cross-validation. We reveal that the spectral boundary adaptation can help improve the performance of sleep/wake detection when actigraphy is absent.

29. Long, X., Fonseca, P., Aarts, R.M., Haakma, R., Rolink, J., and Leonhardt, S., "Detection of nocturnal slow wave sleep based on cardiorespiratory activity in healthy adults," *IEEE Journal of Biomedical and Health Informatics*, 2015, accepted.

Abstract: Human slow wave sleep (SWS) during bedtime is paramount for energy conservation and memory consolidation. This work aims at automatically detecting SWS from nocturnal sleep using cardiorespiratory signals that can be acquired with unobtrusive sensors in a home-based scenario. From the signals, time-dependent features are extracted for continuous 30-s epochs. To reduce the measuring noise, body motion artifacts, and/or within-subject variability in physiology conveyed by the features and thus enhance the detection performance, we propose to smooth the features over each night using a spline fitting method. In addition, it was found that the changes in cardiorespiratory activity precede the transitions between SWS and the other sleep stages (non-SWS). To this matter, a novel scheme is proposed that performs the SWS detection for each epoch using the feature values prior to that epoch. Experiments were conducted with a large data set of 325 overnight polysomnography (PSG) recordings using a linear discriminant classifier and ten-fold cross validation. Features were selected with a correlation-based method. Results show that the performance in classifying SWS and non-SWS can be significantly improved when smoothing the features and using the preceding feature values of 5-min earlier. We achieved a Cohen's Kappa coefficient of 0.57 (at an accuracy of 88.8%) using only six selected features for 257 recordings with a minimum of 30-min overnight SWS that were considered representative of their habitual sleeping pattern at home. These features included the standard deviation, low-frequency spectral power, and detrended fluctuation of heartbeat intervals as well as the variations of respiratory frequency and upper and lower respiratory envelopes. A marked drop in Kappa to 0.21 was observed for the other nights with SWS time of less than 30 min which were found to more likely occur in elderly. This will be the future challenge in cardiorespiratory-based SWS detection.

30. P. Fonseca, R. M. Aarts, X. Long, J. Rolink, and S. Leonhardt, "Estimating actigraphy from motion artifacts in ECG and respiratory effort signals," Physiol. Meas., vol. 37, no. 1, pp. 67–82, 2016.

Abstract: Recent work in unobtrusive *sleep*/*wake* classification has shown that cardiac and respiratory features can help improve classification performance. Nevertheless, actigraphy remains the single most discriminative modality for this task. Unfortunately, it requires the use of actigraphy measuring devices in addition to the sensors used to measure electrocardiogram (ECG) or respiratory effort. This paper proposes a method to estimate actigraphy from the body movement artifacts present in the ECG and respiratory inductance plethysmography (RIP) signals based on the time-frequency analysis of those signals. Using a continuous wavelet transform to analyze RIP signals, and ECG and RIP signals combined, it provides a surrogate measure of actigraphy with moderate correlation (for ECG+RIP, $\rho = 0{:}74$, $p < 0{:}001$) and agreement (mean bias ratio of 0.94 and 95% agreement ratios of 0.11 and 8.45) with reference actigraphy. It can also be used as a replacement of actigraphy in sleep/wake classification: after cross-validation with a data set comprising polysomnographic (PSG) recordings of 15 healthy subjects and 25 insomniacs annotated by an external sleep technician, it achieves a statistically non-inferior classification performance when used together with respiratory features (average κ of 0:64 for 15 healthy subjects, and 0:50 for a dataset with 40 healthy and insomniac subjects), and when used together with respiratory and cardiac features (average κ of 0:66 for 15 healthy subjects, and 0:56 for 40 healthy and insomniac subjects). Since this method eliminates the need for a dedicated actigraphy device, it reduces the number of sensors needed for *sleep*/*wake* classification to a single sensor when using respiratory features, and to two sensors when using respiratory and cardiac features without any loss in performance. This method offers a major benefit in terms of comfort for long-term monitoring and is immediately applicable for legacy ECG and RIP monitoring devices already used in clinical practice and which do not have an accelerometer built-in.

31. P. Fonseca, X. Long, M. Radha, R. Haakma, R. M. Aarts, and J. Rolink, "Sleep stage classification with ECG and respiratory effort," IOP Physiol. Meas., vol. 36, pp. 2027–40, 2015.

Abstract: Automatic sleep stage classification with cardiorespiratory signals has attracted increasing attention. In contrast to the traditional manual scoring based on polysomnography, these signals can be measured using advanced unobtrusive techniques that are currently available, promising the application for personal and continuous home sleep monitoring. This paper describes a methodology for classifying wake, rapid-eye-movement (REM) sleep, and non-REM (NREM) light and deep sleep on a 30 s epoch basis. A total of 142 features were extracted from electrocardiogram and

thoracic respiratory effort measured with respiratory inductance plethysmography. To improve the quality of these features, subject-specific Z-score normalization and spline smoothing were used to reduce between-subject and within-subject variability. A modified sequential forward selection feature selector procedure was applied, yielding 80 features while preventing the introduction of bias in the estimation of cross-validation performance. PSG data from 48 healthy adults were used to validate our methods. Using a linear discriminant classifier and a ten-fold cross-validation, we achieved a Cohen's kappa coefficient of 0.49 and an accuracy of 69% in the classification of wake, REM, light, and deep sleep. These values increased to kappa = 0.56 and accuracy = 80% when the classification problem was reduced to three classes, wake, REM sleep, and NREM sleep

32. Pedro Fonseca, Niek den Teuling, Xi Long, and Ronald M. Aarts, "Cardiorespiratory Sleep Stage Detection Using Conditional Random Fields," accepted for publication in IEEE Journal of Biomedical and Health Informatics.

Abstract: This paper explores the probabilistic properties of sleep stage sequences and transitions to improve the performance of sleep stage detection using cardiorespiratory features. A new classifier, based on conditional random fields, is used in different sleep stage detection tasks (N3, NREM, REM, and wake) in night-time recordings of electrocardiogram and respiratory inductance plethysmography of healthy subjects. Using a dataset of 342 polysomnographic recordings of healthy subjects, among which 135 with regular sleep architecture, it outperforms hidden Markov models and Bayesian linear discriminants in all tasks, achieving an average accuracy of 87.38% and kappa of 0.41 (87.27% and 0.49 for regular subjects) for N3 detection, 78.71% and 0.55 (80.34% and 0.56 for regular subjects) for NREM detection, 88.49% and 0.51 (87.35% and 0.57 for regular subjects) for REM, and 85.69% and 0.51 (90.42% and 0.52 for regular subjects) for wake. In comparison with the state of the art, and having been tested on a much larger dataset, the classifier was found to outperform most of the work reported in the literature for some of the tasks, in particular for subjects with regular sleep architecture. It achieves a comparable accuracy for N3, higher accuracy and kappa for REM, and higher accuracy and comparable kappa for NREM than the best performing classifiers described in the literature.

Heartrate and artifact reduction

1. G Valenti, KR Westerterp "Optical heart rate monitoring module validation study", Consumer Electronics (ICCE), 2013 IEEE International Conference on; 2013/1/11; 195-196.

Abstract: Optical heart rate monitoring (OHRM) offers an unobtrusive solution for continuously measuring heart rate. An OHRM prototype, able to correct for movement artifacts during physical activity, proved to be valid to continuously monitor heart rate during activities including running, allowing monitoring cardiovascular condition in response to fitness and home activities.

2. Zwaag, M. van der, Lenssen, K.M.H. and Denissen, A.J.M., "Reading the human in the driver seat," Proceedings VISION 2014, Versailles, France, 14-15 October 2014.

Abstract: Unobtrusive physiological sensors that seamlessly integrate in the automotive environment provide a novel way to measure driver's health and well-being. A driver pilot was conducted to test the accuracy of some new unobtrusive and wearable physiological sensors that currently are under development. Six drivers drove a pre-defined route twice, once during light and once during dark conditions. The drive included different road types to be able to identify possible differences in measurement precision. Heart rate measurements were done using a standard ECG method as well as via two types of unobtrusive measurements. These included a watch equipped with a photo-plethysmograph and a driver seat that was equipped with capacitive ECG sensors. Skin conductance levels were measured by means of dry sensors at the finger and a wristwatch was used for unobtrusive skin conductance measurements. Results show high intraclass correlation between the unobtrusive and reference physiological measurements while driving in different conditions. This evaluation demonstrates the potential for unobtrusive tracking of driver's health and well-being in the car.

3. Ralph W.C.G.R. Wijshoff, Massimo Mischi, and Ronald M. Aarts. "Reduction of periodic motion artifacts in photoplethysmography", Transactions on Biomedical Engineering, 2016.

Abstract: Periodic motion artifacts affect photoplethysmography (PPG) signals in activities of daily living (ADL), cardiopulmonary exercise testing (CPX), and cardiopulmonary resuscitation (CPR). This hampers measurement of inter-beat-intervals (IBIs) and oxygen saturation (SpO2). Our objective was to develop a generic algorithm to remove periodic motion artifacts, recovering artifact-reduced PPG signals for beat-to-beat analysis.
METHODS: The algorithm was retrospectively evaluated on forehead PPG signals measured while walking on a treadmill. The step rate was tracked in a motion reference signal via a second order generalized-integrator with a frequency-locked loop. Two reference signals were compared: sensor motion relative to the skin (x[n]) measured via self-mixing interferometry, and head motion (av[n]) measured via accelerometry. The step rate was used in a quadrature harmonic model to estimate the artifacts. Quadrature components need only two coefficients per frequency leading to a

short filter, and prevent undesired frequency shifted components in the artifact estimate. Subtracting the estimate from the measured signal reduced the artifacts.
RESULTS: Compared to x[n], av[n] had a better signal-to-noise ratio and more consistently contained a component at the step rate. Artifact reduction was effective for distinct step rate and pulse rate, since the artifact-reduced signals provided more stable IBI and SpO2 measurements.
CONCLUSION: Accelerometry provided a more reliable motion reference signal. The proposed algorithm can be of significance for monitoring in ADL, CPX or CPR, by providing artifact-reduced PPG signals for improved IBI and SpO2 measurements during periodic motion.

4. Ralph Wijshoff. "On photoplethysmography artifact reduction and applications", PhD Thesis at Eindhoven University of Technology, Sept. 6 2016, with supervisors Ronald M. Aarts and M. Mischi.

Energy expenditure

1. Bonomi AG, Plasqui G, Goris AH, Westerterp KR. "Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer", J Appl Physiol (1985). 2009 Sep;107(3):655-61.

Abstract: Accelerometers are often used to quantify the acceleration of the body in arbitrary units (counts) to measure physical activity (PA) and to estimate energy expenditure. The present study investigated whether the identification of types of PA with one accelerometer could improve the estimation of energy expenditure compared with activity counts. Total energy expenditure (TEE) of 15 subjects was measured with the use of double-labeled water. The physical activity level (PAL) was derived by dividing TEE by sleeping metabolic rate. Simultaneously, PA was measured with one accelerometer. Accelerometer output was processed to calculate activity counts per day (AC(D)) and to determine the daily duration of six types of common activities identified with a classification tree model. A daily metabolic value (MET(D)) was calculated as mean of the MET compendium value of each activity type weighed by the daily duration. TEE was predicted by AC(D) and body weight and by AC(D) and fat-free mass, with a standard error of estimate (SEE) of 1.47 MJ/day, and 1.2 MJ/day, respectively. The replacement in these models of AC(D) with MET(D) increased the explained variation in TEE by 9%, decreasing SEE by 0.14 MJ/day and 0.18 MJ/day, respectively. The correlation between PAL and MET(D) (R(2) = 51%) was higher than that between PAL and AC(D) (R(2) = 46%). We conclude that identification of activity types combined with MET intensity values improves the assessment of energy expenditure compared with activity counts. Future studies could develop models to objectively assess activity type and intensity to further increase accuracy of the energy expenditure estimation.

2. Bonomi AG. "Towards valid estimates of activity energy expenditure using an accelerometer: searching for a proper analytical strategy and big data.", J Appl Physiol (1985). 2013 Nov 1;115(9):1227-8.

3. Valenti G, Bonomi AG, Westerterp KR. "Body Acceleration as Indicator for Walking Economy in an Ageing Population", PLoS One. 2015 Oct 29;10(10):e0141431.

Abstract:
BACKGROUND:
In adults, walking economy declines with increasing age and negatively influences walking speed. This study aims at detecting determinants of walking economy from body acceleration during walking in an ageing population.
METHODS:
35 healthy elderly (18 males, age 51 to 83 y, BMI 25.5±2.4 kg/m2) walked on a treadmill. Energy expenditure was measured with indirect calorimetry while body acceleration was sampled at 60Hz with a tri-axial accelerometer (GT3X+, ActiGraph), positioned on the lower back. Walking economy was measured as lowest energy needed to displace one kilogram of body mass for one meter while walking (WCostmin, J/m/kg). Gait features were extracted from the acceleration signal and included in a model to predict WCostmin.
RESULTS:
On average WCostmin was 2.43±0.42 J/m/kg and correlated significantly with gait rate ($r2 = 0.21$, $p<0.01$) and regularity along the frontal (anteroposterior) and lateral (mediolateral) axes ($r2 = 0.16$, $p<0.05$ and $r2 = 0.12$, $p<0.05$ respectively). Together, the three variables explained 46% of the inter-subject variance ($p<0.001$) with a standard error of estimate of 0.30 J/m/kg. WCostmin and regularity along the frontal and lateral axes were related to age (WCostmin: $r2 = 0.44$, $p<0.001$; regularity: $r2 = 0.16$, $p<0.05$ and $r2 = 0.12$, $p<0.05$ respectively frontal and lateral).
CONCLUSIONS:
The age associated decline in walking economy is induced by the adoption of an increased gait rate and by irregular body acceleration in the horizontal plane.

4. Valenti G, Camps SG, Verhoef SP, Bonomi AG, Westerterp KR. "Validating measures of free-living physical activity in overweight and obese subjects using an accelerometer", Int J Obes (Lond). 2014 Jul;38(7):1011-4.

Abstract:
BACKGROUND:
Free-living physical activity can be assessed with an accelerometer to estimate energy expenditure but its validity in overweight and obese subjects remains unknown.
OBJECTIVE:
Here, we validated published prediction equations derived in a lean population with the TracmorD accelerometer (DirectLife, Philips Consumer Lifestyle) in a population of overweight and obese. We also explored possible improvements of new equations specifically developed in overweight and obese subjects.
DESIGN:
Subjects were 11 men and 25 women (age: 41±7 years; body mass index: 31.0±2.5 kg m(-2)). Physical activity was monitored under free-living conditions with TracmorD, whereas total energy expenditure was measured simultaneously with doubly-labeled water. Physical activity level (PAL) and activity energy expenditure (AEE) were calculated from total energy expenditure and sleeping metabolic rate.
RESULTS:
The published prediction equation explained 47% of the variance of the measured PAL (P<0.001). PAL estimates were unbiased (errors (bias±95% confidence interval): -0.02±0.28). Measured and predicted AEE/body weight were highly correlated (r(2)=58%, P<0.001); however, the prediction model showed a significant bias of 8 kJ kg(-1) per day or 17.4% of the average AEE/body weight. The new prediction equation of AEE/body weight developed in the obese group showed no bias.
CONCLUSIONS:
In conclusion, equations derived with the TracmorD allow valid assessment of PAL and AEE/body weight in overweight and obese subjects. There is evidence that estimates of AEE/body weight could be affected by gender. Equations specifically developed in overweight and obese can improve the accuracy of predictions of AEE/body weight.

5. Plasqui G, Bonomi AG, Westerterp KR. "Daily physical activity assessment with accelerometers: new insights and validation studies", Obes Rev. 2013 Jun;14(6):451-62.

Abstract: The field of application of accelerometry is diverse and ever expanding. Because by definition all physical activities lead to energy expenditure, the doubly labelled water (DLW) method as gold standard to assess total energy expenditure over longer periods of time is the method of choice to validate accelerometers in their ability to assess daily physical activities. The aim of this paper was to provide a systematic overview of all recent (2007-2011) accelerometer validation studies using DLW as the reference. The PubMed Central database was searched using the following keywords: doubly or double labelled or labeled water in combination with accelerometer, accelerometry, motion sensor, or activity monitor. Limits were set to include articles from 2007 to 2011, as earlier publications were covered in a previous review. In total, 38 articles were identified, of which 25 were selected to contain sufficient new data. Eighteen different accelerometers were validated. There was a large variability in accelerometer output and their validity to assess daily physical activity. Activity type recognition has great potential to improve the assessment of physical activity-related health outcomes. So far, there is little evidence that adding other physiological measures such as heart rate significantly improves the estimation of energy expenditure.

6. Bonomi AG, Plasqui G. "'Divide and conquer': assessing energy expenditure following physical activity type classification", J Appl Physiol (1985). 2012 Mar;112(5):932.

7. Bonomi AG, Plasqui G, Goris AH, Westerterp KR. "Estimation of free-living energy expenditure using a novel activity monitor designed to minimize obtrusiveness", Obesity (Silver Spring). 2010 Sep;18(9):1845-51.

Abstract: The aim of this study was to investigate the ability of a novel activity monitor designed to be minimally obtrusive in predicting free-living energy expenditure. Subjects were 18 men and 12 women (age: 41 +/- 11 years, BMI: 24.4 +/- 3 kg/m(2)). The habitual physical activity was monitored for 14 days using a DirectLife triaxial accelerometer for movement registration (Tracmor(D)) (Philips New Wellness Solutions, Lifestyle Incubator, the Netherlands). Tracmor(D) output was expressed as activity counts per day (Cnts/d). Simultaneously, total energy expenditure (TEE) was measured in free living conditions using doubly labeled water (DLW). Activity energy expenditure (AEE) and the physical activity level (PAL) were determined from TEE and sleeping metabolic rate (SMR). A multiple-linear regression model predicted 76% of the variance in TEE, using as independent variables SMR (partial-r(2) = 0.55, P < 0.001), and Cnts/d (partial r(2) = 0.21, P < 0.001). The s.e. of TEE estimates was 0.9 MJ/day or 7.4% of the average TEE. A model based on body mass (partial-r(2) = 0.31, P < 0.001) and Cnts/d (partial-r(2) = 0.23, P < 0.001) predicted 54% of the variance in TEE. Cnts/d were significantly and positively associated with AEE (r = 0.54, P < 0.01), PAL (r = 0.68, P < 0.001), and AEE corrected by body mass (r = 0.71, P < 0.001). This study showed that the Tracmor(D) is a highly accurate instrument for predicting free-living energy expenditure. The miniaturized design did not harm the ability of the instrument in measuring physical activity and in determining outcome parameters of physical activity such as TEE, AEE, and PAL.

8. Bonomi AG, Westerterp KR. "Advances in physical activity monitoring and lifestyle interventions in obesity: a review", Int J Obes (Lond). 2012 Feb;36(2):167-77.

Abstract: Obesity represents a strong risk factor for developing chronic diseases. Strategies for disease prevention often promote lifestyle changes encouraging participation in physical activity. However, determining what amount of physical activity is necessary for achieving specific health benefits has been hampered by the lack of accurate instruments for monitoring physical activity and the related physiological outcomes. This review aims at presenting recent advances in activity-monitoring technology and their application to support interventions for health promotion. Activity monitors have evolved from step counters and measuring devices of physical activity duration and intensity to more advanced systems providing quantitative and qualitative information on the individuals' activity behavior. Correspondingly, methods to predict activity-related energy expenditure using bodily acceleration and subjects characteristics have advanced from linear regression to innovative algorithms capable of determining physical activity types and the related metabolic costs. These novel techniques can monitor modes of sedentary behavior as well as the engagement in specific activity types that helps to evaluate the effectiveness of lifestyle interventions. In conclusion, advances in activity monitoring have the potential to support the design of response-dependent physical activity recommendations that are needed to generate effective and personalized lifestyle interventions for health promotion.

9. G. Bonomi, S. Goldenberg, G. Papini, J. Kraal, W. Stut, F. Sartor, H. Kemps "Predicting Energy Expenditure from Photo-Plethysmographic Measurements of Heart Rate under Beta Blocker Therapy: Data Driven Personalization Strategies based on Mixed Models", 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), 2015.

Abstract: Energy expenditure have been often estimated using computational models based on heart rate (HR) and appropriate personalization strategies to account for users cardio-respiratory characteristics. However, medications like beta blockers which are prescribed to treat several cardiac conditions have a direct influence on the cardiovascular system and may impact the relationship between HR and energy expenditure during physical activity (AEE). This study proposes to estimate AEE from HR using mixed models (MIX-REG) by introducing a novel method to personalize the prediction equation. We selected as features to represent the individual random effect in the MIX-REG model those subject characteristics which minimized both estimation error (RMSE) and between-subjects error bias variability. Data from 17 patients post-myocardial infarction were collected during a laboratory protocol. AEE was measured using indirect calorimetry and HR using an innovative wrist worn activity monitor equipped with the Philips Cardio and Motion Monitoring Module (CM3-Generation-1), which is an integrated module including a photo-plethysmographic and accelerometer sensor. The presented method showed large AEE estimation accuracy (RMSE = 1.35 kcal/min) which was comparable to that of models personalized using data from laboratory calibration protocols (HR-FLEX) and was superior to multi-linear regression and MIX-REG models trained using a stepwise features selection procedure.

Cardio fitness

1. G. Bonomi "Protocol-free Assessment of Cardio-Respiratory Fitness from Daily Measurements of Body Movement and Heart Rate, Wearables@Work Conference, 19-20 April 2015, Amsterdam.

Respiration rate

1. Roxana Alexandra Cernat and Constantin Ungureanu and Mihaela Ungureanu and Ronald M. Aarts and Johan Arends, "Real-time extraction of the respiratory rate from photoplethysmographic signal using wearable devices", Workshop on Smart Healthcare and Healing Environments in conjunction with AMI'14, European Conference on Ambient Intelligence http://www.ami-conferences.org/2014 November 11 - 13, 2014, Eindhoven, The Netherlands (Also in Springer's Lecture Notes in Computer Science: Proceedings AmI14), 2014.

Abstract: The respiratory rate is considered one of the vital signs. Its real-time assessment in real-time is valuable information for clinicians, especially in conditions like epilepsy, sleep apnea or pulmonary disease. Current monitoring devices require a nasal cannula, a mask, or a chest band to deter-mine the respiratory rate, which can cause discomfort to the patient, and can be difficult to handle in emergency cases of emergency. We investigated the potential of photoplethysmography (PPG) to determine the respiratory rate, which is an easy-to-use and comfortable method. We investigated the potential of 12 features, derived from infrared and green PPG signals, to determine the respiratory rate. The features include the commonly used pulse width variability (PWV) or pulse amplitude variability (PAV). We showed that a new feature introduced by us provides similar results as PWV and PAV on the Capnobase database. Furthermore, we developed a data fusion model that uses five different PPG features to obtain the respiratory rate in real-time. We evaluated this real-time system against a thermistor respiratory circuit. The error was less than 2 breaths per minute in the 6-30 breaths/minute range. Our system successfully computed the respiratory rate when the sensor was placed on different body locations such as the upper arm, the wrist and the ankles.

2. T. Akkermans, C.Taal, D.Roovers, J.Gelissen, "PPG based Respiration Rate monitoring at the wrist in COPD patients", Wearables@Work Conference, 19-20 April 2015, Amsterdam.

Exercise type

1. Bonomi AG, Goris AH, Yin B, Westerterp KR. "Detection of type, duration, and intensity of physical activity using an accelerometer", Med Sci Sports Exerc. 2009 Sep;41(9):1770-7.

Abstract
OBJECTIVE:
The aim of this study was to develop models for the detection of type, duration, and intensity of human physical activity using one triaxial accelerometer.
METHODS:
Twenty subjects (age = 29 +/- 6 yr, BMI = 23.6 +/- 3.2 kg.m) performed 20 selected activities, including walking, running, and cycling, wearing one triaxial accelerometer mounted on the lower back. Identification of activity type was based on a decision tree. The decision tree evaluated attributes (features) of the acceleration signal. The features were measured in intervals of defined duration (segments). Segment size determined the time resolution of the decision tree to assess activity duration. Decision trees with a time resolution of 0.4, 0.8, 1.6, 3.2, 6.4, and 12.8 s were developed, and the respective classification performances were evaluated. Multiple linear regression was used to estimate speed of walking, running, and cycling based on acceleration features.
|RESULTS:
Maximal accuracy for the classification of activity type (93%) was reached when the segment size of analysis was 6.4 or 12.8 s. The smaller the segment size, the lower the classification accuracy achieved. Segments of 6.4 s gave the highest time resolution for measuring activity duration without decreasing the classification accuracy. The developed models estimated walking, running, and cycling speeds with a standard error of 0.20, 1.26, and 1.36 km.h, respectively.
CONCLUSIONS:
This study demonstrated the ability of a triaxial accelerometer in detecting type, duration, and intensity of physical activity using models based on acceleration features. Future studies are needed to validate the presented models in free-living conditions.

2. Gyllensten IC, Bonomi AG. "Identifying types of physical activity with a single accelerometer: evaluating laboratory-trained algorithms in daily life", IEEE Trans Biomed Eng. 2011 Sep;58(9):2656-63.

Abstract: Accurate identification of physical activity types has been achieved in laboratory conditions using single-site accelerometers and classification algorithms. This methodology is then applied to free-living subjects to determine activity behavior. This study is aimed at analyzing the reproducibility of the accuracy of laboratory-trained classification algorithms in free-living subjects during daily life. A support vector machine (SVM), a feed-forward neural network (NN), and a decision tree (DT) were trained with data collected by a waist-mounted accelerometer during a laboratory trial. The reproducibility of the classification performance was tested on data collected in daily life using a multiple-site accelerometer augmented with an activity diary for 20 healthy subjects (age: 30 ± 9; BMI: 23.0 ± 2.6 kg/m(2)). Leave-one-subject-out cross validation of the training data showed accuracies of 95.1 ± 4.3%, 91.4 ± 6.7%, and 92.2 ± 6.6% for the SVM, NN, and DT, respectively. All algorithms showed a significantly decreased accuracy in daily life as compared to the reference truth represented by the IDEEA and diary classifications (75.6 ± 10.4%, 74.8 ± 9.7%, and 72.2 ± 10.3%; $p < 0.05$). In conclusion, cross validation of training data overestimates the accuracy of the classification algorithms in daily life.

3. J. Margarito, R. Helaoui, A. Bianchi, F. Sartor, and A. G. Bonomi, "User-Independent Recognition of Sports Activities from a Single Wrist-worn Accelerometer: A Template Matching Based Approach", IEEE transactions on Biomedical Engineering, 2015.

Abstract: Goal: To investigate the accuracy of template matching for classifying sports activities using the acceleration signal recorded with a wearable sensor. Methods: A population of 29 normal weight and 19 overweight subjects was recruited to perform eight common sports activities, while body movement was measured using a triaxial accelerometer placed at the wrist. User- and axis-independent acceleration signal templates were automatically extracted to represent each activity category and recognize activity types. Five different similarity measures between example signals and templates were compared: Euclidean distance, dynamic time warping (DTW), derivative DTW, correlation and an innovative index, and combining distance and correlation metrics (Rce). Template-based activity recognition was compared to statistical-learning classifiers, such as Naive Bayes, decision tree, logistic regression (LR), and artificial neural network (ANN) trained using time- and frequency-domain signal features. Each algorithm was tested on data from a holdout group of 15 normal weight and 19 overweight subjects. Results: The Rce index outperformed other template-matching metrics by achieving recognition rate above 80% for the majority of the activities. Template matching showed robust classification accuracy when tested on unseen data and in case of limited training examples. LR and ANN achieved the highest overall recognition accuracy ~85% but showed to be more vulnerable to misclassification error than template matching on overweight subjects' data. Conclusion: Template matching can be used to classify sports activities using the wrist acceleration signal. Significance: Automatically extracted template prototypes from the acceleration signal may be used to enhance accuracy and generalization properties of statistical-learning classifiers.

Atrial Fibrillation

1. Alberto G Bonomi, F Schipper, L M Eerikäinen, J Margarito, R Aarts, S Babaeizadeh, H de Morree, L Dekker, "Atrial fibrillation detection using photo-plethysmography and acceleration data at the wrist", Computing in Cardiology (2016). Accepted.

Abstract:
Background: Atrial fibrillation (AF) is a pathological condition leading to increased risk for embolic stroke and cardiac hospitalizations. Screening for AF represents a technical challenge because of the paroxysmal and frequently asymptomatic nature of the condition.
Objective: The aim was to investigate whether an unobtrusive wrist-wearable device equipped with a photo-plethysmographic (PPG) and acceleration sensors could be used to detect AF.
Methods: Sixteen patients (M = 63%, age: 65.2 ± 14.0 y, BMI: 29.7 ± 7.0 kg/m2) with suspected paroxysmal AF were monitored for 24 hours in outpatient setting using a portable ECG Holter recorder. Simultaneously, a wrist-wearable device equipped with a PPG and acceleration sensor was used to monitor heart rhythm and body movement. PPG data were processed to extract the timing of heart beats and derive inter-beat-intervals (IBI). Acceleration data was used to discard IBI in presence of motion artifacts. An ECG validated first-order Markov model was used to assess the probability of irregularly irregular rhythm of AF being present from PPG-derived IBI. AF detection outcome from the algorithm was compared with adjudications of AF episodes provided by clinical experts after visually inspecting ECG Holter data.
Results: Four patients experienced 100% AF burden, while 1 patient suffered from atrial flutter. The remaining patients showed normal sinus rhythm with several premature beats (808 supraventricular, range 0 – 4879; and 656 ventricular premature beats, range 0 – 4795). AF detection was achieved with 97 ± 2% Sensitivity and 99 ± 3% Specificity. During atrial flutter the algorithm output was non-AF 94.6% of the time. Due to motion artifacts, the algorithm did not provide AF classification for 36 ± 9% of the 24 hours monitoring.
Conclusion: A wrist-wearable device equipped with a PPG and acceleration sensor can provide accurate detection of rhythm irregularities caused by atrial fibrillation in free-living conditions.

Blood Pressure

1. Shaoxiong Sun, Rick Bezemer, Xi Long, Jens Muehlsteff, and Ronald M. Aarts. "Systolic blood pressure estimation using PPG during physical exercise", 38th EMBC 2016, 2016.

Abstract: In this work, a model to estimate systolic blood pressure (SBP) using photoplethysmography (PPG) and electrocardiography (ECG) was proposed for subjects doing physical exercise. After an initialization process for each subject at rest, the model estimated SBP every 30 second for the whole period of exercise. In order to build this model, eighteen features were extracted from PPG signals by means of its waveform, first derivative, second derivative, and frequency spectrum. In addition, pulse arrival time (PAT) was derived as a feature from the combination of PPG and ECG. After evaluating four regression models, we chose multiple linear regression (MLR) to combine all derived features to estimate SBP. The contribution of each feature was quantified using its normalized weight in the MLR. To evaluate the performance of the model, we used a leave-one-subject-out cross validation. In the aim of exploring the potential of the model, we investigated the influences of initialization, the inclusion of PAT, regression models measurement sites (finger and forehead), and posture change. The results showed that the initialization process significantly reduced the root mean square error (RMSE). The inclusion of PAT significantly elevated correlation coefficients. There was no significant difference in the estimation performance between the model using finger- and forehead-derived PPG signals. Separate models have to be built for different postures. The optimized model using finger-derived PPG signals during physical exercise had the performance with a mean difference of -0.56 mmHg, a standard deviation of difference of 13.60 mmHg, and a median correlation coefficients of 0.87. The present work demonstrated promising results of the SBP estimation model during physical exercise.

Epilepsy detection

1. Judith van Andel and Constantin Ungureanu and Johan Arends and Ronald M. Aarts and Frans Leijten,, "Using photoplethysmography in heart rate monitoring of patients with epilepsy", Epilepsy & Behavior, 45, April, pp. 142--145, http://dx.doi.org/10.1016/j.yebeh.2015.02.018,

Abstract: Heart rate is a useful neurophysiological sign when monitoring seizures in patients with epilepsy. In an ambulatory setting, heart rate is measured with ECG involving electrodes on the skin. This method is uncomfortable which is burdensome for patients and is sensitive to motion artifacts, which decrease the usability of measurements. In this study, green light photoplethysmography, an optical technique arising from the fitness industry, was evaluated for usefulness in a medical setting. Simultaneous overnight measurements of HR with a commercially available optical heart rate (OHR) sensor and with ECG (HRECG) were performed in 7 patients with epilepsy. Overall, there was no significant difference between OHR and HRECG in random 10-minute periods during wakefulness (p#=#0.69) and sleep (p#=#1.00). The Bland-Altman analysis showed negligible mean differences. Limits of agreement were higher during wakefulness and during the occurrence of two seizures possibly because of less reliable HRECG measurements due to motion artifacts. Optical heart rate seems less sensitive to these motion artifacts, and measurements are more user-friendly. The optical heart rate sensor may fill the gap of systems for ambulatory heart rate monitoring and can be especially useful in the context of

seizure detection in patients with epilepsy.

2. Constantin Ungureanu and Vih Bui and Wouter Roosmalen and Ronald M. Aarts and Johan B.A.M. Arends and Richard Verhoeven and Johan J. Lukkien, "A wearable monitoring system for nocturnal epileptic seizures", IEEE 8th International Symposium on Medical Information and Communication Technology (ISMICT), Digital Object Identifier: 10.1109/ISMICT.2014.6825205, 2014

Abstract: For people suffering from nocturnal epileptic seizures it is crucial to have a system that can detect such seizures in real-time. In this paper, we present a preclinical demonstrator for real-time detection of nocturnal seizures based on the heart rate. The system is built on the VITRUVIUS body sensor platform, which consists of a body hub (a smart phone) and sensors communicating via Bluetooth. The seizure detection application running on the body hub has an online classifier that is triggered by an adaptable cumulative sum (CUSUM) algorithm. In case of an event, an alarm message can be sent to the seizure monitoring application or to the caregiver's phone. We present the architecture of the system with emphasis on the detection algorithm and the applications. The off-line evaluation of the system on five patients achieved a sensitivity of 95% and a positive predictive value of 85%. The preliminary results are encouraging and the system will be implemented in a clinical trial at Kempenhaeghe epilepsy clinic in the Netherlands.

Other metrics

1. Ralph W.C.G.R. Wijshoff, Antoine M.T.M. van Asten, Wouter H. Peeters, Rick Bezemer, Gerrit Jan Noordergraaf, Massimo Mischi, and Ronald M. Aarts. "Photoplethysmography-based algorithm for detection of cardiogenic output during cardiopulmonary resuscitation", IEEE Transactions on Biomedical Engineering, 62(3):909-921, March 2015.

Abstract: Detecting return of spontaneous circulation (ROSC) during cardiopulmonary resuscitation (CPR) is challenging, time consuming, and requires interrupting chest compressions. Based on automated-CPR porcine data, we have developed an algorithm to support ROSC detection, which detects cardiogenic output during chest compressions via a photoplethysmography (PPG) signal. The algorithm can detect palpable and impalpable spontaneous pulses. A compression-free PPG signal which estimates the spontaneous pulse waveform, was obtained by subtracting the compression component, modeled by a harmonic series. The fundamental frequency of this series was the compression rate derived from the transthoracic impedance signal measured between the defibrillation pads. The amplitudes of the harmonic components were obtained via a least mean-square algorithm. The frequency spectrum of the compression-free PPG signal was estimated via an autoregressive model, and the relationship between the spectral peaks was analyzed to identify the pulse rate (PR). Resumed cardiogenic output could also be detected from a decrease in the baseline of the PPG signal, presumably caused by a redistribution of blood volume to the periphery. The algorithm indicated cardiogenic output when a PR or a redistribution of blood volume was detected. The algorithm indicated cardiogenic output with 94% specificity and 69% sensitivity compared to the retrospective ROSC detection of nine clinicians. Results showed that ROSC detection can be supported by combining the compression-free PPG signal with an indicator based on the detected PR and redistribution of blood volume.