Video-based actigraphy for monitoring wake and sleep in healthy infants

Citation for published version (APA):

DOI:
10.3390/s19051075

Document status and date:
Published: 03/03/2019

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 25. Apr. 2020
Article

Video-Based Actigraphy for Monitoring Wake and Sleep in Healthy Infants: A Laboratory Study

Xi Long 1,2,* , Renée Otte 1, Eric van der Sanden 1, Jan Werth 2 and Tao Tan 3

1 Department of Family Care Solutions, Philips Research, 5656 AE Eindhoven, The Netherlands; renee.otte@philips.com (R.O.); eric.van.der.sanden@philips.com (E.v.d.S.)
2 Department of Electrical Engineering, Eindhoven University of Technology, 5612 AP Eindhoven, The Netherlands; jan.werth@philips.com
3 Department of Mathematics and Computer Science, Eindhoven University of Technology, 5612 AP Eindhoven, The Netherlands; tao.tan911@gmail.com

* Correspondence: xi.long@philips.com or x.long@tue.nl; Tel.: +31-6-3166-0077

Received: 18 January 2019; Accepted: 27 February 2019; Published: 3 March 2019

Abstract: Prolonged monitoring of infant sleep is paramount for parents and healthcare professionals for interpreting and evaluating infants’ sleep quality. Wake-sleep patterns are often studied to assess this. Video cameras have received a lot of attention in infant sleep monitoring because they are unobtrusive and easy to use at home. In this paper, we propose a method using motion data detected from infrared video frames (video-based actigraphy) to identify wake and sleep states. The motion, mostly caused by infant body movement, is known to be substantially associated with infant wake and sleep states. Two features were calculated from the video-based actigraphy, and a Bayesian-based linear discriminant classification model was employed to classify the two states. Leave-one-subject-out cross validation was performed to validate our proposed wake and sleep classification model. From a total of 11.6 h of infrared video recordings of 10 healthy term infants in a laboratory pilot study, we achieved a reliable classification performance with a Cohen’s kappa coefficient of 0.733 ± 0.204 (mean ± standard deviation) and an overall accuracy of 92.0% ± 4.6%.

Keywords: Infant monitoring; wake-sleep pattern; infrared camera; video-based actigraphy; classification

1. Introduction

During the first year of life, the average infant is asleep for a greater part of the day than they are awake [1], making sleep one of the most important activities for their brains [2]. Essential developmental processes take place during sleep, and problems with sleeping, therefore, may inhibit optimal development [3]. Sleep tracking and assessment can offer valuable information on an infant’s mental and physical development, not only for healthcare professionals, but also for parents [4,5]. This information, for instance, could enable tailored coaching for parents and help improve an infant’s sleep quality [6,7].

The most common way to monitor infant sleep is by polysomnography (PSG) [8], including multiple sensor modalities, such as electroencephalography, electrooculography, electromyography, electrocardiography, respiration, and others. A drawback of this method is that it requires obtrusive contact sensing on the body of an infant, which is undesirable in sick or fragile infants, and impractical in a home situation. Moreover, for reliable measures of wake-sleep patterns, prolonged monitoring over multiple nights is recommended [9,10], increasing the need for unobtrusive monitoring.

Over the past decades, a number of unobtrusive or minimally obtrusive techniques have been developed for objective infant (sleep) monitoring, such as actigraphy [9], capacitive sensing [11], ballistocardiography [12], photoplethysmography [13,14], laser Doppler vibrometers [15], and (near)
infrared (IR) or thermal video [16,17]. These techniques have been summarized in previous systematic reviews [18,19]. In particular, video-based monitoring appears to be a promising technique, since, in contrast to wearables, it is entirely unobtrusive, does not require mounting/unmounting, and does not need battery charging, making it convenient and easy to use in real home situations. Video-based approaches are capable of capturing body movements that are highly correlated with infant wake/sleep state [20]. A large number of studies that employed wrist or ankle actigraphy to measure gross body movements have achieved good performance in wake/sleep detection both for adults and children [21,22]. Heinrich et al. [23] demonstrated that video-based solutions are able to replace on-body actigraphic sensors for the application of sleep monitoring and analysis. Furthermore, video-based approaches potentially allow for the measurement of vital signs (e.g., heart rate and breathing rate) [17,24], which are of importance for automatically identifying infant sleep stages (active/quiet sleep).

In this paper, we describe a new method for reliably classifying infant wake and sleep states using motions derived from IR video frames (i.e., video-based actigraphy). The use of an IR video camera allows for the monitoring of infants in a dark environment without visible light. As this was an explorative, proof-of-concept study on the feasibility of using video-based actigraphy for classifying infant wake and sleep states, we started with a small pilot population of healthy infants.

2. Methods

2.1. Subjects and Data

Data from 10 healthy infants aged 5 months on average (range 3–9 months) were included in this study. The infants’ parents had been recruited through word-of-mouth, flyers aimed at young parents living close to Tilburg University, the website of the Tilburg University BabyLab, and meetings for expectant couples in the maternity ward of St. Elisabeth Hospital in Tilburg, The Netherlands. The exclusion criteria were health problems or problems related to feeding, sleeping, or development. Before consenting to have their baby participate, all steps of the study were explained in full to the infant participants’ parents. Both parents signed an informed consent form if they agreed to let their baby participate in the study. The study protocol was approved by both the Internal Ethics Committee for Biomedical Experiments of Philips Research Eindhoven, and the Psychological Ethical Test Committee of Tilburg University, The Netherlands. The study was performed in conformity with the declaration of Helsinki. As the study was observational in nature, no additional approval from an external ethics board was required—the Dutch law on medical scientific research with human beings (WMO) was not applicable.

Video recordings were obtained from an IR camera placed in a “look-down” view above an infant bed placed in the BabyLab at Tilburg University. The whole mattress of the bed was visible. For each infant video, data with an average duration of 1.16 ± 0.43 (mean ± standard deviation) hours was included, resulting in a total of 11.6 h used for the analyses. Synchronized to the IR videos sleep stages were scored as wake, rapid eye movement (REM) sleep, and non-REM (NREM) sleep including N1, N2, and N3 stages by a trained sleep specialist for non-overlapping epochs that lasted 30 s, adhering to the rules of the American Academy of Sleep Medicine [25] and those specifically for infant sleep scoring by Grigg-Damberger et al. [8]. The scoring was based on multichannel PSG signals collected using a commercially available PSG system (Vitaport3, TEMEC Instruments B.V., Heerlen, The Netherlands).

Since this study only focused on automatic classification of wake and sleep states, all annotated REM and N1–N3 sleep epochs were grouped into a single sleep state. Wake and sleep epochs accounted for 28.0% and 71.3% of the data, respectively. “Unable-to-score” epochs made up 0.7% of the data, and pertained to epochs for which the signal quality was poor (caused by poor or no connectivity of the PSG electrodes). The PSG annotations served as the golden standard for automated video-based classification of infant wake/sleep states.
A disadvantage of mACT is that it is challenging to identify wakefulness when body movement is reduced. Therefore, we characterized the ‘possibility’ of being asleep (pSLP) before and after a very high activity level [31]. This was done by quantifying the logarithm of the time difference between each epoch and its nearest epoch with a large amount of body movements (corresponding to a large mACT value). It was then smoothed through a moving average, with an experimentally optimized window of 10 min.

Two features were extracted on a 30 s basis from the raw video-based actigraphy signal. First, we calculated the mean activity count (mACT, i.e., count of non-zero motions) over video frames for each epoch, similar to wrist actigraphy, which is intensively utilized in wake/sleep identification [30]. A disadvantage of mACT is that it is challenging to identify wakefulness when body movement is reduced. Therefore, we characterized the “possibility” of being asleep (pSLP) before and after a very high activity level [31]. This was done by quantifying the logarithm of the time difference between each epoch and its nearest epoch with a large amount of body movements (corresponding to a large mACT value). It was then smoothed through a moving average, with an experimentally optimized window of 10 min. For each

![Figure 1](https://example.com/image1.png)

**Figure 1.** Example of video frames in a 1.35 h recording, and the estimated video-based actigraphy and wake/sleep state. The sampling rate of the video and the video-based actigraphy is 10 Hz. The face of the infant is blurred for privacy purposes.

### 2.3. Feature Extraction and Classification Model

Two features were extracted from the raw video-based actigraphy signal. First, we calculated the mean activity count (mACT, i.e., count of non-zero motions) over video frames for each epoch, similar to wrist actigraphy, which is intensively utilized in wake/sleep identification [30]. A disadvantage of mACT is that it is challenging to identify wakefulness when body movement is reduced. Therefore, we characterized the “possibility” of being asleep (pSLP) before and after a very high activity level [31]. This was done by quantifying the logarithm of the time difference between each epoch and its nearest epoch with a large amount of body movements (corresponding to a large mACT value). It was then smoothed through a moving average, with an experimentally optimized window of 10 min. For each
recording, an epoch with an mACT value larger than the 95th percentile of the mACT values over the entire recording was considered to have a high activity level. pSLP is expected to correctly identify some “low movement” wake epochs when they are very close to wake epochs with a lot of body movements (i.e., a high activity level). The hypothesis, thus, is that epochs with little movement that are close to those with a high level of activity (and thereby with a smaller time difference) are more likely to correspond to the infant being awake than to the infant being asleep.

To diminish the “global” variability between infants conveyed by video-based actigraphy, we needed to normalize the features. For each feature and each infant this was done with a Min–Max method to rescale the feature values within a range between 0 and 1. As can be seen in Figure 2, the two features mACT and pSLP show a high correlation with the infant wake-sleep pattern. Notably, although almost no body movements were observed during the wake period between the 10th and the 30th epochs (corresponding to low mACT values), the possibility of the infant being awake was high, due to the relatively low feature values of pSLP. This is because, for these wake epochs, the time to the epochs with a high activity level (high mACT) was relatively short.

A Bayesian-based linear discriminant classification model (classifier) was deployed, which has been successfully used in sleep classification in previous studies [21,32,33]. The classifier is based on Bayes decision rules for minimizing the probability of error, i.e., to choose the class that maximizes its posterior probability given an observation (feature vector). Note that wake was set to be the positive class and sleep the negative class in this dataset, and the classifier’s prior probability for both classes was equalized.

2.4. Validation and Assessment

For the two video-based actigraphic features statistical analysis (Mann–Whitney U test) was performed to examine the difference between all wake and sleep epochs from all infants.

To show the validity of the proposed classification model, we applied a (subject-independent) leave-one-subject-out cross validation (LOOCV). During each round of the LOOCV, we used data from nine infants to train the classifier, and that from the remaining infant to test the classifier. Ten rounds were conducted and the performance was reported by averaging classification results over all rounds.
Commonly used metrics, including accuracy, precision, sensitivity, and specificity, as well as the area under the “receiver operating characteristic” (ROC) curve (AUC) were considered to assess the wake-sleep classification performance. Since the binary class distribution was imbalanced, as the wake epochs accounted for only 28.0%, the chance-compensated metric Cohen’s kappa coefficient was also employed, and this metric has been widely used for evaluating sleep scoring or sleep staging reliability. In this work, the algorithms were implemented in MATLAB R2016b (MathWorks, Natick, MA, USA).

3. Results and Discussion

Mean and standard deviation of the two proposed features (mACT and pSLP, on a 30 s epoch basis) from video-based actigraphy from all recordings (or infants) during wake and sleep states are shown in Figure 3. A significant difference at $p < 0.0001$ was found between the two states for both features, indicating that they were effective in discriminating between wake and sleep epochs. However, there were still overlaps between wake and sleep states in both features, which would be difficult to correctly identify.

![Figure 3](image.png)

**Figure 3.** Boxplots of the two features (a) mACT and (b) pSLP during wake and sleep (of all epochs pooled over all infants) in normalized unit (n.u.), where wake epochs had higher mACT and lower pSLP than sleep epochs ($p < 0.0001$).

Table 1 compares the LOOCV wake and sleep classification results using different feature sets, i.e., mACT only, pSLP only, and their combination (mACT + pSLP). The best-performing result for accuracy and Cohen’s kappa is in bold. In addition, the performance comparison in the ROC space can be seen in Figure 4. Combining the two features can largely improve the wake and sleep classification performance, where a substantial agreement with a mean Cohen’s kappa coefficient of $0.733 \pm 0.204$, and a mean accuracy of $92.0\% \pm 4.6\%$ across infants was achieved.

Interestingly, the feature pSLP was able to help reduce the number of false negatives (i.e., number of wake epochs incorrectly classified as sleep epochs) and thus increase the precision and sensitivity. After a closer look, we found that these corrected false negatives mostly corresponded to wake epochs with little or even no infant body movements that were very close to epochs with a high mACT value (i.e., a lot of body movements). This can be further observed by comparing the ROC curves in Figure 4. For example, by changing the classifier’s decision-making threshold to have a specificity of ~55%, the sensitivity can go up to 100%, corresponding to zero false negatives when including pSLP, while it is ~90% when excluding pSLP.
Table 1. Comparison of wake and sleep classification results (leave-one-subject-out cross validation—LOOCV) with video-based actigraphy.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>mACT</th>
<th>pSLP</th>
<th>mACT + pSLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>True positive (TP)</td>
<td>304</td>
<td>307</td>
<td>318</td>
</tr>
<tr>
<td>False positive (FP)</td>
<td>36</td>
<td>143</td>
<td>35</td>
</tr>
<tr>
<td>False negative (FN)</td>
<td>85</td>
<td>82</td>
<td>71</td>
</tr>
<tr>
<td>True negative (TN)</td>
<td>954</td>
<td>847</td>
<td>955</td>
</tr>
<tr>
<td>Precision</td>
<td>82.4% ± 18.3%</td>
<td>61.0% ± 22.7%</td>
<td>82.3% ± 16.6%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>71.9% ± 24.4%</td>
<td>82.8% ± 18.9%</td>
<td>77.4% ± 22.7%</td>
</tr>
<tr>
<td>Specificity</td>
<td>95.9% ± 2.6%</td>
<td>84.8% ± 6.5%</td>
<td>95.8% ± 2.6%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>91.3% ± 4.7%</td>
<td>82.0% ± 7.9%</td>
<td>92.0% ± 4.6%</td>
</tr>
<tr>
<td>Cohen’s kappa</td>
<td>0.701 ± 0.227</td>
<td>0.544 ± 0.187</td>
<td>0.733 ± 0.204</td>
</tr>
</tbody>
</table>

1 For confusion matrix elements, results pooled over all epochs from all infants are presented; for the other metrics, mean ± standard deviation results across infants are presented. The highest accuracy and Cohen’s kappa are in bold.

Figure 4. Receiver operating characteristic (ROC) curves of each infant and all epochs pooled over all infants obtained, based on the proposed wake and sleep classification model using three different feature sets (a) mACT, (b) pSLP, and (c) mACT + pSLP, where the area under the curve (AUC) for all infants is 0.902, 0.877, and 0.952.

The classification accuracy and kappa for each infant are presented in Table 2. We noticed that the kappa value for Infant10 is very low (0.220). Video inspection showed this to be because this baby was taken out of bed for around one-fifth of the video recording, during which the PSG was still available with wires connected, whereas the infant was not visible through the video camera. For this infant, when discarding the out-of-bed period, the kappa and accuracy increased to 0.642 and 94.8%, respectively. Nevertheless, this is a drawback of using video camera for infant sleep monitoring: it requires the infant to be in the bed all the time. In a recent work [26], we have shown that it is possible to accurately identify whether an infant is in bed or out of bed (with an accuracy of 96.9%), so that out-of-bed periods can be removed before doing wake and sleep classification. Apart from that, the performance variability across infants remains relatively high (with the kappa ranging from 0.629 to 0.939). One explanation is that for some recordings external disturbances (e.g., from parental activity) also contributed to the motion captured by the videos, likely leading to misclassification of some sleep epochs as wake (false positives). In addition, it could be caused by inter-subject variability in infant weight, age, sleep pattern (in particular, in this dataset with only short recordings), arousals, subtle body movements such as jerks and scratches, and even autonomic regulation [34,35]. We speculate that some misclassifications were between wake and REM sleep with body movements, and between...
sleep and “motionless wake” without, or with fewer, body movements. The feature values of false negative (FN) and false positive (FP) misclassifications are compared in Figure 5. It can be clearly seen that the mACT values of FN (wake) epochs were close to zero, much lower than those of FP epochs and other wake epochs (see Figure 3). On the other hand, the FP (sleep) epochs, in particular during REM sleep, had very high mACT values. With regard to pSLP, on average, FN epochs had even slightly higher values than FP epochs for both REM and NREM sleep. It would be very difficult to correctly identify these misclassifications using video-based actigraphy only. Therefore, besides physical activity, vital signs such as respiration and heart rate (and heart rate variability), that are able to characterize autonomic nervous activity, are then required for further improving the classification.

Table 2. Wake and sleep classification results (LOOCV) for each infant.

<table>
<thead>
<tr>
<th>Infant</th>
<th>Epoch Number</th>
<th>% Wake Epochs</th>
<th>Accuracy</th>
<th>Cohen's kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>216</td>
<td>17.6%</td>
<td>91.7%</td>
<td>0.664</td>
</tr>
<tr>
<td>02</td>
<td>120</td>
<td>46.7%</td>
<td>90.8%</td>
<td>0.816</td>
</tr>
<tr>
<td>03</td>
<td>118</td>
<td>10.2%</td>
<td>93.2%</td>
<td>0.629</td>
</tr>
<tr>
<td>04</td>
<td>237</td>
<td>30.8%</td>
<td>93.3%</td>
<td>0.837</td>
</tr>
<tr>
<td>05</td>
<td>163</td>
<td>58.3%</td>
<td>91.4%</td>
<td>0.826</td>
</tr>
<tr>
<td>06</td>
<td>93</td>
<td>21.5%</td>
<td>97.9%</td>
<td>0.939</td>
</tr>
<tr>
<td>07</td>
<td>123</td>
<td>5.7%</td>
<td>98.4%</td>
<td>0.866</td>
</tr>
<tr>
<td>08</td>
<td>131</td>
<td>38.2%</td>
<td>91.6%</td>
<td>0.817</td>
</tr>
<tr>
<td>09</td>
<td>106</td>
<td>24.5%</td>
<td>89.6%</td>
<td>0.716</td>
</tr>
<tr>
<td>10</td>
<td>72</td>
<td>16.7%</td>
<td>81.9%</td>
<td>0.220</td>
</tr>
</tbody>
</table>

Figure 5. Boxplots of misclassifications for the two features (a) mACT and (b) pSLP in normalized unit (n.u.) where false negatives (FN) are the number of wake epochs misclassified as sleep epochs (FN = 71) and false positives (FP) are the number of sleep epochs incorrectly classified as wake epochs (FP = 7 for rapid eye movement (REM) sleep, FP = 28 for non-REM (NREM) sleep).

In general, the results of our wake and sleep classification model are better than those reported in literature using accelerometer-based actigraphy [20,30]. This indicates that an IR video camera is a feasible instrument to reliably monitor wake and sleep states of healthy term infants in bed. However, the feasibility for infants with sleep disorders should be validated in the future.

It is important to note that this was a laboratory study, where the monitoring took place in a “controlled” environment with a fixed camera placement and relatively stable lighting conditions for all infants. Additionally, only a small dataset was used including 10 infants with an average of 1.15 h recording per infant. In a real application of infant sleep monitoring at home, we would encounter challenges such as various camera placements with different angles and distances to the infant’s bed, changes in brightness or darkness in the room, video occlusion, and other disturbances (e.g., parental activity). These would likely cause a decline in classification performance. Therefore, a home study in a free-living setting with a larger dataset, including more infants, longer and multiple-day recordings that can capture daily wake-sleep rhythms, should be conducted in the future in order to validate the
method proposed in this paper. Advanced video processing and machine learning algorithms to deal with these challenges merit further investigation.

Furthermore, as for this proof-of-point study, we included only healthy infants. The accuracy regarding sick or preterm infants and those with sleep disorders remains unclear. Further research is needed to validate the methodology for these clinical populations. In addition to validation, future work could focus on the type and quality of movements during in bed periods to help in detection, monitoring, and classification of sleep-related movement disorders, such as restless legs syndrome, periodic limb movement, night terrors/nightmares, head banding, excessive somnolence, insomnia, jet lag syndrome, and sleep-disordered breathing. For movement-related disorders, advanced video processing and pattern/texture analysis methods—rather than general motion detection—would potentially be possible to capture the specific type and quality of disorder-related movements. For infants with sleep-disordered breathing such as sleep apnea, as mentioned in the introduction, cardiac and respiratory information could be detected from videos [24]. Combining video-based cardiorespiratory and movement information for wake and sleep classification warrants further research in sleep apnea subjects.

4. Conclusions

This work experimentally validated the feasibility of automatically classifying wake and sleep states in healthy term infants using an infrared video camera, which is considered to be an unobtrusive (or contactless) method of monitoring and assessing infant sleep. From infrared video frames, video-based actigraphy was estimated, from which two discriminative features were extracted. A Bayesian linear discriminant classification was used. A small set of data was collected in a laboratory study, including video recordings of 11.6 h from 10 infants. Leave-one-subject-out cross validation revealed a substantial agreement with a Cohen’s kappa coefficient of 0.733 ± 0.204 (mean ± standard deviation across infants), compared with polysomnography-based human scoring in wake and sleep classification. Further studies with larger datasets collected in free-living environments at home, and with clinical populations, are necessary in the future.

Author Contributions: Conceptualization, methodology, software, validation, formal analysis, visualization, writing—original draft preparation, X.L.; investigation, resources, data curation, writing—original draft preparation, review and editing, R.O.; software, investigation, visualization, E.v.d.S.; resources, software, writing—review and editing, J.W.; methodology, writing—review and editing, T.T.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank Age van Dalfsen for implementing the recursive search motion estimator, and Elly Zwartkruis-Pelgrim, Sebastiaan den Boer, Jos Gelissen, Shannon van Sambeek, and Rohan Joshi for their insightful inspirations and discussions.

Conflicts of Interest: The authors declare no conflict of interest. At the time of writing, X.L., R.O. and E.v.d.S. are employed by Royal Philips, Research.

References


27. Prechtl, H.F. The behavioural states of the newborn infant (a review). *Brain Res.* **1974**, *76*, 185–212. [CrossRef]


