Detection of sleep position by means of a wrist-worn sensor technology

Ricardo Marques Trindade
ricardo.trindade@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

November 2016

Abstract

Sleep position is a clinically relevant parameter, as it affects the quality of sleep and is associated with several pathologies such as sleep apnea. The goal of this work is to answer the question: Is it possible to develop an algorithm that classifies the sleeping position, solely based on a three-axis accelerometer embedded in a smartwatch? For this we built a sleep movement database. To acquire sleep movements we developed an Android application that allowed to record accelerometer readings from a smartwatch. A wrist accelerometer was worn by patients from Hospital de Santa Maria, and volunteers while they slept. The data obtained from these nights was used to assess the algorithm's performance. The developed algorithm uses classifiers such as Neural Networks, Support Vector Machines, and the Naive Bayes classifier. The core task is to identify the sleep movement performed and from there infer the position. When defining only 3 possible sleep positions (supine, lateral and prone) the average algorithm accuracy was 53.2±26.9% for neural networks, 51±20.9% for support vector machines, 44.5±15.9% for the Naive Bayes classifer and 35.6±21.1% with Hidden Markov models. We conclude that we have paved the way for the development of an algorithm that classifies sleep position using only an accelerometer placed on the wrist, as there is a lot of room for improvement, so we suggest using recurrent neural networks and trying other devices such as the newly announced Android Wear devices in January 2017.
Keywords: Sleep Position, Sleep Apnea, Neural Networks, Accelerometer, Support Vector Machine

1. Introduction

Sleep position during the night can be a clinically relevant parameter as it is associated with several sleep disturbances, even further, analyzing sleep movements helps determining the quality of sleep and sleep patterns [1]. Some of these pathologies are the Gastroesophageal Reflux Disease where it has already been demonstrated that an optimal sleep position for individuals with this disturbance is by lying on their left side [2]. Another disturbance is sleep apnea which is associated with several cardiovascular diseases affecting mostly middle-aged men and has been shown to be aggravated by certain sleep positions [3]. Up until now, no application or device on the wrist can detect sleeping position. By having an application or device that could measure this, would make the life of the patient easier, and keeping up with the advances in technology with the incorporation of sensors like the photoplestimography, accelerometer, and gyroscope in a smartwatch would reduce the burden of cables that patients are subjected to when doing a polysomnography(PSG). The main goal of this work is to answer the question : Is it possible to develop an algorithm that classifies the sleeping position, solely based on a three-axis accelerometer embedded in a smartwatch? To do so, a smartwatch application will be developed to allow extracting data from the watch's sensors, we will also build a sleep movement database with the data retrieved from the sensors.

Figure 1: A common piezoelectric sensor used to measure sleeping position taken from [4].

For the classification of the sleeping position several approaches have already been made. For patients doing a PSG the most common way to measure their sleeping position is by placing a piezoelectric sensor in the area of the chest. This sensor, however, can be uncomfortable to wear, specially if the patient tries to sleep on its stomach. A non-invasive approach to measure sleeping position was made by Zachary Beatle et al. in [5], where they placed load cells under the patient's bed. Each load cell was placed under each bed support, giving a total of 4 cells. The experiment conducted consisted in patients lying in bed during 32 minutes where they were instructed to lie in 4 different positions. Any movement of legs or arms were left to the patient own will. The signal retrieved from the load cells allowed to compute the bed's center of pressure (COP) on the x and y directions. For the classification of the sleeping position, they implemented a K-Means classifier with 4 centroids, using the COP in the y direction as a feature. They obtained an accuracy of 0.68, 0.57, 0.69 and 0.33 for the back,right,left and stomach positions respectively [5]. The problem with this approach is that the COP_y values for the stomach and back position are similar and the classifier fails often to distinguish between them.

Jaehoon et al. in [6] developed a system that used a Kinect sensor to monitor the sleeping position. In this paper, they took advantage of the Kinect's infrared camera, and the Kinect's own human body joints algorithm that detects human body movement and builds a model based on that information. The Kinect was placed above the bed and would stream data to a nearby computer which would later process the data. The joint model outputted by the Kinect had 25 points, and they recorded the x,y and z position of all those points. The sleep movement was calculated based on the Euclidean distance between those points through time.

Shinar et al. [7] introduced the algorithm which detects sleep posture characterized by morphological differences in QRS complex of the Lead I, II, and III electrocardiogram ECG. The limitation of this study is that this method is applicable only when the ECG is clearly measured and the QRS complex is accurately detected.

Despite not being available an algorithm for detecting the sleep position, using a wrist accelerometer, these sensors are being used for sleep detection. In [8] Borazio et. al developed an algorithm that was able to classify whether the patient was sleeping or not based on the data of a single wrist-worn accelerometer. They analyzed the signal with a fixed size sliding window. Their method was based on thresholding the standard deviation of the accelerometer signal.

2. Obstructive Sleep Apnea

We will focus on sleep apnea, as its relation with sleep position is well known.

Obstructive sleep apnea (OSA) is a pathology where individuals present recurrent episodes of partial or complete upper airway obstruction during sleep. Despite inspiratory effort, airflow is decreased (hypopnea) or completely interrupted (apnea).

The respiratory disturbance index (RDI) is used in the medical community to classify the severity of the syndrome. This index has the formula

$$RDI = (RERA + Hypopneas + Apneas) \times 60/TST \tag{1}$$

where TST stands for Total Sleep Time in minutes, and RERA are respiratory effort related arousals. Some of the factors that affect the predisposition of obstructive sleep apnea are presented next.

It was defined by Cartwright and Lloyd [3] that patients whose RDI was at least twice as high in the dorsal decubitus position than in the lateral position were classified as positional apnea patients, whereas the other patients were considered nonpositional patients. Some patients have reduced their total RDI by sleeping on their sides [3, 9].

Alterations in the external anatomy of the head and neck can be risk factors for developing OSA, irrespective of obesity. Craniofacial anomalies can involve a delay in the development of the mandible, producing mandibular retroposition. This causes the tongue to be positioned posteriorly and thus reducing the upper airways. In [10] was found that in normal individuals, the upper airways present a lateral diameter larger than the anteroposterior diameter, whereas the inverse occurs in OSA patients. This corresponds to a narrowing at a critical point, the retropalatal airway. The structures that cause this effect are the lateral pharyngeal walls and the lateral pharyngeal fat pad.

The effect of age on OSA prevalence in the general population still needs further clarification. Studies involving men and women between 20 and 100 years of age have been carried out in order to identify age-specific prevalences. Bixler et al. found that among women, those over 65 years presented the highest prevalence of OSA, whereas the highest prevalence in men was seen among those in the 45 to 64 age bracket [11]. Regarding the severity of OSA, the study by Bixler [12] suggests that it is less severe in older individuals. It is also known that age correlates positively with obesity and neck circumference, and these are both risk factors for sleep apnea.

Regarding gender, the prevalence of sleep apnea is greater in men than in women. In [11] they used clinical criteria and polysomnography and obtained the results of 1.2 % and 3.9 % in men. The prevalence of sleep apnea is higher in menopausal women than in premenopausal women [11]. Hormone replacement therapy for postmenopausal women seems to be associated with a lower prevalence of sleep apnea. A difference between genders is the genioglossus muscle tone, which is believed to be greater in women, suggesting a defence mechanism for the maintenance of the upper airway permeability. Therefore, in the dorsal decubitus position, men present a greater reduction in the dimensions of the upper airway lumen than do women. Martin corroborates this hypothesis by saying that when there is an increase in neck circumference due to fat accumulation, the transverse diameter of the upper airway remains greater in women than in men, suggesting once again that women present a better defence mechanism [13].

Obesity is an important pathogenic factor in sleep apnea. Approximately 70% of OSA patients are obese, and obesity is the only significant risk factor that is reversible [14]. Obesity of type 1 is defined as having a BMI greater than 25kg/m^2. Central or visceral obesity has been more frequently associated with apnea than have other forms of obesity. Some authors analysed the distribution of body fat using magnetic resonance imaging of the neck and abdomen in patients under clinical suspicion of having sleep apnea. The results showed a significant correlation between sleep apnea and the quantity of intra-abdominal fat, as well as a borderline correlation between sleep apnea and subcutaneous abdominal fat and no correlation between sleep apnea and parapharyngeal or subcutaneous fat in the neck region [14]. Oksenberg found, in [15], that losing weight causes a drastic reduction of 91.1% of a patient's RDI. This study also showed that as the body mass index increases, the positional dependency decreases, thus showing how critical the weight of a person is related to the probability of having OSA. The shape and dimension of the upper airways depend on the position of the soft tissue structures (palate, uvula and pharyngeal wall), which can be influenced by gravity. In the dorsal decubitus position, the tongue and soft palate project themselves posteriorly, thereby reducing the area of the oropharynx [10]. A study was conducted in order to evaluate the upper airways area and volume during the dorsal decubitus position, which found no statistically significant difference between genders. However, men presented more significant changes in the dorsal decubitus position than did women [13].

3. Data Acquisition

One of the objectives of the work was to build a sleep movement database. For this we defined 4 possible positions in which the patient can be. This was defined according with the classification of the position done by the PSG system used in Hospital de Santa Maria. This nomenclature is the same found in [1].

- Supine Position, where the person is lying on its back;

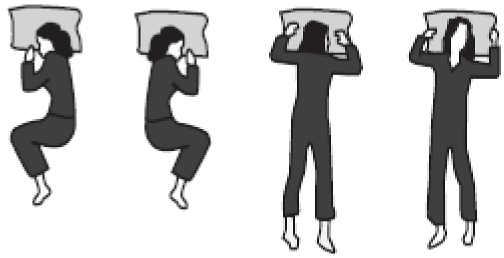

Figure 2: Body positions while lying on bed adapted from [1]. From left to right, we have left lateral position, right lateral position, prone position and supine position.

- Right Lateral Position, where the person is lying on its right side;
- Left Lateral Position, where the person is lying on its left side;
- Prone Position, where the person is lying on its stomach

The approach taken to develop the algorithm was to train a sleep movement classifier, with the assumption of, knowing which movement was performed, the position could be derived from it. We defined 8 possible sleep movements as follows:

- From supine to left lateral position;
- From supine to right lateral position;
- From left lateral to supine position;
- From right lateral to supine position;
- From left lateral to prone position;
- From right lateral to prone position;
- From prone to left lateral position;
- From prone to right lateral position

There can be other movements considered like left lateral to right lateral position but these movements are a composition of the ones stated above. A classifier needs a great amount of data to be trained, so we acquired several samples of each movement. For the recording of these movements, we developed an Android Wear application that allowed to record the values from a Sony Smartwatch 3 accelerometer, gyroscope, magnetometer and rotation vector (also known as orientation sensor) and store them in comma-separated values (CSV) files on the watch. The application usage required two people, one for controlling which movements to perform and another,a volunteer that would put on the watch and perform the desired movements. The person controlling the phone would select the movement to execute within a list. For each volunteer, all movements would be performed several times in no particular order and never doing the same movement twice in a row to allow the data to be as diverse as possible. The watch was always placed on the left hand. After recording the movements all the CSV files would be extracted from the watch by pluging it to a computer. The acquisition was performed in $n = 20$ different individuals, with ages between 18-51 years.

We also recorded volunteer's sleeping at night, while wearing a Sony Smartwatch 3, with a Kinect V2 from Microsoft. The Kinect was used with the purpose of confirming the person's position. The sampling frequency set on the watch was 32Hz, as is the one used in the Empatica E4 wristband.

The camera would be set next to the volunteer's bed and just before the volunteer went to sleep he/she had to run a computer script that prompted for his/her name and after that, the Kinect would start recording. The watch would start recording as well after the volunteer pressed the button on the smartphone application. If the watch started recording successfully a warning would display on the smartphone screen. The same applies if the user was able to stop the recording. As before, the watch was placed on the left hand.

The video files from each night occupied a great amount of space in the computer (around 90Gb/hour) so the script would only allow to record for 8 hours and then shut down the computer. This was done to prevent volunteers from forgetting to stop Kinect's recording and the files from growing into astronomical sizes. During the realization of this work, patients from the Neurology department in Hospital de Santa Maria were asked to take part in this work. These patients went to the Neurology consultation to realize a PSG due to suspicions of sleep disorders as sleep apnea. This procedure started by them entering the Neurology consultation to be prepared for the PSG. If the patient would agree to participate in this work, a consent would be signed and then the device was placed on the left hand. The PSG started recording by pressing a button on the device attached to the chest, and in order to synchronize the recording with the device, the pressing of the PSG button occurred at the same time as the button pressing on the smartphone to trigger the recording in the device. After this the patient would go home and sleep the whole night with the device recording, or stay in the hospital in case the PSG required a video recording as well. The total time from start until stop of the PSG would take around 13-15 hours and the Sony Smartwatch 3 battery did not last that time while recording, so we used an Empatica E4 wristband paired with an iPhone. The disadvantage of this wristband is that it only has an accelerometer when compared to the accelerometer, gyroscope, magnetometer and rotation vector from the Smartwatch. The accelerometer from the wristband has a fixed sampling rate of 32Hz unlike the smartwatch which can be set manually by code.

Movement Type	# Examples
Supine - Left Lateral	64
Supine - Right Lateral	79
Left Lateral - Supine	69
Right Lateral - Supine	64
Left Lateral - Prone	41
Right Lateral - Prone	48
Prone - Left Lateral	45
Prone - Right Lateral	50

Table 1: Number of examples for each movement, of a total of 460 different examples.

When the patient returned to the hospital or woke up the recording would stop(in case the patient slept in the Hospital).

The Empatica E4 streams all data through Bluetooth to the iPhone. For each recording the files written would have a different id associated to each patient (i.e 0001 acceleration.csv, 0002. acceleration.csv). The same id would be present on the PSG signal. The PSG would be later analyzed by medical doctors at Hospital de Santa Maria. The position signal from PSG is an integer-valued signal where each position corresponds to a single integer and has the following notation:

- 2, For upright position, when the patient is standing or sitting;
- 3, For left lateral position;
- 4, For right lateral position;
- 6, For supine position;
- 8, For prone position.

As was done later on, only 3 classes were considered. These classes are an aggregation of the 8 movements mentioned before. The aggregation was done in a way that classes ending in supine position belong to the first class (Supine Class), movements ending in a lateral position belong to the second class (Lateral Class), and finally movements ending in a prone position belong to the third class (Prone Class). This means merging the classes Supine-Left Lateral, Supine-Right Lateral, Prone-Left Lateral, Prone Right Lateral into one class, Left Lateral - Supine, Right Lateral Supine into another and finally Left Lateral - Prone and Right Lateral - Prone into another. The results are in table 3.

Position	# Examples
Supine	133
Lateral	238
Prone	89

Table 2: Number of examples for each movement,considering the position on which they ended,using 3 classes, of a total of 460 different examples.

Position	# Examples
Supine	133
Non Supine	327

Table 3: Number of examples for each movement,considering the position on which they ended,using 2 classes, of a total of 460 different examples.

We also merged the Lateral and Prone classes into a single one, in order to build a binary classification problem (Supine vs Non Supine). The distribution of movements is in 3.

4. Algorithm Implementation

The algorithm's core task is to classify which movement was performed and from there infer the position. The analysis of the accelerometer signal was made with a sliding, non-overlapping time window of 6 seconds. Before identifying the movement, first it had to be detected, where we developed a transition detector. This detector analyzed the accelerometer signal and detected if the person had indeed changed position. Should this detection be positive, then we would present that section of the signal to the classifier. The transition detector, was implemented based on the assumption that changes in sleep position reflect as fast variations on the accelerometer signal. To detect these changes we will resort to the derivative of the accelerometer signal. The derivative of the accelerometer signal will be represented by S. The method used to detect these movements used two signal features that were as follows:

$$\overline{m} = S^T S \tag{2}$$

Here, $\overline{m}$ can be seen as the absolute value of the derivative of the accelerometer signal. For sleeping movement detection this value had to be between were 3 and 10. These values were computed for the whole sleeping movement dataset and by plotting them in MATLAB the empirical values were established.

For the development of the algorithm we used Artificial Neural Networks, Support Vector Machines, Naive Bayes Classifiers, and Hidden Markov Model. The implementation of these models was made with MATLAB's toolboxes [16, 17, 18] respectively. The training set for the classifiers consisted in feature vectors computed for each example from the movement database. Representing the accelerometer signal by a, as a matrix, where each row is the acceleration on the x,y and directions respectively, the initially selected features are listed below.

- Mean of the acceleration signal on each direction at the start of the transition;
- Mean of the acceleration signal on each direction at the end of the transition;
- Acceleration standard deviation on each direction;
- Correlation between xy, yz and xz signals;
- Correlation between xy, yz and xz signals from the gyroscope;
- Median on each direction;
- Interquartile range on each direction;
- Number of samples above zero (positive acceleration) on each direction;
- Number of samples below zero (negative acceleration) on each direction;
- Determinant of $a^T a$.

However this was a large number of features (25), for a model with such a small dataset to be trained on. After comparing the p-values obtained from the pairwise comparison between features for each class we decided to use only 5 that are listed below. The selection was based on the smallest p-values among different classes, with a significance level of 5%.

- Mean of the acceleration signal and the end of the movement, on the x and y directions;
- Correlation between the acceleration signal on directions x and y;
- Correlation between the acceleration signal on directions x and z;
- Number of samples on the y direction with acceleration > 0.

For the implementation of neural networks, we used all three possible algorithm offered by MATLAB's toolbox [16], Scaled Conjugate Gradient Descent, Levenberg-Marquardt and Bayesian Regularization. Gradient descent is the most commonly used algorithm for training neural networks, as it is easy to implement and a first order method that under certain conditions converges fast [19]. Levenberg-Marquardt algorithm is a second order method, that relies on the inversion of the Hessian matrix for updating the network's weights. This method outperforms the gradient descent as long as the number of parameters is not too large ($>$ 100000). Finally the Bayesian regularization is another second order method that excels in preventing overfit [20] by computing the effective number of network parameters and discarding all the irrelevant ones. As for the Support Vector Machines we used a Gaussian Kernel or Radial Basis Function, and the for Naive Bayes classifier we assumed a normal distribution of our features.

MATLAB's Neural Network toolbox is very flexible as it allows to control several parameters of the network. For all the tests we made, we designed a simple 3 layer (input-hidden-output) network, while varying the number of hidden units and training function. The hyperbolic tangent was used as activation function for all units. First we removed 20% of the dataset, that would be used for testing, and from the remaining dataset we split it on a 80/20 proportion. 80% was used to train the network, while the remaining 20 was used for validation.The partition of the dataset was done randomly. The use of a validation set is an important method, as it allows to stop training the network early, moreover it prevents the network from overfitting. To do this, one needs to keep track of both training set and validation set error. With each iteration and adjustment of the network weights, the training set error decreases. This should happen as well for the validation set, if it does not, then the training process is reset as continuing to train, will overfit the parameters to the training set. The toolbox has a parameter that stops the training of the network after the reset has occurred a specific number of times - which can be set manually - in the current case it was set to 20. The validation set should be completely independent and non-related with the test set, as the objective of training the network is to generalize it as better as possible. By generalizing we mean that if the network is trained with a set S, then it should classify correctly any other training example that is completely independent from S. After training we would test the network, using the first 20% that were removed from the dataset. This was done to ensure the test set was never seen by the network during the training process. The network training was done several times, around 100 iterations for each configuration (with the same network structure). For each iteration of the training the process, the dataset was split differently. After the iterations we selected the network with better performance in the test set. The same process applies for the Support Vector Machines and Naive Bayes Classifier.

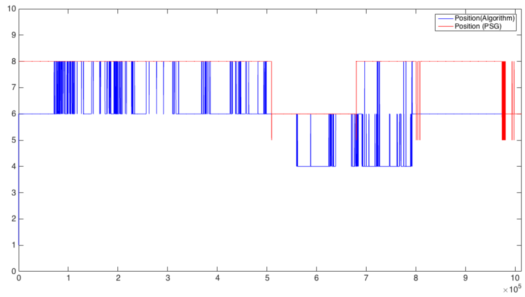

Figure 3: Output of the algorithm output when no transition detector is used

The classification using a Markov follows a different procedure. In this case the computation of a feature vector is not done, as the whole accelerometer signal is analyzed using Viterbi' algorithm. The first step is training a Markov Model using MATLAB's own function *hmmtrain*, that uses the Baum-Welch algorithm [21]. Each signal had to be normalized between 1 and 10, and rounded to the closest integer, as MATLAB's toolbox only works with integer sequences, we also had to pass as input the corresponding position signal. This was done for each of directions of the accelerometer signal. After training we would obtain the transition and emission $(t_x, t_y, t_z, e_x, e_y, e_z)$ matrices that define our model. For the analysis of the accelerometer signal, we used MATLAB's own function *hmmviterbi* that receives as input a model and a vector, and returns the most probable sequence of states, using Viterbi's algorithm.

5. Results

This chapter covers the results obtained from the algorithm on the test set (night recordings from Hospital de Santa Maria and Kinect video recordings) used to assess its performance. The performance measure used was the amount of time, in percentage, that the algorithm's position output matched the real position signal. The values from the magnetometer were not used as they can be biased if the patient is near any electronic device, which we are not able to control. The values from the rotation vector were not used as well because its sampling frequency is not constant during the whole recording. Our dataset had a total of n = 460 movements, distributed over the 8 different movement classes. The full distribution of movements can be consulted in table 3.

Tables 4,5,6 and 7 report the average accuracy of the algorithm, when tested with the recordings from Hospital de Santa Maria ($n = 8$) and recordings using the Kinect ($n = 4$). As we have not found any other work that uses an accelerometer on the wrist to detect the sleeping position, no comparison with the literature can be done at this time. When train-

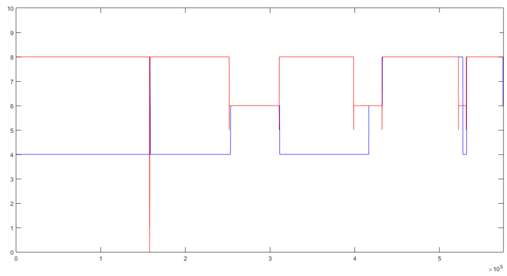

Figure 4: Example of the algorithm output versus the original position signal

#Classes	Accuracy
2	53.4± 29.9%
3	53.2± 26.9%
8	46.3± 18.8%

Table 4: Average algorithm accuracy using neural networks, simple Input-Hidden-Output architecture, with 5 neurons on the hidden layer

#Classes	Accuracy
2	66.3± 27.5%
3	48.2± 23 %
8	43.5± 22.5%

Table 5: Average algorithm accuracy using neural networks and no transition detector, simple Input-Hidden-Output architecture, with 5 neurons on the hidden layer

#Classes	Accuracy
2	29.6± 25.0%
3	51.0± 20.9%
8	39.7± 23.7%

Table 6: Average algorithm accuracy using support vector machines with a Radial Basis Function kernel

#Classes	Accuracy
2	28.4±17.0%
3	44.5±15.9%
8	51.5±19.9%

Table 7: Average algorithm accuracy using naive Bayes classifier assuming normal distribution of features.

ing the network and testing it, we noticed that its worst performance was with movements that ended in a prone position. When recording the movements we noticed that this kind of movements were the most complex, that involved many complex arm movements, whereas movements that ended in a lateral or supine position were way more simple. Therefore, capturing features that can describe this movements is a tough problem. Perhaps the features we are using for the describing movements are not the best ones. A possible improvement could be the use of the rotation vector, present on the Smartwatch, based on the idea that it can capture the rotation movement performed when the person lies on its stomach.

In this work we merged the classes of movements ending in lateral position and prone position into a single class. This was meant as a workaround to improve the classification for movements ending in a prone position. These results show a significant improvement ($p \approx 0$) when compared with the three classes scenario, however having only two classes once more leads to loss of information.

Often the transition detector detected some transitions where there was no change in position. One of the reasons was the presence of spikes on the accelerometer signal. These spikes often coincide with episodes of periodic limb movements. The spasms show up on the accelerometer signal misleading the algorithm, and detecting a transition where there is none. After this event, the algorithm fails to detect the actual position and stays on an incorrect one until the next transition occurs. As a solution, we propose improving the transition detector, by including a spike detector that denies the detection of the transition if it finds an isolated spike on the accelerometer signal. A possible solution would be for the algorithm to prevent itself from evaluating periods of great movement, as happens when the patient gets up or sits in bed. This may require the inclusion of another class on the classifier that represents the patient being standing or sitting. As we already mentioned before, the inclusion of the rotation vector sensor may prove useful in this case, as it may capture the orientation of the device which may be unique for a standing or sitting position. Also here could be useful including the frequency related features used in [22].

Regarding the Bayes classifier it is important to remark that the it is based on the strong assumption that all features are independent [23]. This is certainly not true in our case, and can explain why its classification accuracy is so low in some cases. In what concerns SVMs, the use of a Gaussian kernel may have lead to overfitting the data to the training set, resulting in poor results on the sleep recordings set. A possible solution would be to reduce the number of features used or even try another kernel function like the polynomial or hyperbolic tangent.

The average results for the Hidden Markov Models were of 35.6 ± 21.1 %. This accuracy is lower when compared to the other classifiers. This may have been due to bad combination of weights on each accelerometer signal (we are using 0.33 for each direction). A possible solution would be study around the values that will output better results. Also the normalization procedure, of setting the acceleration values as integers between 1 and 10, leads to a great loss of information on the content of the acceleration signal. This is a disadvantage of MATLAB's toolbox [18]. In this case a possible solution would be to implement the Markov Model algorithm that do not require the emitted symbols to be represented by integers.

6. Conclusions

In this work we implemented a sleep position classifier using different classifiers namely Artificial Neural Networks, Support Vector Machines, Naive Bayes Classifiers and Markov Models. The novelty of this work is that it classifies the position according to data from an accelerometer placed on the wrist, which to our knowledge has not been done before. Two Android applications were developed, one for recording the sleep of volunteers and patients during the night, and the other to record sleep movements in volunteers. The latter allowed to build a sleep movement database with over four hundred different examples. We enfatize the need to increase the size of this database to broaden more variations of the same movement. The Android application that was meant to be used during recordings in Hospital de Santa Maria, was not used as patients would arrive at the hospital around 6pm and return the next day at 9am. The device's battery life does not last more than eight hours recording so we opted for using the Empatica E4 wristband instead. Smartwatches are still in an early stage of development and were not designed for recording sensor values during such a long period of time.

The process of developing the application took a long time until a production version (by production we mean, ready to use by the technicians at Hospital de Santa Maria) was ready. Until we realized that Doze mode [24] was taking place and lowering the sampling frequency from the sensors, several approaches were tried in order to overcome this problem. Fortunately an update from Google arrived in April which allowed to use the PowerManager to overcome this hindrance. Unfortunately, however this was not enough as the device never got to be used by the technicians in Hospital de Santa Maria more than once, due to battery considerations as was already mentioned.

During the development of the work we only mentioned smartwatches running on Android, however Apple has its own device as well, the iWatch. This device was not used for the acquisitions due to software constraints. The watch's operating system stops all the running applications once its screen goes off, which is unfeasible for us as the application needs to keep recording during the entire night and not only when the device's screen is on. We also felt more comfortable at developing in Java for Android, instead of Objective-C for Apple.

We were not able to compare our results with the literature, as to our knowledge, this is a pioneer work. Other works have developed classifiers for sleep position, using accelerometers, placed on the head, hip or chest, but none of them used a device on the wrist. Although, we believe that our method will never be able to surpass the conventional piezoelectric sensor attached to the chest.

A major difficulty of this work was handling all the large video files recorded by the Kinect, which slowed down the acquisition process of data. Despite having a high video quality the Kinect was not designed to record several hours of video, and we were not aware of the necessity of recording the depth camera stream in order to record the infrared one.

We suggest using another toolbox for Hidden Markov Models, as the one used in this work only allows using integer-valued sequences, and this may lead to great loss information on the content of the signals. Regarding SVMs, using another kernel, instead of the Gaussian one, may prove useful and give in to better results. As for the Naive Bayes classifier, the suggestion would be to use completely independent features from the sensor signals, as to avoid violating the strong assumption on which it lies on.

We were not able to use the rotation vector sensor as its sampling frequency is not constant and is dependent on the workload that the device is subjected to. The magnetometer was not used as well, as it output is influenced by the presence of electromagnetic devices.

Acknowledgements

The author would like to thank

To Duarte Mendes de Almeida for giving me the opportunity to develop this work, his constant help, suggestions, and for all he taught me that proved essential to this work.

To Professor Miguel Tavares da Silva for his constant motivation and readiness to discuss and criticize my ideas.

To Professor Carla Bentes, for her help in the medical component of this work and availability in conducting the study in Hospital de Santa Maria.

To the neurology department technicians Rosa and Ligia for always keeping me up to date with the issues with the devices being used.

References

[1] Enamul Hoque, Robert F. Dickerson, and John A. Stankovic. Monitoring body positions and movements during sleep using wisps. In *Wireless Health 2010*, WH '10, pages 44–53, New York, NY, USA, 2010. ACM.

[2] L Khoury, Ramez M Camacho-Lobato, O Katz, Philip, Muhammad A. Mohiuddin, and Donald O. Castell. Influence of spontaneous sleep positions on nighttime recumbent reflux in patients with gastroesophageal reflux disease. *Am J Gastroenterol*, 1999.

[3] RD Cartwright. Effect of sleep position on sleep apnea severity. *Sleep*, 1984.

[4] e-health sensor platform v2.0 for arduino and raspberry pi [biometric / medical applications]. `https://www.cooking-hacks.com/documentation/tutorials/`, 2016. Accessed : 2015-12-03.

[5] Z. T. Beattie, C. C. Hagen, and T. L. Hayes. Classification of lying position using load cells under the bed. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 474–477, Aug 2011.

[6] Jaehoon Lee, Min Hong, and Sungyong Ryu. Sleep monitoring system using kinect sensor. *Int. J. Distrib. Sen. Netw.*, 2015:205:205–205:205, January 2016.

[7] Z. Shinar, A. Baharav, and S. Akselrod. Detection of different recumbent body positions from the electrocardiogram. *Medical and Biological Engineering and Computing*, 41(2):206–210, 2003.

[8] M. Borazio, E. Berlin, N. Kücükyildiz, P. Scholl, and K. V. Laerhoven. Towards benchmarked sleep detection with wrist-worn sensing units. In *Healthcare Informatics (ICHI), 2014 IEEE International Conference on*, pages 125–134, Sept 2014.

[9] Neil B. Kavey, Andrew Blitzer, Stephen Gidro-Frank, and Kathryn Korstanje. Sleeping position and sleep apnea syndrome. *American Journal of Otolaryngology*, 6(5):373 – 377, 1985.

[10] I Ayappa and David M Rapoport. The upper airway in sleep: physiology of the pharynx. *Sleep Medicine Reviews*, 7(1):9–33, 2002.

[11] Edward O. Bixler, Alexandros N. Vgontzas, L. Hung-Mo, T. Ten Have, J. Rein, and A. Kales. Prevalence of sleep-disordered breathing in women. *American Journal of Respiratory and Critical Care Medicine*, 163(3):608–613, 2001.

[12] Edward O Bixler, AN Vgontzas, T Ten Have, K Tyson, and A Kales. Effects of age on sleep apnea in men. *Am J Respir Crit Car Med*, 1998.

[13] SE Martin, R Mathur, I Marshall, and NJ Douglas. The effect of age, sex, obesity and posture on upper airway size. *European Respiratory Journal*, 10(9):2087–2090, 1997.

[14] H. Schäfer, D. Pauleit, Thomas Sudhop, I. Gouni-Berthold, S. Ewig, and HK Berthold. Body fat distribution, serum leptin, and cardiovascular risk factors in men with obstructive sleep apnea. *Chest Journal*, 122(3):829–839, 2002.

[15] Arie Oksenberg, Elena Arons, Henryk Radwan, and Donald S. Silverberg. Positional vs nonpositional obstructive sleep apnea patients: Anthropomorphic, nocturnal polysomnographic and multiple sleep latency test data. *Chest*, 112(3):629 – 639, 1997.

[16] Matlab neural network toolbox. `https://www.mathworks.com/products/neural-network/`, 2016. Accessed : 2016-06-20.

[17] Matlab statistics and machine learning toolbox. `https://www.mathworks.com/help/stats/fitcecoc.html`, 2016. Accessed : 2016-09-20.

[18] Matlab hidden markov model toolbox. `https://www.mathworks.com/help/stats/hidden-markov-models-hmm.html`, 2016. Accessed : 2016-06-20.

[19] Emile Fiesler and Russell Beale, editors. *Handbook of Neural Computation*. Oxford University Press, 1997.

[20] Burden F. Winkler D. *Bayesian Regularization of Neural Networks*, pages 23–42. Humana Press, 2009.

[21] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, Feb 1989.

[22] A Mannini, S Intille, M Rosenberger, A Sabatini, and Ws Haskell. Activity recognition using a single accelerometer placed at the wrist or ankle. *Med Sci Sports Exerc*, 2014.

[23] L. B. Almeida. Statistical aspects of supervised learning, 2015.

[24] Powermanager. `https://developer.android.com/reference/android/os/PowerManager.html`, 2016. Accessed : 2016-15-03.