Automatic detection of sleep apnea using a hidden Markov model and nonlinear analysis of nocturnal oximetry

By
Fatma Z. Abdel-  F. E. Z. Abou Chadi**  Hosam M. Salah ***  Shahira F. Loza**** Mageed*

Abstract:
The aim of this work is to develop an automatic system that can be used as an assistant tool for the detection and diagnosis of different kinds of sleep Apnea (Obstructive, Hypopnea and Central Apnea, respectively). Three nonlinear techniques were used for feature extraction: Central tendency measures (CTM), Lempel-Ziv complexity (LZC) and Approximate Entropy (ApEn) for oxygen saturation signals (SaO2). A statistical Comparison using (t – test) was performed for comparing the population mean of normal group with each of the Sleep Apnea groups for the nonlinear parameters. Three Hidden Markov Models (HMMs), based on Baum–Welch algorithm were proposed to estimate the optimal number of the parameters. The results have showed that the use of HMM and the nonlinear features gave promising results used for classifying Sleep Apnea diseases.
Introduction:

Sleep Apnea [1], is a very common sleep disorder with an estimated prevalence from 1 to 10% in the adult and 11% in children causing: irritability, heart disease, high blood pressure and other physiological dysfunction [2]. It is defined as a recurrent cessation of airflow for 10 seconds or longer more than five times per sleep hour. Health studies affirm that more than 30 of these non breathing episodes per night should be considered abnormal [16].

There exist three kinds of Sleep Apnea: Obstructive sleep apnea (OSA), Hypopnea and Central Apnea [2]. OSA is defined as complete cessation of airflow in the nose and mouth associated with reduction of oxygen levels in arterial blood. The term Hypopnea is used when the breath does not stop but decrease over 50% of its normal value followed by 4% desaturation of haemoglobin level. Central Apnea is defined as partial blockage of airflow associated with lack of respiratory efforts due to the failure of brain to send appropriate signals to the respiratory muscles.

Pulse oximetry is a well-established tool routinely used in many settings of modern medicine to determine a patient's arterial oxygen saturation and heart rate using dual wave length finger probe to give sufficient amount of information about a person's respiratory patterns.

Recently, Nocturnal oximetry arises as an alternative to polysomnography (PSG) since it is readily available, relatively inexpensive and can be performed at home [4,15]. It allows monitoring arterial oxygen saturation (SaO2) during sleep which is considered as a powerful tool for Sleep Apnea detection.

Previous studies showed that: several oximetric indexes including number of oxyhemoglobin desaturations below a certain threshold, and the cumulative time spent below an oxygen saturation (SaO2) of 90 % have been suffered from several limitations [15].

Also, spectral analysis of nocturnal oxygen saturation (SaO2) detecting the presence of a periodogram peak within certain period couldn't be sufficient to give an accurate diagnosis results [13,14].

Nonlinear analysis of blood oxygen saturation (SaO2) has demonstrated to provide useful information that help in Sleep Apnea diagnosis; improving the diagnostic accuracy of classical oximetric indexes and spectral characteristics [6]. Central tendency measure (CTM), Lempel-Ziv complexity (LZC) and approximate entropy (ApEn) were applied to quantify variability, complexity and regularity of SaO2 recordings, respectively [9].

HMMs have proved to be a powerful and flexible class of statistical model for describing different kinds of sequence data and representing individual component states of a dynamic system [3]. The utility of hidden Markov models stems from their ability to offer an effective balance between the twin data modeling issues of complexity and tractability. The trade-off between descriptive modeling power and
practical ease-of-use is perhaps the main reason for the success of hidden Markov models in practice [5].

In this work, we investigate the ability of Hidden Markov Models (HMMs) using nonlinear analysis of SaO2 recordings to discriminate between different types of Sleep Apnea.

2. Subjects:
A total of 128 subjects (32 for Normal, 32 for OSA, 40 for Hypopnea and 24 for Central Apnea, respectively) were collected from Cairo Center For Sleep Disorder (Egypt). The Normal records were extracted from the Apneic records after treatment using Continuous Positive Airway Pressure (CPAP) giving more classification accuracy [2]. CPAP is the most widely recommended treatment to sleep apnea diseases, entails wearing a mask-like device that provides constant air pressure to prevent the airway from collapsing regardless of whether patient is breathing in or out.

SaO2 were recorded using a dual wave length-based finger probe with a sampling frequency of 1 Hz (one sample every second) from midnight to 8:00 AM. An example of the overnight oximetric recordings for Normal, OSA, Hypopnea and Central Apnea patients is shown in Fig. 1.

3. Methods
Three nonlinear analysis methods were applied to SaO2 signals: Central tendency measure (CTM), Lempel-Ziv complexity (LZC) and approximate entropy (ApEn).

3A. Central Tendency Measure
Central tendency measure (CTM) is used to quantify the signal variability using a second-order difference plot. These kinds of scatter diagrams, given by (1), are graphs centered in the origin and used to assess the degree of chaos in a data set, assigning larger values to lower variability [6].

\[ [x(n+2)-x(n+1)] \text{ versus } [x(n+1)-x(n)]. \]

CTM is computed by selecting a circle of radius \( r \) around the origin, counting the number of points that fall within the radius, and dividing by the total number of points which is \( N-2 \) using scatter plot. Then, the CTM can be computed as:

\[
\text{CTM} = \frac{\sum_{i=1}^{n-2} \delta(d_i)}{N - 2},
\]

where

\[
\delta(d_i) = \begin{cases} 
1 & \text{if } [(X(i + 2) - X(i + 1))^2 + (X(i + 1) - X(i))^2]^{1/2} < \rho. \\
0 & \text{otherwise}
\end{cases}
\]

In the present study, CTM is computed with various radii \( \rho = 1, 5, 8 \) using 512 segments and averaging over the overnight SaO2 records which is about 8 hours (28800 samples) length for each subject [6]. Fig. 2 illustrates the second-order difference plots for the SaO2 signals depicted in Fig. 1.

**Figure (2):** Second-order difference plots for (a) a common Normal subject, (b) an OSA patient, (c) Hypopneic patient and (d) a Central Apnea patient.
3.B. Lempel-Ziv complexity

Lempel-Ziv complexity (LZC) is a nonparametric measure of complexity based on a coarse-graining of the Measurements [18]. First, signal must be transformed into a finite (0–1) binary sequence, \( P = s(1), s(2), ..., s(n) \), by comparison with a median threshold \( T_d \), then \( s(i) \) is defined as [17]:

\[
s(i) = \begin{cases} 
0 & \text{if } X(i) < T_d, \text{ for } i = 1, 2, ..., N \\
1 & \text{if } X(i) \geq T_d
\end{cases}
\]

(4)

To compute LZ complexity, the sequence \( P \) has to be scanned from left to right and a complexity counter \( c(n) \) is increased by one unit every time a new subsequence of consecutive characters is encountered. In the case of a (0–1) sequence, \( c(n) \) can be normalized as follows [10]:

\[
\text{LZC} = \frac{c(n)}{b(n)},
\]

(5)

where

\[
b(n) = n / \log_2(n).
\]

(6)

LZC was computed using symbolic sequences of 50 samples length and averaging for the overnight SaO2 records. A typical example of the results of the LZ complexity evaluated from SaO2 records is shown in Fig. 3.

Figure (3): LZ complexity evaluated on SaO2 for (a) a common Normal subject, (b) an OSA patient, (c) Hypopneic patient and (d) a Central Apnea patient.
Approximate Entropy (ApEn) quantifies the regularity in time series by measuring the logarithmic likelihood that runs of patterns that are close (within r) for m contiguous observations remain close (within the same tolerance r) on subsequent incremental comparisons [9], with larger values corresponding to more irregular data. Pincus [8], suggested parameter values of m = 1 and r a fixed value about 0.25 times the standard deviation (S.D.) of the original time series producing good statistical reproducibility.

Thus, in the present study, we computed ApEn with m = 1 and r equal to 0.15, 0.25 and 0.50 times the S.D. of the SaO2 signals, dividing the total night records into 50 samples segments. Table I summarizes the mean ± S.D values for each nonlinear feature derived from Central tendency measure (CTM), Lempel-Ziv complexity (LZC) and Approximate entropy (ApEn) analysis. Normal group shows, in average, large r CTM (r = 1), lower LZC and lower ApEn (r = 0.15, 0.25 and 0.50 times the S.D. of SaO2 signals). On the other hand, Apneic groups (OSA, Hypopnea and Central Apnea) show larger CTM (ρ = 5 and 8), higher LZC and higher ApEn (r = 0.15, 0.25 and 0.50 times the S.D. of SaO2 signals).

<table>
<thead>
<tr>
<th>Feature</th>
<th>Normal</th>
<th>OSA</th>
<th>Hypopnea</th>
<th>Central Apnea</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X ± SD</td>
<td>X ± SD</td>
<td>X ± SD</td>
<td>X ± SD</td>
</tr>
<tr>
<td>Ctm1</td>
<td>0.941 ± 0.104</td>
<td>0.600 ± 0.148</td>
<td>0.266 ± 0.305</td>
<td>0.515 ± 0.169</td>
</tr>
<tr>
<td>Ctm5</td>
<td>0.998 ± 0.014</td>
<td>0.922 ± 0.012</td>
<td>0.584 ± 0.005</td>
<td>0.985 ± 0.036</td>
</tr>
<tr>
<td>Ctm8</td>
<td>0.998 ± 0.001</td>
<td>0.984 ± 0.058</td>
<td>0.806 ± 0.010</td>
<td>0.990 ± 0.037</td>
</tr>
<tr>
<td>LZ</td>
<td>0.216 ± 0.458</td>
<td>0.415 ± 0.644</td>
<td>0.582 ± 0.113</td>
<td>0.194 ± 0.739</td>
</tr>
<tr>
<td>Appr.15</td>
<td>0.459 ± 0.034</td>
<td>1.168 ± 0.040</td>
<td>1.995 ± 0.045</td>
<td>0.097 ± 0.023</td>
</tr>
<tr>
<td>Appr.25</td>
<td>0.270 ± 0.013</td>
<td>0.581 ± 0.030</td>
<td>0.609 ± 0.025</td>
<td>0.035 ± 0.016</td>
</tr>
<tr>
<td>Appr. 5</td>
<td>0.141 ± 0.007</td>
<td>0.452 ± 0.002</td>
<td>0.573 ± 0.021</td>
<td>0.019 ± 0.010</td>
</tr>
</tbody>
</table>

### 4. Statistical analysis

The t-test was used to determine whether there are significant differences in the mean values of the nonlinear parameters between the normal group and each group of Apnea subjects (OSA, Hypopnea and Central Apnea, respectively) at 5% level of significance. For the inter-subject average values \( \bar{X}_A \) and \( \bar{X}_N \) of the parameters concerned, and the corresponding standard deviations SDA and SDN, where NN and NA denote the number of subjects of the normal group and abnormal group, respectively [7]. The t-test statistics is calculated as:
Proceedings of the 8th ICEENG Conference, 29-31 May, 2012

\[ T = \frac{\bar{X}_N - \bar{X}_A}{S\sqrt{\frac{1}{N} + \frac{1}{N_A}}} . \]  \hspace{1cm} (7)

where,

\[ S = \sqrt{\frac{(N_N - 1)SD^2_N + (N_A - 1)SD^2_A}{N_N + N_A - 1}} . \] \hspace{1cm} (8)

The degree of freedom is

\[ df = N_N + N_A - 1. \] \hspace{1cm} (9)

Table II shows the results of the t-test for four different groups of cases: Normal, OSA, Hypopnea and Central Apnea, using nonlinear analysis.

**Table (2): Statistical Differences Between Different Groups Using T-Test**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Between Normal and OSA groups</th>
<th>Between Normal and Hypopnea groups</th>
<th>Between Normal and Central Apnea groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t-value</td>
<td>Probability (p)</td>
<td>t-value</td>
</tr>
<tr>
<td>Ctm1</td>
<td>1.223</td>
<td>0.001</td>
<td>1.563</td>
</tr>
<tr>
<td>Ctm5</td>
<td>3.564</td>
<td>2.774e-004</td>
<td>1.178</td>
</tr>
<tr>
<td>Ctm8</td>
<td>0.946</td>
<td>0.052</td>
<td>3.984</td>
</tr>
<tr>
<td>LZ</td>
<td>1.377</td>
<td>0.016</td>
<td>5.653</td>
</tr>
<tr>
<td>Appr.15</td>
<td>3.908</td>
<td>1.324e-003</td>
<td>1.254</td>
</tr>
<tr>
<td>Appr.25</td>
<td>0.979</td>
<td>0.014</td>
<td>2.701</td>
</tr>
<tr>
<td>Appr.5</td>
<td>0.466</td>
<td>0.044</td>
<td>0.799</td>
</tr>
</tbody>
</table>

5. **Hidden Markov Models**

A hidden Markov model (HMM) is a tool to statistically model time-variant process with the following characteristics[20]:

1- Set of unobserved (hidden) states \(Q=\{q_1,\ldots,q_N\}\), where \(N\) is the number of hidden states in the model.
2- Set of observation symbols \(O=\{o_1,\ldots,o_L\}\) where \(L\) is the number of distinct emission symbol per state.
3- The transition matrix \(A\) whose elements \(a_{ij}\) represent the probability to go from state \(q_i\) to state \(q_j\)
4- The emission matrix B whose elements $b_{jk}$ represent the probability of emission of a symbol $o_k$ when the system state is $q_j$.

5- The set of initial state probability distributions $\Pi = \{\pi_1, \ldots, \pi_N\}$ whose elements $\pi_i$ represent the probability for $q_i$ to be the initial state. For convenience we denote HMM as compact notation notation $\lambda = \{A, B, \Pi\}$.

Under an HMM, there are two conditional independence assumptions made about these random variables that make associated algorithms tractable [19]:

1. The $t$th hidden variable, given the $(t-1)$st hidden variable, is independent of previous variables, or: $P(Q_t | Qt-1, Ot-1, \ldots, Q1, O1) = P(Qt | Qt-1)$.

2. The $t$th observation depends only on the $t$th state.

$P(Ot | Qt, Ot, \ldots, Q1, O1) = P(Ot | Qt)$.

The learning task in HMMs is to find, given an output sequence or a set of such sequences, the best set of state transition and output probabilities. The task is usually to derive the maximum likelihood estimate of the parameters of the HMM given the set of output sequences using Baum–Welch algorithm [11]

5.A. Baum–Welch algorithm

The Baum–Welch algorithm is an example of a forward-backward algorithm, and is a special case of the expectation-maximization algorithm. It can compute maximum likelihood estimates and posterior mode estimates for the parameters (transition and emission probabilities) of an HMM, when given only emissions as training data [5].

Baum-Welch algorithm implies that first initialization the set $\lambda = \{A, B, \Pi\}$ with random initial conditions. The algorithm updates the parameters of $\lambda$ iteratively until convergence, following the procedure below:

The forward procedure: We define: $\alpha_i(t) = p(O1 = o_1, \ldots, Ot = o_t, Qt = i | \lambda)$, which is the probability of seeing the partial sequence $o_1, \ldots, o_t$ and ending up in state $i$ at time $t$.

$\alpha_i(t)$ can be recursively, calculated as:

1. $\alpha_i(t) = \pi_i b_i(o_1), \quad (10)$

2. $\alpha_j(t + 1) = b_j(o_t) \sum_{i=1}^{N} \alpha_i(t) a_{ij}, \quad (11)$

The backward procedure: This is the probability of the ending partial sequence $oT+1, \ldots, oT$ given that we started at state $i$, at time $t$. We can efficiently calculate $\beta_i(t)$ as:

$\beta_i(T) = 1, \quad (12)$

$\beta_i(t) = \sum_{j=1}^{N} \beta_j(t+1) a_{ij} b_j(o_{t+1}), \quad (13)$

The following variables can be calculated using $\alpha$ and $\beta$, such as follows:

\[ y_i(t) = p(O_t = i|O, \lambda) \frac{\alpha_i(t) \beta_i(t)}{\sum_{j=1}^{N} \alpha_j(t) \beta_j(t)} \] 

(14)

\[ \xi_{ij}(t) = p(Q_t = i, Q_{t+1} = j|O, \lambda) \frac{\alpha_i(t) \alpha_j(t+1) \beta_{ij}(t) \beta_j(t+1)}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i(t) \alpha_j(t+1) \beta_{ij}(t) \beta_j(t+1)} \] 

(15)

Having \( \gamma \) and \( \xi \), one can define update rules as follows:

\[ \pi_i = y_i(1), \] 

(16)

\[ a_{ij} = \frac{\sum_{t=1}^{T-1} \xi_{ij}(t)}{\sum_{t=1}^{T-1} y_i(t)}, \] 

(17)

\[ b_i(k) = \frac{\sum_{t=1}^{T} \delta_{O_t, o_k} y_i(t)}{\sum_{t=1}^{T} y_i(t)}, \] 

(18)

(note that the summation in the nominator of \( b_i(k) \) is only over observed symbols equal to \( o_k \). Using the updated values of A, B and \( \pi \), a new iteration is performed until convergence.

**5.B. Training HMM**

The Baum–Welch algorithm was used to train HMMs, one for each type of Sleep Apnea, using the training data set [3]. Three models with different number of hidden state \( N = \{ 5, 7, 13 \} \) were used for each Sleep Apnea type.

A separate Model was defined for each class of patterns. Maximum likelihood classification of an unknown observation sequence can be achieved by calculating the probability of the observations given the model \( P(O|\lambda) \) for each model in turn.

The unknown Pattern is assigned to the class of the model that has the highest probability of generating the observed data; that is for \( M \) classes \( C = c_1, c_2, \ldots, c_M \), where \( c_m \) is represented by model \( \lambda_m \) then \( O \) is assigned to the class \( c_m \) if

\[ P(O|\lambda( \equiv \max_{d=1}^{M} P(O|\lambda_d) ) \] 

(19)

Table 3 illustrates the results of classifications obtained using HMM models of order 5, 7, and 13. It revealed average classification rate reached about 98.43 % using HMM model of order 13 for the nonlinear features.
Table (3): Classification Results Of HMM With Orders 5, 7 And 13

<table>
<thead>
<tr>
<th>HMMs Order</th>
<th>Normal</th>
<th>OSA</th>
<th>Hypopnea</th>
<th>Central Apnea</th>
<th>Average Classification Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>12/16</td>
<td>14/16</td>
<td>14/20</td>
<td>11/12</td>
<td>79.68%</td>
</tr>
<tr>
<td>7</td>
<td>12/16</td>
<td>15/16</td>
<td>18/20</td>
<td>10/12</td>
<td>85.93%</td>
</tr>
<tr>
<td>13</td>
<td>15/16</td>
<td>16/16</td>
<td>20/20</td>
<td>12/12</td>
<td>98.43%</td>
</tr>
</tbody>
</table>

6. Results

Figs. 1 (a), (b), (c) and (d) display the overnight oximetric recordings for a Normal, OSA, Hypopnea and Central Apnea patients, respectively. The mean saturation level of Normal SaO2 signals is about 95 while the mean level of other Apneic records is almost less than 90.

Figs. 2 (a), (b), (c) and (d) show the second-order difference plots for the SaO2 signals depicted in Fig. 1. It is clear that dispersion of Hypopneic signals is greater than OSA which is greater than Central Apnea and Normal subjects, respectively. This confirms the results in Table I where CTM1 for Normal subject has higher values compared with OSA, Central Apnea and Hypopneic patients. This means that: Normal subjects have the lowest variability. Increasing the value of radii \( r \) would noticeably increase values of CTM5 and CTM8 for Hypopneic, OSA and Central Apnea patients.

Figs. 3 (a), (b), (c) and (d) show the results of the LZ complexity evaluated from SaO2 signals. It is clear that, Hypopneic signals have the highest values of LZC while Central Apneic subjects have the lowest one.

Table 1 also indicates that the Approximate Entropy (ApEn) using different values of \( r \) is always greater for Hypopneic and OSA when compared with Central Apnea and Normal subjects. This means that irregularity of Hypopneic and OSA records is greater than Central Apnea and Normal subjects.

A statistical Comparison using (t – test) was performed for comparing the population mean of Normal group with OSA group, Normal group with Hypopneic group, and Normal group with Central Apnea patients group, respectively.

Table 2 shows the results of the t-test for four different groups of cases, using nonlinear features. It can be seen that there are significant differences in some parameters in the case of OSA, Hypopneic and Central Apnea groups.
The parameters CTM5 and Appr.15 are significantly different for OSA group. Therefore, these parameters can be considered as discriminating parameters for OSA cases.

Similarly, the parameters CTM8 and LZ are considered as discriminating parameters for Hypopneic group, and CTM1 and Appr.25 for Central Apnea group.

Baum–Welch algorithm was used for training HMMs, applying the hold-out method [12], where 50% of the 128 SaO2 records were used for training, that is 64 records (16 for Normal, 16 for OSA, 20 for Hypopnea and 12 for Central Apnea), and the other 64 records were used for testing stage.

Table 3 illustrates the results of classifications obtained using HMMs of orders 5, 7, and 13. It revealed average classification rate reached about 98.43% when using HMM model of order 13 for the nonlinear features.

7. Conclusion

Since Sleep Apnea diseases cause a change in the dynamics of the oximetry system and consequently the physical measured observation sequences have different statistical linear and nonlinear prosperity therefore, it is convenient to develop HMM classifiers with different orders to detect three types of Sleep Apnea (OSA, Hypopnea and Central Apnea).

In this the present work, nonlinear features extracted from central tendency measure (CTM), Lempel-Ziv complexity (LZC) and approximate entropy (ApEn) were used to characterize SaO2 recordings from 128 subjects (32 for Normal, 32 for OSA, 40 for Hypopnea and 24 for Central Apnea).

In an attempt to achieve dimensionality reductions of features and to assess the diagnostic ability of each nonlinear feature for discriminating between normal subject and different Apneic groups, a statistical comparison using (t-test) were performed.

It has been shown that the most discriminating parameters for each group of sleep apnea were: CTM5 and Appr.15 for OSA group, CTM8 and LZ for Hypopnea group and CTM1 and Appr.25 for Central Apnea group.

For the classification stage, three HMMs with different number of hidden states \(N=\{5,7,13\}\), were designed and tested. The Baum–Welch algorithm was used to train HMMs, using the hold-out method.

In summary, it has been concluded that the adopted nonlinear analysis and the HMM methodologies are promising as assistant tools in the diagnostic ability of SaO2 from nocturnal oximetry to detect sleep Apnea diseases. Further work is required to apply our methodology to a larger data set with a wide spectrum of sleep-related breathing disorders.
References:


