Automatic Sleep Staging using the Recursive Indexing and the Artificial Neural Networks

1Malcon Anderson Tafner, Dr., 2João Artur de Souza, Dr., 3Roberto T. Raittz, M.D., 4Gertrudes A. Dandolini, M.D.
1Instituto Catarinense de Pós-Graduação
Rua Alexandre Flemming, 59 - Blumenau – SC-Brazil
89010-670 - Fone (047) 322-5545
E-mail: malcon@flynet.com.br
2, 3Universidade Federal de Pelotas - UFPel
Rua Manoel Pedro Vieira, 645 – Florianópolis – SC-Brazil
88066-100 - Fone (048) 237-9297
E-mail: jartur@eps.ufsc.br
4Universidade Federal de Santa Catarina - UFSC
Rua Manoel Pedro Vieira, 645 – Florianópolis – SC-Brazil
88066-100 - Fone (048) 237-9297
E-mail: jartur@eps.ufsc.br

Abstract

The present work describes a study carried out for the implementation of an automatic sleep staging using the electroencephalogram - EEG signals. These signals were properly acquired by a digital poligraphy system. The system in study uses the artificial intelligence technology, specifically the artificial neural networks, for automatic staging.

INTRODUCTION

The knowledge about the vigil-sleep cycle has developed since the 1930's, when Loomis and collaborators (Apud GUIOT [4]), using a unitary and passive process, composed of stages, each one with it's own characteristics. This concept of unitary sleep lasted until the 1950's, when Aserink and Kleitman (Apud GUIOT [4]) observed the existence of another sleep stage characterized by a cortical activation and rapid eye movement outbreaks; from this study, sleep comes to be considered as a process composed of two distinct biological stages. In the following years, new researches by Dement and Kleitman (Apud GUIOT [4]) verified that in this dual process the two sleep stages cyclically alternated, in addition to associating dreams with the rapid eye movements.

With the opening of a new physiology field, and the possibility of registering a considerable number of variables during sleep, the amount of research on the subject grew significantly. In 1968 criteria of sleep monitoring and staging were defined by RECHTSCHAFFEN and KALES [9]. Although revised by CARSKADON and RECHTSCHAFFEN [2] en 1994, the pattern of 68 is still used.

Following the staging criteria, a great number of doctors and technicians in polissonography have done the staging of the polissonographic records manually since the 1970's. This task, when done manually, is considered boring, since the record of one entire night of sleep takes, approximately, 3 hours of taking notes and careful observation. For this type of reason, much effort has been made to project automatic sleep staging.

SLEEP AND ITS STRUCTURE

Sleep

Thanks to the polissonography, it's accepted today that sleep is not a homogenous stage, and that there are two distinct sleep stages. The more surprising sleep, and the last to be discovered, is the sleep when the rapid eye movements occur. Because of its initials in English, "rapid eye movements", this sleep is called REM. Although it takes only 20% of an adult sleep, the REM is so important that the remaining is called NREM (Not REM). The NREM sleep can be divided in 4 other phases called stages, from 1 to 4.

Monitoring

Most part of the knowledge about the sleep and its stages came from the polissonography, which acts as a type of "X-ray" of the sleep. It consists of various types of records, such as the electroencephalogram - EEG (electrical brain waves), the electro-oculogram - EOG (eye movements), the electromiogram - EMG (muscular tension), electrocardiogram - ECG (electrical heart waves), respiratory movements and the blood oxygenation (CARSKADON and RECHTSCHAFFEN [2]). The polissonography is done in sleep laboratories through special equipment, and used by capable technicians.
For the manual staging, according to CARSKADON and RECHTSCHAFFEN [2], the EEG channels C3 and C4 are enough to distinguish the sleep stages. However, many laboratories have also used the records from channels 01 and 02, among others, as auxiliary channels.

**Staging Pattern**

This section describes, shortly, the staging pattern of sleep in adults according to the criteria pointed by RECHTSCHAFFEN and KALES (CARSKADON and RECHTSCHAFFEN [2]).

a) **Wakeful:** Most people show an EEG with alpha rhythm when relaxed with eyes closed. This rhythm changes when the subject is tense, or even with open eyes. The alpha rhythm is also present, even with open eyes, if the subject is excessively sleepy.

b) **Stage 1:** The EEG pattern of the subject when in stage 1, is described as activity of relatively low voltage and mixed frequencies. Mixed because the presence of artifacts is common\(^1\), besides small areas of theta activity.

c) **Stage 2:** The EEG pattern of stage 2 reflects activity of low voltage and mixed frequencies. Basically, the difference between stage 2 and stage 1, is in two specific patterns known as fusos and K complex. The fusos are composed of waves of 12 to 14 Hz with duration of 0.5 to 1.5 sec., while the K complexes are described as "negative sharp waves well defined followed by a positive component" (CARSKADON and RECHTSCHAFFEN [2], p. 950).

d) **Stages 3 and 4:** The EEG patterns of stages 3 and 4 are defined by the presence of waves of slow activity. According to RECHTSCHAFFEN and KALES, in stage 3, there should be, in order to its characterization, a minimum of 20% and a maximum of 50% of waves of 2 Hz and amplitudes higher than 75 µV from pick to pick per epoch (apud CARSKADON and RECHTSCHAFFEN [2]).

\(^1\) An artifact is a product of mechanic art. In this case, artifacts is what waves with sharp vertices are called.
e) **REM Sleep:** The EEG pattern for the REM sleep is also characterized by relative low voltage and mixed frequencies. Although this stage is called REM a simple eye movement is not enough to indicate this state of sleep. This happens once the phenomenon is not universal, since many individuals do not show the rapid eye movement.

![Picture 07 – EEG in REM state (20s).](image)

**Sleep Progress During the Night**

According to CARSKADON and DEMENT [1], the ideal case of an example of a sleep seen as normal, can be configured from an adult youngster (man or woman). The normal sleep in an adult youngster, alternates between the NREM and the REM sleep throughout the night, in around four to six cycles.

The first cycle, in an adult, begins with stage I of NREM sleep, and in a few minutes (1 to 7 minutes) goes to stage 2. Stage 2, signalized by the sleep fusos, or still, by the K complexes detected in the EEG, usually lasts from 10 to 25 minutes. As a natural progress from stage 2, in 20 minutes average, it reaches stage 3. Stage 3, of short duration in this first cycle, serves as transition to stage 4, which usually lasts from 20 to 40 minutes, and then, a return to less profound stages occurs.

Small intermissions of the vigil during normal sleep occur in the form of brief awakenings in which conscience or memory is not recovered. This is manifested through 30 to 60 movements per night. The movements occur in the stage changes and in the stages of more superficial sleep (MARTINEZ [6]).

**Sleep Stagers**

Among the techniques based on IA to automatically stage sleep, researchers have used the intuitive learning to build branches of decisions and classifiers based on rules, as the specialized system for the realization of the staging suggested by RAY, LEE and MORGAN in 1986[8].

Another example of the use of rules is the work presented by NIELSEN et alli [7] which, using EEG and EOG signals, used a probabilistic network (Causal Probabilistic Network - CPN) to do the automatic sleep staging on 6 patients. As comparison, NIELSEN et alli used the manual staging done by two specialists as reference. The results obtained by NIELSEN had a concordance of 68% to 71%.

Artificial Neural Networks have also been used for sleep classifying starting from training done as selected examples, as the work of SCHALLENBRAND [10] and LOSSMANN [5], which used Kohonen neural networks to produce an automatic semi staging. We can say semi staging once LOSSMANN's work did not divide the stages according to the definitions recommended by RECHTSCHAFFEN and KALES [9], but according to his own criteria that reduces the number of stages.

**THE AUTOMATIC STAGING**

The suggested automatic staging is done by an Artificial Intelligence System, specifically, artificial neural networks. This task is done starting from a biological system pre-processed through various digital signal processing techniques.

The biological signal used for the automatic staging comes from the electroencephalogram, specifically channel C3, recommended even by CARSKADON and RECHTSCHAFFEN [2], where they call the attention to the fact that all the waves used to distinguish the stages are well visualized when using channels C3 and C4.

To the acquirement of the electroencephalogram signal, the system of digital poligraphy POLIWIN of the enterprise EMSA [2], which functions through a microcomputer linked to an analogical amplification equipment was used. The sleep signal acquired is done in a sample rate, which can vary from 150 to 300 Hz. The default value for the signal taking is 200 Hz. The quantization rate of the signal is fixed in 12 bits of precision.

**Data Pre-processing**

The data acquired by the POLIWIN system is submitted to a series of processes with the objective of increasing their representativeness in relation to the system. This representativeness increase is necessary because it can separate the noise from the information itself.

Data pre-processing consists then of 3 different phases, each one of them gives to the acquired biological information a new interpretation. The acting of one phase on the results acquired on the previous phase constitutes
the processing for this work. The phases are denominated as follows:

1. Biological Signal Adjustment
2. Recursive Index Generation
3. Characteristics Extraction

**Biological Signal Adjustment**

The signal captured by the polissonography equipment has noises that can interfere with the following phases in the pre-processing. Therefore, it is necessary that a previous processing tries to ease the presence of noise (signals of high amplitude in module) and adjust the signal to the phases that will follow. The easing of the noise is done through a cut on the signals that surpasses the defined limits (superior and inferior limits). After the signals have been adjusted to the new limits, all the negative signals are converted to its symmetric.

![Part of the original biological signal with the superior and inferior limits](image)

**Recursive Index Generation**

A technique called Recursive Indexing (INREC), whose task is, according to SOUZA (1999), to find a transformation that takes a vector $R^n$ to a unidimensional space of numeric values $R$ that keeps the capability to identify the close points in both spaces. The intention is to reduce the characteristics' space of the dimension patterns $p$ of a determined Problem Domain to a numeric value (index) $\rho \in R$, keeping the discriminatory power. Therefore, being an $X_i$ pattern represented by its $p$ characteristics, a $\rho$ index is defined as:

$$\rho = f(x_{a1}, f(x_{a2}, f(x_{a3}, ... f(x_{ap}...))))$$

where $f$ is the function for recursion, $x_i$ are the $p$ characteristics of the pattern $X_i$, and $a \in \{1,-1\}$. This way, we have $\rho$ as the result of a recursion or composition of a mathematics function applied on the original pattern.

Trying to apply this technique as part of the pre-processing, the INREC was used to reduce the amount of signals, without loss of information. This was done applying the INREC on segments of the original biological signals, after the first phase. Each segment was replaced with its respective index, building a new vector representing the original signal. To the building of this new index vector the function $f = \text{tangent hyperbolic}$.

![Index vector](image)

**Characteristics Extraction**

In this work, we call characteristics, the counting of the presence of signals of the index vector in some amplitude ranges. The amplitude ranges are defined using a breakage of the complete and ordered index vector. The breakage is defined according to the vector's size. Then, the ranges are defined as being the image of each breakage component.

![Amplitude range and breakage element for the index vector](image)

**The Artificial Neural Network**

**Architecture**

The network architecture used for this work, called neural network ARC, was projected based on a network structure type feedforward and largely connected whose neurons disposition is bidimensional (TAFNER, [13]).
The neuron of the ARC network has, besides exit y, two other information, the Action Potential's Strength, denoted by δ, and the label, denoted by R. The neural network's entry layer, whose elements are connected with all the neural network's neurons is represented mathematically by an entry vector denoted by $x = (x_1, x_2, x_3, ..., x_p) \in \mathbb{R}^p$. The connections of each neural network's neuron are represented by a vector, denoted by $w_i = (w_{i1}, w_{i2}, w_{i3}, ..., w_{ip}) \in \mathbb{R}^p$. The $w_i$ vector is initiated with values, usually, randomly chosen. The y exit of the i neuron is denoted by the Euclidean norm $|| x - w_i ||$.

### Characteristics

Each neuron of the network ARC has two characteristics that distinguish them from the normal structure of the artificial neurons. The Action Potential’s Strength and the Label. These two attributes participate actively in the network's decisive process. Another remarkable characteristic of the ARC architecture that can be emphasized is the absence of many interactions to train the neural network, occurring the assimilation of a fact with only one presentation of the training group to the network.

### The High Cortical Representation - ARC

This neural network is called ARC because of the characteristic of having many neurons responding to the same stimulus. Therefore, starting from the idea that a neural network is, by simple analogy, the representation of an artificial cortex, we have a high cortical representation for each new stimulus learned by it.

And to avoid that all neurons respond in the same way when presented to a stimulus to the network, it is attributed, to each neuron, a response with different intensity determined by the amount of training according to the spatial disposition, grouping the stimulus in specific regions.

1.1.1.1 The Action Potential’s Strength of the Exit (δ)

The Action Potential's Strength idea, δ, is based on the repetitive neural activity of the biological neurons, whose function is to pass information about the stimulus’ intensity. According to SCHAUFF et alli [11], when a prolonged depolarization results from the action of many synaptic potentials or from receptors on the receptive region of the cell, a series of action potentials can be produced, many times, on the axon. This type of response is known as repetitive neural activity (or simply, impulse train).
\[ \delta_j = f(y_j) = \begin{cases} 100 - \frac{100 \cdot y_j}{\text{Limiar}} & \text{se} \quad 0 \leq y_j \leq \text{Limiar} \\ 0 & \text{se} \quad y_j > \text{Limiar} \end{cases} \]

where the Limiar \( \in [0, 100]\).

**The Neuron’s Label (R)**

The neuron's label carries the stimulus's neural representation, since in this neural network each neuron has the power of representation of a whole stimulus. In other words, we have the label of any given neuron as the wanted exit of a specific stimulus when this neuron is representing it. Therefore, for a training group, \( \overline{X} = (c_i, x_i)_{i=1}^n \), \( x_i \) is the \( i \)-th exit (\( x_i = (x_1, x_2, ..., x_p) \)) and \( c_i \) is its class. The \((i,j)\) neuron label, \( R_{ij} \) once the criteria of the learning algorithm has been met, can be given as \( R_{ij} = c_i \).

**Neural Learning**

The learning phase happens in two stages, first the network learns (assimilate) the facts that are submitted to it through an adjustment process of the neurons’ synaptic connections, and according to the network it labels the neurons with the name of the learning facts. The learning phase is simpler, being responsible only for returning the label of the selected neuron before any stimulus present in the entry layer.

**RESULT ANALYSIS**

**Initial Considerations**

For the due tests, the exams of four male adult patients provided by the EMSA enterprise were used. Professionals in the health field connected to the enterprise manually staged the exams. The provided exams were denominated as the following:

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age</th>
<th>Recording Time (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6501</td>
<td>48</td>
<td>03:53:05</td>
</tr>
<tr>
<td>7706</td>
<td>25</td>
<td>05:47:58</td>
</tr>
<tr>
<td>6740</td>
<td>25</td>
<td>07:35:43</td>
</tr>
<tr>
<td>9300</td>
<td>31</td>
<td>07:36:40</td>
</tr>
</tbody>
</table>

For the testing, the amount of classes (stages, REM and wakeful) of the automatic staging was changed. The change includes, basically, stages 3, 4 and 1. Stages 3 and 4 were grouped in one stage, called stage 3. This decision was made once both stages (3 and 4) are well defined compared to the rest of the group and their difference is not relevant to the study in question.

The ARC neural network was trained using only the exam 6501, when the rest of the exams were part of the tests for the system. For the training, 15 epochs were used for the training of the "wakeful" class, 49 for the "REM" class, 50 epochs for stage 2 and 50 epochs for stage 3.

![Picture 14 – Distribution of the Trained Neurons by Class](image)

For the building of the characteristic vectors, the pre-processing phases were followed, where for the first phase the inferior and superior limits in -1000 and +1000, respectively, were used. For the second phase, the index vector was generated at each 10 points of the original signal, and, finally, for the third phase, the characteristics extraction, 20 amplitude ranges were defined generated at regular intervals of 30 seconds, or 600 points of the index vector.

**Results**

The exam of the patient 7706, submitted to the trained neural network analysis with the exam of the patient 6501, obtained a success rate of 75.54% confronting with the results of the staging done manually. The simulation follows the following distribution:
Chart 02 – Success Distribution for the Patient 7706

<table>
<thead>
<tr>
<th>Class</th>
<th>Automatic</th>
<th>Manual</th>
<th>Success (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wakeful</td>
<td>0</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>REM</td>
<td>71</td>
<td>146</td>
<td>48.63</td>
</tr>
<tr>
<td>Stage 1/2</td>
<td>290</td>
<td>355</td>
<td>81.69</td>
</tr>
<tr>
<td>Stage 3/4</td>
<td>171</td>
<td>181</td>
<td>94.48</td>
</tr>
<tr>
<td>Total</td>
<td>532</td>
<td>695</td>
<td>76.54</td>
</tr>
</tbody>
</table>

This result was obtained in one minute and twelve seconds in a Pentium II computer of 200 MHz.

The hypnogram with the presentation of the automatic staging results (system) plus the manual staging (medical), also reveals the consonance of the results obtained according to its spatial disposition in time, that is, the obtained results are counted taking in consideration also the time where both results coincide. This disposition can be better observed on picture 15.

![Picture 15 – Hypnogram of patient 7706](image)

Another graphic representation that could be very useful is the histogram of each class by method (automatic and manual).

For the other two exams, 6740 and 9300, it's possible to observe the following results, also obtained using the same configurations applied to exam 7706:

Chart 03 – Success Distribution for the Patients 6740 and 9300

<table>
<thead>
<tr>
<th>Patient</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class</td>
<td>Automatic</td>
<td>Manual</td>
</tr>
<tr>
<td></td>
<td>Wakeful</td>
<td>1</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>REM</td>
<td>32</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Stage 1/2</td>
<td>515</td>
<td>659</td>
</tr>
<tr>
<td></td>
<td>Stage 3/4</td>
<td>89</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>637</td>
<td>911</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patient</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class</td>
<td>Automatic</td>
<td>Manual</td>
</tr>
<tr>
<td></td>
<td>Wakeful</td>
<td>4</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>REM</td>
<td>122</td>
<td>232</td>
</tr>
<tr>
<td></td>
<td>Stage 1/2</td>
<td>321</td>
<td>409</td>
</tr>
<tr>
<td></td>
<td>Stage 3/4</td>
<td>111</td>
<td>193</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>558</td>
<td>913</td>
</tr>
</tbody>
</table>

**CONCLUSIONS AND RECOMMENDATIONS**

At first, we can talk about the high speed of the automatic staging compared to the manual. The automatic staging takes up to three machine minutes, while the manual can take up to three hours. Another relevant factor is the staging quality, since the manual method tends to be tiring, therefore causing a variable quality throughout the staging, while an automatic system guaranties the performance throughout the entire process.

The work also reveals that the use of digital signal pre-processing technologies, combined with artificial neural networks show results favorable to the deepest investigations. In the case of the pre-processing, we point out the efficacy of the characteristics extraction technique, also deserving more specific studies.

As subject for future studies, we recommend the use of multiple channels of EEG, as well as the use of other channels, as it is the case of the EOG, not explored in this work for detecting the REM stage. It is also worth using other functions to obtain the index vector, besides exploring other ways of extracting the characteristic vector. Finally, the use of other neural networks' learning architecture models can bring new results.
References


